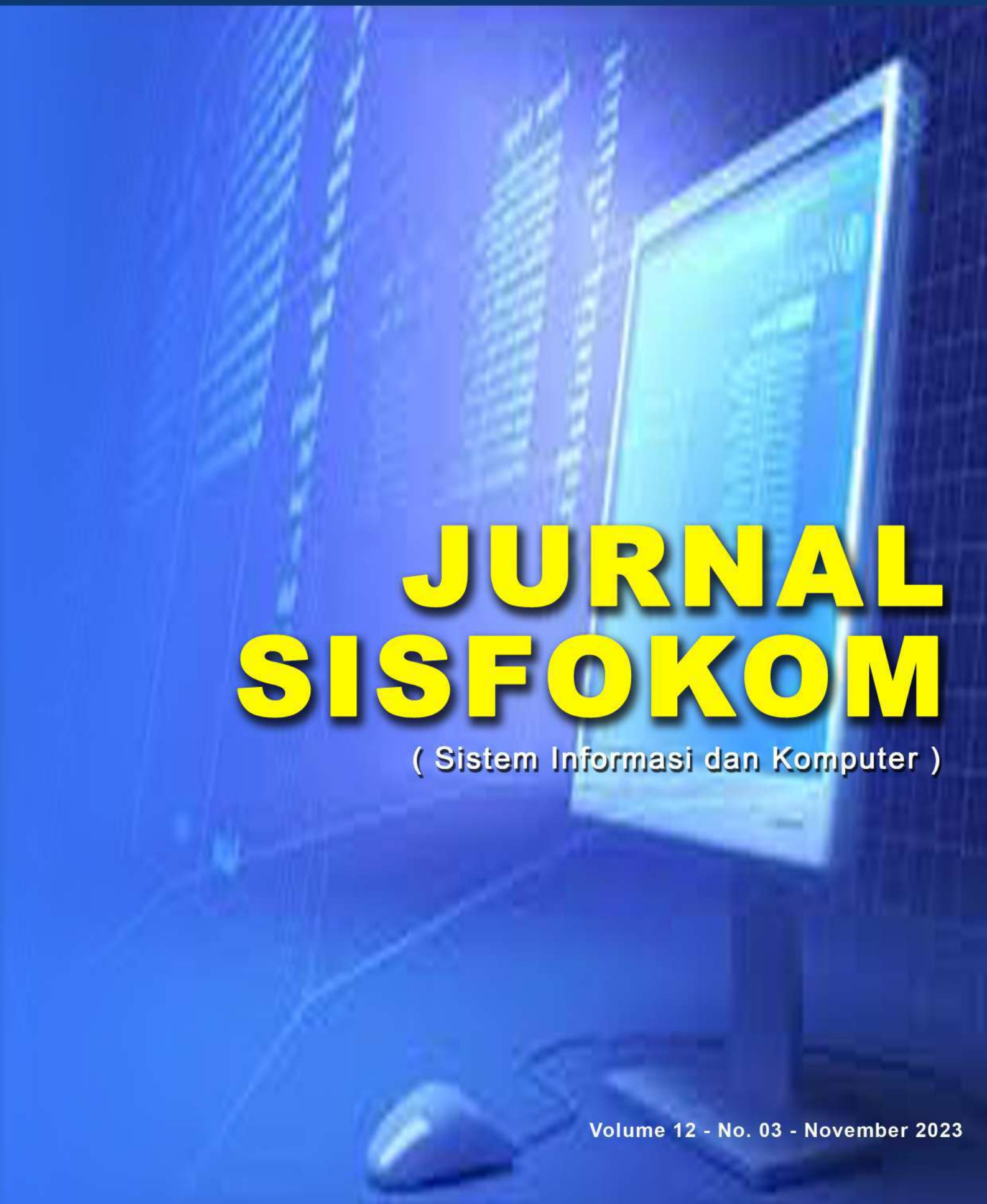
E-ISSN : 2581-0588
P-ISSN : 2301-7988

INSTITUT SAINS DAN BISNIS
ATMA LUHUR



JURNAL SISFOKOM

(Sistem Informasi dan Komputer)

Volume 12 - No. 03 - November 2023

P-ISSN: 2301-7988
E-ISSN: 2581-0588

JURNAL SISFOKOM

(SISTEM INFORMASI DAN KOMPUTER)

Jurnal Sisfokom merupakan singkatan dari Jurnal Sistem Informasi dan Komputer. Jurnal ini merupakan kolaborasi antara sivitas akademika ISB Atma Luhur dengan perguruan tinggi maupun universitas di Indonesia. Jurnal ini berisi artikel ilmiah dari peneliti, akademisi, serta para pemerhati TI. Jurnal Sisfokom diterbitkan 3 kali dalam setahun yaitu pada bulan Maret, Juli, dan November. Jurnal ini menyajikan makalah dalam bidang ilmu sistem informasi dan komputer.

Penanggung Jawab
Ketua LPPM Institut Sains dan Bisnis Atma Luhur

Ketua Dewan Penyunting
Tri Sugihartono

Penyunting Pelaksana
Maxrizal
Laurentinus
Eza Budi Perkasa
Vivi Sahfitri (Universitas Bina Darma, Indonesia)
Fransiskus Panca Juniawan (Universitas Bangka Belitung, Indonesia)
Siraj (Universitas Malikussaleh, Indonesia)
Bambang Arianto (STISIP Banten Raya, Indonesia)
Shoffan Saifullah (Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia))

Mitra Bestari
Prof. Dr. Ir. Joko Lianto Buliali, M.Sc (Institut Teknologi Sepuluh November, Indonesia)
Dr. Ir. Djoko Soetarno, D.E.A (Universitas Bina Nusantara, Indonesia)
Dr. Setiawan Hadi, M.Sc. CS. (Universitas Padjajaran, Indonesia)
Dr. Indra Budi, S.Kom., M.Kom (Universitas Indonesia, Indonesia)
A'ang Subiyakto, Ph.D (UIN Syarif Hidayatullah Jakarta, Indonesia)
Dr. Henderi, S.Kom, M.Kom (Universitas Raharja, Indonesia)
Dr. Heriyanto, A.Md., S.Kom., M.Cs (Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia)
Dr. Ir. Untung Rahardja, MTI., MM (Universitas Raharja, Indonesia)
Novan Wijaya, S.Kom., M.Kom (AMIK MDP, Indonesia)

Silvester Dian Handy Permana, S.T., M.T.I. (Universitas Trilogi, Indonesia)
Zanuar Rifai,S.Kom., M.MSI. (Universitas Amikom Purwokerto, Indonesia)
Anjar Wanto, S.Kom., M.Kom (STIKOM Tunas Bangsa, Pematangsiantar, Indonesia)
Muhammad Nur Faiz, M.Kom (Politeknik Negeri Cilacap, Indonesia)
Molavi Arman, S.Kom., M.Kom (STMIK MDP, Indonesia)
Agus Perdana Windarto, S.Kom., M.Kom (STIKOM Tunas Bangsa, Indonesia)
Usman Ependi, M.Kom (Universitas Bina Darma, Indonesia)
Oris Krianto Sulaiman, S.T., M.Kom (Universitas Islam Sumatera Utara, Indonesia)
Harrizki Arie Pradana, S.Kom, MT (Institut Sains dan Bisnis Atma Luhur, Indonesia)
Harma Oktafia Lingga Wijaya, M.Kom (STMIK Musirawas Lubuklinggau, Indonesia)
Ramos Somya, S.Kom., M.Cs (Universitas Kristen Satya Wacana, Indonesia)
Adil Setiawan, S.Kom., M.Kom (Universitas Potensi Utama, Indonesia)
Bahtiar Imram, S.ST., M.TI (Universitas Teknologi Mataram, Indonesia)
Agus Junaidi, S.Kom., M.Kom (Universitas Bina Sarana Informatika, Indonesia)
Astrid Lestari Tungadi, S.Kom., M.TI (Universitas Atma Jaya Makassar, Indonesia)
Dudih Gustian, ST., M.Kom (Universitas Nusa Putra, Indonesia)
Dr. Nihar Athreyas (Spero Devices, Inc. Springfield, United States)
Dr. Kurniabudi (Universitas Dinamika Bangsa, Indonesia)
Dr. Heny Pratiwi, S.Kom., M.Pd., M.TI (STMIK Widya Cipta Dharma, Indonesia)
Al Hafiz Akbar Maulana Siagian, Ph.D (Badan Riset dan Inovasi Nasional, Indonesia)
Khairun Nisa Meiah Ngafidin, S.Pd., M.Kom (Institut Teknologi Telkom Purwokerto, Indonesia)
Puji Winar Cahyo, S.Kom., M.Cs.(Universitas Jenderal Achmad Yani Yogyakarta, Indonesia)
Adhitia Erfina, S.T., M.Kom (Universitas Nusa Putra, Indonesia)
Muhammad Fadlan, M.Kom (STMIK PPKIA Tarakanita Rahmawati, Indonesia)
Herliyani Hasanah, MT (Universitas Duta Bangsa Surakarta, Indonesia)
Oman Somantri, S.Kom., M.Kom (Politeknik Negeri Cilacap, Indonesia)
Suprianto, S.Kom., M.Kom (STMIK PPKIA Tarakanita Rahmawati, Indonesia)
Hendra Gunawan, S.T., M.Kom (STMIK IM, Indonesia)
Anggara Trisna Nugraha, MT (Politeknik Perkapalan Negeri Surabaya, Indonesia)
Dimas Sasongko, S.Kom., M.Eng (Universitas Muhammadiyah Magelang, Indonesia)
Dr. Hetty Rohayani, ST., M.Kom (Universitas Muhammadiyah Jambi, Indonesia)
Kresna Ramanda, M.Kom (STMIK Nusa Mandiri, Indonesia)
Lusiana, M.Kom (STMIK AMIK Riau, Indonesia)
Resad Setyadi, S. S.I, ST, M. M.S.I (Institut Teknologi Telkom Purwokerto, Indonesia)
Rometdo Muzawi, M.Kom (STMIK AMIK Riau, Indonesia)
Susandri, M.Kom (STMIK AMIK Riau, Indonesia)
Tegar Arifin Prasetyo, M.Si (Institut Teknologi Del, Indonesia)
Titus Kristanto, M.Kom (Institut Teknologi Telkom Surabaya, Indonesia)
Yunita, M.Kom (Universitas Nusa Mandiri Jakarta, Indonesia)
Dr. David, M.Cs., M.Kom (STMIK Pontianak, Indonesia)
Assoc. Prof. Dr. Sandy Kosasi, M.M., M.Kom ((STMIK Pontianak, Indonesia)
Dr. Susanto, M.Kom (Universitas Bina Insan, Indonesia)

P-ISSN: 2301-7988
E-ISSN: 2581-0588

Garno, S.Kom., M.Kom (Universitas Singaperbangsa Karawang, Indonesia)
Febria S Handayani, S.Kom., M.Kom (Institut Teknologi dan Bisnis PalComTech, Indonesia)
Muhammad Rizky Pribadi, S.Kom., M.Kom (Universitas Multi Data Palembang, Indonesia)
Rivalri Kristianto Hondro, S.Kom., M.Kom (Universitas Budi Darma, Indonesia)
Muchamad Rusdan, S.T., M.T (Sekolah Tinggi Teknologi Bandung, Indonesia)
Yoga Priatyanto, S.Kom., M.Eng (Universitas Amikom Yogyakarta, Indonesia)
Ardiansyah, S.Kom., M.Kom (Universitas Muhammadiyah Klaten, Indonesia)
Yogi Isro Mukti, S.Kom., M.Kom (Institut Teknologi Pagar Alam, Indonesia)
M. Agus Syamsul Arifin, S.ST., M.Kom (Universitas Bina Insan, Indonesia)
Yulmaini, S.Kom., M.Cs (Institut Informatika dan Bisnis Darmajaya, Indonesia)

SEKRETARIAT

LPPM Institut Sains dan Bisnis Atma Luhur
Jl. Jend. Sudirman, Selindung Baru, Pangkalpinang
Kepulauan Bangka Belitung - Indonesia
Telp. (0717) 433 506 Fax. (0717) 433 506
e-mail : sisfokom@atmaluhur.ac.id - lppm@atmaluhur.ac.id

P-ISSN: 2301-7988
E-ISSN: 2581-0588

PENGANTAR REDAKSI

Jurnal Sisfokom (Sistem Informasi dan Komputer) adalah jurnal yang dikelola dan diterbitkan oleh LPPM ISB Atma Luhur Pangkalpinang.

Jurnal Sisfokom (Sistem Informasi dan Komputer) Volume 12 Nomor 03 – November 2023 ini merupakan kolaborasi antara sivitas ISB Atma Luhur dengan perguruan tinggi maupun universitas di seluruh Indonesia.

Redaksi mengucapkan terima kasih atas partisipasi dan kerja sama rekan – rekan dosen, sehingga Jurnal Sisfokom (Sistem Informasi dan Komputer) Vol.12 No.03 – November 2023 ini dapat terbit sesuai dengan yang telah kami rencanakan.

Selain itu sejumlah pakar dari dalam maupun dari luar ISB Atma Luhur telah memberikan kontribusi yang sangat berharga dalam menilai dan memberi koreksi perbaikan bagi makalah yang dimuat. Oleh karena itu redaksi mengucapkan banyak terima kasih kepada para pakar tersebut.

Pada kesempatan ini, redaksi juga mengundang dan membuka kesempatan seluas – luasnya bagi para peneliti, rekan – rekan dosen, dan pengamat / pemerhati Sistem Informasi maupun Teknik Informatika untuk ikut serta mempublikasikan hasil penelitiannya melalui jurnal ini.

Akhirnya, redaksi berharap makalah – makalah yang diterbitkan dalam jurnal ini dapat memberikan manfaat bagi seluruh elemen sivitas akademika di ISB Atma Luhur khususnya, dan bagi perkembangan ilmu pengetahuan dan teknologi informasi pada umumnya.

Redaksi,

P-ISSN: 2301-7988
E-ISSN: 2581-0588

DAFTAR ISI

P-ISSN: 2301-7988
E-ISSN: 2581-0588

P-ISSN: 2301-7988
E-ISSN: 2581-0588

Deteksi Botnet IoT Menggunakan Autoencoder dan Decision Tree

Susanto[1]*, M. Agus Syamsul Arifin[2], Harma Oktafia Lingga Wijaya[3]
Program Studi Informatika : Fakultas Ilmu Teknik[1]
Program Studi Rekayasa Sistem Komputer : Fakultas Ilmu Teknik[2]
Program Studi Sistem Informasi : Fakultas Ilmu Teknik[3]
Universitas Bina Insan
Lubuklinggau, Indonesia
susanto@univbinainsan.ac.id[1], mas.arifin@univbinainsan.ac.id [2], harmaoktafialingga@univbinainsan.ac.id [3]

***Abstract*— The use of IoT devices has grown rapidly, leading to an increase in cyber attacks that pose greater security and privacy threats than ever before. One such threat is botnet attacks on IoT devices. An IoT botnet is a group of Internet-connected IoT devices infected with malware and remotely controlled by an attacker. Machine learning techniques can be employed to detect botnet attacks. The use of machine learning-based detection methods has been shown to be effective in identifying cyber attacks. The performance of the detection system in machine learning can be improved by utilizing data reduction methods. The data reduction process in classification is used to overcome the problem of scalability and computation resources in the IoT. This paper proposes a detection system using the Autoencoder reduction method and the Decision tree classification method. The test results demonstrate that the Deep Autoencoder algorithm can reduce data and memory usage from 1.62 GB to 75.9 MB, while also improving the performance of decision tree classification, resulting in a high level of accuracy up to 100%. The Autoencoder approach in conjunction with the Decision Tree exhibits superior capabilities compared to previous studies.**

***Keywords*— Botnet IoT, Dimensionality reduction, Autoencoder, Decision Tree**

***Abstrak*— Seiring dengan pesatnya pertumbuhan perangkat IoT, serangan dunia maya juga semakin meningkat dan menimbulkan ancaman keamanan dan privasi yang lebih serius daripada sebelumnya. Salah satunya adalah serangan botnet pada perangkat IoT. Botnet IoT adalah kumpulan perangkat IoT yang terhubung ke Internet yang telah terinfeksi malware dan dikelola dari jarak jauh oleh penyerang. Salah satu teknik yang dapat digunakan dalam mendeteksi serangan botnet adalah teknik pembelajaran mesin. Penerapan metode deteksi berbasis pembelajaran mesin telah terbukti efisien dalam mendeteksi serangan cyber. Performa kinerja sistem deteksi pada pembelajaran mesin dapat ditingkatkan dengan menggunakan metode reduksi data. Proses reduksi data pada klasifikasi digunakan untuk mengatasi masalah skalabilitas dan sumber daya komputasi di IoT. Pada makalah ini, kami mengusulkan Sistem deteksi menggunakan metode reduksi Autoencoder dan metode klasifikasi Decision tree. Hasil pengujian menunjukkan bahwa algortima Autoencoder dalam mereduksi data sehingga dapat mengurangi penggunaan memori data dari 1.62 GB menjadi 75.9 MB, selain itu juga meningkatkan kinerja klasifikasi decision tree. Hal ini terlihat dari tingkat akurasi yang tinggi hingga mencapai 100%. Pendekatan Autoencoder dengan Decision Tree memiliki kemampuan yang lebih unggul dibandingkan penelitian sebelumnya.**

***Kata Kunci*— Botnet IoT, Pengurangan Dimensi, Autoencoder, Decision Tree**

I. PENDAHULUAN

Pesatnya perkembangan dan penerapan teknologi berbasis *Internet of Things* (IoT) telah memungkinkan berbagai kemungkinan kemajuan teknologi untuk berbagai aspek kehidupan [1]. Peningkatan penggunaan perangkat perangkat IoT, selaras dengan serangan *cyber* juga yang semakin meningkat, sehingga dapat mengganggu kinerja dari perangkat IoT bahkan menjadi ancaman serius pada keamanan dan privasi [2]. Serangan *cyber* pada perangkat IoT yang paling serius salah satunya adalah serangan botnet [3], [4]. Serangan botnet IoT, yang bertujuan untuk melakukan kejahatan dunia maya yang nyata, efisien, dan menguntungkan, sehingga menjadi salah satu ancaman IoT yang paling serius [5].

Salah satu teknik yang dapat digunakan dalam mendeteksi serangan botnet adalah teknik pembelajaran mesin [6]. Penerapan metode deteksi berbasis pembelajaran mesin telah terbukti efisien dalam mendeteksi serangan *cyber* [7], meskipun bukan tanpa keterbatasan [8]. Oleh karena itu performa kinerja sistem deteksi pada pembelajaran mesin dapat ditingkatkan dengan menggunakan metode reduksi data [9]. Tujuan reduksi data bukan untuk menghilangkan informasi yang dapat diekstraksi tetapi untuk meningkatkan efektivitas pembelajaran mesin ketika kumpulan data yang tersedia besar [10]. Data yang direduksi menjadikan penggunaan sumber daya penyimpanan yang lebih sedikit sehingga manajemen penyimpanan lebih efisien [11].

Hal ini terlihat telah banyak penelitian diantaranya Bahsi et al. [12] melakukan reduksi data dengan menggunakan metode fisher score yang kemudian diklasifikasikan dengan menggunakan k-NN dan decision tree. Kemudian Nomm et al. [13] menggunakan metode entropi, variance, dan Hopkins dalam mereduksi data yang kemudian diklasifikasikan dengan one class SVM and isolation forest. Selanjutnya Susanto et al. [14] melakukan proses reduksi data dengan menggunakan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023

metode fastICA, setelah itu diklasifikasikan dengan metode k-NN, Decision Tree, Random Forest, and Gradient Boosting. Selain itu Alqahtani et al. [15] mereduksi data mengunakan metode fisher score, yang selanjutnya diklasifikasikan dengan metode XGBoost.

Pekerjaan yang dilakukan ini berkontribusi terhadap pengembangan model untuk sistem deteksi serangan botnet pada jaringan IoT. Sistem deteksi yang diusulkan dengan melakukan proses reduksi data menggunakan Autoencoder, kemudian dilanjutkan dengan proses klasifikasi menggunakan Decision tree. Metode reduksi Autoencoder dipilih karena memiliki perbedaan dengan metode reduksi dimensi lainnya. Hal ini terlihat dalam beberapa kasus, pembuat enkode otomatis tidak hanya mengurangi dimensi, tetapi juga dapat mendeteksi struktur berulang [16]. Disisi lain metode klasifikasi decision tree memiliki kelebihan antara lain, Pertama, model pohon keputusan sampai taraf tertentu mudah diikuti dan diterapkan. Kedua, model pembelajaran berbasis pohon keputusan kurang memperhatikan proses penyiapan sampel karena pada tingkat tertentu tidak berguna dan memakan waktu. Ketiga, proses verifikasi efektivitas dan ketahanan model pohon keputusan dapat dengan mudah diukur [17].

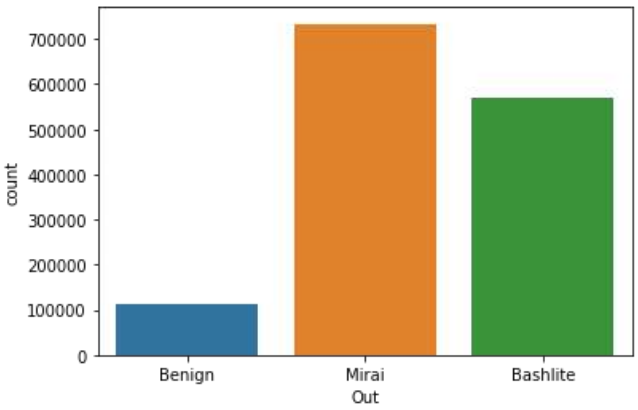
II. METODE PENELITIAN

A. Dataset

Penelitian ini menggunakan menggunakan jenis *dataset* comma-separated value (CSV) yaitu dataset N-BaIoT [18]. Dataset N-BaIoT diekstraksi menggunakan metode incremental statistic [19], yang menghasilkan total 115 fitur. Dataset N-BaIoT terdiri dari sembilan perangkat IoT yang mencakup webcam, monitor bayi, empat kamera keamanan, termostat, dan dua bel pintu. *Dataset* ini memiliki tiga jenis data lalu lintas (yaitu, satu jenis lalu lintas Benign dan dua jenis serangan lalu lintas yaitu Mirai dan Bashlite), yang berjumlah total 7.062.606 data. Dalam pengujian ini menggunakan sekitar 20% dari kumpulan data N-BaIoT, dengan total 1.415.912 data. Proses pengambilan data dilakukan dengan cara mengambil data sebanyak 20% dari setiap file dataset dimulai dari baris pertama sehingga dapat mewakili seluruh datanya. Selain itu pada dataset pelabelan data terletak pada penamaan file maka diperlukan penambahan label baru pada fiturnya untuk proses klasifikasi sehingga total fitur menjadi 116 fitur. Sebaran data tersaji pada tabel I dan gambar 1.

TABEL I. SEBARAN DATASET N-BAIOT

Label Baru	Label File	Jumlah Data
Benign	Benign	111179
Bashlite	Combo	103030
	Junk	52158
	Scan	51022
	TCP	171969
	UDP	192873
Mirai	Ack	128764
	Scan	107596
	Syn	146660
	UDP	246001
	UDPplain	104660
Total		1415912



Gambar 1. Sebaran label pada 20% dataset N-BaIoT

B. Autoencoder

Proses reduksi dimensi data berguna untuk mengurangi jumlah fitur dan memperkecil ukuran dataset sehingga dapat meningkatkan performa deteksi. Penelitian ini menggunakan algoritma Autoencoder sebagai metode reduksi data. Rumus Autoencoder dinyatakan dalam (1) dan (2) [20].

$$Y = f(X) = s(WX + bx) \quad (1)$$

$$X' = g(Y) = s(W'Y + by) \quad (2)$$

Dimana,

$Y = f(X)$ = Fungsi Encoded
W = Fungsi Weighted encoded
$X' = g(Y)$ = Fungsi Decode
bx = Bias encoded
S = Fungsi Activation
W' = Weighted decoded
by = Bias decoded

Arsitektur dan parameter Autoencoder disajikan pada tabel II. Nilai pada parameter tersebut diisi dengan cara manual yang digunakan selama pengujian trial error sampai menemukan hasil yang maksimal.

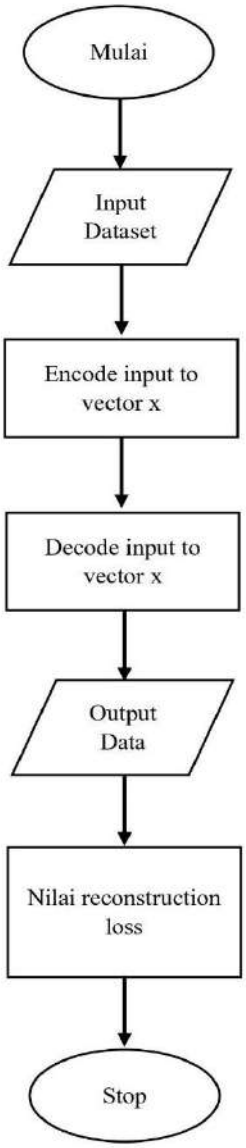
TABEL II. PARAMETER AUTOENCODER

Parameter	Nilai
Node Input layer	74
Node Ouput layer	74
Node on Hidden layer 1	50
Node on Hidden layer 2	30
Node on Hidden layer 3	20

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023

Node on Hidden layer 4	10
Node on Hidden layer 5	5
Activation function of hidden layer	Relu
Activation function of Ouput layer	Sigmoid and Softmax
Learning rate	0.00001
Fungsi Loss	Categorical crossentropy
Fungsi optimasi Optimization	Adam

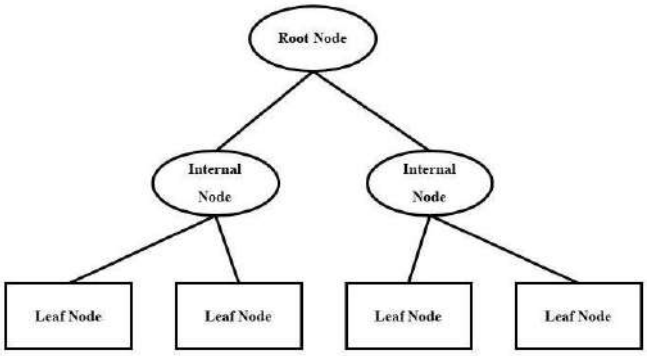
Proses pengurangan dimensi data menggunakan Autoencoder disajikan pada gambar 2.



Gambar 2. Flowchart pengurangan dimensi menggunakan Autoencoder

C. Decision Tree

Decision Tree adalah struktur mirip pohon yang memiliki daun, yang merepresentasikan klasifikasi dan cabang, yang pada gilirannya merepresentasikan konjungsi fitur yang mengarah ke klasifikasi tersebut. Keuntungan dari klasifikasi Decision Tree adalah ekspresi pengetahuan intuitif, akurasi klasifikasi tinggi, dan implementasi sederhana. Kerugian utamanya adalah untuk data, termasuk variabel kategori dengan beberapa level yang berbeda, nilai perolehan informasi cenderung mendukung fitur dengan lebih banyak level [21]. Struktur decision tree [22] disajikan pada gambar 3.



Gambar 3. Struktur decision tree

D. Evaluasi Performa

Proses evaluasi performa klasifikasi decision tree menggunakan *confusion matrix* seperti yang disajikan pada table III [23].

TABEL III. CONFUSION MATRIX

	Predicted Class		
		Possitive (P)	***Negative (N)***
Actual Class	*True* (T)	TP	FP
	False (F)	FN	TN

Dimana, TP= *True Positive*, TN= *True Negative*, FP= *False Positive*, and FN= *False Negative*. Sehingga didapat kalkulasi *classification matrix* untuk evaluasi performa peningkatan klasifikasi (3) – (7),

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (5)$$

$$Specificity = \frac{TP}{TP+FN} \quad (6)$$

$$False\text{-}positive\ rate = \frac{FP}{TN+FP} \quad (7)$$

E. Validasi

Proses validasi dilakukan dengan menggunakan berbagai perbandingan data *testing* dan data *training*, seperti yang disajikan pada tabel IV.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023

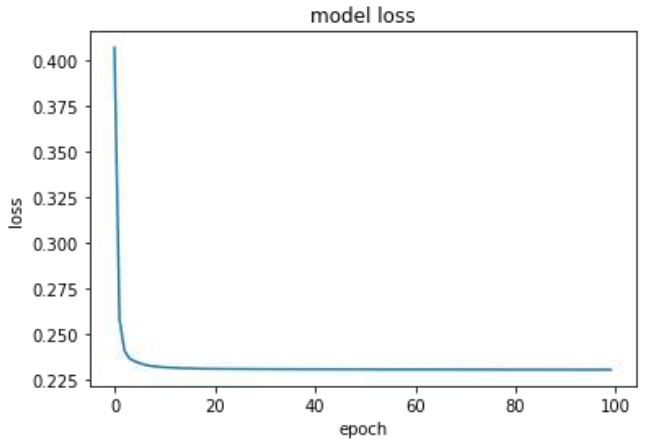
TABEL IV. VALIDASI PEMBAGIAN DATA

Validasi	Data
Validasi pertama	*Training* Data 50% and *Testing* Data 50 %
Validasi kedua	*Training* Data 60% and *Testing* Data 40 %
Validasi ketiga	*Training* Data 70% and *Testing* Data 30 %
Validasi keempat	*Training* Data 80% and *Testing* Data 20 %
Validasi kelima	*Training* Data 90% and *Testing* Data 10 %

III. HASIL PENGUJIAN DAN KOMPARASI

A. *Reduksi Data menggunakan Autoencoder*

Hasil pengujian dalam mereduksi data menggunakan arsitektur *autoencoder* dengan dengan konfigurasi 3 layer, optimasi Adam, nilai *epoch* sebesar 100, ukuran *batch size* sebesar 64, nilai *learning rate* sebesar 0.00001 dan menggunakan fungsi *loss categorial_crossentropy*. Berikut plot grafik untuk model *loss* dengan menggunakan arsitektur *autoencoder* 3 layer dapat dilihat pada gambar 4.



Gambar 4. Grafik loss data multiclass terhadap autoencoder 3 layer

Pada gambar 4 menunjukkan grafik *loss* dari proses reduksi dimensi dataset menggunakan *autoencoder* 3 layer. Nilai loss yang didapatkan pada arsitektur *autoencoder* 3 layer ini sebesar 23.09. Hasil reduksi data yang pada awalnya memiliki ukuran 115 kolom menjadi 5 kolom yang dapat dilihat pada tabel V. *Dataset* awal memiliki ukuran file sebesar 1.62 GB setelah melewati proses reduksi data menjadi sebesar 75.9 MB.

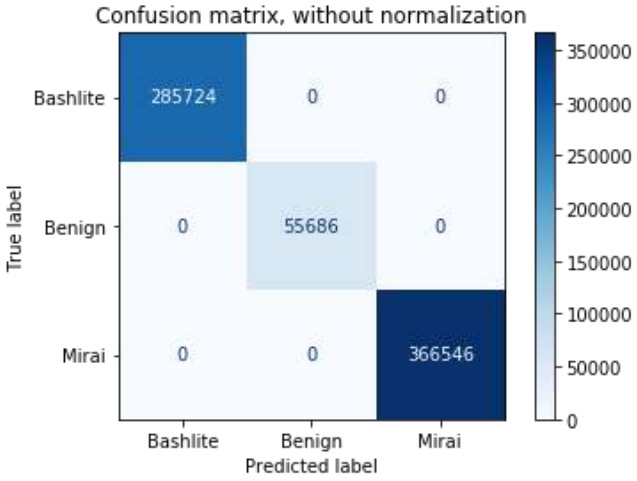
TABEL V. REDUKSI DATA MENGGUNAKAN 3 LAYER AUTOENCODER

Feature_0	Feature_1	Feature_2	Feature_3	Feature_4
0.000000	64.312424	28.513479	27.695612	50.174500
11.541241	5.329647	17.573317	4.191519	19.135214
11.444980	6.075598	36.264290	11.650904	22.337646
0.000000	15.838280	2.586906	39.339481	51.064587
9.027152	25.660934	16.014372	8.803638	8.126516

B. *Hasil Klasifikasi*

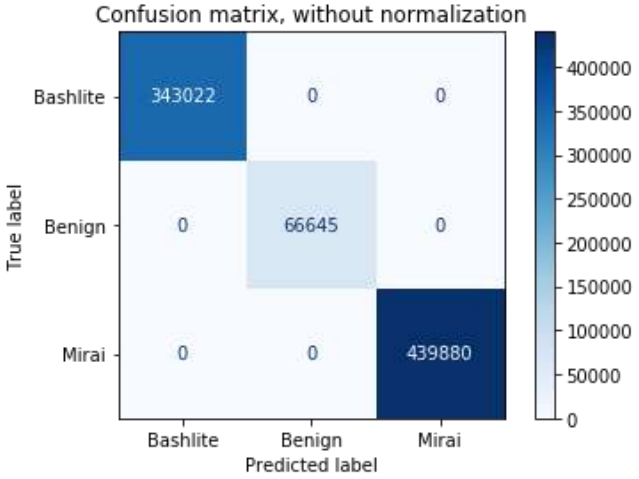
Algoritma klasifikasi decision tree diuji pada dataset kemudian dievaluasi dan dibandingkan. Tujuh metrik kinerja, yaitu: *confusion matrix*, akurasi, *precision*, *sensitifity*, *specificity*, dan *False Positive Rate* (FPR).

Gambar 5 menunjukkan hasil *confusion matrix* dari deteksi botnet dengan perbandingan 50% data *training* dan 50% data *testing*. Hasil pengujian menunjukkan bahwa klasifikasi Decision tree mampu mengklasifikasikan 709956 record data atau 100% dengan benar. Gambar 6 menunjukkan hasil *confusion matrix* dari deteksi botnet dengan perbandingan 60% data *training* dan 40% data *testing*. Hasil pengujian menunjukkan bahwa klasifikasi Decision tree mampu mengklasifikasikan 849547 record data atau 100% dengan benar. Gambar 7 menunjukkan hasil *confusion matrix* dari deteksi botnet dengan perbandingan 70% data *training* dan 30% data *testing*. Hasil pengujian menunjukkan bahwa klasifikasi Decision tree mampu mengklasifikasikan 991138 record data atau 100% dengan benar. Gambar 8 menunjukkan hasil *confusion matrix* dari deteksi botnet dengan perbandingan 80% data *training* dan 20% data *testing*. Hasil pengujian menunjukkan bahwa klasifikasi Decision tree mampu mengklasifikasikan 1132729 record data atau 100% dengan benar. Gambar 9 menunjukkan hasil *confusion matrix* dari deteksi botnet dengan perbandingan 90% data trainig dan 10% data testing. Hasil pengujian menunjukkan bahwa klasifikasi Decision tree mampu mengklasifikasikan 1274320 record data atau 100% dengan benar.

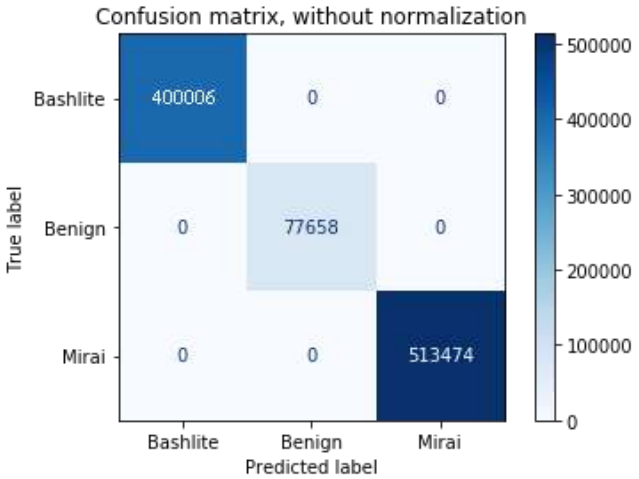


Gambar 5. Confusion matrix decision tree dengan perbandingan 50% data *training* dan 50% data *testing*

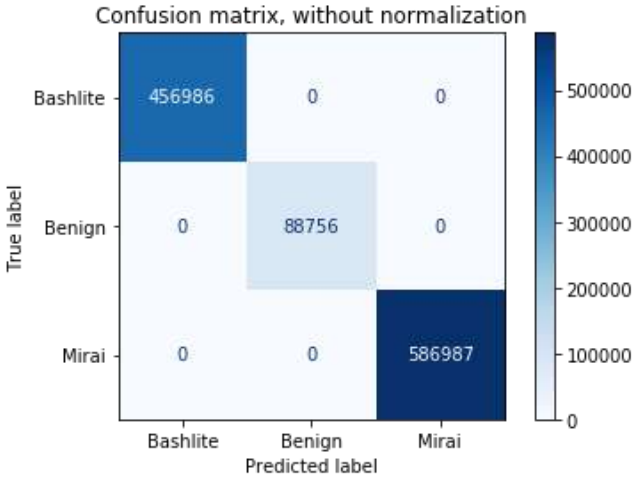
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023



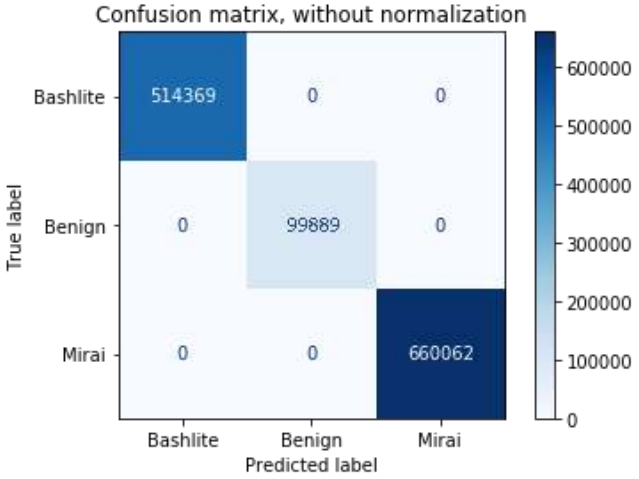
Gambar 6. Confusion matrix decision tree dengan perbandingan 60% data *training* dan 40% data *testing*



Gambar 7. Confusion matrix decision tree dengan perbandingan 70% data training dan 30% data testing



Gambar 8. Confusion matrix decision tree dengan perbandingan 80% data training dan 20% data testing



Gambar 9. Confusion matrix decision tree dengan perbandingan 90% data *training* dan 10% data *testing*

Berdasarkan hasil pengujian, didapatkan performa kinerja deteksi botnet IoT yang direduksi dengan menggunakan autoencoder, kemudian diklasifikasikan menggunakan decision tree. Hasil validasi yang disajikan pada tabel VI dan tabel VII, menunjukkan bahwa tingkat akurasi, *specificity*, *sensitivity*, *precision*, dan *false positive rate* tidak terpengaruh oleh perubahan rasio data training dan data testing. Tingkat akurasi mencapai 100% baik pada data training maupun data testing. Selanjutnya nilai *specificity*, *sensitivity*, dan *precision* tetap dinilai 1.0 baik pada data training maupun data testing. Kemudian nilai false positive rate stabil di angka 0 baik pada data *training* maupun data *testing*.

TABEL VI. HASIL PENGUJIAN TRAINING DATA KLASIFIKASI DECISION TREE

Rasio Data	Training data				
	Akurasi	*Specificity*	*Sensitivity*	*Precision*	*False Positive rate*
50:50	100	1.0	1.0	1.0	0
60:40	100	1.0	1.0	1.0	0
70:30	100	1.0	1.0	1.0	0
80:20	100	1.0	1.0	1.0	0
90:10	100	1.0	1.0	1.0	0

TABEL VII. HASIL PENGUJIAN TESTING DATA KLASIFIKASI DECISION TREE

Rasio Data	Testing data				
	Akurasi	*Specificity*	*Sensitivity*	*Precision*	*False Positive rate*
50:50	100	1.0	1.0	1.0	0
60:40	100	1.0	1.0	1.0	0
70:30	100	1.0	1.0	1.0	0
80:20	100	1.0	1.0	1.0	0
90:10	100	1.0	1.0	1.0	0

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023

C. Komparasi dengan penelitian lainnya

Untuk mengetahui tingkat ke unggulan metode yang diusulkan, penulis mengkomparasikan dengan penelitian sebelumnya yang menggunakan *dataset* sama yaitu N-BaIoT. Hasil komparasi menunjukkan bahwa metode yang diusulkan lebih akurat dalam mendeteksi botnet IoT jika dibandingkan metode lainnya, seperti yang ditunjukkan pada tabel VIII.

TABEL VIII. KOMPARASI DENGAN PENELITIAN LAINNYA

Referensi & Tahun	Metode	Dataset	Jumlah Data	Jumlah Dimensi	Akurasi
[12] (2018)	Fisher score + k-NN	N-BaIoT	1507815 (21.35%)	3	97.24
[15] (2020)	Fisher score + XGBoost with Genetika Algorithm	N-BaIoT	1667796 (23.61%)	3	99.96
This Work	**Autoencoder + Decision Tree**	**N-BaIoT**	**1415912 (20%)**	**3**	**100%**

IV. KESIMPULAN

Kami telah mengimplementasikan algortima Autoencoder dalam mereduksi data sehingga dapat mengurangi memori data dari 1.62 GB menjadi 75.9 MB, selain itu juga meningkatkan kinerja klasifikasi decision tree. Model yang diusulkan memiliki kinerja yang sangat baik pada deteksi botnet IoT, hal ini terlihat dari tingkat akurasi yang tinggi hingga mencapai 100%.. Pada penelitian selanjutnya, kami akan mencoba untuk meningkatkan kinerja deteksi dan pencegahan serangan *cyber* dari jaringan IoT yang lebih kompleks..

REFERENCES

[1] S. Nižetić, P. Šolić, D. López-de-Ipiña González-de-Artaza, and L. Patrono, "Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future," *J. Clean. Prod.*, vol. 274, 2020, doi: 10.1016/j.jclepro.2020.122877.

[2] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, "The effect of IoT new features on security and privacy: New threats, existing solutions, and challenges yet to be solved," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1606–1616, 2019, doi: 10.1109/JIOT.2018.2847733.

[3] Susanto, M. A. Syamsul Arifin, D. Stiawan, M. Y. Idris, and R. Budiarto, "The trend malware source of IoT network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 1, pp. 450–459, 2021, doi: 10.11591/ijeecs.v22.i1.pp450-459.

[4] M. Alshamkhany, W. Alshamkhany, M. Mansour, M. Khan, S. Dhou, and F. Aloul, "Botnet Attack Detection using Machine Learning," in *Proc. 14th International Conference on Innovations in Information Technology, IIT*, 2020, no. November, pp. 203–208.

[5] Z. Shao, S. Yuan, and Y. Wang, "Adaptive online learning for IoT botnet detection," *Inf. Sci. (Ny).*, vol. 574, pp. 84–95, 2021, doi: 10.1016/j.ins.2021.05.076.

[6] S. Srinivasan and D. P, "Enhancing the security in cyber-world by detecting the botnets using ensemble classification based machine learning," *Meas. Sensors*, vol. 25, no. December 2022, p. 100624, 2023, doi: 10.1016/j.measen.2022.100624.

[7] C. Maudoux, S. Boumerdassi, A. Barcello, and E. Renault, "Combined Forest: A New Supervised Approach for a Machine-Learning-based Botnets Detection," *2021 IEEE Glob. Commun. Conf. GLOBECOM 2021 - Proc.*, pp. 1–6, 2021, doi: 10.1109/GLOBECOM46510.2021.9685261.

[8] S. Miller and C. Busby-Earle, "The role of machine learning in botnet detection," *2016 11th Int. Conf. Internet Technol. Secur. Trans. ICITST 2016*, no. December, pp. 359–364, 2017.

[9] Susanto, D. Stiawan, M. A. S. Arifin, J. Rejito, M. Y. Idris, and R. Budiarto, "A Dimensionality Reduction Approach for Machine Learning Based IoT Botnet Detection," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 2021–Octob, no. October, pp. 26–30, 2021, doi: 10.23919/EECSI53397.2021.9624299.

[10] I. Czarnowski and P. Jędrzejowicz, "An approach to data reduction for learning from big datasets: Integrating stacking, rotation, and agent population learning techniques," *Complexity*, vol. 2018, 2018, doi: 10.1155/2018/7404627.

[11] M. H. ur Rehman, C. S. Liew, A. Abbas, P. P. Jayaraman, T. Y. Wah, and S. U. Khan, "Big Data Reduction Methods: A Survey," *Data Sci. Eng.*, vol. 1, no. 4, pp. 265–284, 2016, doi: 10.1007/s41019-016-0022-0.

[12] H. Bahsi, S. Nomm, and F. B. La Torre, "Dimensionality Reduction for Machine Learning Based IoT Botnet Detection," in *Proc. 2018 15th International Conference on Control, Automation, Robotics and Vision, ICARCV*, 2018, pp. 1857–1862.

[13] S. Nomm and H. Bahsi, "Unsupervised Anomaly Based Botnet Detection in IoT Networks," in *Proc.- 17th IEEE International Conference on Machine Learning and Applications, ICMLA*, 2019, pp. 1048–1053.

[14] Susanto *et al.*, "Dimensional Reduction With Fast ICA for IoT Botnet Detection," *J. Appl. Secur. Res.*, vol. 0, no. 0, pp. 1–24, 2022, doi: 10.1080/19361610.2022.2079906.

[15] M. Alqahtani, H. Mathkour, and M. M. Ben Ismail, "IoT botnet attack detection based on optimized extreme gradient boosting and feature selection," *Sensors (Switzerland)*, vol. 20, no. 21, pp. 1–21, 2020, doi: 10.3390/s20216336.

[16] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, 2016, doi: 10.1016/j.neucom.2015.08.104.

[17] Y. Liu and S. Yang, "Application of Decision Tree-Based Classification Algorithm on Content Marketing," *J. Math.*, vol. 2022, 2022, doi: 10.1155/2022/6469054.

[18] Y. Meidan *et al.*, "N-BaIoT-Network-based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervasive Comput.*, vol. 17, no. 3, pp. 12–22, Sep. 2018, doi: 10.1109/MPRV.2018.03367731.

[19] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," *arXiv*, no. February, pp. 18–21, 2018.

[20] Y. N. Kunang, S. Nurmaini, D. Stiawan, A. Zarkasi, and F. Jasmir, "Automatic Features Extraction Using Autoencoder in Intrusion Detection System," in *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science, ICECOS 2018*, 2019, vol. 17, pp. 219–224, doi: 10.1109/ICECOS.2018.8605181.

[21] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494502.

[22] A. Banjongkan, W. Pongsena, N. Kerdprasop, and K. Kerdprasop, "A study of job failure prediction at job submit-state and job start-state in high-performance computing system: Using decision tree algorithms," *J. Adv. Inf. Technol.*, vol. 12, no. 2, pp. 84–92, 2021, doi: 10.12720/jait.12.2.84-92.

[23] A. Tharwat, "Classification assessment methods," *Appl. Comput. Informatics*, vol. 17, no. 1, pp. 168–192, 2021, doi: 10.1016/j.aci.2018.08.003.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1693, Copyright ©2023
Submitted : May 3, 2023, Revised : May 8, 2023, Accepted : June 15. 2023, Published : November 03, 2023

Restorasi Citra Dengan *Image completion* Berbasis Deep Learning

Phie Chyan[1]*, N. Tri Suswanto Saptadi[2]
Program Studi Teknik Informatika, Fakultas Teknologi Informasi [1], [2]
Universitas Atma Jaya Makassar
Makassar, Indonesia
Phie_chyan@lecturer.uajm.ac.id[1], Tri_saptadi@lecturer.uajm.ac.id [2]

***Abstract*— Digital images can experience various disturbances in acquisition and storage, one of which is a disturbance indicated by damage to certain areas of the image field and causes the loss of some of the information represented by the image. One of the ways to restore an image experiencing disturbances like this is with image completion technology. Image completion is an image restoration technology capable of filling in or completing missing or corrupted parts of an image. Various methods have been developed for this image completion, starting from those based on basic image processing to the latest relying on artificial intelligence algorithms. This study aims to develop and implement an image completion model based on deep learning with the transfer learning method from the completion.net architecture. Using the Facesrub training dataset consisting of a collection of unique facial photos allows the model to understand facial attributes better. Compared to conventional image completion based on image patches, the method developed in this study can perform image filling in image gaps with more realistic results. Based on visual tests conducted on respondents, the results obtained enable respondents to understand all the information represented by the restored image, similar to the original image.**

Keywords*— *Image completion, Image restoration, Deep learning, CNN, Completion net

***Abstrak*—Citra digital dapat mengalami berbagai gangguan dalam proses akuisisi maupun penyimpanannya, salah satunya adalah gangguan yang diindikasikan dengan rusaknya area tertentu pada bidang citra dan menyebabkan hilangnya sebagian informasi yang direpresentasikan oleh citra tersebut. Salah satu cara untuk merestorasi citra yang mengalami gangguan seperti ini adalah dengan teknologi *image completion*. *Image completion* ini adalah merupakan teknologi restorasi citra yang mampu melakukan pengisian atau melengkapi bagian hilang atau terkorupsi dari citra. Beragam metode telah dikembangkan untuk *image completion* ini mulai dari yang berbasis pengolahan citra dasar hingga yang terbaru mengandalkan algoritma kecerdasan buatan. Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan model *image completion* berbasis deep learning dengan metode transfer learning dari arsitektur completion net. Dengan menggunakan dataset latih *Facesrub* yang terdiri dari koleksi berbagai foto wajah yang unik, memungkinkan model untuk memahami atribut wajah dengan lebih baik. Dibandingkan dengan *image completion* konvensional berbasis tambalan citra, metode yang dikembangkan dalam penelitian ini mampu untuk melakukan pengisian citra pada celah citra dengan hasil yang lebih realistis. Berdasarkan pengujian visual yang dilakukan kepada responden, hasil yang diperoleh memungkinkan responden dapat memahami keseluruhan informasi yang direpresentasikan oleh citra hasil restorasi serupa dengan citra asli.**

Kata Kunci*— *Image completion, Restorasi citra, Deep learning, CNN, Completion net

I. PENDAHULUAN

Citra digital dapat mengalami berbagai gangguan baik pada proses akuisisi maupun penyimpanannya. Gangguan yang dialami dapat berupa *noise* seperti piksel-piksel yang tersebar secara random dan tidak berkesesuaian dengan piksel disekitarnya hingga gangguan yang mengkorupsi bagian tertentu dari citra sehingga dapat menyebabkan sebagian informasi dari citra tidak dapat dipahami secara lengkap [1]–[3].

Merestorasi citra yang memiliki *noise* yang tinggi merupakan permasalahan penelitian yang sampai saat ini terus dieksplorasi, beragam metode telah dikembangkan untuk mengatasi hal ini [4]–[7]. Salah satu metode yang paling populer adalah menggunakan *Generatice Adversarial Networks* (GAN), selain itu berbagai metode *patch matching* dan pendekatan non pembelajaran mesin yang berbasis pengolahan citra dasar juga banyak digunakan tetapi belum mendapatkan hasil yang memuaskan.

Image completion atau *inpainting* merupakan suatu teknologi restorasi citra yang berkaitan dengan beragam teknik yang digunakan untuk mengisi bagian yang hilang atau terkorupsi dari sebuah citra [8]. Saat ini telah banyak penelitian yang mengkaji mengenai metode dalam melakukan *image completion*, meskipun demikian pengembangan metode *image completion* ini tetap menjadi tantangan yang kompleks dalam implementasinya karena umumnya metode ini membutuhkan pemahaman yang tinggi terhadap konteks citra [9]–[11]. Hal ini diperlukan untuk dapat mengisi pola citra yang hilang dan lebih penting lagi pengenalan terhadap konteks citra yang akan diisi [12]–[14]. Pemanfaatan deep learning dalam *image completion* dapat memberikan hasil yang lebih baik karena pendekatan ini memanfaatkan korelasi yang diperoleh antara berbagai pola dari gambar berderau dan gambar asli [15].

Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan model *image completion* yang dibasiskan pada model berbasis *Fully-Convolutional deep*

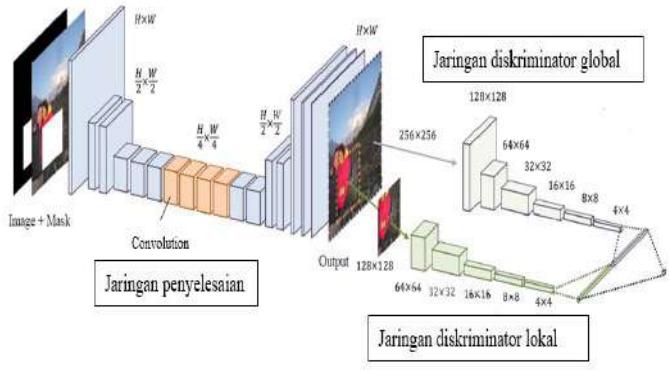
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023

learning network (FCN) yang disebut *completion network* model [8]. Arsitektur ini terdiri dari tiga bagian yaitu, jaringan penyelesaian, diskriminator global dan diskriminator lokal. Jaringan penyelesaian adalah merupakan jaringan konvolusi penuh yang digunakan untuk melakukan *image completion* sedangkan diskriminator global dan lokal merupakan jaringan yang akan digunakan untuk proses latih dan menentukan apakah citra telah dikomplesi secara konsisten. Diskriminator global mengambil citra keseluruhan sebagai input untuk mengenali konsisten global dari suatu *scene* sedangkan diskriminator lokal menjangkau area yang lebih kecil disekitar area yang dikomplesi untuk menilai kualitas tampilan yang lebih detail. Diskriminator ini akan diperbarui selama setiap iterasi dalam pelatihan sehingga dapat membedakan dengan baik antara citra asli dan lengkap dari hasil pelatihan. Kemudian setelah itu jaringan penyelesaian akan diperbarui sehingga mengisi area yang hilang pada citra dengan baik dan memenuhi persyaratan dari jaringan diskriminator.

Hasil dari metode *image completion* berbasis *deep learning* ini berupa citra yang telah dikomplesi yang berasal dari citra dengan bagian yang terkorupsi yang telah diisi dengan bagian yang telah diketahui dari citra [16],[17]. Untuk mengevaluasi keluaran dari metode yang digunakan akan dilakukan perbandingan terhadap citra asli dengan citra terkorupsi yang telah dikomplesi dengan metode yang digunakan. Pengembangan model *image completion* menggunakan *transfer learning* dari *completion net* dan dikombinasikan dengan dataset latih *Facesrub*, merupakan kontribusi dari penelitian ini yang memungkinkan model yang dikembangkan mampu melakukan pengisian terhadap berbagai citra objek termasuk wajah.

II. METODE PENELITIAN

Metode *image completion* yang digunakan berbasis *deep learning* dengan arsitektur *convolutional neural network*. Arsitektur ini terdiri dari tiga buah jaringan yang terdiri dari jaringan penyelesaian, diskriminator lokal dan diskriminator global. Sebuah jaringan penyelesaian digunakan untuk *image completion* dengan dua jaringan tambahan yaitu diskriminator lokal dan diskriminator global digunakan untuk melatih jaringan untuk menyelesaikan proses pengisian citra. Dalam periode latih, jaringan diskriminator disiapkan untuk memastikan apakah kondisi citra sudah terisi secara lengkap, sementara jaringan penyelesaian dilatih untuk mengelabui jaringan diskriminator. Dengan melatih ketiga jaringan bersama-sama dimungkinkan untuk jaringan penyelesaian dapat menyelesaikan pengisian gambar secara realistis. Ilustrasi dari pendekatan yang dibahas ini disajikan pada gambar 1. Secara garis besar penelitian dibagi menjadi beberapa tahap yaitu pertama, studi literatur, dimana dalam tahap ini peneliti meninjau berbagai metode state-of-the-art untuk *image completion* utamanya yang berbasis *machine* dan *deep* learning, tahap kedua dilanjutkan dengan persiapan dataset dan preprocessing terhadap citra yang akan digunakan, tahap ketiga dilanjutkan dengan pelatihan, validasi dan testing model dan tahap terakhir dilakukan pengujian untuk melihat bagaimana kinerja model dalam melakukan *image completion.*



Gambar 1. Arsitektur model Completion.Net

A. Convolutional Neural Network

Model jaringan yang digunakan dibasiskan pada model *Convolutional Neural Network* (CNN). Ada dua model arsitektur CNN yang diterapkan pertama adalah arsitektur standar CNN dan yang kedua adalah Variasi dari standar CNN yang disebut dilated CNN, yang memungkinkan ukuran tiap layer ditingkatkan sebagai input. Hal ini dilakukan tanpa meningkatkan jumlah bobot pembelajaran dengan cara menyebarkan kernel konvolusi pada peta input. Secara matematis operator konvolusi melebar dapat dituliskan untuk setiap piksel sebagai berikut.

$$y_{u,v} = \sigma\left(b + \sum_{i=-k'_h}^{k'_h} \sum_{j=-k'_w}^{k'_w} W_{k'_h+i,k'_w+j}\, x_{u+\eta i, v+\eta j}\right),$$
$$k'_h = \frac{k_h - 1}{2}, \qquad k'_w = \frac{k_w - 1}{2}, \tag{1}$$

Dimana k_w dan k_h adalah lebar dan tinggi kernel, η adalah faktor dilasi, $X_{u,v} \varepsilon R^C$ dan $Y_{u,v} \varepsilon R^{C'}$ adalah komponen piksel dari layer input dan output. $W_{s,t}$ adalah matriks kernel C' x C dan $b \varepsilon R^C$ adalah vektor layer bias. Kedua jaringan konvolusi dilatih untuk meminimalkan fungsi loss dengan propagasi mundur dan dilatih dengan dataset yang terdiri dari pasangan input dan output. Fungsi loss bertujuan untuk meminimalkan jarak jaringan output dengan pasangan output yang bersesuaian pada dataset [18].

B. Jaringan Penyelesaian

Jaringan penyelesaian merupakan jaringan yang berbasis konvolusional penuh. Input dari jaringan penyelesaian adalah citra RGB dengan *channel* biner yang mengindikasikan masking *image completion* dan output juga merupakan citra RGB. Arsitektur umum mengikuti struktur *encoder-decoder* yang dapat mengurangi penggunaan memori dan waktu komputasi dengan menurunkan resolusi di awal pemrosesan sebelum memproses citra. Setelah itu kemudian output dikembalikan ke resolusi asli menggunakan layer dekonvolusi [8]. Berbeda dengan arsitektur umumnya yang memanfaatkan banyak layer pooling untuk menurunkan resolusi, pada arsitektur ini hanya menurunkan resolusi dua kali menggunakan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023

stride konvolusi ¼ kali dari ukuran asli yang penting untuk menghasilkan tekstur yang tidak blur pada area yang hilang pada citra. Layer konvolusi terdilasi juga biasa digunakan pada layer tengah. Layer ini menggunakan kernel yang tersebar dan dapat menghitung setiap piksel output dengan input area yang lebih lebar tetapi tetap dengan pemanfaatan jumlah parameter dan waktu komputasi yang sama. Hal ini penting dalam menghadirkan realisme dalam image completion karena dengan menggunakan konvolusi terdilasi pada resolusi rendah, model dapat menjangkau area input yang lebih luas secara efektif dalam menghitung setiap piksel output dibandingkan dengan layer konvolusi biasa.

C. Jaringan Diskriminator Konteks

Diskriminator konteks lokal dan global bertujuan untuk mengidentifikasi bilamana suatu citra itu masih asli atau telah selesai diisi. Jaringan diskriminator ini didasarkan pada jaringan konvolusional yang memampatkan citra menjadi fitur vektor. Keluaran dari jaringan digabungkan bersama oleh rangkaian layer yang memperkirakan nilai kontinyu yang sesuai dengan probabilitas citra menjadi nyata. Jaringan diskriminator global menerima input dan di skalakan ke 256 x 256 piksel. Jaringan ini juga terdiri dari 6 layer konvolusi dan sebuah layer terkoneksi penuh yang menghasilkan sebuah 1024 vektor dimensional. Jaringan diskriminator lokal memiliki pola yang sama dengan diskriminator global kecuali untuk input menggunakan tambalan citra 128 x 128 piksel yang berpusat pada area citra yang akan diisi.

Output dari jaringan diskriminator lokal dan global dihubungkan bersama kedalam 2048 vektor dimensional yang kemudian diproses oleh sebuah layer terkoneksi penuh untuk menghasilkan nilai kontinyu. Fungsi transfer sigmoid digunakan agar nilai berada dalam range 0 dan 1 yang merepresentasikan probabilitas citra adalah asli atau apakah citra dikomplesi dengan isian.

D. Proses Latih

Secara garis besar proses latih dari jaringan *image completion* dinyatakan melalui diagram alir pada gambar 2. Untuk melatih jaringan dalam mengkomplesi citra input secara realistis maka dua fungsi loss digunakan bersama-sama yaitu *weighted mean squared error* (MSE) untuk stabilitas latih dan *Generative Adversarial Networks* (GAN) untuk meningkatkan realisme pada citra hasil [19], [20]. Dengan menggunakan gabungan dari kedua fungsi *loss* ini memungkinkan pelatihan stabil pada model jaringan berkinerja tinggi. Untuk menstabilkan pelatihan, MSE *loss* yang digunakan memperhitungkan *mask* komplesi *region* yang dipakai. *MSE loss* didefnisikan sebagai

$$L(x, M_c) = \| M_c \odot (C(x, M_c) - x) \|^2 \quad (2)$$

Jaringan diskriminator konteks juga merupakan fungsi loss yang umum dikenal sebagai GAN loss dan merupakan bagian yang penting untuk mengubah *neural network* dengan optimasi standar ke optimalisasi min-max dimana pada setiap iterasi jaringan diskriminator diperbarui bersama dengan jaringan penyelesaian. Optimalisasi kedua jaringan tersebut dinyatakan menjadi

$$\min_C \max_D \mathbb{E}[\log D(x, M_d) + \log(1 - D(C(x, M_c), M_c))] \quad (3)$$

Dengan menggabungkan kedua fungsi loss menjadi

$$\min_C \max_D \mathbb{E}[L(x, M_c) + \alpha \log D(x, M_d) + \alpha \log(1 - D(C(x, M_c), M_c))], \quad (4)$$

Dimana C dan D adalah jaringan diskriminator, M_d adalah *random mask*, M_c adalah *input mask*, α adalah *hyper parameter* dan nilai ekspektasi merupakan nilai rata-rata atas citra latih x. ADADELTA digunakan sebagai algoritma optimasi yang mengatur laju pembelajaran untuk setiap bobot pada jaringan secara otomatis.

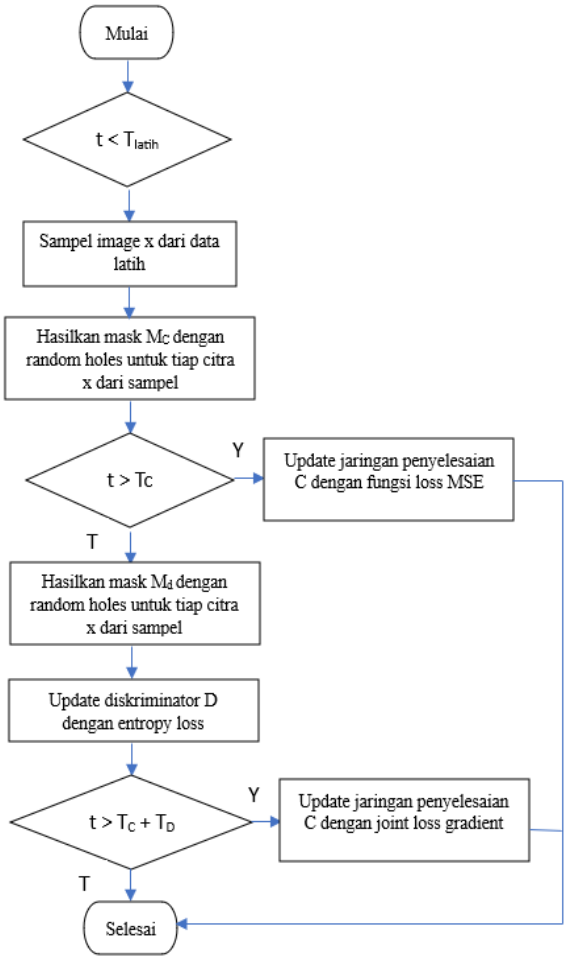
Dalam proses optimasi, jaringan penyelesaian dan diskriminator dinyatakan sebagai perubahan nilai C dan D yang berarti bahwa bias dan bobot jaringan berubah. Pada gradien stokhastis turunan persamaan optimasi min-max berarti bahwa untuk latih C, diambil nilai gradien dari fungsi loss terhadap θc dan memperbarui parameter agar nilai dari fungsi loss berkurang. Persamaan gradien adalah sebagai berikut

$$\mathbb{E}[\nabla_{\theta_C} L(x, M_c) + \alpha \nabla_{\theta_C} \log(1 - D(C(x, M_c), M_c))] . \quad (5)$$

Dalam implementasinya dilakukan cara yang lebih efisien dengan menjaga nilai gradien diskriminator untuk membantu menstabilkan proses pembelajaran.

Dalam proses latih, Diskriminator dilatih untuk membedakan gambar asli dan palsu, sementara jaringan penyelesaian dilatih untuk mengelabui diskriminator. Karena optimisasi terdiri dari menggabungkan fungsi min dan max menyebabkan jaringan ini tidak begitu stabil sehingga untuk meningkatkan kestabilannya maka metode ini tidak dapat digunakan untuk menggenerate citra dari citra yang berderau. Gambaran mengenai prosedur latih seperti yang tersaji pada gambar 2 adalah proses latih terbagi menjadi 3 tahap yaitu pertama, jaringan penyelesaian dilatih menggunakan fungsi MSE loss dari persamaan (2) untuk iterasi sebanyak T_c. Setelah itu jaringan penyelesaian ditetapkan dan jaringan diskriminator yang berikutnya dilatih dari awal sebanyak iterasi T_D. dan fase terakhir atau yang ketiga, jaringan penyelesaian dan diskriminator dilatih bersama-sama hingga akhir proses latih. Proses latih dilakukan dengan mengatur ukuran citra agar titik terkecil merupakan nilai random dalam [256,384] rentang piksel. Setelah itu tambalan random 256x256 piksel diekstraksi dan digunakan sebagai citra input. Untuk *mask*, digunakan lubang random dengan rentang piksel antara [96,128] dan diisi dengan nilai piksel rata-rata dari dataset latih.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023



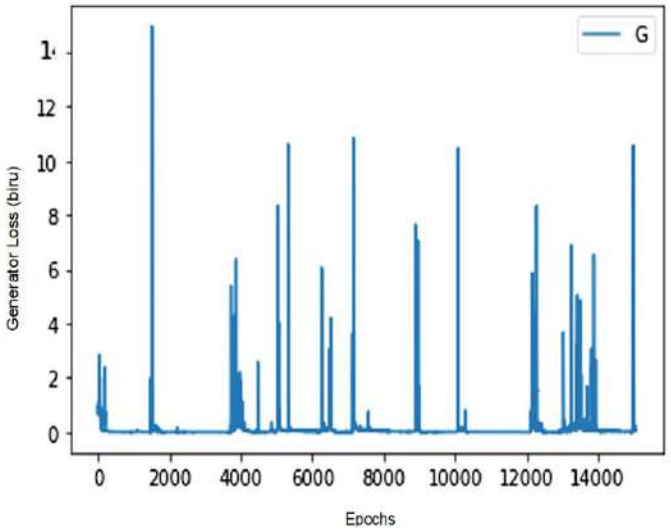
Gambar 2. Diagram alir proses latih jaringan *image completion*

III. HASIL DAN PEMBAHASAN

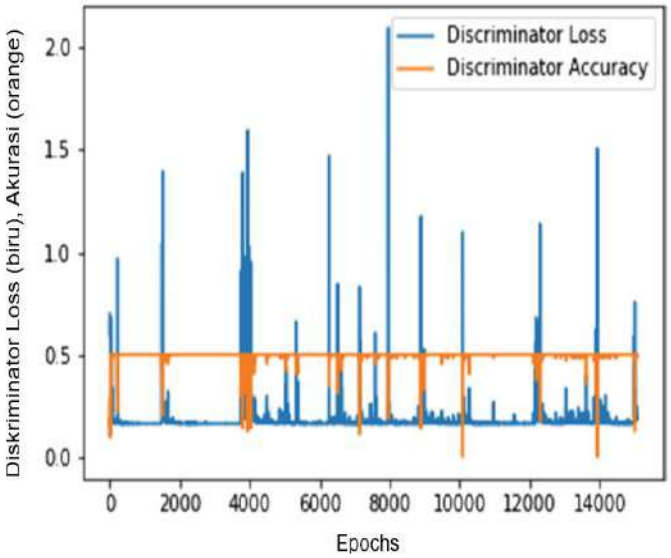
A. *Model dan Dataset*

Untuk *image completion* ini, model yang digunakan adalah model pra-latih *Completion Net* yang diimplementasi melalui proses *transfer learning*. Kemudian model ini ditambahkan dengan dataset open-source *Facescrub* dari winklebros.net [21]. Dataset ini terdiri dari 100.000 citra yang berasal dari 500 wajah *public figure* yang unik. Model dilatih dengan dataset ini dengan tujuan untuk membuat kemampuan model dalam memahami atribut wajah semakin baik. Citra berkualitas tinggi sangat berperan penting dalam proses latih dan dan memberikan generalisasi yang baik. Dari dataset ini citra di*crop* dan di*resize* ke dimensi 64 x 64 piksel untuk membatasi hanya wajah saja yang tersimpan untuk setiap citra. Model akan menerima inputan citra dan mask (sebagai bagian yang hilang dari citra) dan akan mencoba memprediksi dan melengkapi bagian yang hilang tersebut berdasarkan informasi yang diberikan oleh citra yang tidak lengkap tersebut. Dataset dilatih pada komputer dengan GPU NVIDIA GTX 960M 4GB hingga 13000 *epoch* dimana setiap *epoch* terdiri dari 1000 batch citra dari dataset. Gambar 3 menunjukkan hasil dari generator yang dihitung dengan fungsi loss diatas 13000 epoch. Dari gambar tersebut juga terlihat peningkatan seragam pada interval reguler dari nilai *loss* yang cukup sesuai untuk menggambarkan kehandalan dari model. Puncak dari nilai *loss* secara bertahap berkurang yang menandakan pembelajaran model yang terjadi secara bertahap. Menggunakan jumlah *epoch* yang sangat tinggi, proses latih butuh waktu sekitar 38 jam untuk diselesaikan. Generator dan diskriminator mengikuti mekanisme paralel karena waktu training meningkat secara eksponensial.

Nilai akurasi diskriminator konstan disepanjang epoch. Hanya terjadi reduksi minor pada diskriminator yang disebabkan oleh arsitektur model. Nilai kumulatif rata-rata loss meningkat menunjukkan ketika jumlah data meningkat pada sisi latih maka kemampuan model untuk mempelajari pola menjadi berkurang. Gambar 4 menunjukkan hasil yang diperoleh dari nilai loss diskriminator.



Gambar 3. Generator loss terhadap jumlah epochs



Gambar 4. Diskriminator loss dan akurasi terhadap epochs

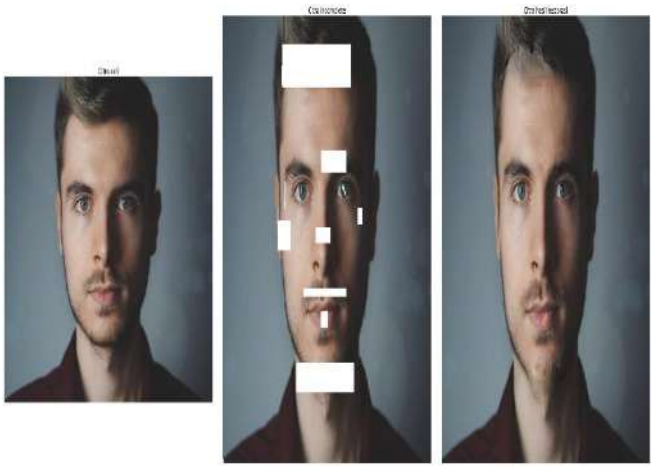
B. *Hasil Perbandingan Visual*

Gambar 5 dan Gambar 6 menampilkan hasil restorasi citra yang menyajikan hasil perbandingan antara gambar asli, *masking* dan gambar hasil *completion*, berdasarkan perbandingan visual yang diperoleh model yang digunakan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023

dapat mengisi atau melengkapi bagian hilang dari citra dengan hasil yang cukup baik. Untuk pengujian visual peneliti menggunakan 10 responden yang dipilih secara acak melalui media sosial, setelah responden menyatakan kesediaan untuk berpartisipasi, maka peneliti mengirimkan 2 sampel citra hasil restorasi yaitu gambar 5 dan gambar 6. Setiap responden diberi pertanyaan apakah dapat mengenali gambar yang diberikan dan apakah terdapat kecacatan visual pada gambar yang mempengaruhi persepsi responden. Dari 10 orang responden yang diberi pertanyaan tersebut, seluruhnya dapat memahami objek pada gambar hasil restorasi dan mengaku dapat mengenali semua objek pada gambar hasil restorasi tanpa ada kecacatan yang mempengaruhi persepsi visual responden.

Gambar 5. Hasil restorasi citra pemandangan dengan *image completion*



Gambar 6. Hasil restorasi citra wajah dengan *image completion*

Dari hasil yang diperoleh, model dapat digunakan untuk merestorasi citra dengan beragam resolusi tetapi berdasarkan pengujian yang dilakukan semakin besar ukuran celah loss informasi pada citra maka akan semakin berkurang kualitas pengisian yang diperoleh bahkan dapat menyebabkan kegagalan untuk melakukan pengisian. Hal ini karena keterbatasan dari model arsitektur CNN yang digunakan, tetapi dengan meningkatkan ukuran layer akan dimungkinkan untuk mendapatkan hasil pengisian yang lebih baik tetapi dengan timbal balik waktu pemrosesan yang lebih lama.

C. Pembahasan dan Limitasi

Berdasarkan pengujian model yang digunakan dapat mengisi celah dengan ukuran bervariasi, tetapi untuk celah yang berukuran besar belum dapat diisi dengan baik. Limitasi ini hanya berlaku pada mask berbentuk persegi, area yang luas pada persegi masih dapat diisi selama tingginya tidak terlalu besar karena informasi dari bagian atas dan bawah citra akan digunakan untuk melengkapi gambar. Sehingga bila celah dalam citra terlalu besar atau terlalu tinggi maka akan sangat membatasi proses ekstrapolasi citra dimana mask akan berada di batas dari citra. Hal ini akan menyebabkan kegagalan dalam pengisian celah citra. Pendekatan lain yang memanfaatkan dictionary learning dapat mengatasi masalah ini sepanjang database citra yang digunakan mengandung kemiripan dengan citra input. Untuk pendekatan ini ekstrapolasi lebih mudah daripada inpainting karena lebih sedikit yang harus dicocokkan pada tepian citra akan tetapi metode ini lebih cocok digunakan untuk menghasilkan ekstensi citra dibanding inpainting yang berfokus untuk melakukan restorasi pada citra yang mengalami kerusakan Beberapa contoh kegagalan pengisian gambar dapat dilihat pada gambar 7 dan 8. Kegagalan pengisian yang terjadi umumnya terjadi pada citra yang memiliki objek yang kompleks yaitu citra yang didalamnya hanya memiliki sedikit pola informasi yang berulang dan atau citra memiliki celah yang menutupi keseluruhan dari bagian objek yang dapat digunakan untuk pembelajaran sehingga akhirnya model gagal untuk memperoleh informasi untuk mengisi celah pada citra tersebut. Pada gambar 7 model gagal untuk mengisi gambar kepala pada gambar anak yg bagian kepalanya terhapus karena lebih memprioritaskan untuk merekonstruksi citra pohon pada latar belakang dibandingkan kepala dari orang dan pada gambar 8 model gagal untuk mengisi gambar anjing.

Gambar 7. Kegagalan pengisian citra manusia

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023

Gambar 8. Kegagalan pengisian citra hewan

IV. KESIMPULAN

Bagian Penelitian yang dilakukan untuk melakukan restorasi citra dengan metode *image completion* berbasis deep learning telah berhasil memperoleh hasil yang baik berdasarkan uji visual yang dilakukan kepada responden. Dengan menggunakan jaringan penyelesaian dan diskriminator, model yang digunakan mampu menghasilkan pengisian celah citra yang realistis. Dibandingkan dengan metode umum berbasis tambalan citra, model ini dapat mengisi bagian citra menggunakan bagian baru yang tidak pernah digunakan dalam citra. Limitasi dari model yang dikembangkan saat ini terkait dengan arsitektur yang digunakan sehingga untuk ukuran celah yang terlalu besar akan mengurangi kualitas pengisian yang dilakukan pada citra sehingga menyebabkan gambar tidak terlengkapi dengan sempurna. Penelitian di masa depan dapat dilakukan untuk mengembangkan arsitektur yang digunakan dengan ukuran layer jaringan yang lebih besar untuk mengatasi limitasi ini.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada LPPM Universitas Atma Jaya Makassar yang telah memberikan dukungan dana untuk terlaksananya penelitian ini melalui hibah penelitian dosen LPPM UAJM

REFERENCES

[1] P. Chyan and N. T. Saptadi, "Pemulihan Citra Berbasis Metode Markov Random Field," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 218, Apr. 2022.

[2] P. Chyan, "PENERAPAN IMAGE ENHANCEMENT ALGORITHM UNTUK MENINGKATKAN KUALITAS CITRA TAK BERGERAK," *Inf. dan Teknol. Ilm.*, vol. 4, no. 3, May 2017.

[3] P. Chyan, "Rancang Bangun Mesin Pencari Citra Dengan Pendekatan Temu Balik Berbasis Konten," *TEMATIKA*, vol. 3, no. 2, 2015.

[4] J. Bin Huang, S. B. Kang, N. Ahuja, and J. Kopf, "*Image completion* using planar structure guidance," *ACM Trans. Graph.*, vol. 33, no. 4, 2014.

[5] G. Chican and M. Tamaazousti, "Constrained PatchMatch for *image completion*," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8887, pp. 560–568, 2014.

[6] J. Sun, L. Yuan, J. Jia, and H. Y. Shum, "*Image completion* with structure propagation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 861–868, Jul. 2005.

[7] C. Zheng, T. J. Cham, J. Cai, and D. Phung, "Bridging Global Context Interactions for High-Fidelity *Image completion*," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 11502–11512, 2022.

[8] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent *image completion*," *ACM Trans. Graph.*, vol. 36, no. 4, 2017.

[9] Q. Li, L. Hu, Q. Shang, Y. Wang, L. Jiang, and W. Long, "Research Progress in the Field of *Image completion*," *Proc. - 2021 3rd Int. Conf. Artif. Intell. Adv. Manuf. AIAM 2021*, pp. 398–402, 2021.

[10] P. Aditya Laksana, "Restorasi Pada Citra Digital Menggunakan Pendekatan Exemplar Based Inpainting," Oct. 2014.

[11] A. H. Hasugian and I. Zufria, "Perancangan Sistem Restorasi Citra Dengan Metode Image Inpainting," *Algoritm. J. ILMU Komput. DAN Inform.*, vol. 2, no. 2, p. 31, Jan. 2019.

[12] M. Failzal and A. Rao, "'RECOGNITION of STRESS USING FACE IMAGE and FACIAL LANDMARKS,'" *Int. Adv. Res. J. Sci. Eng. Technol.*, vol. 8, no. 5, pp. 792–797, 2021.

[13] P. Chyan, "Image Enhancement Based On Bee Colony Algorithm," *J. Eng. Appl. Sci.*, vol. 14, no. 1, pp. 43–49, 2019.

[14] J. Yang, R. Xu, Z. Qi, and Y. Shi, "Visual Anomaly Detection for Images: A Systematic Survey," *Procedia Comput. Sci.*, vol. 199, pp. 471–478, 2021.

[15] P. Chyan, "Sistem Temu Balik Citra Menggunakan Ekstraksi Fitur Citra Dengan Klasifikasi Region Untuk Identifikasi Objek," *TEMATIKA*, vol. 2, no. 2, pp. 63–72, 2014.

[16] N. Akimoto, S. Kasai, M. Hayashi, and Y. Aoki, "360-Degree *Image completion* by Two-Stage Conditional Gans," *Proc. - Int. Conf. Image Process. ICIP*, vol. 2019-September, pp. 4704–4708, Sep. 2019.

[17] Q. Chen, G. Li, Q. Xiao, L. Xie, and M. Xiao, "*Image completion* via transformation and structural constraints," *Eurasip J. Image Video Process.*, vol. 2020, no. 1, pp. 1–18, Dec. 2020.

[18] L. Theis, A. Van Den Oord, and M. Bethge, "A note on the evaluation of generative models," *4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc.*, 2016.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[20] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5967–5976, Nov. 2017.

[21] H. W. Ng and S. Winkler, "A data-driven approach to cleaning large face datasets," *2014 IEEE Int. Conf. Image Process. ICIP 2014*, pp. 343–347, 2014.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1699, Copyright ©2023
Submitted : May 5, 2023, Revised : June 8, 2023, Accepted : July 5, 2023, Published : November 3, 2023

Deteksi Dini Penyakit Alzheimer dengan Algoritma C4.5 Berbasis BPSO (*Binary Particle Swarm Optimization*)

Anistya Rosyida[1], Theopilus Bayu Sasongko[2]*
Program Studi Informatika, Fakultas Ilmu Komputer[1], [2]
Universitas Amikom Yogyakarta
Yogyakarta, Indonesia
anistya.88@students.amikom.ac.id[1], theopilus.27@amikom.ac.id[2]

***Abstract*— Alzheimer's disease is a degenerative disease associated with memory loss, communication difficulties, mental health, thinking skills, and other psychological disorders that affect a person's daily activities. Alzheimer's disease is a disease that causes disability for people aged 70 years and over and is the seventh highest contributor to death in the world. However, until now there has not been found an effective treatment to cure Alzheimer's disease. Thus, early detection of Alzheimer's disease is very important so that sufferers of Alzheimer's disease can immediately receive intensive medical care so as to reduce the death rate from Alzheimer's disease. One method that can be used to detect Alzheimer's disease is by utilizing a machine learning algorithm model. The machine learning model in this study was carried out using the Decision Tree C4.5 algorithm classification method based on Binary Particle Swarm Optimization (BPSO). The C4.5 Decision Tree algorithm is used to classify Alzheimer's disease, while the BPSO algorithm is used to perform feature selection. By performing feature selection with the BPSO algorithm, the results show that the BPSO algorithm can improve accuracy and can increase the performance of the C4.5 algorithm in the Alzheimer's disease classification process. The results of the accuracy of the C4.5 algorithm using the BPSO feature selection are greater, namely 98.2% compared to the C4.5 algorithm without BPSO feature selection, which is only 96.4%.**

Keywords*— *Alzheimer's Disease, Dementia, Classification, Decision Tree C4.5, BPSO

***Abstrak*—Penyakit Alzheimer merupakan suatu penyakit degenerative yang berkaitan dengan penurunan daya ingat, kesulitan berkomunikasi, kesehatan mental, keterampilan berpikir, dan kecakapan psikologis lainnya yang memengaruhi seseorang dalam melakukan aktivitas sehari-hari. Penyakit Alzheimer merupakan penyakit yang menjadi penyebab kecacatan bagi manusia yang berusia 70 tahun ke atas dan merupakan penyumbang kematian tertinggi ketujuh di dunia. Meskipun demikian, sampai detik ini belum juga ditemukan pengobatan efektif untuk menyembuhkan penyakit Alzheimer. Dengan demikian, deteksi dini penyakit Alzheimer sangat penting agar para penderita penyakit Alzheimer dapat segera mendapatkan perawatan medis yang intensif sehingga dapat mengurangi angka kematian akibat penyakit Alzheimer. Salah satu metode yang dapat dipergunakan untuk deteksi penyakit Alzheimer ini yaitu dengan memanfaatkan model algoritma machine learning. Model machine learning pada penelitian ini dilakukan dengan mempergunakan metode klasifikasi algoritma *Decision Tree C4.5* berbasis *Binary Particle Swarm Optimization* (BPSO). Algoritma *Decision Tree* C4.5 dipergunakan untuk melakukan klasifikasi penyakit Alzheimer, sedangkan algoritma BPSO dipergunakan untuk melakukan seleksi fitur. Dengan melakukan seleksi fitur dengan algoritma BPSO, didapatkan hasil bahwa algoritma BPSO dapat meningkatkan hasil akurasi dan dapat menambah performa dari algoritma C4.5 pada proses klasifikasi penyakit Alzheimer. Hasil *accuracy* algoritma C4.5 dengan menggunakan seleksi fitur BPSO lebih besar yaitu 98,2% dibandingkan dengan algoritma C4.5 tanpa seleksi fitur BPSO yaitu hanya sebesar 96,4%.**

Kata Kunci*— *Penyakit Alzheimer, Dementia, Klasifikasi, Decision Tree C4.5, BPSO

I. PENDAHULUAN

Penyakit Alzheimer merupakan sebuah penyakit penyebab kecacatan bagi manusia berusia 70 tahun ke atas dan merupakan penyumbang kematian tertinggi ketujuh di dunia [1]. Penyakit Alzheimer merupakan sebuah penyakit yang dapat memperburuk kondisi seseorang dari waktu ke waktu dan biasanya ditandai dengan penurunan daya ingat seseorang. Penyakit Alzheimer merupakan sebuah bentuk demensia paling umum yang terjadi di dunia yaitu mencapai sekitar 60% - 70% kasus [2]. Dementia merupakan sebuah penyakit yang ditandai dengan penurunan kemampuan dalam meningat, berkomunikasi, dan juga beraktivitas. Penyakit Alzheimer merupakan penyakit yang menyerang system saraf pada otak manusia dimana secara permanen dapat menyebabkan hilangnya sel neuron dan berakibat pada terhalangnya manusia dalam menjalankan aktivitas sehari-hari [3]. Penyakit Alzheimer ini juga dapat menyebabkan lenyapnya kecakapan mengingat, kesulitan dalam berkomunikasi, berpikir jernih, dan lebih parahnya adalah terjadinya peralihan perilaku serta kecakapan untuk mengurusi diri sendiri [4]. Dengan demikian, Alzheimer dan dementia merupakan dua hal yang berbeda akan tetapi saling berkaitan.

Penyakit demensia tipe Alzheimer saat ini memengaruhi sekitar 50 juta orang diseluruh dunia, dimana jumlahnya

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

diperkirakan akan terus bertambah dikemudian hari hingga mencapai 82 juta pada tahun 2030 dan bertambah lagi menjadi 152 juta pada tahun 2050 [5]. Prevalensi dari demensia tipe Alzheimer, prevalensinya akan terus bertambah beriringan dengan pertambahan usia. Prevalensi dementia tipe Alzheimer bagi perempuan lebih tinggi dibandingkan dengan prevalensi dementia tipe Alzheimer bagi laki-laki. Prevalensi demensia tipe Alzheimer bagi laki-laki berumur 65 tahun, prevalensinya sekitar 0,6%, sedangkan pada perempuan prevalensi sekitar 0,8% [6]. Meskipun demikian, bukan berarti bahwa orang yang berusia dibawah 65 tahun tidak dapat menderita penyakit Alzheimer. Factor pemicu Alzheimer bagi orang berusia dibawah 65 tahun didasari pada kelainan genetic, hipertensi, depresi, kebiasaan merokok, obesitas, serta factor aktivitas sosial [7], [8].

Pengobatan penyakit Alzheimer hingga saat ini masih dilakukan dengan cara tradisional yaitu dengan melakukan terapi [9]. Berbagai macam jenis terapi telah dilakukan oleh para tenaga ahli kesehatan sebagai upaya untuk menyembuhkan penyakit Alzheimer. Akan tetapi, hingga saat ini masih belum dapat dipastikan apakah pengobatan dengan cara tersebut sudah efektif atau masih tergolong berbahaya. Saat ini, banyak juga instrumen diagnostic dan pengobatan seperti terapi baru yang dikembangkan, tetapi masih terbatas pada konteks penelitian saja [10]. Hal ini menjadikan bukti bahwa hingga detik ini pengobatan yang tepat untuk mengobati penyakit Alzheimer belum juga ditemukan. Dengan demikian, dibutuhkan sebuah system yang dapat melakukan klasifikasi penyakit Alzheimer yang dapat mendeteksi secara dini penyakit Alzheimer dan mampu untuk mendeteksi perkembangan penyakit Alzheimer selama tahap praklinis [11]. Sehingga harapannya adalah dapat memperlambat gejala keparahan yang dialami oleh para penderita penyakit Alzheimer dan dapat mengurangi angka kematian akibat penyakit Alzheimer.

Penelitian mengenai deteksi penyakit Alzheimer sudah banyak dilakukan sebelumnya. Para peneliti telah berusaha untuk mencari metode-metode yang bisa dipergunakan untuk mendeteksi dengan cepat dan menghasilkan akurasi yang tinggi. Salah satu metode yang dapat digunakan untuk deteksi dini penyakit Alzheimer adalah dengan memanfaatkan model machine learning metode klasifikasi. Penelitian mengenai deteksi dini penyakit Alzheimer dengan model machine learning metode klasifikasi sudah pernah dilakukan sebelumnya. Penelitian tersebut membandingkan algoritma *Decision Tree*, algoritma *Logistic Regression*, algoritma *Naïve Bayes*, algoritma *Random Forest*, dan algoritma *Support Vector Machine* dimana algoritma *Support Vector Machine* menghasilkan nilai akurasi tertinggi dan terbaik dibandingkan dengan empat algoritma lain [12]. Algoritma *Support Vector Machine* dapat lebih baik dibandingkan dengan algoritma *Decision Tree* dikarenakan algoritma tersebut dapat memberikan hasil terbaik secara keseluruhan pada semua metriknya. Akan tetapi, pengklasifikasian pada algoritma *Decision Tree* juga dapat mengidentifikasi fitur yang paling diskriminatif dan penting untuk deteksi penyakit Alzheimer [13]. Selain itu, pada penelitian tersebut juga didapatkan hasil dari model algoritma *Decision Tree* lebih mudah untuk dipahami dan diinterpretasikan [14]. Berdasarkan pada penelitian yang dilakukan oleh Bari Antor di atas, algoritma *Decision Tree* tidak memiliki performa yang cukup baik. Akan tetapi, apabila algoritma *Decision Tree* dilakukan seleksi fitur dengan algoritma PSO, maka hal tersebut dapat meningkatkan performa dari algoritma *Decision Tree* dengan ditandai bertambahnya nilai akurasi yang dihasilkan [15]. Algoritma PSO merupakan algoritma optimasi yang efektif untuk memecahkan masalah pada algoritma *Decision Tree*. Hal ini dikarenakan algoritma PSO dapat memilih beberapa atribut fitur yang relevan dengan iterasi minimum [16].

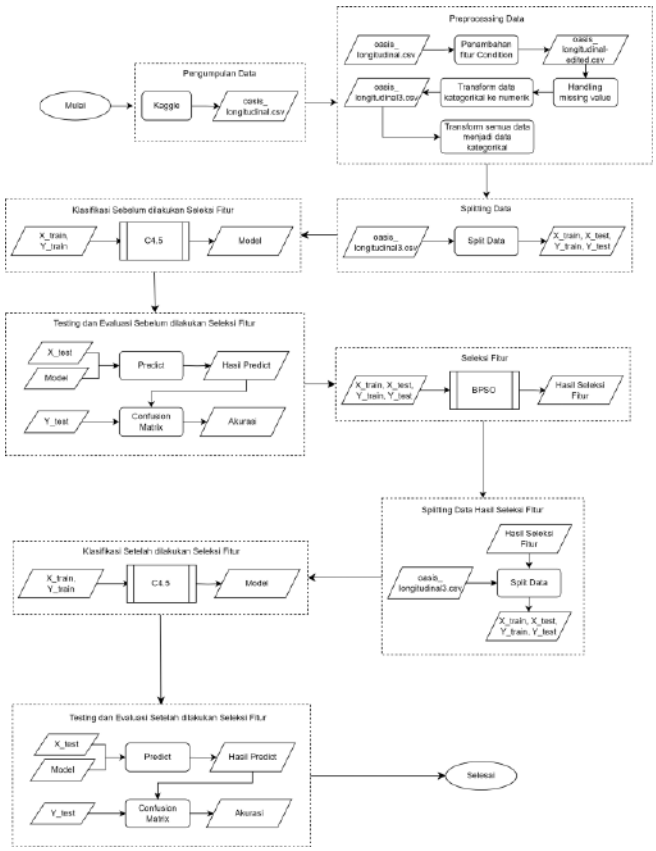
Penelitian tentang klasifikasi penyakit Alzheimer dengan menggunakan beberapa model dari algoritma *Decision Tree* dengan ditambahkan seleksi fitur berbasis PSO juga pernah dilakukan sebelumnya. Penelitian tersebut menggunakan beberapa model dari algoritma *Decision Tree* diantaranya yaitu algoritma C4.5, ID3, CHAID, dan *Random Forest* dengan berbasis PSO. Hasil performa terbaik yang didapatkan dari penelitian tersebut adalah dengan menggunakan algoritma *Decision Tree* model *Random Forest* dengan hasil akurasi sebelum dilakukan seleksi fitur dengan PSO adalah sebesar 91.15%, sedangkan setelah dilakukan seleksi fitur dengan PSO didapatkan hasil akurasi sebesar 93.56% [15]. Meskipun algoritma PSO telah terbukti efektif untuk mengoptimalkan algoritma *Decision Tree*, akan tetapi algoritma PSO perlu dimodifikasi menjadi algoritma *Binary* PSO (BPSO) agar dapat lebih efektif untuk menangani data diskrit dan dapat memilih fitur-fitur yang paling relevan untuk proses klasifikasi penyakit Alzheimer.

Tujuan dari dibuatnya penelitian ini adalah untuk berkontribusi dalam bidang keilmuan mengenai deteksi dini penyakit Alzheimer menggunakan algoritma C4.5 berbasis BPSO dan diharapkan dapat memberikan gambaran untuk pengembangan model sebagai acuan pada penelitian selanjutnya. Pada pembuatan model untuk deteksi dini penyakit Alzheimer ini dilakukan pada dataset *oasis_longitudinal* yang didapatkan dari website Kaggle dengan berupaya untuk mencari nilai akurasi terbaik dengan mempergunakan algoritma *Decision Tree* C4.5 berbasis BPSO. Data-data dari dataset tersebut diolah dan direduksi atributnya dengan menggunakan algoritma BPSO, kemudian dibuat model klasifikasinya dengan mempergunakan algoritma *Decision Tree* C4.5. Penggunaan algroritma BPSO diharapkan dapat meningkatkan performa dari algoritma C4.5 untuk melakukan klasifikasi penyakit Alzheimer. Dengan demikian, diharapkan pasien yang telah didiagnosa menderita penyakit Alzheimer lebih awal dapat segera mendapatkan penanganan medis sedini mungkin, sehingga kematian yang disebabkan oleh penyakit Alzheimer dapat berkurang, mengingat hingga pada saat ini masih belum juga didapati metode penyembuhan yang tepat untuk penyakit Alzheimer.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

II. METODE PENELITIAN

Penelitian ini nantinya akan membuat sebuah model *machine learning* yang digunakan untuk mengklasifikasikan penyakit Alzheimer. Algoritma yang akan digunakan untuk mengklasifikasikan penyakit Alzheimer ini adalah algoritma C4.5 dan BPSO (*Binary Particle Swarm Optimization*). Algoritma C4.5 dipergunakan untuk melakukan klasifikasi penyakit Alzheimer, sedangkan algoritma BPSO dipergunakan pada proses seleksi fitur dengan harapan agar dapat membentuk akurasi yang terbaik dan dapat menambah performa proses klasifikasi. Metode penelitian dilaksanakan dengan beberapa tahapan diantaranya yaitu pengumpulan data, *preprocessing* data, *splitting* data, klasifikasi dengan algoritma C4.5 sebelum dilakukan seleksi fitur, *testing* dan evaluasi sebelum dilakukan seleksi fitur, seleksi fitur mempergunakan algoritma BPSO, *splitting* data hasil seleksi fitur, klasifikasi dengan algoritma C4.5 setelah dilakukan seleksi fitur, dan terakhir adalah testing dan evaluasi setelah dilakukan seleksi fitur. Gambar 1 menampilkan alur dari penelitian yang dilakukan.



Gambar 1. Alur Penelitian

Pada gambar 1 alur penelitian di atas, dapat dilihat bahwa penelitian diawali dari tahapan pengumpulan data yang didapatkan dari website Kaggle. Kemudian, dilanjutkan ke tahapan setelahnya yaitu *preprocessing* data. Pada *preprocessing* data, data dilakukan penambahan satu kolom lagi yaitu kolom *Condition* untuk mempermudah proses klasifikasi, kemudian dilakukan penanganan pada *missing value*-nya, kemudian mengubah data-data yang bertipe objek diubah menjadi numerik, dan yang terakhir adalah mengubah keseluruhan data menjadi data kategorikal dengan interval data. Selanjutnya, masuk pada tahap *splitting* data yaitu memisahkan data menjadikan dua bagian yaitu data *training* dan data *testing*. Kemudian, tahapan selanjutnya adalah klasifikasi dengan algoritma C4.5 sebelum dilakukan seleksi fitur, dan dilanjutkan dengan *testing* dan evaluasi sebelum dilakukan seleksi fitur. Tahapan selanjutnya adalah seleksi fitur dengan algoritma BPSO, kemudian setelah dilakukan seleksi fitur, tahapan selanjutnya adalah membagi data hasil seleksi fitur menjadi data *training* dan data *testing* lagi. Kemudian, baru dijalankan proses klasifikasi mempergunakan algoritma C4.5 setelah dilakukan seleksi fitur, dan yang terakhir adalah tahap *testing* dan evaluasi setelah dilakukan seleksi fitur. Pada tahap *testing* dan evaluasi baik sebelum maupun sesudah dilakukan seleksi fitur, dilakukan dengan menggunakan metode *confusion matrix*.

A. Pengumpulan Data

Dataset yang dipergunakan pada penelitian ini merupakan data sekunder yang bernama *Oasis Longitudinal* yang didapatkan dari *website Kaggle Open Datasets and Machine Learning Project*. Dataset terdiri dari 373 baris dan 15 kolom dengan rincian sebanyak 341 data menderita penyakit Alzheimer dan sisanya sebanyak 32 data tidak menderita penyakit Alzheimer. Dataset ini tergolong pada distribusi data yang tidak seimbang. Hal ini dikarenakan terdapat perbedaan yang signifikan antara jumlah penderita penyakit Alzheimer dengan yang tidak menderita penyakit Alzheimer. Selain itu, dataset ini juga masih tergolong ke dalam jumlah data yang kecil atau sedikit pada sebuah *machine learning.*

B. Preprocessing Data

Setelah dilakukan pengumpulan data, tahap selanjutnya adalah melakukan *preprocessing* data. *Preprocessing* data di sini berfungsi untuk meminimalisir kesalahan dan mengoptimalkan hasil dari proses klasifikasi. Tahap pertama dari *preprocessing* data ini adalah dengan menambahkan satu kolom baru pada dataset yaitu kolom *Condition* untuk memperjelas kolom ASF. Kolom ASF merupakan kolom yang berisi tentang hasil penentuan pasien terdeteksi penyakit Alzheimer atau tidak. Akan tetapi, dikarenakan nilai dari kolom tersebut bertipe data float, maka hal tersebut dapat menyulitkan *machine* pada proses klasifikasi dan menyulitkan para pembaca untuk memahaminya. Oleh karena itu, ditambahkan satu kolom lagi yaitu kolom *Condition* yang bertipe data object dan nantinya akan di-*convert* menjadi tipe data numerik dengan harapan dapat mempermudah *machine* pada proses klasifikasi dan memudahkan para pembaca untuk memahami dataset tersebut. Sehingga pada penelitian ini dataset yang dipergunakan adalah data *Oasis Longitudinal* baru dengan nama *oasis_longitudinal-edited* yang terdiri dari 373 baris dan 16 kolom dengan rincian sebanyak 341 data menderita penyakit Alzheimer dan sisanya sebanyak 32 data tidak menderita penyakit Alzheimer. Dataset ini tergolong pada distribusi data yang tidak seimbang. Hal ini dikarenakan terdapat perbedaan yang signifikan antara jumlah penderita penyakit Alzheimer dengan yang tidak menderita penyakit Alzheimer. Table I

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

menampilkan fitur-fitur dan deskripsi dari fitur-fitur dalam dataset.

TABLE I. DATASET OASIS LONGITUDINAL

Features	Description
Subject ID	ID subject
MRI ID	ID MRI
Group	Demented / Nondemented
Visit	Total visit
MR Delay	Delayed enhancement MRI (DE-MRI)
M/F	Gender
Hand	Right / Left
Age	Person's age
EDUC	Years of education
SES	Socioeconomic status
MMSE	Mini-mental state examination
CDR	Clinical Dementia Rating
eTIV	Estimated total intracranical volume
nWBV	Normalized whole brain volume
ASF	Atlas scaling factor
Condition	Condition patient positive dementia or negative dementia

Tahap *preprocessing* data yang kedua adalah mengatasi *missing value*. Pada dataset *Oasis Longitudinal* ini terdapat 19 *missing value* pada fitur SES dan 2 *missing value* pada fitur MMSE. Cara yang digunakan untuk mengatasi *missing value* ini adalah dengan mengisikan data yang kosong dengan mempergunakan nilai rata-rata dari masing-masing kolom yang terdapat *missing value*. Selanjutnya, tahap ketiga dari *preprocessing* data adalah mengubah fitur-fitur dengan tipe data *object* menjadi *numerik*. Fitur-fitur tersebut diantaranya adalah *Subject ID, MRI ID, Hand, Grup, M/F,* dan *Condition*. Kemudian, tahap terakhir dari *preprocessing* data ini adalah mengubah keseluruhan data dari tipe data *numerik* diubah menjadi data *kategorikal* dengan menggunakan interval data. Meskipun algoritma *Decision Tree* dapat digunakan untuk melakukan klasifikasi baik pada data *kategorikal* maupun *numerik*, akan tetapi agar dapat menghasilkan performa yang lebih optimal, maka data perlu diubah menjadi bertipe kategorikal.

C. Splitting Data

Pada tahap *splitting* data, dataset dibagi dua membentuk data *training* dan data *testing*. Rasio perbandingan pembagian data adalah 70:30. Dengan rinciannya adalah 70% atau sebanyak 261 data dipergunakan untuk data *training*, dan sisanya 30% atau sebanyak 112 data dipergunakan untuk data *testing*. Hasil dari pembagian data *training* dan data *testing* ini diantaranya adalah data X_train, Y_train, X_test, dan Y_test. Data X merupakan data atribut, sedangkan data Y merupakan data target.

Alasan dipilihnya rasio perbandingan data *training* dan data *testing* 70:30 adalah adanya keterbatasan dataset yang dipergunakan. Pada penelitian ini dataset yang dipergunakan merupakan dataset yang tergolong kecil atau sedikit dan hanya berjumlah 373 data saja. Dengan membuat data *testing* yang cukup besar dengan mengambil sebanyak 30% data, dapat menghasilkan performa yang lebih baik jika diperbandingkan dengan rasio perbandingan data *training* dan data *testing* 80:20 serta 90:10. Hal ini juga sesuai dengan penelitian yang telah dilaksanakan sebelumnya dengan mempergunakan dataset yang sama, didapatkan hasil bahwa rasio perbandingan data *training* dan data *testing* 70:30 memberikan kinerja model yang lebih baik dibandingkan dengan rasio 80:20 [12] dan [9].

D. Feature Selection dengan BPSO

Binary Particle Swarm Optimization (BPSO) merupakan hasil modifikasi dari algoritma sebelumnya yaitu *Particle Swarm Optimization* (PSO) dimana algoritma ini menggunakan metode pendekatan meta-heuristik yang terinspirasi dari sekumpulan burung dan ikan [17]. Algoritma PSO dikembangkan menjadi algoritma *Binary* PSO (BPSO) bertujuan untuk memudahkan dalam melakukan seleksi fitur. Hal ini dikarenakan pada data diskrit, seleksi fitur dengan algoritma BPSO lebih efektif karena dilakukan dengan menggunakan nilai biner dan dapat memilih fitur-fitur yang paling relevan. Alasan dipilihnya algoritma BPSO dikarenakan algoritma ini memiliki perilaku konvergensi yang cepat, fleksibel, dan efisien [18]. Algoritma BPSO dapat lebih cepat dan efisien karena algoritma ini menggunakan nilai biner (1 dan 0) sehingga dapat memudahkan *machine* untuk melakukan pemilihan fitur dalam waktu *relative* singkat.

Tahapan seleksi fitur dengan algoritma BPSO adalah sebagai berikut [19]:

1) Menginisiasi nilai parameter yang dibutuhkan, nilai batas iterasi, *cognitive learning* (c1), *social learning* (c2), dan *inertia weight* (w).

2) Menginisiasi posisi dan kecepatan awal dari seluruh partikel. Pembentukan kecepatan awal partikel dibuat dengan nilai *default* 0, kemudian pada penentuan posisi awal partikel ditentukan *random* dengan nilai 0 atau 1. Inisiasi partikel ini dipergunakan untuk menyeleksi fitur-fitur yang akan dipergunakan untuk klasifikasi dengan algoritma C4.5.

3) Memperkirakan nilai *fitness* setiap partikel dengan menjalankan klasifikasi mempergunakan algoritma C4.5. Partikel yang nilainya 1 akan dipergunakan untuk klasifikasi, sedangkan partikel yang bernilai 0 tidak digunakan. Sehingga diperoleh hasil *accuracy* dari algoritma C4.5 merupakan nilai *fitness* nya.

4) Menetapkan nilai Pbest dan Gbest. Nilai Pbest akan di-*design* persis dengan nilai posisi awal partikel. Sedangkan pada iterasi setelahnya, Pbest ditetapkan dengan memilih nilai *fitness* paling tinggi dari posisi partikel pada tiap iterasinya. Kemudian nilai Gbest ditetapkan dengan mengacu partikel pada Pbest dengan nilai *fitness* paling tinggi.

5) Pengecekan kondisi selesai berdasar pada jumlah iterasi yang sudah dilalui. Apabila jumlah perulangan belum sampai batas maksimum iterasi, maka akan diteruskan ke tahapan *update* kecepatan dan penentuan posisi baru. Akan tetapi, apabila jumlah perulangan telah mencapai maksimum

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

iterasi, maka telah diperoleh hasil seleksi fitur paling baik dari nilai Gbest di iterasi paling akhir.

6) Pembaruan kecepatan untuk menetapkan arah kemana partikel akan melakukan perpindahan dan juga untuk membenarkan posisi awal. Persamaan (1) merupakan rumus untuk update kecepatan pada algoritma BPSO.

$$v_{id}^{new} = w * v_{id}^{old} + c_1 r_1 \left(pb_{id}^{old} - x_{id}^{old}\right) + c_2 r_2 \left(gb_{id}^{old} - x_{id}^{old}\right) \quad (1)$$

Keterangan:

v_{id}^{new} : *velocity* partikel baru
v_{id}^{old} : *velocity* partikel lama
x_{id}^{old} : titik partikel lama
$r_1 \; dan \; r_2$: nilai random antara 0 atau 1
c_1 : *factor cognitive learning*
c_2 : *factor social learning*
w : bobot inersia
gb : Gbest (*global best*)
pb : Pbest (*personal best*)

7) Penentuan posisi baru. Posisi partikel pada BPSO sesuai namanya yaitu harus di-*convert* ke dalam nilai biner dengan interval [0,1]. Agar dapat memberikan batasan nilai kecepatan dari setiap partikel, maka perlu dilaksanakan transformasi *limiting*. Persamaan (2) merupakan rumus pada transformasi *limiting*. Sedangkan (3) merupakan rumus sigmoid.

$$x_{id}^{new} = \begin{cases} 1, & sigmoid\ (v_{id}^{new}) > rand \\ 0, & sigmoid\ (v_{id}^{new}) \leq rand \end{cases} \quad (2)$$

Keterangan:

x_{id}^{new} : titik partikel baru
v_{id}^{new} : *velocity* partikel baru
$rand$: nilai random antara 0 dan 1

$$sigmoid\ (v_{id}^{new}) = \frac{1}{1 + e^{-v_{id}^{new}}} \quad (3)$$

8) Apabila telah sampai pada iterasi maksimal, maka hasil dari seleksi fitur paling baik sudah didapatkan dari partikel Gbest di iterasi paling akhir.

E. Klasifikasi dengan Algoritma C4.5

Algoritma C4.5 merupakan penyempurnaan dari algoritma *Decision Tree* sebelumnya yaitu algoritma ID3 yang hanya dapat dipergunakan untuk melakukan klasifikasi data kategorikal saja [20]. Selain itu, algoritma ID3 juga tidak dapat menampilkan hasil akurasi apabila terdapat *noise* pada data. Oleh karena itu, algoritma ID3 disempurnakan menjadi algoritma C4.5. Algoritma C4.5 dapat dipergunakan untuk menjalankan klasifikasi prediktif baik untuk data kategorikal maupun numerik [21]. Selain itu, algoritma C4.5 juga dapat membangun model yang mudah untuk dipahami dan ditafsirkan. Alasan lain dipilihnya algoritma C4.5 ini adalah algoritma ini memiliki keunggulan dalam mengelola data dengan *missing value* dan dapat menciptakan pedoman yang sederhana serta mudah untuk diinterpretasikan dan paling cepat dibandingkan dengan algoritma-algoritma lainnya [22]. Selain itu, algoritma C4.5 juga merupakan sebuah algoritma yang sangat kuat dan terkenal untuk melakukan klasifikasi dan prediksi [23]. Hal ini dikarenakan algoritma C4.5 membuat pohon keputusan secara *default* melalui proses pemangkasan. Sehingga, dapat terbentuk pohon keputusan yang lebih kecil dan sederhana.

Tahapan klasifikasi dengan algoritma C4.5 adalah sebagai berikut [19]:

1) Melakukan data *training* sesuai fitur-fitur yang terpilih pada tahap seleksi fitur dengan algoritma BPSO.

2) Menghitung *gain ratio* dari setiap atribut. Sebelum menghitung *gain ratio*, terlebih dahulu menghitung nilai *entropy* total. Persamaan (4) merupakan rumus untuk menghitung *entropy* total. Persamaan (5) merupakan rumus untuk menghitung nilai *gain*. Persamaan (6) merupakan rumus untuk menghitung *split info*. Kemudian yang terakhir yaitu (7) untuk menghitung nilai *gain ratio* nya.

$$Entopy\ (S) = \Sigma_{i=1}^{n} - pi * log_2 pi \quad (4)$$

Keterangan:

S : kumpulan kasus
pi : perbandingan antara Si terhadap S
n : jumlah partisi S

$$Gain\ (S, A) = Entopy\ (S) - \Sigma_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy\ (Si) \quad (5)$$

Keterangan:

S : kumpulan kasus
A : atribut
$|S_i|$: jumlah kasus pada partisi ke-i
$|S|$: jumlah kasus dalam s
n : jumlah partisi atribut A

$$Split\ Info\ (S, A) = -\Sigma_{i=1}^{n} \frac{S_i}{S} log_2 \frac{S_i}{S} \quad (6)$$

Keterangan:

S : kumpulan kasus
A : atribut
S_i : banyak sampel pada atribut i

$$Gain\ Ratio\ (a) = \frac{gain\ (a)}{split\ (a)} \quad (7)$$

Keterangan:

$gain\ (a)$: nilai gain
$split\ (a)$: nilai split

3) Menentukan *root node* atau *node* pada bagian paling atas dari pohon keputusan dengan cara memilih atribut yang

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

bernilai *gain ratio* paling tinggi. Selanjutnya buat *rule* berdasarkan atribut terpilih tersebut.

4) Melakukan perhitungan kembali *gain ratio* pada seluruh atribut kecuali pada atribut yang telah dipilih atau atribut yang telah menjadi *node* pada perulangan sebelumnya.

5) Menentukan nilai *gain ratio* paling tinggi untuk dijadikan sebagai *internal node* atau *node* dari sebuah percabangan. Selanjutnya buat *rule* berdasarkan pada atribut yang sudah dipilih tersebut.

6) Apabila atribut dari *internal node* belum signifikan dalam menentukan kelas prediksi, maka perlu dilakukan ulang langkah 4) dan 5) hingga *rule* yang dibuat mencapai kriteria yang ditetapkan untuk mencapai kelas prediksi yang signifikan. Akan tetapi, apabila atribut telah memenuhi kriteria, maka perulangan sudah bisa dihentikan dan pohon keputusan sudah terbentuk.

F. Evaluasi

Pada tahap evaluasi ini dilakukan dengan mempernggunakan *confusion matrix*. *Confusion matrix* dapat menggambarkan bagaimana model dapat melakukan prediksi dengan benar dan bagaimana model melakukan prediksi secara salah [24]. Pada *confusion matrix* hal-hal yang akan dievaluasi diantaranya adalah nilai akurasi, *precission, recall*, dan juga *f1-score*. *Confusion matrix* mengkategorikan prediksi benar dengan menggunakan nilai *true positif* dan *true negative*, sedangkan untuk kategori prediksi salah dilakukan dengan menggunakan nilai *false positive* dan *false negative* [25]. Table II menampilkan table dari *confusion matrix*. Table III menampilkan penjelasan dari *confusion matrix*.

TABLE II. CONFUSION MATRIX

		Actual	
		True	*False*
Predicted Class	*True*	*True Positive* (TP)	*False Positive* (FP)
	False	*False Negative* (FN)	*True Negative* (TN)

TABLE III. PENJELASAN CONFUSION MATRIX

Confusion Matrix	Description
True Positive	Banyak data positif terdeteksi dengan benar oleh system
True Negative	Banyak data negative terdeteksi dengan benar oleh system
False Positive	Banyak data positive terdeteksi salah oleh system
False Negative	Banyak data negative terdeteksi salah oleh system

Setelah dilakukan proses evaluasi dengan menggunakan *confusion matrix*, tahap selanjutnya adalah mengetahui performa hasil dari system. Untuk mengetahui performa hasil dari system juga dilakukan dengan *confusion matrix* menggunakan satuan ukur evaluasi diantaranya yaitu *accuracy*, *precision, recall*, dan *f1-score*. Pada penentuan gabungan tolak ukur terbaik dalam algoritma BPSO dilakukan dengan memanfaatkan nilai rata-rata *accuracy* dan waktu komputasi BPSO sebagai satuan ukur evaluasi. Sedangkan pada scenario klasifikasi algoritma C4.5 tanpa seleksi fitur dan dengan seleksi fitur BPSO dilakukan dengan mempergunakan nilai rata-rata dari *accuracy*, *precision*, *recall*, *f1-score*, dan waktu komputasi algoritma C4.5 sebagai satuan ukur evaluasi [19]. Persamaan (8) merupakan rumus menghitung nilai *accuracy*. Persamaan (9) merupakan rumus menghitung nilai *precision*. Persamaan (10) merupakan rumus menghitung nilai *recall*. Persamaan (11) merupakan rumus menghitung nilai *f1-score*.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} x\ 100\% \quad (8)$$

$$precision = \frac{TP}{TP+FP} x\ 100\% \quad (9)$$

$$recall = \frac{TP}{TP+FN} x\ 100\% \quad (10)$$

$$f1-score = 2\frac{(Precision\ x\ Recall)}{(Precision+Recall)} \quad (11)$$

III. HASIL DAN ANALISIS

Pada bagian ini berisi tentang penjelasan-penjelasan mengenai hasil dan analisis dari pemilihan fitur dengan algoritma BPSO, pembuatan model, klasifikasi dengan algoritma C4.5, dan klasifikasi dengan algoritma C4.5 + BPSO.

A. Pemilihan Fitur dengan Algoritma BPSO

Pemilihan fitur dengan algoritma BPSO dilakukan menggunakan nilai konstanta *c1* dan *c2* adalah 0.5, nilai konstanta *w* adalah 0.9, maksimum iterasi adalah 30, dan ukuran populasi partikelnya adalah 2. Dengan rincian tersebut, didapatkan hasil pemilihan fitur yang dapat dilihat pada table IV berikut ini. Dimana nilai 1 merupakan fitur terpilih, sedang nilai 0 merupakan fitur yang tidak terpilih. Table IV menampilkan hasil pemilihan fitur dengan menggunakan algoritma BPSO.

TABLE IV. HASIL PEMILIHAN FITUR

Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1

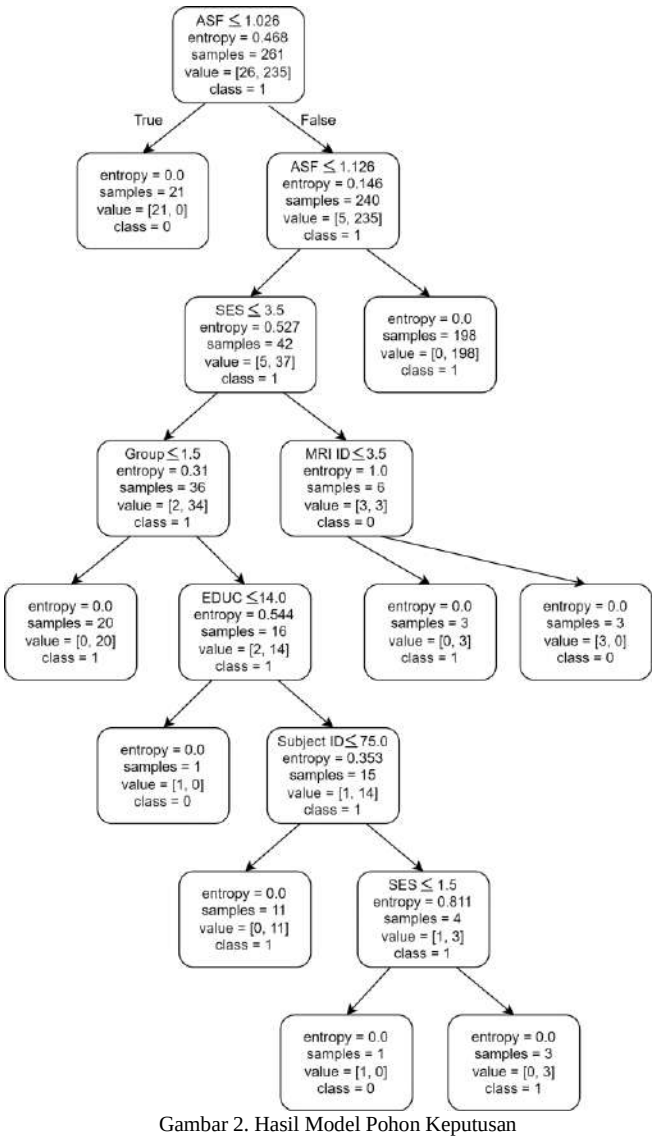
Hasil dari pemilihan fitur menggunakan algoritma BPSO didapatkan hasil bahwa fitur yang terpilih menurut seleksi fitur dengan algoritma BPSO ini diantaranya adalah *Subject ID, MRI ID, Group, Visit, MR Delay, M/F, Hand, Age, EDUC, SES, MMSE, CDR, nWBV*, dan *ASF*. Sedangkan fitur yang tidak terpilih adalah fitur *eTIV*. Fitur-fitur terpilih tersebut

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

merupakan fitur yang akan digunakan untuk klasifikasi dengan algoritma C4.5 dan fitur yang tidak terpilih tidak akan digunakan untuk klasifikasi.

B. Hasil Pembuatan Model

Pada pembuatan model ini menggunakan jumlah data total sebanyak 373 data. Data keseluruhan dibagi menjadi 70% data *training* dan 30% data *testing*. Pada pembuatan model ini, data *training* yang digunakan adalah sebanyak 261 data atau merupakan 70% dari keseluruhan data yang ada. Kemudian, fitur-fitur yang digunakan untuk membuat model ini merupakan fitur-fitur yang berhasil dipilih dari proses pemilihan fitur sebelumnya. Gambar 2 menampilkan hasil model pohon keputusan.



Gambar 2. Hasil Model Pohon Keputusan

Pada gambar 2 hasil model pohon keputusan di atas, dapat dilihat bahwa *root* node pada pohon keputusan tersebut adalah ASF, dimana fitur ASF ini merupakan fitur terpenting yang akan digunakan untuk memprediksi hasil akhir yaitu memprediksi pasien *positif* Alzheimer atau *negative* Alzheimer. Dari *root* node ASF ini memiliki dua cabang *True* dan *False*, dimana cabang *true* ini sudah selesai dan tidak perlu dilakukan perhitungan ulang, sedangkan cabang *false* perlu dilakukan perhitungan ulang kembali. Selanjutnya cabang kelas *false* menghasilkan dua cabang lagi yaitu salah satunya membentuk anak cabang SES. Kemudian, anak cabang SES juga membentuk dua cabang yaitu anak cabang Group dan anak cabang MRI ID. Anak cabang Group membentuk dua cabang yang salah satunya adalah anak cabang EDUC. Sedangkan anak cabang MRI ID membentuk dua cabang yang sama-sama sudah selesai semua perhitungannya. Selanjutnya anak cabang EDUC membentuk dua cabang dimana salah satunya merupakan anak cabang Subject ID. Anak cabang Subject ID membentuk dua cabang lagi dan salah satunya merupakan node SES. Terakhir anak cabang SES juga membentuk dua cabang yang sama-sama sudah selesai perhitungannya.

C. Klasifikasi dengan Algoritma C4.5

Pada tahap klasifikasi dengan algoritma C4.5 ini dilakukan dengan mempergunakan data *testing* dengan jumlah 30% dari data keseluruhan atau sebanyak 112 data. Pengujian model ini dilakukan dengan mempergunakan *confusion matrix*. Dengan menggunakan *confusion matrix* ini, maka nantinya akan menghasilkan nilai akurasi, *precision*, *recall*, dan *f1-score*. Table V menampilkan hasil *confusion matrix* dari proses klasifikasi dengan algoritma C4.5.

TABLE V. HASIL CONFUSION MATRIX ALGORITMA C4.5

		Actual	
		True	*False*
Predicted Class	*True*	5	1
	False	3	103

Pada table V di atas, dapat ditarik kesimpulan bahwa dari 112 data yang digunakan untuk *testing*, sebanyak 108 data dapat diklasifikasikan dengan benar dan sisanya yaitu 4 data diklasifikasikan salah. Berdasarkan hasil dari table *confusion matrix* di atas, dapat diketahui pula nilai akurasi, *precision*, *recall*, dan *f1-score* nya. Table VI menampilkan hasil performa dari model algoritma C4.5.

TABLE VI. HASIL PERFORMA DARI MODEL C4.5

Pengujian Model	Precision	Recall	F1-Score
Tidak Menderita (0)	0.62	0.83	0.71
Menderita (1)	0.99	0.97	0.98
Akurasi		96.4	

Pada table VI di atas, didapatkan hasil perhitungan dari model yang dibuat dengan nilai akurasi sebesar 96.4%, dengan nilai *precision* 0.62 (tidak menderita penyakit Alzheimer) dan 0.99 (menderita penyakit Alzheimer), kemudian nilai *recall* 0.83 (tidak menderita penyakit Alzheimer) dan 0.97 (menderita penyakit Alzheimer), serta nilai *f1-score* 0.71 (tidak menderita penyakit Alzheimer) dan 0.98 (menderita penyakit Alzheimer).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

D. Klasifikasi dengan Algoritma C4.5 + BPSO

Pada tahap klasifikasi dengan algoritma C4.5 + BPSO ini juga dilakukan dengan menggunakan data *testing* dengan jumlah 30% dari data keseluruhan atau sebanyak 112 data. Pengujian model ini dilakukan dengan menggunakan *confusion matrix*. Dengan menggunakan *confusion matrix* ini, maka nantinya akan menghasilkan nilai akurasi, *precision, recall,* dan *f1-score*. Table VII menampilkan hasil *confusion matrix* dari proses klasifikasi dengan algoritma C4.5 + BPSO.

TABLE VII. HASIL CONFUSION MATRIX ALGORITMA C4.5 + BPSO

		Actual	
		True	*False*
Predicted Class	*True*	6	2
	False	0	104

Pada table VII di atas, dapat ditarik kesimpulan bahwa dari 112 data yang digunakan untuk *testing*, sebanyak 110 data dapat diklasifikasikan dengan benar dan sisanya yaitu 2 data diklasifikasikan salah. Berdasarkan hasil dari table *confusion matrix* di atas, dapat diketahui pula nilai akurasi, *precision, recall*, dan *f1-score* nya. Table VIII menampilkan hasil performa dari model algoritma C4.5 + BPSO.

TABLE VIII. HASIL PERFORMA DARI MODEL C4.5 + BPSO

Pengujian Model	Precision	Recall	F1-Score
Tidak Menderita (0)	1	0.75	0.86
Menderita (1)	0.98	1	0.99
Akurasi	98.2		

Pada table VIII di atas, didapatkan hasil perhitungan dari model yang dibuat dengan nilai akurasi sebesar 98.2%, dengan nilai *precision* 1.00 (tidak menderita penyakit Alzheimer) dan 0.98 (menderita penyakit Alzheimer), kemudian nilai *recall* 0.75 (tidak menderita penyakit Alzheimer) dan 1.00 (menderita penyakit Alzheimer), serta nilai *f1-score* 0.86 (tidak menderita penyakit Alzheimer) dan 0.99 (menderita penyakit Alzheimer).

IV. KESIMPULAN

Berdasarkan penjelasan dari penelitian yang sudah dilaksanakan di atas, dapat kita simpulkan bahwa proses klasifikasi penyakit Alzheimer dengan mempergunakan algoritma BPSO untuk melakukan seleksi fitur terbukti dapat meningkatkan performa dari algoritma C4.5 yang ditandai dengan meningkatknya nilai akurasi dari sebesar 96.4% menjadi 98.2%. Hal ini dapat terjadi dikarenakan teknik seleksi fitur dengan algoritma BPSO pada klasifikasi dengan algoritma C4.5 dapat menghindari data yang bersifat *noise* dalam pembentukan pohon keputusan, serta mampu memilih fitur-fitur yang relevan tanpa menghilangkan informasi penting dari data tersebut karena fitur-fitur yang dipilih sudah dapat merepresentasikan data yang diklasifikasikan. Sehingga, hasil performa dari model algoritma C4.5 + BPSO dapat lebih baik dibandingkan dengan performa dari model algoritma C4.5 saja.

Adapun saran yang mungkin dapat dilakukan oleh para peneliti selanjutnya yaitu para peneliti bisa menggunakan dataset dengan jumlah yang lebih besar atau kompleks dengan harapan dapat memeroleh hasil yang semakin baik dan efektif.

DAFTAR PUSTAKA

[1] D. Kestel, "WHO Towards a dementia plan: a WHO guide. Geneva: World Health Organization," 2021.

[2] *WHO Global action plan on the public health response to dementia 2017 - 2025. Geneva: World Health Organization*. 2017. [Online]. Available: http://apps.who.int/bookorders.

[3] S. Khotimatul Wildah, S. Agustiani, M. S. Rangga Ramadhan, W. Gata, H. Mahmud Nawawi, and S. Nusa Mandiri, "Deteksi Penyakit Alzheimer Menggunakan Algoritma Naïve Bayes dan Correlation Based Feature Selection," *JURNAL INFORMATIKA*, vol. 7, no. 2, pp. 166–173, 2020, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/ji

[4] J. S. Birks and R. J. Harvey, "Donepezil for dementia due to Alzheimer's disease," *Cochrane Database of Systematic Reviews*, vol. 2018, no. 6. John Wiley and Sons Ltd, Jun. 18, 2018. doi: 10.1002/14651858.CD001190.pub3.

[5] A. Martin Prince *et al.*, "World Alzheimer Report 2015 The Global Impact of Dementia An AnAlysIs of prevAlence, IncIDence, cosT AnD TrenDs." [Online]. Available: www.alz.co.uk/worldreport2015corrections

[6] K. Nisa Berawi, E. Cania Bustomi, M. Rokok dalam Meningkatkan Risiko Penyakit Alzheimer, and H. Mulya Budiman, "Hasril Mulya Mekanisme Rokok dalam Meningkatkan Risiko Penyakit Alzheimer."

[7] L. Rizzi, I. Rosset, and M. Roriz-Cruz, "Global epidemiology of dementia: Alzheimer's and vascular types," *Biomed Res Int*, vol. 2014, 2014, doi: 10.1155/2014/908915.

[8] K. Nisa and R. Lisiswanti, "Kandita Mahran Nisa dan Rika Lisiswanti| Faktor Risiko Demensia Alzheimer MAJORITY I Volume 5 I Nomor 4 I Oktober," 2016.

[9] F. Akbar, "Jurnal Politeknik Caltex Riau Komparasi Algoritma Machine Learning Untuk Memprediksi Penyakit Alzheimer," 2022. [Online]. Available: https://jurnal.pcr.ac.id/index.php/jkt/

[10] C. Song *et al.*, "Immunotherapy for Alzheimer's disease: targeting β-amyloid and beyond," *Translational Neurodegeneration*, vol. 11, no. 1. BioMed Central Ltd, Dec. 01, 2022. doi: 10.1186/s40035-022-00292-3.

[11] H. Hampel *et al.*, "Precision pharmacology for Alzheimer's disease," *Pharmacological Research*, vol. 130. Academic Press, pp. 331–365, Apr. 01, 2018. doi: 10.1016/j.phrs.2018.02.014.

[12] M. Bari Antor *et al.*, "A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease," *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/9917919.

[13] A. Vyas, F. Aisopos, M. E. Vidal, P. Garrard, and G. Paliouras, "Identifying the presence and severity of dementia by applying interpretable machine learning techniques on structured clinical records," *BMC Med Inform Decis Mak*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12911-022-02004-3.

[14] Z. Abidin, E. Nurhana, Permata, and F. Ulum, "ANALISIS PERBANDINGAN ALGORITMA DECISION TREE C4.5 DAN C5.0 PADA DATA KARYAWAN BERPOTENSI PROMOSI JABATAN," 2023. [Online]. Available: www.kaggle.com

[15] R. A. Saputra *et al.*, "Detecting Alzheimer's Disease by the Decision Tree Methods Based on Particle Swarm Optimization," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Nov. 2020. doi: 10.1088/1742-6596/1641/1/012025.

[16] T. R. Sivapriya, A. R. N. B. Kamal, and P. R. J. Thangaiah, "Ensemble Merit Merge Feature Selection for Enhanced Multinomial Classification in Alzheimer's Dementia," *Comput Math Methods Med*, vol. 2015, 2015, doi: 10.1155/2015/676129.

[17] W. Wiharto, E. Suryani, and V. Cahyawati, "The methods of duo output neural network ensemble for prediction of coronary heart

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

disease," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 7, no. 1, pp. 50–57, Mar. 2019, doi: 10.11591/ijeei.v7i1.458.

[18] J. Too, A. R. Abdullah, N. M. Saad, and W. Tee, "EMG feature selection and classification using a Pbest-guide binary particle swarm optimization," *Computation*, vol. 7, no. 1, 2019, doi: 10.3390/computation7010012.

[19] I. G. A. M. Pratama *et al*., "Diagnosis Penyakit Ginjal Kronis dengan Algoritma C4.5, K-Means dan BPSO," 2022.

[20] M. Ardiansyah, A. Sunyoto, and E. T. Luthfi, "Edumatic: Jurnal Pendidikan Informatika Analisis Perbandingan Akurasi Algoritma Naïve Bayes dan C4.5 untuk Klasifikasi Diabetes," vol. 5, no. 2, pp. 147–156, 2021, doi: 10.29408/edumatic.v5i2.3424.

[21] M. Qois Syafi, "Increasing Accuracy of Heart Disease Classification on C4.5 Algorithm Based on Information Gain Ratio and Particle Swarm Optimization Using Adaboost Ensemble," *Journal of Advances in Information Systems and Technology*, vol. 4, no. 1, 2022, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/jaist

[22] M. Burhan Hanif and G. Guntoro Setiaji, "Meningkatkan Kinerja Decision Tree C4.5 Dengan Seleksi Fitur Korelasi Pearson Pada Deteksi Penyakit Diabetes," *Indonesian Journal of Computer Science*, vol. 11, pp. 685–695, Aug. 2022.

[23] Y. Irawan, "Penerapan Algoritma Decision Tree C4.5 Untuk Prediksi Kelayakan Calon Pendonor Darah Dengan Klasifikasi Data Mining," vol. 2, no. 4, pp. 181–189, 2021.

[24] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," 2021.

[25] F. Rofii, G. Priyandoko, M. I. Fanani, and A. Suraji, "Vehicle Counting Accuracy Improvement By Identity Sequences Detection Based on Yolov4 Deep Neural Networks," *TEKNIK*, vol. 42, no. 2, pp. 169–177, Aug. 2021, doi: 10.14710/teknik.v42i2.37019.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1716, Copyright ©2023
Submitted : May 13, 2023, Revised : July 6, 2023, Accepted : August 24, 2023, Published : November 4, 2023

Emotion Mining Review Pengguna Aplikasi Mobile Banking BRImo Menggunakan Algoritma Decision Tree

Debby E Sondakh [1]*, Raissa C Maringka [2], Ferlien P Ayorbaba [3], Joanne S.C.B.T Mangi [4], Stenly R. Pungus [5]
Fakultas Ilmu Komputer Sistem Informasi [1], [2], [3], [4], [5]
Universitas Klabat Airmadidi, Manado, Indonesia
Debby.sondakh@unklab.ac.id [1], raissam@unklab.ac.id [1],
s21910087@student.unklab.ac.id [3], s11910041@student.unklab.ac.id[4], stenly.pungus@unklab.ac.id[5]

***Abstract*— As consumer transaction preferences shifted from analog to digital, banks were compelled to develop digital transactions in the form of mobile banking. Users of mobile banking provide feedback regarding the application's usability. The opinions of users can be emotive. Emotions influence what a person emits or applies. Emotions are the behavioral response of a person when he is happy or unhappy. Thus, the manifestation of a person's emotions, whether in the form of facial expressions, verbal communication, written text, or judgment, can be used as a source of information to aid in decision making. The objective of this study is to apply emotion mining to the analysis of user evaluations of the BRImo application, one of the three most popular platforms in Indonesia as of August 2022, with a total of 800,000 reviews on the Play Store. Emotion Mining can be used to analyze the four categories of emotions expressed by users in the comments section: happy, angry, sad, and afraid. According to BRImo user evaluations, the decision tree algorithm is used to categorize happy, sad, afraid, and angry feelings. Using a decision tree to manage large data category sets is effective. The obtained dataset included 2959 happy classes, 2196 sad classes, 387 angry classes, and 81 scared classes. According to the findings of the analysis, a significant number of users of the BRImo application express positive sentiments in their evaluations, which are indicative of happy emotions. The Decision Tree algorithm yields results with a performance specification of 84.5%, sensitivity of 85.5%, and precision of 84.4%.**

Keywords*— *Emotion Mining, Classification, BRImo, Decision Tree, Machine Learning, Sentiment Analysis

***Abstrak*— Perpindahan gaya bertransaksi nasabah dari konvensional ke digital membuat perbankan memanfaatkan teknologi untuk mengembangkan transaksi digital dalam bentuk mobile banking. Pengguna mobile banking memberikan pendapat mereka dalam menggunakan aplikasi tersebut. Pendapat yang diberikan oleh pengguna bisa berupa emosi. Emosi seseorang mempengaruhi apa yang mereka keluarkan atau luapkan. Emosi merupakan response perilaku seseorang disaat ia merasa senang atau tidak senang. Sehingga, pengekspresian emosi seseorang, berupa ekspresi wajah, verbal, teks maupun penilaian dapat menjadi informasi untuk membantu suatu individu mengambil keputusan. Penelitian ini bertujuan menerapkan emotion mining dalam menganalisis ulasan pengguna terhadap aplikasi BRImo, salah satu dari tiga platform yang banyak digunakan di Indonesia, dengan jumlah 800.000 ulasan di Play Store per Agustus 2022. Emotion Mining dapat membantu menganalisis emosi yang dikategorikan ke dalam 4 jenis emosi yang dituangkan pengguna pada kolom komentar, yaitu emosi senang, marah, sedih dan takut. Algoritma decision tree digunakan untuk mengklasifikasi emosi senang, sedih, takut dan marah ulasan pengguna BRImo. Decision tree memiliki kinerja yang baik dalam mengelola kumpulan data kategorikal yang besar. *Dataset* yang diperoleh, terdiri dari 2959 kelas senang, 2196 kelas sedih, 387 kelas marah dan 81 kelas takut. Hasil analisi menunjukkan bahwa pengguna aplikasi BRImo banyak memberikan ulasan yang positf yang menandakan emosi senang. Hasil yang diperoleh dari algoritma Decision Tree menghasilkan performa spesifisitas 84.5%, sensitivitas 85.5% dan akurasi 84.4%.**

Kata Kunci*— *Emotion Mining, Klasifikasi, Pohon Keputusan, Machine Learning, Analisis Sentimen

I. PENDAHULUAN

A. *Latar Belakang*

Teknologi saat ini sudah sangat melekat dalam kehidupan masyarakat sekarang, mulai dari anak-anak dalam pendidikan hingga orang dewasa dalam menunjang pekerjaan mereka seperti membuat laporan, memesan transportasi untuk pergi ke kantor, bahkan sampai kepada transaksi perbankan semuanya membutuhkan teknologi. Teknologi sudah diterapkan oleh industri perbankan di Indonesia, termasuk *mobile banking*, aplikasi berbasis mobile yang memungkingkan nasabah melakukan transaksi perbankan melalui telepon selular. *Mobile banking* tidak hanya mendukung nasabah untuk bertransaksi secara online tanpa Batasan waktu, juga membantu penggunanya mendapatkan informasi dan berkomunikasi [1]

Bank Rakyat Indonesia (BRI) meluncurkan aplikasi BRI Mobile (BRImo) pada tahun 2019, untuk melengkapi layanan berbasis teknologi yang telah digunakan sebelumnya yaitu *internet banking*, info BRI, dan call BRI. BRImo menggabungkan *mobile banking*, *internet banking*, dan uang elektronik dalam satu aplikasi. Pada tahun peluncurannya, aplikasi BRImo digunakan oleh sebanyak 2.2 juta pengguna dan jumlah transaksi mencapai hingga 1.6 triliun. Gambar 1 memperlihatkan *mobile banking* di Indonesia dengan kinerja yang terbaik pada tahun 2020 dan 2021. Berdasarkan laporan dari Bank *Service Excellence Monitor* (BSEM) dari *Marketing*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

Research Indonesia menunjukan aplikasi BRImo menduduki peringkat ketiga [2]. Tahun 2022, merujuk data di Play Store, BRImo telah diunduh lebih dari 10 juta kali dengan *rating* 4.5/5 [3]. Bahkan, pada kuartal 1-2023, jumlah pengguna mencapai lebih dari 26.3 juta dengan kenaikan jumlah transaksi finansial BRImo mencapai 99.07%, merupakan pertumbuhan transaksi *mobile banking* terbesar [4]. Disediakannya 100 fitur yang memberikan kemudahan dan kenyamanan bertransaksi menjadi salah satu alasan terjadinya peningkatan ini.

Gambar 1. Kinerja Mobile Banking di Indonesia 2020-2021 [2]

Pengguna aplikasi mobile memberikan ulasan yang mendeskripsikan pendapat, komentar, atau evaluasi subjektif mengenai kelebihan, kekurangan, kualitas atau pengalaman pengguna dengan suatu aplikasi. Dalam konteks aplikasi mobile banking, ulasan pengguna dapat mencakup komentar mengenai kemudahan penggunaan, keamanan, kecepatan transaksi, fitur yang disediakan, kehandalan, dan pelayanan pelanggan. Ulasan yang dibagikan di *platform online* seperti situs resmi aplikasi, forum atau komunitas pengguna, dan toko aplikasi resmi seperti Google Play Store atau Apple App Store, memberikan perspektif pengguna yang dapat membantu calon pengguna lain dalam mengambil keputusan untuk menggunakan aplikasi atau tidak. Selain itu, ulasan dapat membantu mengidentifikasi masalah dalam aplikasi yang berguna bagi pengembang untuk memperbaiki atau melakukan pembaruan dan peningkatan fitur.

Perlu digarisbawahi bahwa ulasan pengguna pada dasarnya bersifat subyektif, karena dipengaruhi oleh pendapat dan perspektif individu. Satu ulasan tidak mencerminkan pengalaman pengguna secara keseluruhan secara akurat. Sehingga, sebelum merumuskan penilaian atau keputusan, sangat penting untuk mempertimbangkan dan menganalisis ulasan secara holistik untuk mendapatkan tren menyeluruh dari umpan balik pengguna. Pendekatan pembelajaran mesin banyak digunakan untuk mendapatkan sentimen (positif atau negatif) terhadap aplikasi *mobile banking*, yang disarikan dari ulasan pengguna.

Sentimen analisis atau *opinion mining* merupakan sebuah bidang pengolahan bahasa alami yang luas. Sentimen analisis memiliki hubungan dengan text mining yang bertujuan untuk menganalisa pendapat, sentimen, sikap, evaluasi, penilaian serta emosi seseorang. Sentimen analisis digunakan untuk mengelompokan polaritas dari teks yang ada dalam sebuah dokumen, kalimat maupun nilai yang bersifat positif, negatif atau netral [5], [6].

Menurut Ranganathan dan Tzacheva [7] analisis sentimen tidak hanya mengidentifikasi opini, tetapi juga emosi. Emosi, sebagai fenomena kognitif yang rumit, bermanifestasi sebagai keadaan mental beraneka segi yang dapat dibentuk oleh berbagai faktor termasuk rangsangan eksternal, perubahan fisiologis, dan dinamika interpersonal. Emosi mempengaruhi apa yang diluapkan seseorang. Emosi dapat berupa ekspresi wajah, secara verbal, dan melalui teks. Emosi mempengaruhi persepsi manusia. Saat ini, persepsi manusia menjadi faktor penting yang mempengaruhi berbagai bidang, termasuk bisnis, pendidikan, seni, dan sebagainya. Dalam konteks bisnis, pemahaman yang lebih komprehensif tentang persepsi pengguna dapat diperoleh dengan menganalisis emosi dari teks ulasan yang diberikan untuk suatu produk atau layanan, menggunakan teknik analisis teks *emotion mining*. *Emotion mining* adalah studi tentang emosi yang tercermin dalam sebuah teks, yang mencakup proses menambang emosi dari teks mencakup deteksi emosi dalam teks, menentukan polaritas emosi yang ada dalam teks, klasifikasi emosi, dan mendeteksi penyebab emosi [8].

Berdasarkan paparan masalah diatas, peneliti akan melakukan penelitian terhadap emosi yang tuangkan pengguna pada aplikasi BRImo di Google Play Store menggunakan teknik *emotion mining*. Beberapa penelitian telah dilakukan untuk menganalisis sentimen pengguna aplikasi BRImo [9]–[11]. Dari 9000 ulasan yang dianalisis menggunakan algoritma Support Vector Machine, diperoleh 29% sentimen positif dan 71% negatif [11]. Penelitian berbeda mendapati BRImo memiliki masalah serius terkait kehandalan (*reliability*) [10], yang dapat berkontribusi pada besarnya persentase sentimen negatif. Hasil yang berbeda diperoleh [9] menganalisis 1906 ulasan pengguna BRImo dari Google Play Store menggunakan Naïve Bayes, dimana persentase sentimen positif (53.09%) lebih besar daripada sentimen negatif (46.9%). Namun, literatur yang mengkaji tentang penggunaan emotion mining masih sangat terbatas. Di sisi lain, *emotion mining* dapat melampaui analisis sentimen (opini) dan mengungkap emosi spesifik yang diungkapkan oleh pengguna dalam ulasan mereka tentang aplikasi *mobile banking*, dalam hal ini BRImo. Dengan mengidentifikasi emosi seperti senang, marah, sedih, atau takut, dapat memberikan wawasan yang lebih dalam tentang pengalaman pengguna. Memahami emosi spesifik yang terkait dengan fitur, fungsi, atau aspek tertentu dari aplikasi dapat membantu pengembang dan penyedia aplikasi untuk mengatasi masalah pengguna, meningkatkan kepuasan pengguna, dan menciptakan pengalaman yang lebih menarik secara emosional.

Penelitian ini menggunakan model klasifikasi algoritma *decision tree* (DT). DT merupakan metode klasifikasi model *supervised learning* untuk menyelesaikan permasalahan ke dalam struktur pohon [12]. DT melihat variabel dalam kumpulan data, menentukan mana yang paling penting, dan kemudian muncul dengan struktur hierarki keputusan yang paling baik mempartisi data. pohon dibuat dengan membagi data berdasarkan variabel dan kemudian menghitung untuk melihat berapa banyak yang ada di setiap keranjang setelah setiap pemisahan [13]. Membangun DT itu mudah dan cepat, dan dapat membuat prediksi dengan cepat. Ini sederhana dan mudah dipahami, dan pengguna tidak perlu melakukan banyak hal untuk menyiapkan data. Aturan dapat dibuat dengan cepat,

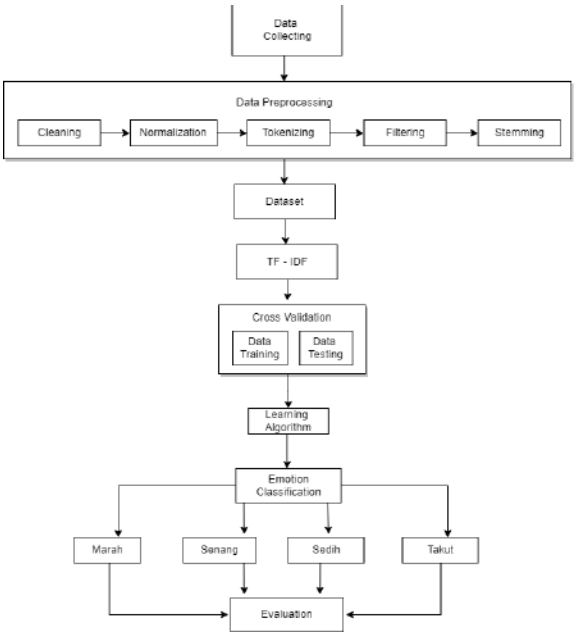
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

dan kompleksitas masalah juga berkurang [13], [14]. DT ini merupakan algoritma non-parametrik sehingga dapat digunakan pada *dataset* yang besar dan kompleks tanpa memaksakan struktur parametrik yang sulit. Ketika sampel yang akan digunakan lumayan besar maka *dataset* tersebut dapat dibagi menjadi *dataset training* dan *dataset* validasi. *Dataset training* digunakan sebagai dasar membangun model dari DT dan *dataset* validasi digunakan untuk memutuskan ukuran *tree* yang sesuai yang dibutuhkan untuk mencapai model akhir yang lebih optimal [15].

Selain itu, melihat dari penelitian sebelumnya teks [7], hasil klasifikasi emosi teks dari data Twitter menggunakan DT memperoleh akurasi yang lebih tinggi dibandingkan penelitian lain yang menggunakan SVM dan KNN. Nurfauzan dan Maharani, juga menambang emosi dari teks Twitter berbahasa Indonesia mendapati DT memberikan nilai akurasi klasifikasi emosi yang baik [15].

II. METODE PENELITIAN

Gambar 2 menjelaskan proses-proses dari mengambil data pada kolom komentar aplikasi BRImo yang tersedia pada *Google Play Store.*



Gambar 2. Langkah-langkah Penelitian

Setelah itu data akan melalui tahap *pre-processing* dan akan diukur bobotnya dengan menggunakan TF-IDF. *Dataset* tersebut kemudian dibagi menjadi 2 bagian yaitu *data training* dan *data testing* dan akan diolah untuk kemudian mendapatkan model klasifikasi. Setelah itu data tersebut akan dievaluasi untuk mendapatkan nilai dari 4 emosi yaitu marah, senang, sedih dan takut.

A. Data Collection

Peneliti menggunakan *web data scrapping* serta *library* python "pandas" dan "numpy" untuk mengambil sebanyak 10.000 ulasan dari Google Play Store. Paramater *scrapping* yang peneliti gunakan yaitu id "id.co.bri.brimo", lang "id", country "id", sort "Sort.NEWEST", count "10000" dan filter score with "None".

B. Text Preprocessing

Data ulasan aplikasi BRImo melewati beberapa tahapan seperti *cleaning, normalization, tokenizing, filtering* dan *stemming.* Seperti yang ditunjukkan pada Gambar 3.

Gambar 3. Langkah-langkah Text Preprocessing

TABEL 1. Contoh Hasil *Preprocessing*

Before	After	Process
Aplikasi sampah! kesalahan sistem kok setiap hari! Error kok setiap hari juga! Sangat tidak berguna!	Aplikasi sampah kesalahan sistem kok setiap hari Error kok setiap hari juga sangat tidak berguna	Cleaning
Aplikasi sampah kesalahan sistem kok setiap hari Error kok setiap hari juga sangat tidak berguna	'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'kok', 'setiap', 'hari', 'Error', 'kok', 'setiap', 'hari', 'juga', 'Sangat', 'tidak', 'berguna'	Tokenizing
'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'kok', 'setiap', 'hari', 'Error', 'kok', 'setiap', 'hari', 'juga', 'Sangat', 'tidak', 'berguna'	'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'kok', 'setiap', 'hari', 'Error', 'kok', 'setiap', 'hari', 'juga', 'Sangat', 'tidak', 'berguna'	Normalization
'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'kok', 'setiap', 'hari', 'Error', 'kok', 'setiap','hari', 'juga', 'Sangat', 'tidak', 'berguna'	'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'Error', 'Sangat', 'berguna'	Filtering
'Aplikasi ', 'sampah', 'kesalahan', 'sistem', 'Error', 'Sangat', 'berguna'	aplikasi sampah kesalahan sistem error berguna	Stemming

Data ulasan yang diperoleh oleh peneliti kemudian melewati proses *cleaning* untuk menghilangkan angka atau simbol yang tidak dibutuhkan, seperti hashtag, emoticon, tanda baca dan simbol-simbol lainnya. Setelah itu masuk ke tahap *tokenizing* untuk memisahkan kalimat menjadi kata per kata. Kemudian pada tahap normalization dilakuan mapping untuk menghilangkan kata slang atau kata informal menjadi formal. Setelah itu pada tahap *filtering* kalimat dilakukan pengurangan kata dalam corpus yang disebut stopwords dan penghilangan kata yang tidak diperlukan dalam penelitian. Dan pada tahap *stemming* akan dilakukan proses penghilangan semua kata imbuhan menjadi kata dasar. Contoh: di-, me-, kan-, meng-, peng-, pe- dan lain-lain. *Stemming* bertujuan untuk mengurangi kata yang memiliki kata dasar yang sama. Seperti pada Tabel 1. Data yang awalnya berjumlah 10.000 telah berkurang sebanyak 3.373 dan tersisa 6.627.

Setelah melewati tahap preprocessing, data dianalisis dengan membagi data ke dalam 3 kategori yaitu positif, negatif

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

dan netral. Selain itu, ditentukan *range polarity* berdasarkan *human judgement*.

C. *Modeling*

Setelah melewati tahap *pre-processing* dan *sentiment analysis*, data kemudian terbentuk menjadi *dataset*. *Dataset* yang telah terbentuk akan melewati tahap TF-IDF untuk membobotkan setiap kata dalam dokumen berdasarkan frekuensi kemunculan kata tersebut. Setelah dilakukan pembobotan terhadap kata dan diperoleh *dataset*, peneliti kemudian membagi data ke dalam 2 bagian, yaitu *data training* 80% dan *data testing* 20% untuk kemudian dilakukan pengujian algoritma.

Pada tahap modeling, peneliti mengimplementasikan *decision tree* menggunakan *Scikit-Learn* yang sudah teruji dan memiliki dokumentasi yang lengkap untuk mengetahui emosi apa yang sering digunakan oleh pengguna aplikasi BRImo berdasarkan ulasan yang diberikan.

D. *Emotion Classfication*

Setelah data telah melewati tahap preprocessing, sentimen analisis dan modeling, maka selanjutnya dari hasil sentimen positif dan negatif akan diklasifikasikan ke dalam 4 emosi yaitu marah, senang, sedih dan takut. Neutral merupakan sentimen analisis yang memiliki hasil *undefined*.

E. *Evaluation*

Setelah membangun model klasifikasi sentimen, penting untuk mengevaluasi kinerjanya. Ini biasanya dilakukan dengan menggunakan data uji berlabel yang tidak digunakan selama fase pelatihan model. Metrik evaluasi umum meliputi akurasi, presisi, daya ingat, skor F1, dan analisis *confusion matrix*.

III. HASIL DAN ANALISA

Dari 10.000 data yang diperoleh, peneliti melakukan preprocessing dan analisis sentiment terhadap data yang ada, dari tahap preprocessing cleaning sampai filtering, data masih berjumlah 10.000 dan pada tahap stemming data berkurang menjadi 6627 data. Pada tahap sentimen analisis data dibagi ke dalam 3 kategori yaitu positif, negatif, dan netral, kode program ditunjukkan pada Gambar 4.

Setelah data telah melewati tahap preprocessing dan sentiment analisis maka kemudian akan dilakukan proses pembobotan pada kata dengan TF-IDF, kode program ditunjukkan pada Gambar 5. Setelah diperoleh *dataset*, data kemudian akan dibagi ke dalam 2 bagian yaitu *data training* 80% dan data testing 20% dan kemudian akan dilakukan pengujian algoritma.

```
def sentiment_analysis_lexicon_indonesia(text):
    score = 0
    for word in text:
        if(word in lexicon_positive):
            score = score + lexicon_positive[word]
    for word in text:
        if(word in lexicon_negative):
            score = score + lexicon_negative[word]
    polarity = ''
    if (score > 0):
        polarity = 'positive'
    elif (score < 0):
        polarity = 'negative'
    else:
        polarity = 'neutral'
    return score,polarity
```

Gambar 4. Kode untuk Analisis Sentimen

```
# Vectorize text data using TF-IDF
vectorizer = TfidfVectorizer(analyzer='word', max_features=2000)
tfidf_matrix_train = vectorizer.fit_transform(X_train)
tfidf_matrix_test = vectorizer.transform(X_test)
```

Gambar 5. Kode untuk Pembobotan Kata

```
# Train decision tree model
clf = DecisionTreeClassifier(random_state=0)
clf.fit(tfidf_matrix_train, y_train)

# Evaluate model on test set
y_pred = clf.predict(tfidf_matrix_test)
accuracy = clf.score(tfidf_matrix_test, y_test)
print('Decision tree accuracy:', accuracy)

# Print classification report
report = classification_report(y_test, y_pred)
print(report)
precision, recall, fscore, support = precision_recall_fscore_support(y_test, y_pred)
sensitivity = recall[1]
specificity = recall[0]
print('Sensitivity (positive recall):', sensitivity)
print('Specificity (negative recall):', specificity)
```

Gambar 6. Kode Algoritma Decision Tree

Pada tahap modeling, peneliti mengimplementasikan DT menggunakan *Scikit-Learn* yang sudah teruji dan memiliki dokumentasi yang lengkap untuk mengetahui emosi apa yang sering digunakan oleh pengguna aplikasi BRImo berdasarkan ulasan yang diberikan.

Tabel 2 menampilkan kinerja model klasifikasi menggunakan algoritma DT, dimana 950 sampel diklasifikasikan dengan benar, masing-masing 486 dan 464 untuk kelas positif dan negatif. Jumlah sampel yang salah klasifikasi adalah 175.

TABEL 2. *Confusion matrix*

True	**Prediksi**		**Total**
	Negatif	***Positif***	
Negatif	TN: 464	FP: 89	553
Positif	FN: 86	TP: 486	572
Total	550	575	1125

Selanjutnya, metrik evaluasi yaitu *sensitivity*, *Specificity*, *precision*, *recall*, *f1-score* dan *accuracy* dihitung, dengan hasil masing-masing memiliki persentase 85.5%, 84%, 84.5, 84%, 84% dan 84%. Perhitungan masing-masing *metric* yang digunakan adalah sebagai berikut:

Sensitivity: Mengukur seberapa baik *classifier* dapat digunakan untuk mengidentifikasi *instance* dari kelas positif.

$$Sensitivity: \frac{486}{486+86} x\ 100\% = 85.5\%$$

Specificity: Mengukur seberapa baik *classifier* dapat digunakan untuk mengidentifikasi *instance* dari kelas negatif.

$$Specificity: \frac{464}{464+89} x\ 100\% = 84\%$$

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

Precision: Mengukur rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif.

$$precision: \frac{486}{486+89} x\ 100\% = 84.5\%$$

Accuracy: Mengukur keakuratan dari klasifikasi.

$$Accuracy: \frac{486+464}{486+464+89+86} x\ 100\% = 84\%$$

Dari hasil penentuan *range polarity* yang telah dilakukan pada tahap sentimen analisis, sentimen positif dan negatif diklasifikasikan ke dalam 4 emosi yaitu marah, senang, sedih dan takut. *Neutral* merupakan sentimen analisis yang memiliki hasil *undefined*. *Range polarity* yang ditentukan oleh peneliti adalah berdasarkan *human judgement* sebagai berikut: -1 sampai -10 *sad*, -11 sampai -20 *angry*, -21 sampai -60 *fear* dan 1 sampai 36 *happy*.

Maka pada tahap *emotion classification*, data yang berjumlah 6627 berkurang 998 karena peneliti tidak menggunakan *neutral* dalam mengklasifikasikan emosi. Gambar 7 menampilkan hasil klasifikasi emosi 5623 data, dimana didapati dari setiap kelas terbagi secara tidak merata. *Dataset* yang peneliti peroleh, terdiri dari 2959 kelas senang, 2196 kelas sedih, 387 kelas marah dan 81 kelas takut. Dapat disimpulkan bahwa pengguna aplikasi BRImo banyak memberikan ulasan yang positif yang ditandai 2959 (53%) emosi senang. Namun, secara keseluruhan, ulasan positif tidak dominan dibandingkan total ulasan negatif yang tercermin dari emosi seidh, marah dan takut sebanyak 2664 (47%).

Gambar 7. Klasifikasi Emosi Pengguna BRImo

IV. KESIMPULAN

Pada penelitian ini, teknik *emotion mining* yang digunakan oleh peneliti untuk mengidentifikasi emosi dari teks ulasan pengguna aplikasi BRImo. Algoritma DT yang digunakan memiliki kemampuan yang baik dalam mengenali dengan tepat emosi yang terkait dengan ulasan pengguna. Dengan tingkat *Sensitivity* sebesar 85.5%, algoritma tersebut mampu mengidentifikasi dengan akurat sebagian besar emosi yang terkandung dalam ulasan. *Specificity* sebesar 84% menunjukkan kemampuan algoritma dalam mengklasifikasikan dengan benar emosi negatif. *Precision* sebesar 84.5% mengindikasikan bahwa sebagian besar prediksi positif yang dilakukan oleh algoritma benar. Dan dengan tingkat *accuracy* sebesar 84%, algoritma decision tree mampu memberikan hasil klasifikasi yang akurat secara keseluruhan.

Hasil analisis klasifikasi sentimen menunjukkan bahwa pengguna aplikasi BRImo memberikan ulasan yang cukup beragam. Terdapat 2959 (53%) ulasan dengan sentimen positif yang ditandai dengan emosi senang. Meskipun demikian, secara keseluruhan, ulasan positif tidak mendominasi jumlahnya dibandingkan dengan total ulasan negatif. Ulasan negatif tercermin dalam emosi sedih, marah, dan takut, yang mencapai 2664 (47%). Hal ini menunjukkan bahwa meskipun ada sejumlah besar ulasan positif, terdapat juga sejumlah signifikan ulasan negatif yang perlu diperhatikan dan ditindaklanjuti oleh pihak yang bertanggung jawab terhadap aplikasi BRImo. Penting untuk mempertimbangkan dan mengatasi masalah yang mendasari ulasan negatif tersebut guna meningkatkan pengalaman pengguna dan kualitas layanan secara keseluruhan.

DAFTAR PUSTAKA

[1] A. C. Mandiri, Efriyanto, and E. Y. Metekohy, "Pengaruh Kualitas Layanan dan Kepercayaan Terhadap Kepuasan Nasabah dalam Menggunakan BRI Mobile (BRImo)," *Account*, vol. 8, no. 1, pp. 1423–1430, Jun. 2021.

[2] "Wow! Ini Mobile Banking Terbaik Se-Indonesia," *detikNews*, May 06, 2021.

[3] L. D. Jatmiko, "Top 10 Aplikasi Bank di Indonesia," *Bisnis.com*, May 24, 2022.

[4] K. Anam, "Ramai Transaksi Mobile Banking, Bank Mana yang Tumbuh Tinggi?," *CNBC Indonesia*, Apr. 28, 2023.

[5] S. Hadi and N. Novi, "Faktor-Faktor Yang Mempengaruhi Penggunaan Layanan Mobile Banking," *Optimum: Jurnal Ekonomi dan Pembangunan*, vol. 5, no. 1, p. 55, 2015, doi: 10.12928/optimum.v5i1.7840.

[6] G. A. Sandag, E. H. E. Soegiarto, L. Laoh, A. Gunawan, and D. Sondakh, "Sentiment Analysis of Government Policy Regarding PPKM on Twitter Using LSTM," in *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*, Prapat, Indonesia: IEEE, Oct. 2022, pp. 1–6. doi: 10.1109/ICORIS56080.2022.10031474.

[7] J. Ranganathan and A. Tzacheva, "Emotion mining in social media data," *Procedia Comput Sci*, vol. 159, pp. 58–66, 2019, doi: 10.1016/j.procs.2019.09.160.

[8] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current State of Text Sentiment Analysis from Opinion to Emotion Mining," *ACM Comput Surv*, vol. 50, no. 2, pp. 1–33, Mar. 2018, doi: 10.1145/3057270.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

[9] M. K. Khoirul Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi BRImo Pada Ulasan Pengguna di Google Play Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 478–483, Mar. 2023, doi: 10.36040/jati.v7i1.6373.

[10] M. A. Ligiarta and Y. Ruldeviyani, "Customer Satisfaction Analysis of Mobile Banking Application Based on Twitter Data," in *2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, Yogyakarta, Indonesia: IEEE, Nov. 2022, pp. 322–327. doi: 10.1109/ICE3IS56585.2022.10010221.

[11] H. I. Harianja, "Analisis Sentimen Ulasan Pengguna terhadap Aplikasi BRImo BRI pada Situs Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," Universitas Sumatera Utara, Medan, 2022.

[12] K. A. Rokhman, B. Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *Journal of Information System Management (JOISM)*, vol. 3, no. 1, pp. 1–7, 2021, doi: 10.24076/joism.2021v3i1.341.

[13] A. Kumar, V. Dabas, and P. Hooda, "Text classification algorithms for mining unstructured data: a SWOT analysis," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1159–1169, Dec. 2020, doi: 10.1007/s41870-017-0072-1.

[14] D. E. Sondakh, "Comparative Study of Classification Algorithms: Holdouts as Accuracy Estimation," *Cogito Smart Journal*, vol. 1, no. 1, pp. 13–23, Dec. 2015.

[15] A. Nurfauzan and W. Maharani, "Klasifikasi Emosi Pada Pengguna Twitter Menggunakan Metode Klasifikasi Decision Tree," *eProceedings …*, 2021.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1721, Copyright ©2023
Submitted : May 15, 2023, Revised : June 22, 2023, Accepted : July 10, 2023, Published : November 4, 2023

Classification of Final Project Titles Using *Bidirectional Long Short Term Memory* at the Faculty of Engineering Nurul Jadid University

Faridatul Warda[1], Fathorazi Nur Fajri[2]*, Abu Tholib[3]
Information Systems Study Program, Faculty of Engineering Nurul Jadid University Probolinggo[1], [2]
Informatics Study Program, Faculty of Engineering Nurul Jadid University Probolinggo [3]
firdalanova@gmail.com[1] , fathorazi@unuja.ac.id[2] , ebuenje@gmail.com[3]

***Abstract*— Every year, the Faculty of Engineering at Nurul Jadid University forms a committee to manage the process of students' final projects from the title selection stage to the final examination process until graduation. The process of selecting the final project title is still done manually, namely by checking the titles one by one, which takes a long time and allows errors because there is a lot of data to check, so human errors can also occur. Therefore, this research proposes to use the Bidirectional Long Short Term Memory (BiLSTM) method to classify the final project title based on its grade category. Several experiments were conducted to generate the most appropriate labels. The first experiment produced 4 labels and the second experiment produced 2 labels. From the results of several experiments, it was concluded that the second experiment had the best accuracy results with the 'good enough' and 'good' classes. The oversampling technique was then applied to overcome overlapping data, and the turning process was then performed on several parameters that could re-optimize the previous accuracy result of 75.24% to 91.15%. With a configuration of 10 random state parameters, using 64 batch sizes and 50 epochs. In addition, model adjustments were made to the hidden layer by adding a dropout layer and relu activation.**

Keywords*— *Final project, Classification text, BiLSTM, Hyperparameter turning.

I. INTRODUCTION

As a requirement for graduation at a college or university, students must complete a final project[1] therefore have an impact on future academic endeavours or jobs that require a higher level of knowledge[2]. Therefore, dissertations are often created to assess a student's ability to apply their knowledge and skills. The final project involves research on a problem within a specific research topic, guided by a supervisor, and by following some related guidelines. The final project involves researching a problem on a specific topic with the guidance of a supervisor and following some related guidelines[3].

The process of completing the final project is not an easy process and requires full commitment in going through several stages, such as submitting the final project title, proposal seminar, preparing the final project report, and final project trial[4]. Therefore, every stage in completing the final project must be undertaken wholeheartedly and with high commitment.

Submitting the final project title is the first step in completing the final project. Determining the right title is very important because the title must meet the standards set, such as the product produced, the theory used, novelty, and a clear object. A title that does not meet the standards can hinder the next process in completing the final project.

To select the final project title and some interests for the final project, the Faculty of Engineering of Nurul Jadid University formed a selection team committee that is responsible for selecting the final project title. However, the selection process for the final project title is currently still done manually by the selection team. This takes a long time in the selection process and risks producing human error. Human error can be caused by a lack of seriousness, physical problems or mental fatigue in completing work, or even rushing to complete work[5].

Therefore, a manual approach to text categorisation is impractical due to high workload and low efficiency[6]. A new innovation is needed to assist the selection team in selecting final project titles efficiently. One technique that can help the selection team in selecting final project titles efficiently is to create a model for text classification. As explained in[7], the application of text classification allows almost anything to be organized, arranged, and categorized based on its class. Text classification, which typically involves assigning labels to documents based on their content, is a challenge in text mining. This can involve single label problems (one text to one label) or multiple label problems (one text to many labels)[8].

The current research proposes a text classification method using machine learning natural language processing (NLP) to address these issues. Because it allows the system to operate more freely, the use of categorisation using Natural Language Processing (NLP) is increasingly being developed. As can be seen above, the use of NLP is considered to be more efficient than the use of conventional techniques. There are various approaches to Natural Language Processing (NLP), including [9]that using the LSTM method to classify the sentiment of online coronavirus discussion topics, Classification of COVID-19 related tweets in Nepal using a set of CNN's[10].

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

This research will use the latest development algorithm of neural networks, namely Bidirectional Long Short-Term Memory (BiLSTM) in making the final project title classification model because from research [11]has shown that the Bi-LSTM method is one of the most successful methods in the context of text classification by comparing the use of Bi-LSTM in sentiment analysis with neural network (RNN), convolutional neural network (CNN), traditional LSTM, and Naïve Bayes (NB) methods. The results show that Bi-LSTM outperforms the other methods, producing better contextual information and higher precision, recall and F1 than other methods. To achieve this goal [12] In this research, data processing will be carried out from before and after so that it can obtain more accurate data. This is done in order to understand the context of the data better and ensure the classification of the final project title based on the class is more precise[13].

There have been many previous studies on the use of text classification especially in classifying final project titles[14][15][16]. Although there have been many previous studies conducted related to classifying final project titles, these studies only classify based on the research topic and have not used the latest algorithms. This time, the researcher attempted different research by classifying the final project title based on its score using the latest method of neural networks by using several experiments in labeling.

In addition, many melting experiments were run on the data to determine which melting was best for this investigation. In the future, it is hoped that this will speed up the process of selecting project titles. The final value recorded as a reference for the grouping of final project title.

The purpose of this research is to create an efficient final project title classification model by developing the Bidirectional Long Short-Term Memory (Bi-LSTM) method with the hope of helping the selection team to classify final project titles more efficiently at the Faculty of Engineering, Nurul Jadid University.

II. METHODOLOGY

In this study, as like Fig. 1 after identifying the problem, namely in the process of selecting the final project title which is still done manually, the next step is for researchers to conduct a literature study of several scientific publications on text classification, deep learning, NLP, BiLSTM, and python.

Only then can an action be taken to achieve the goal of this research, namely to classify the final project title based on the value listed in the dataset into a "grade" label using Bidirectional LSTM. First collect the dataset of the recapitulation of the final project title grades of the previous few years through the SIAMTEK account owned by the lecturer, then dilute the data by doing several experiments.

Several labelling trials were performed on the data, the first of which was to round the values first and then label based on the rounding results. The second trial is to label based on a predefined range of values.

Furthermore, pre-processing the data by lowercasing, removing spaces, removing single characters, and removing irrelevant letters, and performing text tokenization and creating sequence data. Next, divide the pre-processed data into 80% training data and 20% testing data. We train the Bidirectional LSTM model on the training data and evaluate its performance on the testing data.

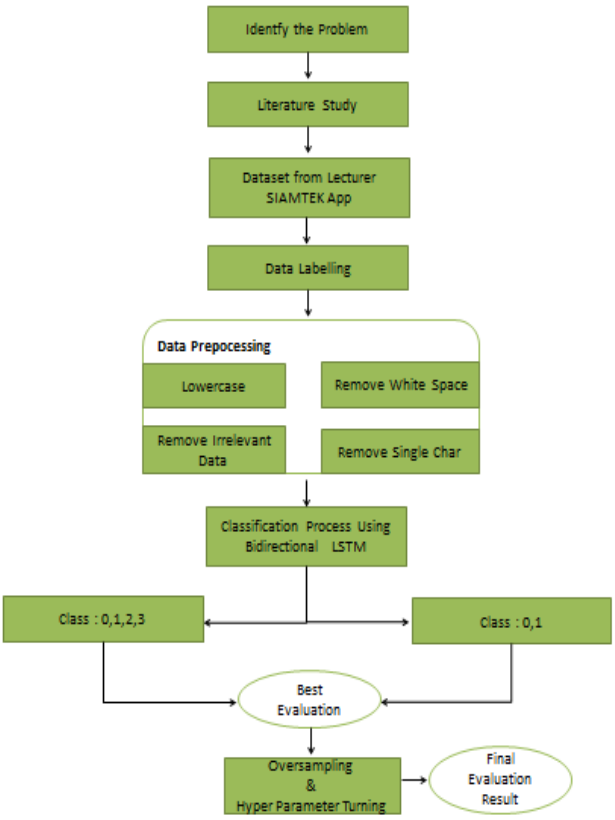


Fig. 1. Flow of Research

A. Dataset

In this study, the dataset used consists of 1149 data on titles and final assignment grades from students in the Faculty of Engineering between 2018 and 2021. The data are divided 80:20 into training and testing data. The training data, totalling 1064 data, is used to build the model, while the testing data, totalling 266 data, is used to evaluate the accuracy of the model's predictions.

Next, the data was labeled with several experiments. first, from the rounding results, poor, fair, good, and excellent classes were produced. The details of the number of each class are as Fig. 2:

```
df['grade'].value_counts()

Baik          665
Cukup         394
Sangat Baik    84
Buruk           6
```

Fig. 2 Details of Value in Each Class

Another experiment is to make a dilation based on the value range 67-79 being the 'Good Enough' category and the value

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

range 80-95 being in the 'Good' category, and to delete data that are very far from the other classes, namely the value range '50-64', so that the total data used is 1143 data with details as shown in Fig. 3.

```
df['grade'].value_counts()

cukup_baik    804
baik          339
```

Fig. 3 Details of Value in Other Experiments

In the initial process of selecting which data to use, in Fig. 4 is the initial data and the results of the data selection process in this study only 'title' and 'value' as shown in Fig. 5 are used as labels.

	judul	nilai	grade
0	Aplikasi perizinan santri di pondok pesantren ...	71	cukup_baik
1	SISTEM MONITORING KEGIATAN BADAN PEMBINAAN KHU...	80	baik
2	APLIKASI PENCATATAN KEUANGAN PADA PANTI ASUHAN...	73	cukup_baik
3	PENERAPAN PENGAJARAN NAHWU DALAM PENGAJARAN BA...	70	cukup_baik
4	Sistem Pendukung Keputusan Penilaian Kinerja K...	70	cukup_baik

Fig. 4 Data before the selection process

	judul	grade
0	Aplikasi perizinan santri di pondok pesantren ...	cukup_baik
1	SISTEM MONITORING KEGIATAN BADAN PEMBINAAN KHU...	baik
2	APLIKASI PENCATATAN KEUANGAN PADA PANTI ASUHAN...	cukup_baik
3	PENERAPAN PENGAJARAN NAHWU DALAM PENGAJARAN BA...	cukup_baik
4	Sistem Pendukung Keputusan Penilaian Kinerja K...	cukup_baik

Fig. 5 Data after the selection process

The next step is to select the data to be used in the next process and convert the target variable to a numeric as in Fig. 6 value so that the system can process it.

	judul	grade
0	aplikasi perizinan santri di pondok pesantren raudlatul fatah di desa maron berbasis web	0
1	sistem monitoring kegiatan badan pembinaan khusus (bpk) putri mts nurul jadid berbasis web dan android	1
2	aplikasi pencatatan keuangan pada panti asuhan nahdlatul ulama probolinggo berbasis web dengan menggunakan codeigniter	0
3	penerapan pengajaran nahwu dalam pengajaran bahasa arab di pondok pesantren nurul islam alhamidy berbasis android	0
4	sistem pendukung keputusan penilaian kinerja karyawan di pt tanjung windu .dengan metode fuzzy dan ahp	0

Fig. 6 Result of convert target data

B. *Prepocessing Data*

Text processing is used to transform the unstructured input into more structured data that can be examined with high accuracy by the model. The text processing techniques performed sequentially include data cleaning and tokenization.

During the data cleaning process, the data in the "title" column will be converted to lowercase, spaces will be removed, single characters will be removed, and unimportant letters will be removed which shown in the Fig. 7.

	judul
0	aplikasi perizinan santri di pondok pesantren raudlatul fatah di desa maron berbasis web
1	sistem monitoring kegiatan badan pembinaan khusus (bpk) putri mts nurul jadid berbasis web dan android
2	aplikasi pencatatan keuangan pada panti asuhan nahdlatul ulama probolinggo berbasis web dengan menggunakan codeigniter
3	penerapan pengajaran nahwu dalam pengajaran bahasa arab di pondok pesantren nurul islam alhamidy berbasis android
4	sistem pendukung keputusan penilaian kinerja karyawan di pt tanjung windu .dengan metode fuzzy dan ahp

Fig. 7 Sample data after preprocessing

The next stage is tokenization, which is the technique of splitting up words in a sentence with the aim of making the words form an array to facilitate analysis[17]. than step includes padding, which is used to balance the length of each sequence.

C. *Bidirectional Long Short Term Memory*

Then the data is divided 80:20 into training and testing data. While training data is used to build the model, testing data is used to evaluate the prediction accuracy of the model. The BI-LSTM model, which belongs to the Deep Learning model, is used in the modeling stage. Deep Learning can be used to classify text and image data. To disable inactive perceptrons and prevent overfitting[18] a dropout layer is also used.

In this stage, the architecture design of the latest development model of the neural network is carried out, which has the best chance of accuracy[19]. BiLSTM can be considered as a neural network with a programmable design, so that its shape can be adjusted to the application[20]. In addition, research[21]discusses the advantages of BiLSTM itself which is able to improve RNN which only has short-term memory and is unable to process long sequential data BiLSTM itself is a variation of the LSTM model that has two layers whose directions are opposite to each other. Fig. 8 illustrates how two LSTM, one moving forward and one moving backward, form a bidirectional long short-term memory (BiLSTM). Data can be stored in both forward and backward directions in this network[22].

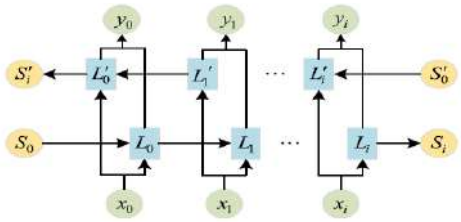


Fig. 8 The BiLSTM Structure

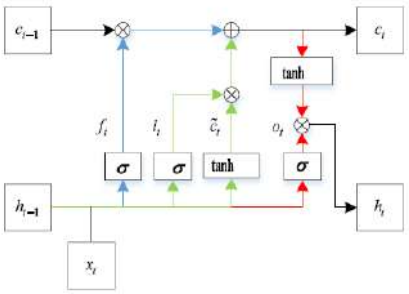


Fig. 9 LSTM's Cell Structure

As shown in Fig. 9, *L* and *L'* in Fig. 8 are the cell architecture

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

of the LSTM. The LSTM consists of a number of unit cells, each of which has an input gate(i_t), an output gate(o_t) and a forgetting gate(f_t).

Foot chooses whatever facts from the past to throw away. It determines which data is fed into the network and the output value is between 0-1, as shown in formula (1); it also determines which layer is used for the output, as illustrated in the formula (2); it also determines which structure is used for the output and the output value is within 0-1, as presented in formula (3); it also determines which of the current units is output to the hidden layer (ht), as shown in formula (4); and it also determines the predicted final output ($\hat{y}_t$) as appears in formula (5).

$$f_t = \sigma\left(w_{fx} \cdot \chi_t + w_{fh} \cdot h_{t-1} + b_f\right) \quad (1)$$

$$i_t = \sigma(w_{ix} \cdot \chi_t + w_{ih} \cdot h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma(w_{ox} \cdot \chi_t + w_{oh} \cdot h_{t-1} + b_o) \quad (3)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (4)$$

$$\hat{y}_t = f\left(w_{yh} \cdot h_t + b_y\right) \quad (5)$$

D. Hyper Parameter Turning

The fast-developing subset of machine learning known as deep learning has proven to be just that. By mimicking the workings of the human brain, it seeks to identify patterns and predict outcomes in the real world. There are applications for models based on this type of neural network topology in almost every sector of the economy. Training such a model to make reliable predictions on any new test data is perhaps the most crucial stage of all the (integral) processes. To ensure that training is efficient in terms of time and fit (whether the model "knows" the training data too well or too poorly; to limit any form of overfitting or underfitting), it is important to choose a model's hyperparameters carefully[23].

The researchers will test a number of variables to see which will improve the accuracy of the model used, having identified which melting tests produced the most accurate results. In most cases, hyperparameters are static (they don't change during training) model variables that are provided by the user prior to training and that control the behaviour of the model and the overall architecture of the ANN[24]. In this study, the researcher also performed a manual hyperparameter turning process to identify good hyperparameters and tune them using different random state variables, number of epochs, dropouts, batch size, activation type and number of epoch. The details are listed in Table I.

TABLE I. MODEL TURNING EXPERIMENT DETAILS

Variables	Trials
Random state	0-100
Epoch	10-100
Bathc size	32
	64
	128
	256

E. Hyper Parameter Turning

For this stage, the Bidirectional Long Short-Term Memory method is tested in classifying the final project title based on its grade. The test data consists of 266 data. In order to determine the degree of success in classifying the final project title, calculations are made [25]using a formula such as Equation (6).

$$Accuracy = \frac{Number\ of\ data\ detected}{Number\ of\ undetected\ data} \quad (6)$$

III. RESULT AND ANALYSIS

In order to assist the selection committee in classifying dissertation titles according to their grade, several experiments were conducted in this study to develop a classification model for dissertation titles. Specifically, two studies with different data labels were considered. In two different tests, the researcher compared the performance of the proposed model. For each of these cases, the experimental results are discussed below. Our technique was implemented in all experiments using the Python programming language and Google Colab.

A. First Experiment

First experiment related to different types of labels used by using the same number of attributes or variables in the model building process, the first type of label uses four (4) labels including; poor, fair, good and very good, resulting in an accuracy value of 56.46% and a loss of 88.21% for training data and for testing data resulted in an accuracy of 63.59% and a loss of 87.29 as in Fig.10.

Fig. 10. Accuracy and loss results from the first trial

B. Second Experiment

Whereas in the results of the second experiment by deleting data whose number of values is very far from other values and producing a dilation with the provisions that values from 67-79 belong to the "good enough" category and values from 80 to 95

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

are in the "good" category. In Fig. 11 conclude this experiment, in an accuracy value of 75.24% and a loss of 51.77% for training data and for testing data resulted in an accuracy of 68.85% and a loss of 60.18.

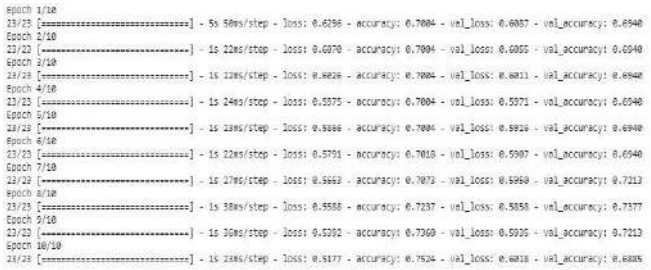


Fig. 11 Accuracy and loss results from the second trial

From the results of these experiments we can conclude that the second experiment has better accuracy results by using two labels, namely 'good enough' and 'good', because in the first experiment 'bad value' has a value that is very far from other data, so it can be one of the causes. But in this second experiment it still needs to be optimised again because the accuracy created is still not good enough to be used as a model. The method used in this research is to balance the data and rotate several parameters in the modelling to further optimise the accuracy.

C. *Oversampling Dataset*

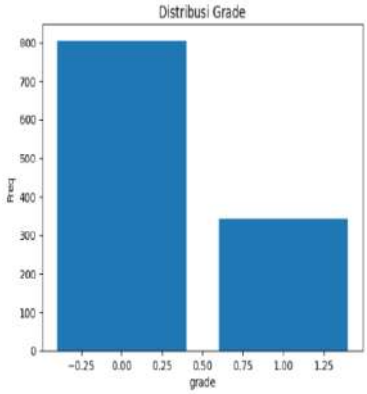


Fig. 12 The proportion of each label in the second experiment

It can be seen in Fig. 12 that the class grouping is not balanced, namely 804 data in the "good enough" class and 339 data in the "good" class, so it is necessary to balance the data. In this case, the researcher took the initiative to perform oversampling techniques on the data, also considering the limitations of this research, because deep learning requires a lot of data for the training process [26].

To solve the problem of classification on unbalanced data sets, this research has claimed [27] that oversampling is a method that is arguably superior to other methods because it can restore data balance while maintaining the characteristics of the original data. It does this by balancing data classes that have less data (minority) with data classes that have more data (majority). So, because the minority data is in the "good" class, the amount of data is balanced to the "good enough" class, which is 804 data as in fig. 13.

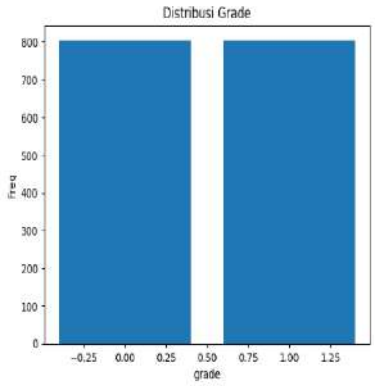


Fig. 13 Data after oversampling process

D. *Turning Parameter*

The next step is for the researchers to do further data processing to re-optimize the best temporary accuracy results by turning on several predetermined parameters, and the following are the best parameter results from several experiments.

The best result from several trials of Hyperparameter Turning on BiLSTM model building is by running trials with random state 10, Bathc size 64 and use 100 epoch as showin Table II. After all models have been built, the model is first configured using "Adam" optimisation and binary crossentropy optimisation to determine the loss value of the developed model.

TABLE II. BEST RESULT PARAMETER

Parameters	Best Results
Epoch	50
Random state	10
Bathc size	64

Adjustments were also made to the construction of the Bidirectional Long-Term Memory model by adding a dropout layer in front of the hidden layer and giving the hidden layer relu activation. The details of the model building architecture are shown in Fig. 14 below.

```
# Sequential Layer
model = Sequential()
# Input Layer, Embedding Layer
model.add(Embedding(len(word_index) + 1, 300, input_length=X_train.shape[1], trainable=False))
# Dropout Layer 1 -> untuk mencegah terjadinya overfitting
model.add(Dropout(0.4))
# Hidden Layer
model.add(Bidirectional(LSTM(64, activation = "relu")))
# Dropout Layer 2 -> untuk mencegah terjadinya overfitting
model.add(Dropout(0.2))
# Dense Layer
model.add(Dense(64, activation = "relu"))
# Dropout Layer 3 -> untuk mencegah terjadinya overfitting
model.add(Dropout(0.4))
# Output Dense Layer -> untuk prediksi
model.add(Dense(1, activation='sigmoid'))
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
print(model.summary())
```

Fig. 14 Model Architecture Details

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

Based on these conclusions, the researchers used the best results from the rotation process of several parameters mentioned above in the final project title classification model using the Bidirectional Long Short-Term Memory method, which resulted in an accuracy value of 90.18% and a loss of 24.72% for training data and for testing data resulted in an accuracy of 83.33% and a loss of 46.72% as in Fig. 15. And the graph of each epoch is as in Fig. 16, where in the graph image the loss model is stable between the loss of the training data and the test data. This shows that the model with the following parameters is a good model.

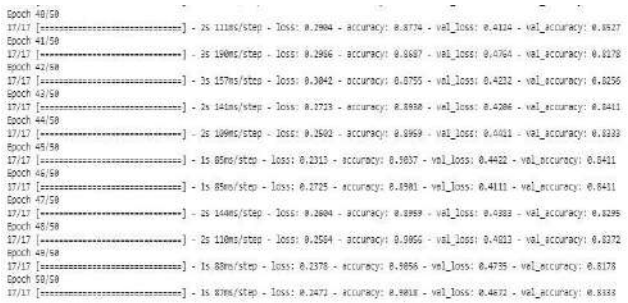


Fig. 14 Results of accuracy and loss values

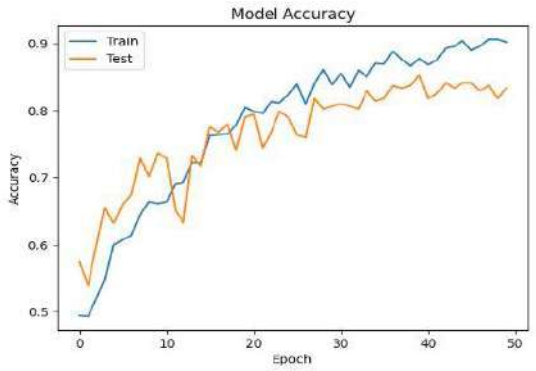


Fig. 15 Graph of best turning performance

And in Fig. 16, where the graph shows that the loss model starts to stabilise between the loss of train data and the test data. This shows that the model with the following parameters is a good model (no overfitting).

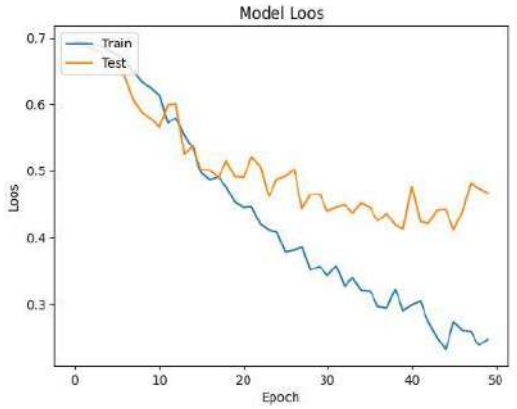


Fig. 16 Graph of loss model

IV. CONCLUSION

This paper presents an improved NLP model for the classification of final project titles based on their grade. The proposed model is based on the Bi-LSTM system, which is tested using two different labelling experiments performed on its dataset obtained from the lecturer's SIAMTEK account. The experimental results show that the two-category labelling has better accuracy and also the turning of some parameters in the model building. The proposed model can be used by the selection committee of Faculty of Engineering, Nurul Jadid University to help them to classify the final project title faster, thus facilitating them in the selection process.

In this study, the researcher performs an oversampling technique to balance the overlapping data and rotated several parameters in the model building, including random state variables, batch size, and number of epoch to improve the accuracy results. Until the best results were obtained by rotating several parameters, namely random state 10 times batch size of 64, and 50 epoch. This resulted can increase the accuracy.

In addition, in this study, we only used the values listed in the dataset in the process of determining the labels. However, this research has its own advantages and differences from previous research, namely in terms of the target classifier.

REFERENCES

[1] A. Arizal, A. N. Puteri, F. Zakiyabarsi, and D. F. Priambodo, "Metode Prototype pada Sistem Informasi Manajemen Tugas Akhir Mahasiswa Berbasis Website," *J. Teknol. Inf. dan Komun.*, vol. 10, no. 1, pp. 1–8, 2022.

[2] T. Tuononen and A. Parpala, "The role of academic competences and learning processes in predicting Bachelor's and Master's thesis grades," *Stud. Educ. Eval.*, vol. 70, p. 101001, Sep. 2021, doi: 10.1016/J.STUEDUC.2021.101001.

[3] N. Nurdin, M. Suhendri, Y. Afrilia, and R. Rizal, "Klasifikasi Karya Ilmiah (Tugas Akhir) Mahasiswa Menggunakan Metode Naive Bayes Classifier (NBC)," *SISTEMASI*, vol. 10, no. 2, pp. 268–279, 2021.

[4] R. Rais, A. Rakhman, and A. H. Sulasmoro, "Rekomendasi Tema Tugas Akhir Menggunakan Metode Clasifikasi Supervised Learning," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 11, no. 3, pp. 535–540, 2022.

[5] B. D. Laraswati, "Human Error: Pengertian dan Jenis-Jenisnya," 2022. https://blog.algorit.ma/human-error-pengertian-dan-jenis-jenisnya/ (accessed Mar. 15, 2023).

[6] M. Liang and T. Niu, "Research on Text Classification Techniques Based on Improved TF-IDF Algorithm and LSTM Inputs," *Procedia Comput. Sci.*, vol. 208, pp. 460–470, Jan. 2022, doi: 10.1016/J.PROCS.2022.10.064.

[7] A. Cindy Rahayu, "Klasifikasi Teks – MTI," 2020. https://mti.binus.ac.id/2020/09/03/klasifikasi-teks/ (accessed Mar. 15, 2023).

[8] H. Ying, C. Deng, Z. Xu, H. Huang, W. Deng, and Q. Yang, "Short-term prediction of wind power based on phase space reconstruction and BiLSTM," *Energy Reports*, vol. 9, pp. 474–482, Sep. 2023, doi: 10.1016/j.egyr.2023.04.288.

[9] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 10, pp.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

2733–2742, 2020.

[10] C. Sitaula and T. B. Shahi, "Multi-channel CNN to classify nepali covid-19 related tweets using hybrid features," *arXiv Prepr. arXiv2203.10286*, 2022.

[11] H. Wu, Z. He, W. Zhang, Y. Hu, Y. Wu, and Y. Yue, "Multi-class text classification model based on weighted word vector and bilstm-attention optimization," in *Intelligent Computing Theories and Application: 17th International Conference, ICIC 2021, Shenzhen, China, August 12–15, 2021, Proceedings, Part I 17*, 2021, pp. 393–400.

[12] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 89–99, 2022.

[13] R. Z. Insani, C. Setianingsih, and T. W. Purboyo, "Analisis Sentimen Komentar Berdasarkan Geo Tagged Menggunakan Algoritma Bilstm," *eProceedings Eng.*, vol. 10, no. 1, 2023.

[14] A. P. Wibowo, W. Widiyono, A. Saifudin, and A. S. Darmawan, "Naïve Bayes, Neural Network dan K-Nearest Neighbor untuk Klasifikasi Topik Tugas Akhir," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 11, no. 4, pp. 661–667, 2022.

[15] N. Andriani and A. Wibowo, "Implementasi Text Mining Klasifikasi Topik Tugas Akhir Mahasiswa Teknik Informatika Menggunakan Pembobotan TF-IDF dan Metode Cosine Similarity Berbasis Web," in *Prosiding Seminar Nasional Mahasiswa Bidang Ilmu Komputer dan Aplikasinya*, 2021, vol. 2, no. 2, pp. 130–137.

[16] S. Sumarni and S. Rustam, "Klasifikasi Topik Tugas Akhir Mahasiswa menggunakan Algoritma Particle Swarm Optimization dan K-Nearest Neighbor," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 168–175, 2020.

[17] F. N. Fajri and S. Syaiful, "Klasifikasi Nama Paket Pengadaan Menggunakan Long Short-Term Memory (LSTM) Pada Data Pengadaan," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, pp. 1625–1633, 2022.

[18] W. Afandi, S. N. Saputro, A. M. Kusumaningrum, H. Adriansyah, M. H. Kafabi, and S. Sudianto, "Klasifikasi Judul Berita Clickbait menggunakan RNN-LSTM," *J. Inform. J. Pengemb. IT*, vol. 7, no. 2, pp. 85–89, 2022.

[19] A. D. D. Wibiyanto and A. Wibowo, "Penerapan Algoritma Multiclass Support Vector Machine dan TF-IDF Untuk Klasifikasi Topik Tugas Akhir," *SKANIKA Sist. Komput. dan Tek. Inform.*, vol. 6, no. 1, pp. 42–50, 2023.

[20] E. Subowo, F. A. Artanto, I. Putri, and W. Umaedi, "BLTSM untuk analisis sentimen berbasis aspek pada aplikasi belanja online dengan cicilan," *J. FASILKOM*, 2022.

[21] M. Kowsher *et al.*, "LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy," *Procedia Comput. Sci.*, vol. 193, pp. 131–140, Jan. 2021, doi: 10.1016/J.PROCS.2021.10.013.

[22] J. Song, "Aspect-Based Sentiment Analysis on Mobile Game Reviews Using Deep Learning," *Grad. Sch. Comput. Inf. Sci. Hosei Univ*, 2020.

[23] K. Chowdhury, "10 Hyperparameters to keep an eye on for your LSTM model — and other tips | by Kuldeep Chowdhury | Geek Culture | Medium," *Geek Culture*, 2021. https://medium.com/geekculture/10-hyperparameters-to-keep-an-eye-on-for-your-lstm-model-and-other-tips-f0ff5b63fcd4 (accessed May 09, 2023).

[24] D. Passos and P. Mishra, "A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks," *Chemom. Intell. Lab. Syst.*, vol. 223, p. 104520, Apr. 2022, doi: 10.1016/J.CHEMOLAB.2022.104520.

[25] F. N. Fajri, A. Tholib, and W. Yuliana, "Penerapan Machine Learning untuk Penentuan Mata kuliah Pilihan pada Program Studi Informatika," *J. Tek. Inform. dan Sist. Inf. p-ISSN*, vol. 2443, p. 2210.

[26] Lakshana, "Artificial Intelligence Vs Machine Learning Vs Deep Learning," *Analytics Vidhya*, 2021. https://www.analyticsvidhya.com/blog/2021/05/ai-ml-dl/ (accessed May 11, 2023).

[27] G. Wei, W. Mu, Y. Song, and J. Dou, "An improved and random synthetic minority oversampling technique for imbalanced data," *Knowledge-Based Syst.*, vol. 248, p. 108839, Jul. 2022, doi: 10.1016/J.KNOSYS.2022.108839.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1723, Copyright ©2023
Submitted : May 15, 2023, Revised : June 16, 2023, Accepted : July 4, 2023, Published : November 4, 2023

Prediksi Kelulusan Mahasiswa Fakultas Teknologi Informasi ISB Atma Luhur Menggunakan Algoritma C4.5

Ine Widyaningrum Mustama Putri[1], Rusdah Rusdah[2], Lis Suryadi[3], Dian Anubhakti[4]
Program Studi Sistem Informasi
Fakultas Teknologi Informasi Universitas Budi Luhur
Jakarta, Indonesia
1912500939@student.budiluhur.ac.id[1], rusdah@budiluhur.ac.id[2], lis.suryadi@budiluhur.ac.id[3],
dian.anubhakti@budiluhur.ac.id[4]

***Abstract*— Higher Education is a level of education after secondary education which includes diploma programs, undergraduate programs, master programs, doctoral programs, professional programs, and specialist programs organized based on the culture of the Indonesian nation. Student graduation is one of the important factors to improve university accreditation. Students who graduate above 5 years and the number of students who drop out are important indicators in determining accreditation which then causes the difficulty of accrediting a college to rise. This research aims as an early warning for students who graduate on time and graduate late from the Faculty of Information Technology, Institute of Science and Business Atma Luhur using the C4.5 decision tree algorithm by implementing the Cross-Industry Standard Process for Data Mining (CRISP- DM) method. The initial data of this research amounted to 1,015 which was taken through a query in the database of the Atma Luhur Institute of Science and Business. However, the data that will be used becomes 694 after preprocessing due to the large number of record contents that do not have a graduation year, with a total of 641 graduates graduating on time and 53 graduates graduating late. Based on the application of the model using the C4.5 decision tree algorithm and the Confusion Matrix method, the accuracy is 93.94%, Recall is 98.59%, and Precision is 95.03%. So it can be concluded that the C4.5 decision tree algorithm is the most effective algorithm for predicting student graduation, because it has a high level of accuracy.**

Keywords*— Prediction, Graduation, Decision Tree C4.5, Early Warning

***Abstrak*— Perguruan Tinggi merupakan jenjang pendidikan setelah pendidikan menengah yang mencakup program diploma, program sarjana, program magister, program doktor, program profesi, dan program spesialis yang diselenggarakan berdasarkan kebudayaan bangsa Indonesia. Kelulusan mahasiswa menjadi salah satu faktor penting untuk meningkatkan akreditasi perguruan tinggi. Mahasiswa yang lulus diatas 5 tahun dan jumlah mahasiswa yang mengalami drop out merupakan indikator penting dalam penentuan akreditasi yang kemudian menyebabkan sulitnya akreditasi sebuah perguruan tinggi naik. Penelitian ini bertujuan sebagai early warning atau peringatan dini untuk kelulusan mahasiswa yang lulus tepat waktu dan lulus terlambat fakultas teknologi informasi, Institut Sains dan Bisnis Atma Luhur menggunakan algoritma decision tree C4.5 dengan mengimplementasikan metode Cross-Industry Standard Process for Data Mining (CRISP- DM). Data awal penelitian ini berjumlah 1.015 yang diambil melalui query dalam database ISB Atma Luhur. Namun, data yang akan digunakan menjadi 694 setelah dilakukan preprocessing dikarenakan banyaknya isi record yang tidak memiliki tahun kelulusan, dengan total 641 wisudawan lulus tepat waktu dan 53 wisudawan lulus terlambat. Berdasarkan penerapan model menggunakan algoritma decision tree C4.5 dan metode Confusion Matrix diperoleh Akurasi sebesar 93.94%, Recall sebesar 98.59%, dan Presisi sebesar 95.03%. Maka dapat disimpulkan bahwa algoritma decision tree C4.5 adalah algoritma yang paling efektif untuk memprediksi kelulusan mahasiswa, karena memiliki tingkat akurasi yang tinggi.**

Kata Kunci*— Prediksi, Kelulusan, Decision Tree C4.5, Early Warning

I. PENDAHULUAN

Perguruan tinggi merupakan suatu instansi yang menyelenggarakan proses belajar, mengajar, penelitian serta pengabdian kepada masyarakat atau biasa disebut lembaga penyelenggara Tri Dharma Perguruan Tinggi. Setiap tahun perguruan tinggi menghasilkan lulusan mahasiswa sesuai bidang yang ditempuhnya, dalam kurun waktu penyelesaian studi tepat waktu. Lulus tepat waktu merupakan salah satu indikator keberhasilan mahasiswa dalam memperoleh gelar sarjana. Mahasiswa lulus tepat waktu apabila mampu menyelesaikan studinya di perguruan tinggi selama kurang dari atau sama dengan empat tahun, sedangkan mahasiswa dikatakan tidak lulus tepat waktu apabila menyelesaikan studinya lebih dari empat tahun. Namun tidak setiap mahasiswa dapat menyelesaikan pendidikan sarjana dalam kurun waktu 4 (empat) tahun di perguruan tingginya, mahasiswa yang telah menyelesaikan pendidikan sarjana kemudian mendaftar sebagai calon wisudawan baru kemudian bisa mengikuti wisuda. Berdasarkan Peraturan Menteri Riset Teknologi dan Pendidikan Tinggi Nomor 44 Tahun 2015 tentang Standar Nasional Pendidikan Tinggi terkait dengan beban belajar dan masa belajar mahasiswa program sarjana paling lama 7 (tujuh) tahun.

ISB (Institut Sains dan Bisnis) Atma Luhur adalah salah satu perguruan tinggi dalam bidang komputer dan bisnis digital di Provinsi Kepulauan Bangka Belitung khususnya di kota Pangkal Pinang. ISB Atma Luhur berdiri pada tahun 2020

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

merupakan pengembangan atau rubah bentuk dari Sekolah Tinggi Manajemen Informatika dan Komputer Atma Luhur yang berdiri sejak tahun 2009. Saat ini telah berdiri dan berkembang beberapa sekolah tinggi dan universitas di Provinsi Kepulauan Bangka Belitung, sehingga lingkungan ISB Atma Luhur Fakultas Teknologi Informasi berada dalam lingkungan yang kompetitif.

Persentase kelulusan tepat waktu pada tahun ajaran 2015/2016 adalah 46%, sedangkan pada tahun ajaran 2016/2017 persentase kelulusan tepat waktu adalah 65% dan pada tahun ajaran 2017/2018 persentase kelulusan tepat waktu adalah 61%. Sehingga rata-rata persentase kelulusan tepat waktu tahun ajaran 2015/2016, 2016/217, 2017/2018 adalah 57.3%. Pada tahun ajaran 2015/2016 mahasiswa baru berjumlah 376, mahasiswa aktif sejumlah 257 sedangkan mahasiswa yang lulus tepat waktu tahun ajaran 2019/2020 berjumlah 233. Pada tahun ajaran 2016/2017 mahasiswa baru berjumlah 295, mahasiswa aktif sejumlah 209 sedangkan yang lulus tepat waktu tahun ajaran 2020/2021 berjumlah 194. Pada tahun ajaran 2017/2018 mahasiswa baru berjumlah 344, mahasiswa aktif sejumlah 342 sedangkan yang lulus tepat waktu tahun ajaran 2021/2022 berjumlah 214. Maka manajemen membutuhkan penelitian yang dapat memprediksi kelulusan mahasiswanya tepat waktu.

Salah satu teknik melakukan prediksi yang dapat dilakukan adalah dengan teknik penggalian data atau *data mining*. *Data mining* adalah proses untuk mendapatkan informasi yang berguna dari basis data yang besar dan perlu diekstraksi agar menjadi informasi baru dan dapat membantu dalam pengambilan keputusan. Dengan salah satu teknik dari *data mining* yaitu klasifikasi dengan algoritma C4.5 [1]. Algoritma C4.5 dipilih karena mampu memprediksi dengan memberikan tingkat nilai akurasi yang ideal dan hasil yang optimal untuk memprediksi yang diharapkan dapat menemukan informasi tingkat kelulusan dan persentase kelulusan mahasiswa sehingga dapat digunakan oleh pihak manajemen untuk mencari solusi dalam proses evaluasi pembelajaran di Fakultas Teknologi Informasi ISB Atma Luhur [2].

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode prediksi yang berguna untuk mengeksplorasi informasi dari sebuah dataset, dalam penelitian ini digunakan dataset kelulusan mahasiswa Fakultas Teknologi Informasi ISB Atma Luhur tahun 2015-2017. Pada penelitian ini, penulis memprediksi kelulusan mahasiswa menggunakan metode data mining menggunakan *tools* RapidMiner untuk menunjang penelitian yang dilakukan.

Penelitian dengan menggunakan algoritma C4.5, tersebut menggunakan perbandingan algoritma C4.5 dan KNN dan menghasilkan data keakuratan yang tinggi diatas 80% [3].

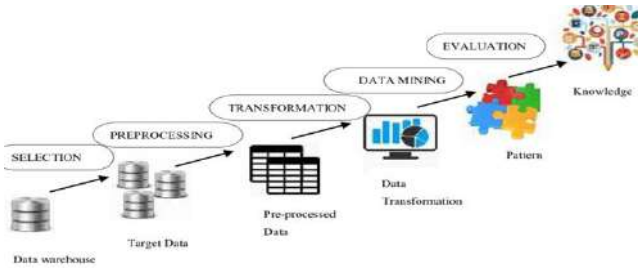
Penelitian lain yang menggunakan algoritma C4.5, tersebut uji coba untuk nilai akurasi dilakukan sebanyak 3 kali dengan jumlah *record data training* yang berbeda. Semakin banyak jumlah *record* data yang digunakan untuk proses data training maka semakin tinggi juga nilai akurasinya[4].

II. STUDI PUSTAKA

A. *Data Mining*

Data mining merupakan suatu proses penggalian atau penambangan informasi untuk menemukan informasi yang tidak ditemukan sebelumnya dari sekumpulan data besar (*big data*) atau *repository database* lainnya. *Data mining* bukan sekedar terkumpul data saja tetapi mencangkup analisis dan prediksi dari informasi yang ingin ditampilkan [5]. Data yang dikumpulkan disimpan dalam *database* kemudian diproses sehingga dapat dijadikan untuk pengambilan keputusan dalam melihat informasi yang akan digunakan [6].

*Knowledge Discovery In Databas*e (KDD) merupakan metode untuk memperoleh pengetahuan dari database yang ada. Dalam *database* terdapat tabel - tabel yang saling berhubungan / berelasi. Hasil pengetahuan yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (*knowledge base*) untuk keperluan pengambilan keputusan. Istilah *Knowledge Discovery in Database* (KDD) dan *data mining* seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* [7].



Gambar 1. Proses KDD

B. Tahapan proses KDD ada 5 [8] yaitu :

1) *Data Selection:* pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai.

2) *Preprocessing:* sebelum proses data mining dapat dilaksanakan perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses *enrichment*, yaitu proses "memperkaya" data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal [9].

3) *Transformation:* yaitu proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam *database*.

4) *Data Mining:* proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu [10].

5) *Interpretation / Evaluation:* pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak.

C. *Metode Decision Tree*

Pohon (*tree*) merupakan sebuah struktur data yang terdiri dari simpul (*node*) dan rusuk (*edge*). Simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar, simpul percabangan dan simpul daun. Pohon keputusan merupakan presentasi sederhana dari teknik klasifikasi untuk sejumlah kelas berhingga, simpul internal maupun simpul akar ditandai dengan nama atribut, rusuk – rusuknya diberi label nilai atribut dan simpul daun ditandai dengan kelas – kelas yang berbeda [11].

Pohon keputusan merupakan sebuah tampilan grafis proses pengambilan keputusan yang mengidentifikasi alternatif yang ada, kondisi alamiah dan peluangnya dan juga imbalan bagi setiap kombinasi alternatif keputusan. Cara kerja pohon keputusan yaitu mengubah bentuk data tabel menjadi model pohon, mengubah model pohon menjadi *rule* (aturan), dan menyederhanakan *rule* (aturan).

D. Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang dapat digunakan untuk mengkonstruksi sebuah pohon keputusan. Algoritma C4.5 merupakan pengembangan algoritma ID3 (Quinlan, 1993), menentukan seberapa informatif sebuah masukan atribut dimana kekurangan yang dimiliki algoritma ID3 ditutupi oleh algoritma C4.5. Empat hal yang membedakan algoritma C4.5 dengan ID3 antara lain: tahan (*robust*) terhadap *data noise*, mampu menangani variabel dengan tipe diskrit maupun kontinu, mampu menangani variabel yang memiliki *missing value*, dan dapat memangkas cabang dari pohon keputusan [12].

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- Pilih atribut sebagai akar
- Buat cabang untuk masing – masing nilai
- Bagi kasus dalam cabang
- Ulangi proses untuk masing – masing cabang sampai semua kasus pada cabang memiliki kelas yang sama. Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut – atribut yang ada, untuk menghitung gain digunakan rumus seperti yang tertera pada persamaan (1).

$$Gain(S,A) = Entropy(\mathrm{S}) - \sum_{n=1}^{n} \frac{|S_i|}{|S|} * Entropy(Si) \qquad \textbf{(1)}$$

Keterangan:

S : Himpunan Kasus
A : Atribut
n : Jumlah partisi atribut A
|Si| : Jumlah kasus pada partisi ke i
|S| : Jumlah kasus dalam S

Sebelum mendapatkan nilai *Gain* adalah dengan mencari nilai Entropi. Entropi digunakan untuk menghasilkan sebuah atribut. Rumus dasar dari Entropi seperti terdapat pada persamaan (2).

$$Entropy(S) = \sum_{i=1}^{n} -pi * log_2\, pi \qquad \textbf{(2)}$$

Keterangan:

S : Himpunan kasus
n : Jumlah partisi S
pi : Proporsi dari Si terhadap S

E. *Cross Validation*

Cross *Validation* atau dapat disebut estimasi rotasi adalah sebuah teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independen, teknik ini utamanya digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model prediktif ketika dijalankan dalam prakteknya. Dalam sebuah masalah prediksi, sebuah model biasanya diberikan kumpulandata (*dataset*) yang diketahui untuk digunakan dalam menjalankan pelatihan (*data training*), serta kumpulan data yang tidak diketahui (*data testing*)terhadap model yang diuji [13].

F. Presisi, *Recall* dan Akurasi

Precission (presisi) dan *recall* digunakan untuk mengukur kinerjasistem. Presisi adalah kecocokan antara bagian data yang diambil dengan informasi yang dibutuhkan. *Recall* merupakan tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. *Accuracy* (akurasi) adalah tingkat kedekatan antara nilai yang didapat terhadap nilai sebenarnya. Presisi,*recall* dan akurasi dapat dihitung menggunakan *confusion matrix*. Ukuran besaran presisi, *recall*, dan akurasi biasanya diberi nilai dalam bentuk presentase antara 1 sampai 100%. Sebuah sistem akan dianggap baik jika tingkat presisi, *recall*, dan akurasi-nya tinggi [14].

G. *Tools* yang Digunakan

Dalam penelitian ini, penulis menggunakan *software* atau *tools*, antara lain

1) RapidMiner versi 9.10: RapidMiner dalam penelitian ini digunakan untuk melakukan tahap *modeling* data. RapidMiner adalah aplikasi atau perangkat lunak yang berfungsi sebagai alat pembelajaran dalam ilmu data mining. Platfrom dikembangkan oleh perusahaan yang didedikasikan untuk semua langkah yang melibatkan sejumlah besar data dalam bisnis komersial, penelitian, pendidikan, pelatihan, dan pembelajaran [15].

2) Microsoft Excel 2011: Program aplikasi Microsoft Excel ini memiliki banyak fitur dan fungsi yang digunakan untuk mengolah angka. Fitur Fungsi dan Formula atau yang lebih dikenal dengan rumus Excel menjadikannya terkenal dan banyak digunakan dalam berbagai bidang dan persoalan seperti membuat, mengedit, mengurutkan, menganalisa, serta meringkas data [16].

Microsoft Excel dalam penelitian ini digunakan untuk mengolah data *reprocessing*. Pengelolaan atau pengolahan data

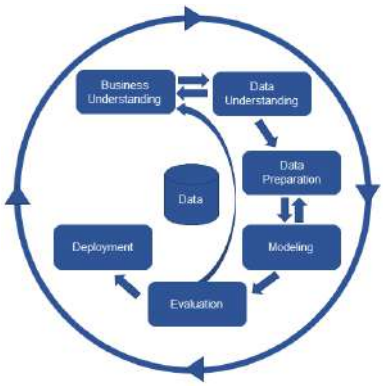
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

berfungsi untuk pengelolaan *database* statistik ,mencari nilai tengah, rata-rata, pencarian nilai maksimum [17].

III. METODOLOGI PENELITIAN

Pada tahapan penelitian dijelaskan apa saja yang dilakukan beserta luaran dari setiap tahapan. Dijelaskan juga bagaimana data dikumpulkan dan luaran dari data tersebut, bagaimana data yang diperoleh dianalisis atau diolah untuk diproses pada tahapan berikutnya.

Pada penelitian ini menggunakan metodologi yakni CRISP-DM untuk melakukan analisis dan mengolah data sebagai pemecah masalah yang umum untuk bisnis dan penelitian. Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment* [18]. Proses metodologi ini terdiri dari 6 tahapan yang dapat dijelaskan sebagai berikut :



Gambar 2. Diagram CRISP-DM

A. Diagram CRISP-DM

1) Pemahaman Bisnis *(Business Understanding)*: pada tahap *business understanding,* dilakukan wawancara langsung dengan narasumber dari Biro Sistem Informasi ISB Atma Luhur. Wawancara ini bertujuan untuk menemukan identifikasi masalah. Hasil dari tahapan ini adalah diketahuinya atau diidentifikasinya rumusan masalah dan tujuan.

2) Pemahaman Data *(Data Understanding)*: pada tahap ini, berdasarkan tujuan penelitian maka dilakukan permintaan data kelulusan mahasiswa tahun ajaran 2015/2016, 2016/2017, dan 2017/2018 melalui *query* pada Biro Sistem Informasi ISB Atma Luhur, setelah itu didapatkan beberapa atribut yang diperlukan oleh penulis untuk penelitian. Data wisudawan yang terkumpul sebanyak sebanyak 1015 *record* dan 57atribut. Pada tahap ini penulis mempelajari data agar dapat mengenal seperti apa data yang akan digunakan untuk keperluan penelitian.

Atribut yang didapatkan adalah program studi, NIM, nama, jenis kelamin, asal SMA, tahun ajaran Lulus, tanggal lulus, keterangan keaktifan perkuliahan semester 1-12, IPS semester 1-12, jumlah SKS yang diambil pada semester 1-12, IPK semester 1-12, jenis kelas, SKS lulus semester 1-12.

3) Persiapan Data *(Data Preparation)*: pada tahap *data preparation*, data yang telah dikumpulkan akan dilakukan *preprocessing* menggunakan Microsoft Excel sebagai berikut:

a) *Data Reduction:* kegiatan *preprocessing* yang dilakukan pada tahapan ini adalah reduksi data yaitu penghapusan atribut yaitu atribut NIM, nama, SKS yang diambil semester 3-12, IPK semester 3-12, keterangan aktif kuliah semester 3-12. Proses penghapusan ini dilakukan karena atribut tersebut sudah tidak dapat digunakan sebagai *early warning* dalam menentukan tepat atau terlambat nya lama masa studi mahasiswa.

b) *Data Cleaning*: terhadap 6 *record missing value* pada atribut SKS semester 2 dan SKS lulus semester 2.

c) *Data Transformation:* tahap ini menggabungkan semua *sheet* tahun ajaran 2015/2016, 2016/2017, dan 2017/2018 ke dalam satu *sheet*, kemudian membentuk data baru yaitu masa studi diperoleh dari tahun lulus dikurangi tahun masuk, dari masa studi ini akan dibuat atribut keterangan lulus.

d) Penentuan Data *Training* dan Data *Testing*: data hasil *preprocessing* dipecah menjadi data *training* (data latih) dan data *testing* (data uji) menggunakan *split* data dengan rasio 60:40, 70:30, dan 80:20.

4) Pemodelan (*Modeling*): pada tahap ini penelitian melakukan pemodelan terhadap dataset yang sudah dilakukan *Preprocessing* dan telah ditentukan *data training* dan *data testing* serta tanpa *split* yaitu *cross validation.* Pada tahapan ini akan dieksplorasi beberapa algoritma untuk klasifikasi untuk mendapatkan model yang terbaik.

5) Evaluasi (*Evaluation)*: pada tahap evaluasi, dilakukan pengukuran kinerja algoritma model, algoritma ini dapat melakukan klasifikasi terhadap prediksi kelulusan mahasiswa menggunakan *confusion matrix* yaitu presisi, *recall*, dan akurasi.

6) Penyebaran (*Deployment*): Tahapan ini dilakukan dengan pembuatan laporan dan artikel jurnal menggunakan model yang dihasilkan.

B. Perbedaan dari penelitian sebelumnya

Dalam penelitian sebelumnya oleh Rizki Muliono, J. H. Lubis, dan Nurul Khairina yang menggunakan *algoritma K-Nearest Neighbor* (KNN) tingkat Akurasi yang didapatkan level K3 lebih baikdari level K1 dan K2 yaitu 98.5% [19].

Dalam penelitian sebelumnya oleh Abdul Rohman dan Anief Rofiyanto yang menggunakan algoritma DecisionTree C4.5 rata-rata tingkat Akurasi nya 65.98% [20].

Dalam penelitian sebelumnya oleh Nurul Khasanah et al. yang menggunakan algoritma *Naïve Bayes* tingkat Akurasi nya rata-rata sebesar 88.16% [21] .

C. *Data Preprocessing*

Data yang digunakan harus melewati *data preprocessing*, karena sumber data yang diperoleh dari database masih bersifat kotor, artinya terdiridari beberapa data yang tidak lengkap, data hilang atau kosong, kekurangan atribut tertentu atau atribut yang sesuai, *data noise*, dan data yang tidak konsisten.

Preprocessing terdiri dari 4 jenis yaitu *data cleaning, data integration, data transformation,* dan *data reduction.* Jumlah data tahun kelulusan 2019, 2020, dan 2021 yang diperoleh sebanyak 694 data. Pada penelitian ini,*preprocessing data* yang akan dilakukan adalah *cleaning data* yaitu digunakan untuk melengkapi atau menghilangkan data yang tidak lengkap, menghapus data atribut yang tidak diperlukan dalam proses

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

klasifikasi, dan memperbaiki data yang tidak konsisten, membentuk atribut baru. *Data preprocessing* dilakukan dengan memfilter secara manual menggunakan microsoft excel.

D. Teknik Pengujian

Pada tahap ini pengujian dilakukan dengan metode *confusion matrix* mempresentasikan hasil evaluasi model. Evaluasi menggunakan *confusion matrix* menghasilkan nilai akurasi, presisi dan *recall*. Akurasi dalam klasifikasi merupakan presentasi ketepatan *record data* diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Presisi merupakan proposikasi yang diprediksi positif yang juga positif benar pada data sebenarnya. *Recall* merupakan proporsi kasus positif yang sebenarnya diprediksi positif secara benar [22].

IV. ANALISIS DAN PEMBAHASAN

Data yang digunakan adalah data sekunder yang didapatkan dari Biro Sistem Informasi ISB Atma Luhur. Data ini ditarik menggunakan *query* dan *function* di Oracle. Data keseluruhan yang digunakan pada penelitian adalah 694. Atribut yang digunakan dari data penelitian adalah program studi, jenis kelamin, asal SMA, keterangan masa aktif kuliah semester 1 dan 2, IPS semester 1 dan 2, jumlah SKS yang diambil pada semester1 dan 2, IPK semester 1 dan 2, jenis kelas,jumlah SKS lulus semester 1 dan 2, dan keterangan lulus.

A. *Data Preprocessing* atau Pra Pemrosesan Data

1) Data *Reduction:* kegiatan *preprocessing* yang dilakukan pada tahapan ini adalah reduksi data yaitu penghapusan atribut yaitu atribut NIM, nama, SKS yang diambil Semester 3-12, SKS Lulus Semester 3-12, IPK Semester 3-12, IPS 3-12, keterangan aktifkuliah semester 3-12. Proses penghapusan ini dilakukan karena atribut tersebut sudah tidak dapat digunakan sebagai *early warning* dalam menentukan tepat atau terlambatnya lama masa studi mahasiswa.

Kemudian dilakukan proses penghapusan data mahasiswa tahun ajaran tahun ajaran 2015/2016, 2016/2017, dan 2017/2018 yang tidak memiliki data tanggal TA dan tanggal lulus. Artinya mahasiswa tersebut tidak lulus, sejumlah 320 mahasiswa. Terdapat 1 *record* yang tidak memiliki nilai pada atribut SKS lulussemester 2, namun memiliki nilai IPS semester 2, sehingga diputuskan untuk menghapus satu *record* tersebut. Sehingga *dataset* setelah proses reduksi data adalah 16 atribut, 694 *record*.

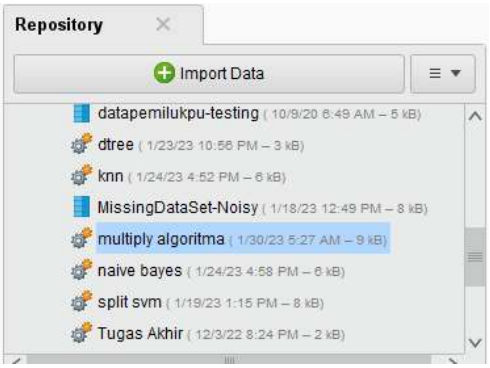
2) Data Cleaning: terhadap 6 *record missing value* pada atribut SKS semester 2 dan SKS lulus semester 2. Karena atribut aktif semester 2 bernilai tidak, maka atribut SKS semester 2 dan SKS lulus semester 2 diisi dengan nilai 0 (nol).

3) Data Transformation: tahap ini menggabungkan semua sheet tahun ajaran 2015/2016, 2016/2017, dan 2017/2018 ke dalam satu sheet, kemudian membentuk data baru yaitu masa studi diperoleh dari tahun lulus dikurangi tahun masuk. Sehingga jumlah atribut menjadi 17 atribut. Berdasarkan atribut masa studi dibuat atribut keterangan lulus. Bila masa studi = 4 tahun, maka keterangan lulus berisi tepat. Bila masa studi > 4 tahun, maka keterangan lulus berisi telat.

4) *Data Reduction* tahap kedua: pada tahap ini atribut tahun masuk, tahun lulus dan masa studi dihapus. Sehingga dataset terdiri atas 14 atribut.

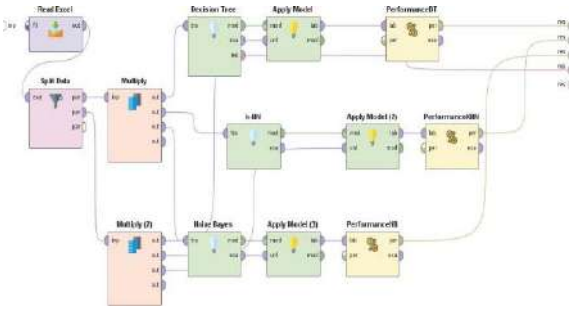
B. Penentuan Data Latih dan Data Uji

Setelah *preprocessing* data, selanjutnya akan dilakukan penentuan data *training* (data latih) dan data *testing* (data uji). Data *training* adalah data yangdigunakan untuk melatih mesin agar dapat mengenali pola, sedangkan data testing adalah data yang digunakan untuk menguji hasil dari pelatihan yang telah dilakukan terhadap mesin [23].



Gambar 3. Penentuan Data Latih dan Data Uji

Pemodelan dilakukan dengan memanfaatkan aplikasi RapidMiner. Proses diawali dengan mengambil data training yang telah disediakan sebelumnya. Ditambahkan dengan operator *Read Excel*, untuk melakukan import data karena data training berupa file excel; operator *Split Data* untuk membagi data; operator *Multiply*, untuk membuat salinan objek data; operator Algoritma C4.5, KNN dan *Naïve Bayes*; operator *Apply Model*, untuk menerapkan model yang telah dilatih sebelumnya menggunakan *data training* pada *unlabeled data* (*data testing*) dan operator *Performance*, untuk mengevaluasi kinerja model [24].



Gambar 4.Proses Pemodelan menggunakan Teknik Split Data.

Penentuan *data training* dan *data testing* dilakukan dengan cara membagi data *(split data)* dengan perbandingan.Dalam penelitian ini data *training* dan data *testing* dibagi menjadi tiga perbandingan yaitu 80:20, 70:30, dan 60:40 [25]. Sebaran data *training* seperti pada tabel 1.

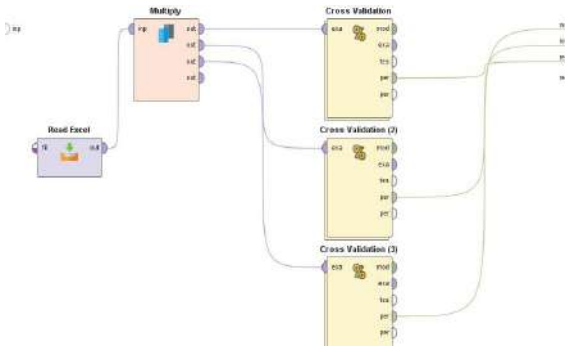
Tabel 1. Tabel Sebaran Data Training

Komposisi *Dataset*	**Data *Training***	**Data *Testing***

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

60 : 40	416	278
70 : 30	486	208
80 : 20	555	139

Teknik *sampling* yang dipilih adalah *stratified random sampling*. Setelah menentukan pembagian perbandingan untuk data *training* dan data *testing*, kemudian membuat persebaran data untuk masing-masing bagian tersebut. Sebagai contoh perbandingan 80:20 yang berarti 80% merupakan data *training* dan 20% merupakan data *testing*.



Gambar 5. Eksplorasi dengan parameter uji akurasi

Pada Gambar 5 dieksplorasi menggunakan teknik penentuan data *training* dan *testing* menggunakan *10-fold cross validation* dengan teknik *stratified random sampling.* Hasil eksplorasi menggunakan *split* data dan *10-fold cross validation* disajikan pada Tabel 2 dengan parameter uji yaitu akurasi.

Tabel 2. Parameter Uji Akurasi

Dataset	**Akurasi**		
	C4.5	**KNN**	***Naïve Bayes***
60:40	92.45%	94.60%	92.09%
70:30	91.83%	93.75%	91.35%
80:20	92.81%	94.24%	92.81%
10-Fold Cross Validation	93.95%	93.67%	91.79%

Setelah melakukan eksplorasi, dapat dilihat pada tabel II bahwa komposisi data *training* dan *testing* yang menghasilkan akurasi tertinggi dengan *classifier* C4.5 diperoleh dari dataset *cross validation* yaitu sebesar 93,95%. Sedangkan pada *classifier* KNN diperoleh dari dataset dengan komposisi 60:40, yaitu sebesar 94,60%. Dan pada *classifier Naïve Bayes* akurasi tertinggi diperoleh dari dataset dengan komposisi 80:20, yaitu sebesar 92,81%.

C. Hasil Komparasi Model

Pada tahap awal pemodelan, dilakukan eksplorasi beberapa algoritma untuk mengetahui performa model terbaik. Parameter uji yang digunakan adalah nilai akurasi, presisi, *recall* dan AUC. Tabel 3 merupakan hasil akurasi algoritma *decision tree C4.5* menggunakan metode 10-*fold cross validation.*

Tabel 3. Hasil Akurasi

Algoritma	**Akurasi**	**Presisi**	***Recall***	**AUC**
C4.5	93.95%	69.00%	39.67%	68.2%
KNN	93.67%	68.33%	30.33%	71.5%
Naïve Bayes	91.79%	46.78%	40.00%	78.5%

D. Hasil Komparasi Model

Setelah dilakukan hasil komparasi model, selanjutnya pada bagian ini, disajikan hasil pemrosesan dari model atau algoritma *Decision Tree* C4.5. Implementasi algoritma *Decision Tree* C4.5 terdiri dari, pembentukan *decision rules* dan visualisasi keputusan.

Gambar 6. Visualisasi Pohon Keputusan *Decision Tree*

Dalam visualisasi pohon keputusan terdapat 2 kesimpulan yaitu:

- Atribut atau faktor yang paling berpengaruh adalah IPS semester 1 dan 2, SKS lulus semester 1, kelas, IPK semester 1 dan 2, jenis kelamin, program studi.
- Atribut atau faktor yang tidak berpengaruh adalah asal SMA, keterangan aktif semester 1 dan 2, SKS lulus semester 2.

E. Pengujian

Pada tahap ini pengujian menggunakan metode *confusion matrix* dan parameter uji yang terdiri dari akurasi, presisi, *recall*. Hasil pengujian disajikan dalam bentuk tabulasi, menggunakan perbandingan *confusion matrix*. Pengujian *confusion matrix* untuk dataset yang diolah menggunakan algoritma *decision tree* yang dapat dilihat pada tabel 4.

Tabel 4. Pengujian *confusion matrix*

	***True* Tepat Waktu**	***True* Telat**	***Class Precission***
Prediction Tepat Waktu	631	33	95.03%
Prediction Telat	9	21	70.00%
Class *Recall*	98.59%	38.89%	

Pada *prediction* tepat waktu menghasilkan nilai *true* tepat waktu senilai 631, *true* telat senilai 33 akan menghasilkan nilai *class precission* 95.03% sedangkan untuk *prediction* telat menghasilkan *true* tepat waktu senilai 9, *true* telat senilai 21 akan di dapatkan hasil *class precission* 70.00%. Hasil dari *class recall* true tepat waktu 98.59%, *true* telat 38.89%. Perhitungan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

akurasi, presisi, dan *recall* yang dilakukan juga secara manual untuk membandingkan hasil perhitungan RapidMiner seperti pada persamaan (3), (4), dan (5) :

$$Akurasi = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} = \frac{631 + 21}{631 + 21 + 33 + 9} = 93.94\% \quad \textbf{(3)}$$

$$Presisi = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{631}{631 + 33} = 95.03\% \quad \textbf{(4)}$$

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{631}{631 + 9} = 9 \quad \textbf{(5)}$$

Pengolahan data yang dilakukan pada tahap *preprocessing* adalah memilih atribut yang tidak memiliki *record*, atribut yang tidak diperlukan sebagai *early warning*, dan atribut yang terdapat *missing value*. Dataset yang awal mulanya berjumlah 1.015 setelah dilakukannya *preprocessing* menjadi berjumlah 694. terkumpul selanjutnya ditentukan data *training* dan data *testing* menggunakan *split* data dan *10-fold cross validation* dengan *tools* RapidMiner. Selanjutnya dilakukan pengujian untuk mengetahui hasil uji akurasi, presisi, dan *recall* menggunakan *confussion matrix*. Setelah mendapatkan hasil uji akurasi, maka dapat diambil kesimpulan bahwa algoritma C4.5 mermpunyai hasil uji akurasi paling tinggi dibandingkan dengan algortima KNN dan *Naïve Bayes.*

V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan untuk memprediksi kelulusan mahasiswa pada Fakultas teknologi Informasi ISB Atma Luhur maka dapat ditarik kesimpulan bahwa algoritma C4.5 terbukti lebih akurat dalam memprediksi kelulusan mahasiswa karena mempunyai nilai akurasi paling tinggi dibandingkan dengan algoritma KNN dan *Naïve Bayes* senilai 93.94%. Algoritma C4.5 merupakan algoritma yang paling akurat mengingat kelebihan algortima ini paling efisien dalam menangani banyak tipe atribut diskret dab numerik dalam dataset. Untuk penelitian selanjutnya sebaiknya menggunakan atribut yang lebih banyak agar menghasilkan data yang lebih akurat.

Penelitian selanjutnya dapat dilakukan dengan menambahkan metode *feature selection* untuk mendapatkan hasil yang lebih baik dalam klasifikasi.

REFERENCES

[1] A. K. Wahyudi, N. Azizah, and H. Saputro, "Data Mining Klasifikasi Kepribadian Siswa SMP Negeri 5 Jepara Menggunakan Metode Decision Tree Algoritma C4.5," *JISTER (Journal of Information System and Computer Science),* vol. 2, no. 2, 2022.

[2] R. H. Pambudi and B. D. Setiawan, "Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal," vol. 2, no. 7, 2018.

[3] E. Purwanto, "Prediksi Kelulusan Tepat Waktu Menggunakan Metode C4.5 Dan K-NN (Studi Kasus : Mahasiswa Program Studi S1 Ilmu Farmasi, Fakultas Farmasi, Universitas Muhammadiyah Purwokerto)," vol. 20, no. 2, pp. 131–142, 2019.

[4] A. F. A. Rahman, S. Sorikhi, and S. Wartulas, "Prediksi Kelulusan Mahasiswa Menggunakan Aalgoritma C4.5 (Studi Kasus Di Universitas Peradaban)," vol. 1, no. 2, 2020.

[5] A. S. Sungae, "Optimasi Algoritma C4.5 Dalam Prediksi Web Phishing Menggunakan Seleksi Fitur Genetic Algoritma," *Paradigma,* vol. XX, no. 2, pp. 27–32, 2018.

[6] E. Elisa, " Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Konstruksi PT.Arupadhatu Adisesanti," *Jurnal Online Informasi*, vol. 2, no. 1, 2017.

[7] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. 2.i2 (213-219), no. 2, 2016.

[8] S. Z. Harahap, A. Nastuti "Teknik Data Mining Untuk Penentuan Paket Hemat Sembako Dan Kebutuhan Harian Dengan Menggunakan Algoritma FP-GROWTH (Studi Kasus di Ulfamart Lubuk Alung)," *Jurnal Ilmiah Fakultas Sains dan Teknologi*, vol. 7, no. 3, 2019.

[9] F. Elfaladonna and A. Rahmadani, "Analisa Metode Classification-Decission Tree Dan Algoritma C.45 Untuk Memprediksi Penyakit Diabetes Dengan Menggunakan Aplikasi Rapid Miner," *SINTECH (Science and Information Technology) Journal*, vol. 2, no. 1, 2019.

[10] A. Y. Saputra and Y. Primadasa, "Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbour," vol. 17, no. 4, 2018.

[11] S. Bahri and A. Lubis, "Metode Klasifikasi Decision Tree Untuk Memprediksi Juara English Premier League," vol. 2, no. 1, 2020.

[12] P. B. N. Setio, D. R. S. Saputro, and B. Winarno, "Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5," *PRISMA (Prosiding Seminar Nasional Matematika)*, vol. 3, pp. 64–71, 2020.

[13] R. N. Dessy et al, "Kajian Machine Learning Dengan Komparasi Klasifikasi Prediksi Dataset Tenaga Kerja Non-Aktif," vol. 7, no. 1, 2019.

[14] F. Satria, Zamhariri, M. Apun Syaripudin, "Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung," *Jurnal Ilmiah MATRIK* , vol. 22, no. 1, 2020.

[15] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 1, pp. 8–17, 2021.

[16] M. O. Odja, F. J. Likadja, W. T. Ina, and S. I. Pella, "Penggunaan Microsoft Excel untuk Kemudahan Pengolahan Data Nilai Hasil Belajar Siswa," Vol. XV, no. 2, 2021.

[17] A. R. Putri, "Optimalisasi Penggunaan Microsoft Excel Untuk Pengolahan Nilai Raport Di SMAN 1 Ngunut Tulungagung," vol. 3, no. 1, 2015.

[18] M. A. Hasanah, S. Soim, and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," vol.5, no. 22, 2021.

[19] R. Muliono, J. H. Lubis, and N. Khairina, "Analysis K-Nearest Neighbor Algorithm for Improving Prediction Student Graduation Time," *SinkrOn*, vol. 4, no. 2, p. 42, 2020.

[20] A. Rohman and A. Rufiyanto, "Implementasi Data Mining Dengan Algoritma Decision Tree C4.5 Untuk Prediksi Kelulusan Mahasiswa Di Universitas Pandanaran", vol. 3, 2019.

[21] N. Khasanah, A. Salim, and et al, "Prediksi Kelulusan Mahasiswa Dengan Metode Naive Bayes," vol. 13, 2022.

[22] D. Marlina and M. Bakrie, "Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, no. 1, 2021.

[23] W. Musu and A. Ibrahim, H. Heriadi "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5," vol. 10, no.1, 2021.

[24] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data Untuk Klasifikasi Wne Menggunakan Algoritma K-NN," vol. 4, no. 1, 2019.

[25] R. S. Putri and I. Waspada, "Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika," *Khazanah Informatika*, vol. 4, no. 1, 2018.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1731, Copyright ©2023
Submitted : May 18, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 4, 2023

Priority Recommendations for Residential Road Improvement Using the SMART Analysis Method

Lilik Sumaryanti[1]*, Syaiful Nugraha[2], Lusia Lamalewa[3]
Department of Informatic Engineering
Universitas Musamus
Merauke, Indonesia
lilik@unmus.ac.id[1], syaiful_ft@unmus.ac.id [2], lusia@unmus.ac.id [3]

***Abstract*— Roads are infrastructure for organizing transportation, which are places for traffic to flow both for people and goods to reach destinations safely, securely, comfortably, quickly, smoothly, orderly, and efficiently, especially roads in residential areas. Setting priorities for the road improvement program is the responsibility of the Public Housing, Settlement Areas, and Land Affairs Office, which handles technical planning, development, arrangement, supervision, and control of development in residential areas. Recommendations for road proposals for the currently running improvement program, based on an assessment of their physical condition, are carried out by experts. This prioritization certainly takes a long time because experts have to compare the physical conditions of the roads one by one to make a decision. A decision support system is specifically designed for the decision-making process that can be applied in various aspects of the decision-making field. Recommendations for alternative roads in the road improvement program were analyzed using the SMART method to find alternatives with the highest preference value and the advantage that they can be used for all weighting techniques. Accuracy testing shows that the priority recommendation output presented by the application has an accuracy rate of 80%. This value is obtained by comparing the results of recommendations from experts.**
Keywords*— *Priorities, SMART, Decision Support Systems, Road.

***Abstrak*— Jalan menjadi prasarana untuk penyelenggaraan transportasi, yang menjadi tempat arus lalu lintas baik angkutan orang maupun barang untuk mencapai tujuan dengan selamat, aman, nyaman, cepat, lancar, tertib dan teratur secara efisien, terutama jalan pada lingkungan permukiman penduduk. Penentuan prioritas pada program peningkatan jalan menjadi tanggung jawab Dinas Perumahan Rakyat, Kawasan Permukiman dan Pertanahan, yang bertugas menangani perencanaan teknis, pembangunan, penataan, pengawasan, dan pengendalian pembangunan di kawasan permukiman. Rekomendasi usulan jalan untuk program perbaikan yang sedang berjalan saat ini, dengan menyeleksi jalan berdasarkan penilaian kondisi fisiknya, dilakukan oleh ahli, penentuan prioritas ini tentunya membutuhkan waktu yang lama, karena ahli harus membandingkan satu persatu kondisi fisik jalan, guna menentukan keputusannya. Sistem pendukung keputusan merupakan suatu sistem yang dirancang khususnya untuk proses pengambilan keputusan yang dapat diterapkan diberbagai aspek bidang pengambilan keputusan. Rekomendasi alternatif jalan pada program peningkatan jalan dianalisis menggunakan metode SMART, untuk mencari alternatif yang memiliki nilai prefensi tertinggi dan memiliki kelebihan yaitu dapat digunakan untuk semua jenis teknik pemberian bobot. Pengujian akurasi menunjukkan bahwa output rekomendasi prioritas yang disajikan aplikasi memliki tingkat akurasi 80%, nilai ini diperoleh dengan membandingkan hasil rekomendasi yang bersumber dari ahli.**
Kata Kunci*— *Prioritas, SMART, Sistem Pendukung Keputusan, Jalan.

I. INTRODUCTION

The public facility is the most widely used transportation infrastructure by the community for carrying out their daily mobility, so the volume of vehicles passing through a road will affect its capacity and carrying capacity. The strength and durability of road pavement construction are largely determined by the load that the road receives, the characteristics of the bearing capacity of the subgrade, and the quality of the pavement's raw materials [1]. The good transportation operations, roads must serve the flow of traffic for people and goods to reach destinations safely, securely, comfortably, quickly, smoothly, orderly, and efficiently, especially roads in residential areas. The function of the road network is differentiated based on the nature and movement of traffic and road transport, which is divided into arterial, collector, local and environmental.

Setting priorities for the road improvement program is the responsibility of the Public Housing, Settlement Areas, and Land Affairs Office, which handles technical planning, construction, arrangement, supervision, and control of development in residential areas. Technical planning for development in residential areas for priority determination of roads to be repaired/constructed, through direct observation in the field by experts, by assessing the feasibility of the road development plan by considering several factors/criteria that determine the selection of the location of the recommended road to be the priority for proposed improvements road. The criteria for determining priority for road repairs in residential areas include; road damage level, population, public facilities, road length, and road age [2][3]. Recommendations for road proposals for the currently running improvement program, by selecting roads based on an assessment of their physical condition, are carried out by experts. This prioritization certainly takes a long time because experts have to compare the

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

physical conditions of the roads individually to make a decision.

Decision Support System (DSS) application technology can be used to assist in recommending priority alternatives for determining roads to be repaired. A decision support system is specifically designed for the decision-making process that can be applied in various aspects of the decision-making field [4]. DSS is a computer-based interactive system that presents and processes information [5], which enables decision-making to be more productive, dynamic, and innovative [6]. The decision support system can be used to determine the right priority for road improvement which is analyzed using the VIKOR method [7], DSS for determining road repair priorities is a process for generating alternative decisions that have been obtained from processing decision-making models using the AHP-SAW-TOPSIS method [2]. Alternative recommendations use the Simple Multi-Attribute Rating Technique (SMART) method by looking for alternatives that have the highest preference value [8], so that the recommendations obtained are accurate based on the specified criterion values [9]. The advantage of this method is that it is simple and can be used for all types of weighting techniques [10], so it can help make decisions in application areas in environmental, construction, logistics transportation, military, manufacturing, and assembly issues [11]. One of the roles of a decision support system is that it can assist decision-makers in obtaining alternatives that are by the objectives of multi-criteria decision-making [12], where each alternative has criteria and has value and weight to facilitate decision-making with special cases [13]. Determination of building material suppliers uses the SMART method as an analytical model that presents information on the feasibility of the selected supplier as an alternative solution [14]. DSS is a decision-making tool for imposing a lockdown during the Covid-19 pandemic [15]. Its application is also used to determine a government program and feasibility analysis using the AHP method [16].

II. LITERATURE REVIEW

A. Road Function

According to their designation based on Law No. 38 of 2004 concerning Roads, roads consist of public and special roads. Public roads are further grouped according to system, function, status, and road class [17]. Environmental road network management programs include road maintenance and improvement programs and new road construction programs [18]. Road handling consists of routine maintenance, periodic maintenance, rehabilitation, and reconstruction.

1. Routine maintenance includes road shoulder maintenance; filling surface gaps/cracks; asphalt paving; and hole patching.
2. Periodic maintenance activities include road shoulder repairs; filling surface gaps/cracks; hole patching; wave repair; non-structural resurfacing; and thin asphalt coating.
3. Rehabilitation activities include road shoulder repair, asphalt paving, and structural resurfacing.

B. Decision Support System

DSS allows decision makers to produce decisions in a faster time-saving analysis time [15], due to system support that can process large amounts of data [19] quickly and can produce decisions that are by the objectives [20].

C. Stages Analysis Use the SMART method

The Simple Multi-Attribute Rating Technique (SMART) method was developed for decision-making with many criteria and theoretical basis alternatives and had a weight [21]. The basis of the theory can be seen in how important the value of the weight is compared to other criteria [22]. The stages of analysis using this method are described as follows [11][23]:

1. Determine the criteria, determine the criteria used to solve problems or cases, with discussions with experts, to obtain information on the criteria used in the decision-making system.
2. Determine the weight of the criteria by giving weight to each criterion using a certain rating scale, taking into account the importance of each criterion.
3. Normalization weight of the criteria normalization is done by changing the value of the numeric column in the data set to use a common scale without distorting the differences in the range of values or losing information. Normalization is also required for some algorithms to model the data properly. The following equation is used to normalize weights [14]:

$$w_i' = \frac{W_i}{\sum_{j=1}^{n} w_j} \quad (1)$$

Description :
W_i': weight of normalized criteria for criterion i
W_i: weight criteria i-th
W_j: weight criteria jth _
j : 1,2,3, ... , n amount criteria

4. Provide parameter values for each criterion, assigning criteria values for each alternative using quantitative data (numbers) or qualitative data. If the criterion value is used in qualitative form, we need to convert it into quantitative data by making value parameters use a certain scale.
5. Determining the value of utility is done by assigning an importance value based on the usefulness of the criteria, which depends on the nature of the value.
   (a) Cost criteria have a value in a currency, such as costs to be incurred (e.g., price criteria). This type of criterion is calculated using the following equation [24]:

$$u_i(a_i) = \frac{C_{max} - C_{out}}{C_{max} - C_{min}} \quad (2)$$

Description :
$u_i(a_i)$: utility criterion value i for the alternative -i
c_{max}: maximum criterion value
c_{min}: minimum criterion value

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

c_{out}: output criterion value

(b) Criteria for benefits (benefit criteria) are criteria that can be an advantage (e.g., criteria for size or quality and others) and are calculated by equation (2).

6. Calculating the final value, analyzing alternative values based on each criterion using the following equation [21]:

$$u_i(a_i) = \sum_{j=1}^{n} w_j * u_j(a_i) \quad (3)$$

Description :
$u_i(a_i)$: total value for alternative i
w_j: mark weight criteria to -j already normalized
$u_j(a_i)$: utility criterion value to -j for alternative i-

7. The alternative ranking is the sorting of alternative solutions, which are the analysis results from the largest to the smallest value. The alternative with the highest final value shows the best alternative.

III. RESEARCH METHODS

The development of a decision support system aims to recommend priority alternative roads to be repaired in residential areas, and the application presents alternative roads with priority values indicated by ranking. The results of the recommendations presented can be used by the local government, more specifically for the Public Housing, Settlement Areas, and Land Affairs Office of Merauke Regency. The stages of research activities for developing the application are shown in the research flowchart.

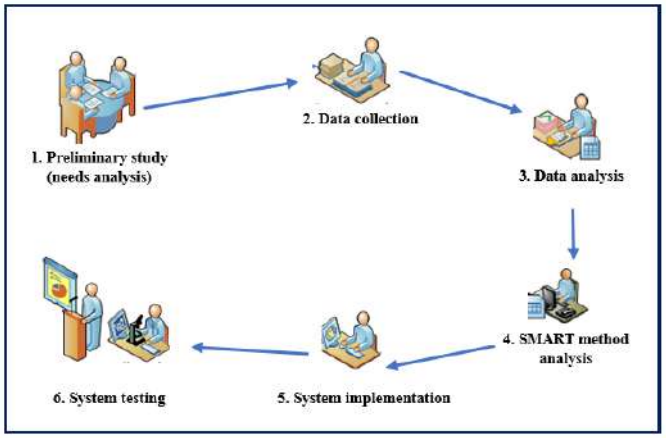


Fig. 1. Stages of research activities

The stages of analysis using the SMART method in a decision support system to recommend alternative roads as priority road improvement programs are shown in Figure 1, with the stages of activities described as follows:

1. Needs Analysis is an activity carried out to find information about system requirements (input/output) and analysis of user needs so that the resulting information can meet the needs of system users. The results of this activity are the data needed by system management and information needs for users.
2. Data collection is an activity to obtain information needed to achieve research objectives. The data related to the research are; road data, road improvement criteria data, and road physical condition assessment data through field survey activities. The data used is a sample of ten damaged road data in the Merauke District which is the result of an environmental road survey by the Public Housing, Settlement Areas and Land Affairs Office.
3. Data analysis is the stage for modeling the problem by determining the priority of the road improvement program, such as determining the weight for each criterion and calculating the criteria to determine the recommended alternative solutions. This activity was done through joint discussions with government authorities, namely the Public Housing, Settlement Areas, and Land Affairs Office employees.
4. SMART method analysis is an analysis of the results of assessing the physical condition of the road with the stages of the SMART algorithm to find information on alternative solutions as a result of recommendations presented to decision-makers through the application. At this stage, it will rank all alternative street name solutions recommended as priority road improvement programs.
5. System implementation is the stage of implementing analysis into a decision support system to determine whether the results/output presented by the system are by the information needs of decision-makers.
6. System testing is an activity to measure the truth of alternative solutions the system recommends, compared to alternative recommendations from experts. This activity will produce an accuracy value for the accuracy of the recommendation output.

IV. RESULT AND DISCUSSION

A. *Modeling of Road Improvement Priority Recommendation System*

The flow of data and information from users of decision support systems to recommend roads, which are priority road improvement programs, is shown in the flowchart in Figure 2. User involvement is explained as follows:

1. Admin has access to add system user data, add and update criteria/sub-criteria data, add and update road data, add and update road condition assessments sourced from field survey results, obtain information on the results of the SMART method analysis in the form of alternative rankings recommended roads for the road improvement program.
2. Tim Survei has access to add road condition assessment data from the survey results that have been conducted and can obtain information related to the list of road data.
3. Pimpinan, have access to obtain information on the list of road data, the results of alternative j

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

recommendations, which are a priority for the road improvement program, are proven accurately based on the results of an analysis with rankings.

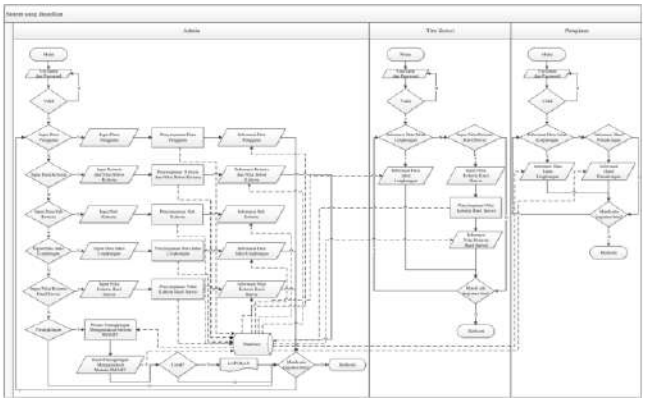


Fig. 2. Flowchart System for recommending priority for road improvement

The context diagram of the decision support system for determining the priority of environmental roads can be seen in Figure 3, which shows three external entities related to the system, namely admin, survey team, and leader.

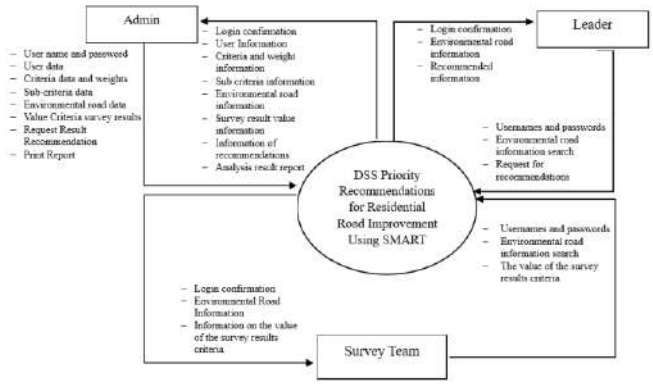


Fig. 3. Context Diagram of System for recommending priority for road improvement

The Data Flow Diagram represents the system built graphically and shows the various processes, data flows, and information shown in Figure 4. The processing of data into information is explained as follows:

1. Process 1.0 is a login process where the admin, survey team, and leadership must enter a username and password to gain access rights to the system.
2. Process 2.0 is a data management process where the admin can view, add, edit, or delete master data.
3. Process 3.0 is a ranking process where the admin can carry out an analysis of environmental road priorities. The system uses criteria weights and sub-criteria, along with survey results criteria values using the SMART method; the results of the analysis are in the form of a ranking of recommended alternative solutions. Admin and leaders can access ranking result data.
4. Process 4.0 is the process of making a report through a printed facility. The results of alternative recommendations for the road improvement program, the admin can be accessed.

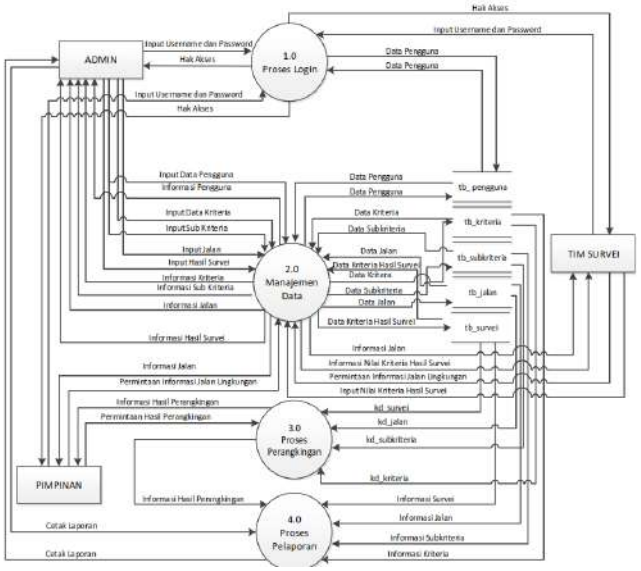


Fig. 4. Data flow diagram of the priority recommendation system for road improvement

B. Analysis Stages Using the SMART Method

Determining priority recommendations for road improvement programs using the SMART method begins with determining the criteria used to solve problems or cases, with discussions with experts at the Public Housing, Settlement Areas, and Land Affairs Office to obtain information on the criteria used in the decision-making system. SMART analysis stages are described as follows:

1. Determination of the criteria used consists of the physical condition of road damage, population, number of public facilities such as places of worship, health facilities, educational facilities, road length, and age. Each criterion has different unit values and sub-criteria, shown in Table 1.

TABLE 1. PRIORITY CRITERIA FOR ROAD IMPROVEMENT

No	*Criteria*	*Unit*	*Sub Criteria / Value*
1	Condition physique damaged road	condition state	• Damaged lightweight (1) • Damaged medium (2) • Damaged weight (3)
2	Amount Resident	Person	• <40 people (1) • 40 to 60 people (2) • 70 to 100 people (3) • > 100 people (4)
3	Amount facility general	units	• No There are facilities (1) • One Facility (2) • Two facilities (3) • >2 facilities (4)
4	Long road	Meters	• <100m (1) • 100 to 300 m (2) • 301 to 600 m (3) • 601 to 1000m (4) • > 1000m (5)

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

No	Criteria	Unit	Sub Criteria / Value
5	Age road	Year	• < 1 year (1) • 1 to 3 years (2) • 4 to 5 years (3) • >5 years (4)

2. The weighting of the criteria, taking into account the value of the importance of the criteria, with the results of the weighting is shown in Table 2.

TABLE II. THE WEIGHTING OF ROAD IMPROVEMENT PRIORITY CRITERIA

Code	Criteria	Weight	Type
Kr_1	The physical condition of road damage	30%	Benefit
Kr_2	Number of inhabitants	25%	Benefit
Kr_3	Public facilities	20%	Benefit
Kr_4	Road length	15%	Cost
Kr_5	Road age	10%	Benefit

3. Normalizing the criteria weight values by applying equation 1, which is explained as follows:

The physical condition of road damage = $\frac{30}{100} = 0{,}3$

Number of inhabitants = $\frac{25}{100} = 0{,}25$

Public facilities = $\frac{25}{100} = 0{,}25$

Road length = $\frac{15}{100} = 0{,}15$

Road age = $\frac{10}{100} = 0{,}1$

4. Assessment of criteria for alternatives using quantitative data. The examples of cases used for the analysis of the SMART method are shown in Table 3 below:

TABLE III. ROAD IMPROVEMENT RECOMMENDATION CASE

Case	Street Name	Code				
		Kr_1	Kr_2	Kr_3	Kr_4	Kr_5
X_1	Jln Pisang	3	3	2	2	3
X_2	Jln Gambit	2	2	1	2	2
X_3	Jln Kanguru	1	1	1	2	2
X_4	Jln Arafura	3	3	2	2	4
X_5	Jln Cikombong	2	1	1	4	2

5. Calculate the utility value of each criterion using equation 2, which is explained as follows:
(a) Road damage utility value (Kr_1)

$$X_{11} = \frac{3-1}{3-1} = 1 \qquad X_{12} = \frac{2-1}{3-1} = 0{,}5$$

$$X_{13} = \frac{1-1}{3-1} = 0 \qquad X_{14} = \frac{3-1}{3-1} = 1$$

(b) The utility value of the number of inhabitants (Kr_2)

$$X_{21} = \frac{3-1}{4-1} = 0{,}67 \qquad X_{22} = \frac{2-1}{4-1} = 0{,}33$$

$$X_{23} = \frac{1-1}{4-1} = 0 \qquad X_{24} = \frac{3-1}{4-1} = 0{,}67$$

This stage is done for all alternatives based on each criterion.

6. Analyze the alternative final values based on each criterion by multiplying the weight of the criteria using Equation 3.
(a) Calculate the final value of road damage (Kr_1)

$X_{11} = 1 * 0{,}3 = 0{,}3$ $\qquad$ $X_{12} = 0{,}5 * 0{,}3 = 0{,}15$

$X_{13} = 0 * 0{,}3 = 0$ $\qquad$ $X_{14} = 1 * 0{,}3 = 0{,}3$

(b) Calculate the final value of the number of inhabitants (Kr_2)

$X_{21} = 0{,}67 * 0{,}25 = 0{,}17$ $\qquad$ $X_{22} = 0{,}33 * 0{,}25 = 0{,}08$

$X_{23} = 0 * 0{,}25 = 0$ $\qquad$ $X_{24} = 0{,}67 * 0{,}25 = 0{,}17$

Analysis of the final results by adding up the value of each criterion is exemplified as follows:
Calculating the analysis results of Jln Pisang (X_1)
$X_1 = 0.3 + 0.17 + 0.07 + 0.11 + 0.07 = 0.71$

7. Ranking of alternative solutions is sorting the results of the calculation of stage 6 from the highest value to the minimum.

TABLE IV. ALTERNATIVE RATING RESULTS

Ranking	Street Name	Preference results
1	Jln Arafura	0,75
2	Jln Pisang	0.71
3	Jln Gambit	0,38
4	Jln Cikombong	0,22
5	Jln Kanguru	0,15

C. *Results of SMART Implementation for Analysis of Alternative Solutions*

Determination of sub-criteria is carried out to scale the criteria unit to a certain value. In contrast, the results of the preparation of the sub-criteria used are shown in Figure 5.

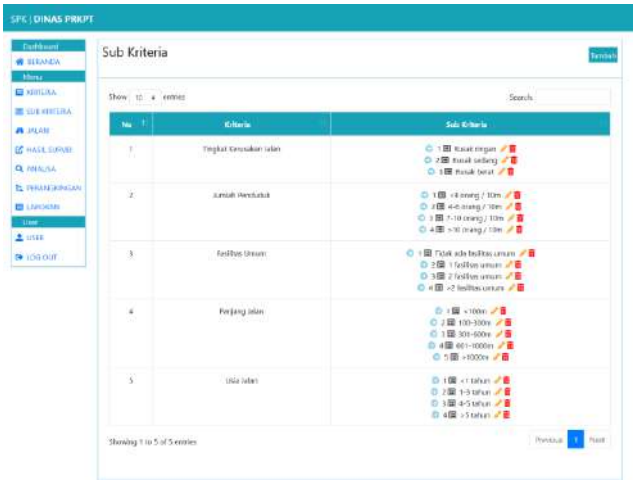


Fig. 5. Sub-criteria for road improvement recommendation system

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

The results of normalizing the criterion weights use equation 1, shown in Figure 6. The weight values use a scale of 0 to 100.

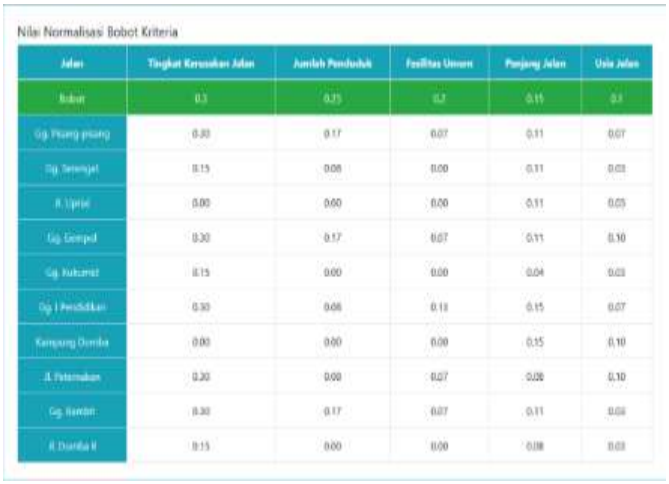


Fig. 6. Results of the normalization of the criterion weights

Assessment of road conditions based on survey results is used to add data to each alternative's criteria value, namely the name of the road. The implementation of adding alternative data is shown in Figure 7.

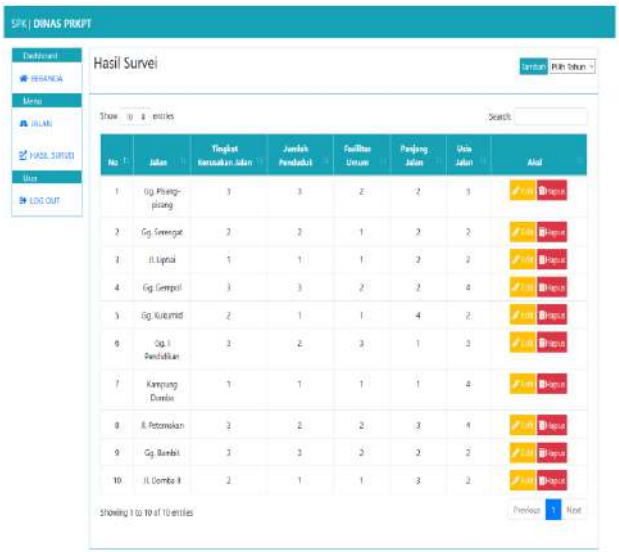


Fig. 7. Facility to add alternatives

The SMART method analysis begins by providing the criterion values for each alternative/case, as shown in Figure 8.

Fig. 8. Facilities for assessing criteria for each alternative

The system facility for presenting information on the results of calculating the utility value of each criterion uses equation 2, shown in Figure 9.

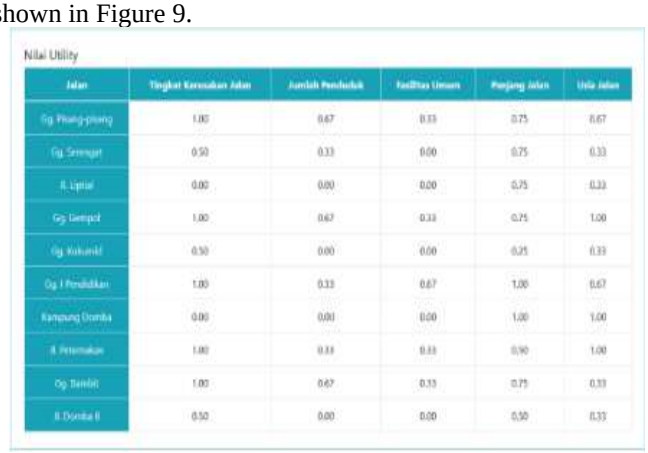


Fig. 9. Calculation results of the utility value

The final stage of the analysis uses the SMART method, namely ranking the alternatives from the highest to the smallest preference value. The ranking results show the priority order of alternatives for road improvement program recommendations for decision-makers. The ranking page on the recommendation system can be seen in Figure 10.

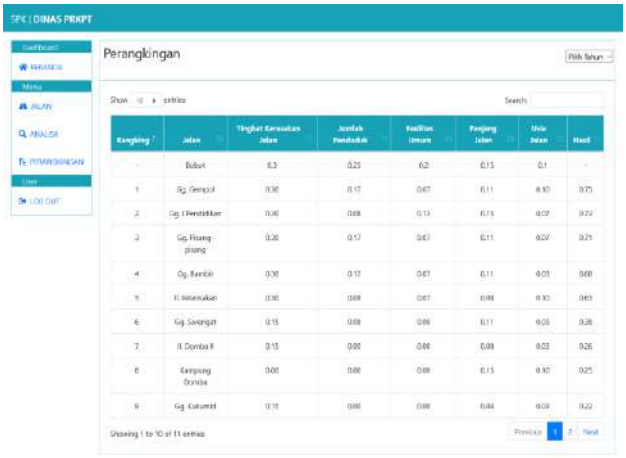


Fig. 10. SMART method analysis results

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

D. Testing the Accuracy of the Results of System Recommendations

The accuracy level analysis technique is carried out by comparing the results of recommendations from the system and the results of the selection independently by experts. The cases used to determine the accuracy value are shown in the following table:

TABLE V. COMPARISON OF EXPERT RECOMMENDATION RESULTS AND SYSTEM

No	*Street Name*	*Criteria Value*					*S*	*A*
		Kr1	*Kr2*	*Kr3*	*Kr4*	*Kr5*		
1	Gg. Pisang	Damaged weight	135	1	150	4	3	3
2	Gg. Serengat	Damaged medium	123	0	238	2	6	6
3	Jl Liptiai	Damaged lightweight	25	0	118,5	3	10	10
4	Gg. Gempol	Damaged weight	120	1	150	6	1	1
5	Gg. Kukumid	Damaged medium	275	0	1000	3	9	9
6	Gg. Pendidikan	Damaged weight	18	2	30	5	2	2
7	Kampung Domba	Damaged lightweight	23	0	80	6	8	7
8	Jl. Peternakan	Damaged weight	201	1	345	3	5	5
9	Gg. Bambit	Damaged weight	140	1	200	3	4	4
10	Jl Domba 2	Damaged medium	137	0	245	3	7	8

Description:
S : System output ranking/priority
A : Ranking / priority of experts

The measurement of the accuracy value is based on the results of the ranking in Table 5, which shows that there are differences in the output recommendations sourced from the system and experts in cases number 7 and 10, so the accuracy value is obtained using the following equation [25]:

$$pr = \frac{test\ data\ is\ correct}{Total\ of\ sample\ data} * 100\% \quad (4)$$

Based on equation 4, the system accuracy value is 80%, which is explained as follows:

$$pr = \frac{8}{10} * 100\% = 80\%$$

V. CONCLUSION

The road improvement recommendation system can be used as a tool for decision-makers in considering options. Alternative recommendations use the SMART analysis method to find alternatives with the highest preference value so that the recommendations obtained are appropriate based on the importance value of the specified criteria. The advantage of this method is that it can be used for all types of weighting techniques to help decision-makers in application areas such as; construction, transportation, logistics, and others.

REFERENCES

[1] Musthofa, "Analisa Kerusakan dan Perbaikan Jalan Aspal," *D'Teksi J. Tek. Sipil*, vol. 4, no. 2, pp. 13–24, 2019.

[2] A. P. Regitha, N. Hidayat, and A. W. Widodo, "Rekomendasi Prioritas Perbaikan Jalan Dengan Metode AHP-SAW-TOPSIS (Studi Kasus: Dinas Pekerjaan Umum dan Penataan Ruang Kota Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 2960–2969, 2019, [Online]. Available: http://j-ptiik.ub.ac.id

[3] R. A. Pratama and W. Harianto, "IMPLEMENTASI METODE AHP-TOPSIS DALAM," *SMARTICS J.*, vol. 7, no. 2, pp. 78–93, 2021.

[4] R. P. Sari and A. M. Alliandaw, "Sistem Penentuan Penerima Bidikmisi UNTAN Dengan Menggunakan Metode MOORA," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 2, pp. 242–250, 2022, doi: 10.32736/sisfokom.v11i2.1420.

[5] D. H. Alahmadi and A. A. Jamjoom, "Decision support system for handling control decisions and decision-maker related to supply chain," *J. Big Data*, vol. 9, no. 1, 2022, doi: 10.1186/s40537-022-00653-9.

[6] H. Apriyansyah, "Literature Review of: Decision Support System: Organization, Human Resources and Knowledge Management," *Dinasti Int. J. Econ. Financ. Account.*, vol. 3, no. 2, pp. 136–148, 2022.

[7] S. Daulay, "Sistem Pendukung Keputusan Menentukan Prioritas Perbaikan Jalan Pada Dinas Pekerjaan Umum Kabupaten Padang Lawas Menggunakan Metode Vikor," *JISTech (Journal Islam. Sci. Technol.*, vol. 5, no. 2, pp. 1–17, 2020.

[8] Y. E. Windarto, I. P. Windasari, and M. A. M. Arrozi, "Implementasi Simple Multi Attribute Rating Technique Untuk Penentuan Tempat Pembuangan Akhir," *J. Pengemb. Rekayasa dan Teknol.*, vol. 15, no. 1, p. 12, 2019, doi: 10.26623/jprt.v15i1.1484.

[9] Y. Septiana, F. Nuraeni, and K. Anisa, "Decision Support System for The Program Indonesia Pintar Scholarship Using Simple Additive Weighting Method," *J. dan Penelit. Tek. Inform.*, vol. 7, no. 4, pp. 2311–2316, 2022, doi: 10.33395/sinkron.v7i4.11786.

[10] A. Jayady *et al.*, "Decision Support System with Multi Criteria Decision Making Technique," *J. Phys. Conf. Ser.*, vol. 1933, no. 1, pp. 0–8, 2021, doi: 10.1088/1742-6596/1933/1/012017.

[11] L. Sumaryanti, T. K. Rahayu, A. Prayitno, and Salju, "Comparison study of SMART and AHP method for paddy fertilizer recommendation in decision support system," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 343, no. 1, 2019, doi: 10.1088/1755-1315/343/1/012207.

[12] M. Saputra and L. Bachtiar, "Analisis Penerimaan Karyawan Pada Pt. Srikandi Diamond Indah Motors Sampit Dengan Metode Analitycal Hierarchy Process (Ahp) Dan Simple Additive Weighting (Saw)," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 10, no. 3, pp. 312–319, 2021, doi: 10.32736/sisfokom.v10i3.1239.

[13] Sukamto, Y. Andriyani, and D. Oktaviani, "Penerapan Metode VIKOR untuk Penilaian Kinerja Karyawan," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, pp. 187–194, 2022.

[14] Aulia Rizky Muhammad Hendrik Noor Asegaff, M Dedy Rosyadi, and Budi Ramadhani, "Implementation of the Smart Methods (Simple Multi-Attribute Rating Technique) for Location Selection of Industrial Work Practice and Monitoring in Vocational School Students," *J. Teknol. Inf. Univ. Lambung Mangkurat*, vol. 7, no. 2, pp. 141–150, 2022, doi: 10.20527/jtiulm.v7i2.140.

[15] T. S. Almulhim and I. Barahona, "Decision support system for ranking relevant indicators for reopening strategies following COVID-19 lockdowns," *J. Qual. Quant.*, vol. 56, no. 2, pp. 463–491, 2022, doi: 10.1007/s11135-021-01129-3.

[16] A. Nugraha, A. Widianto, M. Irfan, M. Nasar, and M. Lestandy, "Decision Support System for Community Housing Subsidy Recipients," *J. Tek. Ind.*, vol. 21, no. 1, pp. 104–114, 2020, doi: 10.22219/jtiumm.vol21.no1.104-114.

[17] Pemerintah Repulik Indonesia, *UNDANG-UNDANG REPUBLIK INDONESIA NOMOR 38 TAHUN 2004 TENTANG JALAN*. Indonesia, 2004, pp. 1–61. [Online]. Available: https://jdih.pu.go.id/internal/assets/assets/produk/UU/2014/10/UU38-2004.pdf

[18] Pemerintah Repulbik Indonesian, *PERATURAN PEMERINTAH REPUBLIK INDONESIA NOMOR 34 TAHUN 2006*. Indonesia: https://jdihn.go.id/files/4/2006pp034.pdf, 2006, pp. 1–38. [Online]. Available: https://jdihn.go.id/files/4/2006pp034.pdf

[19] M. Farkas-Kis, "Decision making in the shadow of mathematical

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

education," *J. Decis. Syst.*, vol. 31, no. S1, pp. 168–180, 2022, doi: 10.1080/12460125.2022.2087417.

[20] J. Peignot, A. Peneranda, and S. Amabile, "Strategic Decision Support Systems for Local Government: A Performance Management Issue? The Use of Information Systems on the Decision-making and Performance Management of Local Government," *Int. Bus. Res.*, vol. 6, no. 2, pp. 92–100, 2012, doi: 10.5539/ibr.v6n2p92.

[21] L. Sumaryanti and N. Nurcholis, "Analysis of Multiple Criteria Decision Making Method for Selection the Superior Cattle," *INTENSIF J. Ilm. Penelit. dan Penerapan Teknol. Sist. Inf.*, vol. 4, no. 1, pp. 131–141, 2020, doi: 10.29407/intensif.v4i1.13863.

[22] L. Sumaryanti, Nurcholis, and L. Lamalewa, "Aplication of Hybrid Method for Superior cattle selection using Decision Support System," *E3S Web Conf.*, vol. 328, 2021, doi: 10.1051/e3sconf/202132803003.

[23] Supardi and D. Mahdiana, "Sistem Penerimaan Asisten Laboratorium Komputer Dengan Menggunakan Metode Analytical Hierarcy Process (AHP) Dan Simple Multi Attribute Rating Technique (SMART)," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, pp. 90–95, 2023.

[24] M. Hisjam, N. Octyajati, W. Sutopo, and A. Ali, "A Decision Support System to Achieve Self-Sufficiency of Soybean (Case: Central Java Province, Indonesia)," *J. Optimasi Sist. Ind.*, vol. 19, no. 2, pp. 144–156, 2020, doi: 10.25077/josi.v19.n2.p144-156.2020.

[25] W. Wahyudi, J. Santony, and G. W. Nurcahyo, "Akurasi Keputusan dalam Penentuan Guru Berprestasi dengan Menggunakan Metode Simple Additive Weighting," *J. Sistim Inf. dan Teknol.*, vol. 2, no. 2017, pp. 9–14, 2020, doi: 10.37034/jsisfotek.v2i1.15.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1738, Copyright ©2023
Submitted : May 21, 2023, Revised : June 14, 2023, Accepted : July 10, 2023, Published : November 4, 2023

Towards Human-Level Safe Reinforcement Learning in Atari Library

Afriyadi Afriyadi [1] *, Wiranto Herry Utomo [2]
Faculty of Computing, President University [1], [2]
Cikarang, Indonesia
afriyadi.it@gmail.com[1], wiranto.herry@president.ac.id[2]

***Abstract*— Reinforcement learning (RL) is a powerful tool for training agents to perform complex tasks. However, from time-to-time RL agents often learn to behave in unsafe or unintended ways. This is especially true during the exploration phase, when the agent is trying to learn about its environment. This research acquires safe exploration methods from the field of robotics and evaluates their effectiveness compared to other algorithms that are commonly used in complex videogame environments without safe exploration. We also propose a method for hand-crafting catastrophic states, which are states that are known to be unsafe for the agent to visit. Our results show that our method and our hand-crafted safety constraints outperform state-of-the-art algorithms on relatively certain iterations. This means that our method is able to learn to behave safely while still achieving good performance. These results have implications for the future development of human-level safe learning with combination of model-based RL using complex videogame environments. By developing safe exploration methods, we can help to ensure that RL agents can be used in a variety of real-world applications, such as self-driving cars and robotics.**

Keywords*— *reinforcement learning, videogame environment, safety constraint, safe reinforcement learning

I. INTRODUCTION

Reinforcement Learning (RL) is a subfield of machine learning that is widely used in solving decision-making problems [42]. In RL, an agent learns to interact with an environment to achieve a specific goal by taking actions that maximize a reward signal. This technique has been successfully applied in various fields, such as robotics, gaming, and finance [12, 29, 30]. However, unlike humans, ensuring safety while solving a problem is still a big challenge in RL, it is to make the agent learn without causing harm to itself or the environment [27].

Traditional RL algorithms optimize for the maximum reward, which can lead to unsafe behaviors, especially in environments with potentially harmful states. To address this issue, researchers have proposed several approaches [8, 9, 14, 15, 21, 41, 45, 52, 53], all of the research are in the scope of Safe Reinforcement Learning (*Safe-RL*), including adding safety constraints or modifying the reward function to incorporate safety. These approaches aim to maintain a balance between achieving the maximum reward and ensuring safety.

Despite the advancements in RL algorithms, there is still a gap in developing *Safe-RL* agents that can balance between achieving the maximum reward and ensuring safety. While there have been studies on safe exploration during training and implementation in RL [49], it is mainly focused on simple environments and tasks, such as the *cart-pole* or *mountain-car* tasks. There is a need to evaluate the effectiveness of Safe-RL in more complex environments. In addition, there is a need to investigate the results of various training iteration quantities on the reward and safety violations and see the minimum iteration to achieve optimal results for the algorithm.

We have several motivations in this research. First is the fact that in RL research, especially in robotics, researchers are required to perform initialization of the agent and environment state after finishes each iteration or each episode. This initialization is consuming a lot of work for the researcher [55]. Second, the proposed solution to reduce labor by creating simulated environment [55] if effective, but the process of creating simulated environment from scratch itself also required significant amount of labor [10, 54]. In other hand, the videogame today, have a diverse set of almost-realistic environments that can be the learning environment for RL agents. And third, the solution of using simulated environment, and creating simulated environments does not yet include safety concerns.

We propose to develop a framework to perform robotic tasks in videogame environments, instead of developing each environment for each specific experiment. To do this we propose using Atari library as the example videogame environment, specifically with *Super-Mario-Bros* game, which is an easy-to-understand game, the game have complex environments, and require the algorithm to perform near-human like intelligence. And within the game, there are also several states that can end the game, that can be declared as unsafe state. We are expecting this research as a steppingstone to utilizing videogame environments for research purposes and reduce researcher additional work for creating simulated environments for every research. And lastly, we want to explore the impact of using the *Safe-RL* algorithm which usually implemented for

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

robotics and autonomous driving in videogame environment.

The goal is to explore the potentials from incorporating a safety mechanism in RL algorithms, seeking towards human-like behavior to RL agents when they are introduced with unsafe states and safety concerns, and still maintain optimal rewards during training and deployment. And we want to simulate this safety-constrained behavior in a complex videogame environment, to show the potential of utilizing videogame environments for RL research.

This research proposes a novel approach for Safe-RL by combining modified reward concepts and hand-crafted safety constraints in state-of-the-art RL algorithms, specifically *Double Deep Q-Network (DDQN)* [19], to achieve high rewards while ensuring safe exploration in the presence of unsafe states. Furthermore, this research explores the effectiveness of various training iteration quantities. Monitoring closely the rewards, and safety violations through customized evaluation metrics.

The experiment involves simulating the safety-constrained algorithm in the *Atari Library* and *gym* (*OpenAI Gym* [10]) environment. The *Atari Library* is a collection of 57 classic *Atari 2600* games that are used as benchmark environments for RL research. *OpenAI Gym* [10] is a toolkit provided by the OpenAI company for developing and comparing reinforcement learning algorithms. It provides several environments that can be used to train RL agents, as well as several tools for evaluating the performance of agents. The usage of *gym* in this research is to investigate whether the agent will act more safely and towards human-like behavior after being introduced with safety constraint and unsafe states compared with the default *gym* settings.

This research also includes developing and evaluating safety mechanisms in agent's exploration method during training and implementation. The evaluation will be based on the performance of the agent in achieving maximum rewards while demonstrating safe exploration in the presence of unsafe states, as well as the number of safety violations.

The research will also include the development of safety violation measurements to be applied to the agents trained using DDQN algorithm. The training iteration quantity will be varied to determine the impact of the number of iterations on the performance of the agent. Additionally, the research will explore the use of modified rewards as a method of enforcing safety constraints during training.

II. PRELIMINARIES

A. Problem Statement

The field of Reinforcement Learning (RL) has gained significant attention in recent years due to its potential for enabling agents to learn from experience and improve decision-making [42]. However, ensuring safety in RL environments remains a crucial challenge. Developing safety mechanisms is essential to ensure the safety of both the agent and the environment. While researchers in the field of robotics and autonomous driving primarily perform Safe-RL training in simulated environments created specifically for their research, this approach adds extra work for researchers in developing *Safe-RL* agents.

In contrast, the gaming industry has been expanding rapidly and has developed many games that emulate real-world situations. Utilizing video game environments as an alternative to creating new simulation environments from scratch for each research case shows promise. This research aims to investigate the effectiveness of *Safe-RL* algorithms, commonly used in robotics and autonomous driving, in complex video game environments. By incorporating hand-crafted safety constraints and reward modification, the research aims to explore the performance and safety capability of RL agents in these complex video game environments. This initial investigation is expected to provide valuable insights into the potential of using video game environments as simulated environments for the development of human-level RL agents in real-world implementations such as robotics, autonomous driving, and other applicable domains.

The goal of this experiment is to assess whether the implementation of hand-crafted safety constraints to the algorithm will generate safer, human-like behavior while maintaining high rewards [42]. Specifically, the research will use state-of-the-art RL algorithms, such as *Double Deep Q-Network (DDQN)*, as a baseline without safety constraints. The effectiveness of these algorithms in achieving maximum rewards and the frequency of safety violations will be evaluated. The research will then develop and apply hand-crafted safety constraints to the algorithm, incorporating awareness of unsafe states to prevent agents from repeating unsafe actions or revisiting unsafe states during training and exploration. The performance of the agent with safety constraints will be compared to the performance of the baseline algorithms without safety constraints to determine the effectiveness of the safety constraints in reducing safety violations.

To measure the implementation of the hand-crafted safety constraints, a set of unique evaluation metrics will be developed. These metrics will assess the tradeoff between maximum rewards and safety violations for the RL agents in the video game environment. The performance of the safety-constrained algorithms will be compared to that of the non-safety-constrained algorithms, and the results will be presented in an easy-to-understand report.

In summary, this research aims to address the challenge of safety in RL by investigating the effectiveness of Safe-RL algorithms in complex video game environments. By incorporating hand-crafted safety constraints and reward modification, the research aims to explore the performance and safety capability of RL agents. The research objectives include evaluating the effectiveness of safety constraints, developing unique evaluation metrics,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

and comparing the performance of safety-constrained and non-safety-constrained algorithms.

B. Literature Review

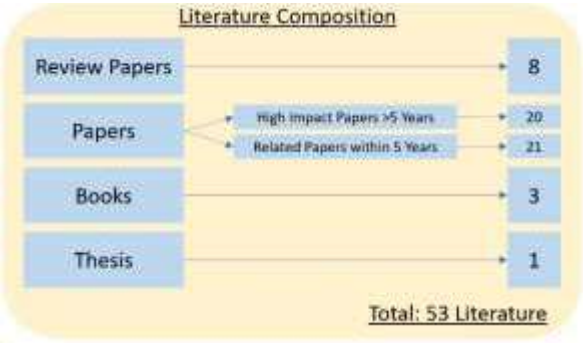


Fig. 1. We use eight recent review papers, 20 high impact papers within more than 5 years range, and relevant papers within 5 years range. We also use three books and one high impact Ph.D. thesis related to Safe-RL development. *(Literature Composition)*

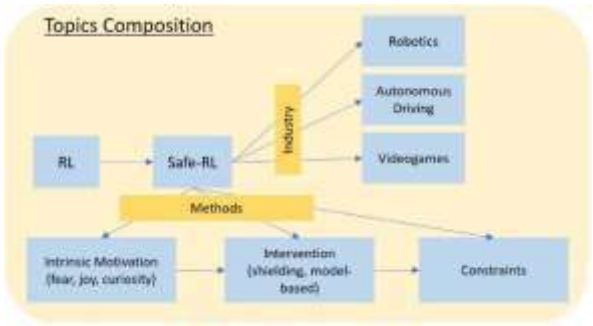


Fig. 2. Start with RL topics, then immediately pivot to Safe-RL. Categorize based on Industry, including Robotics, Autonomous Driving and Videogames. And categorize based on Safe-RL Methods, including Intrinsic Motivations, Intervention and Constraints. *(Topics Composition)*

We are starting by collecting the latest review papers within the field of RL and Safe-RL. Then continued with collecting information from high-impact papers mentioned in the review papers. And for deeper fundamentals, we also collect and analyze from major RL books and theses. The architecture of the literature is as follows:

We start investigating RL in general, also the challenge and opportunities that still exist in the field. Then we focused on safety concerns in RL, we investigated the safety concerns in the field of robotics, autonomous driving, and videogames. Then we investigate the common methods to implement safety within those fields, including intrinsic motivations, intervention with shielding, and putting safety constraints on the algorithm.

We also investigate the drawbacks of implementing RL in the field of robotics, autonomous driving and videogames, including the labor required to perform RL training in real-world, and how the researchers reduce this labor by performing training in simulated environments. And how researchers started to use videogame environments to perform RL experiments.

C. Reinforcement Learning (RL)

TABLE I. The difference between Supervised Learning, Unsupervised Learning and Reinforcement Learning.

Classes of Learning Problems		
Supervised Learning	**Unsupervised Learning**	**Reinforcement Learning**
Data: (**x**, **y**) **x** is data, **y** is label	**Data**: **x** **x** is data, no labels	**Data**: state-action pairs
Goal: Learn function to map **x** → **y**	**Goal**: Learn underlying structure	**Goal**: Maximize future rewards over many time steps
Strawberry Example:  This is strawberry	**Strawberry Example**:  These two items are similar, and should put into same category	**Strawberry Example**:  Eat this item to stay alive longer

Reinforcement learning (RL) is a popular machine learning approach that allows an agent to learn through interaction with an environment via trial-and-error [42]. The objective of RL is to learn a policy that maximizes the cumulative reward over time, based on state transitions and rewards observed through the agent's actions. The policy is a mapping from states to actions, and the goal is to learn the optimal policy that maximizes the expected cumulative reward. [42, 20]

Compared to supervised and unsupervised learning, RL differs in that RL learns from interaction with the environment rather than a labeled dataset. Supervised learning uses labeled examples from a human expert, while unsupervised learning learns from the underlying structure of data. [42].

RL has gained popularity due to its wide-ranging applications, including robotics, gaming, finance, healthcare, and transportation [46]. In robotics, RL has been used to train robots to perform complex tasks such as grasping and manipulation. In gaming, RL has been used to train agents to play games such as *Chess*, *Go*, and *Poker* at a superhuman level [40]. RL has also been applied in finance to optimize trading strategies and in healthcare to develop personalized treatment plans.

RL is composed of various components, such as the agent, the environment, the reward signal, the policy, and the value functions. The agent is responsible for taking actions in the environment based on its policy. The environment is the external system that the agent interacts with. The reward signal is a numerical feedback signal that the agent receives from the environment, indicating how well it is performing in the task. The policy is the strategy that the agent uses to select its actions in a given state. The

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

value function is a function that estimates the long-term value of a state or a state-action pair, which the agent uses to make decisions. These components are fundamental to RL and are necessary for the learning process to occur [42]

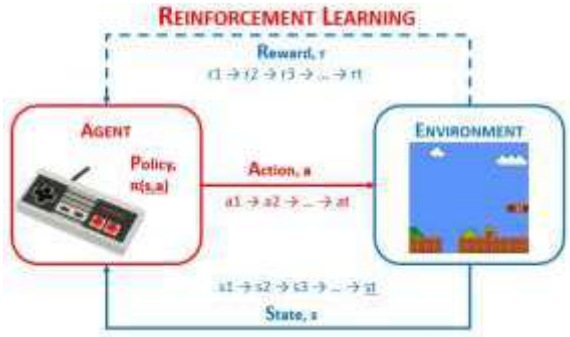


Fig. 3. RL agent in the state s, and then performing action *a* based on policy *π*, resulting in reward *r* and new state *s'*. *(Components of RL)*

D. Markov Decision Process

Markov Decision Processes (MDP) is a widely used framework in RL for modeling decision-making problems with uncertainty [42]. MDP provides a mathematical formulation [16] for solving sequential decision-making problems in stochastic environments.

The Markov Property assumes that the current state s has all the relevant information to predict the future and does not require any information from the past [42, 48]. A MDP is a decision process based on these assumptions. It is represented by a tuple (S, A, P, R, γ) where S denotes the set of states, A denotes the set of actions, P denotes the state transition probability function, R denotes the reward function, and γ denotes the discount factor.

The MDP model assumes that the agent interacts with the environment over a sequence of discrete time steps, t=1,2,3,...,T. At each time step, the agent observes the current state of the environment and selects an action to perform based on its policy. The environment then transitions to a new state according to the transition probability function, and the agent receives a reward based on the reward function [42, 48]. State transition probability is denoted mathematically as:

$$P_{ss'} = \mathbb{P}\left[s_{t+1} = s' \mid s_t = s\right] \quad (1)$$

s_t denotes the current state of the agent and s_{t+1} denotes the next state. What this equation means is that the transition from state S_t to S_{t+1} is entirely independent of the past.

When working with MDPs, the agent interacts with the environment by selecting actions that lead to a change of state and a corresponding reward. The state refers to the current state of the environment, while the action represents the decision made by the agent. The reward function is responsible for determining the reward the agent receives for taking a specific action in a particular state. Meanwhile, the transition probability function determines the probability of moving to a new state based on the current state and the action taken by the agent [42].

Returns (G_t) is the total reward expected from the environment, and returns is denoted mathematically as:

$$G_t = r_{t+1} + r_{t+2} + \cdots + r_T \quad (2)$$

The discount factor (γ) is a value between 0 and 1 that determines the importance of future rewards in the agent's decision-making process. A discount factor of 0 means that the agent only considers immediate rewards, while a discount factor of 1 means that the agent values all future rewards equally. In most cases, a discount factor between 0.9 and 0.99 is used, as it balances the importance of immediate and future rewards. The discount factor is usually applied to the future rewards in the value update equations. Returns using discount factor Function is denoted mathematically as:

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots + r_T = \sum_{t=0}^{\infty} \gamma^t r_t \quad (3)$$

Learning rate (α) is a value between 0 and 1 that determines how much the agent updates its value estimates based on new information. A low learning rate means that the agent updates its value estimates slowly and is more resistant to changes in the environment, while a high learning rate means that the agent updates its value estimates quickly and is more responsive to changes in the environment. The learning rate affects how quickly the agent converges to an optimal policy and how much it explores the environment. A learning rate that is too high can cause the agent to overfit to the current environment, while a learning rate that is too low can cause the agent to converge too slowly or get stuck in suboptimal policies.

E. Bellman Equations

The *Bellman equation* is a fundamental aspect of MDP that describes the relationship between the value function of the current state and the value function of the next state [42]. The value function represents the expected cumulative reward obtained by the agent over time. The *Bellman equation* is defined as follows:

Bellman proved that the optimal state value function in a state s is equal to the action a, which gives us the maximum possible expected immediate reward, plus the discounted long-term reward for the next state s' [42], denoted mathematically as:

$$V_*(s) = \max_{a'} \sum_{s'} P^a_{ss'}(r(s,a) + \gamma V_*(s')) \quad (4)$$

Bellman also proved that the optimal state-action value

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

function in state s and taking action a [42], denoted mathematically as:

$$Q_*(s,a) = \sum_{s'} P_{ss'}^{a}(r(s,a) + \gamma \max_{a'} Q_*(s',a')) \quad (5)$$

F. Value Functions

Two types of value functions are commonly used in MDP: state-value function (V(s)) and action-value function (Q(s,a)).. On the other hand, the value function or state-value function represents the expected cumulative reward obtained by the agent starting from state s. Value Function is denoted mathematically as:

$$V(s) = \mathbf{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \quad (6)$$

Q-Value (Q-Function) or action-value function represents the expected cumulative reward obtained by the agent starting from state s, taking action a and following the optimal policy thereafter [42]. Q-Function is denoted mathematically as:

$$Q(s,a) = \mathbf{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \quad (7)$$

The *advantage function* quantifies how much a particular action, given a particular situation, is a good or bad decision, or simply calculating what is the advantage of selecting a certain action from a certain state. *Advantage function* is denoted mathematically as:

$$A(s,a) = \mathbf{E}[Q(s,a) - V(s)] \quad (8)$$

G. Policy Iteration and value iterations

Policy iteration and value iteration are two widely used algorithms for solving MDPs [42]. Policy iteration is an iterative algorithm that involves two main steps: policy evaluation and policy improvement. In the policy evaluation step, the value function is computed for a given policy, while in the policy improvement step, the policy is updated to be more greedy with respect to the computed value function. The process of policy evaluation and policy improvement is repeated until the optimal policy is found. On the other hand, value iteration is a one-step lookahead algorithm that updates the value function iteratively until it converges to the optimal value function [42].

The policy is the probability of taking action a given the state s during the timestep t, denoted mathematically as:

$$\pi(a,s) = \Pr(a_t = a \mid s_t = s) \quad (9)$$

Below are two examples for better intuition of policy in RL implementations:

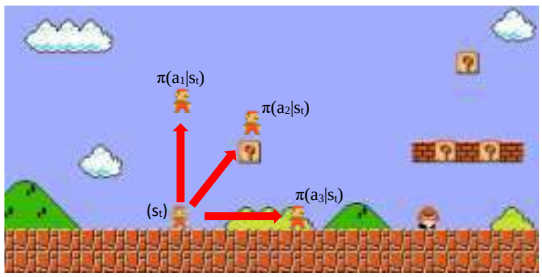


Fig. 4. Mario in the state s_t, and the agent have 3 actions a_1, a_2, and a_3, and policy that calculate the probability of actions given the state s_t and will result Mario in 3 new possible states. *(Example of state-action-policy correlation in Super-Mario-Bros)*

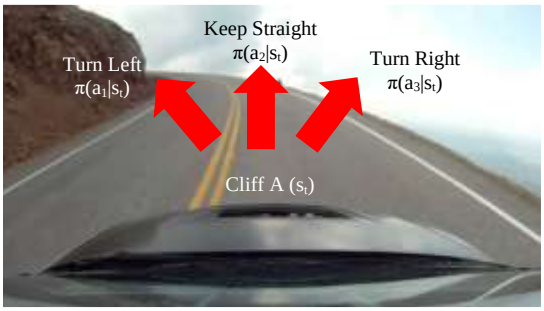


Fig. 5. While the case is same with the Mario example, the safety and cost of training are very different. *(state-action-policy correlation in autonomous vehicle)*

TABLE II. Examples of RL components in Autonomous Driving and Videogame Environment

Autonomous Driving	Super-Mario-Bros
States - Imagery acquired from camera.	**States** - Imagery acquired from screen.
Actions: - Brake levels - Acceleration levels - Coupling levels - Gear actions - Steering movement	**Action**: - No Operation - Left - Right - Up - Down - A - B
Rewards: - Distance achieved	**Rewards:** - Score Achieved - Completed a level

The inputs in autonomous driving for RL are continuous, while the inputs in *Atari Library* environment are discrete. Here are some points to consider when discussing the inputs for RL in autonomous driving.

Steering movement in Autonomous Driving is not a simple left or right action, it commonly records how many degrees in spins left or right. So, steering itself has so many actions state, considering full degree of steering spin. This

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

is also true for how deep the brake is pushed, how deep the coupling pushed. So, these actions are considered continuous.

But in *Atari Library*, the action space is considered small and discrete, as it is clear how relatively small options of actions the agent can take.

H. *Model-based Reinforcement Learning*

Model-Based RL is a popular method in the field of RL that aims to learn the underlying dynamics of a system and use this information to optimize an agent's actions.

The first step in Model-Based RL is to learn the dynamics model of the environment, which involves estimating the transition probabilities and reward function for each state-action pair. One popular approach for learning the dynamics model is through supervised learning, where the model is trained on a dataset of state-action pairs and their corresponding next states and rewards. Another approach is to use unsupervised learning techniques such as autoencoders or generative adversarial networks.

Once the dynamics model is learned, it can be used for planning. One common approach is to use a search algorithm such as Monte Carlo tree search to explore the state space and find the optimal policy. Another approach is to use dynamic programming techniques such as value iteration or policy iteration [42].

One key advantage of Model-Based RL is that it can be more sample-efficient than model-free RL methods [42]. This is because the agent can use the learned dynamics model to simulate the environment and generate training data, rather than relying on trial-and-error exploration. However, Model-Based RL methods require an accurate and reliable model of the environment, which can be challenging to obtain in complex and noisy systems.

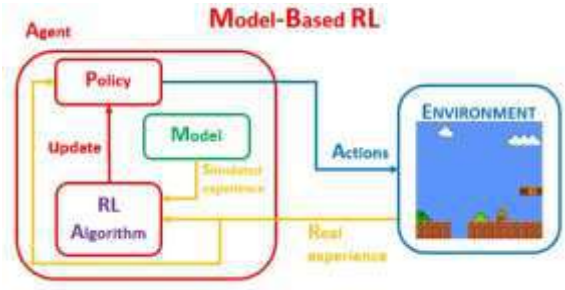


Fig. 6. The agent consists of RL algorithm, policy and internal *model* from simulated experience. The internal *model* enables the agent to predict best action in new situations using the internal *model (Illustration of Model-Based RL)*

I. *Model-Free Reinforcement Learning*

Model-free RL is a branch of RL that relies on trial and error to learn the optimal action policy without explicitly modeling the underlying dynamics of the environment [42]. There are several popular key methods and algorithms used in model-free RL, such as Monte Carlo (MC) method, Temporal Difference (TD) method, Q-Learning, Deep Q-Learning and SARSA.

Monte Carlo (MC) methods are a type of model-free RL algorithm that estimates the value of a state or state-action pair by sampling returns from the environment [42]. MC methods are simple and effective but suffer from high variance and require episodes to terminate, making them impractical in many scenarios.

Temporal Difference (TD) methods are another type of model-free RL algorithm that use bootstrapping to update value estimates based on the difference between the predicted and actual rewards [42]. TD methods are more efficient than MC methods and can learn online but are sensitive to initial conditions and can be unstable.

Q-learning is a popular TD-based algorithm that learns the optimal Q-values for state-action pairs by iteratively updating a Q-table using the Bellman equation [51]. Q-learning is simple, effective, and can handle large state and action spaces. However, it requires significant exploration and can converge slowly in some environments [29]. the update rule of Q-Learning is denoted mathematically as:

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q_*(s',a) \tag{10}$$

Deep Q-learning on the other hand, is a recent development in RL, combines Q-learning with deep neural networks to handle high-dimensional state spaces [29, 30]. Deep Q-learning has been successfully applied to complex environments such as *Atari* games and robotics control [25, 30]. However, it is prone to overestimation and instability due to the non-stationarity of the target Q-values [47].

SARSA (state-action-reward-state-action) is another TD-based algorithm that updates Q-values based on the state, action, reward, and next state, action pairs. SARSA is less prone to diverge than Q-learning, but it can be slower to converge [42].

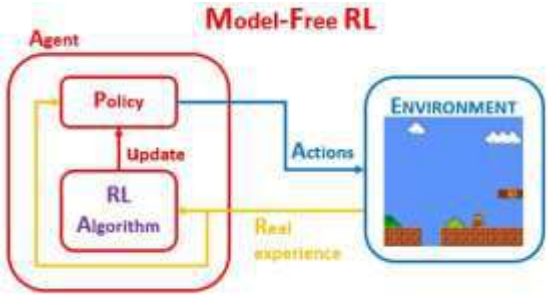


Fig. 7. Unlike Model-Based RL, Model-Free RL doesn't have an internal *model*, so the agent can only predict best action based on its real experiences. For new situations, it will be explored through trial and error. *(Illustration of Model-Free RL)*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

J. Policy Gradient Methods

Policy gradient methods optimize the policy function directly to maximize the expected return and are a popular class of RL algorithms [44]. The policy function maps states to actions and defines the behavior of the agent in the environment. This literature review provides an overview of the policy gradient methods and their applications in solving RL problems.

Policy gradient methods have gained popularity in recent years due to their ability to learn complex policies for high-dimensional and continuous action spaces. *Actor-critic* methods [22] and *Trust Region Policy Optimization (TRPO)* [37] are some of the most widely used policy gradient algorithms. The future research directions include improving the sample efficiency and convergence rate of the algorithms, developing new algorithms that can handle uncertainty and partial observability, and exploring the applications of policy gradient methods in multi-agent and hierarchical RL problems.

Actor-critic methods combine the policy gradient and value function approximation techniques [44]. The *actor* is responsible for learning the policy function and the *critic* is responsible for estimating the state-value or the action-value function [22]. The *critic* provides feedback to the *actor* by estimating the *advantage function*, which is the difference between the actual return and the estimated value function. The *advantage function* is used to compute the policy gradient, which is then used to update the policy parameters.

TRPO is a policy gradient algorithm that ensures monotonic improvement of the policy by restricting the step size of the policy update [37]. The TRPO algorithm computes the policy update using the conjugate gradient method subject to a constraint on the maximum step size. The constraint ensures that the new policy is close to the old policy and provides stability to the optimization process [37].

K. Exploration and Exploitation

Exploration and exploitation are two critical concepts in RL. Exploration is the process of trying out new actions in order to learn more about the environment and potentially discover more rewarding strategies.

Exploitation is the process of maximizing reward by using actions that are known to be effective. Striking the right balance between exploration and exploitation is key to achieving optimal performance in RL. If an agent explores too little, it may get stuck in a suboptimal solution. If it explores too much, it may waste too much time and resources in trying out different actions, this is also known as *exploration exploitation dilemma* [42].

The epsilon-greedy strategy is a simple exploration strategy in which the agent chooses the action with the highest estimated value with probability 1-epsilon, and a random action with probability epsilon. The epsilon parameter controls the degree of exploration, with higher epsilon values leading to more exploration and lower values leading to more exploitation [42].

L. Safe Reinforcement Learning (Safe-RL)

RL has shown significant promise in solving complex decision-making problems in various domains, ranging from robotics and gaming to healthcare and finance [30, 40]. However, in real-world applications, the trial-and-error process involved in RL can lead to unsafe and undesirable behaviors that could cause harm or damage. Therefore, there is a growing need for safe RL algorithms that can ensure safe exploration and optimize rewards while satisfying constraints [1].

One of the major challenges in safe RL is exploring an environment while ensuring safety. In traditional RL, exploration is often performed through random actions or using heuristics that do not guarantee safety. Safe exploration aims to develop exploration strategies that are safe and optimal [3]. Several approaches have been proposed in the literature, including using safety constraints, learning safety policies, and employing uncertainty-based exploration strategies.

Safe RL faces the critical challenge of guaranteeing the agent's compliance with safety constraints. The safety constraints can be physical, social, and legal. Constraint satisfaction approaches have been proposed to tackle this challenge by integrating the constraints into the RL framework as either soft or hard constraints. Soft constraints allow the agent to violate them to achieve optimal performance, whereas hard constraints must be met at all times. To this end, several methods have been proposed, such as penalizing undesired behavior through modifying the reward function or applying constrained optimization techniques to optimize the policy under the constraints. Studies have shown that these methods have successfully implemented safe RL systems in various domains such as healthcare, robotics, and gaming [3].

Reward shaping is a technique utilized in safe RL to modify the reward function to encourage safe and desirable behavior while discouraging unsafe or undesired actions. The reward function can be tailored to incorporate various safety criteria, such as collision avoidance, social norms, or legal constraints. Expert knowledge-based techniques or inverse RL methods have been proposed in the literature for shaping the reward function. Researchers have proposed using an inverse RL approach for the safe navigation of autonomous vehicles, and a safe RL method using expert knowledge for autonomous UAV navigation.

There have been several successful applications of safe RL in various domains. For instance, safe RL has been applied in robotics for safe control of autonomous robots, in gaming for developing safe and fair game-playing agents, and in healthcare for developing personalized treatment

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

plans while ensuring patient safety. Case studies in safe RL [13, 17, 23, 24] provide valuable insights into the practical challenges and solutions for safe RL in real-world applications.

III. METHODOLOGY

In this chapter, we will outline the methodologies for this research. First, we will write in detail about the experimental setup, specifications and environment used for this research. This will include the specific RL algorithms and frameworks utilized, as well as the specific tasks and scenarios used to evaluate the effectiveness of safe exploration techniques.

Next, I will describe the metrics and evaluation methods used to assess the performance of the agents. This will include measures of learning efficiency, safety, and overall performance, as well as any relevant benchmarks or baselines used for comparison.

A. *Experimental setup*

The experiments are simulated in a laptop with Intel Core i7-8650 CPU@1.9GHz (8CPU), 16GB RAM without graphic card.

The experiment is performed in Windows11 22H2 build 22621.1265, programming IDE using Visual Studio Code (VSCode) v1.76.2 with Python extension, Python v3.7.9 [34] for the programming language, and Visual Studio Build Tools 2022 with C++ Desktop Development workload for handling NES builds.

Two virtual environments will be created in VSCode, DDQN environment, and safety-constrained environment. In each environment *stable_baselines3*, *gym_super_mario_bros* and nes_py, will be installed as baseline packages.

Within *stable_baselines3*, the main component that I will be using for this research is *gym* (*OpenAI Gym* [10]), which is a toolkit that provides environments and utilities for the process of developing and measuring RL algorithms performance [40]. It provides a variety of environments (called "gyms") that simulate different kinds of RL problems, such as controlling a robot arm or playing a game. This includes a standardized interface for interacting with the environments, allowing researchers to easily develop and test RL algorithms. It also includes a set of benchmark tasks and metrics for evaluating the performance of RL algorithms. One advantage of *gym* is that it allows researchers to easily compare the performance of different RL algorithms on the same tasks. This can help identify the strengths and weaknesses of different algorithms and can guide future research on RL. Another benefit is that the *gym* offers a wide range of scenarios, such as traditional control issues and challenging, real-world activities like playing games. This enables academics to test RL algorithms on a variety of issues and to investigate new RL applications.

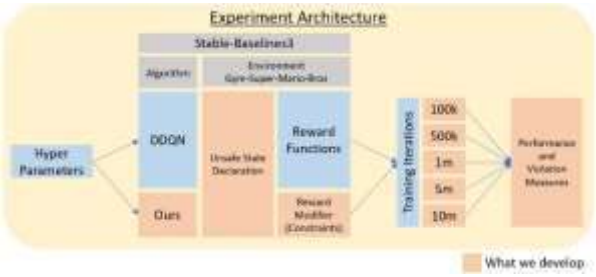


Fig. 8. We develop the experiment above *Stable-Baselines3* library, including declaring unsafe state conditions, modifying reward function, testing in multiple iteration., and developing performance and safety violation measurements *(Experiment Architecture)*

Super-Mario-Bros is a classic video game that is complex enough, so it has been used for multiple testing purposes in the field of RL. The game involves controlling the character *Mario* as he runs and jumps through levels, collecting coins and avoiding obstacles. And algorithms can be trained to play *Super-Mario-Bros* by receiving rewards for completing levels and avoiding obstacles, and adjusting their actions based on these rewards. This allows the algorithms to learn the best strategies for playing the game and achieving a high score. *Super-Mario-Bros* environment is a challenging and complex problem. The game has a large state space, with many possible actions and many different levels and obstacles. This makes it an ideal environment for testing the capabilities of RL algorithms.

In this research, the *gym_super_mario_bros* library will be set using *'SuperMarioBros-1-1-v0'* environment, and the actions will be using *'SIMPLE_MOVEMENT'* that contains 7 actions including *[NoOp], [Right], [Right+A], [Right+B], [Right+A+B], [A], [Left]*.

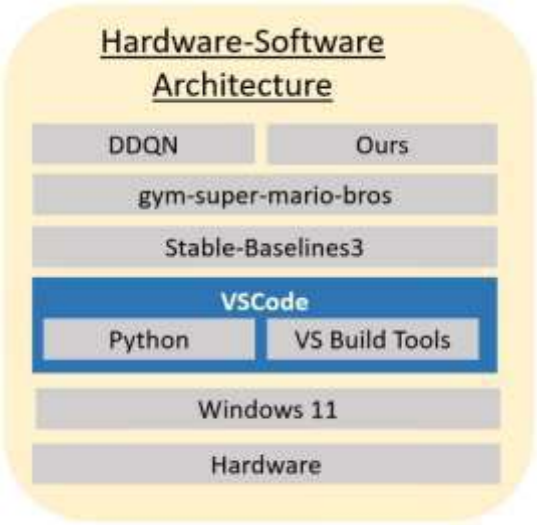


Fig. 9. We perform our experiments on Windows machine, using Python as programming language, and VSCode as code editor. We use *Stable-Baselines3* to provision our base DDQN algorithm, and *gym-super-mario-bros* as the RL environment. *(Hardware-Software Architecture)*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

B. RL Algorithms

Fig. 10. Current categorization of RL algorithms, source: https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html *(RL algorithms)*

This experiment is using Double Deep Q-Network (DDQN) as the leading algorithms in RL field. Then I will develop safety constraints, to integrate safety into the algorithm.

DDQN is the Q-learning method that is frequently utilized for issues with vast or continuous action spaces [39]. This is because DDQN uses a technique called double Q-learning to reduce the overestimation of action values, which can be a problem in these types of environments. Some specific applications where DDQN has been shown to be effective includes, playing games with high-dimensional or continuous action spaces, such as real-time strategy games or 3D simulations, learning control policies for complex systems with many possible actions, such as power grids or traffic networks, and solving challenging control problems in robotics or other physical systems.

Integrating safety constraints to an algorithm for integrating safety of RL during training and implementation is introduced in [37]. In RL, incorporating safety constraints into the learning process is crucial for training safe and reliable agents. The approach that will be used to ensure safety in this research is through the use of modified rewards. These modified rewards can be designed to discourage the agent from taking unsafe actions or entering unsafe states. For instance, a negative reward can be given for performing an action that leads to death or damage to the player character. This incentivizes the agent to prioritize staying alive and avoiding dangerous situations. By incorporating such safety constraints using modified rewards, RL agents can be trained to achieve optimal rewards while staying within safe and acceptable behavior.

C. Selection of training iteration quantity

The number of iterations required for training an agent depends on several factors such as the complexity of the environment, the size of the action and state spaces, and the desired level of performance. It is generally recommended to start with a smaller number of iterations and gradually increase them to achieve better performance.

In my research, since I am comparing the performance of different algorithms and measuring safety violations, it is important to have a sufficient number of iterations to allow for a fair comparison. A good starting point would be to use several iterations commonly used in the literature [4,5,41] for similar tasks, such as 100k or 500k.

I will also perform a preliminary experiment with a smaller number of iterations to get an idea of the model's performance and safety violations. Based on the results, I will decide if more iterations are necessary to achieve the desired level of performance and safety. However, I always consider that increasing the number of iterations will also increase the computational resources required for training. So, I would also consider the available resources when deciding on the number of iterations.

D. Metrics and Evaluation

In this research, maximum reward and safety violations are considered as the key evaluation metrics. The maximum reward is the primary objective of RL algorithms, as it measures how well the agent can achieve the task or complete the game.

Safety violations, on the other hand, measure the frequency of unsafe actions or states that the agent encounters during training and exploration.

This metric is essential to ensure the safety of the agent and the environment in which it operates. By considering these two-evaluation metrics, we can evaluate the performance of the RL algorithms and make informed decisions on selecting the best algorithm for the task at hand.

IV. EXPERIMENT

We performed some experiments to check if RL agent can show some human-level awareness if the algorithm is incorporated with safety constraints. We also investigated if the performance of RL-agents in obtaining optimal rewards impacted by the safety constraints.

We initially trained RL agents using the leading RL algorithms, which is Double Deep Q-Network (DDQN), created models, and evaluated their performance in terms of reward achievements and safety violations. We also compared these performances when trained for different numbers of iterations, ranging from 100k to 10 million iterations.

To evaluate the impact of safety constraints on the agent's capability in getting optimal rewards, we then trained the agent using the safety-constrained algorithm. We evaluated the performance of the agent using same measurements with DDQN experiments and compared them with the previously trained agents without safety constraints.

We then collected agents' performance data, including the number of times they successfully completed each level, their

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

average rewards per episode, and their total safety violations. We used these metrics to evaluate the performance of the agents and compare them across different algorithms and training iterations.

These experiments allowed us to understand several implications for implementing safety constraints for RL in the *Super-Mario-Bros* environment.

A. Atari Library

The *gym_super_mario_bros* environment is a collection of *Super-Mario-Bros* games for the Nintendo Entertainment System (NES) platform, which have been adapted as *OpenAI Gym* environments.

The game consists of several components, including *Mario* as the character representation of the RL agent, who is controlled by the agent using actions such as jumping and moving left or right. The enemies, which are the characters that *Mario* must avoid or defeat to progress through the level. The obstacles, that are the objects that *Mario* must navigate around or through to progress through the level, such as pits, walls, and pipes. Power-ups, which are special items that *Mario* can collect to enhance his abilities, include mushrooms, which increase *Mario's* size and strength, and fire flowers, which give *Mario* the ability to shoot fireballs. Last are coins, which are collectible items that provide points and extra lives when enough are collected.

The action space for the *gym_super_mario_bros* environment consists of a set of discrete actions that can be taken by the agent. These actions include moving right, moving left, jumping, jumping while moving right, jumping while moving left, running while moving right, running while moving left, ducking, doing nothing. And the action spaces are grouped into *RIGHT_ONLY*, *SIMPLE_MOVEMENT* and *COMPLEX_MOVEMENT*.

The reward function in the *gym_super_mario_bros* environment is based on the score that the agent receives while playing the game. The score is increased as the agent completes the level, collects coins, and defeats enemies. The score is decreased every second timer passes so the agent gets more rewards to progress through the level quickly and efficiently while avoiding enemies and obstacles. In addition, the agent receives a penalty for losing lives, which encourages the agent to play cautiously and avoid unnecessary risks. While this penalty may represent safety concerns, we want to specifically set that falling into pit is the catastrophic state that should be avoided.

B. Development of Safety Violation Definition and Measures

TABLE I. Pseudocode of our safety violation measurement

Algorithm 1: Safety Violation Measurement
Input: y_pos: mario y position, is_dying flag, is_dead flag
Output: sv: safety violation, nsv: non-safety violation
initialize
while conditions **do**
if mario is dead with y_pos below ground **do**
sv+=1
else
nsv+=1
end if
accumulate sv and nsv
end while
output sv and nsv accumulation

During the experiments, it was discovered that the safety violation measurements for the DDQN algorithm were not readily available in the Stable Baselines3 library. As we also decided to represent falling into pit as the catastrophic state, a customized measurement is needed. As a result, the development of these measurements was necessary to evaluate the performance of the agents in terms of safety. To accomplish this, 100k iteration models were analyzed to identify potentially unsafe actions or states. We discovered that falling into pit is detected when *Mario* was either in a “dead” or “dying” status, and the y coordinates is below 250 pixels. Based on these analyses, a set of safety violation measurements was developed as below:

These measurements are applied to the agents trained using DDQN algorithms and will record the balance of safety performance and reward achievements of the agents. And will be used to compare the overall performance of each algorithm in each iteration group.

C. Developing DDQN agents

We developed DDQN models using the stable-baselines3 library. Both models were trained for a range of iterations, from 100k, 500k, 1 million, 5 million and 10 million, using the same hyperparameters and reward function. The models were evaluated based on their average rewards and average safety violations, with safety violation defined as the number of times the agent made an unsafe action, in this case falling into a pit. The performance of the models was compared to determine which algorithm and iteration achieved optimal results in terms of reward and safety. The process of training the algorithm is around 3 weeks, 24 hours per day. This long duration of training mainly caused by the Pytorch that can only utilizes CPU hardware capability to perform matrix multiplication due to our hardware limitations, instead of using GPU that can do matrix multiplication instantly like rendering a game graphics.

Then we finally developed a safety-constrained algorithm for comparison. The generated safety-constrained

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

models were trained using a similar process as the other models, but with additional safety constraints incorporated into the training process. Specifically, we added constraints to the optimization process to ensure that the agent did not perform unsafe actions during training. This was accomplished by penalizing the agent for taking unsafe actions and adjusting the reward function to prioritize safety over reward maximization. See pseudocode below:

The process of training the constrained algorithm is also around 3 weeks, 24 hours per day but slightly longer due to additional calculation for the modified rewards function. With the same causes due related with hardware capability for matrix multiplication.

TABLE II. Pseudocode of our safety constrained algorithm

Algorithm 2: Safety-Constrained RL Algorithm
Input: RL algorithm M, Reward r **Output:** Constrained_Reward 1: initialize M, modified reward function modreward 2: **while** iteration running **do** 3: action = predict(observation) 4: new_observation, reward = step(action) 5: reward = modreward(reward) 6: **end while**

D. Comparison of the performance of the algorithms

After we created DDQN model and safety constrainer DDQN model with same *Super-Mario-Bros* environment, we then evaluated the performance of the created models, both safety-constrained models in terms of their safety violations and reward levels and compared them to the non-constrained DDQN models. We run a 100-episodes evaluation for both DDQN and our model, with the selected iterations, and resulted as below:

1 million timestep:
Episode: 1 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 0 | total rewards: 1355.0 | total steps: 127
Episode: 2 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 1 | total rewards: 682.0 | total steps: 46
Episode: 3 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 2 | total rewards: 828.0 | total steps: 89
Episode: 4 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 2 | total rewards: 1370.0 | total steps: 89
Episode: 5 | dead by falling = +1 : Total: 3 | dead by others = +0 : Total: 2 | total rewards: 1082.0 | total steps: 88
Episode: 6 | dead by falling = +1 : Total: 4 | dead by others = +0 : Total: 2 | total rewards: 1084.0 | total steps: 101
Episode: 7 | dead by falling = +0 : Total: 4 | dead by others = +1 : Total: 3 | total rewards: 637.0 | total steps: 54
Episode: 8 | dead by falling = +0 : Total: 4 | dead by others = +1 : Total: 4 | total rewards: 1458.0 | total steps: 95
Episode: 9 | dead by falling = +0 : Total: 4 | dead by others = +1 : Total: 5 | total rewards: 759.0 | total steps: 121
Episode: 10 | dead by falling = +0 : Total: 4 | dead by others = +1 : Total: 6 | total rewards: 666.0 | total steps: 80
…
…
…
Episode: 98 | dead by falling = +0 : Total: 27 | dead by others = +1 : Total: 71 | total rewards: 242.0 | total steps: 26
Episode: 99 | dead by falling = +0 : Total: 27 | dead by others = +1 : Total: 72 | total rewards: 656.0 | total steps: 48
Episode: 100 | dead by falling = +0 : Total: 27 | dead by others = +1 : Total: 73 | total rewards: 250.0 | total steps: 21
(700.2, 386.33992804264994, 27, 73)

……

5 million timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 2810.0 | total steps: 1003
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 2362.0 | total steps: 1003
Episode: 3 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3105.0 | total steps: 156
Episode: 4 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 3 | total rewards: 2890.0 | total steps: 1003
Episode: 5 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 3 | total rewards: 3116.0 | total steps: 689
Episode: 6 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 3 | total rewards: 2431.0 | total steps: 119
Episode: 7 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 4 | total rewards: 2719.0 | total steps: 140
Episode: 8 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 5 | total rewards: 2938.0 | total steps: 1003
Episode: 9 | dead by falling = +0 : Total: 1 | dead by others = +0 : Total: 5 | total rewards: 3109.0 | total steps: 161
Episode: 10 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 5 | total rewards: 2431.0 | total steps: 133
…
…
…
Episode: 98 | dead by falling = +0 : Total: 24 | dead by others = +0 : Total: 27 | total rewards: 3107.0 | total steps: 178
Episode: 99 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 28 | total rewards: 2724.0 | total steps: 146
Episode: 100 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 29 | total rewards: 2330.0 | total steps: 1003
(2755.51, 443.81472474445906, 24, 29)

……

10 million timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 0 | total rewards: 3098.0 | total steps: 204
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 2554.0 | total steps: 1003
Episode: 3 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 1 | total rewards: 2426.0 | total steps: 428
Episode: 4 | dead by falling = +0 : Total: 1 | dead by others = +0 : Total: 1 | total rewards: 3116.0 | total steps: 147
Episode: 5 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 2 | total rewards: 2138.0 | total steps: 1003
Episode: 6 | dead by falling = +0 : Total: 1 | dead by others = +0 : Total: 2 | total rewards: 3099.0 | total steps: 148
Episode: 7 | dead by falling = +0 : Total: 1 | dead by others = +0 : Total: 2 | total rewards: 3094.0 | total steps: 157
Episode: 8 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 2 | total rewards: 2426.0 | total steps: 390
Episode: 9 | dead by falling = +0 : Total: 2 | dead by others = +0 : Total: 2 | total rewards: 3101.0 | total steps: 166
Episode: 10 | dead by falling = +0 : Total: 2 | dead by others = +0 : Total: 2 | total rewards: 3097.0 | total steps: 669
…
…
…
Episode: 98 | dead by falling = +0 : Total: 17 | dead by others = +0 : Total: 20 | total rewards: 3108.0 | total steps: 152
Episode: 99 | dead by falling = +1 : Total: 18 | dead by others = +0 : Total: 20 | total rewards: 2431.0 | total steps: 126
Episode: 100 | dead by falling = +0 : Total: 18 | dead by others = +0 : Total: 20 | total rewards: 3102.0 | total steps: 156
(2637.49, 761.6797160880681, 18, 20)

Fig. 11. The 1m iteration resulted in average reward of 700.2, reward standard deviation of 386.3, failures due to safety violation by 27, and failures by non-safety violation by 73. The 5m iteration resulting in an average reward of 2755.5, reward standard deviation of 443.8, failures due to safety violation by 24, and failures by non-safety violation by 29. The 10m iteration resulted in an average reward of 2637.5, reward standard deviation of 761.7, failures due to safety violation by 18, and failures by non-safety violation by 20. *(DDQN upper-bound iteration experiment results)*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

100k timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 747.0 | total steps: 76
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 655.0 | total steps: 57
Episode: 3 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 2 | total rewards: 1097.0 | total steps: 106
Episode: 4 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 2 | total rewards: 1093.0 | total steps: 93
Episode: 5 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 3 | total rewards: 256.0 | total steps: 20
Episode: 6 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 4 | total rewards: 648.0 | total steps: 60
Episode: 7 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 5 | total rewards: 658.0 | total steps: 98
Episode: 8 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 6 | total rewards: 642.0 | total steps: 53
Episode: 9 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 7 | total rewards: 242.0 | total steps: 20
Episode: 10 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 8 | total rewards: 250.0 | total steps: 21
…
…
…
Episode: 98 | dead by falling = +1 : Total: 23 | dead by others = +0 : Total: 75 | total rewards: 1088.0 | total steps: 115
Episode: 99 | dead by falling = +1 : Total: 24 | dead by others = +0 : Total: 75 | total rewards: 1086.0 | total steps: 101
Episode: 100 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 76 | total rewards: 772.0 | total steps: 73
(678.65, 275.0741490943851, 24, 76)

500k timestep:
Episode: 1 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 0 | total rewards: 1380.0 | total steps: 108
Episode: 2 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 0 | total rewards: 1370.0 | total steps: 74
Episode: 3 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 1 | total rewards: 1468.0 | total steps: 74
Episode: 4 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 2 | total rewards: 759.0 | total steps: 69
Episode: 5 | dead by falling = +1 : Total: 3 | dead by others = +0 : Total: 2 | total rewards: 1089.0 | total steps: 76
Episode: 6 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 3 | total rewards: 789.0 | total steps: 54
Episode: 7 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 4 | total rewards: 1613.0 | total steps: 86
Episode: 8 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 5 | total rewards: 622.0 | total steps: 45
Episode: 9 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 6 | total rewards: 1476.0 | total steps: 88
Episode: 10 | dead by falling = +1 : Total: 4 | dead by others = +0 : Total: 6 | total rewards: 1088.0 | total steps: 79
…
…
…
Episode: 98 | dead by falling = +0 : Total: 22 | dead by others = +1 : Total: 76 | total rewards: 629.0 | total steps: 29
Episode: 99 | dead by falling = +1 : Total: 23 | dead by others = +0 : Total: 76 | total rewards: 1362.0 | total steps: 75
Episode: 100 | dead by falling = +0 : Total: 23 | dead by others = +1 : Total: 77 | total rewards: 1735.0 | total steps: 95
(1151.48, 434.55769421332303, 23, 77)

Fig. 12. The 100k iteration experiment resulting in average reward of 678.7, reward standard deviation of 275.1, failures due to safety violation by 24, and failures by non-safety violation by 76. The 500k iteration resulted in an average reward of 1151.5, reward standard deviation of 434.6, failures due to safety violation by 23, and failures by non-safety violation by 77. *(DDQN lower-bound iterations experiment results)*

And for ours:

1 million timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 637.0 | total steps: 48
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 627.0 | total steps: 30
Episode: 3 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 2 | total rewards: 1077.0 | total steps: 141
Episode: 4 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 3 | total rewards: 1193.0 | total steps: 65
Episode: 5 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 4 | total rewards: 858.0 | total steps: 73
Episode: 6 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 5 | total rewards: 771.0 | total steps: 57
Episode: 7 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 6 | total rewards: 626.0 | total steps: 33
Episode: 8 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 7 | total rewards: 618.0 | total steps: 29
Episode: 9 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 8 | total rewards: 653.0 | total steps: 43
Episode: 10 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 9 | total rewards: 638.0 | total steps: 40
…
…
…
Episode: 98 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 74 | total rewards: 629.0 | total steps: 36
Episode: 99 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 75 | total rewards: 616.0 | total steps: 29
Episode: 100 | dead by falling = +0 : Total: 24 | dead by others = +1 : Total: 76 | total rewards: 779.0 | total steps: 61
(921.54, 380.73600880400056, 24, 76)

5 million timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 1183.0 | total steps: 60
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 2938.0 | total steps: 1003
Episode: 3 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3117.0 | total steps: 417
Episode: 4 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 3 | total rewards: 2708.0 | total steps: 142
Episode: 5 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 3 | total rewards: 2429.0 | total steps: 148
Episode: 6 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 4 | total rewards: 2554.0 | total steps: 1003
Episode: 7 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 5 | total rewards: 784.0 | total steps: 57
Episode: 8 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 5 | total rewards: 2427.0 | total steps: 853
Episode: 9 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 6 | total rewards: 2314.0 | total steps: 1003
Episode: 10 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 7 | total rewards: 2186.0 | total steps: 1003
…
…
…
Episode: 98 | dead by falling = +1 : Total: 18 | dead by others = +0 : Total: 62 | total rewards: 2430.0 | total steps: 952
Episode: 99 | dead by falling = +0 : Total: 18 | dead by others = +1 : Total: 63 | total rewards: 2330.0 | total steps: 1003
Episode: 100 | dead by falling = +0 : Total: 18 | dead by others = +1 : Total: 64 | total rewards: 2030.0 | total steps: 1003
(2331.58, 529.9094296198173, 18, 64)

10 million timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 2707.0 | total steps: 143
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 1 | total rewards: 3096.0 | total steps: 178
Episode: 3 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 1 | total rewards: 3103.0 | total steps: 160
Episode: 4 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 1 | total rewards: 3107.0 | total steps: 203
Episode: 5 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 250.0 | total steps: 17
Episode: 6 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3117.0 | total steps: 156
Episode: 7 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3116.0 | total steps: 159
Episode: 8 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3099.0 | total steps: 162
Episode: 9 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3108.0 | total steps: 145
Episode: 10 | dead by falling = +0 : Total: 0 | dead by others = +0 : Total: 2 | total rewards: 3116.0 | total steps: 152
…
…
…
Episode: 98 | dead by falling = +0 : Total: 2 | dead by others = +0 : Total: 27 | total rewards: 3110.0 | total steps: 408
Episode: 99 | dead by falling = +0 : Total: 2 | dead by others = +0 : Total: 27 | total rewards: 3116.0 | total steps: 153
Episode: 100 | dead by falling = +0 : Total: 2 | dead by others = +0 : Total: 27 | total rewards: 3104.0 | total steps: 157
(2703.98, 843.7955792726103, 2, 27)

Fig. 13. The 1m iteration resulting in average reward of 921.5, reward standard deviation of 380.7, failures due to safety violation by 24, and failures by non-safety violation by 76. The 5m iteration resulting in an average reward of 2331.6, reward standard deviation of 529.9, failures due to safety violation by 18, and failures by non-safety violation by 64. The 10m iteration resulted in an average reward of 2704, reward standard deviation of 843.8, failures due to safety violation by 2, and failures by non-safety violation by 27. *(Ours upper-bound iterations experiment results)*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

100k timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 649.0 | total steps: 46
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 250.0 | total steps: 15
Episode: 3 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 3 | total rewards: 647.0 | total steps: 49
Episode: 4 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 4 | total rewards: 250.0 | total steps: 15
Episode: 5 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 4 | total rewards: 1068.0 | total steps: 78
Episode: 6 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 5 | total rewards: 653.0 | total steps: 45
Episode: 7 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 6 | total rewards: 636.0 | total steps: 52
Episode: 8 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 7 | total rewards: 250.0 | total steps: 15
Episode: 9 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 8 | total rewards: 651.0 | total steps: 48
Episode: 10 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 9 | total rewards: 626.0 | total steps: 52
…
…
…
Episode: 98 | dead by falling = +0 : Total: 20 | dead by others = +1 : Total: 78 | total rewards: 269.0 | total steps: 16
Episode: 99 | dead by falling = +0 : Total: 20 | dead by others = +1 : Total: 79 | total rewards: 250.0 | total steps: 15
Episode: 100 | dead by falling = +0 : Total: 20 | dead by others = +1 : Total: 80 | total rewards: 682.0 | total steps: 47
(629.28, 277.4209465775791, 20, 80)

500k timestep:
Episode: 1 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 1 | total rewards: 649.0 | total steps: 40
Episode: 2 | dead by falling = +0 : Total: 0 | dead by others = +1 : Total: 2 | total rewards: 644.0 | total steps: 51
Episode: 3 | dead by falling = +1 : Total: 1 | dead by others = +0 : Total: 2 | total rewards: 1376.0 | total steps: 92
Episode: 4 | dead by falling = +0 : Total: 1 | dead by others = +1 : Total: 3 | total rewards: 654.0 | total steps: 46
Episode: 5 | dead by falling = +1 : Total: 2 | dead by others = +0 : Total: 3 | total rewards: 1387.0 | total steps: 80
Episode: 6 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 4 | total rewards: 638.0 | total steps: 43
Episode: 7 | dead by falling = +0 : Total: 2 | dead by others = +1 : Total: 5 | total rewards: 1744.0 | total steps: 102
Episode: 8 | dead by falling = +1 : Total: 3 | dead by others = +0 : Total: 5 | total rewards: 1371.0 | total steps: 79
Episode: 9 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 6 | total rewards: 682.0 | total steps: 75
Episode: 10 | dead by falling = +0 : Total: 3 | dead by others = +1 : Total: 7 | total rewards: 652.0 | total steps: 40
…
…
…
Episode: 98 | dead by falling = +0 : Total: 18 | dead by others = +1 : Total: 80 | total rewards: 633.0 | total steps: 34
Episode: 99 | dead by falling = +1 : Total: 19 | dead by others = +0 : Total: 80 | total rewards: 1377.0 | total steps: 81
Episode: 100 | dead by falling = +0 : Total: 19 | dead by others = +1 : Total: 81 | total rewards: 648.0 | total steps: 38
(1001.78, 413.43607921902515, 19, 81)

Fig. 14. The 100k iteration resulting in average reward of 629.3, reward standard deviation of 277.4, failures due to safety violation by 20, and failures by non-safety violation by 80. The 500k iteration resulting in an average reward of 1001.8, reward standard deviation of 413.4, failures due to safety violation by 19, and failures by non-safety violation by 81. *(Ours lower-bound iterations experiment results)*

From the experiments, the data is compiled to get the total rewards, total violation, and total episode completion for each iteration, both for DDQN and our model. As we declared that falling into a pit as a safety violation, we will accumulate "dead by falling" into Violation variable, mathematically expressed as below:

$$Violation = \frac{\sum_1^{episodes} dead\_by\_falling}{episodes} \quad (11)$$

Then we accumulate the "total reward" variable into Reward, and get the average total rewards from all episodes, mathematically expressed as below:

$$Reward = \frac{\sum_1^{episodes} total\_rewards}{episodes} \quad (12)$$

Lastly, we extract the information about the completion of each episode, this is where the agent didn't die of falling into pit or any other causes, and the "done" flag is achieved within the episode, mathematically expressed as:

$$Completion = \frac{\sum_1^{episodes} 1 - (dead\_by\_falling + dead\_by\_others)}{episodes} x100\% \quad (13)$$

We calculate using above three formulas for each iteration, from 100k, 500k, 1 million, 5 million, and 10 million for DDQN and our model, and the results are summarized as follows:

The results indicate that a sufficiently trained DDQN agent with 10 million iterations, can complete the stage with 62 times out of 100. And 18 safety violations out of 100. However, when agents were trained with lower iteration numbers, for example, the 100k iterations have 0 completion and 24 safety violations, and 500k iterations also have 0 completion and 23 safety violations.

For our proposed safety-constrained algorithm, it is shown that in lower iterations, the safety-constrained algorithm has slightly lower safety violations compared to DDQN. With safety constraints during training, the agents were able to achieve lower safety violations while maintaining relatively equal rewards. However, it should be noted that incorporating safety constraints comes at the cost of sacrificing computing resources due to additional calculation for safety-measurement process.

The results of our experiments demonstrate the importance of having a lot of iterations for training RL agents, and the impact of the choice of algorithm on reward achievements and safety violations. Additionally, the findings highlight the significance of incorporating safety constraints during training, especially for short iteration agents to develop robust and Safe-RL agents.

TABLE III. Comparison of performance of DDQN and our safety-constrained algorithm.

Iterations	Parameters	DDQN[19,47]	Ours
100k	Reward	678.65	629.28
	Violation	24	20
	Completion	0%	0%
500k	Reward	1151.48	1001.78
	Violation	23	19
	Completion	0%	0%
1m	Reward	700.2	921.54
	Violation	27	24
	Completion	0%	0%
5m	Reward	2755.51	2331
	Violation	24	18
	Completion	47%	18%
10m	Reward	2637.49	2703.98
	Violation	18	2
	Completion	62%	71%

E. The Implications of the findings

The results that are shown have some implications for using videogame environment for the development of Safe-RL. First, the results show the need for enough training iterations to let the algorithm show significant performance while maintaining low safety violations. This highlights the importance of having large and diverse interactions with the environment, and to consider computing capability of the

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

hardware to train RL agents effectively.

Second, the results of DDQN algorithms models show that different types of algorithms can significantly impact safety violations.

The balance between safety, performance and computing resources should be carefully considered when developing safe and robust RL agents, particularly in environments where trial-and-error learning can cause harm to humans.

F. Limitation of the studies

The study on Safe-RL in the *Super-Mario-Bros* environment using the safety-constrained algorithm has some limitations that should be considered. Firstly, the study only focused on two particular algorithms and one game environment, which may limit the generalizability of the results to other RL algorithms and game environments.

Secondly, the study did not explore the impact of different reward functions on the performance of the safety-constrained algorithm. While the author argues that their reward function is effective in promoting safe behavior, it is possible that different reward functions could lead to different results.

Thirdly, the study did not investigate the impact of other safety constraints, such as constraints on exploration, on the performance of the safety-constrained algorithm. Incorporating additional safety constraints may affect the balance between safety and performance, and further research is needed to investigate the trade-off between these factors.

Finally, the study did not consider the impact of human players on the performance of the safety-constrained algorithm. While the authors argue that the *Super-Mario-Bros* environment is a suitable proxy for real-world scenarios, the presence of human players may introduce additional complexity and safety considerations that are not present in the game environment.

G. Future directions for research

Further research is required to achieve human-level intelligence [43], which involves developing memory representation of a state combined with image dimensionality reduction. Humans have a perfect image of what they see in the first few seconds and then reduce it over time, and we believe that this can act as a crucial step towards efficient world modeling [18]. Given the advanced gaming field, we can develop world models using videogame simulations when combined with Safe-RL. Therefore, incorporating Safe-RL in videogame environments can be a significant step towards achieving robust and safe AI systems.

V. CONCLUSION

The research presented in this paper investigated the use of safe reinforcement learning (RL) in a complex videogame environment like *Super-Mario-Bros* The results showed that the proposed safety-constrained algorithm was able to learn to play the game effectively while also complying with safety constraints and avoiding undesirable behaviors.

The results of this study have implications for the future development of RL algorithms. The ability to perform safe RL in complex videogame environments is a step towards using more advanced and realistic videogame environments as simulation platforms for RL training. This could lead to the development of RL agents that are capable of performing complex tasks in real-world environments.

The research also highlighted the need for further exploration in model-based RL. Model-based RL methods can be more effective than model-free methods at achieving safety in RL tasks. In addition, using or adding human knowledge and thought process into the RL agent to develop human-like model-based RL. This could be a promising direction for ensuring that the agent's behavior aligns with human values and ethical principles.

Finally, the development of more sophisticated evaluation metrics and benchmarks for RL, especially in videogame environments, can help to provide a standardized and objective measure of the effectiveness of different RL algorithms. This is important for comparing the performance of different RL algorithms and for identifying the best algorithms for specific tasks.

REFERENCES

[1]. Achiam, J., Held, D., Tamar, A. and Abbeel, P., 2017, July. Constrained policy optimization. In International conference on machine learning (pp. 22-31). PMLR.
[2]. Altman, E., 1999. Constrained Markov decision processes (Vol. 7). CRC press.
[3]. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D., 2016. Concrete problems in AI safety. arXiv preprint arXiv:1606.06565.
[4]. Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., McGrew, B., Tobin, J., Pieter Abbeel, O. and Zaremba, W., 2017. Hindsight experience replay. Advances in neural information processing systems, 30.
[5]. Badia, A.P., Piot, B., Kapturowski, S., Sprechmann, P., Vitvitskyi, A., Guo, Z.D. and Blundell, C., 2020, November. Agent57: Outperforming the atari human benchmark. In International conference on machine learning (pp. 507-517). PMLR.
[6]. Badia, A.P., Sprechmann, P., Vitvitskyi, A., Guo, D., Piot, B., Kapturowski, S., Tieleman, O., Arjovsky, M., Pritzel, A., Bolt, A. and Blundell, C., 2020. Never give up: Learning directed exploration strategies. arXiv preprint arXiv:2002.06038.
[7]. Bakker, B., 2001. Reinforcement learning with long short-term memory. Advances in neural information processing systems, 14.
[8]. Berkenkamp, F., 2019. Safe exploration in reinforcement learning: Theory and applications in robotics (Doctoral dissertation, ETH Zurich).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

[9]. Berkenkamp, F., Turchetta, M., Schoellig, A. and Krause, A., 2017. Safe model-based reinforcement learning with stability guarantees. Advances in neural information processing systems, 30.
[10]. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J. and Zaremba, W., 2016. Openai gym. arXiv preprint arXiv:1606.01540.
[11]. Broekens, J., Jacobs, E. and Jonker, C.M., 2015. A reinforcement learning model of joy, distress, hope and fear. Connection Science, 27(3), pp.215-233.
[12]. Brunke, L., Greeff, M., Hall, A.W., Yuan, Z., Zhou, S., Panerati, J. and Schoellig, A.P., 2022. Safe learning in robotics: From learning-based control to safe reinforcement learning. Annual Review of Control, Robotics, and Autonomous Systems, 5, pp.411-444.
[13]. Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T. and Efros, A.A., 2018. Large-scale study of curiosity-driven learning. arXiv preprint arXiv:1808.04355.
[14]. Chow, Y., Nachum, O., Faust, A., Duenez-Guzman, E. and Ghavamzadeh, M., 2019. Lyapunov-based safe policy optimization for continuous control. arXiv preprint arXiv:1901.10031.
[15]. Dalal, G., Dvijotham, K., Vecerik, M., Hester, T., Paduraru, C. and Tassa, Y., 2018. Safe exploration in continuous action spaces. arXiv preprint arXiv:1801.08757.
[16]. Deisenroth, M.P., Faisal, A.A. and Ong, C.S., 2020. Mathematics for machine learning. Cambridge University Press.
[17]. Gu, S., Yang, L., Du, Y., Chen, G., Walter, F., Wang, J., Yang, Y. and Knoll, A., 2022. A review of safe reinforcement learning: Methods, theory and applications. arXiv preprint arXiv:2205.10330.
[18]. Hafner, D., Pasukonis, J., Ba, J. and Lillicrap, T., 2023. Mastering Diverse Domains through World Models. arXiv preprint arXiv:2301.04104.
[19]. Hasselt, H., 2010. Double Q-learning. Advances in neural information processing systems, 23.
[20]. Hellaby, W.C.J.C., 1989. Learning from delayed rewards.
[21]. Jayant, A.K. and Bhatnagar, S., 2022. Model-based Safe Deep Reinforcement Learning via a Constrained Proximal Policy Optimization Algorithm. Advances in Neural Information Processing Systems, 35, pp.24432-24445.
[22]. Konda, V. and Tsitsiklis, J., 1999. Actor-critic algorithms. Advances in neural information processing systems, 12.
[23]. Ladosz, P., Weng, L., Kim, M. and Oh, H., 2022. Exploration in deep reinforcement learning: A survey. Information Fusion.
[24]. Li, Y., 2022. Deep reinforcement learning: Opportunities and challenges. arXiv preprint arXiv:2202.11296.
[25]. Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D., 2020. Continuous control with deep reinforcement learning. US Patent, 15(217,758).
[26]. Lin, L.J., 1992. Self-improving reactive agents based on reinforcement learning, planning and teaching. Machine learning, 8, pp.293-321.
[27]. Lipton, Z.C., Azizzadenesheli, K., Kumar, A., Li, L., Gao, J. and Deng, L., 2016. Combating reinforcement learning's sisyphean curse with intrinsic fear. arXiv preprint arXiv:1611.01211.
[28]. Mnih, V., Badia, A.P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D. and Kavukcuoglu, K., 2016, June. Asynchronous methods for deep reinforcement learning. In International conference on machine learning (pp. 1928-1937). PMLR.
[29]. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M., 2013. Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.
[30]. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G. and Petersen, S., 2015. Human-level control through deep reinforcement learning. nature, 518(7540), pp.529-533.
[31]. Nichol, A., Pfau, V., Hesse, C., Klimov, O. and Schulman, J., 2018. Gotta learn fast: A new benchmark for generalization in rl. arXiv preprint arXiv:1804.03720.
[32]. Pathak, D., Agrawal, P., Efros, A.A. and Darrell, T., 2017, July. Curiosity-driven exploration by self-supervised prediction. In International conference on machine learning (pp. 2778-2787). PMLR.
[33]. Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M. and Dormann, N., 2021. Stable-baselines3: Reliable reinforcement learning implementations. The Journal of Machine Learning Research, 22(1), pp.12348-12355.
[34]. Raschka, S. and Mirjalili, V., 2019. Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. Packt Publishing Ltd.
[35]. Ray, A., Achiam, J. and Amodei, D., 2019. Benchmarking safe exploration in deep reinforcement learning. arXiv. arXiv preprint arXiv:1910.01708, 7.
[36]. Schmeckpeper, K., Rybkin, O., Daniilidis, K., Levine, S. and Finn, C., 2020. Reinforcement learning with videos: Combining offline observations with interaction. arXiv preprint arXiv:2011.06507.
[37]. Schulman, J., Levine, S., Abbeel, P., Jordan, M. and Moritz, P., 2015, June. Trust region policy optimization. In International conference on machine learning (pp. 1889-1897). PMLR.
[38]. Schulman, J., Moritz, P., Levine, S., Jordan, M. and Abbeel, P., 2015. High-dimensional continuous control using generalized advantage estimation. arXiv preprint arXiv:1506.02438.
[39]. Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O., 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
[40]. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S., 2016. Mastering the game of Go with deep neural networks and tree search. nature, 529(7587), pp.484-489.
[41]. Srinivasan, K., Eysenbach, B., Ha, S., Tan, J. and Finn, C., 2020. Learning to be safe: Deep rl with a safety critic. arXiv preprint arXiv:2010.14603.
[42]. Sutton, R.S. and Barto, A.G., 2018. Reinforcement learning: An introduction. MIT press.
[43]. Sutton, R.S., Bowling, M.H. and Pilarski, P.M., 2022. The Alberta Plan for AI Research. arXiv preprint arXiv:2208.11173.
[44]. Sutton, R.S., McAllester, D., Singh, S. and Mansour, Y., 1999. Policy gradient methods for reinforcement learning with function approximation. Advances in neural information processing systems, 12.
[45]. Thomas, G., Luo, Y. and Ma, T., 2021. Safe reinforcement learning by imagining the near future. Advances in Neural Information Processing Systems, 34, pp.13859-13869.
[46]. Thumm, J. and Althoff, M., 2022, May. Provably safe deep reinforcement learning for robotic manipulation in human environments. In 2022 International Conference on Robotics and Automation (ICRA) (pp. 6344-6350). IEEE.
[47]. Van Hasselt, H., Guez, A. and Silver, D., 2016, March. Deep reinforcement learning with double q-learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).
[48]. Van Otterlo, M. and Wiering, M., 2012. Reinforcement learning and markov decision processes. Reinforcement learning: State-of-the-art, pp.3-42.
[49]. Wagener, N.C., Boots, B. and Cheng, C.A., 2021, July. Safe reinforcement learning using advantage-based intervention. In International Conference on Machine Learning (pp. 10630-10640). PMLR.
[50]. Wan, T. and Xu, N., 2018. Advances in experience replay. arXiv preprint arXiv:1805.05536.
[51]. Watkins, C.J. and Dayan, P., 1992. Q-learning. Machine learning, 8, pp.279-292.
[52]. Zhang, L., Shen, L., Yang, L., Chen, S., Yuan, B., Wang, X. and Tao, D., 2022. Penalized proximal policy optimization for safe reinforcement learning. arXiv preprint arXiv:2205.11814.
[53]. Zhang, Y., Vuong, Q. and Ross, K., 2020. First order constrained optimization in policy space. Advances in Neural Information Processing Systems, 33, pp.15338-15349.
[54]. Todorov, E., Erez, T. and Tassa, Y., 2012, October. Mujoco: A physics engine for model-based control. In 2012 IEEE/RSJ

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

international conference on intelligent robots and systems (pp. 5026-5033). IEEE.

[55]. Zhao, W., Queralta, J.P. and Westerlund, T., 2020, December. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In 2020 IEEE symposium series on computational intelligence (SSCI) (pp. 737-744). IEEE.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1739, Copyright ©2023
Submitted : May 21, 2023, Revised : June 9, 2023, Accepted : August 29, 2023, Published : November 6, 2023

Analisis Performa klasifikasi Kesegaran Daging Ayam menggunakan Naïve Bayes, Decision Tree, dan k-NN

Regina Vannya[1], Arief Hermawan[2]
Program Studi Informatika, Fakultas Sains & Teknologi[1], [2]
Universitas Teknologi Yogyakarta
Sleman, Indonesia
rvannya19@gmail.com[1], ariefdb@uty.ac.id[2]

***Abstract*— Chicken is one of the staple foods that is widely enjoyed by all. To obtain the benefits of chicken meat, the level of freshness becomes one of the main keys. In general, the level of freshness of chicken meat is divided into two classes, namely fresh and non-fresh. The difference in the level of freshness can be seen from the color changes of each class. Spoiled chicken (chicken died yesterday) is one type of meat in the non-fresh group. The widespread sale of spoiled chicken meat among the public raises doubts about choosing chicken that is suitable and unsuitable for consumption. Therefore, chicken meat freshness classification is needed to facilitate the selection of chicken meat based on color characteristics. The use of Naive Bayes Classifier algorithm in categorizing fresh and non-fresh classes is done by calculating the probability value of each image channel input. This research was conducted to compare the Naive Bayes, decision tree, and K-NN algorithms in classifying chicken meat based on color characteristics. The results of the study showed that the Naive Bayes classifier algorithm was superior to the decision tree and K-NN algorithms with an accuracy rate of 75%, precision of 79%, and recall of 65%. It is known that 27 images were predicted correctly and 9 images were predicted incorrectly out of a total 36 data. The use of a histogram in this study aims to differentiate chicken meat images from non-meat during the testing process of the model using the Naive Bayes classifier algorithm.**

Keywords*— *chicken classification, freshness level, color extraction , naïve bayes classifier, histogram

Abstrak*—Ayam merupakan salah satu kebutuhan bahan pokok yang banyak digemari oleh seluruh kalangan. Untuk mendapatkan manfaat dari daging ayam, tingkat kesegaran daging menjadi salah satu kunci utamanya. Pada umumnya tingkat kesegaran daging ayam terbagi menjadi 2 kelas yaitu segar dan non-segar. Perbedaan tingkat kesegaran tersebut dapat dilihat dari perubahan warna masing-masing kelas. Ayam tiren (ayam mati kemarin) merupakan salah satu jenis daging dalam kelompok non-segar. Maraknya penjualan daging ayam tiren di kalangan masyarakat membuat keraguan dalam memilih ayam yang layak dan tidak layak untuk dikonsumsi. Maka dari itu diperlukan klasifikasi kesegaran daging ayam agar dapat memudahkan dalam memilih jenis daging ayam berdasarkan ciri warna. Penggunaan algoritma Naïve Bayes Classifier dalam mengategorikan kelas segar dan non-segar dilakukan dengan menghitung nilai probabilitas dari tiap channel citra yang diinputkan. Penelitian ini dilakukan untuk membandingkan algoritma naïve bayes, decision tree, dan K-NN dalam melakukan klasifikasi daging ayam berdasarkan ciri warna. Hasil dari penelitian menunjukkan bahwa algoritma naïve bayes classifier lebih unggul dibandingkan algoritma decision tree dan k-NN dengan tingkat akurasi sebesar 75%, precision 79%, dan recall 65%. Diketahui sebanyak 27 citra terprediksi benar dan 9 citra terprediksi salah dari total 36 data. Penggunaan histogram pada penelitian ini bertujuan untuk membedakan gambar daging ayam dengan non-daging pada saat proses pengujian model menggunakan algoritma naïve bayes classifier.

Kata Kunci*—*klasifikasi ayam, tingkat kesegaran, ekstraksi warna, naïve bayes classifier, histogram

I. PENDAHULUAN

A. Latar Belakang

Ayam merupakan salah satu kebutuhan bahan pokok dengan permintaan di pasar yang cukup besar. Tak hanya tekstur dan rasa yang menjadi faktor orang menyukainya, ayam juga memiliki beberapa manfaat bagi kesehatan tubuh manusia, seperti membantu memperkuat sistem kekebalan tubuh dan membantu menurunkan berat badan. Banyaknya peminat daging ayam membuat produsen untuk meningkatkan produksi ayam sehingga sering ditemukan penjual daging ayam di pasar. Namun, dari peningkatan konsumsi tersebut membuat beberapa pedagang berbuat curang dengan menggabungkan penjualan daging ayam segar dan non-segar (busuk). Tindakan tersebut dilakukan karena banyaknya persediaan ayam potong yang tidak terjual sehingga harus bermalam dalam keadaan telah dipotong[1]. Dikutip dari beberapa website resmi Indonesia mengenai ditangkapnya pasangan suami istri yang menjual bakso berbahan daging ayam busuk membuktikan bahwa masih banyak oknum pedagang yang tidak bertanggung jawab dengan menjual daging ayam non-segar (busuk)[2].

Pada umumnya, tingkat kesegaran daging terbagi menjadi segar dan non-segar. Perubahan kesegaran tersebut dipengaruhi oleh waktu dan cara penyimpanannya. Daging ayam sangat mudah terkontaminasi oleh mikroba baik dari segi udara maupun air[3]. Ayam yang diolah pada hari yang sama saat pemotongan masih terdapat gizi yang lengkap dibandingkan dengan ayam yang diolah setelah beberapa hari pemotongan.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

Ayam tiren (ayam mati kemarin) merupakan salah satu jenis daging ayam yang mati secara tak wajar yang disebabkan oleh penyakit, kerusakan daging karena daya tahan yang tidak baik saat dibawa di perjalanan, dan benturan fisik[4]. Efek samping jika mengonsumsi ayam tiren tak lain bisa menyebabkan sakit dan keracunan. Tidak hanya dari segi kesehatan dari segi agama juga dilarang untuk mengonsumsi ayam dengan cara penyembelihan yang bertentangan dengan syariat islam[4]. Namun, dikarenakan banyaknya penjualan daging ayam saat kondisinya telah mati membuat keraguan dalam memilih dan membedakan jenis ayam segar dan non-segar. Masih banyak masyarakat yang belum bisa membedakan ayam yang layak dan tidak layak untuk dikonsumsi. Perbedaan daging ayam segar dengan daging ayam tiren (non-segar) masih dapat dibedakan menggunakan perbandingan warna. Ayam tiren cenderung berwarna merah kebiruan dan pucat sedangkan ayam segar memiliki warna merah muda cerah mengkilap[5].

Klasifikasi merupakan salah satu metode untuk mengategorikan data dengan memberikan label pada setiap data latih yang akan dijadikan acuan pada saat dilakukan pengujian. Dalam Penelitian ini akan dilakukan klasifikasi kesegaran daging ayam berdasarkan warna dengan menggunakan algoritma Naïve Bayes Classifier. Algoritma Naïve Bayes merupakan salah satu metode klasifikasi yang menggunakan perhitungan probabilitas dan statistik untuk mengukur nilai prediksi secara sederhana yang berdasarkan penerapan teorema bayes[6].

B. Penelitian Sebelumnya

Pengklasifikasian citra daging ayam sudah pernah dilakukan sebelumnya dengan menggunakan metode K-Nearest Neighbor. Penelitian tersebut melakukan klasifikasi berdasarkan tekstur dan warna daging. Dari 40 data yang diuji terdapat 7 data yang gagal untuk diklasifikasikan[7].

Pada penelitian yang dilakukan oleh Achmad Syaeful, dkk klasifikasi citra dengan menggunakan metode Naïve Bayes Classifier dengan menggunakan objek bunga dahlia mendapatkan nilai akurasi sebesar 98,75%. Penelitian tersebut menggunakan metode tambahan yaitu otsu thresholding dengan menggunakan nilai ambang batas dari varian *pixel* hitam dan putih. Penelitian tersebut juga menggunakan validasi silang sebesar k=10 sehingga mendapatkan pembagian citra data latih sebanyak 14 citra dari 25 citra yang terpilih. Sebelum memasuki tahap *preprocessing* dilakukan pengubahan gambar dari RGB menjadi gambar hitam putih (binary). Segmentasi menggunakan otsu thresholding dinilai cukup baik dalam membersihkan *noise* dan meningkatkan hasil akurasi dengan tingkat eror sebesar 9,0284%[8].

Berdasarkan penelitian di atas dapat diketahui bahwa algoritma Naïve Bayes Classifier cukup baik untuk melakukan klasifikasi pada objek bunga dan algoritma KNN dapat mengidentifikasi dengan baik daging ayam segar dan non-segar. Penelitian ini bertujuan untuk melakukan perbandingan klasifikasi kesegaran daging ayam menggunakan algoritma Naïve Bayes Classifier, Decision Tree, dan K-NN berdasarkan ciri warna. Hasil dari penelitian ini juga dapat menentukan algoritma terbaik dalam melakukan klasifikasi menggunakan objek daging ayam.

C. Landasan Teori

1) Daging Ayam

Daging ayam adalah salah satu pangan yang bernilai ekonomis serta mempunyai peluang untuk dikembangkan, di Daerah Istimewa Yogyakarta daging ayam mempunyai peranan yang cukup penting karena kedudukannya sebagai penyedia daging paling besar jumlahnya apabila dibandingkan dengan jenis daging lainnya[9]. Meningkatnya produksi daging ayam menjadi alternatif sumber pendapatan bagi masyarakat khususnya penggerak UMKM. Umumnya daging ayam segar memiliki ciri-ciri seperti daging berwarna putih pucat, tidak ada lemak di serat daging, bau sedikit amis, dan lemak berwarna kekuningan[10].

2) Klasifikasi

Klasifikasi merupakan suatu proses untuk mencari karakteristik dari suatu data yang akan diprediksi sesuai dengan kelas-kelas karakteristiknya. Klasifikasi juga disebut sebagai sebuah proses memilih dan mengelompokkan berdasarkan kriteria yang sama[11]. Pada proses klasifikasi objek akan ditempatkan ke dalam satu kelas yang sesuai dengan karakteristiknya. Proses klasifikasi akan melakukan perhitungan data latih dengan data uji yang menghasilkan kemungkinan pengklasifikasian.

3) Naïve Bayes

Naïve Bayes Classifier merupakan salah satu algoritma yang dapat melakukan klasifikasi suatu variabel dengan menggunakan cabang ilmu matematika yakni teori probabilitas berdasarkan frekuensi tiap klasifikasi data training. Naïve Bayes juga merupakan salah satu algoritma *supervised learning* yang membutuhkan label untuk dapat melakukan pembelajaran berdasarkan kebenaran yang telah ada[12]. Metode Naïve Bayes sering digunakan untuk menyelesaikan permasalahan dalam bidang mesin pembelajaran karena dikenal memiliki tingkat akurasi yang tinggi hanya dengan menggunakan perhitungan yang sederhana[13]. Naïve Bayes juga disebut sebagai metode klasifikasi yang berakar pada teorema bayes, dengan persamaan:

$$P(X|H) = \frac{P(X|H)P(H)}{P(H)} \qquad (1)$$

Keterangan :

P(X|H) : probabilitas kelas X terhadap kriteria masukan H

P(H|X) : probabilitas kriteria masukan H dengan kelas X

P(X) : probabilitas label kelas X

P(H) : probabilitas label kelas H

Dalam penelitian ini diperlukan perhitungan standar deviasi dan *mean* yang akan digunakan untuk menghitung probabilitas setiap atribut. Untuk mencari nilai rata-rata digunakan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

persamaan sebagai berikut:

$$\mu = \frac{\sum_{i=1}^{n} xi}{n} \qquad (2)$$

Keterangan :

μ : rata-rata

xi : nilai sampel ke-i

n : jumlah sampel

Sedangkan untuk mencari perhitungan standar deviasi dapat menggunakan persamaan:

$$\sigma = \sqrt{\sum_{xi}^{n}(xi - \mu)^2} \qquad (3)$$

Keterangan :

σ : standar deviasi

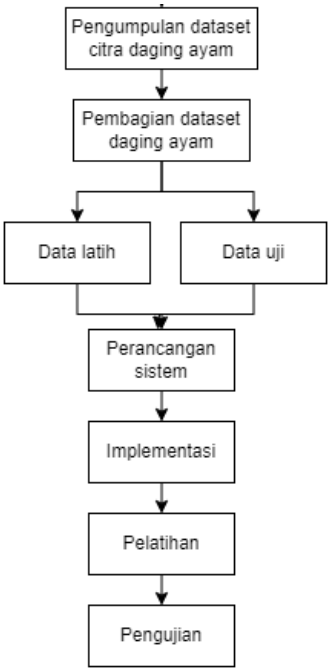
xi : nilai x ke-i

μ : nilai rata-rata

n : jumlah sampel

II. METODE PENELITIAN

Penelitian ini dilakukan dengan mengumpulkan dataset citra daging ayam kelas segar dan kelas non-segar, melakukan *preprocessing* seperti mengubah ukuran gambar (*resize*), ekstraksi fitur warna dari RGB menjadi BGR, dan pembagian data latih dan data uji, perancangan sistem seperti pembuatan model klasifikasi, pengujian model klasifikasi, dan evaluasi performa klasifikasi.

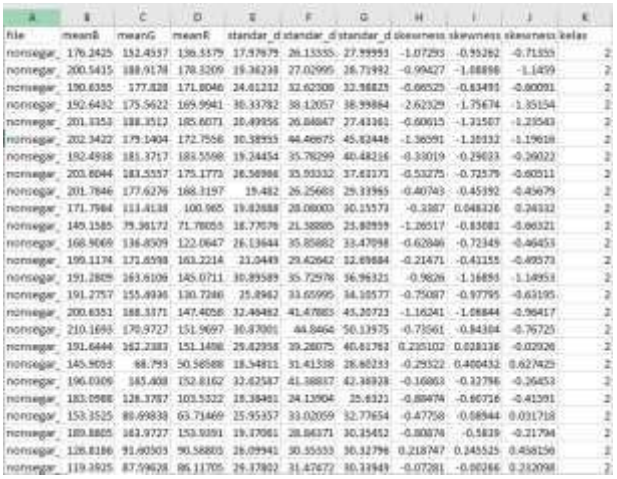


Gambar 1. Alur Klasifikasi

A. Pengumpulan Data

Data penelitian dengan menggunakan objek daging ayam diambil secara langsung di rumah potong ayam menggunakan *smartphone* vivo dan mendapatkan data kelas segar sebanyak 140 sampel dan non-segar sebanyak 60 sampel. Total keseluruhan dataset dari kelas segar dan non-segar sebesar 200 data. Adanya ketidakseimbangan jumlah data antar kelas sangat mempengaruhi terhadap nilai akurasi model yang akan dibangun. Hasil prediksi juga cenderung dominan terhadap kelas segar dibandingkan kelas non-segar. Maka dari itu penelitian ini hanya menggunakan 60 sampel untuk kelas segar dan 60 sampel untuk kelas non-segar. Pelabelan data untuk kelas segar dan non-segar dilakukan secara manual dengan menggunakan bantuan jasa ahli dibidang pemotongan ayam.

Gambar 2. Dataset Gambar



file	meanR	meanG	meanR	standar_d	standar_d	standar_d	skewness	skewness	skewness	kelas
nonsegar_	176.2425	152.4557	136.5579	17.97679	26.13555	27.99993	-1.07295	-0.95262	-0.71359	2
nonsegar_	200.5415	188.9178	178.3209	19.36238	27.02995	28.71992	-0.99427	-1.08898	-1.1499	2
nonsegar_	190.6355	177.828	171.8046	24.61232	32.62308	32.98823	-0.66525	-0.83491	-0.80091	2
nonsegar_	192.6432	175.5622	169.9941	30.33782	38.12057	38.99864	-2.62329	-1.75674	-1.35154	2
nonsegar_	201.3353	188.3513	185.6073	20.49956	26.84847	27.43161	-0.60615	-1.23507	-1.23583	2
nonsegar_	202.3422	179.1404	172.7558	30.38933	44.46875	45.82448	-1.36991	-1.20332	-1.19618	2
nonsegar_	192.4918	181.3717	183.5598	19.24454	35.78299	40.48216	-0.33019	-0.29023	-0.29022	2
nonsegar_	203.8044	183.5557	175.1773	26.56986	35.93332	37.63171	-0.53275	-0.72579	-0.60511	2
nonsegar_	201.7846	177.6276	168.3197	19.482	26.25683	29.33969	-0.40743	-0.45192	-0.45679	2
nonsegar_	171.7984	113.4138	100.965	19.82888	28.08003	30.15573	-0.3387	0.048126	0.24332	2
nonsegar_	145.1585	79.36172	71.78055	18.77579	21.58885	25.80959	-1.26517	-0.83081	-0.66321	1
nonsegar_	168.9069	136.8509	122.0647	26.13644	35.85882	33.47098	-0.62846	-0.72349	-0.46453	2
nonsegar_	199.1174	171.6598	163.2214	21.0449	29.42642	32.69884	-0.21471	-0.42155	-0.49573	2
nonsegar_	191.2809	163.6109	145.0711	30.89589	35.72976	36.96325	-0.9826	-1.16893	-1.14953	2
nonsegar_	191.2757	155.4936	130.7246	25.8942	33.65995	34.10577	-0.75087	-0.97795	-0.63195	2
nonsegar_	200.6351	188.3171	147.4058	32.46482	41.47885	45.20723	-1.16241	-1.08844	-0.96417	2
nonsegar_	210.1693	170.9727	151.9697	30.87001	44.8466	50.13975	-0.73561	-0.84304	-0.76725	2
nonsegar_	191.6444	162.2383	151.1408	29.42958	39.28075	40.61782	0.235102	0.028136	-0.02926	2
nonsegar_	145.9053	88.793	50.58588	18.54811	31.41338	28.80233	-0.29522	0.400432	0.627425	2
nonsegar_	196.0309	165.408	152.8162	32.62587	41.38837	42.38928	-0.16863	-0.32796	-0.26453	2
nonsegar_	183.0988	128.3787	103.5322	18.38481	24.13904	25.8121	-0.89474	-0.60716	-0.41591	2
nonsegar_	153.3525	89.69838	63.71469	25.95357	33.02059	32.77654	-0.47758	-0.08944	0.031718	2
nonsegar_	189.8805	163.9727	153.9391	19.37081	28.84371	30.35452	-0.80878	-0.5819	-0.21794	2
nonsegar_	128.8186	91.60503	90.56803	26.09941	30.35533	30.32796	0.218747	0.245525	0.458156	2
nonsegar_	119.3925	87.59628	86.11705	26.17802	31.47672	30.33949	-0.07281	-0.00266	0.232098	2

Gambar 3. Dataset CSV

B. Preprocessing

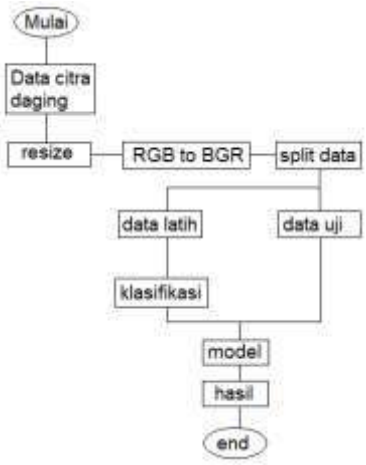
Preprocessing merupakan representasi fitur pembersihan yang dilakukan agar data yang digunakan tidak terdapat *noise* sehingga tidak mempengaruhi hasil pada model yang akan dibangun[14]. Pada penelitian ini tahap *preprocessing* yang dilakukan adalah:

- Melakukan pengubahan ukuran citra menjadi 256x256 pixel.
- Mengubah format citra dari RGB menjadi BGR.
- Melakukan pembagian data latih dan data uji terhadap dataset yang digunakan.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

C. *Perancangan Sistem*

Tahap perancangan sistem dilakukan dengan melakukan pembuatan model klasifikasi menggunakan algoritma Naïve Bayes Classifier, melakukan pengujian terhadap model yang dibangun, dan menampilkan hasil evaluasi performa sistem berdasarkan tingkat akurasi, presisi, dan recall.



Gambar 4. Flowchart sistem

III. HASIL DAN PEMBAHASAN

Berdasarkan pengujian yang telah dilakukan, pengubahan dataset menjadi format csv melewati beberapa perhitungan seperti menghitung nilai rata-rata, standar deviasi, dan skewness sepanjang nilai array pada citra tersebut. Dari perhitungan yang dilakukan menghasilkan 9 *features* yakni: meanB, meanG, meanR, standar_devB, standar_devG, standar_devR, skewness_B, skewness_G, skewness_R.

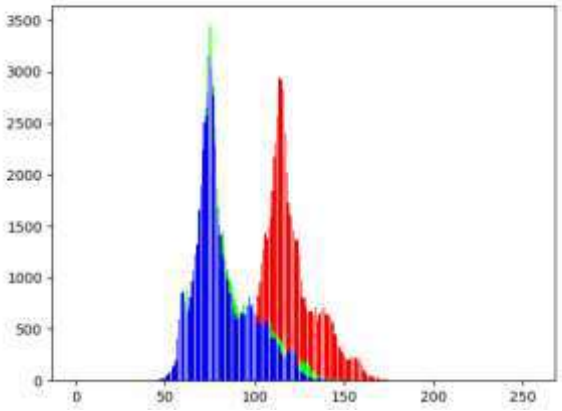
A. *Ekstraksi Fitur Warna*

Ekstraksi fitur warna dilakukan dengan mengubah gambar dari RGB menjadi BGR. Proses ekstraksi warna dan bentuk dinilai cukup baik terhadap hasil penelitian yang telah diuji sebelumnya[15]. Proses *resize* dilakukan sebelum ekstraksi warna agar seluruh gambar memiliki jumlah komponen pixel yang sama. Berikut merupakan perbedaan dari citra RGB dan citra BGR menggunakan sampel citra daging ayam kelas segar. Perubahan warna dari citra RGB dan BGR dilakukan untuk setiap dataset yang diunggah ke dalam sistem klasifikasi.

Gambar 5. Citra RGB

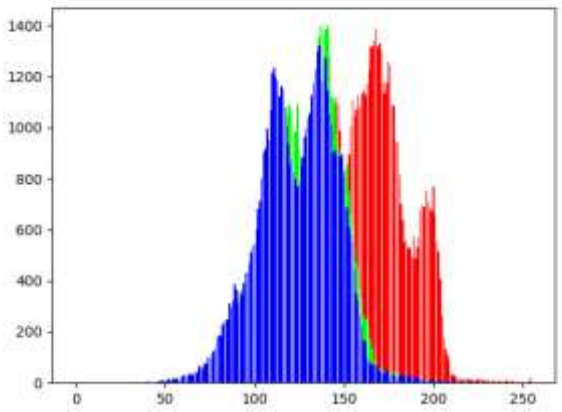
Gambar 6. Citra BGR

Pada Gambar 6 dan Gambar 7, ditampilkan perbandingan histogram dari kelas segar dan non-segar. Penggunaan histogram bertujuan untuk membedakan inputan citra daging ayam dengan non-daging ayam dengan nilai yang digunakan adalah nilai *maximal* dari layer biru. Setelah dilakukan pengujian terhadap 10 sampel kelas segar dan non-segar, didapatkan hasil rata-rata layer biru pada kelas segar dan non-segar tidak kurang dari 1000 dan tidak lebih dari 10000.



Gambar 7. Histogram nonsegar

Dapat dilihat pada Gambar 7 yakni histogram kelas segar memiliki persebaran intensitas warna RGB lebih panjang dibandingkan kelas non-segar. Pada kelas non-segar rentang nilai yang muncul dimulai dari 50 hingga 150, sedangkan kelas segar memiliki rentang nilai yang dimulai dari 50 hingga 200 dengan jumlah yang lebih bervariasi dibandingkan kelas non-segar.



Gambar 8. Histogram segar

B. *Perhitungan Parameter*

Pada penelitian ini perhitungan dilakukan untuk seluruh layer red, green, dan blue. Atribut yang dihasilkan akan dijadikan karakteristik untuk klasifikasi kesegaran daging ayam. Pada tabel di bawah ini dapat dilihat perbedaan yang cukup signifikan terhadap perhitungan ketiga atribut. Pada tabel *mean* dan standar deviasi kelas non-segar memiliki nilai yang cukup tinggi dibandingkan kelas segar. Namun pada

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

perhitungan skewness kelas non-segar memiliki nilai yang rendah dibandingkan kelas segar. Seluruh citra yang ada di dalam dataset akan dilakukan perhitungan dengan tiga parameter yang selanjutnya akan diubah ke dalam bentuk csv. Pada TABEL I dilakukan perhitungan parameter nilai rata-rata untuk 2 sampel yakni dari kelas segar dan non-segar. Pada hasil tersebut dapat diketahui bahwa kelas non-segar mendominasi nilai tertinggi daripada kelas segar terutama pada *layer red*.

TABEL I. Nilai Mean

JENIS DAGING	MEAN		
	Red	***Green***	***Blue***
Segar	153.92881	121.7711	116.9732
Non segar	187.5041	182.1008	175.5018

Perhitungan nilai standar deviasi dapat dilihat pada TABEL II dengan menggunakan sampel citra yang sama dengan perhitungan nilai rata-rata. Dapat dilihat pada tabel tersebut kelas non-segar masih mendominasi nilai tertinggi untuk setiap perhitungan pada layer *red*, *green*, dan *blue*. Hal tersebut menunjukkan bahwa pada kelas non-segar memiliki tingkat variasi yang lebih besar dibandingkan kelas segar terutama pada layer *green* dan *blue*. Perhitungan standar deviasi dilakukan dengan menghitung selisih setiap data dengan nilai rata-rata yang didapatkan seperti pada persamaan 3.

TABEL II. Nilai Standar Deviasi

JENIS DAGING	STANDAR DEVIASI		
	Red	***Green***	***Blue***
Segar	23.4170	27.1168	27.9057
Non segar	23.5158	31.0131	31.7900

Perhitungan nilai skewness juga dilakukan pada 2 sampel citra tersebut dengan mendapatkan hasil citra dengan kelas segar lebih mendominasi nilai tertinggi dibandingkan sampel kelas non-segar. Perhitungan tiap parameter dilakukan untuk setiap dataset citra yang diunggah sebagai karakteristik klasifikasi kelas yang akan dilakukan.

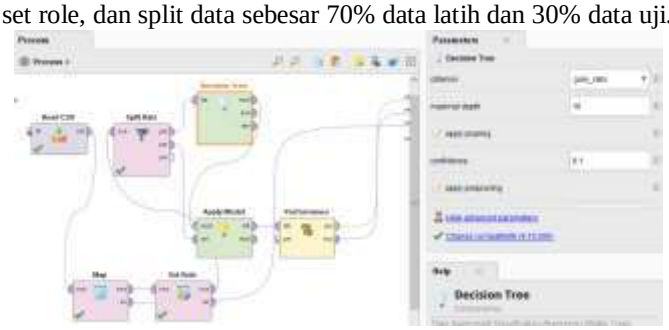
TABEL III. Nilai Skewness

JENIS DAGING	SKEWNESS		
	Red	***Green***	***Blue***
Segar	0.4169	0.5706	0.6471
Non segar	-1.2140	-1.0891	-0.8787

C. Pembuatan Model

1) Decision Tree

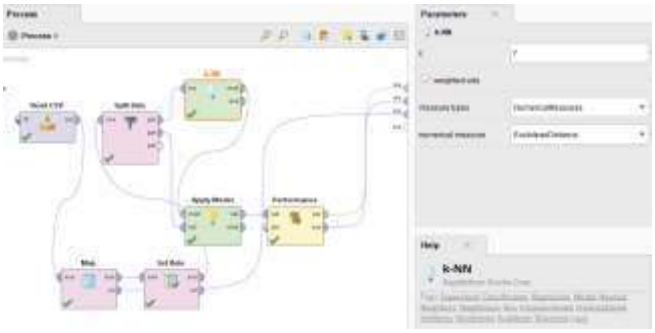
Pembuatan model dengan menggunakan algoritma decision tree dilakukan perubahan format untuk atribut target dari *real* menjadi polynominal. Adapun parameter yang digunakan adalah *criterion, confidence*, dan *minimal gain*. *Criterion* pada algoritma ini menggunakan *gain ratio* dengan nilai *confidence* sebesar 0.1 dan *minimal gain* adalah 0.01. Pembuatan model decision tree dengan *tools* rapidminer mendapatkan nilai akurasi data testing sebesar 63.89%, presisi 72.73%, dan *recall* 44.44%. Proses yang dilakukan menggunakan algoritma decision tree dimulai dari penggunaan operator read csv, map, set role, dan split data sebesar 70% data latih dan 30% data uji.



Gambar 9. Proses Decision Tree

2) k-Nearest Neighbor

Pembuatan model dengan algoritma k-NN mendapatkan akurasi data testing sebesar 61.11%, presisi 62.50%, dan *recall* 55,56% dengan menggunakan nilai k=7 , *measure types* menggunakan *numerical measure*, dan perhitungan jarak menggunakan *Euclidean distance.*



Gambar 10. Proses k-NN

3) Naïve Bayes Classifier

Pembuatan model naïve bayes classifier mendapatkan akurasi data testing sebesar 75%, presisi 79%, dan *recall* 65%. Parameter yang digunakan pada algoritma ini berupa nilai *test_size* sebesar 0.3 dan *random state*=0.

	precision	recall	f1-score	support
1.0	0.79	0.65	0.71	17
2.0	0.73	0.84	0.78	19
accuracy			0.75	36
macro avg	0.76	0.74	0.75	36
weighted avg	0.75	0.75	0.75	36

Gambar 11. Performance Report Naïve Bayes

D. Pengujian Sistem

Berdasarkan hasil pembuatan model dengan algoritma naïve bayes classifier, decision tree, dan k-NN didapatkan hasil bahwa algoritma naïve bayes lebih unggul dibandingkan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

dengan dua algoritma lainnya. Pembagian dataset dari ketiga algoritma tersebut dibagi menjadi 70% data latih dan 30% data uji dari keseluruhan dataset kelas segar dan nonsegar. Hasil pengujian model menggunakan algoritma naïve bayes terhadap data testing didapatkan akurasi 75%, *precision* 79%, dan *recall* 65%. Pengujian menggunakan algoritma decision tree untuk data testing mendapatkan akurasi 63.89%, *precision* 72.73%, dan *recall* 44.44%. Sedangkan algoritma k-NN dengan nilai k=7 mendapatkan akurasi 61.11%, *precision* 62.50%, dan *recall* 55.56%. Hasil confusion matrix untuk model klasifikasi dengan algoritma k-NN didapatkan 22 citra terprediksi benar dan 14 citra terprediksi salah. Berikut merupakan hasil perhitungan FPR dan sensitifitas untuk algoritma k-NN.

FPR (False Positive Rate) = FP / (FP + TN)
= 6 / (6 + 12)
= 0.333 atau 33%

Sensitifitas = TP / (TP + FN)
= 10 / (10 + 8)
= 0.555 atau 55%

	true tidak segar	true segar
pred. tidak segar	12	8
pred. segar	6	10

Gambar 12. Matriks Model k-NN

Pada model klasifikasi decision tree didapatkan sebanyak 23 citra terprediksi benar dan 13 citra terprediksi salah. Hasil dari confusion matrix digunakan untuk menghitung FPR dan sensitifitas model algoritma decision tree.

FPR (False Positive Rate) = FP / (FP + TN)
= 3 / (3 + 15)
= 0.166 atau 16%

Sensitifitas = TP / (TP + FN)
= 8 / (8 + 10)
= 0.444 atau 44%

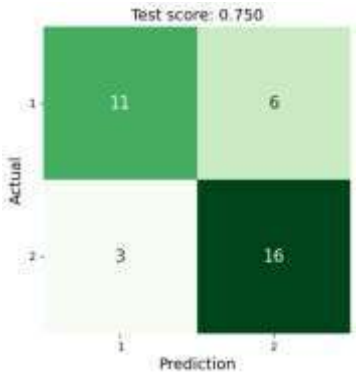
	true tidak segar	true segar
pred. tidak segar	15	10
pred. segar	3	8

Gambar 13. Matriks Model Decision Tree

Sedangkan untuk model klasifikasi menggunakan algoritma naïve bayes classifier mendapatkan 27 citra terprediksi benar dan 9 citra terprediksi salah dari total 36 citra daging ayam. Dilakukan juga perhitungan nilai FPR dan sensitifitas model algoritma naïve bayes sebagai pembanding performa tiap algoritma.

FPR (False Positive Rate) = FP / (FP + TN)
= 6 / (6 + 16)
= 0.272 atau 27%

Sensitifitas = TP / (TP + FN)
= 11 / (11 + 3)
= 0.785 atau 78%



Gambar 14. Matriks Naïve Bayes

Hasil pengujian data testing tersebut menunjukkan bahwa model algoritma naïve bayes classifier memiliki performa yang lebih baik. Dari hasil pengujian tersebut juga dilakukan perhitungan nilai FPR (False Positive Rate) dan sensitifitas seperti yang dapat dilihat pada TABEL IV. Perhitungan dilakukan untuk melihat evaluasi performa dari algoritma k-NN, decision tree, dan naïve bayes classifier.

TABEL IIV. Evaluasi Performa Algoritma

Algoritma	Akurasi	Presisi	Recall	FPR	Sensitifitas
K-NN	61.11%	62.50%	55.56%	33%	55%
Decision Tree	63.89%	72.73%	44.44%	16%	44%
Naïve Bayes	75%	79%	65%	27%	78%

Penelitian ini menggunakan 120 data dengan 60 kelas segar dan 60 kelas non-segar. Dapat dilihat pada tabel iv akurasi tertinggi didapat oleh naïve bayes classifier dengan total citra yang terprediksi benar sebanyak 27 dan terprediksi salah sebanyak 9 citra. Keunggulan naïve bayes dalam mengklasifikasikan citra daging ayam karena penggunaan dataset yang sedikit. Naïve bayes classifier dikenal sebagai algoritma yang cukup sederhana karena bisa mendapatkan hasil yang cukup baik hanya dengan menggunakan data latih yang relative kecil[16]. Naïve bayes melakukan klasifikasi dengan mencari peluang terbesar berdasarkan frekuensi data training yang berakar pada teorema bayes. Pengujian sistem dilakukan dengan memasukkan gambar yang akan diklasifikasi menggunakan algoritma naïve bayes classifier. Percobaan pertama dilakukan dengan objek bukan daging ayam. Pada gambar 17 dan 18 dapat dilihat hasil percobaan dengan dua gambar yang berbeda yakni pada gambar 15 dan 16. Berikut merupakan tampilan *interface* dengan menggunakan *framework* flask.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

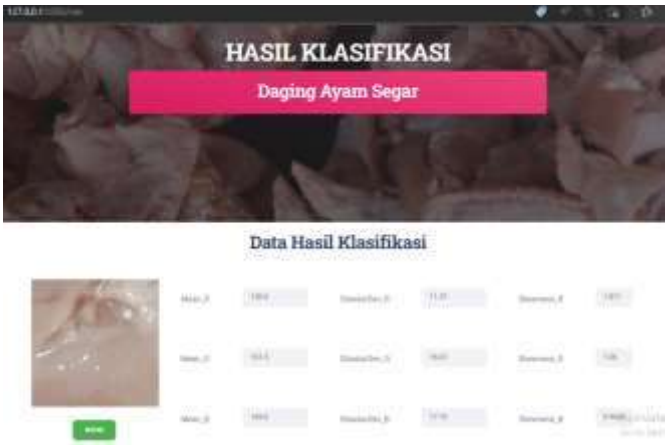
Gambar 15. Input Pertama

Gambar 16. Input Kedua

Berikut merupakan tampilan sistem pengujian klasifikasi jika gambar yang diunggah bukan merupakan citra daging ayam.

Gambar 17. Hasil Percobaan Pertama

Pada Gambar 17 ditampilkan halaman sistem klasifikasi jika gambar yang diunggah adalah citra daging ayam. Sistem akan melakukan klasifikasi menggunakan model yang telah dibangun dan menampilkan nilai dari perhitungan *mean*, standar deviasi, dan *skewness*.



Gambar 18. Hasil Percobaan Kedua

IV. KESIMPULAN

Klasifikasi kesegaran daging ayam dengan algoritma naïve bayes classifier mendapatkan hasil akurasi, presisi, dan *recall* yang tinggi diantara algoritma decision tree dan k-NN. Hasil tersebut menggunakan parameter *test size* sebesar 0.3 dan *random state* 0. *Preprocessing* dilakukan dengan mengubah ukuran citra menjadi 256x256 pixel dan mengubah seluruh dataset dari RGB menjadi BGR yang selanjutnya dilakukan perhitungan histogram *channel* biru untuk membedakan citra daging ayam dengan non-daging ayam. Penelitian yang dilakukan dengan algoritma naïve bayes terhadap 120 dataset mendapatkan hasil yang cukup baik dengan akurasi sebesar 75%, presisi 79%, dan *recall* 65%.

REFERENCES

[1] T. Sajekti, D. Supiyadi, and A. H. Saputro, "PENGOLAHAN DAGING AYAM FROZEN SEBAGAI PENINGKATAN PEMASARAN AYAM POTONG," *Jurnal Abdi Insani*, vol. 9, no. 2, 2022, doi: 10.29303/abdiinsani.v9i2.591.

[2] A. Husein, A. Ramzi, N. I. Muzakki, and R. Hasanah, "Quameaty: Aplikasi Pendeteksi Kualitas Daging Ayam Mentah Berbasis Pengolahan Citra Menggunakan Model InceptionV3," *JTERA (Jurnal Teknologi Rekayasa)*, vol. 7, no. 1, 2022, doi: 10.31544/jtera.v7.i1.2022.107-114.

[3] A. Syahadat, "Kualitas Mikrobiologi Daging Ayam Mati Kemarin 'Tiren' dan Ayam Segar Strain Cobb 500 Ditinjau dari Total Plate Count, Salmonella sp. dan Eschericia coli," *Αγαη*, vol. 8, no. 5. p. 55, 2015. [Online]. Available: http://repository.ub.ac.id/id/eprint/137566/

[4] Alwi Musa Muzaiyin, "PERILAKU PEDAGANG UNGGAS DITINJAU DARI PERSPEKTIF ETIKA BISNIS ISLAM (The Behavior of Poultry Traders Viewed from Islamic Business Ethics Perspective)," *Qawãnïn Journal of Economic Syaria Law*, vol. 5, no. 1, pp. 33–52, Jan. 2021, doi: 10.30762/qawanin.v5i1.2945.

[5] V. P. Bintoro, B. Dwiloka, and D. A. Sofyan, "PERBANDINGAN DAGING AYAM SEGAR DAN DAGING AYAM BANGKA DENGAN MEMAKAI UJI FISIKO KIMIA DAN MIKROBIOLOGI (The Comparison of the Slaughtered and Nonslaughtered Chicken Meat Using Physico-chemical and Microbiological Test)."

[6] H. Zhang, "The Optimality of Naive Bayes Naive Bayes and Augmented Naive Bayes," *Aa*, vol. 1, no. 2, 2004.

[7] M. Nasir, "Klasifikasi Citra Daging Ayam Dengan Menggunakan Metode K -Nearest Neighbor," *E-Jurnal.Pnl.Ac.Id*, vol. 1, no. 1, 2017.

[8] syaeful achmad, ilham fadilah muhammad, muftadi imam, and iskandar dadang, "Klasifikasi Citra Bunga Dahlia Berdasarkan Warna Menggunakan Metode Otsu Thresholding Dan Naïve Bayes," *sains komputer & Informatika*, vol. 6, 2022.

[9] S. N. Hidayah, H. I. Wahyuni, and S. Kismiyati, "Kualitas Kimia Daging Ayam Broiler dengan Suhu Pemeliharaan yang Berbeda," *Jurnal Sains dan Teknologi Peternakan*, vol. 1, no. 1, 2019, doi: 10.31605/jstp.v1i1.443.

[10] R. Indarti, "Pengaruh Pemilihan Daging Ayam terhadap Pembuatan Dim Sum di Restaurant Tang Palace Hotel JW Marriot Surabaya," *Tourism, hospitality and culinary journal*, vol. 2, no. 1, 2018.

[11] Y. Withasari, "Pengaruh Media Big Book Terhadap Kemampuan Mengklasifikasi Pada Anak Usia Dini," *NOURA: Jurnal Kajian Gender*, vol. 3, no. 2, 2019, doi: 10.32923/nou.v3i2.1046.

[12] Annissa Widya Davita, "Mengenal Naive Bayes Sebagai Salah Satu Algoritma Data Science," *DQLAB.COM*, 2022.

[13] C. C. Aggarwal and C. X. Zhai, "A survey of text classification algorithms," in *Mining Text Data*, 2012. doi: 10.1007/978-1-4614-3223-4_6.

[14] J. Chaki and N. Dey, *A Beginner's Guide to Image Preprocessing Techniques*. 2018. doi: 10.1201/9780429441134.

[15] Maulana Fansyuri and O. Hariansyah, "Pengenalan Objek Bunga dengan Ekstraksi Fitur Warna dan Bentuk Menggunakan Metode Morfologi dan Naïve Bayes," *Jurnal Sistem dan Informatika (JSI)*, vol. 15, no. 1, 2020, doi: 10.30864/jsi.v15i1.338.

[16] F. Howedi and M. Mohd, "Text Classification for Authorship Attribution Using Naive Bayes Classifier with Limited Training Data," *Computer Engineering and Intelligent Systems*, vol. 5, no. 4, 2014.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1740, Copyright ©2023
Submitted : May 22, 2023, Revised : June 16, 2023, Accepted : August 28, 2023, Published : November 6, 2023

Analisis Kinerja Model Klasifikasi Pada Multiclass *Facial Expression Recognition* Berbasis Fitur *Eigenface*

Syefrida Yulina[1]*, Heni Rachmawati[2]
Program Studi Sistem Informasi, Jurusan Teknologi Informasi, Politeknik Caltex Riau[1], [2]
Jl. Umbansari, Pekanbaru – Riau, 28265, Indonesia
syefrida@pcr.ac.id[1] , heni@pcr.ac.id[2]

Abstract*— Facial Expression Recognition (FER) is currently widely explored by researchers in the field of Computer Vision. The application of Machine Learning and Deep Learning methods is useful in developing an intelligent system that is accurate in recognizing facial expressions such as emotions. This is inseparable from the type of dataset and classification method used which certainly affects the desired results. To choose the right method, it is necessary to compare the performance of these methods. This study focuses on comparing the performance results of four classification methods namely, Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes Classifier (NBC) on a multiclass dataset for seven classes of facial emotion labels based on Eigenface feature selection uses the Personal Component Analysis (PCA) algorithm. The test parameters used to perform method comparisons are accuracy, recall, precision, f1-score, as well as the Receiving Operating Characteristic (ROC) and Area Under Curve (AUC) curves. The results of the analysis state that the SVM method has the highest accuracy value, while other methods show varying performance based on recall, precision, f1-score, and ROC and AUC analysis. This research was conducted on the FER 2013 dataset which showed that the classification method tested had quite good performance according to the test parameters.

***Keywords*— Facial expression, FER dataset, Method comparison, Performance analysis**

Abstrak*— *Facial Expression Recognition* (FER) saat ini banyak dieksplorasi oleh para peneliti dalam bidang *Computer Vision*. Penerapan metode *Machine Learning* dan *Deep Learning*, bermanfaat dalam mengembangkan sebuah sistem cerdas yang akurat dalam mengenali ekspresi wajah seperti emosi. Hal ini tidak terlepas dari jenis dataset dan metode klasifikasi yang digunakan yang pastinya mempengaruhi hasil yang diinginkan. Untuk pemilihan metode yang tepat, maka diperlukan sebuah perbandingan kinerja terhadap metode-metode tersebut. Penelitian ini berfokus pada perbandingan hasil kinerja empat metode klasifikasi yaitu, *Convolutional Neural Network* (CNN), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), *Naïve Bayes Classifier* (NBC) pada dataset multiclass untuk tujuh kelas label emosi wajah berdasarkan seleksi fitur *Eigenface* menggunakan algoritma *Personal Component Analysis* (PCA). Parameter uji yang digunakan untuk melakukan perbandingan metode adalah seperti nilai akurasi, *recall, precision, f1-score*, serta kurva *Receiving Operating Characteristic* (ROC) dan *Area Under Curve* (AUC). Hasil analisis menyatakan bahwa metode SVM memiliki nilai akurasi tertinggi, sedangkan metode lainnya menunjukan kinerja bervariasi berdasarkan nilai *recall, precision, f1-score*, serta analisis ROC dan AUC. Penelitian ini dilakukan pada dataset FER 2013 yang menunjukan metode klasifikasi yang diujikan memperoleh kinerja cukup baik sesuai dengan parameter uji.

***Kata Kunci*— Analisis Performa, Dataset FER, Ekspresi Wajah, Perbandingan Metode**

I. PENDAHULUAN

Pengenalan emosi pada wajah menggunakan teknologi *Facial Expression Recognition* (FER) sedang aktif dieksplorasi dalam penelitian *Computer Vision*. Banyaknya metode *Machine Learning* dan *Deep Learning*, berpotensi untuk membangun sebuah sistem cerdas yang akurat dalam mengenali emosi [1]. Sebagai bagian penting dari Computer Vision, FER semakin menarik perhatian, sehingga penelitian dibidang ini mengadopsi berbagai metode untuk meningkatkan FER tersebut [2-3]. Saat ini terdapat penelitian tentang penerapan serta perbandingan metode klasifikasi pada berbagai jenis dataset [4-5] terutama pada klasifikasi emosi wajah [1][6-8]. Dengan hadirnya *library-library* yang telah disediakan oleh bahasa pemrograman tertentu seperti Python, sehingga memungkinkan para peneliti dengan mudah dan cepat untuk melakukan klasifikasi terhadap citra wajah. Kinerja berbagai metode klasifikasi perlu diuji pada sebuah dataset agar menjadi salah satu cara dalam penentuan metode yang tepat. Masing-masing dari metode tersebut memiliki kelebihan dan kekurangan, hal ini dapat dilihat pada keunggulan dalam mengolah objek dataset. Gap analisis pada penelitian ini dapat dilihat dari Tabel I.

Metode klasifikasi emosi wajah yang umumnya banyak diterapkan adalah seperti: *Convolutional Neural Network* (CNN), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), *Naïve Bayes Classifier* (NBC). Metode tersebut merupakan algoritma klasifikasi citra yang efektif untuk masalah dua kelas atau multi kelas [9-10]. Penggunaan sejumlah dataset sangat berpengaruh terhadap akurasi sebuah metode klasifikasi. Hal ini dapat menjadikan acuan bagi peneliti untuk memperoleh nilai akurasi yang bervariasi serta karakteristik yang dimiliki oleh sebuah metode [11].

Pada penelitian ini melakukan analisis kinerja dari metode-metode klasifikasi tersebut diatas dalam mengelola dataset yang memiliki lebih dari dua label (*multiclass*) emosi wajah yaitu: bahagia (*happiness*), tidak bahagia (*sadness*), terkejut (*surprise*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

netral. Dataset yang akan digunakan untuk penelitian ini adalah FER-2013 yang memiliki banyak data citra wajah berisikan tujuh emosi wajah. Tahapan pengenalan emosi wajah (FER) yang dilakukan yaitu, seperti: mendeteksi dan mengakuisi bagian wajah, *pre-processing* data, ekstraksi fitur wajah berupa Eigeface menggunakan algoritma Personal Component Analysis (PCA), dan mengklasifikasikan emosi pada wajah tersebut. Kinerja setiap metode diukur melalui pengujian nilai akurasi, *recall*, *precision*, *f1-score*, serta *Receiving Operating Characteristic* (ROC) dan *Area Under Curve* (AUC) menggunakan dataset yang sama. Dengan adanya perbandingan beberapa metode klasifikasi yang berbeda, maka dapat dijadikan sebagai acuan dalam menentukan metode yang terbaik.

TABEL I. TABEL GAP ANALISIS PENELITIAN

Tahun	Peneliti	Metode	Dataset	Ekstraksi Fitur	Performa				ROC	AUC
					Akurasi	Presisi	Recall	F-Measure		
2023	Penelitian Sekarang	SVM KNN NBC CNN	FER 2013	Eigenface	✓	✓	✓	✓	✓	✓
2021	Song [4]	CNN	FER 2013	ROI/Gabor	✓	-	-	-	-	-
2019	Canedo & Neves [1]	SVM KNN CNN	CK+/JAFFE/FER 2013	LBP/ROI/Gabor	✓	-	-	-	-	-
2019	Utami et al. [6]	CNN	CK+	LBP/Gabor	✓	-	-	-	-	-
2017	Tarnowski et al. [5]	KNN/MLP - NN	KDEF	AU	✓	-	-	-	-	-

II. PENELITIAN TERKAIT

A. Penelitian Terdahulu

Saat ini terdapat banyak penelitian mengenai penerapan serta perbandingan metode klasifikasi emosi wajah serta penerapannya pada berbagai jenis dataset. Pada tahun 2021, [8] menerapkan metode CNN dalam mengekstraksi fitur penting pada ekspresi wajah serta meningkatkan akurasi pada pengenalan emosi. Tahun 2020, [6] melakukan perbandingan performa beberapa metode klasifikasi pada dataset citra busur panah, dengan hasil penelitian menyatakan bahwa metode SVM memiliki performa yang lebih baik dibandingkan metode lainnya. Kemudian selanjutnya [1] dan [12]ditahun yang sama melakukan studi analisis untuk pengenalan ekpresi wajah dalam berdasarkan perspektif dataset, metode klasifikasi, dan variasi pose wajah, lalu [13] menerapkan metode KNN dan MLP dalam pengenalan emosi wajah untuk model wajah 3 dimensi.

Menurut [9] *face*/wajah merupakan kunci pembeda yang digunakan secara luas sebagai identitas personal yang memiliki keunikan atau ciri-ciri khusus. Bidang kajian *face detection* khususnya *Facial Expression Recognition* telah menjadi perhatian yang dipertimbangkan di dalam pengembangan dan peningkatan interaksi antara manusia dengan mesin atau manusia dengan manusia, dengan pertimbangan dan alasan kealamian serta efisiensi kinerja di dalam berkomunikasi melalui sebuah ekspresi wajah. *Facial Expression Recognition* secara esensial melokasikan dan mengekstraksikan suatu bagian wajah dari latarbelakangnya yang kemudian melakukan pengenalan suatu ekspresi [2][12][14]. Hal ini tampak suatu permasalahan atau tugas yang mudah, tetapi memerlukan suatu teknik/metode dalam penyelesaiannya. Wajah manusia merupakan suatu objek yang dinamis dan mempunyai derajat kevariasian yang sangat tinggi, sehingga *Facial Expression Recognition* menjadi suatu permasalahan yang sangat rumit di dalam *Computer Vision* [15]. Sebuah citra wajah manusia dapat dijadikan sebagai sampel pada suatu kajian dalam pengklasifikasian emosi wajah manusia.

Facial Expression Recognition terdiri dari beberapa langkah, yaitu: *image acquisition*, *pre-processing*, *feature extraction*, *classification/regression* [1][5]. Berbagai metode/teknik/pendekatan telah banyak dibahas pada suatu kajian ilmiah tentang face detection meliputi: segmentasi warna, citra wajah, eigenface, jaringan syaraf tiruan, dll. Pada *Facial Expression Recognition*, algoritma klasifikasi digunakan untuk memprediksi/mengklasifikasikan label emosi tertentu dengan inputan berupa citra wajah.

B. Metode Klasifikasi

[9-10] mendeskripsikan algoritma klasifikasi *K-Nearest Neighbor, Naïve Bayes Classifier, Support Vector Machine, dan Convolutional Neural Network* merupakan algoritma klasifikasi citra yang efektif untuk masalah dua kelas atau multi kelas.

K-Nearest Neighbor (KNN) adalah sebuah algoritma pembelajaran berbasis *instance* yang menggunakan teknik non-parametrik saat melakukan klasifikasi atau regresi. KNN mencari nilai vektor fitur k dengan sifat termirip, lalu mengelompokkan data baru ke kelompok vektor tersebut. Data pelatihan terdiri dari vektor dalam ruang fitur multidimensi, masing-masing dengan label kelas. Fase pelatihan algoritma hanya terdiri dari penyimpanan vektor fitur ini dan kelas milik mereka. Pada langkah klasifikasi, fitur input atau kumpulan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

fitur diprediksi dengan menetapkan kelas yang memiliki fitur terdekat dengan input. Metrik jarak yang umum digunakan untuk menghitung fitur mana yang lebih dekat ke input adalah jarak Euclidean (ED) dan jarak Hamming. Kekuatan metode ini terletak pada implementasi yang sederhana dan pada langkah pelatihan yang cepat.

Naïve Bayes Classifier adalah pengklasifikasi probabilistic berdasarkan penerapan teorema Bayes untuk probabilitas bersyarat. Asusmsinya adalah bahwa semua fitur independent dan tidak terkait satu sama lain (*naïve*). Pengklasifikasi dibangun dengan mengalikan probabilitas bersyarat individu dari setiap fitur untuk mendapatkan peluang total suatu kelas. Kemudian kelas dengan tertinggi probabilitas dipilih.

Support Vector Machine (SVM) adalah algoritma *Machine Learning* yang banyak digunakan untuk masalah klasifikasi/regresi. Model SVM adalah representasi fitur dalam ruang, dipetakan sehingga fitur milik setiap kelas dibagi dengan celah yang jelas yang selebar mungkin. Fitur input kemudian dipetakan ke dalam ruang yang sama dan diprediksi menjadi milik kelas berdasarkan sisi celah mana mereka berada. Fase pelatihan membuat peta ini yang digunakan setelah untuk prediksi.

Convolutional Neural Network (CNN) merupakan salah satu jenis jaringan syaraf tiruan yang banyak digunakan dalam *Computer Vision (Deep Learning)*. CNN memiliki kemampuan untuk mendeteksi dan mengidentifikasi pola dasar yang terlalu rumit. Sebuah citra inputan diproses melalui beberapa lapisan tersembunyi CNN yang akan diuraikan menjadi fitur-fitur. Fitur tersebut kemudian digunakan untuk klasifikasi, umumnya melalui fungsi Softmax yang mengambil nilai probabilitas tertinggi dari distribusi probabilitas kelas sebagai kelas yang diprediksi.

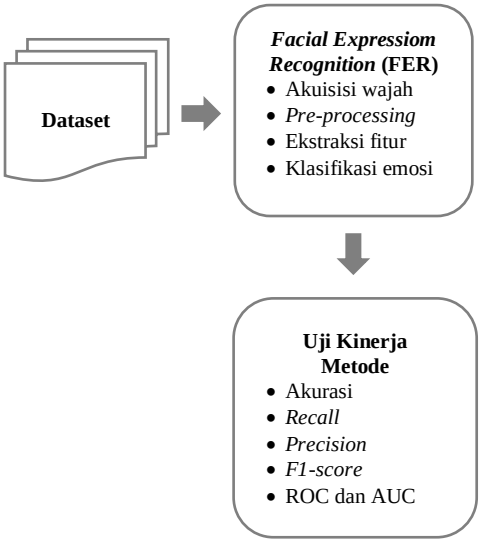
C. Python dan Library Computer Vision

Python merupakan salah satu Bahasa pemrograman yang dapat digunakan untuk berbagai masalah dalam *Artificial Inteligent* dan *Machine Learning*. Python banyak diterapkan pada *Computer Vision* (CV), yang merupakan bagian dari *Artificial Intelligent* terkait aplikasi-aplikasi CV. Python menyediakan *library-library* yang memudahkan para *developer* untuk mengembangkan aplikasi CV, seperti OpenCV, Keras, Matplotlib, Scikit-Image, dll.

OpenCV (*Open-Source Computer Vision Library*) adalah library perangkat lunak Machine Learning dan Computer Vision yang bersifat open-source [16]. OpenCV diciptakan untuk menyediakan infrastruktur umum pada aplikasi *Computer Vision* dan untuk mempercepat penggunaan persepsi mesin dalam produk komersial. *Library* ini memiliki lebih dari 2500 algoritma yang dioptimalkan. Algoritma pada OpenCV dapat digunakan untuk mendeteksi wajah, mengidentifikasi objek, klasifikasi, dll. OpenCV-Python memanfaatkan Numpy, yang merupakan perpustakaan yang sangat dioptimalkan untuk operasi numerik dengan sintaks gaya MATLAB. Semua struktur array OpenCV dikonversi ke dan dari array Numpy. Ini juga memudahkan integrasi dengan library lain yang menggunakan Numpy seperti SciPy dan Matplotlib.

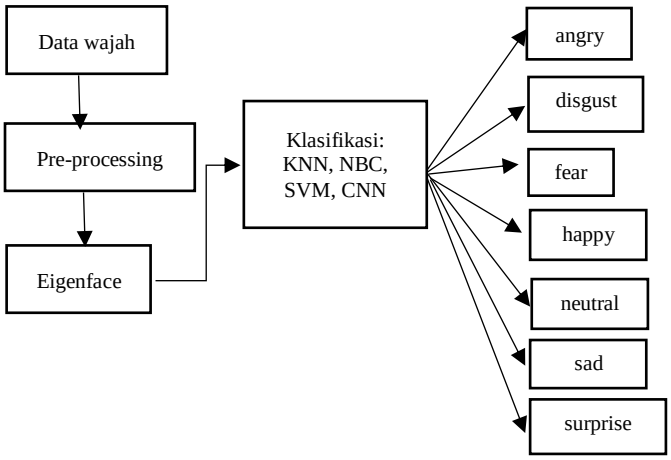
III. METODE PENELITIAN

Metode penelitian terdapat pada Gambar 1. Metode penelitian mencakup tiga tahapan, yaitu dataset citra wajah, *Facial Expression Recognition*, dan uji kinerja metode klasifikasi.



Gambar 1. Metodologi Penelitian

Tahapan *Facial Expression Recognition* (FER) digunakan untuk proses klasifikasi emosi yang dimulai dari proses deteksi dan akuisisi wajah, *pre-processing* data, ekstraksi fitur, kemudian klasifikasi emosi pada wajah menggunakan metode tertentu (Gambar 2). Pada penelitian ini metode yang diterapkan adalah algoritma *k-Nearest Neighbor* (KNN), *Naïve Bayes Classifier* (NBC), *Support Vector Machine* (SVM), dan *Convolutional Neural Network* (CNN). Class yang diklasifikasikan yaitu *angry*, *disgust*, *fear*, *happy*, *neutral*, *sad* dan *surprise*.



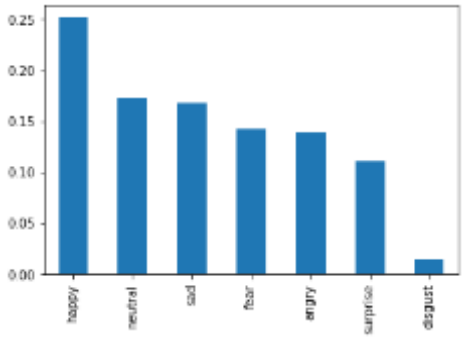
Gambar 2. Tahapan FER

Tahapan uji kinerja setiap metode klasifikasi diukur melalui pengujian nilai akurasi, *recall*, *precision*, *f1-score*, serta

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

Receiving Operating Characteristic (ROC) dan *Area Under Curve* (AUC) menggunakan dataset yang sama.

Dataset yang digunakan untuk klasifikasi adalah data FER 2013 yang disediakan oleh Kaggle, dimana data ini terdiri dari berbagai citra ekspresi wajah yang memiliki tujuh label emosi wajah seperti Gambar 3.



Gambar 3. Distribusi Data untuk Emosi Wajah

Pada Gambar 3 setiap emosi wajah memiliki label *multiclass* seperti emosi bahagia (*happiness*), tidak bahagia (*sadness*), terkejut (*surprise*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan netral. Jumlah total ekspresi wajah pada dataset ini sejumlah 28709 wajah dengan distribusi tiap emosi.

IV. HASIL DAN PEMBAHASAN

Tahap pengujian dilakukan pada dataset FER 2013 dengan distribusi data dapat dilihat pada Gambar 3. Data ini telah melewati tahapan akuisisi wajah, *pre-processing*, seleksi fitur wajah, dan klasifikasi ekspresi wajah dengan metode KNN, NBC, SVM dan CNN.

A. Hasil

1) Akuisisi Wajah

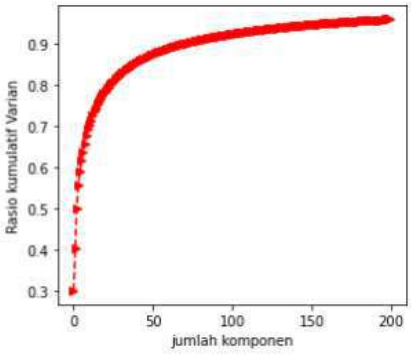
Proses deteksi wajah adalah tahapan awal agar proses klasifikasi dalam dilakukan. Untuk penelitian ini, algoritma yang diterapkan pada proses akuisisi wajah adalah Haar Cascade Clasifier.

2) Pre-processing Data

Pada tahapan *pre-processing* dilakukan normalisasi data citra wajah untuk menghilangkan serta mengurangi redudansi data. Normalisasi memetakan derajat keabuan citra berdimensi n dengan rentang nilai [min, maks] ke citra dengan rentang nilai [minBaru, maksBaru]. Hasil dari normalisasi ini berupa data array dengan ekstensi file .npz. Setelah data dinormalisasi maka kemudian dilanjut proses reduksi dimensi data menggunakan *Principal Component Analysis* (PCA).

3) Ekstraksi Fitur

PCA mengekstraksi fitur wajah yang optimal dengan membentuk suatu arah menurut varian maksimumnya. Algoritma yang digunakan pada PCA ini adalah *Eigenface*, dimana komponen yang berkontribusi dipertahankan. Pemilihan fitur dilakukan sebelum mengekstraksi fitur menggunaan PCA. Proses pemilihan fitur dapat dilihat pada Gambar 4.



Gambar 4. Seleksi Fitur Wajah Menggunakan *Eigen Face*

Gambar 4 memperlihatkan jumlah komponen yang digunakan pada penelitian ini adalah sebanyak **50 fitur**. Pada citra wajah dapat direpresentasikan sebagai sebuah vektor yang terdiri dari panjang (*w*) dan lebar (*h*) piksel. Jumlah komponen dari vektor tersebut adalah *w* * *h*. Sementara varian mengukur penyebaran nilai komponen yang ada didalam himpunan. Analisis pemilihan fitur berdasarkan jumlah komponen dan jumlah varian tiap data lalu mencari nilai dan vektor eigen matrik kovarians nya. Matriks kovarians ini mengukur kekuatan hubungan antar komponen satu sama lain pada himpunan data.

Pemilihan fitur ini untuk memilih fitur yang akan banyak berpengaruh terhadap nilai n fitur yang ada. Fitur menjadi parameter pembeda atau karakteristik dari wajah. Dataset yang dihasilkan setelah menerapkan PCA adalah sebuah citra *eigenface* dapat dilihat pada Gambar 5.

Gambar 5. Sampel Citra Wajah *Eigenface*

4) Klasifikasi Emosi

Tahapan selanjutnya adalah melakukan proses klasifikasi pada ekspresi wajah. Pada setiap metode klasifikasi, dilakukan perbandingan uji kinerja untuk nilai akurasi, *recall*, *precision*, *f1-score*, dan ROC AUC.

a) Akurasi Metode Klasifikasi

Akurasi setiap metode klasifikasi dilakukan pada dataset

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

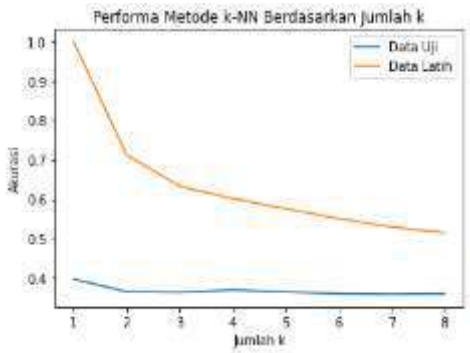
yang sama dengan perbandingan untuk data latih dan data uji adalah 80:20 persen. Rasio perbandingan ini umum digunakan untuk mengetahui kesalahan minimum yang terjadi seperti kondisi *overfitting* maupun *underfitting pada* validasi metode. Hasil akurasi dari setiap metode klasifikasi dapat dilihat pada Tabel II.

TABEL II. AKURASI METODE KLASIFIKASI

Metode	Skor Data Latih (%)	Akurasi Data Uji (%)
KNN	64	37
NBC	34	33
SVM	52	43
CNN	69	62

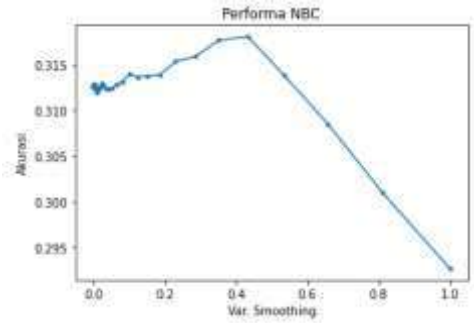
Pada metode KNN didapatkan akurasi data latih sebesar 64% dan data uji sebesar 37%. Pada metode NBC mendapatkan akurasi 34% pada data latih dan 33% pada data uji. Akurasi metode SVM menghasilkan 52% untuk data latih dan 43% untuk data uji. Sedangkan akurasi metode CNN didapatkan 69% untuk data latih dan 62% untuk data uji.

Cross validation dilakukan untuk meningkatkan hasil akurasi pada setiap metode. C*ross validation* yang merupakan evaluasi kinerja metode dimana data dipisahkan menjadi dua subset yaitu data latih dan data uji.



Gambar 6. *Cross Validation* Metode KNN

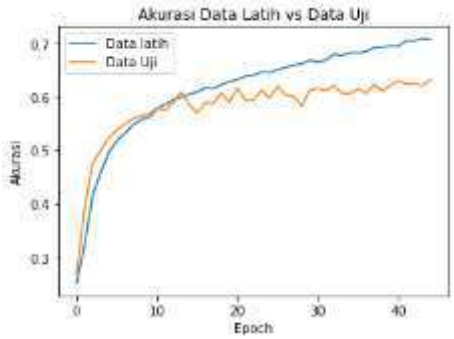
Pada Gambar 6 memperlihatkan hasil akurasi metode KNN pada nilai k=1 hingga k=jumlah data, nilai akurasi terbaik didapatkan pada nilai k=1 dengan akurasi 39% pada data uji.



Gambar 7. Distribusi Variabel *Smoothing* Metode NBC

Pada Gambar 7 menunjukan hasil klasifikasi metode NBC, dimana nilai akurasi tertinggi adalah 33% pada variabel *smoothing* 0,4. Variabel smoothing digunakan untuk meningkatkan hasil klasifikasi pada NBC sehingga nilai yang dihasilkan lebih besar.

Hasil akurasi pada metode SVM setelah dilakukan *hyper* parameter adalah 99% dengan parameter terbaik adalah C bernilai 1, gamma bernilai 0.1 dan kernel bertipe poly. Uji parameter ini dilakukan dengan memberikan beberapa nilai pada nilai C, nilai gamma dan tipe kernel. C merupakan parameter regularisasi yang bernilai positif, rentang nilai C yang digunakan adalah 1 dan 10. Gamma merupakan nilai coefisien Kernel, pada pengujian ini nilai gamma yang dipakai adalah 0.1. Tipe kernel yang diujikan adalah rbf dan poly.



Gambar 8. Akurasi CNN pada nilai Epoch

Pada Gambar 8 memperlihatkan hasil akurasi pada metode CNN dengan nilai epoch hingga 45 untuk setiap data latih dan data uji. Jumlah layer yang digunakan terdiri dari 32, 64, 128, 512 filter, menggunakan *max pooling* dalam mencari nilai maksimum untuk setiap dimensi vektor. Setelah tahapan convalotion dan pooling, maka menghasilkan satu vektor yang kemudian dilewatkan pada *multilayer perceptron* (*fully connected*) untuk melakukan klasifikasi. Akurasi tertinggi metode CNN adalah 62%.

b) Precision, Recall, F1-Score

Detail hasil *performa precision, recall*, dan *f1-score* dapat dilihat pada Tabel III. Dalam menguji kinerja metode klasifikasi, nilai *precision* menggambarkan ukuran ketepatan pengklasifikasian untuk setiap kelas, sedangkan *recall* merupakan ukuran kelengkapan metode klasifikasi, dan *f1-score* menunjukan perbandingan rata-rata *precision* dan *recall* yang dibobotkan.

TABEL III. UJI KINERJA METODE KLASIFIKASI

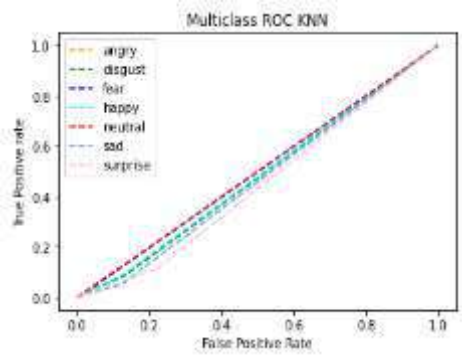
Metode	Variabel	*Precision* (%)	*Recall* (%)	*F1-Score* (%)
KNN	*angry*	46	45	45
	disgust	34	31	32
	fear	36	41	38
	happy	38	35	36
	neutral	32	31	32
	sad	56	58	57
	surprise	32	38	35
	macro avg	39	40	39
	weighted avg	39	39	39
NBC	*angry*	38	66	48
	disgust	22	9	13
	fear	5	2	3

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

	happy	20	14	17
	neutral	27	24	25
	sad	45	40	42
	surprise	30	27	28
	macro avg	27	26	25
	weighted avg	30	33	30
SVM	*angry*	45	75	56
	disgust	36	16	22
	fear	100	3	6
	happy	36	18	24
	neutral	34	36	35
	sad	58	48	53
	surprise	37	38	38
	macro avg	49	33	33
	weighted avg	41	42	39
CNN	*angry*	16	16	16
	disgust	0	0	0
	fear	14	9	11
	happy	27	24	25
	neutral	18	22	20
	sad	17	22	19
	surprise	13	13	13
	macro avg	15	15	15
	weighted avg	18	18	18

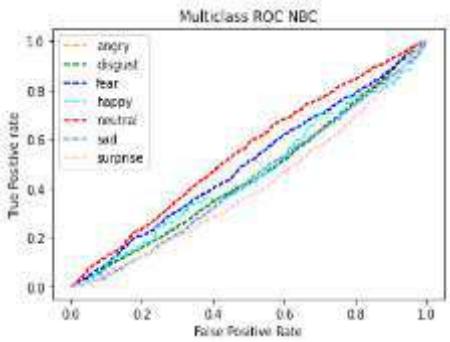
c) ROC dan AUC

Hasil visualisasi kurva ROC dan nilai AUC untuk metode KNN, NBC, SVM dan CNN dapat dilihat pada Gambar 9.a hingga 9.d. *Receiving Operation Characteristic* (ROC) merupakan kurva probabilitas sedangkan *Area Under Curve* (AUC) mewakili derajat/ukuran keterpisahan. Ini memberitahukan seberapa besar metode mampu membedakan antar kelas.



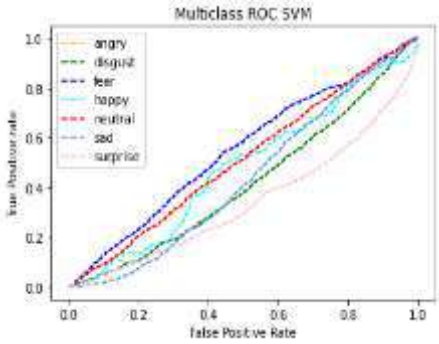
Nilai AUC KNN = 44%

Gambar 9.a. Kurva ROC Metode KNN



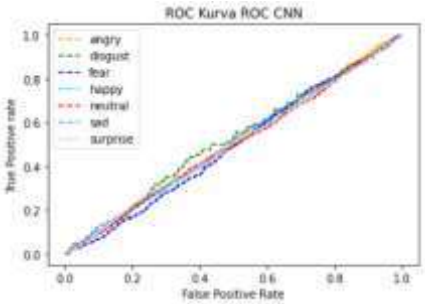
Nilai AUC NBC = 40%

Gambar 9.b. Kurva ROC Metode NBC



Nilai AUC SVM = 34%

Gambar 9.c. Kurva ROC Metode SVM



Nilai AUC CNN = 50%

Gambar 9.d. Kurva ROC Metode CNN

B. Pembahasan

Akurasi pada setiap metode memperoleh nilai yang bervariasi. Hasil dari akurasi didapatkan dari nilai akurasi data training dan data testing dari dataset FER-2013. Analisis berdasarkan nilai akurasi ini adalah terdapatnya kondisi overfitting pada metode KNN yang cukup signifikan dimana akurasi data training lebih besar daripada akurasi data testing. Sementara untuk metode NBC, SVM, dan CNN tidak memiliki perbedaan akurasi yang signifikan. Uji validasi akurasi terbaik dilakukan *cross validation* untuk setiap metode. Maka didapatkan nilai akurasi terbaik untuk KNN = 39% dan SVM = 99%, untuk NBC dan CNN tidak memiliki perbedaan akurasi setelah *cross validation*.

Berdasarkan hasil yang diperoleh dari nilai *precision*, *recall*, dan *f1-score* untuk setiap klasifikasi emosi menghasilkan nilai *macro avg* dan *weighted avg*. Penelitian ini menghasilkan nilai yang berbeda pada setiap metode klasifikasi untuk *precision*, *recall*, dan *f1-score*. Nilai macro avg dan weighted avg juga dihitung untuk tiap metode klasifikasi.

Berdasarkan kurva ROC yang dihasilkan oleh metode KNN untuk klasifikasi ketujuh emosi wajah, memperlihatkan bahwa semua nilai berada dibawah baseline yang mendekati keadaan *false positive rate* dengan nilai AUC = 44%, sedangkan metode NBC, SVM, CNN menunjukan kurva ROC yang bervariasi untuk setiap emosi yang diklasifikasikan, dimana nilai AUC NBC = 40%, AUC SVM = 34%, dan AUC CNN = 50%.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

V. KESIMPULAN

Pada penelitian ini membandingkan metode klasifikasi emosi wajah: KNN, NBC, SVM, dan CNN pada dataset FER-2013 telah melewati tahapan preprocessing, ekstraksi fitur dan klasifikasi. Penelitian ini juga menerapkan algoritma PCA dan Eigenface dalam menentukan fitur wajah. Analisis kinerja metode-metode klasifikasi berdasarkan nilai akurasi, *precision*, *recall*, *f1-score, dan AUC.* Akurasi tertinggi dihasilkan oleh metode SVM sebesar 99% setelah dilakukan cross *validation* dalam mengklasifikasikan emosi seperti bahagia (*happiness*), tidak bahagia (*sadness*), terkejut (*surprise*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan netral. Masing-masing kinerja metode klasifikasi memiliki karakteristik nilai optimal yang berbeda-beda. *Cross validation* dilakukan untuk menganalisis setiap metode agar dapat meningkatkan akurasi terhadap klasifikasi dengan membagi data menjadi data latih dan data uji. ROC dan AUC dilakukan untuk mengukur kinerja metode klasifikasi pada berbagai ambang batas (*threshold*), semakin tinggi AUC, semakin baik model dalam memprediksi. Nilai AUC tertinggi dihasilkan oleh metode CNN sebesar 50%, kemudian metode KNN sebesar 44%, metode NBC sebesar 40% dan nilai AUC terendah dihasilkan oleh metode SVM sebesar 34%. Meskipun hasil analisis FER untuk citra wajah telah diperoleh berdasarkan kinerja masing-masing metode, penelitian selanjutnya diharapkan dapat menggabungkan dengan metode lainnya yang saling terintegrasi sehingga dapat memberikan informasi tambahan tentang kehandalan metode. Kemudian untuk kompresi data patut dipertimbangkan agar informasi tentang fitur wajah yang tidak relevan dapat dikurangi agar memperoleh akurasi yang optimal.

DAFTAR PUSTAKA

[1] D. Canedo and A. J. R. Neves, "Facial Expression Recognition Using Computer Vision: A Systematic Review," *Appl. Sci.*, vol. 9, no. 21, p. 4678, Nov. 2019, doi: 10.3390/app9214678.

[2] H. Y. Cao and C. Qi, "Facial Expression Study Based on 3D Facial Emotion Recognition," in *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)*, Dec. 2021, pp. 375–381. doi: 10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00067.

[3] Y. Huang, F. Chen, S. Lv, and X. Wang, "Facial Expression Recognition: A Survey," *Symmetry (Basel).*, vol. 11, no. 10, p. 1189, Sep. 2019, doi: 10.3390/sym11101189.

[4] H. Azis, F. Tangguh Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," *Techno.Com*, vol. 19, no. 3, pp. 286–294, 2020, doi: 10.33633/tc.v19i3.3646.

[5] V. Cherepanova, S. Reich, S. Dooley, H. Souri, M. Goldblum, and T. Goldstein, "A Deep Dive into Dataset Imbalance and Bias in Face Identification," Mar. 2022, [Online]. Available: http://arxiv.org/abs/2203.08235

[6] Z. Song, "Facial Expression Emotion Recognition Model Integrating Philosophy and Machine Learning Theory," *Front. Psychol.*, vol. 12, no. September, Sep. 2021, doi: 10.3389/fpsyg.2021.759485.

[7] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Emotion recognition using facial expressions Emotion recognition using facial expressions," *Procedia Comput. Sci.*, vol. 108, pp. 1175–1184, 2017, doi: 10.1016/j.procs.2017.05.025.

[8] P. Utami, R. Hartanto, and I. Soesanti, "A Study on Facial Expression Recognition in Assessing Teaching Skills: Datasets and Methods," *Procedia Comput. Sci.*, vol. 161, pp. 544–552, 2019, doi: 10.1016/j.procs.2019.11.154.

[9] J. Erik Solem, *Programming Computer Vision with Python*. O'Reilly, 2012.

[10] J. Wira Gotama Putra, "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," 2020.

[11] I. M. Revina and W. R. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 6, pp. 619–628, Jul. 2021, doi: 10.1016/j.jksuci.2018.09.002.

[12] S. Bayrakdar, D. Akgün, and İ. Yücedağ, "A Survey on Automatic Analysis of Facial Expressions," *Res. Artic.*, vol. 20, no. 2, p. 383, Dec. 2016, doi: 10.16984/saufenbilder.92940.

[13] F. Fadlisyah, *Computer Vision dan Pengolahan Citra*. ANDI Yogjakarta, 2007.

[14] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial Expression Analysis," in *Handbook of Face Recognition*, New York: Springer-Verlag, 2005, pp. 247–275. doi: 10.1007/0-387-27257-7_12.

[15] B. K. Durga and D. V Rajesh, "Review of Facial Emotion Recognition System," *Int. J. Pharm. Res.*, vol. 10, no. 03, Jul. 2018, doi: 10.31838/ijpr/2018.10.03.056.

[16] "Introduction to OpenCV."
https://docs.opencv.org/4.x/da/df6/tutorial_py_table_of_contents_setup.html

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1742, Copyright ©2023
Submitted : May 22, 2023, Revised : July 6, 2023, Accepted : July 26, 2023, Published : November 10, 2023

Analisis Penerimaan Pengguna *E-Wallet* Aplikasi Dana Menggunakan UTAUT 2 dan UX TAM

Putri Meliana[1]*, Nurul Mutiah[2], Syahru Rahmayuda[3]
Program Studi Sistem Informasi[1], [2], [3]
Universitas Tanjungpura Pontianak
Pontianak, Indonesia
putrimelianasisfo@student.untan.ac.id[1], nurul@sisfo.untan.ac.id[2], yudarahma@sisfo.untan.ac.id [3]

***Abstract*— DANA is a digital wallet application that has an open platform concept, meaning it can be used on different platforms but is integrated with one another. However, there were complaints that were felt by DANA application users which were conveyed in Google Playstore reviews, namely frequent errors and delays in the transaction process. This is the basis for measuring the level of acceptance of DANA application users based on the user's experience. This research model is an integration of the Unified Theory of Acceptance and Use of Technology 2 (UTAUT 2) model and the User Experience technology Acceptance Model (UX TAM). The data analysis technique used Partial Least Square-Structural Equation Modeling (PLS-SEM) and used SmartPLS 3 tools Data collection was carried out by randomly distributing questionnaires to 100 respondents, namely the Pontianak community with an age range of 15-40 years. Data collection was carried out by distributing questionnaires to 100 respondents, namely the Pontianak community. Of the 21 hypotheses proposed, 10 hypotheses stated there was a relationship between the two variables and the other 11 hypotheses had no relationship. The hypotheses that have an influence are Effort Expectancy on Behavior Intention, Habit on Behavior Intention, Efficiency on Effort Expectancy, Efficiency on Performance Expectancy, Output Quality on Performance Expectancy, Dependability on Habit, Stimulation on Hedonic Motivation, Output Quality on Perceived Usefulness, Dependability on Perceived Ease Of Use, and Behavioral Intention to Use Behavior. The results of the research are in the form of recommendations that are expected to improve the performance of the DANA application.**

Keywords*— *E-Wallet, DANA, UTAUT 2, UX TAM, PLS-SEM

***Abstrak*—DANA merupakan aplikasi dompet digital yang memiliki konsep open platform yaitu dapat digunakan pada platform berbeda namun saling terintegrasi satu sama lain. Namun, terdapat keluhan yang dirasakan para pengguna aplikasi DANA yang disampaikan pada ulasan google playstore yaitu sering terjadi eror dan keterlambatan proses transaksi. Hal tersebut menjadi landasan mengapa perlunya mengukur tingkat penerimaan pengguna aplikasi DANA berdasarkan pengalaman yang rasakan oleh pengguna. Model penelitian ini merupakan pengintegrasian model Unified Theory of Acceptance and Use of Technology 2 (UTAUT 2) dan User Experience technology Acceptance Model (UX TAM). Teknik analisis data menggunakan Partial Least Square-Structural Equation Modelling (PLS-SEM) dan menggunakan tools SmartPLS 3. Pengumpulan data dilakukan dengan menyebarkan kuesioner kepada 100 responden yaitu masyarakat Kota Pontianak dengan rentang usia 15-40 tahun secara acak. Dari 21 Hipotesis yang diajukan, 10 hipotesis dinyatakan adanya hubungan atau pengaruh antara dua variabel dan 11 hipotesis lainnya tidak ada hubungan atau pengaruh. Adapun Hipotesis yang memiliki pengaruh yaitu Effort Expectancy terhadap Behavior Intention, Habit terhadap Behavior Intention, Efficiency terhadap Effort Expectancy, Efficiency terhadap Performance Expectancy, Output Quality terhadap Performance Expectancy, Dependability terhadap Habit, Stimulation terhadap Hedonic Motivation, Output Quality terhadap Perceived Usefulness, Dependability terhadap Perceived Ease Of Use, dan Behavior Intention terhadap Use Behavior. Hasil dari penelitian berupa rekomendasi yang diharapkan dapat meningkatkan performa aplikasi DANA.**

Kata Kunci*— *E-Wallet, DANA, UTAUT 2, UX TAM, PLS-SEM

I. PENDAHULUAN

Pertumbuhan pada bidang teknologi bisnis, dan ekonomi yang secara pesat berkembang memberikan kemudahan bagi manusia dalam segala bidang aktivitas [1]. Pertumbuhan teknologi pada saat ini diiringi dengan adanya internet serta penggunaan *smartphone* yang hampir semua orang miliki. Hal tersebut memberikan dampak yang signifikan bagi aktivitas masyarakat terutama pada bidang ekonomi. Untuk mendorong aktivitas bisnis yang lebih efisien dan efektif maka para pelaku bisnis mulai melakukan sebuah inovasi yang dapat memberikan layanan terbaik dan memenuhi kebutuhan pengguna yaitu dengan memberikan sebuah layanan dompet digital (*e-wallet*) [1]. Popularitas *e-wallet* mulai berkembang sejak munculnya sistem belanja online yang meningkatkan minat belanja pada masyarakat. Berdasarkan data RedSeer yang dimuat dalam situs databoks.katadata.co.id pada Januari 2022, Indonesia saat melakukan transaksi pada E-Commerce sebanyak 29% dilakukan dengan e-wallet [10]. Dengan memanfaatkan *e-wallet* pada *smartphone* yang terhubung pada jaringan internet proses pembayaran dapat dilakukan dengan mudah kapanpun dan dimanapun. Hingga saat ini telah banyak aplikasi dompet digital (*e-wallet*) yang tersedia di Indonesia tak terkecuali kota Pontianak yaitu adalah aplikasi DANA. Aplikasi DANA memiliki konsep *open platform* yang berarti dapat digunakan pada berbagai *platform* yang berbeda namun tetap saling

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

terintegrasi satu sama lain baik secara offline maupun online. Adanya konsep ini diharapkan agar aplikasi DANA dapat memberikan kemudahan sebagai alat pembayaran di berbagai bidang. Aplikasi DANA memiliki ulasan dengan rating sebesar 4,5 pada *google play store*. Namun terdapat beberapa keluhan yang dirasakan pengguna saat menggunakan aplikasi DANA seperti sering terjadi eror, dan keterlambatan dalam proses transaksi hal ini disampaikan dalam kolom ulasan pada *google play store*. Dengan adanya keluhan tersebut maka dilakukan penelitian tentang penerimaan pengguna yang berdasarkan dari pengalaman pengguna dan judul penelitian yang dilaksanakan adalah "Analisis Penerimaan Pengguna E-Wallet Aplikasi DANA Menggunakan UTAUT 2 dan UX TAM" yang berstudi kasus pada masyarakat Kota Pontianak yang menggunakan aplikasi DANA.

II. LANDASAN TEORI

A. Penerimaan Pengguna

Tujuan dari sebagian organisasi yang berbasis sistem informasi adalah untuk meningkatkan kinerja pekerjaan [2]. Secara umum penerimaan pengguna adalah keinginan seseorang atau sekelompok orang untuk menerapkan teknologi informasi dalam melakukan sebuah pekerjaan sehingga dapat mendukung pekerjaan yang sedang orang tersebut lakukan.

Penerimaan pengguna sering dijadikan sebagai faktor yang dapat menunjukan nilai keberhasilan maupun nilai kegagalan suatu proyek sistem informasi. Karena jika suatu sistem ditolak oleh pengguna maka dampak kinerja dapat hilang.

B. E-Wallet

Dompet digital atau *e-wallet* merupakan aplikasi yang berguna dalam proses penyimpanan data instrumen pembayaran, yaitu alat pembayaran yang dapat berupa uang elektronik atau dengan menggunakan kartu, yang bisa menampung dana, untuk terlaksananya sebuah proses pembayaran [3]. *E-wallet* adalah suatu bentuk penyimpanan uang digital berbasis server yang harus koneksi dahulu dengan server penerbit saat akan digunakan.

E-wallet adalah sebuah apliksi yang berfungsi sebagai alat menyimpan uang secara digital dan dapat pula melakukan transaksi secara digital atau non tunai. *E-wallet* atau *mobile wallet* merupakan sebuah layanan pembayaran yang beroperasi melalui sebuah perangkat mobile dan dinaungi di bawah regulasi keuangan.

C. UTAUT 2

UTAUT 2 (*Unified Theory of Acceptance and Use of Technology 2)* adalah suatu model penerimaan pengguna yang banyak digunakan sebagai metode dalam melakukan penelitian tentang penerimaan pengguna terutama dalam konteks konsumen pada penggunaan suatu teknologi [3]. UTAUT 2 berisi pemaparan rinci tentang penggunaan teknologi informasi oleh seseorang atau kelompok [1]. 7 variabel *independent* pada UTAUT 2 terdiri dari *performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value,* dan *habit,* serta dua variabel *dependent,* yaitu *behavioral intention* dan *use behavior* [11].

D. UX TAM

Menurut Venkatesh [4], *Technology Acceptance Model* (TAM) didefinisikan sebagai sebuah konsep yang membahas tentang sikap seorang pengguna terhadap penerimaan suatu teknologi informasi. Terdapat 3 konstruktur utama pada TAM yaitu *perceived usefulness* (persepsi kegunaan), *perceived ease of use* (persepsian kemudahan penggunaan), dan *actual technology use* (penggunaan teknologi sesungguhnya) [5]. Sedangkan *User Experience* menurut definisi ISO 9241-210 (2009) merupakan suatu keadaan yang menggambarkan perasaan pengguna setelah menggunakan sebuah produk, sistem, ataupun jasa. Pengalaman pengguna sendiri adalah sebuah perasaan yang dirasakan pengguna setelah menggunakan sebuah produk, baik itu perasaan puas senang maupun sebaliknya.

Untuk mengukur pengalaman pengguna dapat menggunakan *User Experience questionnaire (UEQ)* oleh Laugwitz,B & Schrepp [6]. *User Experience Questionnaire* terdiri atas 6 skala penilaian yaitu *Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, Novelty*.

Mlekus dalam "*How to raise technology acceptance: user experience characteristics as technology-inherent determinants*" melakukan suatu penelitian tentang penerimaan pengguna dengan menggabungkan dua model penelitian yaitu *Technology Acceptance Model* (TAM) dan *User experience* dan dikenal sebagai UX TAM [7]. UX TAM adalah sebuah model penerimaan pengguna yang merupakan perluasan dari model TAM yang berdasarkan karakteristik pengalaman pengguna.

E. PLS-SEM

SEM (*Structural Equation Modelling*) merupakan analisis statistik yang tak jarang dipergunakan untuk menguji beberapa korelasi atau interaksi yang dibentuk oleh satu atau beberapa peubah tak bebas dengan satu atau beberapa peubah bebas yang relatif sulit secara simultan. Pada SEM terdapat teknik analisis yaitu berdasarkan *covariance* (CBSEM) dan SEM yang berbasis *variance* (PLS) [8].

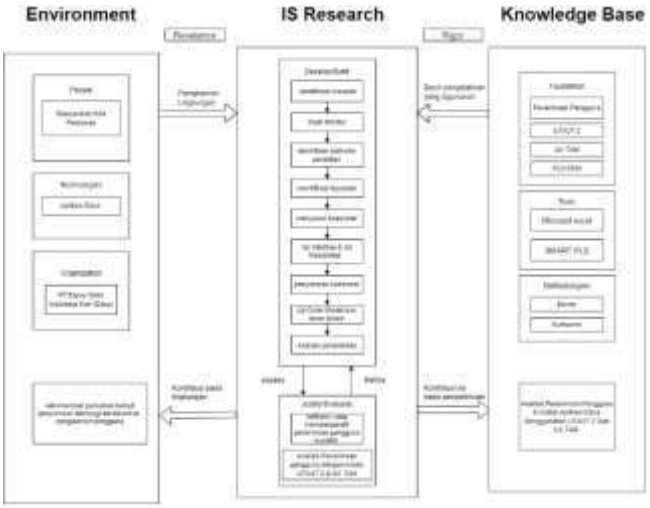
Pada tahun 1982 Joreskog dan Wold mengembangkan PLS yaitu metode umum dalam peramalan model laten yang diukur oleh peubah penjelas secara tidak langsung [8]. PLS-SEM digunakan untuk memprediksi variabel-variabel utama pada penelitian yang merupakan riset perluasan dari teori struktural yang telah ada.

III. METODE PENELITIAN

Pada penelitian ini menggunakan pendekatan kuantitatif untuk mengukur tingkat penerimaan pengguna e-wallet aplikasi DANA. Populasi yang akan diteliti saat dilaksanakannya penelitian adalah masyarakat Kota Pontianak dan sampel yang digunakan adalah masyarakat Kota Pontianak yang menggunakan *e-wallet* aplikasi DANA. Dikarenakan populasi pengguna *e-wallet* tidak diketahui maka dalam pengambilan sampel akan menggunakan rumus lemeshow dengan usia yang ditentukan pada rentang 15-40 tahun.

Metodologi penelitian yang digunakan adalah kerangka kerja IS Research seperti pada Gambar 1.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023



Gambar 1. IS Research Framework

Berdasarkan kerangka kerja IS Research maka diketahui bahwa terdapat tiga komponen yaitu *Environment*, *IS Research*, dan *Knowledge Base* [13]. *Environment* pada penelitian ini adalah masyarakat Kota Pontianak yang akan menjadi responden dan teknologi yang ada berupa aplikasi DANA, serta organisasi adalah PT Espay Debit Indonesia Koe selaku perusahaan pengembang aplikasi DANA. Pada *IS Research* adalah gambaran besar tahap penelitian dan terdapat 2 fase yaitu *Develop/Build* dan *Justify/Evaluate*. *Knowledge Base* yang digunakan sebagai landasan teori adalah Penerimaan Pengguna, UTAUT 2, UX TAM dan PLS-SEM. Metodologi yang digunakan untuk mengumpulkan data adalah berupa survey dan penyebaran kuesioner. Terdapat pula tool Microsoft excel dan SmartPLS yang dimanfaatkan untuk melakukan pengolahan data hasil kuesioner yang telah disebar.

IV. PEMBAHASAN

A. *Pementuan sampel*

Dalam melakukan penentuan ukuran sampel yang akan digunakan maka menggunakan rumus Lemeshow untuk mengetahui jumlah populasi yang tidak diketahui secara pasti [9]. Adapun perhitungan rumus Lemeshow dapat dilihat pada persamaan dibawah ini:

$$n = \frac{1{,}96^2 \times 0{,}5(1-0{,}5)}{0{,}1^2}$$

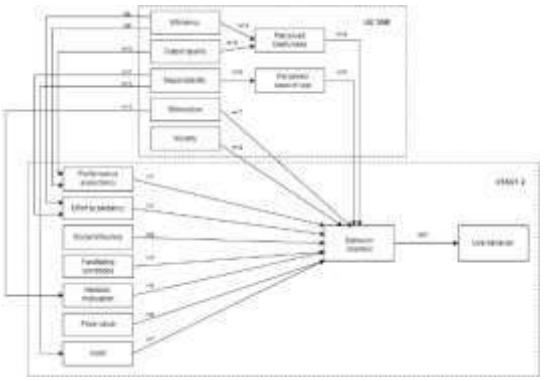
$$n = \frac{1{,}96^2 \times 0{,}5(0{,}5)}{0{,}1^2}$$

$$n = 96{,}04$$

Berdasarkan perhitungan rumus diatas maka jumlah sampel untuk penelitian ini adalah sebanyak 96,04 dan akan dibulatkan menjadi 100 orang responden.

B. *Pemetaan Model*

Pemetaan model dilakukan dalam pengembangan model penelitian baru yaitu dengan melakukan pengintegrasian pada model UTAUT 2 dan model UX TAM. Variabel yang terdapat pada model UTAUT 2 dan UX TAM selanjutnya dianalisis guna didapatkan variabel yang sesuai dengan kebutuhan penelitian. Semua variabel yang dimiliki pada model UTAUT akan digunakan pada penelitian ini namun, pada model UX TAM terdapat satu variabel yang akan dibuang atau tidak digunakan yaitu *Perspicuity*. Variabel *Perspicuity* tidak digunakan pada penelitian ini karena variabel *Perspicuity* memiliki definisi yang sama dengan variabel *Effort Expectancy* yaitu kemudahan dalam menggunakan sistem. Oleh sebab itu hanya variabel *Effort Expectancy* yang akan digunakan sebagai variabel dalam mengukur tingkat kemudahan penggunaan sistem. Berdasarkan variabel-variabel yang telah dipilih dari model UTAUT 2 dan UX TAM maka selanjutnya akan dibuat sebuah model penelitian berikut ini



Gambar 2. Model Penelitian UTAUT2 dan UX TAM

Berdasarkan model penelitian yang telah dipetakan pada Gambar 2 maka hipotesis yang dirumuskan adalah sebagai berikut:

1) H1. *Performance Expectancy* memiliki pengaruh terhadap *Behavior Intention*:
2) H2. *Effort Expectancy* memiliki pengaruh terhadap *Behavior Intention*
3) H3. *Social Influence* memiliki pengaruh terhadap *Behavior Intention*
4) H4. *Facilitating Conditions* memiliki pengaruh terhadap *Behavior Intention*
5) H5. *Hedonic Motivation* memiliki pengaruh terhadap *Behavior Intention*
6) H6. *Price Value* memiliki pengaruh terhadap *Behavior Intention*
7) H7. *Habit* memiliki pengaruh terhadap *Behavior Intention*
8) H8. *Efficiency* memiliki pengaruh terhadap *Effort Expectancy*
9) H9. *Efficiency* memiliki pengaruh terhadap *Performance Expectancy*
10) H10. *Output Quality* memiliki pengaruh terhadap *Performance Expectancy*
11) H11. *Dependability* memiliki pengaruh terhadap *Effort Expectancy*
12) H12. *Dependability* memiliki pengaruh terhadap *Habit*

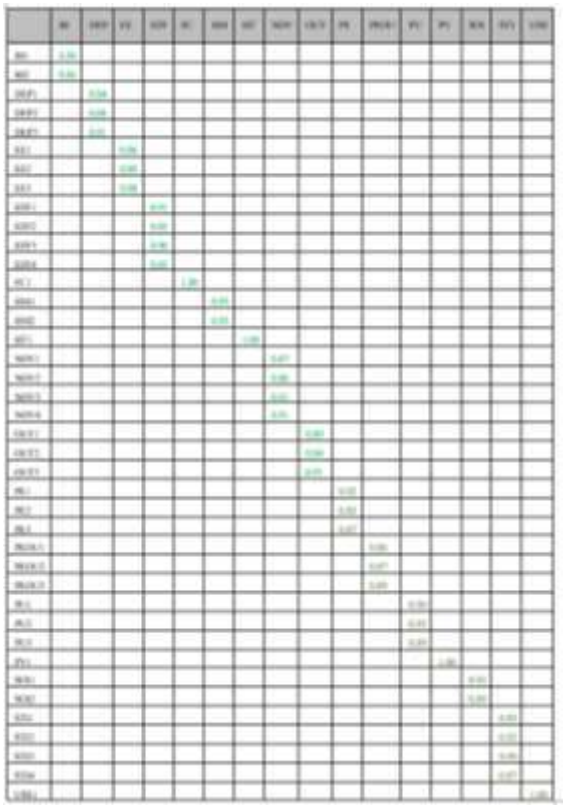
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

13) H13. *Stimulation* memiliki pengaruh terhadap *Hedonic Motivation*
14) H14. *Efficiency* memiliki pengaruh terhadap *Perceived Usefulness*
15) H15. *Output Quality* memiliki pengaruh terhadap *Perceived Usefulness*
16) H16. *Dependability* memiliki pengaruh terhadap *Perceived Ease Of Use*
17) H17. *Stimulation* memiliki pengaruh terhadap *Behavior Intention*
18) H18. *Novelty* memiliki pengaruh terhadap *Behavior Intention*
19) H19. *Perceived Usefulness* memiliki pengaruh terhadap *Behavior Intention*
20) H20. *Perceived Ease Of Use* memiliki pengaruh terhadap *Behavior Intention*
21) H21. *Behavior Intention* memiliki pengaruh terhadap *Use Behavior*

C. Evaluasi Model

Evaluasi model penelitian ini menggunakan teknik analisis PLS-SEM dengan *tools* SmartPLS versi 3.0. Tahap evaluasi model yang dilakukan adalah dengan melakukan menguji pada outer model dan inner model.

1. Uji Outer Model
   1) Uji Validitas Konvergen
      Menggambarkan apakah indikator-indikator mampu mewakili variabel laten yang dibangun. Pengujian dilakukan dengan mengamati nilai pada *loading factor* yang didapat dari nilai tiap indikator terhadap variabel. Jika nilai *loading factor* > 0,7 maka indikator bernilai VALID [15]. Adapun dengan mengunakan *software* SmartPLS 3.0 nilai *loading factor* pada uji validitas konvergen disajikan seperti pada gambar berikut.



Gambar 3 *Loading Factor*

Setelah dilakukan uji validitas konvergen diperoleh bahwa semua nilai indikator *loading factor* (>0,7) dan dilanjutkan dengan melihat nilai AVE (*Average Variance Extracted*) jika nilai (>0,5) maka tiap indicator pada variabel dapat dikatakan akurat.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

Tabel 1 Nilai AVE

Average Variance Extracted (AVE)	
BI	0.783
DEP	0.769
EE	0.781
EFF	0.771
FC	1.000
HM	0.872
HT	1.000
NOV	0.801
OUT	0.784
PE	0.784
PEOU	0.771
PU	0.827
PV	1.000
SOI	0.826
STI	0.784
USE	1.000

2) Uji Validitas Diskriminan

Digunakan untuk menguji dan mengetahui besarnya perbedaan satu konstruk dengan konstruk lainnya dengan membandingkan nilai *cross loading* dari tiap-tiap indikator terhadap konstruknya dengan nilai *cross loading* dari indikator konstruk lainnya. Uji validitas diskriminan bernilai baik jika nilai suatu *cross loading* pada indikator pada konstruknya (>) dari nilai *cross loading* pada indikator konstruk lainnya.

3) Uji Reliabilitas

Variabel dapat dikatakan bernilai reliabel jika nilai pada *Composite Reliability* dan *Cronbach alpha* lebih dari 0,7 [14]. Adapun nilai *Composite Reliability* dan nilai *Cronbach alpha* dapat dilihat pada tabel

Tabel 2 nilai *Composite Reliability* dan *nilai Cronbach alpha*

	Cronbach's Alpha	Composite Reliability
BI	0.723	0.878
DEP	0.849	0.909
EE	0.859	0.914
EFF	0.901	0.931
FC	1.000	1.000
HM	0.854	0.932
HT	1.000	1.000
NOV	0.917	0.942
OUT	0.864	0.916
PE	0.862	0.916
PEOU	0.851	0.910
PU	0.896	0.935
PV	1.000	1.000
SOI	0.790	0.904
STI	0.908	0.935
USE	1.000	1.000

Berdasarkan nilai yang terdapat pada tabel 2 maka didapatkan nilai bahwa *cronbach alpha* dan *composite reliability* pada semua variabel bernilai reliabel.

2. Uji Inner Model

Digunakan untuk melihat keterkaitan antara variabel satu dengan variabel lainnya.

1) Uji *R-Square*

Uji *R-Square* adalah pengujian yang dilakukan untuk mengukur pengaruh variabel *independent* terhadap variabel *dependent*. Hasil *R-Square* bernilai kuat apabila sebesar 0,67, bernilai moderat apabila sebesar 0,33, dan bernilai lemah apabila sebesar 0,19.

Tabel 3 Nilai *R-Square*

	R-Square	Keterangan
BI	0.731	Kuat
EE	0.549	Moderat
HM	0.657	Moderat
HT	0.261	lemah
PE	0.629	Moderat
PEOU	0.447	Moderat
PU	0.514	Moderat
USE	0.549	Moderat

Berdasarkan hasil pada Tabel 3 nilai *R-Square* untuk variabel BI adalah 0,731 (73,1%) dan dikategorikan sebagai kuat. Variabel EE dengan nilai *R-Square* sebesar 0,549 (54,9%) dikategorikan sebagai moderat. Variabel HM dengan nilai *R-Square* sebesar 0,657 (65,7%) dikategorikan sebagai moderat Variabel HT dengan nilai *R-Square* sebesar 0,261 (26,1%) dikategorikan sebagai lemah. Variabel PE dengan nilai *R-Square* sebesar 0.629 (62,9%) dikategorikan sebagai moderat. Variabel PEOU dengan nilai *R-Square* sebesar 0,447 (44,7%) dikategorikan sebagai moderat. Variabel PU dengan nilai *R-Square* sebesar 0.514 (51,4%) dikategorikan sebagai moderat. Variabel USE dengan nilai *R-Square* sebesar 0.549 (54,9%) dikategorikan sebagai moderat.

2) Path Coefficient

Pengujian yang berfungsi untuk menganalisa hubungan antar variabel pada model penelitian, serta melihat hasil sebuah hipotesis apakah akan diterima ataupun ditolak dengan meninjau nilai *P-Values* dan *T-Statistics*. Jika pada suatu hubungan variabel X terhadap Y memiliki nilai P-Values lebih kecil dari 0,05 atau nilai *T-Statistics* lebih besar dari 1,96 maka dapat ditarik kesimpulan bahwa suatu variabel X mempengaruhi variabel Y

Tabel 4 *Uji Path Coefficient*

Hx	Hubungan	T-Statistics	P-Values	Hasil
H1	PE -> BI	0.099	0.921	Ditolak
H2	EE -> BI	2.233	0.026	Diterima
H3	SOI -> BI	1.546	0.123	Ditolak
H4	FC -> BI	1.118	0.264	Ditolak
H5	HM -> BI	0.786	0.433	Ditolak
H6	PV -> BI	0.580	0.562	Ditolak
H7	HT -> BI	3.760	0.000	Diterima
H8	EFF -> EE	5.494	0.000	Diterima
H9	EFF -> PE	4.938	0.000	Diterima
H10	OUT -> PE	2.373	0.018	Diterima
H11	DEP -> EE	1.175	0.241	Ditolak
H12	DEP -> HT	5.657	0.000	Diterima
H13	STI -> HM	14.580	0.000	Diterima
H14	EFF -> PU	1.685	0.093	Ditolak
H15	OUT -> PU	3.854	0.000	Diterima

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

H16	DEP -> PEOU	6.165	0.000	Diterima
H17	STI -> BI	0.580	0.562	Ditolak
H18	NOV -> BI	1.052	0.293	Ditolak
H19	PU -> BI	0.480	0.631	Ditolak
H20	PEOU -> BI	0.007	0.994	Ditolak

Berdasarkan Tabel 4 maka dinyatakan bahwa hipotesis yang diterima adalah H2, H7, H8, H9, H10, H12, H13, H15, H16.

3) Uji *Predictive Relevance*
Digunakan untuk melihat nilai observasi yang dilakukan dan untuk menilai kecocokan relevansi model penelitian secara struktural. Nilai *predictive relevance* dikatakan sudah baik atau memiliki *predictive relevance* jika bernilai lebih dari 0 (>0) [12].

Tabel 5 *Predictive Relevance*

	Q^2predict
BI	0.490
EE	0.366
HM	0.559
HT	0.250
PE	0.413
PEOU	0.296
PU	0.386
USE	0.538

D. Rekomendasi
Berdasarkan pada analisis data pada model penerimaan pengguna pada model UTAUT 2 dan UX TAM maka diperlukan rekomendasi-rekomendasi terhadap evaluasi faktor-faktor yang memiliki pengaruh pada minat pengguna aplikasi DANA. Rekomendasi ini diharapkan dapat menjadi acuan dalam meningkatkan performansi kinerja pada aplikasi dana untuk kedepannya.

1) *Effort Expectancy* terhadap *Behavior intention*. Pada variabel *Effort Expectancy* yaitu sebuah tingkat usaha serta kemudahan yang dirasakan pengguna saat sedang menggunakan aplikasi DANA memiliki pengaruh terhadap *Behavior Intention* yaitu keinginan menggunakan aplikasi DANA berdasarkan manfaat yang dirasakan. Sehingga rekomendasi yang dapat diberikan agar pengguna dalam menggunakan aplikasi dana merasakan kemudahan adalah dengan melakukan riset secara berkala berdasarkan keluhan yang mungkin dirasakan oleh pengguna.
2) *Habit* terhadap *Behavior Intention*. Pada variabel *Habit* yaitu kecenderungan yang dirasakan pengguna dalam menggunakan aplikasi secara terus menerus memiliki pengaruh terhadap *Behavior Intention* yaitu keinginan menggunakan aplikasi DANA berdasarkan manfaat yang dirasakan. Rekomendasi yang diberikan agar pengguna dapat menggunakan aplikasi DANA secara terus menerus adalah dengan meningkatkan performansi sistem karena dengan adanya peningkatan pada performansi maka pengguna akan merasa kan manfaat jika menggunakan aplikasi DANA dan secara tidak langsung menjadikan DANA sebagai alat pembayaran pada banyak transaksi pembayaran.
3) *Efficiency* terhadap *Effort Expectancy*. Pada variabel *efficiency* yaitu kondisi dimana pengguna tidak memerlukan upaya lebih dalam menggunakan aplikasi DANA memiliki pengaruh terhadap *Effort Expectancy*. Sehingga rekomendasi yang diberikan adalah dengan melakukan riset kepada pengguna dan menganalisis permasalah yang menyebabkan kesulitan yang dirasakan pengguna saat menggunakan aplikasi dana dan selanjutkan dilakukan evaluasi sehingga pengguna kedepannya dapat menggunakan dan mengoperasikan aplikasi DANA tanpa adanya usaha lebih.
4) *Efficiency* terhadap *Performance Expectancy*. Pada variabel *efficiency* yaitu kondisi dimana pengguna tidak memerlukan upaya lebih dalam menggunakan aplikasi DANA memiliki pengaruh terhadap *Performance Expectancy* yaitu kepercayaan yang pengguna rasakan bahwa dengan menggunakan aplikasi DANA maka akan meningkatkan harapan kinerja. Sehingga rekomendasi yang diberikan adalah dengan menjaga kinerja sistem agar tetap memberikan nilai *efficiency* seperti dapat digunakan secara praktis dan cepat saat digunakan dalam transaksi pembayaran sehingga nantinya pengguna dapat merasakan peningkatan pada harapan kinerja dan mencapai tujuan kerja yang diharapkan.
5) *Output Quality* terhadap *Performance Expectancy*. Pada variabel *Output Quality* yaitu sistem dapat beroperasi dan menghasilkan keluaran yang baik memiliki pengaruh terhadap *Performance Expectancy* yaitu kepercayaan yang pengguna rasakan bahwa dengan menggunakan aplikasi DANA maka akan meningkatkan harapan kinerja. Sehingga rekomendasi yang diberikan adalah perlunya evaluasi yang berkelanjutan pada alur operasi sistem sehingga saat digunakan tidak ada kendala dan dapat menghasilkan keluaran yang baik sehingga pengguna dapat mencapai harapan kerja yang ingin dicapai.
6) *Dependability* terhadap *Habit*. Pada variabel *Dependability* yaitu kemampuan mengendalikan dan memprediksi perilaku sistem memiliki pengaruh terhadap *Habit* yaitu pengguna menggunakan aplikasi DANA karena telah menjadi sebuah kebiasaan. Sehingga rekomendasi yang diberikan adalah aplikasi DANA sebaiknya dalam alur operasinya memberikan kemudahan sehingga jika digunakan oleh user baru maka user dapat memprediksi dan mengendalikan sistem dengan mudah dan nantinya akan berdampak pada penggunaan aplikasi yang berkelanjutan dan dapat menjadi sebuah kebiasaan.
7) *Stimulation* terhadap *Hedonic Motivation*. Pada variabel *Stimulation* yaitu perasaan tertarik untuk menggunakan aplikasi memiliki pengaruh terhadap *Hedonic motivation* yaitu perasaan senang saat menggunakan aplikasi DANA. Sehingga rekomendasi yang diberikan adalah pihak DANA perlu memberikan sesuatu yang dapat menarik perhatian pelanggan seperti memberikan diskon atau *cashback* saat melakukan transaksi tertentu sehingga pelanggan akan mendapatkan perasaan senang saat menggunakan aplikasi DANA.
8) *Output Quality* terhadap *Perceived Usefulness*. Pada

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

variabel *Output Quality* yaitu sistem dapat beroperasi dan menghasilkan keluaran yang baik memiliki pengaruh terhadap *Perceived Usefulness* yaitu dengan menggunakan aplikasi DANA akan meningkatkan performansi kerja. Sehingga rekomendasi yang diberikan adalah perlunya menjaga alur operasi yang baik serta memastikan agar sistem dapat menghasilkan keluaran yang baik pula sehingga saat pengguna menggunakan aplikasi DANA maka performansi kerja dapat meningkat.

9) *Dependability* terhadap *Perceived Ease Of Use*. Variabel *Dependability* yaitu kemampuan pengguna dalam mengendalikan dan memprediksi perilaku sistem memiliki pengaruh terhadap variabel *Perceived Ease Of Use* yaitu pengguna dalam menggunakan aplikasi DANA terbebas dari upaya yang tidak perlu. Sehingga rekomendasi yang diberikan adalah aplikasi DANA sebaiknya dalam alur operasinya memberikan alur yang mudah untuk dipahami dan dikendalikan sehingga pengguna saat menggunakan aplikasi tidak memerlukan upaya yang berlebih untuk menyelesaikan transaksi pembayaran.
10) *Behavior Intention* terhadap *Use Behavior*. Variabel *Behavior Intention* yaitu keinginan menggunakan aplikasi DANA berdasarkan manfaat yang dirasakan memiliki pengaruh terhadap *Use behavior* yaitu frekuensi pengguna dalam menggunakan sistem. Sehingga rekomendasi yang diberikan adalah dengan memperhatikan fitur dan layanan yang diberikan kepada pengguna dan menganalisa kebutuhan pengguna sehingga aplikasi DANA dapat memenuhi kebutuhan pengguna.

V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan maka didapatkan beberapa hasil dari analisis penerimaan pengguna *e-wallet* aplikasi DANA menggunakan UTAUT 2 dan UX TAM. Adapun hasil penelitian tersebut adalah sebagai berikut.

1) Penelitian ini menggunakan responden sebanyak 100 responden yang di ambil secara acak dengan rantang usia 15-40 tahun.
2) Analisis data yang dilakukan dengan menggunakan *tools* SmartPLS dengan terdapat dua 12 variabel pada model UTAUT 2 dan UX TAM yang akan diuji apakah memiliki mempengaruhi terhadap penerimaan pengguna *e-wallet* aplikasi DANA yaitu *Performance Expectancy, Effort Expectancy, Hedonic Motivation, Habit, Efficiency, Output Quality, Dependability, Stimulation, Perceived Usefulness, Perceived Ease Of Use, Behavior Intention, Use Behavior*.
3) Berdasarkan hasil analisis yang dilakukan maka didapat faktor atau variabel yang mempengaruhi penerimaan pengguna sehingga rekomendasi yang diberikan adalah dengan melakukan peningkatan pada beberapa faktor yang dapat meningkatkan kualitas layanan aplikasi DANA yaitu mampu memberikan keuntungan sehingga dapat mencapai tujuan kerja (*Performance Expectancy*) , Mampu memberikan kemudahan dalam menggunakan aplikasi DANA *(Effort Expectancy*), mampu menciptakan pengalaman yang menyenangkan saat menggunakan aplikasi DANA (*Hedonic Motivation*), menambah perluasan pada berbagai *platform* sehingga pengguna dapat menggunakan sistem menjadi sebuah kebiasaan (*Habit*) , menyediakan layanan yang lebih efisien (*Efficiency*), mampu memberikan hasil akhir yang maksimal (*Output Quality*), menyediakan layanan yang mudah diprediksi oleh pengguna baru (*Dependability*), memberikan tawaran yang dapat menarik para pengguna (*Stimulation*).

REFERENCES

[1] Mahwadha, W. I. (2019). BEHAVIORAL INTENTION OF YOUNG CONSUMERS TOWARDS E-WALLET ADOPTION: AN EMPIRICAL STUDY AMONG INDONESIAN USERS. RJOAS, 79-93.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[2] Davis. (1993). User Acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Academic Press Limited, 475-487.

[3] Hidayat, M. T., Aini, Q., & Fetrina, E. (2020). Penerimaan Pengguna E-Wallet Menggunakan UTAUT 2 (Studi Kasus). Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 9(3), 239-247.

[4] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology. Toward a Unified View, 27(3), 425-478.

[5] Hamrul H, Soedijono B, Amborowati A. (2013). Analisis Perbandingan Metode TAM Dan UTAUT dalam Mengukur Kesuksesan Penerapan Sistem Informasi Akademik (Studi Kasus Penerapan Sistem Informasi STMIK Dipanegara Makassar). Semin. Nas. Inform, 140-146.

[6] Laugwitz,B., Held T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience questionnaire. In Symposium of the Austrian HCI and usability engineering group, 63-76.

[7] Mlekus,L., Bantler,D., Paruzel, A., et al. (2020). How to raise technology acceptance: user experience characteristics as technology-inherent determinants. Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie, 51(3), 273-283.

[8] Ningsi, B. (2018). Analisis Kepuasan Pelanggan Atas Kualitas Produk dan Pelayanan Dengan Metode SEM-PLS. Jurnal Statistika dan Aplikasinya, 2(2), 8-16.

[9] Lemeshow, S., Hosmer, DW., Klarr, J., & Lwanga, S.K. (1990). Adequacy of sample size. Chichester: John Wiley & Sons Ltd.

[10] Pahlevi, R. (2022). Survei DailySocial: OVO Jadi Dompet Digital Paling Banyak Dipakai Masyarakat. dalam https://databoks.katadata.co.id/datapublish/2022/01/12/survei-dailysocial-ovo-jadi-dompet-digital-paling-banyak-dipakai-masyarakat, 15 Juni 2022

[11] Watmah, S., Fauziah, S., & Herlinawati, N. (2020). Identifikasi Faktor Pengaruh Penggunaan Dompet Digital Menggunakan Metode TAM Dan UTAUT2. Indonesian Journal on Software Engineering (IJSE), 6(2), 261-269.

[12] Anuraga, G. S. (2017). Structural equation modeling–partial least square untuk pemodelan indeks pembangunan kesehatan masyarakat (IPKM) di Jawa Timur. In Seminar Nasional Matematika dan Aplikasinya, Surabaya, 257-263.

[13] Hevner, A. R., March, S. T., Park, J., & Ram, S. . (2004). Design science in information systems research. Management Information Systems Quarterly, 75-105.

[14] Sugiyono. (2019). Metode Penelitian Kuantitatif, Kualitatif, Dan R&D. Bandung: Alfabeta.

[15] Juningsih, E. H., Aziz, F., Ismunandar, D., Sarasati, F., Irmawati, I., & Yanto, Y. (2020). Penggunaan Model UTAUT2 Untuk Memahami Persepsi Pengguna Aplikasi G-Meet. Indonesian Journal on Software Engineering (IJSE), 6(2), 289-295.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1750, Copyright ©2023
Submitted : May 22, 2023, Revised : July 5, 2023, Accepted : October 8, 2023, Published : November 6, 2023

Comparison Analysis of Graph Theory Algorithms for Shortest Path Problem

Yosefina Finsensia Riti[1], Jonathan Steven Iskandar[2], Hendra[3]
Informatics Study Program, Faculty of Engineering[1], [2], [3]
Darma Cendika Catholic University, Surabaya, Indonesia
yosefina.riti@ukdc.ac.id[1], jonathan.iskandar@student.ukdc.ac.id[2], hendra@student.ukdc.ac.id[3]

***Abstract*— The Sumba region, Indonesia, is known for its extraordinary natural beauty and unique cultural richness. There are 19 interesting tourist attractions spread throughout the area, but tourists often face difficulties in planning efficient visiting routes. From this case, it can be solved by applying graph theory in terms of searching for the shortest distance which is completed using the shortest path search algorithm. Then these 19 tourist objects are used to build a weighted graph, where the nodes represent the tourist objects and the edges of the graph describe the distance or travel time between these objects. Therefore, this research aims to compare the shortest path search algorithm with parameters to compare the shortest distance results, algorithm complexity and execution time for tourism in the Sumba area. The results of this research involve a comparison of several shortest path search algorithms, with the aim of finding the shortest distance results, algorithm complexity, and execution time for tourism in the Sumba area. Based on the test results of the five algorithms with the parameters that have been prepared, and the findings show that each algorithm has its own characteristics, the results are as follows: Dijkstra's algorithm can be used to calculate the shortest route for single-source and single-destination types. This resembles the Bellman-Ford algorithm, only the Bellman-Ford algorithm can be used simultaneously on graphs that have negative weight values. Meanwhile, the Floyd-Warshall algorithm is suitable for use on the all-pairs type. Then, the Johnson Algorithm can be used to determine the shortest path from all pairs of paths where the destination node is located in the graph. Finally, the Ant Colony algorithm to compute from a node to each pair of destination nodes.**

Keywords*— *Dijkstra Algorithm, Bellman-Ford Algorithm, Floyd-Warshall Algorithm, Johnson Algorithm, Ant Colony Algorithm

I. INTRODUCTION

Graph theory is one of the topics in the computer field that is studied and is a branch of discrete mathematics [1] that studies related graphs, namely the relationship between one object and another object, where the connection consists of a set of vertices/points (vertex/nodes) and sides/lines (edges) that connect one point to another point. Graph theory is used in studies related to relations or relationships between objects that are implemented to solve various problem models, one of which is the optimal path or route search problem [2-3] so that it can minimize costs [4] and time efficiency [5], optimization scheduling [6-8], communication network modeling [9], and other issues [9]. Searching for the shortest path is the process of finding a path between two vertices, namely from the source node to the destination node in a graph by minimizing the number of weights so that the path with the smallest weight is obtained.

The Sumba region is a tourist destination rich in attractive natural and cultural tourist attractions, and 19 attractive tourist attractions are distributed throughout the region. However, for many tourists, the challenge often faced is to plan efficient visits to these various tourist attractions on a single trip. This is becoming increasingly important as their time and resources are limited. To address this problem, this study applied the implementation of graph theory in terms of the shortest distance search between tourist attractions in Sumba area through testing of several algorithms that have often been used to solve related problems. With this effort, it is also possible for each algorithm to perform efficiently on what kind of graph and path models, as well as what the advantages and disadvantages of each algorithm are. The hope can also be a recommendation for readers to know the most efficient algorithms for researching or solving the same problem.

Implementation of graph theory in the case of finding the shortest distance can be solved using the shortest path search algorithm, which is classified into a single source and multi-source [10] where the single source algorithm is the shortest path search algorithm from one source node to various other destination nodes. These algorithms are among the Greedy [11], Dijkstra, and Bellman-Ford algorithms. Meanwhile, multi-source algorithms look for the shortest path by calculating all pairs of vertices in a graph, including the Floyd Warshall, Ant Colony, and Johnson algorithms. Several algorithms have made it possible for tourists to visit as many tourist attractions as possible in one visit, while minimizing travel time.

Some research using the Greedy algorithm includes determining the best tour packages for tour communicatists [12] and determining the fastest routes for public transportation [13]. The application of Dijkstra's algorithm

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

includes determining the shortest path [14-21], that there is the application of the Bellman-Ford algorithm in finding the best path including the network [22], a comparison of the Bellman-Ford algorithm with other algorithms such as Prim's algorithm [23], Dijkstra [24-27]. Regarding the application of the Floyd-Warshall algorithm in finding the shortest route, they include [16][28]. The application of other shortest route search algorithms such as Ant Colony includes [1][20-21][29-30]. Meanwhile, the application of Johnson's algorithm in finding the shortest path [6]. There are also other studies that make comparisons between the Dijkstra, Bellman-Ford, and Floyd-Warshall algorithms [10].

The shortest path search algorithms that have been mentioned have their own advantages and disadvantages which can be measured through several parameters including, the results of the shortest path or route given [24-25], negative sides [26], negative cycles [26][10], the memory allocation used [10], the complexity of the algorithm [24], as well as the execution time required for an algorithm [26][10].

This study has the primary purpose of testing and comparing from five different shortest path search algorithms in optimizing the travel of tourists in the Sumba region. This area is a focal point because it presents its own challenges in determining tourist attractions to visit, along with the condition of road infrastructure that still suffers a lot of damage. Therefore, this study will focus on a number of relevant parameters, including shortest distance results, algorithm complexity, and execution time.

II. METHODOLOGY

The data used in this research is about the distance data between tourist objects in Southwest Sumba and West Sumba. The data is taken from google maps and consists the distance between TMC Airport to Kita Beach, and also to Kabisu Lobo Oro Site, Lendongara Hill, Waikelo Beach, Kawona Beach, Karakar Indah Beach, Waikuri Lagoon, Mandorak Beach, Tanjung Karoso Beach, Tanjung Karoso Resort, Rotenggaro Beach, Bawana Beach, Watu Malandong Beach, Sumba Culture House, Praijing Village, Lailiang Beach, Rua Beach, and Mata Yangu Waterfall. The route data was taken with the consideration of good road access, paved and two-way without any damage.

The following is an explanation of the initialization of the dataset to be used. To facilitate the writing of data tables, each location name will be represented by numbering as shown in Table 1 of initializing variables. Then, continue with Fig. 1 that displays information about the distance values in each row and column that have been adjusted by initializing the location variables and distance values using units of kilometers.

TABLE I. VARIABLE INITIALIZATION FOR DATASET

Variable Initialization	Sumba Regional Destination	Variable Initialization	Sumba Regional Destination
1	TMC Airport	11	Tanjung Karoso Resort
2	Kita Beach	12	Rotenggaro Beach
3	Kabisu Lobo Oro Site	13	Bawana Beach
4	Lendongara Hill	14	Watu Malandong Beach
5	Waikelo Beach	15	Sumba Culture House
6	Kawona Beach	16	Praijing Village
7	Karakat Indah Beach	17	Lailiang Beach
8	Waikuri Lagoon	18	Rua Beach
9	Mandorak Beach	19	Mata Yangu Waterfall
10	Tanjung Karoso Beach		

FROM / TO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0	19	9	10	8,3	8,7	22	42	40	51	43	47	58	63	6,5	43	66	59	65
2	19	0	20	13	22	26	40	59	57	67	59	63	75	79	23	43	68	61	67
3	9	20	0	11	12	16	30	49	47	57	49	53	64	69	13	49	72	65	71
4	10	13	11	0	13	17	31	50	49	58	50	54	66	70	14	50	73	66	73
5	8.3	22	12	13	0	5.7	19	40	38	54	46	50	62	66	11	47	70	63	70
6	8.7	26	16	17	5.7	0	14	34	32	49	42	45	56	61	12	49	72	64	71
7	22	40	30	31	19	14	0	20	18	35	28	31	43	47	17	57	80	73	79
8	42	59	49	50	40	34	20	0	2	11	9.1	20	31	36	38	78	101	94	100
9	40	57	47	49	38	32	18	2	0	9	7.1	18	29	34	36	76	99	92	98
10	51	67	57	58	54	49	35	11	9	0	5.7	14	26	30	43	83	106	81	105
11	43	59	49	50	46	42	28	9.1	7.1	5.7	0	18	30	35	37	77	101	86	100
12	47	63	53	54	50	45	31	20	18	14	18	0	19	24	41	81	95	75	104
13	58	75	64	66	62	56	43	31	29	26	30	19	0	7.9	53	72	79	59	95
14	63	79	69	70	66	61	47	36	34	30	35	24	7.9	0	57	67	74	54	90
15	6.5	23	13	14	11	12	17	38	36	43	37	41	53	57	0	41	64	57	64
16	43	43	49	50	47	49	57	78	76	83	77	81	72	67	41	0	19	21	24
17	66	68	72	73	70	72	80	101	99	106	101	95	79	74	64	19	0	20	51
18	59	61	65	66	63	64	73	94	92	81	86	75	59	54	57	21	20	0	43
19	65	67	71	73	70	71	79	100	98	105	100	104	95	90	64	24	51	43	0

Fig. 1. Distance Data between Tourist Objects

The data above are initial data showing information on the distance values between one point to each other obtained at the data collection stage. The dataset will be used in the testing process, as well as to see if a particular algorithm can find a solution so that a point can have a smaller distance value while also finding the shortest route between the starting point and the destination point. The data consists of 19 location points with each distance value written with a unit of kilometers. This study also aims to optimize the travel of tourists in the Sumba region. This study wanted to contribute to facilitating the determination of tourist attractions to be visited, especially when talking about road infrastructure conditions and sometimes there is a potential route mismatch that certain applications can produce when they want to know the route on their way to a location.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

The algorithms used in this study include Dijkstra, Bellman-Ford, Floyd-Warshall, Johnson, and Ant Colony algorithms. Dijkstra's algorithm is an algorithm that is used to find the shortest distance or route from the initial node to the final node in a weighted graph. Dijkstra's algorithm is that it is the most important and useful algorithm for optimal solutions in a large class of shortest path problems [14]. Bellman-Ford algorithm can also be used to solve problems related to the shortest path problem where this algorithm can work even though there are negative edge weight values. This algorithm itself is a refinement of Dijkstra's algorithm. Bellman-Ford algorithm offers a dynamic solution for all nodes in a graph to determine the minimum route with edge weights that can be negative, but still does not contain negative cycles [23]. Floyd-Warshall algorithm is a dynamic algorithm and is often used to determine the shortest route between all points on a directed graph without negative cycles. This algorithm is presented to find solutions to the shortest-path problem between certain fixed nodes and other related nodes [16]. Johnson's algorithm is also one of the commonly used algorithms to solve the shortest route problem. This algorithm is a combination of the Bellman-Ford and Dijkstra algorithms. Johnson's algorithm, among other things, can be used on negative weighted graphs. [6] In addition, Johnson's algorithm is also an appropriate solution method for solving scheduling problems such as the research by M. Okwu and I. Emovon [7] and by M. Redi and M. Ikram [8]. Ant Colony Optimization (ACO) algorithm is an optimization algorithm that uses probabilistic techniques and is used to solve computational problems and find optimal paths. The Ant Colony Optimization strategy will choose the shortest path based on the path most frequently traveled by ants, using mechanisms that mimic behavior or social strategies that exist in nature [31]. In testing the five algorithms, it is also assessed based on the type of shortest-path, regarding which algorithm is suitable for use in finding solutions based on a particular type of shortest-path. There are several types such as single-pair, then single-source, single-destination, and all-pairs shortest path problem.

The parameter used as an indicator of the assessment of an algorithm being tested. The first is related to the calculation of the shortest distance, then also related to complexity, execution time, and the final result of the distance that being traveled. Calculation of the shortest distance is to find the path between the initial node and the final node so that the number of edges with minimum weight is found. For example, that a node can be connected directly to another node through one edge or a series of edges. So this research has observed visually, that there may be shorter 'distances' between some vertices and others in the graph [17]. Complexity is a parameter that provides information regarding how complicated aspects of the algorithm used are. Complexity here actually has a broader definition, in the sense that this parameter can cover several aspects such as the complexity of memory usage, space, running time [18][19][20], and so on. Execution time is used to find out how efficient and effective an algorithm is in terms of time to process data. In research by Abusalim, et al. [27]Studying that the number of nodes that make up the graph has an influence on the fast or slow running time when executing certain algorithms. Then, for the final result is to talk about how the output is generated in the program.

The research was carried out through several stages, such as receiving input in the form of data used, then proceeding with the determination of the starting point and destination point used during the process. Implementation was carried out on five algorithms and through these implementations, followed by an analysis based on the parameters that have been compiled to obtain results. The following Fig. 2 is to describe the flow of the research through a flowchart.

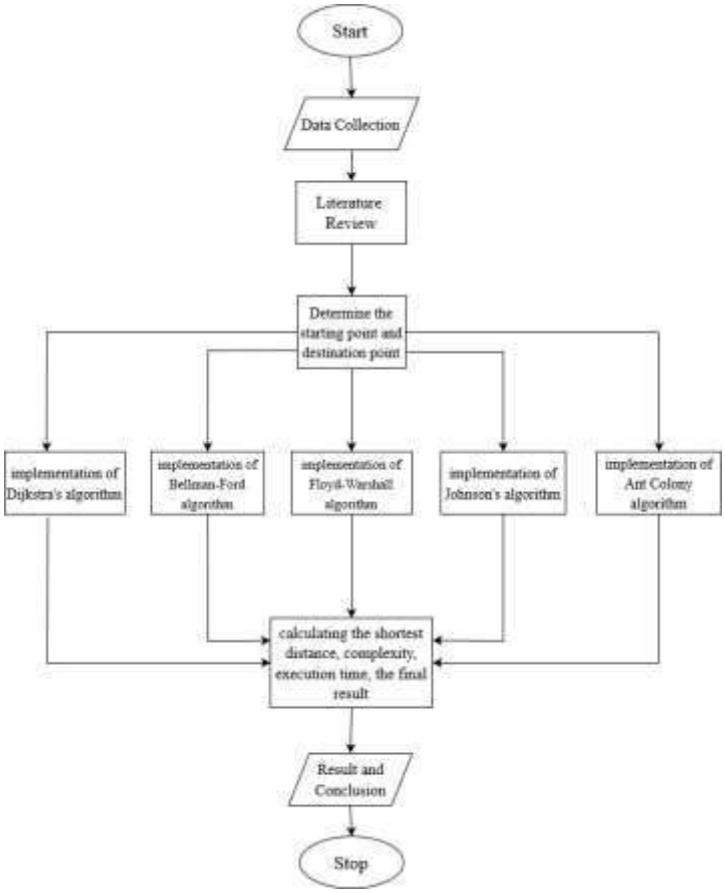


Fig. 2. Distance Data between Tourist Objects

Based on the flowchart, the first step is to collect the data. In this case, the dataset that will be used in the research. The stage continues with a literature review regarding each algorithm that will be used as material in testing. And also from this dataset, the starting point and destination point are determined as a starting point for starting the test. After determining the point of reference, the implementation is

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

carried out with five algorithms with the help of the Python program. From the tests carried out, then the results and comparisons are analyzed referring to the parameters that have been compiled, related to the calculation of the shortest distance, complexity, execution time, and the final result of the distance traveled. Then the stage is continued by determining the results and conclusions obtained through the tests that have been carried out.

III. RESULT AND ANALYSIS

The results and analysis of the shortest-path search calculation use 19 data from the Sumba Region nodes with five algorithms, including the Dijkstra, Bellman-Ford, Floyd-Warshall, Johnson, and Ant Colony algorithms. From these algorithms, modeling is implemented using Python programming language with Spyder application (Python 3.7). For the explanation as follows.

A. Dijkstra's Algorithm

1) Calculation of the shortest distance

The steps for calculating Dijkstra's algorithm use an approach in determining the shortest path starting from a TMC Airport node to several other destination nodes in a graph, including the following :

(a) The first step is to mark all the nodes to be visited,
(b) Mark the selected initial node with current distance 0, and other nodes with infinity "∞",
(c) Then update the initial node as the current node,
(d) For the current node, analyze all unvisited neighbor nodes and measure their distance by adding the current distance from the current node to the edge weight connecting the neighbor node and the current node,
(e) Compare the distance measured recently with the currently assigned distance to the neighboring nodes and take that as the new current distance from the neighboring node,
(f) After that, consider all unvisited neighbors of the current node, mark the current node as visited,
(g) If the marked destination node is visited then stops, then the algorithm has ended. If not, select the unvisited node marked with the closest distance, fix it as the current new node, and repeat the process again from step (d).

2) Complexity

The complexity when running Dijkstra's algorithm program in Python is 70 lines.

3) Execution Time

Dijkstra's Algorithm is carried out 5 times to run the Python language program, so that it can better determine the amount of time the program is running, and can take the average time. The following Table II is for the test results related to the execution time of the program.

TABLE II. DIJKSTRA'S ALGORITHM EXECUTION TIME

Execution	Time
First Execution	0.0013222999999982221 seconds
Second Execution	0.0003200000000020964 seconds
Third Execution	0.00033470000000335176 seconds
Fourth Execution	0.0004951999999960321 seconds
Fifth Execution	0.0007295999999996639 seconds
Average Time	0.0006403599999999873252 seconds

For the first execution, the program takes 0.0013222999999982221 seconds to run the algorithm to the stage of displaying the results obtained, which in this case is route information. This is also the same for the second execution up to the fifth execution stage. From the Table II, it can be seen that the time used to run the algorithm can be said to be very fast and does not have a significant difference in each process. The average execution time obtained was 0.0006403599999999873252 seconds, meaning that programs that implement the Dijkstra's Algorithm can be executed by the console lightly and quickly.

4) Final Result

The results of the shortest distance that being traveled are calculated from TMC Airport to all destinations in the Sumba Region, with the results of the Table III data as follows.

TABLE III. THE RESULT OF DIJKSTRA'S ALGORITHM

No.	Distance	Path Traversed	Value (km)
1	TMC Airport	TMC Airport – TMC Airport	0
2	Kita Beach	TMC Airport – Pantai Kita	19
3	Kabisu Lobo Oro Site	TMC Airport – Kabisu Lobo Oro Site	9
4	Lendongara Hill	TMC Airport – Lendongara Hill	10
5	Waikelo Beach	TMC Airport – Waikelo Beach	8.3
6	Kawona Beach	TMC Airport – Kawona Beach	8.7
7	Karakat Indah Beach	TMC Airport – Karakat Indah Beach	22
8	Waikuri Lagoon	TMC Airport – Waikuri Lagoon	42
9	Mandorak Beach	TMC Airport – Mandorak Beach	40
10	Tanjung Karoso Beach	TMC Airport – Tanjung Karoso Resort – Tanjung Karoso Beach	48.7 (from; 51)
11	Tanjung Karoso Resort	TMC Airport – Tanjung Karoso Resort	43
12	Rotenggaro Beach	TMC Airport – Rotenggaro Beach	47
13	Bawana Beach	TMC Airport – Bawana Beach	58
14	Watu Malandong Beach	TMC Airport - Watu Malandong Beach	63
15	Sumba Culture House	TMC Airport – Sumba Culture House	6.5
16	Praijing Village	TMC Airport – Praijing Village	43

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

17	Lailiang Beach	TMC Airport – Praijing Village – Lailiang Beach	62 (from; 66)
18	Rua Beach	TMC Airport – Rua Beach	59
19	Mata Yangu Waterfall	TMC Airport – Mata Yangu Waterfall	65

Through the implementation of the program, it was found that Dijkstra's algorithm was able to generate a path from the starting point of TMC Airport to a particular location with a more optimal distance value, such as on the TMC Airport route to Tanjung Karoso Beach, the algorithm can provide route information that can be passed along the route. The total distance value is 48.7 km from the initial data of 51 km. Then, the following example on the TMC Airport path to Lailiang Beach obtained a route with a total distance value of 62 km, is smaller than the initial data of 66 km. Dijkstra's algorithm also computes other destinations until results are shown in Table III.

B. Bellman-Ford Algorithm

1) Calculation of the shortest distance

Calculations with the Bellman-Ford algorithm can be used to find the shortest distance between source nodes to each node. The Bellman-Ford algorithm can be used to find the shortest route solution of the all-pairs type, but in terms of implementation, the program is executed in the form of a single-source shortest path problem, namely from a certain node to all other nodes [24]. Calculation results are displayed with good and accurate results.

2) Complexity

For the complexity of the Bellman-Ford algorithm, it is considered quite complex with 409 lines of source code.

3) Execution Time

Based on 5 experiments on the Python program, the following data obtained for testing results on program execution as follows.

TABLE IV. BELLMAN-FORD ALGORITHM EXECUTION TIME

Execution	Time
First Execution	0.0000004 seconds
Second Execution	0.0000009 second
Third Execution	0.0000005 seconds
Fourth Execution	0.0000006 seconds
Fifth Execution	0.0000004 seconds
Average Time	0.00000056 seconds

Results on testing show that programs with Bellman-Ford algorithms can run at very fast execution times, and can display fast execution results. Through these tests, the average execution time was 0.000056 seconds only.

4) Final Result

For the final results to be represented through rows and columns, the following is a Table V for variable initialization for each tourist attraction based on the dataset used.

TABLE V. VARIABLE INITIALIZATION

Variable Initialization	Sumba Regional Destination	Variable Initialization	Sumba Regional Destination
1	TMC Airport	11	Tanjung Karoso Resort
2	Kita Beach	12	Rotenggaro Beach
3	Kabisu Lobo Oro Site	13	Bawana Beach
4	Lendongara Hill	14	Watu Malandong Beach
5	Waikelo Beach	15	Sumba Culture House
6	Kawona Beach	16	Praijing Village
7	Karakat Indah Beach	17	Lailiang Beach
8	Waikuri Lagoon	18	Rua Beach
9	Mandorak Beach	19	Mata Yangu Waterfall
10	Tanjung Karoso Beach		

Based on the table above, the final results in the Fig. 3 below are adjusted for the variable initialization in rows and columns with their respective distance values.

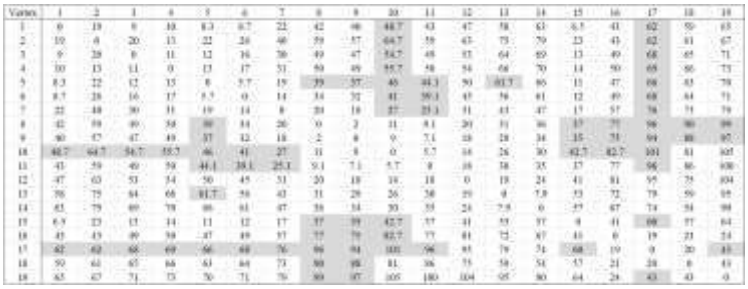


Fig. 3. The result of Bellman-Ford algorithm

The rows and columns in the table are sorted based on the initialization of the data that has been compiled. The Bellman-Ford algorithm works by carrying out calculations on each row and column and obtaining changes at several points with a smaller total distance traveled, changes are highlighted by the colored part. It was found that from all the existing data, the Bellman-Ford algorithm was able to analyze the data well and provide the shortest route solution correctly at each point.

So, by implementing the Bellman-Ford algorithm with a complexity of 409 lines of source code and several tests, it was found that the Bellman-Ford algorithm was considered effective in being able to find the minimum distance value of each node in the dataset. However, the efficiency of the program line is considered sufficient, because when compared to the other algorithms, the application of the source code to the Bellman-Ford algorithm is executed in large numbers.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

C. Floyd-Warshall Algorithm

1) Calculation of the shortest distance

Calculations on the Floyd-Warshall algorithm are indeed more suitable for use for the all-pairs type [16], calculations are carried out using a Python program with 19 nodes which are then written in matrix form. From the implementation, can obtain that the calculations can run well and can find the most optimal solution, in this case the shortest route between all pairs of vertices.

2) Complexity

For the complexity of the program lines used, there are as many as 59 lines of source code.

3) Execution Time

Based on 5 experiments on the Python program, the following data obtained for testing results on Table VI program execution as follows.

TABLE VI. FLOYD-WARSHALL ALGORITHM EXECUTION TIME

Execution	Time
First Execution	0.0000009 seconds
Second Execution	0.0000005 seconds
Third Execution	0.0000004 seconds
Fourth Execution	0.0000006 seconds
Fifth Execution	0.0000005 seconds
Average Time	0.00000058 seconds

Table VI depicts the results of each program execution process that applies the Floyd-Warshall algorithm. It can be seen that the execution time used to run the program is also very fast with the average execution time from 5 tests being 0.00000058 seconds.

4) Final Result

The final result of applying the algorithm based on each row and column is initialized as follows.

TABLE VII. VARIABLE INITIALIZATION

Variable Initialization	Sumba Regional Destination	Variable Initialization	Sumba Regional Destination
1	TMC Airport	11	Tanjung Karoso Resort
2	Kita Beach	12	Rotenggaro Beach
3	Kabisu Lobo Oro Site	13	Bawana Beach
4	Lendongara Hill	14	Watu Malandong Beach
5	Waikelo Beach	15	Sumba Culture House
6	Kawona Beach	16	Praijing Village
7	Karakat Indah Beach	17	Lailiang Beach
8	Waikuri Lagoon	18	Rua Beach
9	Mandorak Beach	19	Mata Yangu Waterfall
10	Tanjung Karoso Beach		

The table above represents the variable initialization for each tourist attraction, then the results will be made in the form of rows and columns represented by the initialization as in Table VII. The following is a Fig. 4 for the results after applying the Floyd-Warshall algorithm to the dataset used.

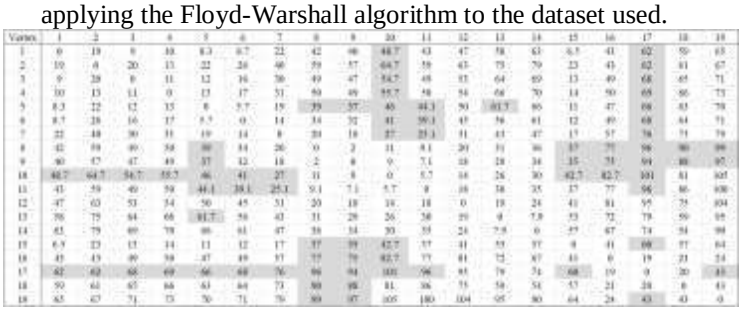


Fig. 4. The result of Floyd-Warshall algorithm

Rows and columns are also written based on the initialization of data that represents each location. Calculations are carried out at each point and obtained using the Floyd-Warshall algorithm calculation method which is also able to display route results with more optimal distance values at several points compared to the initial data, depicted in colored sections. It was found that this algorithm can provide quite significant changes to the problem of finding the shortest route.

So, by applying the Floyd-Warshall algorithm with a source code line complexity of 59 lines and several tests, it can be obtained that the Floyd-Warshall algorithm is considered effective and efficient in finding the minimum distance value for each node in the dataset. This algorithm is also very effective for use in the problem of determining the shortest route of the all-pairs type.

D. Johnson's Algorithm

1) Calculation of the shortest distance

The stages of the Johnson algorithm calculation steps use an approach in determining the shortest path of all pairs of paths where the destination vertices in a graph are among others [32], as follows.

(a) In the first step, add a vertex S to all path points on graph G by giving it a set of 0.

(b) Run the Bellman-Ford algorithm on G' with node S as the source by finding the minimum weight, then calculating the heuristic or h[v-1] for every number of nodes in graph G.

(c) When calculating h[…], then recalculate the edges of the graph using the formula: w(u, v) = w(u, v) + h[u] – h[v].

(d) If the iteration is complete, then delete the added S node and all weights are now positive on the graph. If the calculation iteration has not been completed, repeat step (c).

(e) Run Dijkstra's shortest path algorithm for each node as the source and calculate the shortest route (u,v).

(f) If the iteration is complete, the shortest route u and v will

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

be found. If the calculation iteration has not been completed, repeat step (e).

2) Complexity

The complexity of source code when running Johnson's Algorithm program in Python is 86 lines.

3) Execution Time

The time to run the Johnson's algorithm in Python program implemented 5 times, so that it can better determine the amount of time the program is running and can take the average time. The following Table VIII is the result of the execution of the program.

TABLE VIII. JOHNSON'S ALGORITHM EXECUTION TIME

Execution	Time
First Execution	0.013958700000000768 seconds
Second Execution	0.01639690000000016 seconds
Third Execution	0.013379900000000333 seconds
Fourth Execution	0.019444300000003523 seconds
Fifth Execution	0.01641610000000071 seconds
Average Time	0.0159191800000010988 seconds

Table VIII explains the test results of running the program with the Johnson's algorithm and is obtained from the first execution trial to the last trial stage. The time required to run the program has a very slight time difference, the time used is only around 0.01 seconds at each stage. With this time, it can be seen that the execution time for this algorithm is also very fast and from the five experiments, the average time required to run the program and display the results is 0.0159191800000010988 seconds on the console.

4) Final Result

The results are managed by Johnson's algorithm which then finds the results of the shortest distance weights with those that have been modified from vertex-1 to vertex-19 to several other vertices. The results of the distance when running the Johnson's algorithm program with the Python programming language, found 72 shorter path data from each node that was originally searched for all destination nodes.

So, in the data Johnson's algorithm looks for the distance from vertex-1 and vertex-19 to all destination data, by obtaining 86 lines of program complexity and 0.0159191800000010988 program time in seconds. The results obtained found 72 shorter path data from each initial node that was searched for all destination nodes.

E. Ant Colony Algorithm

1) Calculation of the shortest distance

The steps for calculating the Ant Colony algorithm use an approach in determining the shortest path starting from a TMC Airport node to every one pair of destination nodes on a graph, including the following:

(a) The first step, set the value of Ant Colony optimization on the initial node. With conditions, nodes that have been visited are included in the tabu list, so they will not be visited again.

(b) After performing probabilistic node calculations, then select the next node to be assigned the Ant Colony value.

(c) Move to the next node and the visited node becomes tabu list.

(d) View all visited nodes, if not repeat step (b).

(e) Then it will record the length of the side or the distance of the route taken, and delete the entire tabu list.

(f) Determine the shortest route from the current nodes and update the pheromone, if the Ant Colony travels one full route back to the beginning.

(g) If the limit for the number of Ant Colony is the maximum iteration that has been reached and will be completed, thus determining the shortest route. If the maximum iteration has not been completed, repeat step (a).

2) Complexity

The complexity of source code when running the Python language of Ant Colony algorithm program is 84 lines.

3) Execution Time

When running the Python language program, the Ant Colony algorithm implemented 5 times, so that it can better determine the amount of time the program is running and can take the average time. In testing the execution of the Python language program, it can be seen in the Table IX below the results of testing the program.

TABLE IX. ANT COLONY ALGORITHM EXECUTION TIME

Execution	Time
First Execution	1.2428267000000233 seconds
Second Execution	0.8022379999999885 seconds
Third Execution	0.8272481999999854 seconds
Fourth Execution	1.2401753999999983 seconds
Fifth Execution	0.8539233000000195 seconds
Average Time	0.993282320000003 seconds

In contrast to several algorithms that have been tested previously, the execution time for running the Ant Colony program actually varies. From Table IX it can be seen that in the first and fourth experiments, the time used to run the program was more than 1 second. However, it is different for other experiments, which only take less than 1 second. Even so, each stage still has a slight time difference and is still considered very fast to be able to execute an Ant Colony program according to the program specifications used. Also obtained was the average program execution time of 0.993282320000003 seconds.

4) Final Result

In this algorithm, tests were carried out on 19 ants from 19 data on the destinations in the Sumba area used. Then on the 19 data used variable initialization is applied, as in the following Table X.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

TABLE X. VARIABLE INITIALIZATION

Variable Initialization	Sumba Regional Destination	Variable Initialization	Sumba Regional Destination
1	TMC Airport	11	Tanjung Karoso Resort
2	Kita Beach	12	Rotenggaro Beach
3	Kabisu Lobo Oro Site	13	Bawana Beach
4	Lendongara Hill	14	Watu Malandong Beach
5	Waikelo Beach	15	Sumba Culture House
6	Kawona Beach	16	Praijing Village
7	Karakat Indah Beach	17	Lailiang Beach
8	Waikuri Lagoon	18	Rua Beach
9	Mandorak Beach	19	Mata Yangu Waterfall
10	Tanjung Karoso Beach		

Then, testing was carried out and the shortest distance was found from the iteration of the ant in the form of a circuit. The following Table XI represents the test results data with the Ant Colony algorithm.

TABLE XI. ANT COLONY ALGORITHM RESULTS

Ant Iteration	Path Traversed	Value (km)
1st	1 - 15 - 3 - 4 - 2 - 16 - 17 - 18 - 19 - 13 - 14 - 12 - 10 - 8 - 9 - 11 - 7 - 6 - 5 - 1	385.3
2nd	1 - 15 - 3 - 2 - 4 - 5 - 6 - 7 - 8 - 9 - 11 - 10 - 12 - 13 - 14 - 18 - 17 - 16 - 19 - 1	342.0
3rd	1 - 15 - 5 - 6 - 7 - 9 - 8 - 11 - 10 - 12 - 14 - 13 - 17 - 16 - 19 - 18 - 2 - 4 - 3 - 1	376.0
4th	1 - 15 - 5 - 6 - 7 - 9 - 8 - 11 - 10 - 12 - 13 - 14 - 19 - 16 - 18 - 17 - 3 - 4 - 2 - 1	382.0
5th	1 - 15 - 6 - 5 - 3 - 4 - 2 - 18 - 17 - 16 - 19 - 13 - 14 - 10 - 11 - 12 - 8 - 9 - 7 - 1	402.0
6th	1 - 6 - 5 - 3 - 4 - 2 - 16 - 17 - 18 - 19 - 7 - 15 - 9 - 8 - 11 - 10 - 12 - 13 - 14 - 1	428.0
7th	1 - 6 - 5 - 15 - 3 - 4 - 2 - 16 - 19 - 18 - 17 - 13 - 14 - 12 - 10 - 11 - 9 - 8 - 7 - 1	374.0
8th	1 - 5 - 6 - 7 - 19 - 16 - 17 - 18 - 14 - 13 - 12 - 9 - 8 - 10 - 11 - 15 - 4 - 2 - 3 - 1	380.0
9th	1 - 15 - 6 - 5 - 3 - 4 - 2 - 11 - 10 - 9 - 8 - 7 - 12 - 13 - 14 - 18 - 17 - 16 - 19 - 1	395.0
10th	1 - 3 - 4 - 2 - 5 - 6 - 7 - 8 - 9 - 11 - 10 - 12 - 13 - 14 - 18 - 17 - 16 - 19 - 15 - 1	337.5
11th	1 - 15 - 3 - 4 - 2 - 6 - 5 - 9 - 10 - 11 - 12 - 13 - 14 - 17 - 18 - 16 - 19 - 7 - 8 - 1	452.0
12th	1 - 5 - 6 - 15 - 3 - 4 - 2 - 16 - 19 - 7 - 8 - 9 - 11 - 10 - 12 - 14 - 13 - 17 - 18 - 1	447.0
13th	1 - 15 - 3 - 4 - 2 - 5 - 6 - 7 - 9 - 8 - 10 - 11 - 12 - 13 - 14 - 16 - 17 - 18 - 19 - 1	380.0
14th	1 - 15 - 3 - 5 - 6 - 7 - 9 - 8 - 11 - 10 - 12 - 14 - 13 - 18 - 17 - 16 - 19 - 2 - 4 - 1	343.0
15th	1 - 15 - 6 - 5 - 7 - 8 - 9 - 11 - 10 - 12 - 14 - 13 - 19 - 16 - 17 - 18 - 3 - 4 - 2 - 1	389.0
16th	1 - 5 - 6 - 7 - 8 - 9 - 11 - 10 - 12 - 14 - 13 - 19 - 18 - 17 - 16 - 2 - 15 - 3 - 4 - 1	385.0
17th	1 - 15 - 7 - 6 - 5 - 3 - 4 - 2 - 18 - 17 - 16 - 19 - 13 - 14 - 12 - 10 - 11 - 9 - 8 - 1	400.0
18th	1 - 15 - 5 - 6 - 7 - 9 - 8 - 11 - 10 - 12 - 13 - 14 - 18 - 17 - 19 - 16 - 3 - 4 - 2 - 1	353.0
19th	1 - 5 - 6 - 15 - 7 - 8 - 9 - 11 - 10 - 12 - 13 - 14 - 18 - 16 - 17 - 19 - 3 - 4 - 2 - 1	377.0

The display of running iterations is accompanied by path traversed, and how many total distance values are taken on the corresponding route. In this Ant Colony data algorithm, find the distance from TMC Airport by visiting all 18 destinations and then returning to TMC Airport, obtaining 84 lines of program complexity and 0.993282320000003 program time in seconds. The most optimal route results obtained are as far as 337.5 km, with the following routes: TMC Airport - Kabisu Lobo Oro Site - Lendongara Hill - Kita Beach - Waikelo Beach - Kawona Beach - Karakat Indah Beach - Waikuri Lagoon - Mandorak Beach - Tanjung Karoso Resort - Tanjung Beach Karoso - Rotenggaro Beach - Bawana Beach - Watu Malandong Beach - Rua Beach - Lailiang Beach - Praijing Village - Mata Yangu Waterfall - Sumba Cultural House - TMC Airport.

F. Discussion

Through the results described, the analysis performed on the performance of each algorithm on the test. Through testing algorithms judged based on the parameters that were indicators of this study, each algorithm had its own advantages and disadvantages. As such, the Dijkstra's algorithm can provide the most optimal shortest route results, but it can only effectively work to solve single-source and single-destination-type problems. The Bellman-Ford algorithm also has its own advantages, such as being able to hold negative weight values on graphs, but in this study it has the most complexity on source code among other algorithms. The Bellman-Ford algorithm can be used for almost the same path as Dijkstra in single-source and single-destination types. The Floyd-Warshall algorithm is more effective for use in all-pairs path types, and can provide the most optimal value at any point. Johnson's algorithm is a combination of Dijkstra's algorithm and Bellman-Ford's algorithm. Meanwhile, for the Ant Colony algorithm, it works with the longest execution time among other algorithms, but after analysis it turns out that the Ant Colony algorithm has a high degree of effectiveness in resolving the problem of route lookup for all nodes.

Speaking of which algorithms are best suited to implement, it is also necessary to see how the graph picture will be examined and which solutions will be sought for the problem. Thus, the search can be performed more effectively and efficiently and have the most optimal distance value, in the sense that it does not focus only on which route with the least weight value. As described, Dijkstra's and Bellman-Ford algorithms are suitable for use in single-source and single-destination graphs, Floyd-Warshall algorithms for all-pairs type, Johnson's algorithm can also be used for all-pairs type searches and Ant Colony can be used to find routes that can visit all destinations at once at a time well and the minimum possible distance value.

IV. CONCLUSION

Based on testing of the five algorithms and adjusted to the parameters that have been compiled, it is found that each

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

algorithm has its own type. Dijkstra's algorithm can be used to calculate the shortest route for single-source and single-destination types. It's the same as the Bellman-Ford algorithm, except that the Bellman-Ford algorithm can be used at the same time on graphs that have negative weight values. Meanwhile, the Floyd-Warshall algorithm is suitable for use on the all-pairs type. For Johnson's algorithm it can be used to determine the shortest path from all pairs of paths where the destination node is on a graph, and Ant Colony for calculating from a node to every one pair of destination nodes. Through the implementation of the Python program, it was found that there was a change in the dataset used, namely the Southwest Sumba data, to become data with a more optimum distance value, in this case the minimum distance value from a starting point to a destination point.

Through the Python program that has been researched, the complexity of each of the various algorithms is obtained. Dijkstra's algorithm has 70 lines of source code, Bellman-Ford has 409 source codes, Floyd-Warshall has 59 source codes, Johnson has 86 source codes, and Ant Colony has 84 source codes. Thus, the algorithm with the lowest complexity is owned by Floyd-Warshall algorithm.

Then, Dijkstra's algorithm with execution time running in the Python program obtained an average time of 0.00064035999999873252 seconds, Bellman-Ford's algorithm for 0.00000056 seconds, Floyd-Warshall's algorithm for 0.00000058 seconds, Johnson's algorithm for 0.0159191800000010988 seconds, and Ant Colony's algorithm for 0.993282320000003 seconds. Thus, the Bellman-Ford and Floyd-Warshall algorithms are considered to have the fastest execution time.

Based on all the parameters that have been examined, it can be concluded that the Ant Colony algorithm has high effectiveness for solving the most efficient route finding for all nodes. And the Floyd-Warshall algorithm is considered to be the most effective algorithm in finding all-pairs type routes, and is able to find the most optimal distance value between one node and another node. Also, Bellman-Ford has the fastest execution time when a Python program is run.

Through such implementations it is also obtained about how the algorithm performs if it is based through certain parameters. Each algorithm has its own advantages and disadvantages when implemented. The journal focuses on how to test the subject, then provides a new perspective on which algorithms have the most efficient performance with good accuracy when implemented on specific graph models and path models.

ACKNOWLEDGMENT

Collate Thanks to LPPM Darma Catholic University of Cendika for its contribution to support this research in the form of internal grant funding.

References

[1] B. S. N Assistant professor GFGC, "A Study on Graph Coloring," *Int J Sci Eng Res*, vol. 8, no. 5, 2017, [Online]. Available: http://www.ijser.org

[2] Tirastittam Pimploi and Waiyawuththanapoom Phutthiwat, "Public Transport Planning System by Dijkstra Algorithm Case Study Bangkok Metropolitan Area," *International Journal of Computer and Information Engineering*, vol. 08, 2014.

[3] M. Iqbal, K. Zhang, S. Iqbal, and I. Tariq, "A Fast and Reliable Dijkstra Algorithm for Online Shortest Path," *International Journal of Computer Science and Engineering*, vol. 5, no. 12, pp. 24–27, Dec. 2018, doi: 10.14445/23488387/IJCSE-V5I12P106.

[4] Z. Jiang, V. Sahasrabudhe, A. Mohamed, H. Grebel, and R. Rojas-Cessa, "Greedy algorithm for minimizing the cost of routing power on a digital microgrid," *Energies (Basel)*, vol. 12, no. 16, Aug. 2019, doi: 10.3390/en12163076.

[5] A. Kejriwal and A. Temrikar, "Graph Theory and Dijkstra's Algorithm:A solution for Mumbai's BEST buses," *The International Journal of Engineering and Science (IJES)* ||, pp. 23–42, 2019, doi: 10.9790/1813-0810014047www.theijes.com.

[6] Umamaheswari K., Pavithra A., Srinivashini S., and Subipriya S., "Tackling a Shortest Path Having a Negative Cycle by using Johnson_s Calculation," *TEST Engineering & Management*, vol. 81, 2019.

[7] M. Okwu and I. Emovon, "Application of Johnson's algorithm in processing jobs through two-machine system," *Journal of Mechanical and Energy Engineering*, vol. 4, no. 1, pp. 33–38, Aug. 2020, doi: 10.30464/jmee.2020.4.1.33.

[8] M. Redi and M. Ikram, "Dimension Reduction and Relaxation of Johnson's Method for Two Machines Flow Shop Scheduling Problem," *Sultan Qaboos University Journal for Science [SQUJS]*, vol. 25, no. 1, p. 26, Jun. 2020, doi: 10.24200/squjs.vol25iss1pp26-47.

[9] C. Dalfó and M. À. Fiol, "Graphs, friends and acquaintances," *Electronic Journal of Graph Theory and Applications*, vol. 6, no. 2, pp. 282–305, 2018, doi: 10.5614/ejgta.2018.6.2.8.

[10] X. Z. Wang, "The Comparison of Three Algorithms in Shortest Path Issue," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Oct. 2018. doi: 10.1088/1742-6596/1087/2/022011.

[11] A. A. B. A. Homaid, A. R. A. Alsewari, K. Z. Zamli, and Y. A. Alsariera, "Adapting the elitism on greedy algorithm for variable strength combinatorial test cases generation," *IET Software*, vol. 13, no. 4, pp. 286–294, Aug. 2019, doi: 10.1049/iet-sen.2018.5005.

[12] A. M. Benjamin, S. A. Rahman, E. M. Nazri, and E. A. Bakar, "Developing A Comprehensive Tour Package Using An Improved Greedy Algorithm With Tourist Preferences," 2019. [Online]. Available: https://www.researchgate.net/publication/335541687

[13] M. Elizabeth, B. Gani, M. A. Safitri, C. Lusiana, and M. Dewi, "Real Time Public Transportation Navigator System in Jakarta by Using Greedy Best First Search Algorithm," 2018, [Online]. Available: http://www.emarketer.com/Article/Asia-Pacific-

[14] S. Orhani, "Finding the Shortest Route for Kosovo Cities Through Dijkstra's Algorithm," *Middle European Scientific Bulletin*, vol. 25, 2022, [Online]. Available: https://www.researchgate.net/publication/361443814

[15] R. Chen, "Dijkstra's Shortest Path Algorithm and Its Application on Bus Routing," 2022.

[16] V. Sakharov, S. Chernyi, S. Saburov, and A. Chertkov, "Automatization Search for the Shortest Routes in the Transport Network Using the Floyd-warshell Algorithm," in *Transportation Research Procedia*, Elsevier B.V., 2021, pp. 1–11. doi: 10.1016/j.trpro.2021.02.041.

[17] N. Syuhada, M. Pazil, N. Mahmud, S. H. Jamaluddin, N. Farasyaqirra, and B. Mustafa, "Shortest Path from Bandar Tun Razak to Berjaya Times Square using Dijkstra Algorithm," 2020. [Online]. Available: https://jcrinn.com

[18] G. Deepa, P. Kumar, A. Manimaran, K. Rajakumar, and V. Krishnamoorthy, "Dijkstra Algorithm Application: Shortest Distance between Buildings," *International Journal of Engineering & Technology*, vol. 7, no. 4.10, p. 974, Oct. 2018, doi: 10.14419/ijet.v7i4.10.26638.

[19] Sari I. P., Fahroza M. F., Mufit M. I., and Qathrunad I. F., "Implementation of Dijkstra's Algorithm to Determine the Shortest Route in a City," *Journal of Computer Science, Information Technologi and Telecommunication Engineering*, vol. 02, pp. 134–138, Mar. 2021, doi: 10.30596/jcositte.v2i1.6503.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

[20] O. Khaing, H. Htight Wai, and E. Ei Myat, "Using Dijkstra's Algorithm for Public Transportation System in Yangon Based on GIS," 2018. [Online]. Available: www.ijsea.com442

[21] T. Sari, A. S. Zain, and A. N. Handayani, "Application Of DIJKSTRA Algorithm For Network Troubleshooting In SMK Telkom Malang," 2019.

[22] O. K. Sulaiman, A. M. Siregar, K. Nasution, and T. Haramaini, "Bellman Ford algorithm - In Routing Information Protocol (RIP)," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Apr. 2018. doi: 10.1088/1742-6596/1007/1/012009.

[23] P. Gupta and V. Pathak, "A Minimum Spanning Tree-based Routing Technique of FAT Tree for Efficient Data Center Networking," *Mathematical Statistician and Engineering Applications*, vol. 71, no. 1, Jan. 2022, doi: 10.17762/msea.v71i1.44.

[24] F. Mukhlif and A. Saif, "Comparative Study On Bellman-Ford And Dijkstra Algorithms," 2020. [Online]. Available: https://www.researchgate.net/publication/340790429

[25] Botsis D. and Panagiotopoulos E., "Determination of the shortest path in the university campus of Serres using the Dijkstra and Bellman-Ford algorithms," 2020.

[26] A. Manan, S. Imran, and A. Lakyari, "Single source shortest path algorithm Dijkstra and Bellman-Ford Algorithms: A Comparative study," *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND EMERGING TECHNOLOGIES (IJCET)*, vol. 3, no. 2, pp. 25–28, 2019, [Online]. Available: https://ijcet.salu.edu.pk

[27] S. W. G. Abusalim, R. Ibrahim, M. Zainuri Saringat, S. Jamel, and J. Abdul Wahab, "Comparative Analysis between Dijkstra and Bellman-Ford Algorithms in Shortest Path Optimization," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Sep. 2020. doi: 10.1088/1757-899X/917/1/012077.

[28] O. Yu Lavlinskaya, T. V. Kurchenkova, and O. V. Kuripta, "Shortest path algorithm for graphs in instances of semantic optimization," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, May 2020. doi: 10.1088/1742-6596/1479/1/012036.

[29] Rachmad A., Syarief M., Rochman E. M. S., and R. G. Husni, "Ant colony optimization model for determining the shortest route in Madura-Indonesia tourism places," *Journal of Mathematical and Computational Science*, 2021, doi: 10.28919/jmcs/7078.

[30] D. R. Anamisa, A. Rachmad, and E. M. S. Rochman, "Ant Colony System Based Ant Adaptive for Search of the Fastest Route of Tourism Object Jember, East Java," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, 2020. doi: 10.1088/1742-6596/1477/5/052053.

[31] T. Herianto, "Implementation of the Ant Colony System Algorithm in the Lecture Scheduling Process," *Instal : Jurnal Komputer*, vol. 12, 2020.

[32] I. G. Ivanov, G. V. Hristov, and V. D. Stoykova, "Algorithms for optimizing packet propagation latency in software-defined networks," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Feb. 2021. doi: 10.1088/1757-899X/1031/1/012072.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1756, Copyright ©2023
Submitted : May 24,2023, Revised : August 8, 2023, Accepted : Oktober 1, 2023, Published : November 6, 2023

Systematic Literature Review: Metode Machine Learning dalam Klasifikasi Emosi pada Data Tekstual

Putu Widyantara Artanta Wibawa[1], Cokorda Pramartha[2]*
Program Studi Informatika[1]
Net-Centric Computing Research Lab[2]
Fakultas MIPA, Universitas Udayana
Jimbaran, Bali, Indonesia
putuwaw973@gmail.com[1] , cokorda@unud.ac.id[2]

***Abstract*— Emotions are a person's response to an event. Emotions can be expressed verbally or nonverbally. Over time people can express their emotions through social media. Considering that emotion is a reflection of society's response, it is important to classify emotions in society to find out the community's response as information for consideration in decision-making. This study is aimed to identify and analyze the datasets, methods, and evaluation metrics that are being used in the classification of emotional texts in textual data from research data from 2013 to 2022. Based on the inclusion and exclusion design in selecting literature, a total of 50 kinds of literature were used in extracting and synthesizing data. Analysis of the data shows that out of 50 pieces of literature, there are 36 works of literature that use public datasets while 14 kinds of literature use private datasets. In the method of developing models for classifying, the SVM and Naive Bayes models are the most popular among the other models. In evaluating the model, the F-measure or F1-score metric is the most widely used metric compared to other metrics. There are three main contributions identified in this study, namely methods, models, and evaluation**

Keywords*— *systematic literature review, emotion classification, text data

***Abstrak*—Emosi adalah respon seseorang terhadap suatu peristiwa. Emosi dapat diekspresikan secara verbal dan non-verbal. Seiring dengan perkembangan zaman, orang dapat mengekspresikan emosinya melalui media sosial. Mengingat emosi adalah cerminan dari respon seseorang, maka penting untuk melakukan klasifikasi terhadap emosi dalam masyarakat sehingga dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan. Penelitian ini bertujuan untuk mengidentifikasi serta menganalisis dataset, metode, dan metrik evaluasi yang digunakan dalam penelitian mengenai klasifikasi emosi pada data tekstual dari tahun 2013 sampai tahun 2022. Berdasarkan desain inklusi dan eksklusi dalam pemilihan literatur, didapatkan sebanyak 50 literatur yang akan digunakan dalam ekstraksi dan sintesis data. Hasil analisis menunjukkan bahwa dari 50 literatur, 36 literatur menggunakan dataset publik dan 14 literatur menggunakan dataset privat. Pada metode dalam pengembangan model, SVM dan Naive Bayes merupakan model yang paling populer diantara lainnya. Dalam melakukan evaluasi model, metrik F-measure atau F1-score adalah metrik yang paling banyak digunakan dibandingkan dengan metrik lainnya. Terdapat tiga kontribusi utama yang diidentifikasi dalam penelitian ini, yaitu metode, model, serta evaluasi.**

Kata Kunci*—*systematic literature review, klasifikasi emosi, data tekstual

I. PENDAHULUAN

Emosi didefinisikan sebagai respon seseorang terhadap suatu peristiwa atau situasi secara sadar yang terjadi pada durasi tertentu [1]. Emosi dapat diekspresikan menggunakan komunikasi verbal maupun secara non-verbal [2]. Secara verbal, emosi dapat diekspresikan secara lisan maupun menggunakan tulisan. Sementara secara non-verbal, emosi dapat diekspresikan melalui bahasa tubuh seperti gerakan tangan dan raut wajah. Berdasarkan representasinya, emosi terbagi ke dalam dunia jenis yaitu dimensional dan kategorikal. Emosi dimensional merepresentasikan emosi ke dalam sebuah ruang spasial, semenara emosi kategorikal menempatkan emosi ke dalam beberapa kelas atau kategori yang berbeda [3].

Dengan kemajuan teknologi saat ini, seseorang dapat dengan mudah mengekspresikan emosi yang dimilikinya secara verbal melalui tulisan pada media sosial. Untuk dapat mengetahui emosi tersebut perlu dilakukan klasifikasi emosi. Klasifikasi emosi adalah proses untuk memetakan atau mengkategorikan dokumen ke dalam sekumpulan emosi yang telah ditentukan sebelumnya [4]. Mengingat emosi adalah respon seseorang terhadap suatu peristiwa, klasifikasi emosi menjadi sangat penting untuk dilakukan di berbagai sektor kehidupan. Pada sektor bisnis, respon dari pengguna sangat penting sebagai bahan pertimbangan dalam pengambilan keputusan untuk meningkatkan kualitas produk yang dihasilkan. Selain itu, dalam sektor pendidikan klasifikasi emosi juga dapat membantu guru untuk meningkatkan metode pembelajaran berdasarkan respon dari siswa [5].

Saat ini, klasifikasi emosi pada data tekstual adalah salah satu cabang dari pemrosesan bahasa alami yang berkembang paling cepat [6]. Perkembangan yang cepat ini juga menyebabkan tersebarnya dataset, metode, serta skema evaluasi yang digunakan klasifikasi emosi, sehingga mempersulit untuk mendapatkan gambaran menyeluruh dari klasifikasi emosi. Oleh karena itu, diperlukan sebuah *literature review* yang bertujuan untuk mengidentifikasi dan menganalisis dataset, metode, serta skema evaluasi dari klasifikasi emosi.

Dalam beberapa tahun terakhir, telah dilakukan penelitian-penelitian terkait untuk mendapatkan gambaran dari klasifikasi emosi khususnya pada data tekstual. Rabeya, et al melakukan penelitian untuk menyelidiki pendekatan leksikon terhadap klasifikasi emosi pada teks Bengali [14]. Penulis menemukan bahwa secara khusus terdapat tiga pendekatan yang digunakan, yaitu leksikon, machine learning, dan hybrid. Penelitian lainnya

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

juga melakukan analisis terhadap metode dalam klasifikasi emosi yaitu berdasarkan kata kunci atau keyword, leksikon, machine learning, dan hybrid [70]. Penelitian tersebut juga membahas beberapa dataset yang digunakan dalam klasifikasi emosi. Penelitian yang dilakukan oleh Alswaidan, N., & Menai, M melakukan survei terhadap pendekatan yang digunakan dalam klasifikasi emosi yaitu rule-based, classical learning, deep learning, dan hybrid [71]. Penggunaan features pada setiap pendekatan juga dibahas pada penelitian tersebut. Mohammed R. Elkobaisi et al juga melakukan survei yang berfokus pada ontologi emosi, dataset, serta sistem yang digunakan dalam klasifikasi emosi [72]. Berdasarkan penelitian-penelitian sebelumnya, belum ditemukan *systematic literature review* yang berfokus pada metode machine learning termasuk deep learning dalam klasifikasi emosi pada data tekstual. Oleh karena itu, penelitian ini diharapkan dapat membantu mengurangi kesenjangan penelitian.

Literatur yang digunakan pada penelitian ini berasal dari tahun 2013 sampai tahun 2022. Menggunakan string pencarian yang telah ditentukan, sebanyak 250 literatur berhasil dikumpulkan. Melalui kriteria inklusi dan eksklusi, terpilih 50 penelitian yang akan digunakan dalam proses literatur review yang berkaitan dengan pertanyaan penelitian yang diajukan.

Penelitian ini diharapkan mampu untuk memberikan kontribusi dalam beberapa aspek seperti pembuatan SLR seksama terkait klasifikasi emosi pada data tekstual, analisis dan diskusi yang detail berdasarkan pertanyaan penelitian, dan identifikasi kemungkinan serta tantangan penelitian dalam hal klasifikasi emosi pada data tekstual di masa mendatang.

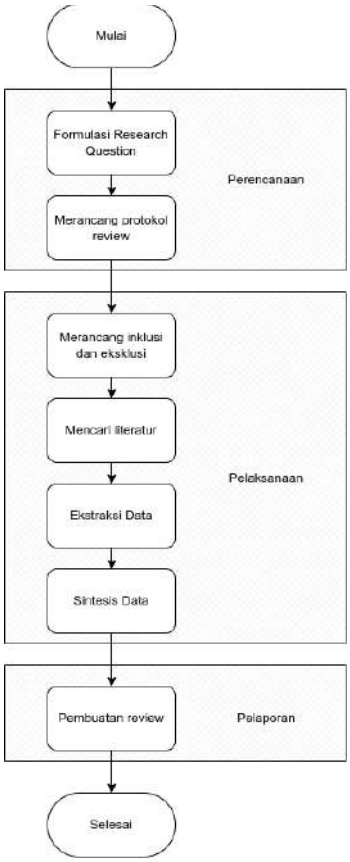
II. METODE PENELITIAN

A. Desain Penelitian

Pada penelitian ini, metode *literature review* yang digunakan adalah *systematic literature review* atau tinjauan pustaka sistematis. *Systematic literature review* atau *systematic review* merupakan cara untuk mengidentifikasi, mengevaluasi, menginterpretasikan semua penelitian yang tersedia terkait dengan pertanyaan penelitian tertentu [7]. Terdapat tiga tahapan utama dalam melakukan *systematic literature review*, ketiga tahap tersebut adalah perencanaan, pelaksanaan, dan pelaporan. Seluruh tahapan utama dalam *systematic literature review* ditunjukkan pada Gambar 1.

B. Pertanyaan Penelitian

Penentuan pertanyaan penelitian (research question) adalah bagian terpenting dalam tahap perencanaan karena akan memandu keseluruhan proses dari *systematic literature review* [8]. Pertanyaan penelitian dibangun menggunakan kriteria PICOC (Population, Intervention, Comparison, Outcome, Context) [7]. Tabel I menunjukkan formulasi pertanyaan penelitian yang dibangun dari PICOC.



Gambar 1. Tahapan *Systematic Literature Review*

TABEL I FORMULA RESEARCH QUESTION BERDASARKAN PICOC

PICOC	Formulasi Pertanyaan
Population	NLP, klasifikasi teks, machine learning, deep learning
Intervention	klasifikasi emosi, klasifikasi, model, metode, dataset, metrik evaluasi
Comparison	-
Outcome	prediksi akurasi dari klasifikasi emosi, metode klasifikasi emosi
Context	dataset berukuran besar dan kecil, penelitian pada bidang industri dan akademik

Berdasarkan formulasi dari Tabel I, adapun pertanyaan penelitian yang digunakan dalam *systematic literature review* ini ditunjukkan pada Tabel II.

TABEL II RESEARCH QUESTION

	Research Question
RQ1	Apakah jenis kontribusi yang diberikan terhadap klasifikasi emosi pada data tekstual?
RQ2	Apakah jenis dataset yang digunakan dalam klasifikasi emosi pada data tekstual?
RQ3	Apakah metode yang digunakan dalam klasifikasi emosi pada data tekstual?
RQ4	Apakah metrik evaluasi yang digunakan dalam klasifikasi emosi pada data tekstual?

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

C. Strategi Pencarian

Pada tahapan ini akan dilakukan pemilihan perpustakaan digital, merancang *string* pencarian, melakukan pencarian, dan mengambil literatur yang sesuai dengan *string* pencarian. Sebelum melakukan pencarian, akan dilakukan pemilihan perpustakaan digital. Perpustakaan atau *database* digital harus dipilih untuk meningkatkan kemungkinan dalam menemukan literatur yang relevan. Pencarian akan dilakukan pada perpustakaan digital yang paling populer akan dicari untuk menemukan literatur dengan cakupan yang luas [9]. Adapun daftar perpustakaan digital yang akan digunakan dalam *systematic literature review* ini ditunjukkan pada Tabel III.

TABEL III. TABEL PERPUSTAKAN DIGITAL

Nama	Situs Web
Google Scholar	https://scholar.google.com/
IEEE Xplore	https://ieeexplore.ieee.org/
ACM Digital Library	https://dl.acm.org/
ScienceDirect	https://www.sciencedirect.com/
Springer	https://www.springer.com/

Setelah penentuan perpustakaan digital, selanjutnya adalah merancang *string* pencarian, adapun *string* pencarian yang digunakan adalah:

(emotion) AND (detection OR classification OR recognition OR analysis) AND (textual or text)

Proses pencarian dilakukan dengan mengambil sampel sebanyak 50 artikel teratas pada setiap perpustakaan digital yang ada, sehingga akan didapatkan 250 artikel.

D. Kriteria Pemilihan

Setelah mendapatkan daftar referensi, selanjutnya adalah melakukan seleksi terhadap setiap literatur untuk menentukan apakah literatur tersebut harus disertakan untuk ekstraksi dan analisis data [10]. Untuk melakukan seleksi, akan digunakan kriteria inklusi dan eksklusi yang ditunjukkan pada Tabel IV.

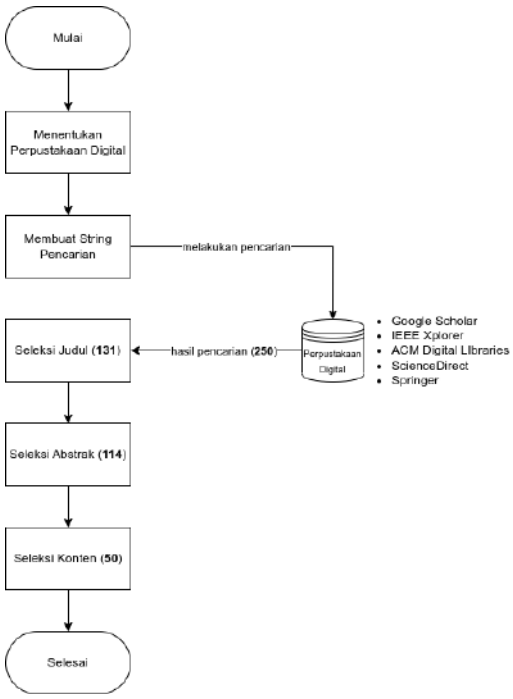
TABEL IV TABEL KRITERIA INKLUSI DAN EKSKLUSI

Inklusi	Eksklusi
Literatur mengembangkan model atau metode komputasi untuk melakukan klasifikasi emosi	Literatur tidak mengembangkan model atau hanya melakukan survei atau studi literatur
Menggunakan dataset data tekstual semi terstruktur atau tidak terstruktur	Menggunakan dataset data tekstual terstruktur atau diluar data tekstual
Literatur memiliki judul serta abstrak yang sesuai dengan string pencarian	Literatur memiliki judul serta abstrak yang tidak sesuai dengan string pencarian
Literatur ditulis dalam bahasa Inggris	Literatur tidak ditulis dalam bahasa Inggris
Literatur menggunakan metrik untuk evaluasi model	Literatur tidak menggunakan metrik untuk evaluasi model

Tahapan seleksi akan dilakukan dalam tiga tahap, yaitu seleksi terhadap judul, seleksi terhadap abstrak dan kata kunci, serta seleksi terhadap konten atau isi dari artikel.

Setelah dilakukan seleksi pada judul artikel, artikel yang awalnya berjumlah 250 artikel berkurang menjadi 131 artikel. Kemudian dilakukan seleksi terhadap abstrak dan kata kunci pada artikel, sehingga didapatkan 114 artikel.

Tahap terakhir dalam seleksi adalah seleksi pada konten yang menghasilkan 50 artikel. Dengan demikian, terdapat 50 literatur yang akan digunakan untuk tahap ekstraksi serta sintesis data. Seluruh tahapan dalam pencarian serta seleksi artikel ditunjukkan pada Gambar 2.



Gambar 2. Tahapan Pencarian dan Seleksi Literatur

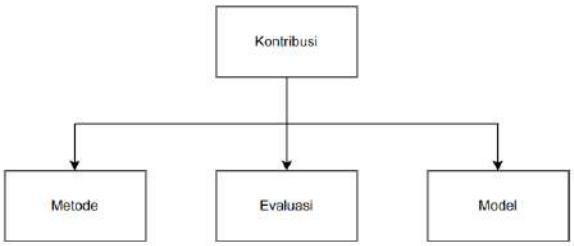
E. Ekstraksi Data

Literatur yang sudah melewati proses inklusi dan eksklusi akan diekstraksi untuk mendapatkan data yang nantinya digunakan untuk menjawab pertanyaan penelitian. Tujuan dari tahap ini adalah merancang formulir ekstraksi data untuk secara akurat merekam informasi yang diperoleh peneliti dari studi utama [7]. Ekstraksi data dilakukan terhadap beberapa kategori yang diidentifikasi berdasarkan pertanyaan penelitian. Terdapat tiga kategori yang digunakan untuk menjawab pertanyaan penelitian yang ditunjukkan pada Tabel V.

TABEL V TABEL KATEGORI EKSTRAKSI DATA

Kategori	Pertanyaan Penelitian
Kontribusi	RQ1
Dataset	RQ2
Metode	RQ3
Metrik Evaluasi	RQ4

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023



Gambar 3 Kerangka Kontribusi dari Literatur

F. Sintesis Data

Sintesis data melibatkan proses proses penyusunan dan peringkasan hasil literatur utama yang telah diekstraksi [7]. Tujuan dari sintesis data adalah untuk mengumpulkan bukti-bukti dari literatur utama untuk menjawab pertanyaan penelitian [9]. Tipe sintesis data yang digunakan dalam *systematic literature review* ini adalah sintesis naratif (deskriptif). Data akan ditabulasikan dengan cara yang konsisten dengan pertanyaan penelitian. Beberapa jenis visualisasi seperti diagram batang dan diagram lingkaran juga digunakan dalam penyajian data.

III. HASIL DAN PEMBAHASAN

A. RQ1: Apakah Jenis Kontribusi yang Diberikan terhadap Klasifikasi Emosi pada Data Tekstual?

Penelitian ini menghasilkan tiga kontribusi, yaitu kontribusi terhadap metode, evaluasi, dan model dalam klasifikasi emosi pada data tekstual. Kontribusi terhadap model memiliki persentase paling besar, yaitu sebanyak 34 literatur dari 50 literatur, kontribusi metode sebanyak 9 literatur dan evaluasi sebanyak 4 literatur. Penelitian untuk model seperti yang dilakukan oleh Deshpande M et al menghasilkan sebuah model machine learning SVM dan Multinomial Naive Bayes dengan akurasi diatas 80% [27].

Penelitian yang dilakukan oleh Batbaatar, et al mengusulkan sebuah metode baru dalam klasifikasi emosi, yaitu SENN (Semantic Emotion Neural Network) [20]. Metode yang berbeda juga diusulkan pada penelitian yang dilakukan oleh Chatterjee et. al. Pada penelitian tersebut peneliti menggunakan SS-BED yang menghasilkan F1 lebih tinggi dibandingkan dengan CNN serta RNN [23]. Liu, et al juga melakukan penelitian dengan mengembangkan sebuah metode baru yaitu SERR (Semantic and Emoticon Based Emotion Recognizer) untuk melakukan klasifikasi emosi [45]. Penelitian lainnya mengembangkan metode IDS-ICM (Interactive Double States Emotion Cell Model) [50]. Metode-metode baru lainnya yang berbasis deep learning juga dikembangkan seperti MHA-BCNN [55], HAN-ReGRU [66], dan ELReLUWL [69] .

Kontribusi lainnya adalah dalam evaluasi yang digunakan. Penelitian yang dilakukan oleh Angel et al menggunakan Pearson Score dan Mean Absolute Error [68]. Selain itu, pada penelitian lainnya yang dilakukan oleh K. Anoop et al menggunakan evaluasi berupa Average Pearson dan Wasserstein Distance dalam melakukan evaluasi terhadap model deep learning yang digunakan [64]. Sebaran kontribusi pada literatur SLR ini ditunjukkan pada Tabel VI.

TABEL VI TABEL REKAPITULASI KONTRIBUSI DARI LITERATUR

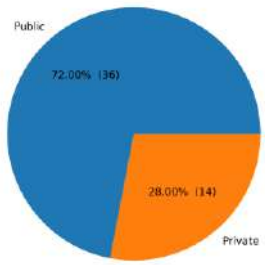
Id	Referensi	Tahun	Sumber	Kontribusi
1	[20]	2019	Google Scholar	Metode
4	[21]	2019	Google Scholar	Model
5	[22]	2018	Google Scholar	Evaluasi
6	[23]	2019	Google Scholar	Metode
8	[24]	2014	Google Scholar	Model
12	[25]	2016	Google Scholar	Evaluasi
14	[26]	2014	Google Scholar	Model
19	[27]	2017	Google Scholar	Model
52	[28]	2020	IEEE Xplore	Evaluasi
61	[29]	2018	IEEE Xplore	Model
62	[30]	2018	IEEE Xplore	Model
64	[31]	2013	IEEE Xplore	Model
..		..	..	..
109	[44]	2019	ACM Digital Library	Model
111	[45]	2020	ACM Digital Library	Metode
133	[46]	2017	ACM Digital Library	Metode
154	[47]	2021	Science Direct	Model
155	[48]	2021	Science Direct	Metode
156	[49]	2021	Science Direct	Model
158	[50]	2020	Science Direct	Metode
162	[51]	2020	Science Direct	Model
163	[52]	2022	Science Direct	Metode
171	[53]	2021	Science Direct	Model
172	[54]	2021	Science Direct	Model
174	[55]	2021	Science Direct	Metode
184	[56]	2016	Science Direct	Model
189	[57]	2020	Science Direct	Model
192	[58]	2020	Science Direct	Model
197	[59]	2020	Science Direct	Model
206	[60]	2019	Springer	Model
212	[61]	2019	Springer	Model
216	[62]	2022	Springer	Model

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

222	[63]	2022	Springer	Model
225	[64]	2022	Springer	Model
229	[65]	2019	Springer	Model
238	[66]	2021	Springer	Metode
240	[67]	2021	Springer	Model
248	[68]	2021	Springer	Evaluasi
249	[69]	2022	Springer	Metode

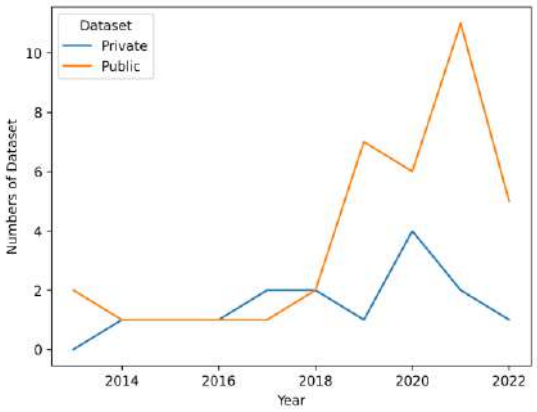
B. RQ2: Apakah Jenis Dataset yang Digunakan dalam Klasifikasi Emosi pada Data Tekstual?

Dataset yang digunakan dalam *systematic literature review* ini adalah dataset tekstual baik terstruktur maupun tidak terstruktur. Dari 50 literatur, terdapat 36 literatur yang menggunakan dataset yang bersifat publik sementara 14 literatur menggunakan dataset yang bersifat privat. Dataset yang bersifat publik sebagian besar menggunakan dataset ISEAR (International Survey on Emotion Antecedents and Reactions) [20], [21] maupun SemEval (Semantic Evaluation) [22], [23]. Sementara untuk dataset privat umumnya didapatkan dari teks pada media sosial [32], [39]. Gambar 3 menunjukkan perbandingan jumlah dataset publik dan privat yang digunakan dari 50 literatur yang digunakan pada *systematic literature review* ini.



Gambar 3.
Gambar 4. Grafik Distribusi Dataset

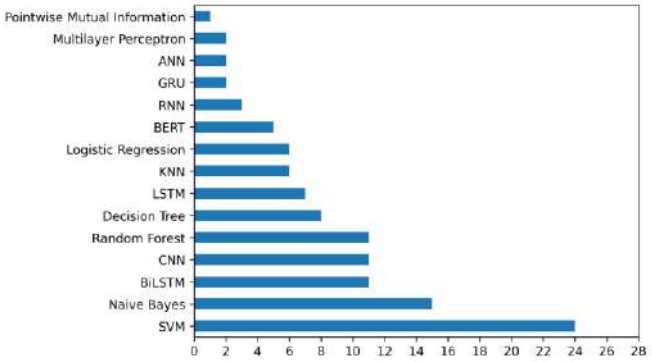
Pada Gambar 4 menunjukkan bagaimana tren penggunaan dataset publik dan privat dari tahun ke tahun, mulai dari tahun 2013 sampai tahun 2022. Dari Gambar 4 terlihat bawah penggunaan dataset publik cenderung mengalami peningkatan dibandingkan dengan dataset privat. Hal ini dimungkinkan mengingat tiap tahunnya semakin banyak studi yang dipublikasikan, sehingga ada semakin banyak data yang bisa untuk digunakan.



Gambar 5. Grafik Tren Penggunaan Dataset

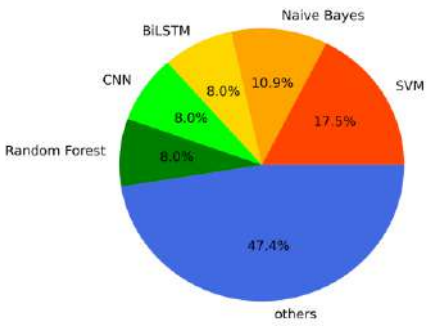
C. RQ3: Apakah Metode Yang Digunakan Dalam Klasifikasi Emosi Pada Data Tekstual?

Dari 50 literatur yang digunakan, terdapat berbagai macam metode yang digunakan untuk membangun model dalam melakukan klasifikasi terhadap emosi. Lima belas metode dengan penggunaan terbanyak ditunjukkan pada Gambar 5.



Gambar 6. Grafik Distribusi Metode Pada Klasifikasi Emosi

Diantaranya lima belas metode tersebut, metode dengan jumlah penggunaan terbanyak adalah SVM (Support Vector Machine). Pada Gambar 6 ditunjukkan lima metode teratas, yaitu SVM, Naive Bayes, BiLSTM, CNN, dan Random Forest. Kelima metode tersebut digunakan pada lebih dari setengah dari literatur yang dianalisis pada *systematic literature review* ini.



Gambar 7. Grafik Penggunaan Lima Metode Teratas pada Literatur

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

Dalam *systematic literature review* ini, terdapat cukup banyak metode yang digunakan untuk membangun model dalam melakukan klasifikasi emosi. Walaupun terdapat beberapa literatur yang membandingkan performa antar metode seperti [11], namun tidak dapat dinyatakan secara jelas bahwa metode tersebut lebih baik diantaranya metode yang lainnya ketika dilihat secara individual. Akan tetapi, berdasarkan grafik pada Gambar 4 dan Gambar 5, dapat dilihat bahwa beberapa metode populer yang digunakan adalah SVM, Naive Bayes, Random Forest, serta Neural Network seperti BiLSTM dan CNN.

Naive Bayes dan SVM dalam *systematic literature review* ini merupakan dua metode *machine learning* yang paling populer dalam melakukan klasifikasi emosi pada data tekstual. Naive Bayes sendiri didasari pada teorema Bayes. Kelebihan dari Naive Bayes adalah metode ini bekerja sangat baik pada data teks dan mudah untuk diimplementasikan serta cenderung cepat jika dibandingkan dengan algoritma yang lain [12]. Akan tetapi Naive Bayes juga memiliki beberapa kekurangan, salah satunya adalah Naive Bayes memiliki asumsi yang kuat terhadap distribusi data [13]. Sehingga, diperlukan dataset dengan distribusi yang imbang agar dapat membangun model yang baik.

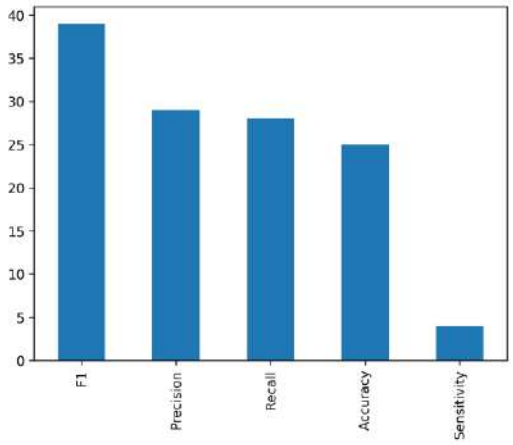
SVM atau Support Vector Machine adalah klasifikasi dan regresi dengan menggunakan konsep kernel dalam dimensi tinggi untuk memecahkan masalah non-linear. SVM beberapa kali menunjukkan performa yang lebih baik dibandingkan model lainnya seperti pada [11] dan [15]. SVM memiliki kelebihan untuk membangun *hyperlane* yang bersifat non-linear sehingga dapat digunakan untuk memaksimalkan jarak margin antar kelas. Kekurangan dari SVM salah satunya adalah kurangnya transparansi dalam hasil klasifikasi yang disebabkan karena tingginya dimensi pada data teks [12].

Neural Network atau Deep Neural Network adalah model yang meniru bagaimana otak manusia bekerja [16]. Arsitektur dari Neural Network terbentuk dari lapisan yang saling terhubung yang terdiri dari tiga tingkatan yang berbeda, yaitu *input layer*, *hidden layer*, dan *output layer* [17]. Secara umum terdapat dua jenis arsitektur dari Neural Network, yaitu *feed-forward network* dan *recurrent network*. Contoh dari *feed-forward network* adalah CNN sementara contoh dari *recurrent network* adalah RNN, LSTM, dan GRU. BERT atau juga menggunakan basis Neural Network untuk membangun *pre-trained natural language* model [18]. Neural Network memiliki kelebihan dalam mengolah data dengan fitur yang kompleks, namun memiliki kelemahan dalam komputasi yang berat dan membutuhkan data yang banyak [12].

Menentukan model yang memiliki kinerja terbaik harus disesuaikan dengan teknik dan data yang tepat. Setiap model akan memberikan hasil yang berbeda dengan perlakuan yang berbeda dan jenis dataset yang berbeda. Oleh karena itu diperlukan pemahaman terhadap setiap algoritma untuk dapat menentukan model yang paling cocok untuk digunakan dalam permasalahan klasifikasi emosi.

D. Apakah Metrik Evaluasi yang Digunakan dalam Klasifikasi Emosi pada Data Tekstual?

Performa dari sebuah model dapat diketahui dengan melakukan evaluasi pada model. Kinerja dari sebuah model direpresentasikan oleh metrik tertentu. Metrik-metrik evaluasi ini memberikan informasi yang berbeda-beda sesuai dengan informasi yang ingin ditunjukkan oleh metrik tersebut [12]. Dalam *systematic literature review* ini, metrik yang paling umum digunakan adalah *accuracy*, *precision*, *recall*, dan *F-measure* yang didasarkan pada *confusion matrix*. *Confusion matrix* sendiri adalah sebuah matriks yang menggambarkan hasil prediksi atau klasifikasi dari model (predicted classification) dan klasifikasi yang sebenarnya (actual classification). Gambar 7 menunjukkan sebaran penggunaan metrik untuk evaluasi.



Gambar 8. Grafik Distribusi Metrik Evaluasi

Penggunaan metrik evaluasi yang paling banyak adalah metrik *F-measure* atau *F1-score*. Metrik *F-measure* digunakan untuk mengetahui *harmonic-mean* dengan menghitung agregat antara metrik precision dan recall [19].

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, melalui tahapan desain inklusi dan eksklusi dalam pemilihan literatur didapatkan sebanyak 50 literatur yang digunakan dalam ekstraksi dan sintesis data. Analisis dari data tersebut menunjukkan bahwa dari 50 literatur, terdapat 72% literatur menggunakan dataset publik sementara 28% literatur menggunakan dataset privat. Pada metode dalam pengembangan model untuk melakukan klasifikasi, model SVM dan Naive Bayes merupakan yang paling populer diantara model lainnya dengan akumulasi penggunaan metode 28% dari jumlah literatur. Dalam menentukan model, diperlukan pemahaman dan analisis terhadap permasalahan agar mendapatkan model yang optimal untuk klasifikasi emosi. Dalam melakukan evaluasi terhadap model, metrik *F-measure* atau *F1-score* adalah metrik yang paling banyak digunakan dibandingkan dengan metrik lainnya dimana metrik ini digunakan pada 78% dari total literatur. Tantangan dari penelitian ini adalah klasifikasi emosi melalui teks termasuk sulit, sehingga untuk kedepannya dapat dilakukan penelitian yang tidak hanya terbatas pada data tekstual, namun bisa melalui data audio maupun citra.

UCAPAN TERIMA KASIH

Penelitian ini mendapat pendanaan dari Hibah Penelitian

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

Universitas Udayana Tahun 2023 [hibah no. B/1.43/UN14.4.A/PT.01.03/2023]

DAFTAR PUSTAKA

[1] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human Emotion Recognition: Review of Sensors and Methods," Sensors, vol. 20, no. 3, 2020

[2] L. Schoneveld, A. Othmani, and H. Abdelkawy, "Leveraging Recent Advances in Deep Learning for Audio-Visual Emotion Recognition," Pattern Recognition Letters, vol. 46, pp. 1-7, 2021

[3] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," Engineering Reports, vol. 2, no. 7, 2020

[4] Oberländer, L.A.M. and Klinger, R., "An analysis of annotated corpora for emotion classification in text," in Proceedings of the 27th International Conference on Computational Linguistics, USA, 2018

[5] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," Social Network Analysis and Mining, vol. 11, no. 1, 2021

[6] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, "Emotion Detection of Textual Data: An Interdisciplinary Survey," 2021 IEEE World AI IoT Congress (AIIoT), USA. 2021

[7] Keele, S, "Guidelines for performing systematic literature reviews in software engineering", In: Technical report, Ver. 2.3 EBSE Technical Report, UK, 2007

[8] V. Torres-Carrion, C. S. Gonzalez-Gonzalez, S. Aciar, and G. Rodriguez-Morales, "Methodology for systematic literature review applied to engineering and education," 2018 IEEE Global Engineering Education Conference (EDUCON), Spain, 2018

[9] Wahono RS, "A systematic literature review of software defect prediction", Journal of software engineering, vol. 1, no. 1, pp. 1-6, 2015

[10] Xiao, Y. and Watson, M., "Guidance on conducting a systematic literature review". Journal of planning education and research, vol. 39, no. 1, pp. 93-112, 2019

[11] T. Parvin, O. Sharif, and M. M. Hoque, "Multi-class Textual Emotion Categorization using Ensemble of Convolutional and Recurrent Neural Network," SN Computer Science, vol. 3, no. 1, 2021

[12] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text Classification Algorithms: A Survey," Information, vol. 10, no. 4, p. 150, 2019

[13] Y. Wang, R. Khardon, and P. Protopapas, "Nonparametric Bayesian estimation of periodic light curves," The Astrophysical Journal, vol. 756, no. 1, pp. 67–67, 2012

[14] T. Rabeya, S. Ferdous, H. S. Ali and N. R. Chakraborty, "A survey on emotion detection: A lexicon based backtracking approach for detecting emotion from Bengali text," 2017 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2017, pp. 1-7, doi: 10.1109/ICCITECHN.2017.8281855.

[15] Ardiada, I.M.D., Sudarma, M. and Giriantari, D., 2019. "Text Mining pada Sosial Media untuk Mendeteksi Emosi Pengguna Menggunakan Metode Support Vector Machine dan K-Nearest Neighbour". Maj. Ilm. Teknol. Elektro, vol. 18, no. 1, 2019

[16] Fudholi, D.H., "Klasifikasi Emosi pada Teks dengan Menggunakan Metode Deep Learning". Syntax Literate; Jurnal Ilmiah Indonesia, vol. 6, no. 1, pp.546-553, 2021

[17] Osinga, D., Deep learning cookbook: practical recipes to get started quickly, Sebastopol: O'Reilly Media, Inc., 2018

[18] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., "Bert: Pre-training of deep bidirectional transformers for language understanding", arXiv (Cornell University), 2018

[19] M. Grandini, Enrico Bagli, and Giorgio Visani, "Metrics for Multi-Class Classification: an Overview," arXiv (Cornell University), 2020

[20] Batbaatar E, Li M, Ryu KH. Semantic-emotion neural network for emotion recognition from text. IEEE access. 2019 Aug 12;7:111866-78.

[21] Sailunaz K, Alhajj R. Emotion and sentiment analysis from Twitter text. Journal of Computational Science. 2019 Sep 1;36:101003.

[22] Kratzwald B, Ilić S, Kraus M, Feuerriegel S, Prendinger H. Deep learning for affective computing: Text-based emotion recognition in decision support. Decision Support Systems. 2018 Nov 1;115:24-35.

[23] Chatterjee A, Gupta U, Chinnakotla MK, Srikanth R, Galley M, Agrawal P. Understanding emotions in text using deep learning and big data. Computers in Human Behavior. 2019 Apr 1;93:309-17.

[24] Li W, Xu H. Text-based emotion classification using emotion cause extraction. Expert Systems with Applications. 2014 Mar 1;41(4):1742-9.

[25] Perikos I, Hatzilygeroudis I. Recognizing emotions in text using ensemble of classifiers. Engineering Applications of Artificial Intelligence. 2016 May 1;51:191-201.

[26] Kiritchenko S, Zhu X, Mohammad SM. Sentiment analysis of short informal texts. Journal of Artificial Intelligence Research. 2014 Aug 20;50:723-62.

[27] Deshpande M, Rao V. Depression detection using emotion artificial intelligence. In2017 international conference on intelligent sustainable systems (iciss) 2017 Dec 7 (pp. 858-862). IEEE.

[28] M. Karna, D. S. Juliet and R. C. Joy, "Deep learning based Text Emotion Recognition for Chatbot applications," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 2020, pp. 988-993, doi: 10.1109/ICOEI48184.2020.9142879.

[29] S. Chawla and M. Mehrotra, "An Ensemble-Classifier Based Approach for Multiclass Emotion Classification of Short Text," 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2018, pp. 768-774, doi: 10.1109/ICRITO.2018.8748757.

[30] M. -H. Su, C. -H. Wu, K. -Y. Huang and Q. -B. Hong, "LSTM-based Text Emotion Recognition Using Semantic and Emotional Word Vectors," 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia), Beijing, China, 2018, pp. 1-6, doi: 10.1109/ACIIAsia.2018.8470378.

[31] T. Patil and S. Patil, "Automatic generation of emotions for social networking websites using text mining," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 2013, pp. 1-6, doi: 10.1109/ICCCNT.2013.6726704.

[32] U. Rashid, M. W. Iqbal, M. A. Sikandar, M. Q. Raiz, M. R. Naqvi and S. K. Shahzad, "Emotion Detection of Contextual Text using Deep learning," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Istanbul, Turkey, 2020, pp. 1-5, doi: 10.1109/ISMSIT50672.2020.9255279.

[33] J. De Silva and P. S. Haddela, "A term weighting method for identifying emotions from text content," 2013 IEEE 8th International Conference on Industrial and Information Systems, Peradeniya, Sri Lanka, 2013, pp. 381-386, doi: 10.1109/ICIInfS.2013.6732014.

[34] H. Al-Omari, M. A. Abdullah and S. Shaikh, "EmoDet2: Emotion Detection in English Textual Dialogue using BERT and BiLSTM Models," 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 2020, pp. 226-232, doi: 10.1109/ICICS49469.2020.239539.

[35] K. P. -Q. Nguyen and K. Van Nguyen, "Exploiting Vietnamese Social Media Characteristics for Textual Emotion Recognition in Vietnamese," 2020 International Conference on Asian Language Processing (IALP), Kuala Lumpur, Malaysia, 2020, pp. 276-281, doi: 10.1109/IALP51396.2020.9310495.

[36] R. Majid and H. A. Santoso, "Conversations Sentiment and Intent Categorization Using Context RNN for Emotion Recognition," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 46-50, doi: 10.1109/ICACCS51430.2021.9441740.

[37] H. A. Ruposh and M. M. Hoque, "A Computational Approach of Recognizing Emotion from Bengali Texts," 2019 5th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, 2019, pp. 570-574, doi: 10.1109/ICAEE48663.2019.8975417.

[38] H. Al Huzali and S. Ananiadou, "Improving Textual Emotion Recognition

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

Based on Intra- and Inter-Class Variation," in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2021.3104720.

[39] R. J. Hasudungan and M. L. Kodhra, "Detecting Emotion on Indonesian Online Chat Text Using Text Sequential Labeling," 2018 International Symposium on Advanced Intelligent Informatics (SAIN), Yogyakarta, Indonesia, 2018, pp. 167-172, doi: 10.1109/SAIN.2018.8673342.

[40] L. Canales, W. Daelemans, E. Boldrini and P. Martínez-Barco, "EmoLabel: Semi-Automatic Methodology for Emotion Annotation of Social Media Text," in IEEE Transactions on Affective Computing, vol. 13, no. 2, pp. 579-591, 1 April-June 2022, doi: 10.1109/TAFFC.2019.2927564.

[41] A. Yousaf et al., "Emotion Recognition by Textual Tweets Classification Using Voting Classifier (LR-SGD)," in IEEE Access, vol. 9, pp. 6286-6295, 2021, doi: 10.1109/ACCESS.2020.3047831.

[42] Jonathan Herzig, Michal Shmueli-Scheuer, and David Konopnicki. 2017. Emotion Detection from Text via Ensemble Classification Using Word Embeddings. In Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR '17). Association for Computing Machinery, New York, NY, USA, 269–272. https://doi.org/10.1145/3121050.3121093

[43] Adil Majeed, Hasan Mujtaba, and Mirza Omer Beg. 2021. Emotion detection in Roman Urdu text using machine learning. In Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering (ASE '20). Association for Computing Machinery, New York, NY, USA, 125–130. https://doi.org/10.1145/3417113.3423375

[44] Marco Polignano, Pierpaolo Basile, Marco de Gemmis, and Giovanni Semeraro. 2019. A Comparison of Word-Embeddings in Emotion Detection from Text using BiLSTM, CNN and Self-Attention. In Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (UMAP'19 Adjunct). Association for Computing Machinery, New York, NY, USA, 63–68. https://doi.org/10.1145/3314183.3324983

[45] Taiao Liu, Yajun Du, and Qiaoyu Zhou. 2020. Text Emotion Recognition Using GRU Neural Network with Attention Mechanism and Emoticon Emotions. Proceedings of the 2020 2nd International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI '20). Association for Computing Machinery, New York, NY, USA, 278–282. https://doi.org/10.1145/3438872.3439094

[46] Panpan Li, Jun Li, Feiqiang Sun, and Peng Wang. 2017. Short Text Emotion Analysis Based on Recurrent Neural Network. In Proceedings of the 6th International Conference on Information Engineering (ICIE '17). Association for Computing Machinery, New York, NY, USA, Article 6, 1–5. https://doi.org/10.1145/3078564.3078569

[47] T. Parvin and M. M. Hoque, "An Ensemble Technique to Classify Multi-Class Textual Emotion," Procedia Computer Science, vol. 193, pp. 72–81, 2021, doi: 10.1016/j.procs.2021.10.008.

[48] L. Kang, J. Liu, L. Liu, Z. Zhou, and D. Ye, "Semi-supervised emotion recognition in textual conversation via a context-augmented auxiliary training task," Information Processing & Management, vol. 58, no. 6, p. 102717, Nov. 2021, doi: 10.1016/j.ipm.2021.102717.

[49] R. Kumari, N. Ashok, T. Ghosal, and A. Ekbal, "Misinformation detection using multitask learning with mutual learning for novelty detection and emotion recognition," Information Processing & Management, vol. 58, no. 5, p. 102631, Sep. 2021, doi: 10.1016/j.ipm.2021.102631.

[50] D. Li, Y. Li, and S. Wang, "Interactive double states emotion cell model for textual dialogue emotion prediction," Knowledge-Based Systems, vol. 189, p. 105084, Feb. 2020, doi: 10.1016/j.knosys.2019.105084.

[51] T. T. Sasidhar, P. B, and S. K. P, "Emotion Detection in Hinglish(Hindi+English) Code-Mixed Social Media Text," Procedia Computer Science, vol. 171, pp. 1346–1352, 2020, doi: 10.1016/j.procs.2020.04.144.

[52] N. Shelke, S. Chaudhury, S. Chakrabarti, S. L. Bangare, G. Yogapriya, and P. Pandey, "An efficient way of text-based emotion analysis from social media using LRA-DNN," Neuroscience Informatics, vol. 2, no. 3, p. 100048, Sep. 2022, doi: 10.1016/j.neuri.2022.100048.

[53] Chowanda, A. et al. (2021) 'Exploring text-based emotions recognition machine learning techniques on social media conversation', Procedia Computer Science, 179, pp. 821–828. doi:10.1016/j.procs.2021.01.099.

[54] [1] Z. Li, H. Xie, G. Cheng, and Q. Li, "Word-level emotion distribution with two schemas for short text emotion classification," Knowledge-Based Systems, vol. 227, p. 107163, 2021. doi:10.1016/j.knosys.2021.107163

[55] K. Dheeraj and T. Ramakrishnudu, "Negative emotions detection on online mental-health related patients texts using the deep learning with MHA-BCNN model," Expert Systems with Applications, vol. 182, p. 115265, 2021. doi:10.1016/j.eswa.2021.115265

[56] D. Yasmina, M. Hajar, and A. M. Hassan, "Using YouTube comments for text-based emotion recognition," Procedia Computer Science, vol. 83, pp. 292–299, 2016. doi:10.1016/j.procs.2016.04.128

[57] A. Gupta and S. M. Srinivasan, "Constructing a heterogeneous training dataset for emotion classification," Procedia Computer Science, vol. 168, pp. 73–79, 2020. doi:10.1016/j.procs.2020.02.259

[58] F. M. Plaza-del-Arco, M. T. Martín-Valdivia, L. A. Ureña-López, and R. Mitkov, "Improved emotion recognition in Spanish social media through incorporation of lexical knowledge," Future Generation Computer Systems, vol. 110, pp. 1000–1008, Sep. 2020, doi: 10.1016/j.future.2019.09.034.

[59] Z. Halim, M. Waqar, and M. Tahir, "A machine learning-based investigation utilizing the in-text features for the identification of dominant emotion in an email," Knowledge-Based Systems, vol. 208, p. 106443, Nov. 2020, doi: 10.1016/j.knosys.2020.106443.

[60] Shrivastava, K., Kumar, S. & Jain, D.K. An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network. Multimed Tools Appl 78, 29607–29639 (2019). https://doi.org/10.1007/s11042-019-07813-9

[61] Hasan, M., Rundensteiner, E. & Agu, E. Automatic emotion detection in text streams by analyzing Twitter data. Int J Data Sci Anal 7, 35–51 (2019). https://doi.org/10.1007/s41060-018-0096-z

[62] Dhar, S., Gour, V. & Paul, A. Emotion recognition from lyrical text of Hindi songs. Innovations Syst Softw Eng (2022). https://doi.org/10.1007/s11334-022-00520-z

[63] Parvin, T., Sharif, O. & Hoque, M.M. Multi-class Textual Emotion Categorization using Ensemble of Convolutional and Recurrent Neural Network. SN COMPUT. SCI. 3, 62 (2022). https://doi.org/10.1007/s42979-021-00913-0

[64] K., A., P., D., Sam Abraham, S. et al. Readers' affect: predicting and understanding readers' emotions with deep learning. J Big Data 9, 82 (2022). https://doi.org/10.1186/s40537-022-00614-2

[65] Ghanbari-Adivi, F., Mosleh, M. Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE). Neural Comput & Applic 31, 8971–8983 (2019). https://doi.org/10.1007/s00521-019-04230-9

[66] Ma, H., Wang, J., Qian, L. et al. HAN-ReGRU: hierarchical attention network with residual gated recurrent unit for emotion recognition in conversation. Neural Comput & Applic 33, 2685–2703 (2021). https://doi.org/10.1007/s00521-020-05063-7

[67] Khalil, E.A.H., Houby, E.M.F.E. & Mohamed, H.K. Deep learning for emotion analysis in Arabic tweets. J Big Data 8, 136 (2021). https://doi.org/10.1186/s40537-021-00523-w

[68] Angel Deborah, S., Mirnalinee, T.T. & Rajendram, S.M. Emotion Analysis on Text Using Multiple Kernel Gaussian.... Neural Process Lett 53, 1187–1203 (2021). https://doi.org/10.1007/s11063-021-10436-7

[69] Yang, H., Alsadoon, A., Prasad, P.W.C. et al. Deep learning neural networks for emotion classification from text: enhanced leaky rectified linear unit activation and weighted loss. Multimed Tools Appl 81, 15439–15468 (2022). https://doi.org/10.1007/s11042-022-12629-1

[70] Sailunaz, K., Dhaliwal, M., Rokne, J. et al. Emotion detection from text and speech: a survey. Soc. Netw. Anal. Min. 8, 28 (2018). https://doi.org/10.1007/s13278-018-0505-2

[71] Alswaidan, N., Menai, M.E.B. A survey of state-of-the-art approaches for emotion recognition in text. Knowl Inf Syst 62, 2937–2987 (2020). https://doi.org/10.1007/s10115-020-01449-0

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

[72] Elkobaisi, M.R., Al Machot, F. & Mayr, H.C. Human Emotion: A Survey focusing on Languages, Ontologies, Datasets, and Systems. SN COMPUT. SCI. 3, 282 (2022). https://doi.org/10.1007/s42979-022-01116-x

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1787, Copyright ©2023
Submitted : June 13, 2023, Revised : July 7, 2023, Accepted : September 4, 2023, Published : November 7, 2023

Pendekatan *Macine Learning* dalam Evaluasi Label Berita Berdasarkan Judul: Studi Kasus Media Online

Rezky Yuranda[1]*, Tata Sutabri[2], Delpiah Wahyuningsih[3]
Program Studi Magister Teknik Informatika Universitas Bina Darma[1], [2],
Program Studi Teknik Informatika ISB Atma Luhur [3]
Palembang[1][2], Indonesia
Pangkalpinang[3], Indonesia
yurandarezky@gmail.com[1], tata.sutabri@gmail.com[2], delphibabel@atmaluhur.ac.id[3]

***Abstract*— In the current digital era, information availability is abundant, and news serves as a primary source of up-to-date and reliable information for the public. However, with the increasing volume of information, a robust evaluation method is necessary to ensure accurate and dependable news labeling. This research employs a machine learning approach, utilizing three common classification algorithms: Naive Bayes, SVM, and Random Forest, to evaluate news labels based on their titles. The dataset utilized in this study is obtained from Jakarta AI Research and consists of 10,000 samples covering various news topics. Evaluation is conducted using accuracy, precision, recall, and F1-Score metrics to gain a comprehensive understanding of the classification algorithm's performance. The results of this research demonstrate that the SVM algorithm exhibits the best performance, achieving an accuracy rate of 92.92%. Random Forest follows with an accuracy rate of 91.21%, and Naive Bayes with an accuracy rate of 89.61%. These findings provide deep insights into the effectiveness of the machine learning approach in evaluating news labels based on their titles. Furthermore, the study highlights the importance of considering other evaluation metrics such as precision, recall, and F1-Score to obtain a more holistic understanding of the algorithm's performance. Further research is encouraged to involve additional classification algorithms and more diverse and extensive datasets to enhance the comprehension of news label evaluation comprehensively. Such endeavors can significantly contribute to the development of automated systems for classifying news with higher accuracy and reliability in the future**

Keywords: News label evaluation, machine learning, Naive Bayes, SVM, Random Forest

***Abstrak*—Dalam era digital saat ini, ketersediaan informasi sangat melimpah. Berita menjadi sumber informasi terkini dan terpercaya bagi masyarakat. Namun, dengan informasi yang semakin banyak, diperlukan sebuah metode evaluasi yang baik untuk memastikan label berita yang akurat dan dapat diandalkan. Penelitian ini menggunakan pendekatan machine learning dengan menggunakan tiga algoritma klasifikasi umum: Naive Bayes, SVM, dan Random Forest untuk mengevaluasi label berita berdasarkan judul. Dataset yang digunakan berasal dari Jakarta AI Research dan mencakup berbagai topik berita sebanyak 8754 sample. Evaluasi dilakukan dengan metrik akurasi, presisi, recall, dan F1-Score guna mendapatkan gambaran komprehensif tentang performa algoritma klasifikasi. Hasil penelitian ini menunjukan bahwa algoritma SVM memiliki kinerja terbaik, dengan tingkat akurasi terbaik mencapat 92.92%. Diikuti Random Forest dengan tingkat akurasi 91.21% dan Naive Bayes dengan tingkat akurasi 89.61%. Temuan ini memberikan wawasan mendalam tentang efektivitas pendekatan machine learning dalam evaluasi label berita berdasarkan judul. Selain itu, hasil penelitian ini juga menekankan pentingkanya mempertimbangkan metrik evaluasi lainnya seperti presisi, recall, dan F1-Score untuk mendapatkan pemahaman yang lebih holistik mengenai kinerja algoritma. Penelitian lebih lanjut dianjurkan untuk melibatkan lebih banyak algortima klasifikasi serta dataset yang lebih luas dan beragam guna meningkatkan pemahaman evaluasi label berita secara lebih komprehensif. Hail ini dapat memberikan sumbangan penting dalam pengembangan sistem otomatis untuk mengklasifikasi berita dengan akurasi dan keandalan yang lebih tinggi di masa depan.**

Kata Kunci: Evaluasi label berita, machine learning, naive bayes, svm, random forest

I. PENDAHULUAN

Dalam era digital saat ini, ketersediaan informasi sangat melimpah. Berita menjadi bagian penting dari informasi tersebut, karena berita menjadi sumber informasi terkini dan terpercaya bagi masyarakat [1].

Bersamaan dengan pertumbuhan teknologi saat ini, terjadi pula peningkatan jumlah informasi tekstual yang tersedia di dalamnya. Data dari *Indonesia Digital Association* (IDA) menunjukan bahwa *96%* masyarakat Indonesia mengkonsumsi berita secara online [2].

Pada sebuah berita, judul utama atau headline merupakan elemen penting dari setiap halaman web. Judul tersebut menentukan apakah pembaca akan tetap berinteraksi dengan konten atau pergi ke halaman lain. Begitu juga dengan kategori atau label yang membantu pembaca untuk menemukan konten berdasarkan topik dan memberikan konteks pada artikel yang di baca [3]. Oleh sebab itu, maka penting untuk memiliki metode evaluasi yang baik dalam mengelola label berita untuk memastikan informasi yang disajikan akurat dan dapat diandalkan bagi konsumen [2].

Dalam hal ini, penggunaan dataset berkualitas dan repesentatif sangat penting dalam melakukan penelitian dan pengembangan algoritma *AI* untuk evaluas label berita. Salah satu dataset yang digunakan dalam penelitian ini berasal dari *Jakarta AI Research*. *Jakarta AI Research* atau yang biasa disebuat *Jakarta Research* adalah sebuah komunitas riset dibidang kecerdasan buatan yang meliputi natural language

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023

processing, computer vision, dan speech processing dan juga intersection antara field tersebut dengan machine learning dan deep learning [13].

Dataset yang disediakan oleh *Jakarta AI Reserach* menawarkan akses ke beberapa media berita online yang mencakut topik yang beragam. Kumpulan data ini mencakup judul-judul berita yang dapat dijadikan dasar untuk melakuka analisis dan evaluasi label dari berita dengan menggunakan pendekatan machine learing. Dalam penelitian ini, akan memanfaatkan dataset tersebut untuk melihat sejauh mana pendekatan machine learning dapat memberikan hasil yang akurat dalam mengevaluasi label berita berdasarkan judul.

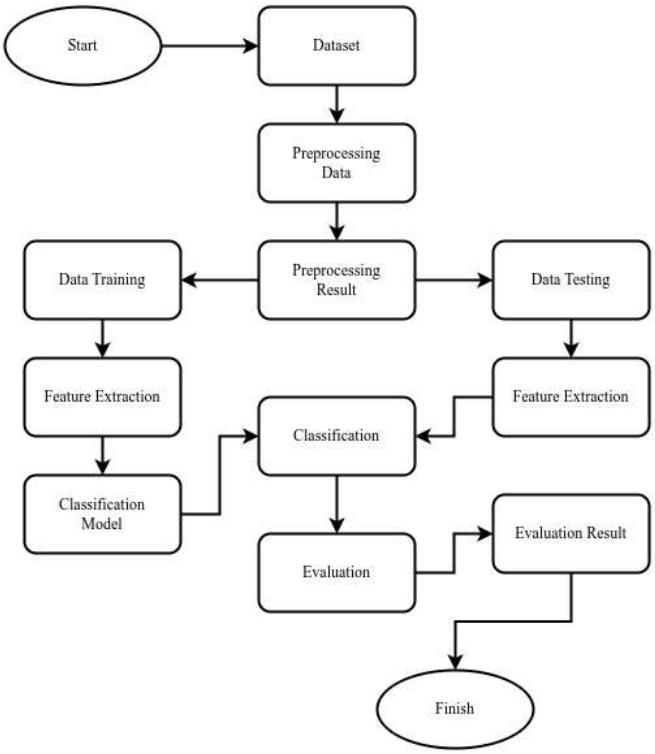
https://journals.usm.ac.id/index.php/transformatika/article/view/2317/1668

Pada penelitian ini, digunakan beberapa algoritma klasifikasi yang umum digunakan, seperti *Naive Bayes*, *Support Vector Machines* (*SVM*), dan *Random Forest*, untuk membangun model klasifikasi label berita berdasarkan judul. Algoritma-algoritma ini memiliki keunggukan masing-masing dalam memahami pola dan hubungan dalam data judul berita, yang kemudian akan digunakan untuk mengklasifikasi berita menjadi berbagai label atau kategori.

Dalam eksperimen ini akan dibandingkan performa ketiga algoritma klasifikas tersebut berdasarkan berbagai metrik evaluasi, seperti *akurasi*, *presisi*, *recall*, dan *F1-Score*. Hasil penelitian ini diharapkan dapat memberikan pemahaman lebih baik tentang efektivitas pendekatan machine learning dalam proses evaluasi label berita berdasarkan judul.

II. METODELOGI PENELITIAN

Metodelogi penelitian untuk pendekatan machine learning dalam evaluasi label berita berdasarkan judul, ditampilkan pada gambar 1.



Gambar 1. Alur penelitian

A. *Dataset*

Langkah pertama mempersiapkan dataset yang akan digunakan untuk pelatihan dan pengujian. Dataset yang digunakan pada penelitian ini diambil dari *Jakarta AI Reserach* yang berisikan informasi judul berita beserta label atau kategorinya sebanyak 8754 data. Pada dataset ini, terdapat didefinisikan 5 label, yaitu bola, news, bisnis, tekno, dan otomotif. Untuk distribusi data per label dapat dilihat pada table 1.

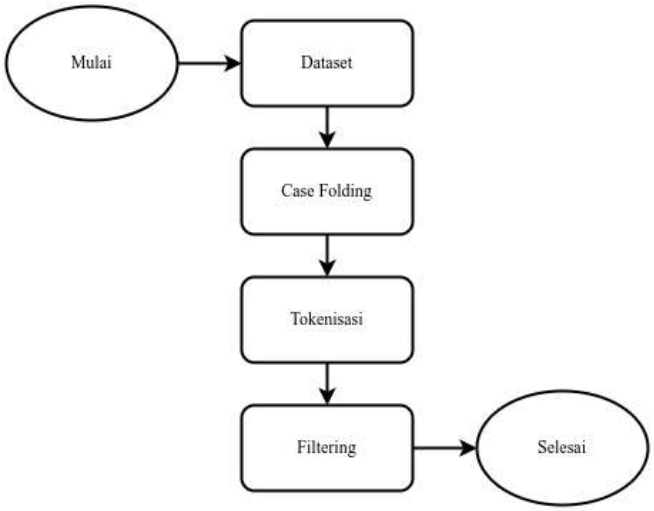
TABEL 1. DISTRIBUSI DATA SETIAP LABEL

No	***Label***	***Jumlah***
1	Bola	2986
2	News	2897
3	Bisnis	1881
4	Tekno	790
5	Otomotif	200
Jumlah Total		**8754**

B. *Preprocessing Data*

Tahapan preprocessing yang sangat penting pada data mining, alasan utamanya adalah karena kualitas dari input data sangatmempengaruhi kualitas output analisis yang dihasilkan [4]. Tahapan Preprocessing data dapat dilihat pada gambar 2.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023



Gambar 2. Preprocessing data

C. Feature extraction

Pada tahap *feature extraction*, dilakukan extraksi fitur dar teks judul berita menggunakan metode *TF-IDF* (*Term Frequency-Inverse Document Frequency*). Metode ini digunakan untuk mengukur tingkat pentingnya sebuah kata dalam suatu dokumen atau dataset berdasarkan frekuensi kemunculan kata dalam dokumen tersebut [6]. Kata-kata terpilih dianggap lebih informatif dan efektif untuk dieksekusi pada model classifier [7][8].

Penerapan teknik *TF-IDF* dilakukan setelah *preprocessing* data dan dilakukan sebelum proses klasifikasi data. Secara lebih detail, *TF-IDF* menghitung skor atau bobot untuk setiap kata dalam dokumen dengan mempertimbangkan dua faktor yaitu *Term Frequency* (*TF*) dan *Inverse Document Frequency* (*IDF*) [9].

$$TF = \frac{fk}{ft} \qquad (1)$$

Pada persamaan (1), *fk* adalah jumlah kemunculan kata dalam suatu kalimat dan *ft* adalah total kata dalam suatu dokumen. *IDF* menggambarkan seberapa umum atau jarang kata tersebut muncul dalam seluruh dataset. *Inverse Document Frequency* dihitung dengan menggunakan rumus

$$IDF = log \qquad (2)$$

Dimana *n* pada konteks ini merujuk pada total jumlah dokumen pada suatu korpus, sementara *df* adalah singkatan dari jumlah dokumen yang mengandung kata tertentu pada suatu korpus.

D. Classification model

Dalam metode pendekatan machine learing dalam evaluasi label berita berdasarkan judul pada penelitan ini menggunakan tiga algoritma klasifikasi yang berbeda yaitu *Naive Bayes*, *Support Vector Machine* (*SVM*), dan *Random Forest*.

Naive Bayes merupakan salah satu algoritma klasifikasi yang berdasarkan pada *Teorema Bayes* [10]. Algoritma ini memprediksi label berita berdarkan probalitas kelas yang diestimasikan dari fitur-fitur yang ada.

SVM (*Support Vector Machines*) adalah algoritma klasifikasi yang bekerja dengan mencari *hyperplane* terbaik yang memisahkan dua kelas berita berdasarkan fitur-fitur nya. *SVM* memiliki kemampuan untuk menangani data yang kompleks dan memiliki kelebihan dalam menangani dataset yang memiliki dimensi yang tinggi [11].

Random Forest merupakan algoritma klasifikasi ensemble yang membangun sejumlah pohon keputusan secara acak dan mengkombinasikan hasil prediksi dari setiap pohon untuk menghasilkan prediksi akhir. Dengan cara ini, *Random Forest* dapat mengatasi overfitting pada dataset dan menghasilkan hasil prediksi yang lebih stabil [11].

Dengan menggunakan kombinasi dari *Naive Bayes*, *SVM*, dan *Random Forest*, dapat dibandingkan dan dilakukan evaluasi performa dari ketiga algoritma ini dalam memprediksi dan mengklasifikasi label berita berdasarkan fitur-fitur yang diekstraksi.

E. Classification

Pada tahap ini, model klasifikasi yang telah di latih akan digunakan untuk memprediksi label berita berdasarkan fitur-fitur yang telah di ekstraksi sebelumnya. Proses ini melibatkan penggunaan rumus atau algoritma tertentu yang dapat memberikan prediksi yang akurat.

Pada algoritma *Naive Bayes*, prediksi label berita dilakukan dengan menggunakan *Teorema Bayes*. Untuk setiap kategori *k*, nilai probalitas posterior dapat dihitung dengan persamaan 3.

$$P(k \vee x) = \frac{(p(x \vee k) * p(k))}{p(x)} \qquad (3)$$

Dimana P(*k*/*x*) adalah probabilitas kondisional dari *k* terhadap *x*, P(*x*/*k*) adalah probabilitas kondisional dari *x* terhadap *k*, P(*k*) adalah probabilitas prior dari *k*, dan P(*x*) adalah probabilitas margin dari *x*.

Pada *SVM*, prediksi label berita dilakukan dengan mencari hyperplane terbaik yang memisahkan dua kelas berdasarkan fitur-fitur yang ada. Rumus untuk memprediksi label pada SVM terdapat pada persamaan 4.

$$f(x) = sign(\sum_{i=1}^{n} \alpha_i \, y_i K(x_i, x) + b) \qquad (4)$$

Dimana *x* adalah input sample yang ingin di prediksi, *xi* adalah training sample ke-*i*, *yi* adalah class label untuk training sample ke-*i* (mengambil nilai -1 atau 1), α*i* adalah koefisien untuk traiing sample ke-*i* yang dihitung oleh algoritma SVM selama proses training, K(*xi*,*x*) adalah kernel function yang mengukur kesamaan anatar input sample *x* dan training sample ke-*i*, *b* adalah bias term yang juga dihitung oleh algoritma *SVM* selama proses training, dan *n* adalah jumlah total training sample.

Dalam *Random Forest*, prediksi label berita dilakukan dengan menggabungkan hasil prediksi dari banyak pohon

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023

keputusan yang terbentuk. Setiap pohon memberikan suara dalam memprediksi label, dan label yang paling banyak terpilih menjadi prediksi akhir. Rumus yang digunakan dalam prediksi Random Forest tidak memiliki matematis khusus, tetapi melibatkan agregasi hasil prediksi dari setiap cabang. .

F. Evaluasi

Pada tahap ini, dilakukan pengukuran performa model klasifikasi yang telah dibangun menggunakan metrik-metrik seperti *akurasi*, *presisi*, *recall*, dan *F1-Score*.

Akurasi adalah metrik yang mengukur sejauh mana model mampu memberikan prediksi yang benar secara keseluruhan. *Presisi* mengukur sejauh mana prediksi positif yang diberikan oleh model adalah benar, sementara *recall* mengukur sejauh mana model mampu mendeteksi semua kasus positif yang sebenarnya. *F1-Score* adalah metrik yang menggabungkan presisi dan recal untuk memberikan gambaran yang lebih holistik tentang performa model.

Dalam tahapan evaluasi, kita menganalisis hasil prediksi model klasifikasi dengan membandingkan prediksi dengan label sebenarnya. Dengan menghitung *akurasi*, *presisi*, *recal* dan *F1-score*, kita dapat memperoleh pemahaman yang lebih mendalam tentang kualitas model klasifikasi dalam tugas evaluasi label berita berdasarkan judul. Metrik-metrik ini akan membantu kita dalam hal mengevaluasi tingkat keberhasilan model dalam memprediksi label berita dengan akurasi tinggi, dan juga membantu kita memahami seberapa baik model dapat menghindari kesalahan prediksi yang mungkin terjadi [12].

III. HASIL DAN PEMBAHASAN

A. Preprocessing Data

Case-folding, yaitu tahapan untuk mengkonversi semua karakter dengan huruf kapital menjadi huruf kecil. Proses l*owercasing* dapat dilihat pada table 2.

TABEL 2. PROSEES CASE-FOLDING

Data Asli	***Case-folding***
London - Lee Dixon khawatir Arsenal tak bisa merekrut Denis Suarez secara permanen musim panas nanti!!	london - lee dixon khawatir arsenal tak bisa merekrut denis suarez secara permanen musim panas nanti!!

Punctuation Removal, karakter khusus atau tanda baca seperti titik, koma, tanda tanya, tanda seru, dan karakter khusus lainnya dihapus dari teks. Proses punctuation removal dapat dilihat pada table 3.

TABEL 3. PROSES PUNCTUATION REMOVAL

Data Hasil Case-folding	***Punctuation Removal***
london - lee dixon khawatir arsenal tak bisa merekrut denis suarez secara permanen musim panas nanti!!	london lee dixon khawatir arsenal tak bisa merekrut denis suarez secara permanen musim panas nanti

Tokenisasi, adalah proses pemotongan kata menjadi satu kata pada setiap data. Proses tokenisasi dapat dilihat pada table 4.

TABEL 4. PROSES TOKENISASI

Data Hasil Punctuation Removal	***Tokenisasi***
london - lee dixon khawatir arsenal tak bisa merekrut denis suarez secara permanen musim panas nanti!!	['london', 'lee', 'dixon', 'khawatir', 'arsenal', 'tak', 'bisa', 'merekrut', 'denis', 'suarez', 'secara', 'permanen', 'musim', 'panas', 'nanti']

Filtering, adalah proses terakhir dalam *preprocessing* data, pada tahap ini akan menyaring kata penghubung yang bersifat umum dan tidak berpengaruh terhadap proses pembagian kelas sehingga kata penghubung tersebut tidak diperlukan. Pada tahap ini juga hasil tokenisasi akan digabungkan kembali menjadi kalimat utuh kembali. Proses filtering dapat dilihat pada table 5.

TABEL 5. PROSES TOKENISASI

Data Hasil Tokenisasi	***Filtering***	***Finishing***
['london', 'lee', 'dixon', 'khawatir', 'arsenal', '**tak**', '**bisa**', 'merekrut', 'denis', 'suarez', '**secara**', 'permanen', 'musim', 'panas', '**nanti**']	['london', 'lee', 'dixon', 'khawatir', 'arsenal', 'merekrut', 'denis', 'suarez', 'permanen', 'musim', 'panas']	Lodon lee dixon khawatir arsenal merekrut denis suarez permanan musim panas

B. Feature extraction

Berikut ini adalah sebagian hasil output dari pembobotan kata menggunakan metode *Term Frequency - Inverse Document Frequency* (TF-IDF) yang dapat dilihat pada Gambar 3.

```
(0, 43669)    0.057985814282321524
(0, 9269)     0.08402188022875495
(0, 37360)    0.08161037736900699
(0, 6297)     0.038826642368675224
(0, 7200)     0.04984697959846997
(0, 21190)    0.051249282037551386
(0, 35941)    0.08402188022875495
(0, 34644)    0.09151269230153071
(0, 2301)     0.061719635833768875
(0, 34640)    0.07359229206906777
(0, 2165)     0.07964003606623181
:     :
(7002, 41888) 0.07923955782363404
(7002, 526)   0.05869590141705259
(7002, 20763) 0.05009213389423332
(7002, 2744)  0.06877615482504694
(7002, 2300)  0.05522019326981233
(7002, 33195) 0.04301106169596901
(7002, 22879) 0.05789511454843924
(7002, 28947) 0.04496987612275906
```

Gambar 3. *Feature extraction*

C. Evaluasi

Naive bayes.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023

Pada tahap klasifikasi menggunakan Multinomial Naïve Bayes, pengujian akan dilakukan sebanyak 5 (lima) kali dengan membagi data menjadi data pelatihan (train) dan data uji (test) yang berbeda secara acak. Hal ini bertujuan untuk melihat hasil terbaik dari pengujian dengan variasi kombinasi data train dan data test. Hasil pengujian dapat dilihat pada table 6.

TABEL 6. HASIL PENGUJIAN NAIVE BAYES

Bobot data		***Score***			
Latih	**Uji**	**Akurasi**	**Presisi**	**Recall**	**F1-Score**
90%	10%	0.89612	0.89612	0.89612	0.88800
80%	20%	0.88007	0.88963	0.88007	0.87062
70%	30%	0.87514	0.88572	0.87514	0.86347
60%	40%	0.86465	0.87673	0.86465	0.85094
50%	50%	0.85995	0.85088	0.85995	0.84505

Berdasarkan hasil klasifikasi *Naive Bayes* diatas, nilai akurasi terbaik adalah 89.61% dengan pembagian data latih 90% dan data uji 10% dari jumlah data

Support Vector Machines (SVM)

Pada tahap klasifikasi dengan *Support Vector Machine* dilakukan juga pengujian sebanyak 5 (lima) kali seperti pengeujian pada naive bayes sebelumnya. Hasil pengujian dapat dilihat pada table 7.

TABEL 7. HASIL PENGUJIAN SUPPORT VECTOR MACHINES (SVM)

Bobot data		***Score***			
Latih	**Uji**	**Akurasi**	**Presisi**	**Recall**	**F1-Score**
90%	10%	0.92922	0.93098	0.92922	0.92932
80%	20%	0.92404	0.92606	0.92404	0.92416
70%	30%	0.91740	0.92005	0.91740	0.91719
60%	40%	0.91034	0.91375	0.91034	0.90962
50%	50%	0.90199	0.90709	0.90199	0.90108

Berdasarkan hasil pada table 7, nilai akurasi terbaik pada *Support Vector Machines* adalah 92.92% dengan pembagian data latih 90% dan data uji 10% dari jumlah data.

Random Forest

Pada tahap klasifikasi dengan *Random Forest*, dilakukan juga pengujian sebanyak 5 (lima) kali seperti pengeujian pada naive bayes sebelumnya. Hasil pengujian dapat dilihat pada table 8.

TABEL 8. HASIL PENGUJIAN RANDOM FOREST

Bobot data		***Score***			
Latih	**Uji**	**Akurasi**	**Presisi**	**Recall**	**F1-Score**
90%	10%	0.91210	0.91582	0.91210	0.91027
80%	20%	0.90520	0.90876	0.90520	0.90386
70%	30%	0.89912	0.90397	0.89912	0.91719
60%	40%	0.89063	0.89630	0.89063	0.88762
50%	50%	0.88805	0.89397	0.88805	0.88468

Berdasrakan hasil pengujian pada table 8, nilai akurasi terbaik pada *Random Forest* adalah 91.21% dengan pembagian data latih 90% dan data uji 10% dari jumlah data.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan terhadap tiga model kelasifikasi, yaitu *Naive Bayes*, *SVM*, dan *Random Forest*, Hasil penelitian menunjukkan bahwa algoritma *SVM* (*Support Vector Machine*) memiliki kinerja terbaik dalam evaluasi label berita, dengan tingkat akurasi mencapai 92.92%. Disusul oleh algoritma Random Forest dengan tingkat akurasi 91.21%, dan algoritma Naive Bayes dengan tingkat akurasi 89.61%.

Dengan demikian, penggunaan algoritma *SVM* dalam mengelola label berita berdasarkan judul utama menunjukkan hasil yang paling mengesankan, memberikan tingkat akurasi tertinggi dalam klasifikasi label berita. Hasil ini menunjukkan bahwa *SVM* adalah pilihan yang baik untuk memastikan label berita yang akurat dan dapat diandalkan bagi konsumen. Penelitian ini memberikan sumbangan penting bagi pengelolaan informasi berita dalam era digital yang penuh dengan informasi tekstual. Dengan metode evaluasi yang baik, diharapkan informasi berita yang disajikan dapat menjadi lebih relevan dan bermanfaat bagi pembaca, membantu mereka menemukan konten yang sesuai dengan minat dan kebutuhan mereka. Selain itu, penelitian ini juga dapat menjadi dasar bagi penelitian selanjutnya dalam mengembangkan dan meningkatkan kinerja algoritma Machine Learning lainnya dalam pengelolaan label berita

Penelitian ini memiliki kelebihan dalam melibatkan tiga model klasifikasi yang berbeda, memberikan perbandingan yang komprehensif dalam hal performa. Metode pemrosesan teks yang baik, termasuk tokenisasi, pengahpusan *stop words*, dan vektorisasi *TF-IDF*, juga digunakan dalam penelitian ini.

Namun penelitian ini memiliki beberapa kekurangan. Keterbatasan pada dataset yang digunakan, seperti jumlah sample, distribusi label yang tidak seimbang dan representasi kelas, sehingga tidak dapat memberikan pemahaman yang lebih baik tentang performa model. Penelitian ini hanya menggunakan algoritma klasifikasi yang umum digunakan dan tidak mempertimbangkan model kasifikasi lainnya yang mungkin memberikan performa yang lebih baik.

Dalam penelitian selanjutnya, disarankan untuk memperluas analisis dengan melibatkan lebih banyak model klasifikasi dan mempertimbangkan variasi parameter untuk meningkatkan performa. Selain itu, dataset yang lebih beragam dan representatif juga dapat membantu untuk mengevaluasi dan membandingkan model secara lebih akurat.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023

REFERENCES

[1] D. Rani dan S. D. Setiawati, "Penyajian Jurnalistik Online Infobdg untuk Menjadi Sumber Informasi Kredibel," J. Jurnalisa, vol. 6, no. 2, pp. 233–247, 2020.

[2] F. S. Nurfikri dan M. S. Mubarok, "Klasifikasi Topik Berita Menggunakan," vol. 5, no. 1, pp. 1579–1588, 2018.

[3] Briggs, Mark. Journalism next: A practical guide to digital reporting dan publishing. CQ Press, 2013.

[4] H. Junaedi, H. Budianto, I. Maryati, dan Y. Melani, "Data Transformation pada Data Mining," Pros. Konf. Nas. Inov. dalam Desain dan Teknol., vol. 7, pp. 93–99, 2011.

[5] T. Jamaluddin dan dkk, "Perbdaningan Algoritma Sentencepiece BPE dan Unigram Pada Tokenisasi Artikel Bahasa Indonesia Pendahuluan Studi Terkait," e-Proceeding Eng., vol. 7, no. 2, pp. 8323–8331, 2020.

[6] M. J. Lavin, Z. Leblanc, dan Q. Dombrowski, "The Programming Historian Analyzing Documents with TF-IDF," Program. Hist., pp. 1–21, 2020.

[7] D. Wahyuningsih dan E. Patima, "Penerapan Naive Bayes Untuk Penerimaan Beasiswa," Telematika, vol. 11, no. 1, p. 135, 2018, doi: 10.35671/telematika.v11i1.665.

[8] Y. Zhai, W. Song, X. Liu, L. Liu, dan X. Zhao, "A Chi-square Statistics Based Feature Selection," 2018 IEEE 9th Int. Conf. Softw. Eng. Serv. Sci., pp. 160–163, 2018.

[9] Schutze, Hinrich, Christopher D. Manning, dan Prabhakar Raghavan. Introduction to information retrieval. Cambridge University Press, 2008.

[10] R. Wati, "Penerapan Algoritma Naive Bayes Dan Particle Swarm Optimization Untuk Klasifikasi Berita Hoax Pada Media Sosial," JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer), vol. 5, no. 2, pp. 159–164, 2020, doi: 10.33480/jitk.v5i2.1034.

[11] Han, Jiawei, Jian Pei, dan Hanghang Tong. Data mining: concepts and techniques. Morgan kaufmann, 2022.

[12] S. Sudianto, A. D. Sripamuji, I. Ramadhanti, R. R. Amalia, J. Saputra, dan B. Prihatnowo, "Penerapan Algoritma Support Vector Machine dan Multi-Layer Perceptron pada Klasisifikasi Topik Berita," J. Nas. Pendidik. Tek. Inform. JANAPATI, vol. 11, no. 2, pp. 84–91, 2022.

[13] A. Candra, "Jakarta Artificial Intelligence Research is Now Open!" [Online]. Tersedia: https://medium.com/data-folks-indonesia/jakarta-artificial-intelligence-research-is-now-open-f404763867b1. Diakses pada: July 15, 2023.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1808, Copyright ©2023
Submitted : June 19, 2023, Revised : July 20, 2023, Accepted : July 26, 2023, Published : November 7, 2023

Perbandingan Ciri Parameter Tapis Gabor untuk Otentikasi *Dorsal Hand Vein* Menggunakan *Artificial Neural Network*

Wahyu Irwan Putra[1]*, Muchtar Ali Setyo Yudono[2], Alun Sujjada[3]
Program Studi Teknik Elektro[1], [2], Program Studi Teknik Informatika[3]
Universitas Nusa Putra
Sukabumi, Indonesia
wahyu.irwan_te19@nusaputra.ac.id[1] , muchtar.alisetyo@nusaputra.ac.id[2] , alun.sujjada@nusaputra.ac.id[3]

***Abstract*—The importance of digital security in today's technological era requires various innovations in creating a reliable security system for humans. Biometrics is an authentication method and the most effective system for performing personal recognition because biometrics have unique characteristics. Dorsal hand vein become biometrics for the individual recognition process in this study using feature extraction of gabor filters and neural network backpropagation to classify recognition into five classes of human individuals, which are expected to be able to provide a higher accuracy value when compared to research on the introduction of dorsal hand vein. This classification process has several stages, namely input image, image pre-processing, segmentation, feature extraction, and image classification. The test results show that the percentage of success based on the five test scenarios has an average value of 75%. In this study, the results of the greatest test accuracy in the fourth scenario were 91%.**

Keywords—Backpropagation Neural Netwotrk, Biometric, Camera NIR LED, Dorsal Hand Vein, Gabor Filter

***Abstrak*—Pentingnya keamanan digital di era teknologi saat ini menuntut berbagai inovasi dalam menciptakan sistem keamanan yang handal bagi manusia. Biometrik merupakan metode otentikasi dan sistem yang paling efektif untuk melakukan pengenalan personal karena biometrik memiliki karakteristik yang unik. Pembuluh darah punggung tangan menjadi biometrik untuk proses pengenalan individu pada penelitian ini menggunakan fitur ekstraksi filter gabor dan neural network backpropagation untuk mengklasifikasikan pengenalan menjadi lima kelas individu manusia, yang diharapkan mampu memberikan nilai akurasi yang lebih tinggi. jika dibandingkan dengan penelitian tentang pengenalan pembuluh darah dorsal tangan sebelumnya. Proses klasifikasi ini memiliki beberapa tahapan, yaitu citra masukan, pra-pengolahan citra, segmentasi, ekstraksi ciri, klasifikasi citra, dan hasil klasifikasi citra pembuluh darah. Hasil pengujian menunjukkan bahwa persentase keberhasilan berdasarkan lima skenario pengujian memiliki nilai rata-rata 75%. Pada penelitian ini didapatkan hasil akurasi pengujian terbesar pada skenario keempat sebesar 91%.**

Kata Kunci— Biometrik, Jaringan Syaraf Tiruan Perambatan Balik, Kamera NIR LED, Pembuluh Darah Punggung Tangan, Tapis Gabor

I. PENDAHULUAN

Pentingnya keamanan digital di era teknologi saat ini membutuhkan berbagai inovasi dalam menciptakan sistem keamanan yang andal untuk manusia, keamanan digital berfungsi agar dapat melindungi informasi pribadi, data privasi, dan izin akses yang dapat membahayakan jika di akses oleh Orang lain. Terdapat berbagai macam metode otentikasi dalam mengidentifikasi Personal pengguna, diantaranya user id, kartu identitas, kata kunci, kode pin, pengetahuan Personal dan biometrik. Pada penggunaan user id dan kartu identitas memiliki karakteristik yang dapat dibagikan dan bisa di duplikasikan sehingga dari segi keamanan metode ini sangat rentan mengalami perampasan hak akses. Begitu pun dengan kata kunci, kode pin, dan pengetahuan pribadi memiliki karakteristik yang bisa di lupakan dan di tebak sehingga orang yang mengetahui informasi tersebut dapat memiliki akses untuk menggunakan izin akses personal penggunan, biometrik merupakan metode otentikasi dan sistem yang paling efektif untuk melakukan otentikasi personal karena biometrik memiliki karakteristik yang unik dan setiap manusia pasti memilikinya [1].

Teknologi biometrik terus dikembangkan untuk menjadi sistem keamanan yang dapat menjaga data privasi penggunan, namun sebagai alat otentikasi identitas dalam penggunaannya banyak ditemukan berbagai kelemahan. Seperti pada sistem pengenalan wajah yang pada awal pengembanganya dapat terbuka hanya dengan foto dan topeng wajah realistis [2], dalam sistem pengenal sidik jari yang dapat di duplikasi dengan menggunakan sidik jari yang tertinggal pada permukaan dapat diambil dan di duplikasi dengan cepat[3]. Sedangkan penggunaan sistem iris mata memiliki nilai akurasi yang cukup tinggi namun memiliki kekurang dari segi kenyamanan pengguna karena iris terletak di dalam mata [4], [5].

Biometrik merupakan sebuah metode otentikasi dan menjadi sistem yang paling efektif untuk melakukan pengenalan Personal, karena biometrik memiliki karakteristik yang unik, tidak dapat dibagikan, tidak bisa dilupakan, dan tidak bisa hilang ataupun dirampas [6]. Teknologi saat ini telah menggunakan berbagai bentuk sistem biometrik

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

diantaranya yaitu Pengenal Wajah [7], Sidik Jari [8], Iris Mata [9], Garis Telapak Tangan [10], Pembuluh Darah Punggung Tangan [11], untuk melakukan otentikasi pengenalan Personal, seperti sistem keamanan sidik jari dan pengenalan wajah pada smartphone. Dalam penggunaan otentikasi sebuah biometrik harus memenuhi beberapa persyaratan diantaranya yaitu universal setiap manusia harus memiliki biometrik tersebut, memiliki keunikan dan mampu membedakan banyak manusia, tidak bisa hilang atau pun berubah [12]. Serta memiliki kinerja akurasi pengenalan yang baik dan cepat [13].

Pengenalan pembuluh darah punggung tangan atau disebut juga Dorsal Hand Vein (DHV) merupakan salah satu biometrik, pembuluh darah mempunyai karakteristik yang unik letaknya berada di dalam kulit manusia sehingga tidak mudah dipalsukan, dimodifikasi dan rusak. struktur pembuluh darah tidak akan berubah, penggunaan pembuluh darah punggung tangan untuk sistem otentikasi biometrik juga memiliki kelebihan dalam segi kebersihan karena dapat digunakan tanpa menyentuh permukaan sensor [14], [15].

Dalam bidang kecerdasan buatan menggunakan pengenalan pola sebagai alat untuk mengolah data dan dapat mengambil keputusan. Pengenalan pola merupakan sistem yang mampu bereaksi menghasilkan otentikasi data atau dapat mendefinisikan sesuatu di dasarkan pengukuran kuantitatif ciri atas sifat utama dari pada objek. Pola sendiri adalah hal yang bisa dikenali oleh sifat-sifatnya [16]. Citra digital terbentuk dari setiap elemen gambar yang disebut piksel, dimana piksel merupakan matriks baris x dan matriks kolom y dari sebuah citra. Parameter sebuah piksel adalah koordinat dan intensitas. dinyatakan pada bidang 2 dimensi f (x,y), dimana nilai pada koordinat ditunjukan (x,y) dan f merupakan amplitude [17]. Tujuan dari pengolahan citra untuk meningkatkan kualitas citra sehingga dapat dilihat oleh manusia ataupun komputer dengan mudah. Pengolahan citra mengubah sebuah citra menjadi citra lain, sehingga masukan dan keluaran dari pengolahan citra adalah citra, tetapi citra keluaran memiliki kualitas lebih tinggi[18]–[21].

Teknik ekstraksi ciri Gabor wavelet ditemukan oleh Dennis Gabor pada tahun 1946. Terdapat beberapa parameter pada filter tapis Gabor 2D yang dapat mempengaruhi hasil dari ekstraksi ciri, parameter θ yang mengontrol orientasi dari tapis gabor dengan variasi orientasi filter Gabor ini akan memungkinkan ekstraksi ciri dengan orientasi yang berbeda-beda, ukuran tapis (kernel) dengan skala yang lebih besar akan menangkap ciri yang lebih kasar dan skala yang lebih kecil akan menangkap ciri yang lebih halus. Parameter u (frekuensi dari gelombang sinusoida) yang dapat mempengaruhi deteksi tepi, dan parameter σ (standar deviasi dari *Gaussian envelope*) yang menyesuaikan dengan ukuran tapis. Sejumlah respon filter untuk titik tertentu pada citra akan didapatkan tergantung dari jumlah frekuensi dan sudut orientasi yang digunakan filter Gabor dengan frekuensi (u) berbeda seperti 2, 3, 4, 5, 6, dan 7 dan sudut orientasi (θ) diterapkan ke titik tersebut dalam citra. Citra yang menjadi subjek ekstraksi ciri digabungkan dengan setiap respon filter yang dibuat. Konvolusi akan menghasilkan ciri spesifik yang merepresentasikan atribut dari citra [22]–[24].

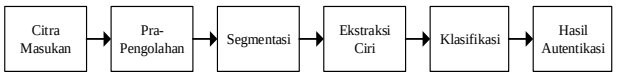
Artificial Neural Network (ANN) merupakan sebuah algoritma pengolahan informasi yang mempunyai karakteristik menyerupai sistem jaringan syaraf biologi pada otak manusia. Jaringan syaraf tiruan atau ANN dapat melakukan proses perhitungan komputasi berdasarkan klasifikasi serta memiliki kemampuan memorisasi dan menggeneralisasi suatu objek. Memorisasi ANN dapat mengingat kembali secara utuh dari sebuah data masukan yang sudah di latih, generalisasi merupakan kemampuan ANN untuk mendapatkan sebuah masukan yang belum pernah di dapatkan pada proses pelatihan [25]–[30].

Berdasarkan penelitian terdahulu yang dilakukan mengenai pengenalan pembuluh punggung tangan yang telah dilakukan berjudul "Teknik pengenalan pembuluh darah punggung tangan berbasis fitur local binary pattern" tahun 2021 [31]. Pada penelitian ini menggunakan 140 citra dengan 35 kelas, yang terbagi menjadi 105 citra sebagai data base dan 35 citra untuk data testing. Pra-Pengolahan citra menggunakan deteksi ROI dan mengubah citra menjadi citra keabuan. Ekstraksi dilakukan menggunakan metode Local Binary Pattern (LBP), tahapan klasifikasi menggunakan metode Fuzzy k-NN, mendapatkan hasil akurasi terbaik sebesar 90% dengan K-fold = 4.

Berdasarkan permasalahan yang ada maka perlu dibuat sebuah sistem pengenalan Pembuluh Darah Punggung Tangan untuk mengotentikasi biometrik yang lebih akurat dan optimal. Menggunakan sistem pengolahan citra yang didalamnya terdapat beberapa tahapan yaitu, pra-pengolahan, segmentasi, ekstraksi ciri, dan identifikasi. Diharapkan dengan sistem ini mampu menyelesaikan permasalahan yang ada. Pada penelitian ini akan menggunakan ekstraksi ciri Tapis Gabor dan metode otentikasi menggunakan ANN Perambatan Balik diharapkan mampu menghasilkan akurasi terbaik.

II. METODOLOGI PENELITIAN

Pada penelitian ini, dilakukan pengambilan data citra pembuluh darah punggung tangan menggunakan Kamera NIR-LED. Sistem pengenalan dan otentikasi pembuluh darah punggung tangan melibatkan empat tahap utama. Tahap-tahap tersebut mencakup pra-pengolahan, segmentasi, ekstraksi ciri menggunakan tapis Gabor, dan klasifikasi menggunakan metode *Artificial Neural Network* (ANN) perambatan balik, seperti yang terlihat pada Gambar 1.

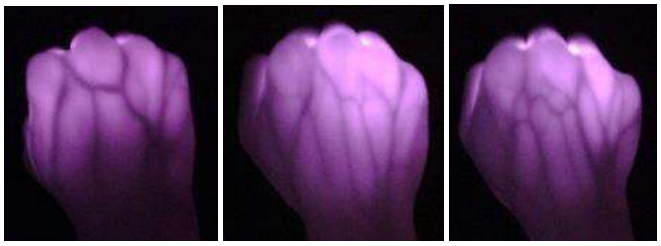


Gambar 1. Diagram Alir Proses Otentikasi

A. *Dataset Citra Pembuluh Darah Punggung Tangan*

Citra yang digunakan untuk penelitian ini adalah citra pembuluh darah punggung tangan manusia berjenis kelamin laki-laki berumur 18 tahun sampai 25 tahun dengan berat badan rata-rata 60 kg. Dengan jumlah 5 individu, setiap individu dilakukan pengambilan 220 citra di punggung tangan kanan dan kiri, citra pembuluh darah diambil dengan menggunakan Kamera NIR (Near-Infrared) LED. Gambar 2 merupakan citra pembuluh darah punggung tangan.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

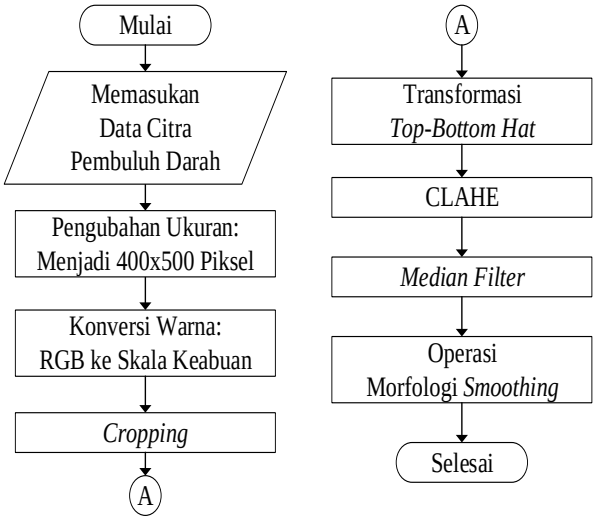


(a) (b) (c)

Gambar 2. Citra Pembuluh Darah Punggung Tangan: (a) Orang 1, (b) Orang 2, (c) Orang 3

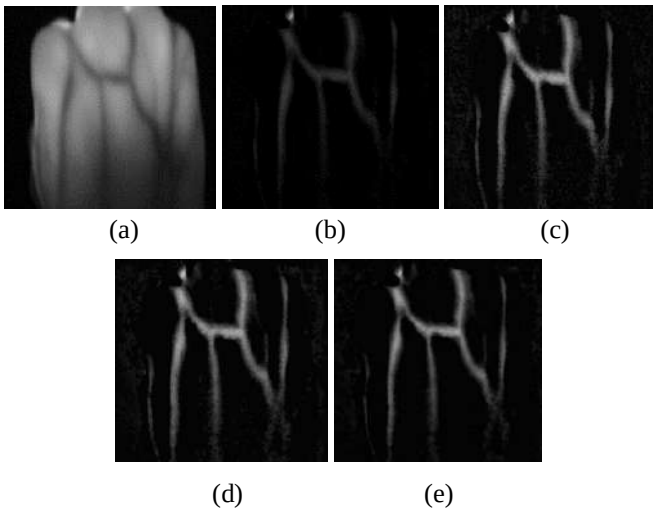
B. Pra-Pengolahan

Pada proses Pra-pengolahan citra pembuluh darah punggung tangan di awali dengan mengubah ukuran citra dengan format menjadi berukuran 400x500 piksel, Diagram alir pada proses pra-pengolahan dapat dilihat pada Gambar 3.



Gambar 3. Diagram Alir Proses Pra-Pengolahan

Proses berikutnya mengubah citra menjadi skala keabuan pada warna RGB. Proses selanjutnya dilakukan proses Morfologi *Opening* dan Morfologi *Closing* untuk menghasilkan citra yang baik pada proses *Top-bottom Hat Transformation* yang berfungsi untuk mempertajam pembatas antara kulit dan pembuluh darah dengan cara mengurangi nilai *Top Hat Transformation* menggunakan citra skala keabuan pada proses morfologi *Opening*, sedangkan *Bottom Hat Transformation* menggunakan citra skala keabuan pada proses morfologi *Closing* untuk mengurangi nilainya. Dengan melakukan pengurangan pada *bottom hat transformation* dan *top hat transformation* akan mendapatkan nilai transformasi *top-bottom hat transformation*. Namun keluaran dari *top-bottom hat transformation* belum memiliki histogram yang rata untuk menghasilkan citra yang cukup jelas, maka perlu dilakukan langkah selanjutnya yaitu CLAHE. Proses selanjutnya morfologi *smoothing* perlu dilakukan untuk pengurangan derau (noise) menggunakan Filter Median agar menghasilkan citra pembuluh darah punggung tangan yang jelas.



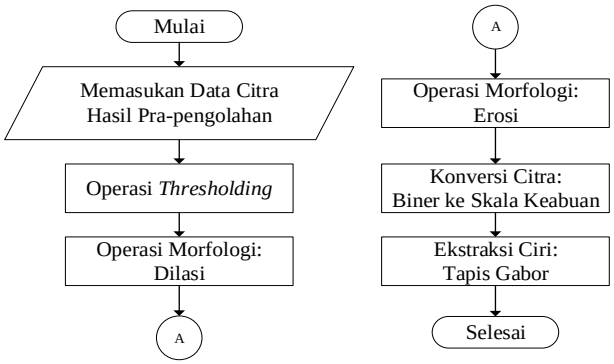
(a) (b) (c)

(d) (e)

Gambar 4. Citra Proses Pra-Pengolahan Pembuluh Darah Orang 1: (a) Citra RGB ke Keabuan, (b)). Citra Hasil Top-Bottom Hat, (c). Citra Hasil CLAHE, (d). Citra Proses Median Filter, (e). Citra Proses Morfologi Smoothing

C. Segmentasi

Proses ini diawali dengan menggunakan proses operasi pengambangan untuk mengubah konversi citra hasil morfologi *smoothing* menjadi sebuah citra biner. Gambar 5 menunjukan diagram alir proses segmentasi pada penelitian ini.



Gambar 5. Diagram Alir Proses Segmentasi

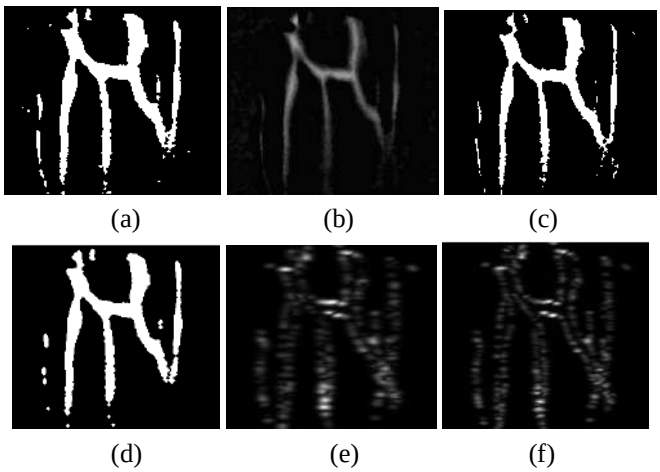
Proses selanjutnya adalah morfologi erosi agar mengurangkan citra yang tidak berguna pada pola citra pembuluh darah, selanjutnya adalah morfologi dilasi untuk memperjelas pola citra pembuluh darah, selanjutnya dilakukan proses pemotongan citra untuk memisahkan antara objek yang digunakan dan tidak di perlukan. Proses berikutnya dilakukan pengubahan warna citra menjadi citra keabuan dengan tujuan agar pada proses selanjutnya yaitu ekstraksi ciri tapis gabor dapat dilakukan dengan citra keabuan.

D. Ekstraksi Ciri

Pada tahap ekstraksi ciri, metode yang digunakan berdasarkan tekstur tapis Gabor dengan beberapa kombinasi dari nilai panjang gelombang, sudut orientasi dan parameter lainnya. Pada penelitian ini menggunakan panjang gelombang 3 dan sudut orientasi (θ) yaitu 0^c, 45^c, 90^c, 135^c, dan 180^c. Parameter yang digunakan dalam kombinasi metode ekstraksi ciri berdasarkan tekstur tapis Gabor adalah *mean*, *variance*,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

dan *entropy*.



Gambar 6. Citra Proses Segmentasi: (a). Citra Proses Pengambangan, (b). Citra Hasil Operasi Erosi Pada Pembuluh Darah, (c). Citra Hasil Operasi Dilasi Pada Pembuluh Darah, (d). Citra Hasil Konversi Berskala Keabuan, (e). Citra Hasil Tapis Gabor Dengan Nilai Panjang Gelombang 3 dan Orientasi Sudut 90, (f). Citra Dengan Nilai Panjang Gelombang 3 dan Orientasi Sudut 135

Tingkat keabuan dapat ditunjukkan oleh histogram pada sumbu X, suatu citra memiliki derajat piksel keabuan di mulai dari titik nol menuju tingkat keabuan yaitu 255. Frekuensi tingkat keabuan citra tersebut terlihat pada sumbu Y. Untuk dapat menghitung parameter ekstraksi ciri orde pertama, dengan menggunakan nilai-nilai pada histogram [67], beberapa parameter diantaranya:

1) Rerata: Menampilkan rata-rata nilai citra untuk intensitas di dalam citra.

$$\mu = \sum_n f_n\, p(f_n) \quad (1)$$

2) Varians: Menampilkan nilai bervariatif dari komponen pada histogram sebuah citra.

$$\sigma^2 = \sum_n (f_n - \mu)^2 p(f_n) \quad (2)$$

3) Entropi*:* Mengukur ukuran bentuk citra yang tidak memiliki keakuratan.

$$H = -\sum_n p(f_n)^2 \log p\,(f_n) \quad (3)$$

E. Identifikasi Menggunakan ANN Perambatan Balik

Algoritma Perambatan Balik sering dipakai untuk memodifikasi bobot yang terhubung ke neuron di lapisan tersembunyi oleh perceptron banyak lapisan. Hasil galat digunakan sebagai pengubah nilai bobot pada pembelajaran arah mundur pada algoritma perambatan balik. terlebih dahulu harus di lakukan pembelajaran arah maju untuk menghasilkan galat ini, fungsi aktivasi *logsig*, *tansig*, dan *purelin* perlu diaktifkan untuk menggunakan neuron yang bisa mendiferensiasikan. Untuk proses otentikasi pembuluh darah punggung tangan ini menggunakan Algoritma *Levenberg-Marquardt* (*Trainlm*). Proses selanjutnya pelatihan, kemudian dilakukan proses uji data latih dan di dapatkan hasil pelatihan serta di dapatkan nilai bobot optimal. Proses berikutnya adalah pengujian yaitu memuat data pengujian, lalu dilakukan proses uji data pengujian dan di dapatkan hasil data pengujian. Proses pelatihan akan optimal jika telah mendapatkan parameter arsitektur yang sesuai. Pengaturan parameter yang dapat dilihat pada Tabel I agar menghasilkan akurasi terbaik.

TABEL I. PARAMETER ANN

Parameter	Spesifikasi
Banyak neuron pada lapisan masukan	3
Banyak neuron pada lapisan tersembunyi 1	30
Banyak neuron pada lapisan tersembunyi 2	50
Banyak neuron pada lapisan keluaran	5
Fungsi Aktivasi	*Tansing-Logsig-Purelin*
Algoritma	*Trainlm*
Galat	10^{-5}
Iterasi	150
Learning Rate	0.5
Nilai Momentum Unit (*Mu*)	0.7
Gradient Minimum	10^{-8}
Set Maksimum Momentum Unit	10^{13}
Unit Momentum *Decrease*	0.1
Unit Momentum *Increase*	10

Pengujian dilakukan dengan berbagai skenario dengan parameter orde pertama ekstraksi ciri tapis gabor yang telah ditentukan dan dengan berbagai nilai panjang gelombang, orientasi sudut, dan jumlah iterasi yang berbeda-beda untuk mendapatkan hasil kombinasi terbaik, skenario kombinasi pengujian dapat dilihat pada Tabel II.

TABEL II. Skenario Pengujian

Skenario	Tapis Gabor	
	Panjang Gelombang	Sudut Orientasi
1	3	0°
2	3	45°
3	3	90°
4	3	135°
5	3	180°

F. Efektivitas Sistem

Dalam mengevaluasi kinerja model Otentikasi, penting untuk mempertimbangkan pengujian kemampuan sistem dalam melakukan prediksi yang akurat dan mampu memisahkan kelas yang berbeda. Saat melihat kesalahan yang dibuat oleh model otentikasi, diperlukan pendekatan statistik yang terkait dengan efektivitas sistem ini untuk menghitung nilai sensitivitas, spesifikasi, dan akurasi. Hal ini memungkinkan untuk menilai keakuratan sistem dengan lebih baik dan memahami sejauh mana sistem tersebut bekerja secara efektif.

Sensitivitas memberikan hasil seberapa baik sistem mampu dalam mengenali sampel yang ada pada kelasnya. Spesifikasi menunjukkan seberapa baik sistem dapat mengidentifikasi sampel di luar kelasnya.

Akurasi dihitung sebagai jumlah identifikasi akurat dibagi dengan jumlah total data identifikasi, dan ini menunjukkan hasil tingkat akurasi dalam sistem terhadap semua model.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

Untuk menampilkan bagaimana hasil prediksi yang dibuat oleh sistem atau dapat disebut juga tabel kontigensi karena matriksnya bisa di acak dibutuhkan sebuah tabel *Confusion matrix*, baris menunjukan kelas yang terdapat pada data. Tabel III adalah tabel confusion matrix pada proses otentikasi dengan 5 kelas klasifikasi.

TABEL III. CONFUSION MATRIX PADA OTENTIKASI 5 KELAS

Kelas Prediksi	Kelas Dikenali				
	A	**B**	**C**	**D**	**E**
A	tp_A	e_{AB}	e_{AC}	e_{AD}	e_{AE}
B	e_{BA}	tp_B	e_{BC}	e_{BD}	e_{BE}
C	e_{CA}	e_{CB}	tp_C	e_{CD}	e_{CE}
D	e_{DA}	e_{DB}	e_{DC}	tp_D	e_{DE}
E	e_{EA}	e_{EB}	e_{EC}	e_{ED}	tp_E

Sensitivitas (A):

$$\frac{TP}{TP+FN} = \frac{tp_A}{tp_A + e_{AB} + e_{AC} + e_{AD} + e_{AE}} x100\% \quad (4)$$

Spesifikasi (A):

$$\frac{TN}{TN+FP} =$$

$$\frac{tp_B + e_{BC} + e_{BD} + e_{BE} + e_{CB} + e_{CC} + e_{CD} + e_{CE} + e_{DB} + e_{DC} + e_{DD} + e_{DE} + e_{EB} + e_{EC} + e_{ED} + e_{EE}}{tp_B + e_{BC} + e_{BD} + e_{BE} + e_{CB} + e_{CC} + e_{CD} + e_{CE} + e_{DB} + e_{DC} + e_{DD} + e_{DE} + e_{EB} + e_{EC} + e_{ED} + e_{EE} + e_{BA} + e_{CA} + e_{DA} + e_{EA}} x100\% \quad (5)$$

Sensitivitas (B):

$$\frac{TP}{TP+FN} = \frac{tp_B}{tp_B + e_{BA} + e_{BC} + e_{BD} + e_{BE}} x100\% \quad (6)$$

Spesifikasi (B):

$$\frac{TN}{TN+FP} =$$

$$\frac{tp_A + e_{AC} + e_{AD} + e_{AE} + e_{CA} + e_{CC} + e_{CD} + e_{CE} + e_{DA} + e_{DC} + e_{DD} + e_{DE} + e_{EA} + e_{EC} + e_{ED} + e_{EE}}{tp_A + e_{AC} + e_{AD} + e_{AE} + e_{BC} + e_{CA} + e_{CC} + e_{CD} + e_{CE} + e_{DA} + e_{DC} + e_{DD} + e_{DE} + e_{EA} + e_{EC} + e_{ED} + e_{EE} + e_{AB} + e_{CB} + e_{DB} + e_{EB}} x100\% \quad (7)$$

Sensitivitas (C):

$$\frac{TP}{TP+FN} = \frac{tp_C}{tp_C + e_{CB} + e_{CB} + e_{CD} + e_{CE}} x100\% \quad (8)$$

Spesifikasi (C):

$$\frac{TN}{TN+FP} =$$

$$\frac{tp_A + e_{AB} + e_{AD} + e_{AE} + e_{BA} + e_{BB} + e_{BD} + e_{BE} + e_{DA} + e_{DB} + e_{DD} + e_{DE} + e_{EA} + e_{EB} + e_{ED} + e_{EE}}{tp_A + e_{AB} + e_{AD} + e_{AE} + e_{BA} + e_{BB} + e_{BD} + e_{BE} + e_{DA} + e_{DB} + e_{DD} + e_{DE} + e_{EA} + e_{EB} + e_{ED} + e_{EE} + e_{AC} + e_{BC} + e_{DC} + e_{EC}} x100\% \quad (9)$$

Sensitivitas (D):

$$\frac{TP}{TP+FN} = \frac{tp_D}{tp_D + e_{DA} + e_{DB} + e_{DC} + e_{DE}} x100\% \quad (10)$$

Spesifikasi (D):

$$\frac{TN}{TN+FP} =$$

$$\frac{tp_A + e_{AB} + e_{AC} + e_{AE} + e_{BA} + e_{BB} + e_{BC} + e_{BE} + e_{CA} + e_{CB} + e_{CC} + e_{CE} + e_{EA} + e_{EB} + e_{EC} + e_{EE}}{tp_A + e_{AB} + e_{AC} + e_{AE} + e_{BA} + e_{BB} + e_{BC} + e_{BE} + e_{CA} + e_{CB} + e_{CC} + e_{CE} + e_{EA} + e_{EB} + e_{EC} + e_{EE} + e_{AD} + e_{BD} + e_{CD} + e_{ED}} x100\% \quad (11)$$

Sensitivitas (D):

$$\frac{TP}{TP+FN} = \frac{tp_E}{tp_E + e_{EA} + e_{EB} + e_{EC} + e_{ED}} x100\% \quad (12)$$

Spesifikasi (D):

$$\frac{TN}{TN+FP} =$$

$$\frac{tp_A + e_{AB} + e_{AC} + e_{AD} + e_{BA} + e_{BB} + e_{BC} + e_{BD} + e_{CA} + e_{CB} + e_{CC} + e_{CD} + e_{DA} + e_{DB} + e_{DC} + e_{DD}}{tp_A + e_{AB} + e_{AC} + e_{AD} + e_{BA} + e_{BB} + e_{BC} + e_{BD} + e_{CA} + e_{CB} + e_{CC} + e_{CD} + e_{DA} + e_{DB} + e_{DC} + e_{DD} + e_{AE} + e_{BE} + e_{CE} + e_{DE}} x100\% \quad (13)$$

Akurasi:

$$\frac{tp_A + tp_B + tp_C + tp_D + tp_E}{tp_A + e_{AB} + e_{AC} + e_{AD} + e_{AE} + tp_B + e_{BA} + e_{BC} + e_{BD} + e_{BE} + tp_C + e_{CA} + e_{CB} + e_{CD} + e_{CE} + tp_D + e_{DA} + e_{DB} + e_{DC} + e_{DE} + tp_E + e_{EA} + e_{EB} + e_{EC} + e_{EL}} \quad (14)$$

III. HASIL DAN PEMBAHASAN

Proses Otentikasi citra pembuluh darah punggung tangan untuk mengkalsifikasikan kedalam 5 kelas individu. Hasil ini di dapatkan dengan menggunakan ekstraksi ciri tapis gabor dengan tiga parameter: rerata, varians, dan entropi. Ketiga ciri tersebut merupakan data masukan pada data pelatihan dan data pengujian, yang diproses menggunakan metode ANN perambatan balik Tabel IV merupakan hasil dari ekstraksi ciri.

TABEL IV. EKSTRAKSI CIRI FITUR ORDE PERTAMA

Ekstraksi Ciri	Orang 1	Orang 2	Orang 3	Orang 4	Orang 5
Rerata	378.207.863	414.321.343	344.967.152	585.433.813	353.681.555
Varians	372.626.413	393.389.86	387.657.715	336.329.892	122.534.622
Entropi	139.835.429	168.525.877	116.388.090	336.309.899	386.720.549

Terlihat pada Tabel IV hasil dari ekstraksi tapis gabor berdasarkan parameter nilai panjang gelombang dan sudut orientasi menghasilkan nilai yang berbeda, yang merupakan hasil dari ekstraksi ciri gabor dengan sudut orientasi 135° dengan panjang gelombang 3 serta parameter yang digunakan adalah rerata, varian, dan entropi. Dapat terlihat juga range pada tiap kelas ini saling bersebrangan terhadap beberapa kelas. Secara penglihatan manusia nilai-nilai tersebut cukup menyulitkan untuk dapat di identifikasikan. dengan menggunakan ANN perambatan balik pada penelitian ini akan dapat mengotentikasikan keempat kelas menggunakan nilai ekstraksi ciri yang saling bersebrangan tersebut.

TABEL V. HASIL AKURASI PELATIHAN

Skenario	Akurasi					
	Orang 1	Orang 2	Orang 3	Orang 4	Orang 5	Sistem
1	91%	81%	85%	75%	91%	84%
2	94%	94%	97%	94%	97%	95%
3	93%	85%	92%	94%	85%	89%
4	100%	97%	99%	100%	96%	98%
5	91%	86%	85%	83%	91%	87%
Rata - Rata	94%	89%	92%	89%	92%	91%

Dapat dilihat pada Tabel V bahwa hasil terbaik pada proses pengujian dicapai oleh skenario ke empat dengan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

akurasi sistem sebesar 98%, pada skenario ke empat menggunakan panjang gelombang bernilai 3 dan sudut orientasi 135 dengan parameter rerata, varian, dan entropi. Pada skenario pertama mendapatkan akurasi terkecil yaitu 84% dengan panjang gelombang bernilai 3 dan sudut orientasi 135.

TABEL VI. HASIL AKURASI PENGUJIAN

Skenario	Akurasi					
	Orang 1	Orang 2	Orang 3	Orang 4	Orang 5	Sistem
1	80%	70%	60%	75%	70%	71%
2	80%	75%	80%	75%	80%	78%
3	85%	65%	65%	70%	70%	71%
4	95%	85%	90%	95%	90%	91%
5	85%	65%	45%	65%	60%	64%
Rata-Rata	85%	72%	68%	76%	74%	75%

Pada Tabel VI dapat diketahui hasil terbaik saat proses pengujian dicapai oleh skenario ke empat dengan akurasi sistem sebesar 91%, pada skenario ke empat menggunakan panjang gelombang bernilai 3 dan sudut orientasi 135 dengan parameter rerata, varian, dan entropi. Pada skenario ke lima mendapatkan akurasi terkecil yaitu 64% dengan panjang gelombang bernilai 3 dan sudut orientasi 135.

TABEL VII. HASIL SENSITIVITAS PENGUJIAN

Skenario	Sensitivitas				
	Orang 1	Orang 2	Orang 3	Orang 4	Orang 5
1	80%	70%	60%	75%	70%
2	80%	75%	80%	75%	80%
3	85%	65%	65%	70%	70%
4	95%	85%	90%	95%	90%
5	85%	65%	45%	65%	60%
Rata-Rata	85%	72%	68%	76%	74%

TABEL VIII. HASIL SPESIFIKASI PENGUJIAN

Skenario	Spesifikasi				
	Orang 1	Orang 2	Orang 3	Orang 4	Orang 5
1	92,50%	93,80%	92,50%	90,00%	95,00%
2	92,50%	95,00%	93,80%	96,30%	95,00%
3	92,50%	90,00%	93,80%	90,00%	97,50%
4	98,70%	98,70%	98,70%	97,50%	95,00%
5	91,30%	92,50%	90,00%	86,20%	95,00%
Rata-Rata	93,50%	94,00%	93,70%	92,00%	95,50%

Tabel VII dan Tabel VIII menyajikan nilai sensitivitas dan spesifikasi yang di dapatkan pada setiap skenario pengujian klasfikasi otentikasi pembuluh darah punggung tangan menggunakan ANN perambatan balik pada setiap kelas individu.Sensitivitas dan spesifikasi yang memiliki nilai mendekati 100% adalah sebuah indikator yang menyatakan bahwa sebuah sistem otentikasi dalam mengidentifikasi data pembuluh darah punggung tangan yang benar pada kelasnya dan sistem tidak mengidentifikasi data pembuluh darah punggung tangan yang bukan berada pada kelasnya.

Pada pengujian ke empat dan kelas Orang 1 memiliki nilai sensitivitas sebesar 95% dan spesifikasi sebesar 98,7%, pada kelas Orang 2 memiliki nilai sensitivitas sebesar 85% dan spesifikasi sebesar 98,7%, pada kelas Orang 3 memiliki nilai sensitivitas sebesar 90% dan spesifikasi sebesar 98,7%, dan pada kelas Orang 4 memiliki nilai sensitivitas sebesar 55% dan spesifikasi sebesar 97,5% dan pada kelas terakhir kelas Orang 5 memiliki nilai sensitivitas sebesar 90% dan spesifikasi sebesar 95%. Maka dapat disimpulkan bahwa kelas Orang 1 dan Orang 3 merupakan kelas yang memiliki nilai sensitivitas dan nilai spesifikasi paling baik diantra kelas lainnya karena memiliki nilai terbesar mendekati 100% yaitu sensitivitas bernilai 95% dan spesifikasi bernilai 98,7%.

Waktu pelatihan pada setiap skenario yang berjumlah sebanyak 5 skenario memiliki waktu yang bervariasi dan jika dihitung nilai rata-rata waktu pelatihan maka memiliki waktu pelatihan sebesar sekon atau detik dengan menggunakan algoritma perambatan balik pelatihan trainlm.

TABEL IX. DURASI WAKTU PELATIHAN

Skenario	Waktu Pelatihan (Sekon)
Skenario 1	357
Skenario 2	360
Skenario 3	350
Skenario 4	411
Skenario 5	348
Rata - Rata	365,2

Berdasarkan Tabel IX rata-rata durasi waktu pelatihan 365,2 detik, pada pengujian ke empat durasi waktu pelatihan adalah 411 detik, ke lima skenario pengujian memiliki durasi waktu pelatihan yang berbeda dikarenakan parameter pelatihan ANN yang memiliki 2 lapisan tersembunyi, jumlah neuron pada lapisan tersembunyi satu sebanyak 30 dan lapisan tersembunyi sebanyak 50 neuron, jumlah iterasi yang digunakan, nilai target galat yang kurang tepat. Pada penelitian ini pengujian pada skenario ke empat memiliki durasi waktu pelatihan terlama karena proses pelatihan yang berjalan dengan optimal dalam melakukan perhitungan data numerik sehingga menghasilkan nilai akurasi tertinggi jika dibandingkan terhadap skenario lainnya.

Faktor yang mempengaruhi nilai hasil akurasi, sensitivitas, spesifikasi dan waktu yang dihasilkan dari klasifikasi 5 kelas individu dari 5 skenario pengujian karena menggunakan parameter yang berbeda pada setiap skenario pada sistem otentikasi pembuluh darah punggung tangan.

IV. KESIMPULAN

Dengan hasil penelitian, analisa dan perancangan sistem ANN perambatan balik untuk otentikasi individu berbasis tekstur menggunakan tapis gabor berdasarkan objek pembuluh darah punggung tangan maka dapat diambil kesimpulan, Ekstraksi ciri menggunakan Tapis Gabor pertama dilakukan dengan 3 parameter yaitu rerata, varian, dan entropy yang kemudian dibagi menjadi 5 skenario pengujian. Pada setiap parameter ekstraksi ciri orde pertama dengan panjang gelombang dan orientasi sudut meghasilkan nilai pada tiap kelas menghasilkan nilai yang cukup bervariasi.

Sitem otentikasi pengenalan biometrik pembuluh darah punggung tangan menggunakan ANN perambatan balik ini menggunakan skenario ke empat dan mendapatkan hasil terbaik dengan menggunakan 3 masukan dari hasil ekstraksi

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

ciri tapis gabor, lapisan tersembunyi 1 berjumlah 30 node, lapisan tersembunyi 2 berjumlah 50 node dengan keluaran berjumlah 5. Menggunakan fungsi aktivasi *Tansig*, *Logsig*, dan *Purelin*, algoritma pembelajaran Trainlm, iterasi berjumlah 150 menghasilkan akurasi pelatihan sebesar 98,1% dan akurasi pengujian sebesar 91%.

Waktu yang di butuhkan oleh setiap skenario yang berjumlah 5 skenario tersebut memiliki rata-rata durasi 365,2 detik, waktu paling optimal didapatkan pada pengujian ke empat dengan waktu selama 411 detik. Waktu rata-rata pada setiap skenario pelatihan dapat dikatakan cukup cepat untuk sebuah pelatihan klasifikasi pengenalan menggunakan ANN perambatan balik.

Hasil presentase di atas menunjukan bahwa sistem ini dapat melakukan klasifikasi pengenalan pada pembuluh darah punggung tangan pada setiap kelas individu. Sistem ini dapat melakukan klasifikasi pengenalan pembuluh darah punggung tangan dalam waktu yang relatif singkat. Orientasi Sudut dan Panjang Gelombang pada ekstraksi ciri tapis Gabor cukup berpengaruh dalam keberhasilan sistem dan tingkat akurasi sistem. Orientasi Sudut dengan arah terbaik pada penelitian ini menggunakan arah 135° dengan Panjang Gelombang sebesar 3 dengan menggunakan 3 parameter utama yaitu rerata, varian, dan entropy.

REFERENCES

[1] I. Junaedi, "Pengembangan Teknologi Informasi Berbasis Access Id Card," *J. Inf. Syst. Informatics …*, vol. 1, no. 1, 2017.

[2] M. Arsal, bheta agus Wardijono, and D. Anggraini, "Face Recognition Untuk Akses Pegawai Bank Menggunakan Deep Learning Dengan Metode CNN," *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 1, pp. 55–63, 2020, doi: 10.1109/UBMK52708.2021.9559031.

[3] A. Andreansyah, R. F. Gusa, and M. Jumnahdi, "Pengenalan Pola Sidik Jari Menggunakan Multi-Class Support Vector Machine," *J. ELKHA*, vol. 11, no. 2, pp. 79–84, 2019.

[4] F. E. Alfian, I. G. P. S. Wijaya, and F. Bimantoro, "Identifikasi Iris Mata Menggunakan Metode Wavelet Daubechies dan K-Nearest Neighbor," *J. Teknol. Informasi, Komputer, dan Apl. (JTIKA)*, vol. 2, no. 1, pp. 1–10, 2020, doi: 10.29303/jtika.v2i1.76.

[5] R. P. Saidah, Sofia, B. Novria, Aulia, G. Shekinah, and F. Wahid, "Analisis Perbandingan Metode LBP dan CLBP pada Sistem Pengenalan Individu Melalui Iris Mata," *JEPIN (Jurnal Edukasi dan Penelit. Inform.*, vol. 6, no. 3, pp. 285–290, 2020.

[6] D. S. Wita and D. Y. Liliana, "Klasifikasi Identitas Dengan Citra Telapak Tangan Menggunakan Convolutional Neural Network (CNN)," *J. Rekayasa Teknol. Inf.*, vol. 6, no. 1, pp. 1–7, 2022.

[7] C. Susim, Theresia dan Darujati, "Pengolahan Citra Untuk Pengenalan Wajah (Face Recognition) Menggunakan OpenCV," *J. Syntax Admiration*, vol. 2, no. 3, pp. 534–545, 2021.

[8] M. N. Ikhsan and R. Rahmadewi, "Sistem Keamanan Sepeda Motor Dengan Teknologi Biometrik Sidik Jari Menggunakan Sensor Fingerprint," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 7, no. 2, pp. 144–153, 2022.

[9] M. S. Purba, "Perancangan Sistem Identifikasi Biometrik Iris Mata Menggunakan Metode Transformasi Hough," vol. 7, no. 2, pp. 117–122, 2020.

[10] H. Setiawan, "Telapak Tangan Menggunakan Learning Vector Quantization Palm Vein Image Identification Using Learning," 2016.

[11] K. Kim, H. W. Jeong, and Y. Lee, "Performance evaluation of dorsal vein network of hand imaging using relative total variation-based regularization for smoothing technique in a miniaturized vein imaging system: A pilot study," *Int. J. Environ. Res. Public Health*, vol. 18, no. 4, pp. 1–12, 2021, doi: 10.3390/ijerph18041548.

[12] M. Simanjuntak, K. Abdi Sinuraya, Irhamna, and J. Panjaitan, "Analisis simulasi dan evaluasi teknik pengenalan tanda tangan menggunakak rbf dan knn," *JUTISAL J. Tek. Inform. Univers.*, vol. 2, no. 1, pp. 54–60, 2022.

[13] M. S. Simanjuntak and J. Panjaitan, "Sistem Information Retrieval Menggunakan K- Nearest Neighbour Dalam Klasifikasi Jurnal Bahasa Inggris," *JUTISAL J. Tek. Inform. Univers.*, vol. 1, no. 2, pp. 1–8, 2021.

[14] K. M. Alashik and R. Yildirim, "Human Identity Verification from Biometric Dorsal Hand Vein Images Using the DL-GAN Method," *IEEE Access*, vol. 9, pp. 74194–74208, 2021, doi: 10.1109/ACCESS.2021.3076756.

[15] S. Bantun, J. Y. Sari, N. Z, M. Mardianto, and A. Achban, "Sistem Absensi Mahasiswa Berbasis Dorsal Hand Vein Menggunakan Local Binary Patterns dan Fuzzy k-NN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 384–396, 2022, doi: 10.35957/jatisi.v9i1.1496.

[16] N. Fajriani, "Pengenalan Pola Garis Telapak Tangan Menggunakan Metode Fuzzy K-Nearest Neighbor," *Edutic - Sci. J. Informatics Educ.*, vol. 4, no. 1, 2017, doi: 10.21107/edutic.v4i1.3385.

[17] P. Nurtanto Andono, T. Sutojo, and Muljono, *Pengolahan Citra Digital*, 1st ed. Yogyakarta: Penerbit Andi, 2018.

[18] A. Kadir and A. Susanto, *Teori dan Aplikasi Pengolahan Citra*, 1st ed. Yogyakarta: Penerbit Andi, 2013.

[19] I. Setiawan, W. Dewanta, H. A. Nugroho, and H. Supriyono, "Pengolah Citra Dengan Metode Thresholding Dengan Matlab R2014A," *J. Media Infotama*, vol. 15, no. 2, 2019, doi: 10.37676/jmi.v15i2.868.

[20] A. Susanto, "Penerapan Operasi Morfologi Matematika Citra Digital Untuk Ekstraksi Area Plat Nomor Kendaraan Bermotor," *Pseudocode*, vol. 6, no. 1, pp. 49–57, 2019, doi: 10.33369/pseudocode.6.1.49-57.

[21] D. N. Rahmah, H. Tjandrasa, and A. Yuniarti, "Implementasi Segmentasi Pembuluh Darah Retina Pada Citra Morfologi Adaptif," no. September 2016, pp. 1–6, 2011.

[22] M. A. S. Yudono, R. R. Isnanto, and A. Triwiyatno, "Comparison of Cataract Classification System Based on Retinal Blood Vessels Objects and Retinal Optic Disc Using Backpropagation Neural Network," *Int. J. Innov. Eng. Technol.*, vol. 18, no. 2, pp. 1–8, 2021, doi: 10.13140/RG.2.2.16638.46408.

[23] admi syarif, A. R. TANJUNG, R. ANDRIAN, and F. R. LUMBANRAJA, "Implementasi Metode Ekstraksi Fitur Gabor Filter dan Probablity Neural Network (PNN) untuk Identifikasi Kain Tapis Lampung," *J. Komputasi*, vol. 8, no. 2, 2020, doi: 10.23960/komputasi.v8i2.2641.

[24] T. Wulandari and N. H. Waryanto, "Identifikasi Iris Mata dengan menggunakan Metode Hidden Markov Model dan Tapis Gabor Wavelet," *J. Kaji. dan Terap. Mat.*, vol. 7, no. 4, pp. 10–11, 2018.

[25] R. N. Hidayat, R. R. Isnanto, and O. D. Nurhayati, "Implementasi Jaringan Syaraf Tiruan Perambatan Balik untuk Memprediksi Harga Logam Mulia Emas Menggunakan Algoritma Lavenberg Marquardt," *J. Teknol. dan Sist. Komput.*, vol. 1, no. 2, p. 49, 2013, doi: 10.14710/jtsiskom.1.2.2013.49-55.

[26] S. Kusumadewi, *Membangung Jaringan Syaraf Tiruan Menggunakan Matlab & Excel link*. Yogyakarta: Graha Ilmu, 2004.

[27] G. Ramadhona, B. D. Setiawan, and F. A. Bachtiar, "Prediksi Produktivitas Padi Menggunakan Jaringan Syaraf Tiruan Backpropagation," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 12, pp. 6048–6057, 2018.

[28] G. Z. Muflih, S. Sunardi, and A. Yudhana, "Jaringan Saraf Tiruan Backpropagation untuk Prediksi Curah Hujan di Wilayah Kabupaten Wonosobo," *MUST J. Math. Educ. Sci. Technol.*, vol. 4, no. 1, p. 45, 2019, doi: 10.30651/must.v4i1.2670.

[29] R. Riyanda, A. H. H. Pardede, and R. Saragih, "Jaringan Syaraf Tiruan Memprediksi Kebutuhan Obat-Obatan Menggunakan Metode Backpropagation (Studi Kasus : UPTD Puskesmas Bahorok)," *Semin. Nas. Inform.*, pp. 47–55, 2021.

[30] M. E. Al Rivan and T. Juangkara, "Identifikasi Potensi Glaukoma dan Diabetes Retinopati Melalui Citra Fundus Menggunakan Jaringan Syaraf Tiruan," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 1, pp. 43–48, 2019, doi: 10.35957/jatisi.v6i1.158.

[31] F. Pontoh, H. V. F. Kainde, and Y. V. Akay, "Teknik pengenalan pembuluh darah punggung tangan berbasis fitur local binary pattern," *J. Widya*, vol. 2, no. 2, pp. 198–203, 2021, doi: 10.54593/awl.v2i2.15.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1819, Copyright ©2023
Submitted : June 21, 2023, Revised : September 11, 2023, Accepted : October 8, 2023, Published : November 7, 2023

Penentuan Rekomendasi Paket Promosi Menggunakan Algoritma *Frequent Pattern Growth* Di *The Javanese Cafe*

Dwi Astuti[1], Samsinar[2]*
Program Studi Sistem Informasi, Fakultas Teknologi Informasi [1], [2]
Universitas Budi Luhur
Jakarta, Indonesia
dwiastuti0603@gmail.com[1] , samsinar@budiluhur.ac.id[2]*

Abstract*— Data analysis and processing is very important to support business development. One example is The Javanese Café which requires analysis and processing to determine promotional menu package recommendations. To carry out data analysis and processing, of course you need technology to make these activities easier. The technology that can be used to overcome this problem is data mining. Data mining has an association rule method which functions to form association patterns. Researchers also use the FP-Growth algorithm to speed up the data processing process. The sales transaction data processing resulted in 14 association patterns with the highest confidence values and 9 menu items with the lowest support values. Then the results were analyzed again and produced 4 recommendations for promotional menu packages that could be used to support product marketing strategies.

***Keywords*— Data Mining, Association Rule, FP-Growth, Purchase Pattern**

***Abstrak*— Analisis dan pengolahan data sangatlah penting untuk mendukung perkembangan bisnis. Salah satu contohnya pada *The Javanese Café* yang membutuhkan analisa dan pengolahan untuk menentukan rekomendasi paket menu promosi. Untuk melakukan analisis dan pengolahan data tentunya membutuhkan sebuah teknologi agar dapat mempermudah kegiatan tersebut. Teknologi yang dapat dimanfaatkan untuk mengatasi permasalahan tersebut adalah data mining. Data mining mempunyai metode *association rule* yang berfungsi untuk membentuk pola asosiasi. Peneliti juga memanfaatkan algoritma *FP-Growth* untuk mempercepat proses pengolahan data. Dari proses pengolahan data transaksi penjualan, menghasilkan 14 pola asosiasi dengan nilai *confidence* tertinggi dan 9 item menu dengan nilai *support* terendah. Kemudian hasil tersebut dianalisa kembali dan menghasilkan 4 rekomendasi paket menu promosi yang dapat digunakan untuk menunjang strategi pemasaran produk.**

***Kata Kunci*— Data Mining, Association Rule, FP-Growth, Pola Pembelian**

I. PENDAHULUAN

Era globalisasi saat ini telah melahirkan banyak bisnis baru dan tingkat persaingan yang tinggi, salah satunya adalah usaha di bidang kuliner. Usaha ini sangat menjanjikan karena kebutuhan pangan konsumen semakin hari semakin meningkat [1]. Untuk menghadapi persaingan bisnis yang tinggi, pelaku bisnis harus merancang berbagai taktik pemasaran agar kelangsungan bisnis terjaga. Salah satu cara agar bisnis tetap kompetitif dengan para pesaing adalah melalui promosi. Sebagai salah satu bentuk pemasaran [2], promosi bertujuan untuk mewujudkan peningkatan kesadaran terhadap produk, mengembangkan preferensi merek di pasar sasaran, mengoptimalkan pemasaran dan pangsa pasar, mendorong pembelian berulang produk yang sama, memperkenalkan produk baru serta memikat pembeli baru [3].

The Javanese Café merupakan suatu usaha yang lahir akibat perkembangan bisnis di bidang kuliner. Tingginya tingkat persaingan menyebabkan penurunan penjualan pada *The Javanese Café*, sehingga pihak *supervisor cafe* ingin melakukan promosi secara berkala untuk meningkatkan penjualan. Namun, untuk melakukan tindakan tersebut, manajemen tidak mempunyai acuan untuk memilih item menu yang akan ditawarkan kepada pelanggan. Oleh karena itu, untuk mengatasi permasalahan tersebut perlu dilakukan pengolahan dan analisis terhadap data transaksi penjualan. Dalam melakukan pengolahan dan analisis data membutuhkan sebuah teknologi untuk mempermudah serta mempercepat kegiatan tersebut. Teknologi yang dapat dimanfaatkan sebagai solusi dari permasalahan tersebut adalah *data mining*.

Data mining adalah serangkaian proses yang menghasilkan wawasan baru dengan menggali sejumlah data besar untuk menemukan pola dan aturan tertentu [4]. Teknik dalam *data mining* yang dapat dimanfaatkan untuk mengatasi masalah pada penelitian ini adalah *association rule.* Teknik asosiasi berfungsi untuk mendeteksi pola yang paling kerap muncul dalam himpunan *itemset* [5]. Proses analisis menggunakan bantuan algoritma *FP-Growth* untuk membuat aturan asosiasi dalam bentuk *"If Then".* Algoritma *FP-Growth* merupakan salah satu algoritma yang bisa diimplementasikan pada teknik *association rules*. Fungsi algoritma *FP-Growth* yaitu untuk mengidentifikasi item yang paling kerap muncul pada himpunan data dalam sebuah *database* [6].

Terdapat beberapa penelitian terdahulu terkait penetapan promosi produk dengan memanfaatkan algoritma

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

FP-Growth, yaitu penelitian yang dilakukan oleh [7] menghasilkan 2 *rule* yang mempunyai nilai *confidence* 79,1 % dan 58,8%. Sedangkan penelitian yang dilakukan [8] mempunyai hasil 3 rule yang semua nilai *confidence* 100%. Algoritma *FP-Growth* diterapkan pada penelitian ini karena untuk memperoleh *frequent itemset* tidak membutuhkan *generate candidate*, melainkan memanfaatkan konsep penyusunan *tree*, sehingga proses pengolahan data lebih cepat.

Tujuan penelitian ini adalah mengidentifikasi pola pembelian pada data transaksi penjualan *The Javanese Cafe* menggunakan algoritma *Frequent Pattern Growth* (*FP-Growth*) sebagai acuan untuk menentukan rekomendasi menu promosi. Sehingga pola yang dihasilkan diharapkan dapat membantu perusahaan untuk mendapatkan rekomendasi menu promosi.

II. METODE PENELITIAN

A. *Metode Pengumpulan Data*

Data yang digunakan pada penelitian ini berasal dari data sekunder, yang bersumber dari data transaksi penjualan *The Javanese Café* periode Bulan Maret 2022 sampai Bulan Maret 2023 dengan jumlah 6.674 *record*. Data sekunder merupakan data yang sengaja dikumpulkan untuk memberikan informasi berupa dokumentasi seperti catatan atau gambar [9]. Pada data transaksi penjualan *The Javanese Café* terdapat 9 atribut, yang terdiri dari *id* transaksi, tanggal, waktu, nama produk, jumlah, kategori produk, nominal, pembayaran dan nama *merchant* yang dapat dilihat pada Tabel 1 berikut:

TABEL I. SAMPEL DATA TRANSAKSI PENJUALAN

ID Transaksi	Tanggal	Waktu	Nama Produk	Jumlah	Kategori Produk	Nominal	Pembayaran	Nama Merchant
46766829	01/03/2022	14:22:07	*Blood Vampire*	1	*Fresh*	Rp15.000	Cash	Javanese Cafe
46766829	01/03/2022	14:22:07	*Javanese Sampler*	1	Camilan	Rp11.000	Cash	Javanese Cafe
46769033	01/03/2022	14:47:50	*Blue Ocean*	1	*Fresh*	Rp12.000	Cash	Javanese Cafe

B. Data *Preprocessing*

Sebelum dataset diolah menggunakan *tools*, terdapat beberapa kegiatan yang harus dilakukan, yaitu :

a. Membersihkan Data (*Data Cleaning*)

Pada tahap ini dilakukan pembersihan terhadap data pada Tabel 1. dengan menghapus beberapa atribut yaitu id transaksi, waktu, jumlah, kategori produk, nominal, pembayaran, dan nama merchant, sehingga menghasilkan data sepeti Tabel 2.

TABEL II. DATA HASIL PEMBERSIHAN ATRIBUT

Tanggal	Nama Produk
01/03/2022	*Blood Vampire, Javanese Sampler*
01/03/2022	*Blue Ocean*

b. Transformasi Data (*Data Transformation*)

Sebelum melakukan proses transormasi data, dilakukan pengkodean terhadap nama produk, karena nama produk yang terlalu panjang. Daftar kode produk bisa dilihat pada Tabel 3 berikut:

TABEL III. KODE PRODUK

Kode	Nama Produk	Kode	Nama Produk
M1	Air Mineral	M26	Taro
M2	Teh	M27	*Coklat Hazelnut*
M3	*Lemon Tea*	M28	*Thai Tea*
M4	*Lemon Tea* Jumbo	M29	*Blue Ocean*
M5	Es Teh Jumbo	M30	*Selasih Fantasy*
M6	Wedang Uwuh	M31	*Red Delight*
M7	Kopi Tubruk Susu	M32	*Blood Vampire*
M8	Kopi Tubruk	M33	*Cinamond Crush*
M9	*Vietnam Drip*	M34	*French Fries*
M10	*Caffe Latte*	M35	Jamur Krispi
M11	*Cappucchino*	M36	Tahu Krispi
M12	*Black Americano*	M37	*Javanese Sampler*
M13	*Espresso*	M38	Siomay
M14	*Mocha Latte*	M39	Roti Bakar
M15	V-60	M40	*Croffle*
M16	*CoffeeMilk Original*	M41	Cireng Krispi
M17	*CoffeeMilk Brown Sugar*	M42	Donat Kentang
M18	*CoffeeMilk Hazelnut*	M43	Nasi Ayam Matah
M19	*CoffeeMilk Matcha*	M44	Nasi Ayam Geprek
M20	*Americano Ice*	M45	Nasgor Krengsengan
M21	*Japanese Ice*	M46	Nasgor Terasi
M22	*CoffeeMilk Vanila*	M47	Nasgor Rawon
M23	Coklat	M48	Indomie Rebus
M24	*Red Velvet*	M49	Indomie Goreng
M25	*Matcha*		

Transformasi data dilakukan untuk mengubah tipe data *categorical* menjadi *binominal.* Proses transformasi menghasilkan data pada Tabel 4.

TABEL IV. HASIL TRANSFORMASI DATA

Tanggal	TID	M1	M2	M3	M4	M5	M6	M7	M8	…	M49	Jumlah Item
01/03/2022	1	0	0	0	0	0	0	0	0	…	0	2
01/03/2022	2	0	0	0	0	0	0	0	0	…	0	1
…	…	…	…	…	…	…	…	…	…	…	…	…
31/03/2023	6674	0	0	0	0	0	0	0	1		0	2

c. Reduksi Data (*Data Reduction*)

Pada tahap reduksi data, aktivitas yang dilakukan adalah menyeleksi data dengan jumlah item per transaksi lebih dari satu. Karena data yang akan diolah adalah data yang memiliki item dengan jumlah minimal dua item dalam satu transaksi, seperti data pada Tabel 5.

TABEL V. DATA HASIL REDUKSI

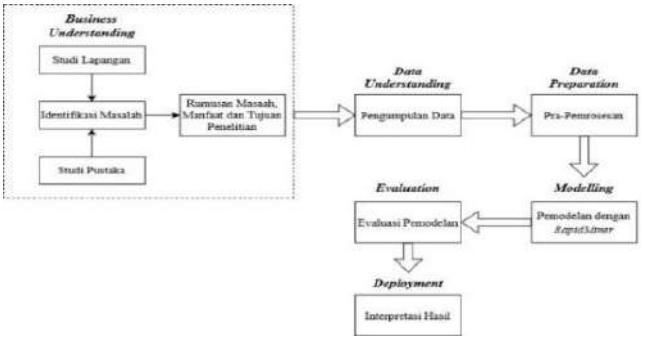
Tanggal	TID	M1	M2	M3	M4	M5	M6	M7	M8	…	M49	Jumlah item
01/03	1	0	0	0	0	0	0	0	0	…	0	2

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

/2022											
01/03 /2022	2	0	0	0	0	0	0	0	0	…	1	3
…	…	…	…	…	…	…	…	…	…	…	…	…
31/03 /2023	47 40	0	0	0	0	0	0	0	1	…	0	2

C. Tahapan Penelitian

Pelaksanaan penelitian mengimplementasikan metode pengembangan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai tahapan penelitian. CRISP-DM merupakan konsorsium perusahaan yang dibangun pada tahun 1996 oleh Komisi Eropa yang ditetapkan sebagai prosedur standar penambangan data sehingga bisa digunakan di beragam bidang industri [10]. Gambar 1. berikut merupakan gambaran alur proses CRISP-DM beserta penjelasannya.



Gambar 1. Tahap CRISP-DM

D. Association Rule

Aturan asosiasi adalah teknik penambangan data yang digunakan untuk mengidentifikasi item satu dengan item lainnya dalam database [11]. Untuk menetapkan hubungan antar item, harus mencakup persyaratan minimum support dan minimum *confidence*, yang mempunyai rentang nilai dari 0% hingga 100% [12], berikut merupakan penjelasan mengenai nilai *support* dan *confidence* [13].

a. Support

Support adalah representasi dari gabungan item pada database. Misalnya, jika ditemukan item A dan item B dalam database, maka nilai supportnya adalah persentase transaksi pada database yang berisi item A dan B. Berikut rumus yang digunakan untuk menentukan nilai *support*:

$$Support\ \mathrm{A} = \frac{\sum Transaksi\ mengandung\ A}{\sum Transaksi} \quad (1)$$

Sedangkan nilai *support* untuk 2 item diperoleh dengan rumus:

$$Support\ (\mathrm{A,B}) = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi} \quad (2)$$

b. Confidence

Confidence pada aturan asosiasi adalah ukuran untuk menetapkan aturan, yaitu penyajian transaksi dalam basis data yang berisi A dan B. Nilai *confidence* bisa diketahui dengan rumus berikut ini:

$$Confidence = P(B|A) = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi\ \ mengandung\ A} \quad (3)$$

E. Algorithm FP-Growth

Algoritma *FP- Growth* adalah metode alternatif untuk mendeteksi kumpulan data yang paling kerap muncul (*frequent itemset*) tanpa memerlukan generasi kandidat [14]. Sebelum melakukan proses pada fase utama, terdapat beberapa langkah khusus untuk menyiapkan *dataset*, yaitu mencari *frequent itemset*, membuat susunan berlandaskan prioritas, *dataset* diatur berlandaskan prioritas, dan membuat *fp-tree*. Setelah langkah khusus terpenuhi, selanjutnya menyelesaikan tiga fase utama dari algoritma *FP-Growth*, yaitu *Fase* Pembangkitan *Conditional Pattern Base*, *Fase* Pembangkitan *Conditional FP-Tree*, dan *Fase* Pembangkitan *Frequent Itemset* [15].

F. Lift Ratio

Nilai *lift ratio* berguna untuk menyatakan keabsahan mekanisme transaksi dalam memberikan informasi tentang benar tidaknya produk A dan produk B dibeli dalam waktu yang bersamaan. Jika nilai *lift ratio* > 1 maka transaksi dikatakan sah, artinya produk A sebenarnya dibeli bersamaan dengan produk B [16]. Untuk menghitung nilai *lift ratio* bisa dilakukan dengan rumus berikut:

$$Lift\ Ratio = \frac{Confidence}{Expected\ Confidence} \quad (4)$$

Sedangkan *expected confidence* dihitung dengan rumus berikut ini:

$$Expected\ Confidence = \frac{\sum Banyak\ Transaksi\ Mengandung\ B}{\sum Transaksi} \quad (5)$$

Jika perhitungan menghasilkan nilai kurang dari 1 maka terdapat korelasi negatif. Sedangkan hasil perhitungan yang nilainya lebih besar dari 1, terdapat korelasi positif. Namun ketika hasil perhitungan nilainya sama dengan 1 maka tidak ada korelasi antara X dan Y [17].

III. HASIL DAN PEMBAHASAN

A. Perhitungan FP-Growth

Perhitungan algoritma FP-Growth akan dilakukan menggunakan 10 sampel item menu dan 20 sampel data transaksi penjualan. Sampel item menu yang digunakan yaitu Vietnam Drip, Caffe Latte, Cappucchino, Black Americano, Mocha Latte, Coklat Hazelnut, French Fries, Tahu Krispi, Roti Bakar dan Croffle. Sedangkan sampel data transaksi dapat dilihat pada Tabel 6:

TABEL VI. SAMPEL DATA TRANSAKSI

TID	Item Transaksi
6	M14, M34

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

TID	Item Transaksi
10	M10, M36
11	M11, M12, M34
16	M11, M12, M36
25	M9, M36
32	M27, M34, M36
41	M10, M12, M39
44	M11, M12, M40
71	M9, M12, M36
73	M9, M11, M36
139	M10, M34, M39
148	M10, M34
189	M9, M12, M36, M39
215	M10, M11, M40
228	M9, M34
243	M10, M36, M39
297	M9, M11, M34, M36, M40
306	M12, M34
307	M10, M14, M39
316	M12, M34

Setelah menentukan sampel data, langkah selanjutnya melakukan empat fase khusus algoritma *FP-Growth* yaitu:

a. Mencari frekuensi kemunculan menggunakan persamaan (1). Nilai minimum *support* yang ditetapkan adalah 0.1, item menu akan dihapus jika tidak memenuhi nilai minimum *support*, seperti yang terlihat pada Table 7.

TABEL VII. FREKUENSI KEMUNCULAN ITEM MENU

Item Menu	Frekuensi	Proses *Support*	*Support*
M9	6	(6/20)	0,3
M10	7	(7/20)	0,35
M11	6	(6/20)	0,3
M12	8	(8/20)	0,4
M14	2	(2/20)	0,1
M27	1	(1/20)	0,05
M34	9	(9/20)	0,45
M36	9	(9/20)	0,45
M39	5	(5/20)	0,25
M40	3	(3/20)	0,15

b. Membuat susunan item menu berdasarkan prioritas. Tabel 8 merupakan hasil penyusunan item menu berdasarkan frekuensi kemunculan.

TABEL VIII. FREKUENSI KEMUNCULAN ITEM MENU BERDASRKAN PRIORITAS

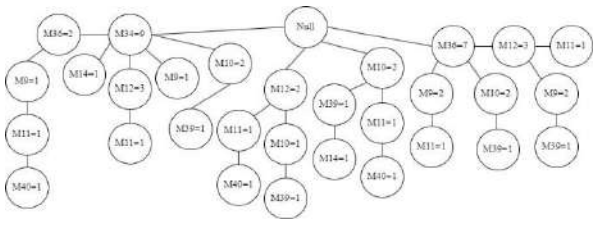
Item Menu	Frekuensi	*Support*
M34	9	0,45
M36	9	0,45
M12	8	0,4
M10	7	0,35
M9	6	0,3
M11	6	0,3
M39	5	0,25
M40	3	0,15
M14	2	0,1

c. *Dataset* diatur berdasarkan prioritas. Setiap transaksi disusun berdasarkan item menu yang mempunyai frekuensi tertinggi menuju item menu yang mempunyai frekuensi rendah.

TABEL IX. DATASET BERDASARKAN PRIORITAS

TID	Item Transaksi
6	M34, M14
10	M36, M10
11	M34, M12, M11
16	M36, M12, M11
25	M36, M9
32	M34, M36
41	M12, M10, M39
44	M12, M11, M40
71	M36, M12, M9
73	M36, M9, M11
139	M34, M10, M39
148	M34, M10
189	M36, M12, M9, M39
215	M10, M11, M40
228	M34, M9
243	M36, M10, M39
297	M34, M36, M9, M11, M40
306	M34, M12
307	M10, M39, M14
316	M34, M12

d. Membuat *fp-tree* berdasarkan pembacaan dari setiap TID. Hasil pembacaan semua TID dapat dilihat pada Gambar 2.



Gambar 2. Pembentukan FP-Tree Setelah Pembacaan Seluruh TID

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

Setelah pembentukan *Fp-Tree* selesai langkah berikutnya melakukan tiga fase utama algoritma *frequent pattern growth* yaitu:

a. *Fase* Pembangkitan *Conditional Pattern Base*

Dibentuk berdasarkan *FP-Tree* pada Gambar 2. Tabel 10 berikut merupakan hasil dari proses *conditional pattern base*:

TABEL X. CONDITIONAL PATTERN BASE

Suffix	*Conditional Pattern Base*
M14	{M34=1}, {M10,M39=1}
M40	{M34,M36,M9,M11=1}, {M12,M11=1}, {M10,M11=1}
M39	{M34,M10=1}, {M12,M10=1}, {M10=1}, {M36,M10=1}, {M36,M12,M9=1}
M11	{M34,M36,M9=1}, {M34,M12=1}, {M12=1}, {M10=1}, {M36,M9=1}, {M36,M12=1}
M9	{M34,M36=1}, {M34=1}, {M36=2}, {M36,M12=2}
M10	{M34=2}, {M12=1}, {M36=2}
M12	{M34=3}, {M36=3}
M36	{M34=2}

b. *Fase* Pembangkitan *Conditional FP-Tree*

Pada *fase* ini, support count dari setiap item dalam conditional pattern base akan diakumulasi. Jika hasil akumulasi support count lebih besar atau sama dengan nilai minimum support yang telah ditentukan, maka item tersebut akan di-generate kembali. Tabel 11 berikut adalah hasil conditional FP-Tree yang terbentuk:

TABEL XI. CONDITIONAL FP-TREE

Suffix	*Conditional FP-Tree*
M40	{M11=3}
M39	{M36=2}, {M10=4}
M11	{M34=2}, {M36=3}, {M36, M9=2}, {M12=3}
M9	{M34=2}, {M36=5, M12=2}
M10	{M34=2}, {M36=2}
M12	{M34=3}, {M36=3}
M36	{M34=2}

c. *Fase* Pembangkitan *Frequent Itemset*

Kombinasi *item conditional FP-Tree* menjadi dasar pembentukan frequent itemset. Sehingga menghasilkan pembentukan frequent itemset pada Tabel 12 berikut:

TABEL XII. FREQUENT ITEMSET

Suffix	*Frequent Itemset*
M40	{M11, M40=3}
M39	{M36, M39=2}, {M10, M39=4}
M11	{M34, M11=2}, {M36, M11=3}, {M36, M9, M11=2}, {M12, M11=3}
M9	{M34, M9=2}, {M36, M9=5}, {M12, M9=2}, {M36, M12, M9=2}
M10	{M34, M10=2}, {M36, M10=2}
M12	{M34, M12=3}, {M36, M12=3}
M36	{M34, M36=2}

Nilai minimum *support* yang ditetapkan adalah 0,1 serta nilai minimum *confidence* 0,5. Perhitungan nilai *support* memanfaatkan persamaan (2), sedangkan nilai *confidence* dihitung menggunakan persamaan (3). Tabel 13 berikut menunjukkan pola asosiasi yang mencakup nilai minimum *support* dan minimum *confidence*:

TABEL XIII. POLA YANG MENCAKUP NILAI MINIMUN SUPPORT DAN NILAI MINIMUM CONFIDENCE

No	*Rule*	*Support*	*Confidence*
1	*If* {M11} *Then* {M40}	3/20 = 0,15	3/5 = 0,5
2	*If* {M40} *Then* {M11}	3/20 = 0,15	3/3 = 1
3	*If* {M10} *Then* {M39}	4/20 = 0,2	4/7 = 0,57
4	*If* {M39} *Then* {M10}	4/20 = 0,2	4/5 = 0,8
5	*If* {M11} *Then* {M39}	3/20 = 0,15	3/6 = 0,5
6	*If* {M36, M11} *Then* {M9}	2/20 = 0,1	2/3 = 0,67
7	*If* {M11, M9} *Then* {M36}	2/20 = 0,1	2/2 = 1
8	*If* {M11} *Then* {M12}	3/20 = 0,15	3/6 = 0,5
9	*If* {M36} *Then* {M9}	5/20 = 0,25	5/9 = 0,56
10	*If* {M9} *Then* {M36}	5/20 = 0,25	5/6 = 0,83
11	*If* {M36, M12} *Then* {M9}	2/20 = 0,1	2/3 = 0,67
12	*If* {M12, M9} *Then* {M36}	2/20 = 0,1	2/2 = 1

B. Pengujian Lift Ratio

Pengujian dengan lift ratio dilakukan guna mengetahui kekuatan pola asosiasi, yang dihitung menggunakan persamaan (4), sedangkan nilai expected confidence akan dihitung menggunakan persamaan (5):

TABEL XIV. HASIL PERHITUNGAN NILAI LIFT RATIO

No	*Rule*	*Support*	*Confidence*	*Expected Confidence*	*Lift*
1	*If* {M11} *Then* {M40}	0,15	0,5	0,15	3,333
2	*If* {M40} *Then* {M11}	0,15	1	0,3	3,333
3	*If* {M10} *Then* {M39}	0,2	0,57	0,25	2,280
4	*If* {M39} *Then* {M10}	0,2	0,8	0,35	2,285
5	*If* {M11} *Then* {M39}	0,15	0,5	0,45	1,111
6	*If* {M36, M11} *Then* {M9}	0,1	0,67	0,3	2,233
7	*If* {M11, M9} *Then* {M36}	0,1	1	0,45	2,222
8	*If* {M11} *Then* {M12}	0,15	0,5	0,4	1,250
9	*If* {M36} *Then* {M9}	0,25	0,56	0,3	1,866
10	If {M9} *Then* {M36}	0,25	0,83	0,45	1,844
11	*If* {M36, M12} *Then* {M9}	0,1	0,67	0,3	2,233
12	*If* {M12, M9} *Then* {M36}	0,1	1	0,45	2,222

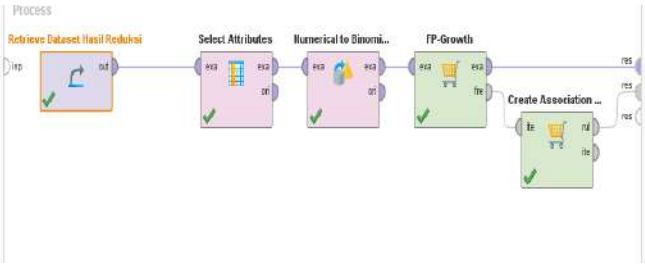
Semua pola menghasilkan nilai lift ratio > 1, sehingga semua pola tersebut memiliki korelasi positif. Pada 12 rule yang terbentuk terdapat beberapa rule yang memiliki item menu yang sama. Oleh karena itu rule tersebut harus dipilih salah satu, dengan mempertimbangkan nilai confidence yang lebih besar untuk dipilih.

C. Implementasi RapidMiner

Software RapidMiner digunakan untuk melakukan pengujian terhadap dataset, sehingga korelasi antar rule yang

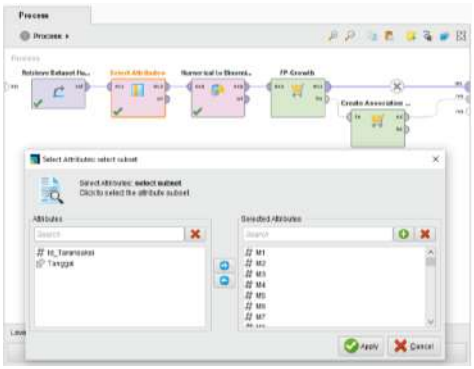
p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

terbentuk dapat diketahui. Desain pengujian ditunjukkan pada Gambar 3 berikut:



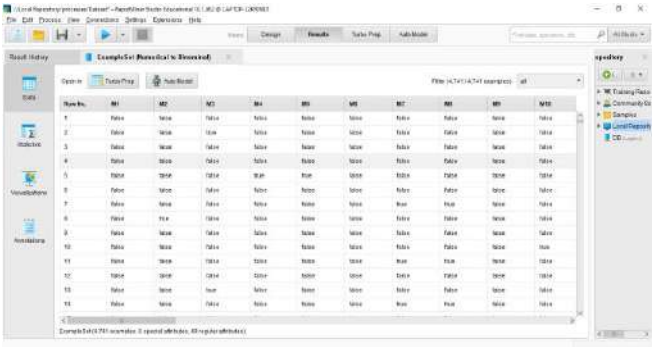
Gambar 3. Desain Pengujian pada RapidMiner

Terdapat 5 operator yang digunakan pada desain pengujian, dimana dalam setiap operator mempunyai fungsi masing-masing. Operator pertama berfungsi untuk mengimport dataset yang sudah terbentuk pada proses persiapan data. Setelah dataset diimport, atribut data akan diseleksi menggunakan operator *select attribute* seperti yang terlihat pada Gambar 4 berikut.



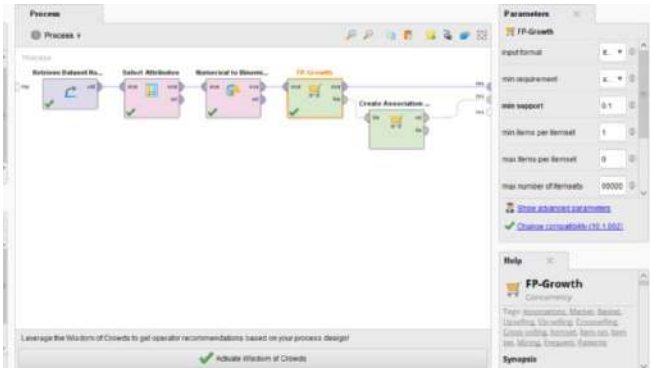
Gambar 4. Operator Select Attribute

Proses selanjutnya yaitu mengubah tipe data yang semula mempunyai tipe numerik akan diubah menjadi tipe binominal menggunakan operator *Numerical to Binominal*, sehingga akan menghasilkan data seperti Gambar 5 berikut.



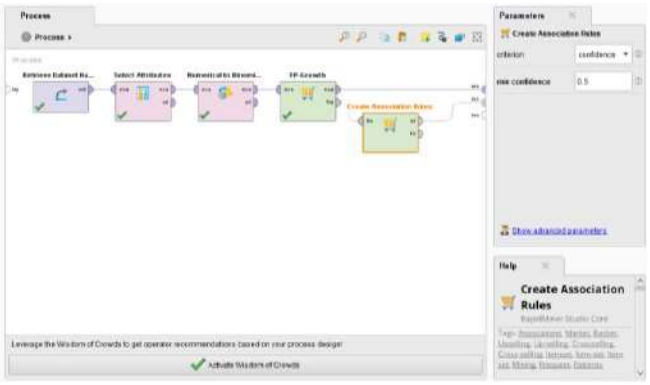
Gambar 5. Hasil Data berdasarkan Operator Numerical to Binominal

Tahap berikutnya adalah penerapan algoritma FP-Growth. Pada operator FP-Growth, kegitan yang dilakukan adalah menentukan nilai minimum *support* untuk menyeleksi dataset. Nilai minimum *support* yang digunakan pada penelitian ini adalah 0.1 seperti yang terlihat pada Gambar 6 berikut :



Gambar 6. Operator FP-Growth

Selain menetukan nilai minimum *support* diperlukan pula penentuan nilai minimum *confidence* untuk mengukur kekuatan kombinasi pola yang terbentuk. Nilai minimum *confidence* yang digunakan pada penelitian ini yaitu 0.5 dan dapat di*setting* pada operator *Create Association Rule*.



Gambar 7. Creat Association Rule

Dari desain pengujian tersebut menghasilkan 14 pola asosiasi dan 9 menu dengan nilai support terendah. Hasil tersebut kemudian dianalisa sehingga menghasilkan rekomendasi paket menu yang dapat dilihat pada Tabel 16.

AssociationRules

```
Association Rules
[M40] --> [M11] (confidence: 0.513)
[M23] --> [M34] (confidence: 0.535)
[M19] --> [M36] (confidence: 0.556)
[M39] --> [M10] (confidence: 0.567)
[M44] --> [M3] (confidence: 0.580)
[M22] --> [M34] (confidence: 0.584)
[M30] --> [M34] (confidence: 0.592)
[M11] --> [M40] (confidence: 0.598)
[M20] --> [M34] (confidence: 0.635)
[M27] --> [M34] (confidence: 0.673)
[M14] --> [M34] (confidence: 0.684)
[M10] --> [M39] (confidence: 0.773)
[M26] --> [M34] (confidence: 0.781)
[M31] --> [M34] (confidence: 0.837)
```

Gambar 6. Hasil Pola Asosiasi

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

TABEL XV. ITEM MENU DENGAN SUPPORT TERENDAH

No	Kode Produk	*Support*
1	M4	15%
2	M13	18%
3	M33	18%
4	M1	18%
5	M21	22%
6	M47	23%
7	M49	23%
8	M46	25%
9	M6	26%

TABEL XVII. HASIL PAKET PROMOSI

No	Paket Promosi yang terbentuk
1	Produk yang sering dibeli dengan tingkat kepercayaan 87,3% di*bundling* promo dengan produk *Cinamond Crush* yang mempunyai presentase *support* 18%
2	Produk yang sering dibeli dengan tingkat kepercayaan 77,3% di*bundling* promo dengan produk *Espresso* yang mempunyai presentase *support* 18%
3	Produk yang sering dibeli dengan tingkat kepercayaan 63,5% di*bundling* promo dengan produk *Japanese Ice* yang mempunyai presentase *support* 22%
4.	Produk yang sering dibeli dengan tingkat kepercayaan 58% di*bundling* promo dengan produk Nasgor Rawon yang mempunyai presentase *support* 23% atau dengan produk Nasgor Terasi yang mempunyai presentase *support* 25%

I. KESIMPULAN

Berdasarkan analisis dan pengujian menggunakan *software RapidMiner* dan Algoritma *FP-Growth* terhadap data transaksi penjualan *The Javanese Café*, menghasilkan 4 rekomendasi paket menu promosi yang dapat menunjang strategi pemasaran yaitu; (a) Produk yang sering dibeli dengan tingkat kepercayaan 87,3% di*bundling* promo dengan produk *Cinamond Crush* yang mempunyai presentase *support* 18%; (b) Produk yang sering dibeli dengan tingkat kepercayaan 77,3% di*bundling* promo dengan produk *Espresso* yang mempunyai presentase *support* 18%; (c) Produk yang sering dibeli dengan tingkat kepercayaan 63,5% di*bundling* promo dengan produk *Japanese Ice* yang mempunyai presentase *support* 22%; (d) Produk yang sering dibeli dengan tingkat kepercayaan 58% di*bundling* promo dengan produk Nasgor Rawon yang mempunyai presentase *support* 23% atau dengan produk Nasgor Terasi yang mempunyai presentase *support* 25%.

Pada penelitian selanjutnya dapat menggunakan *dataset* terbaru dan periode waktu lebih lama, sehingga jumlah transaksi yang lebih besar tersebut dapat menghasilkan nilai data transaksi penjualan dengan tingkat akurasi yang lebih tinggi. Serta dapat menggunakan algoritma *association rule* yang berbeda sehingga dapat mengetahui algoritma mana yang lebih cocok untuk diterapkan.

REFERENCES

[1] E. E. P. Wulandari, "Pengaruh Lokasi, Inovasi Produk, Dan Cita Rasa Terhadap Keputusan Pembelian Pada Eleven Cafe Di Kota Bengkulu," *J. Entrep. dan Manaj. Sains*, vol. 2, no. 1, pp. 74–86, 2021.

[2] R. P. C. Atmojo and C. Herdinata, "Pengaruh Harga, Promosi Dan Lokasi Terhadap Kepuasan Konsumen Pada Perusahaan Cv. Andindo Duta Perkasa," *PERFORMA J. Manaj. dan Start-Up Bisnis*, vol. 5, no. 5, pp. 379–388, 2020.

[3] A. E. Nasution, L. P. Putri, and M. T. Lesmana, "Analisis Pengaruh Harga, Promosi, Kepercayaan dan Karakteristik Konsumen Terhadap Keputusan Pembelian Konsumen Pada 212 Mart di Kota Medan," in *Proseding Seminar Nasional Kewirausahaan*, 2019, pp. 194–199.

[4] Tarigan Fahrul, Azanuddin, and Yanti Nur, "Implementasi Data Mining Menentukan Pola Penjualan Produk Toko Perabot Dua Bersaudara Kutalimbaru Dengan Menggunakan Fp-Growth," *J. CyberTech*, vol. 1, no. 2, pp. 115–129, 2021.

[5] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic J. Pendidik. Inform.*, vol. 4, no. 1, pp. 37–46, 2020.

[6] A. P. Sandia and V. W. Ningsih, "Implementasi Data Mining Sebagai Penentu Persediaan Produk Dengan Algoritma Fp-Growth Pada Data Penjualan Sinarmart," *JUPIKOM*, vol. 1, no. 2, pp. 111–122, 2022.

[7] K. T. Wijaya and I. Pratama, "Penerapan Algoritma FP-Growth Untuk Analisis Data Transaksi Penjualan Di Internet Learning Cafe Kaliurang," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 4, pp. 642–651, 2022.

[8] I. Afriyani and I. Ali, "Implementasi Data Mining Terhadap Data Penjualan Pada Industri Kuliner Menggunakan Algoritma Fp-Growth," *E-Link J. Tek. Elektro dan Inform.*, vol. 18, no. 1, pp. 40–49, 2023.

[9] S. Suhada, D. Ratag, G. Gunawan, D. Wintana, and T. Hidayatulloh, "Penerapan Algoritma Fp-Growth Untuk Menentukan Pola Pembelian Konsumen Pada Ahass Cibadak," *J. Swabumi*, vol. 8, no. 2, pp. 118–126, 2020.

[10] R. M. Thamrin and R. Andriani, "Perancangan Sistem Informasi Pendaftaran Dan Pengelolaan Data Skripsi Mahasiswa Berbasis Web Design Registration Information System And Web-Based Data Management of Student Thesis," *J. Sisfotenika*, vol. 11, no. 1, pp. 101–110, 2021.

[11] A. R. Wibowo and A. Jananto, "Implementasi Data Mining Metode Asosiasi Algoritma FP-Growth Pada Perusahaan Ritel," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 2, p. 200, 2020, doi: 10.35585/inspir.v10i2.2585.

[12] F. Febriantho, D. Renald, Y. Yakub, E. Edy, and I, "Penerapan Association Rule Data Mining Untuk Rekomendasi Produk Kosmetik Pada Pt. Fabiando Sejahtera Menggunakan Algoritma Apriori," *J. Algor*, vol. 2, no. 1, pp. 1–11, 2020, doi: 10.31253/algor.v2i1.437.

[13] J. L. Putra, M. Raharjo, T. A. A. Sandi, Ridwan, and R. Prasetyo, "Implementasi Algoritma Apriori Terhadap Data Penjualan Pada Perusahaan Retail," *J. PILAR Nusa Mandiri*, vol. 15, no. 1, pp. 87–90, 2019.

[14] N. Hutabarat and J. R. Sagala, "Penerapan Data Mining Untuk Pemesanan Pemesanan Percetakan Bahan Ajar Dengan Algoritma Apriori (Studi Kasus Percetakan Wendy)," *J. Method.*, vol. 4, no. 2, pp. 25–32, 2018.

[15] C. E. Firman, "Penentuan Pola Yang Sering Muncul Untuk Penjualan Pupuk Menggunakan Algoritma Fp-Growth," *Inform. J. Inform. Manaj. dan Komput.*, vol. 9, no. 2, pp. 1–8, 2017.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

[16] N. D. Andriai, T. W. Utami, and R. Wasodo, "Implementasi Algoritma Fp-Growth Dalam Market Basket Analysis Untuk Menganalisis Pola Belanja Konsumen Pada Data Transaksi Penjualan," 2019.

[17] R. P. Karina, S. Lestanti, and F. Febrinita, "Penerapan Algoritma Apriori Dalam Seleksi Penjurusan Calon Peserta Didik Baru Di Smak Diponegoro Blitar," *JATI(Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 716–724, 2022.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1904, Copyright ©2023
Submitted : August 3, 2023, Revised : August 18, 2023, Accepted : October 5, 2023, Published : November 7, 2023

Segmentasi Ruang Jantung Dalam Kondisi Kardiomegali Menggunakan Metode CNN Dengan Arsitektur U-Net

Tommy Saputra[1], Siti Nurmaini[2]*, Muhammad Taufik Roseno[3], Hadi Syaputra [4]
Magister Ilmu Komputer, Universitas Sriwijaya[1], [2]
Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Sumatera Selatan [3], [4]
Jl. Raya Palembang - Prabumulih Km. 32 Indralaya, OI, Sumatera Selatan 30662, Telp. 0711-58069 [1], [2]
Jl. Letnan Murod Nomor 55 Talang Ratu Km5, Palembang 30128, Telp. 0811-720-8110 [3], [4]
09012682226012@students.ilkom.unsri.ac.id[1], siti_nurmaini@unsri.ac.id[2],
mtaufikroseno@uss.ac.id[3], hadisyaputra@uss.ac.id[4]

***Abstract*— Cardiomegaly is a disease in which sufferers show no symptoms and have symptoms such as shortness of breath, abnormal heartbeat and edema. Cardiomegaly will cause the sufferer's heart to pump harder than usual. Early diagnosis of cardiomegaly can help make decisions about whether the heart is abnormal or normal. In addition, due to the problem that manual examination takes time and requires human interpretation and experience, tools are needed to automatically develop and identify normal and abnormal hearts. Therefore, this study proposes cardiac chamber segmentation using 2D (two-dimensional) ultrasound convolutional neural networks for rapid cardiomegaly screening in clinical applications based on heart ultrasound examination. The proposed approach uses a CNN with a U-Net architecture model with abnormal and normal heart data. The research results obtained used the pixel matrix evaluation Avg_accuracy of 99.50%, Val_accuracy of 97.98% and Mean_IoU of 90.01%.**

Keywords*— *Segmentation, Cardiomegaly, Convolutional Neural Network, U-Net

***Abstrak*—Kardiomegali merupakan penyakit yang tidak memiliki tanda atau gejala pada beberapa penderitanya dan kemungkinan memiliki gejala seperti sesak napas, detak jantung tidak normal, dan edema. Kardiomegali akan menyebabkan jantung penderita memompa lebih keras dari biasanya atau secara bertahap merusak otot jantung seperti jantung berdebar, sesak dada, dan sesak napas. Mendiagnosis dini kardiomegali dapat membantu membuat keputusan jantung dalam keadaan abnormal atau normal. Selain itu, sehubungan dengan masalah dalam pemeriksaan secara manual akan memakan waktu dan kebutuhan interpretasi serta pengalaman manusia, maka diperlukan alat bantu untuk secara otomatis mengembangkan dan mengidentifikasi jantung normal dan jantung yang abnormal. Oleh karena itu, penelitian ini mengusulkan segmentasi ruang jantung dengan menggunakan jaringan saraf konvolusional 2D (dua dimensi) *ultrasound* untuk *skrining* kardiomegali secara cepat dalam aplikasi klinis berdasarkan pemeriksaan USG jantung. Pendekatan yang diusulkan menggunakan CNN dengan model arsitektur *U-Net* dengan data jantung abnormal dan normal. Hasil penelitian yang diperoleh menggunakan evaluasi matriks piksel Avg_akurasi sebesar 99,50%, Val_akurasi 97,98% dan Mean_IoU sebesar 90,01%.**

Kata Kunci*— *Segmentasi, Kardiomegali, Convolutional Neural Network, U-Net

I. PENDAHULUAN

Jantung yang membesar, yang dikenal sebagai kardiomegali, bukanlah penyakit serius, dan mungkin tidak memiliki tanda dan memiliki gejala, seperti sesak napas, detak jantung tidak normal (aritmia), dan edema. Kardiomegali akan menyebabkan jantung penderita memompa lebih keras dari biasanya atau secara bertahap merusak otot jantung. Pada detak jantung yang tidak normal dapat menyebabkan pembesaran jantung, mengakibatkan tekanan darah tinggi, penyakit katup jantung, dan kardiomiopati. Risiko komplikasi kardiomegali termasuk gagal jantung, pembekuan darah, murmur jantung, dan serangan jantung [1].

Berdasarkan data Riskesdas 2019 menyebutkan bahwa 17 dari 1100 orang atau sekitar 2.774.135 orang di Indonesia dan salah satu pulau di jawa terdapat 365.427 orang yang menderita penyakit jantung dalam kondisi kardiomegali [2]. Mayoritas penderita kardiomegali adalah seseorang yang usia lebih dari 55 tahun, sedangkan minoritasnya penderita kardiomegali lebih ke bayi dan anak-anak.

Pada penelitian yang dilakukan sebelumnya [3] dalam melakukan klasifikasi citra jantung janin yang belum sampai objek didapatkan jantung janin, maka perlu dilakukan segmentasi.

Segmentasi citra adalah proses pemisahan antara objek atau area yang menarik (*foreground*) dengan latar belakang (*background*) dalam citra digital. Proses ini sangat penting dalam pemrosesan citra karena membantu dalam memisahkan objek yang ingin dianalisis atau dikenali dari latar belakangnya. Ada dua jenis utama segmentasi citra, yaitu *full segmentation* dan *partial segmentation.* [4]

Dalam pemrosesan citra medis, pemilihan model dan teknik *deep learning* yang sesuai dapat memainkan peran penting dalam mencapai hasil yang baik. Seringkali, jenis citra dan masalah yang ingin dipecahkan akan mempengaruhi pilihan model yang tepat.[5]

CNN dengan arsitektur U-Net adalah pilihan yang umum digunakan dalam segmentasi citra medis, termasuk segmentasi ruang jantung dalam kondisi kardiomegali. U-Net adalah arsitektur jaringan saraf tiruan yang khusus dirancang untuk

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

tugas segmentasi, dan ia memiliki lapisan konvolusi yang dapat mengambil fitur-fitur dari berbagai tingkat abstraksi dalam citra. Ini membuatnya sangat efektif dalam memisahkan area yang berbeda dalam citra medis [6].

Pada beberapa penelitian yang telah dilakukan sebelumnya dalam segmentasi jantung dalam kondisi kardiomegali yang menggunakan arsitektur U-Net dengan menghasilkan nilai akurasi 93,00% [7], selain itu pada penelitian terdahulu yang menggunakan metode CNN dengan arsitektur U-Net+CRF dengan menghasilkan nilai akurasi 95,03% [8], maka memilih CNN menggunakan arsitektur U-Net bisa menjadi pilihan yang lebih baik. Penting untuk diingat bahwa dalam pengembangan model *deep learning*, sangat penting untuk melakukan eksperimen, menyesuaikan arsitektur, dan menyesuaikan parameter untuk mendapatkan hasil terbaik. Setiap dataset dan masalah bisa memiliki karakteristik yang berbeda. Oleh karena itu, perlu adanya adaptasi khusus untuk mencapai hasil yang optimal. Terus berkembangnya teknologi *deep learning* akan terus membantu dalam meningkatkan kemampuan analisis citra medis untuk diagnosis penyakit jantung dan banyak masalah medis lainnya sehingga pada penelitian ini mengusulkan untuk melakukan proses segmentasi ruang jantung dalam kondisi kardiomegali dengan teknik segmentasi yang diusulkan adalah menggunakan *Convolutional Neural Network* dengan arsitektur *U-Net*.

II. LANDASAN TEORI

A. Kardiomegali

Kardiomegali adalah sebuah kondisi dalam keadaan anatomis pada besarnya jantung lebih besar dari ukuran jantung normal yaitu dengan ukuran diameter transversal dari gambar jantung ≥ 50% dari besar rongga dada. Kardiomegali terjadi ketika besar jantung >50% dari diameter internal rongga dada. Kardiomegali biasanya manifestasi dari proses patologis. Pembesaran jantung dapat melibatkan pembesaran dari atrium atau ventrikel kanan atau kiri ataupun keduanya, namun pada umumnya kardiomegali diakibatkan oleh pembesaran bilik jantung kiri, sehingga terjadinya pengecilan ukuran rongga torak jantung. Ukuran jantung dapat beravariasi bergantung pada bentuk tubuh, namun batas teratas dari ukuran jantung dewasa adalah 15.5 cm untuk wanita dan 16 cm untuk pria[2].

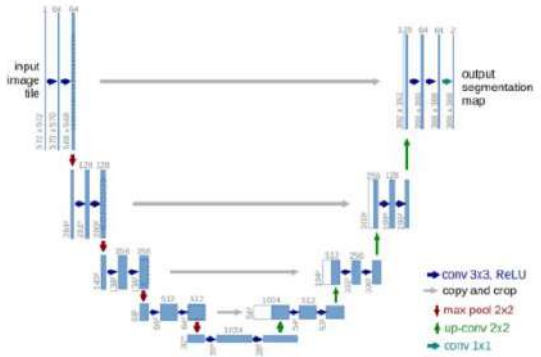
Pada proses klinis, *Ultrasound* merupakan suatu tahapan utama yang digunakan untuk mengidentifikasikan ukuran jantung karena hasil citra yang didapatkan bersifat real time. Adapun tahapan dalam pemerosesan pemindahan sudut pandang ruang jantung yaitu dengan cara ukuran jantung yang hampir mendekati rongga dada dari video pemindaian *Ultrasound* untuk proses segmentasi. [2].

B. Clahe

Contrast Limited Adaptive Histogram Equalization (CLAHE) merupakan bagian dari proses *enhancement* yang mana digunakan untuk memperbaiki kualitas citra dan juga menghilangkan noise. *Clahe* adalah generalisasi dari *Adaptive Histogram Equalization (AHE*). Clahe beroperasi pada tile (region kecil pada citra graysclale), dimana kontras pada disetiap tile akan diperbaiki sehingga menghasilkan menghasilkan histogram yang ditentukan

C. Arsitektur U-Net

U-Net adalah salah satu arsitektur jaringan saraf tiruan yang sangat populer dalam tugas segmentasi citra medis dan pemrosesan citra lainnya. Arsitektur ini dirancang untuk mengatasi masalah segmentasi citra, di mana tujuannya adalah untuk memisahkan objek dari latar belakang dalam citra. Arsitektur *U-Net* memiliki struktur yang khas berbentuk "U" dapat dilihat pada gambar 2 berikut ini:



Gambar 1. Arsitektur U-Net

Berdasarkan gambar diatas bahwa bagian utama dari U-Net terdiri dari dua bagian utama yaitu *Contracting Path* dan *Expansive Path.*

III. METODE PENELITIAN

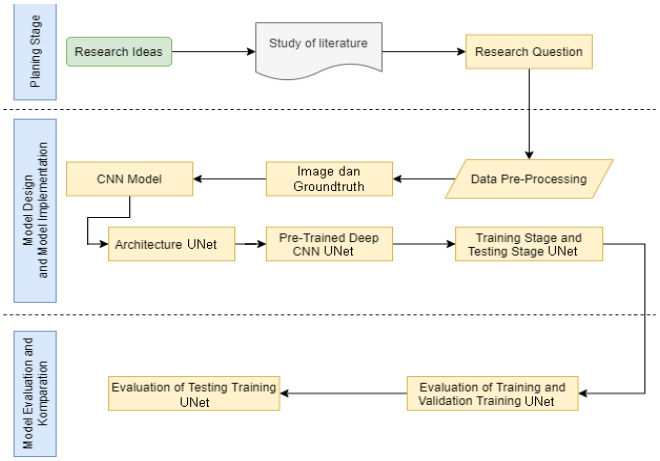
Metode penelitian yang digunakan adalah berfokus pada perancangan model sistem segmentasi ruang jantung dalam kondisi kardiomegali menggunakan citra 2D *ultrasound image* pada *convolutional neural network*. Pada proses pra-pengolahan data untuk dapat dilakukan analisis dengan cara memproses pra-pengolahan data sebelum dilakukan ke tahap selanjutnya. Data secara acak dibagi untuk proses pelatihan, validasi dan pengujian. Jumlah data yang digunakan untuk proses pelatihan adalah 80% dan untuk pengujian sebanyak 20% dari total gambar citra yang digunakan pada penelitian ini. Proses segmentasi dengan menggunakan metode *CNN* dengan menerapkan arsitektur *U-Net* untuk melakukan proses segmentasi yang menunjukan hasil kinerja baik dalam tugas Segmentasi. Dalam proses segmentasi ini, ada tiga tahapan yaitu proses pelatihan, proses validasi dan proses pengujian.

A. Kerangka Berfikir

Pada kerangka berpikir dan tahapan penelitian yang akan dilakukan, serta perencanaan model yang akan dibuat dan rencana evaluasi model menggunakan metode *Convolutional Neural Network (CNN)* dengan arsitektur *U-Net*.Dalam melakukan penelitian ini, untuk mempermudahnya maka dijabarkan langkah-langkah apa saja yang akan diambil dalam melakukan penelitian ini. Kerangka pikir dari penelitian ini

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

di representasikan pada Gambar 3.



Gambar 2. Kerangka Berfikir Penelitian

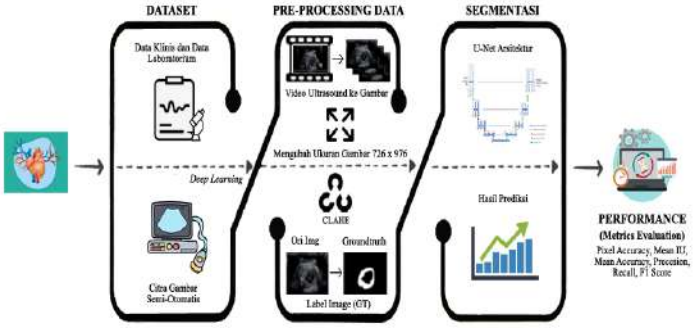
Tahap pertama adalah menentukan tahap penelitian. dalam tahap penelitian yang pertama adalah menentukan sebuah ide suatu penelitian dari analisis literatur, berupa paper/makalah dari jurnal berkualitas yang berkaitan dengan metode *Convolutional Neural Network (CNN)* dengan arsitektur *U-Ne*t. Dari masalah tersebut peneliti menemukan sebuah solusi untuk menyelesaikan atau memperbaiki masalah tersebut sehingga dibuatlah rumusan masalah penelitian atau reseach question. Tahap selanjutnya adalah tahap implementasi model terhadap permasalahan jantung dalam kondisi kardiomegali dengan model *CNN* arsitektur *U-Ne*t. Peneliti melakukan pengumpulan data vidio *Ultrasound / USG* dari Rumah Sakit Bunda Palembang. Kemudian data vidio yang telah didapatkan dikelola menjadi gambar, setelah itu Dataset ini akan dibagi menjadi tiga bagian yaitu *training, testing* dan *validiting*. Sebelum citra masuk ke dalam sebuah arsitektur model, dilakukan proses *image processing*. Prosesing gambar ini bertujuan untuk mempertajam gambar agar bisa di label menggunakan *tools labelme*, sebelum memulai ngelabel. Terlebih dahulu mempelajari permasalahan Jantung dalam kondisi kardiomegali dengan dokter dr.Putri Mirani,Sp.OG setelah mengetahui jantung dalam kondisi kardiomegali kemudian hasil gambar *usg/ultrasound* jantung dilabel dan di *ground truth* dari hasil gambar yang telah dilabel.

Convolutional Neural Network (CNN) dan model *pre-trained Deep Convolutional Neural Network*. Dimana setiap model *pre-trained Deep CNN* yang padat pada layer bawahnya bekerja sebagai *feature extractor* untuk mengenali citra yang dimasukkan. lapisan yang padat ini berisi banyak layer-layer proses konvolusi pada tiap lapisannya yang berfungsi sebagai dimension reduction tanpa mengurangi informasi yang penting dari citra. Setelah gambar melalui proses *feature extractor* kemudian dilakukan *fine-tuning* dimana lapisan atas arsitektur *pre-trained* khusus label rongga torak, proses ini disebut *feature classification.* setelah itu peneliti akan melakukan proses pelatihan dan pengujian.

Kemudian langkah selanjutnya adalah *evaluation* and *analysist resul*t. dalam tahap ini peneliti menggunakan *perfomance metric* untuk mengukur akurasi performa dari model. Di mana peneliti menggunakan *classification metrics* untuk menghitung rasio prediksi benar dibagi dengan rasio prediksi salah. Jika akurasi tercapai maka model arsitektur akan disimpan untuk digunakan dalam pengujian yang disebut *validation test.*.

B. *Tahapan Penelitian*

Convolutional Neural Networks (CNNs) merupakan metode yang akan digunakan dalam penelitian ini menggunakan Arsitektur *U-Net*. Gambar 4 menunjukkan tahapan penelitian yang akan dilakukan :



Gambar 3. Tahapan Penelitian

Berdasarkan gambar diatas dapat dijelaskan bahwa tahapan penelitian yang akan dilakukan memiliki 3 tahapan yakni:

1. Persiapan Data

Pada tahap persiapan data yang diperlukan adalah Dataset yang digunakan sebagai input citra yang dibuat merupakan data gambar kardiomegali dari video ultrasound yang terkena kardiomegali dan normal. Jumlah video yang digunakan terdiri dari 10 video berupa 5 video normal dan 5 video abnormal. Adapun sumber video yang didapatkan di jelaskan pada tabel 1 berikut ini:

Tabel 1. Kardiomegali dan Normal

Kardiomegali				
				
Normal				
				

Berdasarkan tabel diatas sumber video yang digunakan terdiri dari 10 video yang siap diolah untuk ketahap selanjutnya. Video didapatkan dari sumber Rumah Sakit Bunda Palembang dan untuk lebih jelas deskripsi raw data yang digunakan pada penelitian ini dapat dilihat pada Tabel 2.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

Tabel 2. Deskripsi raw data video

Data		Type of file	Size	Duration
Abnormal	Pasien 1	.mp4	41,966kb	42 d/s
	Pasien 2	.mp4	2,748kb	14 d/s
	Pasien 3	.mp4	57,062kb	29 d/s
	Pasien 4	.mp4	9,158kb	45 d/s
	Pasien 5	.mp4	1,543kb	4 d/s
Normal	Pasien 1	.mp4	12,804kb	40 d/s
	Pasien 2	.mp4	8,150kb	31 d/s
	Pasien 3	.mp4	11,547kb	39 d/s
	Pasien 4	.mp4	9,646kb	29 d/s
	Pasien 5	.mp4	7,388kb	26 d/s

Jumlah data seperti yang telah dijelaskan pada tabel diatas. Raw data dilakukan pengolahan lebih lanjut untuk mengambil beberapa bagian citra untuk dikelola sebagai tolak ukur yang ditetapkan untuk perbandingan dengan *expert* dan mesin. Hasil jumlah dari proses pemindahan video ke gambar (foto) maka dapat dilihat pada tabel 3.

Tabel 3. Ekstrasi Video Ke Gambar (Foto)

Data		Hasil Ekstrasi Video ke gambar (foto)
Abnormal	Pasien 1	93
	Pasien 2	85
	Pasien 3	96
	Pasien 4	88
	Pasien 5	96
Normal	Pasien 1	87
	Pasien 2	85
	Pasien 3	97
	Pasien 4	105
	Pasien 5	94

Berdasarkan tabel diatas, dapat dijelaskan bahwa seluruh hasil ekstrasi video abnormal ke gambar berjumlah 458 gambar dan normal berjumlah 468 gambar yang kemudian data tersebut dikelolah ketahap selanjutnya.

2. Preprocessing

Proses selanjutnya dalam melakukan segmentasi citra yaitu pada pre-processing yang digunakan untuk meningkatkan hasil kualitas citra yang diinput. Ada tahapan *pre-processing* yaitu pengubahan ukuran gambar, peningkatan kontras gambar dan *ground trurth*. Berikut rangkuman dalam table 4 dan akan dibahas secara terperinci.

Tabel 4. Tahapan Prepocessing

Tahapan Preprocessing	**Target Data**
Input Video	Gambar (Foto)
Ukuran Gambar	Foto(Ukuran)
Image Procesing	Clahe
Ground Truth	Label Gambar

Input Video

Pada tahap ini, data video *ultrasound* yang didapatkan akan diolah menjadi beberapa gambar, dalam proses *framing* yang dilakukan pada proses ini dilakukan dengan menggunakan bantuan tool OpenCV di bahasa pemrograman python. *Frame* yang dihasilkan didapatkan dari video yang diinput. Langkah selanjutnya akan diproses ketahap penyamarataan ukuran gambar.

Ukuran Gambar

Tahap selanjutnya adalah mengubah ukuran gambar pada hasil frame yang didapatkan sebelumnya dengan cara resize gambar menjadi 976x726 piksel. Hal ini dilakukan untuk dapat menyeimbangkan semua bingkai foto yang akan digunakan untuk ketahap selanjutnya. Proses tahapan resize sama seperti frame diatas menggunakan tool OpenCV di pemrograman python

Image Processing

Tahap selanjutnya *CLAHE (Contrast Limited Adaptive Histogram Equalization)* adalah teknik yang digunakan untuk meningkatkan kontras dan menghilangkan noise pada citra dengan kontras rendah. Ini adalah salah satu teknik pengolahan citra yang populer untuk meningkatkan hasil kualitas citra yang akan dikelolah.

Ground Truth

Pada tahapan ini mengatur data citra pada segmentasi manual citra yang dilihat ukuran jantung terhadap rongga torak. Tahap ini dilakukan segmentasi manual dengan menggunakan labelme. Tabel 4 merupakan hasil dari segmentasi manual.

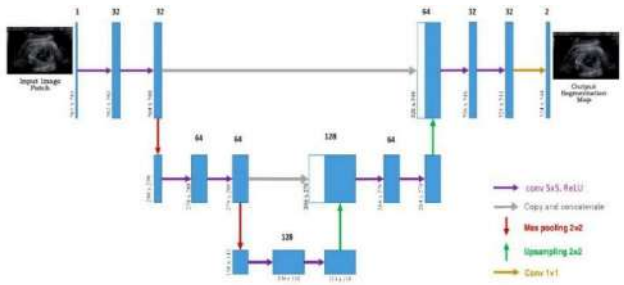
Tabel 5. Hasil Ground Truth

Abnormal		**Normal**	
Image	**Ground Truth**	**Image**	**Ground Truth**
			
			
			
			
			

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

3. Segmentasi Arsitektur U-Net

Dalam proses segmentasi jantung yang dalam kondisi kardiomegali menggunakan proses arsitektur U-Net. Arsiktetur U-Net adalah arsitektur yang memiliki bentuk U dan bentuknya simetris yang terdiri dari 2 bagian yaitu bagian kiri dan bagian kana. Pada segmentasi jantung dalam kondisi kardiomegali ini arsitektur U-Net yang dipakai adalah sebagai berikut :



Gambar 4. Arsitektur U-Net Digunakan Untuk Segmentasi

Pada gambar diatas menunjukkan bahwa arsitektur U-Net memiliki dua lapisan *convolution*. Pada lapisan *convolution* pertama adalah *contracting path* yang berjumlah satu saluran pada inputan gambar kemudian terjadi perubahan jumlah saluran dari satu menjadi tiga puluh dua karena meningkatnya kedalaman gambar dalam proses konvolusi, fungsi aktivasi yang digunakan *ReLU* dan kernel size 5x5. Langkah selanjutnya dilakukan penggabungan maksimal yang mana membagi ukuran gambar menjadi separuh ditunjukkan oleh tanda panah merah. Ukuran diperkecil dari 564x584 menjadi 560x580 disebabkan oleh masalah *padding*, *padding* yang digunakan yaitu "*same*".

Proses selanjutnya pada lapisan layer ketiga dilakukan penggabungan dua layer dengan Ukuran diperkecil dari 138x143 menjadi 134x139 disebabkan oleh masalah *padding*, *padding* yang digunakan yaitu "*same*". Langkah Selanjutnya yaitu tenik *upsampling* untuk dapat memperbesar gambar keukuran aslinya dengan mengali duakan ukuran citra tersebut yang ditunjukkan oleh panah hijau.

Pada proses *expensive path* pada lapisan pertama dari hasil gabungan gambar dengan lapisan *contracting path* dikarenakan untuk mendapatkan hasil prediksi yang lebih akurat. Setelah dilakukan teknik *upsampling* ukuran citra berubah dari 134x139 menjadi 268x278. Ukuran diperkecil dari 268x278 menjadi 264x274 disebabkan oleh masalah *padding*, *padding* yang digunakan yaitu "*same*".Langkah Selanjutnya yaitu tenik *upsampling* untuk dapat memperbesar gambar keukuran aslinya dengan mengali duakan ukuran citra tersebut yang ditunjukkan oleh panah hijau.

Pada proses *expensive path* pada lapisan kedua dari gabungan gambar dengan lapisan *contracting path* pertama. Setelah dilakukan teknik *upsampling* ukuran citra berubah dari 264x274 menjadi 528x548. Jumlah saluran berubah dari 64 menjadi 32 dengan menggunakan fungsi aktivasi *ReLU* dan kernel size 5x5. Ukuran diperkecil dari 528x548 menjadi 524x544 disebabkan oleh masalah *padding*, *padding* yang digunakan yaitu "*same*".

IV. HASIL DAN PEMBAHASAN

Hasil dari penelitian segmentasi setelah pelatihan model U-net pada gambar *USG/Ultrasound*. U-Net dapat menggunakan *hyperparameter* dengan *Activation Function (ReLU), Sigmoid Activation Function*, dan *Loss Using Dice Coefficient Function*[10]. Pada bagian ini dilakukan proses segmentasi jantung menggunakan gambar USG yang terdeteksi oleh pembengkakan jantung (kardiomegali) dan normal. Teknik yang digunakan dalam penelitian ini menggunakan metode *Convolutional Neural Network* berdasarkan arsitektur U-Net.

Pada proses pelatihan data dilakukan dengan menggunakan dataset yang telah dilakukan *preprocessing* sebanyak 466 citra data yang terdiri dari data *citra original* dan *ground truth*. Data yang digunakan berisi gambar kardiomegali dan normal. Data informasi yang digunakan dapat dilihat pada Tabel 5.

a. *Training Paramaters*: Kami melakukan proses pelatihan *Convolutional Neural Network* berbasis U-Net model dengan epoch 150 dan *batch size* 8. Penggunaan epoch 150 dikarenakan untuk mendapatkan hasil yang tepat untuk proses segmentasi khususnya pada analisis data medis membutuhkan proses yang sangat lama dalam pengolahan citra komputasi dilakukan proses pelatihan. Kami menggunakan tingkat pembelajaran 1e-5 dengan *Adam optimizer* dan *smooth loss* 1−5, *threshold* 0,5.

b. *Post-processing*: Apa yang didapat dari data prediksi dari suatu model terkadang memiliki beberapa wilayah dengan label yang berbeda, tidak seperti segmentasi ground truth yang telah dilakukan sebelumnya [11]. Studi ini menggunakan postprocessing untuk membantu mendapatkan hasil yang maksimal, sedangkan *post processing* digunakan dengan menggunakan *thresholding* algoritma. *Thresholding* adalah proses membagi gambar menjadi dua atau lebih kelas piksel, seperti dalam hal ini kasus, itu adalah "latar depan" dan "latar belakang". Algoritma *thresholding* dapat membantu dalam pengolahan citra di hal menghilangkan kebisingan dan memungkinkan untuk meningkatkan akurasi yang tinggi [12]. Hasil prediksi diperoleh pada tahap selanjutnya arsitektur U-Net mengolah hasil prediksi citra dengan menggunakan algoritma thresholding yang digunakan dalam penelitian ini menggunakan *fixed thresholding*. Tahapan selanjutnya adalah proses validasi dan evaluasi untuk melihat seberapa baik metode CNN bekerja dengan arsitektur U-Net aktif data gambar USG dipengaruhi oleh pembengkakan jantung.

C. *Metrics Evaluation*: Segmentasi merupakan langkah-langkah penting untuk pra-pemrosesan dalam aplikasi analisis citra[13]. Dari hasil segmentasi tersebut, maka akan dapat diidentifikasi area terpenting dan objek terjadinya suatu peristiwa yang sangat berguna dalam analisis citra selanjutnya[14].

Pada proses segmentasi menggunakan *metrics evaluation* pada Average Accuracy, Val-Accuracy, Mean IoU, IoU_Coef, Loss, Val_Loss, Val_IoU_Coef. Hasil kinerja dari segmentasi dengan menggunakan Arsitektur U-Net yang telah dilakukan dapat dilihat pada Tabel 7.

Pelatihan pada model U-Net dengan data acak 458 gambar jantung dalam kondisi kardiomegali dengan lima objek pasien dan data acak berjumlah 468 gambar jantung normal dengan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

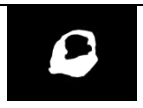
lima objek pasien yang dapat dilihat pada tabel 3 dan tabel 5. Pada tabel 6 menunjukkan hasil sampel segmentasi dari Arsitektur U-Net.

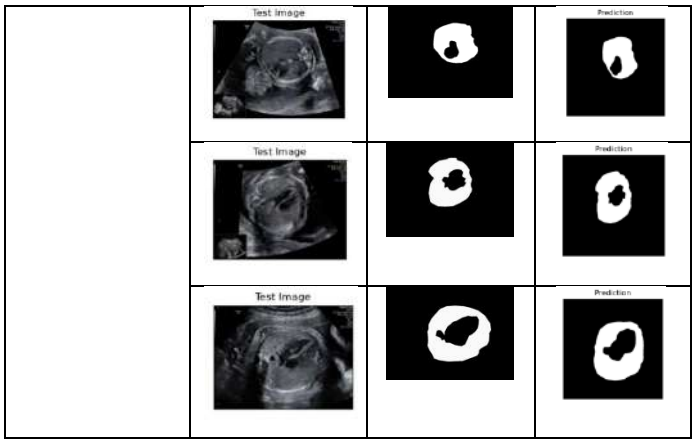
Hasil pengujian pelatihan U-Net berdasarkan plot grafik akurasi hasil sebagai ekstraktor fitur dengan 150 epoch ditunjukkan pada Gambar 6. Berdasarkan hasil prediksi dari model U-Net yang telah dilakukan dan langkah selanjutnya adalah menguji model tersebut dengan berbasis evaluasi metrik. Hasil kinerja untuk segmentasi dapat dilihat pada Tabel 7.

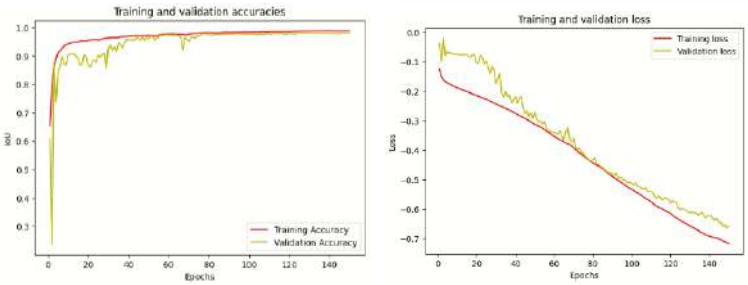
Tabel 6. Split Data Train, Testing, dan Validating

Data	Data Train		Data Testing		Data Validating		Total
	Org Img	Ground Truth	Org Img	Ground Truth	Org Img	Ground Truth	
Kardiomegali	161	34	34	34	34	34	458
Normal	164	35	35	35	35	35	468

Tabel 7. Hasil Prediksi U-Net

Data	Image	Ground Truth	Prediksi U-Net
Kardiomegali			
			
			
			
			
Normal			
			





Gambar 5. Hasil (a) Accuracy (b) loss dari ujicoba menggunakan U-Net

Tabel 8. Hasil Kinerja Segmentasi menggunakan Arsitektur U-Net

Metrics Evaluation	**U-Net**
Avg_Accuracy	99,50 %
Accuracy	98,87 %
Val_Accuracy	97,98 %
Mean IoU	90,01%
IoU_coef	72,24 %
Loss	- 72,24 %
Val_loss	- 66,76 %
Val_IoU_coef	66,86 %

Berdasarkan tabel 7, untuk menilai kinerja model maka dapat diprediksi akurasi secara keseluruhan pada citra jantung kondisi kardiomegali.Hasil average accuracy yang didapatkan adalah 99,50%, Val Accuracy yang didapatkan 97,98% dan Mean IoU yang didapatkan sebesar 90,01%, maka dapat didefiniskan keberhasilan model dalam mensegmentasi citra jantung kondisi kardiomegali dan normal. Terlihat pada tabel 6, bahwa kinerja arsitektur yang digunakan memiliki hasil diatas 90%. Selanjutnya dilakukan interpretasi hasil dengan membandingkan terhadapat penelitian terdahulu yang sudah dilakukan. Adapun perbandingan hasil segmentasi jantung kondisi kardiomegali dengan penelitian lain dapat dilihat pada tabel 8.

Tabel 9 Perbandingan Hasil Evaluasi Kinerja dengan Penelitian lain

Metode	Avg Akurasi (%)	Val Akurasi (%)	Mean IoU (%)

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

CNN (J. X. Wu et al, 2022)	98,00 %	-	-
U-Net+CRF (Z. Li et al, 2019)	95,30 %	97,20%	-
U-Net (A.Bouslama et al, 2020)	93,00 %	-	83,05 %
CadioNet (A.Jafat et al, 2022)	92,01 %	-	-
Metode yang diusulkan	99,50 %	97,98%	90,01 %

Berdasarkan pada tabel 8, dapat dilihat bahwa nilai pengukuran performa untuk dataset kardiomegali menggunakan metode yang memiliki nilai yang paling tinggi pada semua aspek. Dapat dilihat bahwa nilai ini merupakan yang paling baik untuk nilai Avg Akurasi sebesar 99,50 % dan Mean IoU sebesar 90,01 % dari penelitian sebelumnya. Berdasarkan perbandingan hasil penelitian yang telah dilakukan bahwa metode yang diusulkan telah mendapat hasil kinerja yang optimal dalam mensegmentasikan citra kardiomegali. Penelitian ini hanya membahas tentang segmentasi pada jantung kondisi kardiomegali dan normal. Hasil segmentasi pada penelitian ini dapat digunakan untuk mendeteksi adanya kelainan pada jantung. Penelitian ini selanjutnya dapat dimanfaatkan hasil segmentasi penelitian ini untuk pada masalah klasifikasi atau prediksi jenis kelainan pada jantung dalam kondisi kardiomegali.

V. KESIMPULAN

Hasil dari penelitian ini dapat disimpulkan mendapatkan hasil yang sangat baik dalam mensegmentasi gambar jantung dengan kondisi kardiomegali dan normal. Hal ini bisa dilihat dari nilai Avg Akurasi, Val Akurasi dan Mean IoU yang tinggi. Kemampuan model yang diusulkan untuk segmentasi jantung kondisi kardiomegali sudah mendapatkan hasil sangat baik, dilihat dari Avg_akurasi sebesar 99,50%, Val_akurasi 97,98% dan Mean_IoU sebesar 90,01%.

Penelitian ini menggunakan metode *Convolutional Neural Network* dengan arsitektur U-Net untuk melakukan segmentasi citra jantung dalam kondisi kardiomegali dan normal. Sehingga dalam penelitian ini dapat membantu bidang medis untuk mensegmentasi ukuran jantung pasien secara akurat dan diharapkan juga untuk penelitian kedepannya bahwa segmentasi citra jantung dapat dilakukan dengan jaringan semantic untuk melatih node karena sangat bermanfaat terutama dalam bidang medis dan untuk proses klasifikasi penyakit jantung kondisi kardiomegali.

REFERENSI

[1] Semsarian, C.; Ingles, J.; Maron, M.S.; Maron, B.J. New Perspectives on the Prevalence of Hypertrophic Cardiomyopathy. J. Am. Coll. Cardiol. 2015, 65, 1249–1254.

[2] Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan RI. 2019. Riset Kesehatan Dasar

[3] Roseno, M. T., & Syaputra, H. (2022). Classification of Fetal Heart Based on Images Augmentation Using Convolutional Neural Network Method. Journal of Information Systems and Informatics, 4(2), 252–265. https://doi.org/10.51519/journalisi.v4i2.247.

[4] Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. arXiv 2017, arXiv:1606.00915.

[5] D. Ravi et al., "Deep learning for health informatics," IEEE J. Biomed. Heal. informatics, vol. 21, no. 1, pp. 4–21, 2017.

[6] Hatamizadeh, A.; Terzopoulos, D.; Myronenko, A. Edge-Gated CNNs for Volumetric Semantic Segmentation of Medical Images. arXiv 2020, arXiv:2002.04207.

[7] H.-C. Shin, et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning.," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.

[8] Wu, J. X., Pai, C. C., Kan, C. D., Chen, P. Y., Chen, W. L., & Lin, C. H. (2022). Chest X-Ray Image Analysis with Combining 2D and 1D Convolutional Neural Network Based Classifier for Rapid Cardiomegaly Screening. *IEEE Access*, *10*, 47824–47836. https://doi.org/10.1109/ACCESS.2022.3171811

[9] Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.-M.; Larochelle, H. Brain Tumor Segmentation with Deep Neural Networks. Med. Image Anal. 2017, 35, 18–31.

[10] J. Dai, K. He, and J. Sun, "Convolutional feature masking for joint object and stuff segmentation," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 3992–4000, 2015, doi: 10.1109/CVPR.2015.7299025.

[11] T. Wiatowski and H. Bölcskei, "A mathematical theory of deep convolutional neural networks for feature extraction," IEEE Trans. Inf. Theory, vol. 64, no. 3, pp. 1845–1866, 2018.

[12] S. Rueda et al., "Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: A grand challenge," IEEE Trans. Med. Imaging, vol. 33, no. 4, pp. 797–813, 2014, doi: 10.1109/TMI.2013.2276943.

[13] G. Padmavathi, P. Subashini, and A. Sumi, "Empirical Evaluation of Suitable Segmentation Algorithms for IR Images," Int. J. Comput. Sci. Issues, vol. 7, no. 4, pp. 22–29, 2010.

[14] Mittal, A.; Hooda, R.; Sofat, S. LF-SegNet: A Fully Convolutional Encoder-Decoder Network for Segmenting Lung Fields from Chest Radiographs. Wirel. Pers. Commun. 2018, 101, 511–529.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1976, Copyright ©2023
Submitted : September 28, 2023, Revised : October 5, 2023, Accepted : October 16, 2023, Published : November 7, 2023

Analisis Perilaku Penggunaan Sistem Informasi Akademik Dengan Menerapkan Integrasi UTAUT 2 Pada Institut Ilmu Kesehatan dan Teknologi Muhammadiyah Palembang

Hendri Donan[1], Edi Surya Negara[2], Tata Sutabri [3], Firdaus [4]
Magister Komputer, Universitas Bina Darma[1], [2], [3], [4]
Palembang, Indonesia
Hendridn123@gmail.com[1] , e.s.negara@binadarma.ac.id[2] , tata.sutabri@gmail.com[3] , firdaus.dr@binadarma.ac.id[4]

Abstract *- The utilization of Information Technology (IT) in higher education setting aims to enhance the quality of education, and this initiative is realized through the implementation of Information Technology at the Institute of Health Sciences and Technology Muhammadiyah Palembang (IKesT MP) in the form of an Academic Information System (SIMAKAD). SIMAKAD is a vital role as a tool to manage internal data and serves as an information hub for students. This research is conducted to evaluate the acceptance level of the UTAUT2 model and the impact of both the main and target variables within the UTAUT2 model.*

This research utilizes a quantitative method with 150 respondents, analyzed using SMART PLS 3.0 software." *software. The research findings indicate that the acceptance level of the UTAUT2 model reaches 74%, signifying a high adoption rate. Variables like Perceived Value (p-Value: 0.019) and Habit (p-Value: 0.009) significantly influence Behavioral Intention, with a p-Value < 0.05, indicating that their hypotheses are accepted. On the other hand, variables such as Performance Expectancy (p-Value: 0.660), Effort Expectancy (p-Value: 0.417), Social Influence (p-Value: 0.652), and Facilitating Conditions (p-Value: 0.292) There is no substantial influence on Behavioral Intention as a result of using Information Technology (IT), indicating that their hypotheses have not been endorsed.. Additionally, the variable Hedonic Motivation (p-Value: 0.978) also does not can significantly impact one's inclination toward a behavior Intention. However, variables Facilitating Conditions (p-Value: 0.000) and Behavioral Intention (p-Value: 0.000) have a positive impact on Use Behavior, indicating that their hypotheses are accepted. Conversely, the variable Habit (p-Value: 0.915) Does not exert a significant impact on Uss Behavior, resulting in the rejection of its*

hypothesis.Keywords: behavioral intention, use behavior, SIMAKAD, UTAUT, UTAUT2

Abstrak **- Pemanfaatan Teknologi Informasi (TI) dalam lingkungan Perguruan Tinggi bertujuan untuk meningkatkan kualitas pendidikan, dan inisiatif ini diwujudkan melalui implementasi Teknologi Informasi di Institut Ilmu Kesehatan dan Teknologi Muhammadiyah Palembang (IKesT MP) dalam bentuk Sistem Informasi Akademik (SIMAKAD). SIMAKAD memiliki peran penting sebagai alat bantu dalam mengelola data internal dan sebagai pusat informasi bagi mahasiswa. Penelitian ini dilakukan untuk mengevaluasi tingkat penerimaan model UTAUT2 serta dampak dari variabel utama dan variabel tujuan dalam model UTAUT2.**

Penelitian ini menggunakan metode kuantitatif dengan sampel 150 responden, dianalisis menggunakan *software SMART PLS 3.0.* Temuan penelitian menunjukkan bahwa tingkat penerimaan model UTAUT2 mencapai 74%, menunjukkan tingkat adopsi yang tinggi. Variabel Perceived Value (p-Value: 0,019) dan Habit (p-Value: 0,009) memiliki pengaruh signifikan terhadap Behavioral Intention, dengan nilai p-Value < 0,05, sehingga hipotesisnya dapat diterima. Di sisi lain, variabel Performance Expectancy (p-Value: 0,660), Effort Expectancy (p-Value: 0,417), Social Influence (p-Value: 0,652), dan Facilitating Conditions (p. - value: 0,292) Tidak memiliki pengaruh yang berarti pada Behavioral Intention, yang mengindikasikan bahwa hipotesisnya tidak dapat diterima. Selain itu, variabel Hedonic Motivation (p-Value: 0,978) juga tidak memiliki dampak yang signifikan pada Behavioral Intention. Namun, variabel Facilitating Conditions (p-Value: 0,000) dan Behavioral Intention (p-Value: 0,000) memiliki pengaruh positif terhadap Use Behaviour, sehingga hipotesisnya dapat diterima. Sebaliknya, variabel Habit (p-Value: 0,915) tidak memiliki pengaruh signifikan terhadap Use Behaviour, sehingga hipotesisnya tidak dapat diterima.

Kata Kunci: Niat perilaku, perilaku menggunakan, SIMAKAD, UTAUT, UTAUT2

I. PENDAHULUAN

Teknologi informasi maupun sistem informasi telah mengalami perkembangan yang sangat maju dan memberikan dampak yang signifikan pada berbagai bidang. Dalam konteks

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

bisnis dan operasional organisasi yang semakin kompleks, serta dukungan dalam menghadapi perkembangan, organisasi perlu perencanaan dan pengembangan sistem sehingga meningkatkan efektivitas dan efisiensi operasional mereka [1]. Penggunaan teknologi informasi (TI) di institusi pendidikan, seperti perguruan tinggi, merupakan suatu kebutuhan mutlak dengan tujuan meningkatkan kualitas pendidikan [2].

Salah satu aspek penggunaan TI di perguruan tinggi adalah melalui Sistem Informasi Akademik (SIMAKAD), yang sangat mendukung kegiatan belajar-mengajar dan proses administratif. Contohnya, Institut Ilmu Kesehatan dan Teknologi Muhammadiyah Palembang (IKesT MP) telah mengadopsi SIMAKAD untuk meningkatkan mutu dalam bidang Pendidikan sebagai upaya untuk melayani mahasiswa dengan pelayanan terbaik. Namun, untuk menjalankan sistem ini secara efisien dan efektif, diperlukan perencanaan, pengaturan pelaksanaan, dan analisis sebelum mengevaluasi simakad [3].

SIMAKAD berperan sebagai alat pendukung untuk mengelola data baik secara internal dan menjadi pusat layanan informasi terintegrasi antara dosen dan mahasiswa . Sistem ini dilengkapi dengan berbagai fitur yang mendukung administrasi akademik, termasuk KRS, KHS, transkrip, pemantauan perkuliahan, pemantauan, dan pembayaran berbasis online. Meskipun sistem telah diimplementasikan, penting untuk menilai sejauh mana penggunaan sistem ini telah berjalan dengan baik oleh Civitas Akademika IKesT MP.

Untuk mengoptimalkan biaya yang telah dikeluarkan, sangat penting dilakukan analisis terhadap penerimaan sistem informasi akademik di IKesT MP. Hingga saat sistem informasi akademik IKesT MP, sama sekali belum dilakukan analisis dan evaluasi. SIMAKAD IKesT MP penting untuk dilakukan analisis dan evaluasi agar sesuai dengan ekspektasi pengguna. Selain itu dapat membantu dalam menentukan apakah tujuan penggunaan simakad telah terlaksana atau masih perlu perbaikan. Kelemahan yang ada dalam simakad belum bisa diidentifikasi tanpa analisis yang menyeluruh dan hambatan-hambatan yang muncul dalam memenuhi kebutuhan pengguna juga perlu diidentifikasi untuk memahami bagaimana pengguna menerima dan berinteraksi dengan sistem informasi akademik tersebut. Oleh karena itu, diperlukan pendekatan empiris dengan menggunakan model UTAUT2 yang dapat digunakan untuk menganalisis dan mengevaluasi simakad IKest MP.

II. KAJIAN TEORI

A. Informasi Akademik berbasis Web-SIMAK

Simakad pada umumnya dikembangkan dengan mengadopsi model berbasis web, yang berarti sistem ini dapat diakses melalui internet secara onlineamaupun jaringan intranet lokal. Selain itu, sistem ini memiliki kemampuan terintegrasi dengan berbagai sistem seperti SMS Akademik, E-learning, Sistem Informasi Kepegawaian dan Sistem Informasi Perpustakaan [4].

Website juga menjadi salah satu fasilitas layanan yang dimiliki oleh institusi pendidikan sebagai media informasi berbasis online [1]. Selain berfungsi sebagai sarana penyedia informasi, website juga digunakan sebagai alat penunjang kegiatan administrasi akademik yang terintegrasi dalam suatu sistem informasi akademik. Upaya untuk mencapai kemudahan dan efisiensi dalam administrasi akademik memerlukan manajemen layanan yang efektif. Tujuan strategis dapat tercapai apabila strategi yang telah direncanakan, dibuat, diterapkan, dan ada pengelolaan yang baik. untuk mencapainya adalah dengan mengembangkan sistem informasi akademik[5].

Menurut [6], Simakad merupakan sistem yang dirancang mengelola data-data akademik dengan tujuan agar aktivitas administrasi dalam perkuliahan mahasiswa berjalan lancar, sehingga mahasiswa mudah dalam menjalankan tugas-tugas administrasi akademik mereka.

B. UTAUT 2

UTAUT (Unified Theory of Acceptance and Use of Technology) merupakan suatu model terpadu yang dirumuskan dengan tujuan menggabungkan berbagai teori dan penelitian yang berhubunga dengan penerimaan suatu teknologi informasi [7]. Model UTAUT ini mengintegrasikan elemen-elemen dari sembilan teori dan model yang berbeda, termasuk TAM, TRA, TPB, MM, MPCU , C-TAM-TPB, IDT, SCT dan Usability [8]. Venkatesh dan rekan-rekannya di tahun 2012, mengembangkan suatu teori penerimaan teknologi yang dikenal sebagai Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). Teori UTAUT2 berfokus pemahaman diterimanya teknologi informasi serta mengukur sejauh mana variabel-variabel dalam UTAUT2 memengaruhi penggunaan teknologi informasi [9]. Model UTAUT2 pernah dilakukan penelitian terhadap pelanggan dengan digantinya variabel utama, yaitu Price Value diganti dengan Perceived Value di tahun 2019, Meskipun ada penelitian serupa dalam konteks konsumen, penelitian dengan pendekatan yang serupa, dibidang akademik belum dilakukan oleh peneliti sebelumnya [10].

Adapun variabel utama UTAUT 2 adalah sebagai berikut [7].

1. Espektasi Kinerja *(Performance Expectancy)*

Ekspektasi kinerja adalah merupakan tingkat keyakinan pengguna bahwa dalam pemanfaatan suatu sistem informasi bisa membantu mereka dalam meningkatkan kualitas dan efisiensi dalam melaksanakan pekerjaan mereka.. Ekspektasi kinerja adalah salah satu faktor penentu terpenting dalam adopsi sistem TI. Lima konstruk yang termasuk dalam ekspektasi kinerja dari model sebelumnya adalah motivasi ekstrinsik pada MM, job fit 14 pada MPCU, persepsi kegunaan pada TAM + TPB dan TAM/TAM2, ekspektasi hasil pada SC dan keunggulan relatif pada IDT.

2. Espektasi Usaha *(Effort Expectancy)*

Upaya yang diharapkan didefinisikan sebagai suatu tingkat kemudahan seseorang untuk menggunakan suatu sistem. Sudah ada tiga konstruksi yang diturunkan dari model sebelumnya

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

yang mengadopsi ekspektasi ini. Ketiga struktur ini dipertimbangkan kemudahan penggunaan di TAM/TAM2, kompleksitas di MPCU, dan di IDT

3. Pengaruh Sosial *(Social Influence)*

Ekspektasi usaha mengacu pada tingkat persepsi pengguna terkait dengan mudahannya pemakaian di sistem. Apabila sistem dianggap lebih mudah dipakai, Suatu usaha yang diperlukan untuk menggunakannya cenderung lebih rendah. Sebaliknya, diperlukan suatu usaha yang lebih besar jika sistemnya sulit digunakan. Konsep ini tercermin dalam tiga konstruk yang berasal dari model-model sebelumnya, yaitu persepsi kemudahan penggunaan dalam IDT dan TAM/TAM2, kompleksitas dalam MPCU.

4. Kondisi Fasilitas *(Facilitating Conditions)*

Kondisi fasilitas (*facilitating conditions*) dapat diartikan sebagai tingkat keyakinan seseorang terhadap ketersediaan infrastruktur organisasional dan teknis menyokong suatu sistem. Definisi ini sejalan dengan ide yang mirip yang dapat ditemukan dalam kerangka kerja pengendalian persepsi perilaku (perceived behavioral control) dalam TAM+TPB, TPB/DTPB, faktor-faktor yang memungkinkan dalam MPCU dan kesesuaian. (compatibility) dalam IDT. Setiap kerangka kerja ini digunakan dengan metode serupa untuk menggambarkan bagian-bagian dari lingkungan teknologi atau organisasi yang telah dirancang untuk mengatasi hambatan-hambatan dalam penggunaan suatu sistem

5. Motivasi Kesenangan *(Hedonic Motivation)*

Merupakan tingkat kesenangan yang dirasakan seseorang dari penggunaan teknologi tertentu, dan telah terbukti menjadi faktor penting dalam menentukan penggunaan maupun penerimaan teknologi. Konsep ini pertama kali dikemukakan *oleh Brown dan Venkatesh pada tahun 2005*, sebagaimana dijelaskan dalam referensi [7] .

6. Nilai Harga *(Price Value)*

Harga nilai (Price Value) merupakan tingkat perbandingan yang dipersepsikan oleh seseorang antara hasil atau manfaat yang diperoleh dari suatu produk atau layanan dengan harga penggunaan teknologi dengan biaya yang dikeluarkan. Nilai harga dianggap positif ketika seseorang percaya bahwa manfaat yang diperoleh dari teknologi tersebut melebihi biayanya, dan nilai harga yang positif ini berdampak positif pada niat untuk menggunakan teknologi tersebut. Komponen biaya dan persepsi nilai harga bisa berdampak besar pada konsumen untuk memakai teknologi, seperti yang dijelaskan dalam referensi [7] .

7. Kebiasaan *(Habit)*

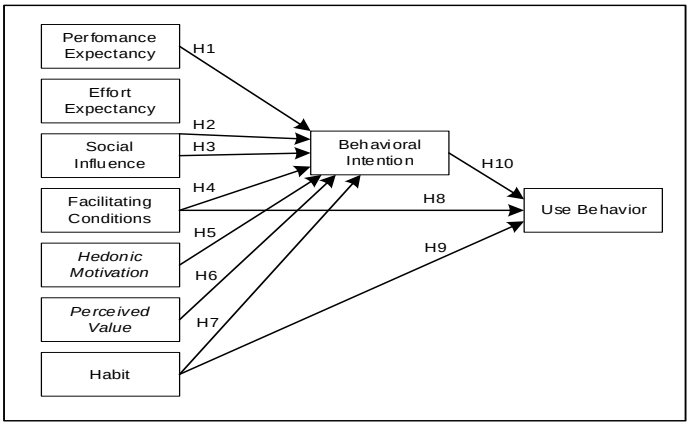
Menurut [11] dalam [7], kebiasaan diartikan seseorang lebih cenderung melakukan suatu perilaku dengan cara terus menerus sebagai hasil dari pembelajaran, sedangkan menurut [12] dalam [7] , asimilasi kebiasaan membiasakan diri dengan otomatisme. Walaupun konsepnya mirip, kebiasaan dioperasionalisasikan dalam dua cara yang berbeda. Pertama, kebiasaan dianggap suatu perilaku yang sudah terjadi dahulunya. Adapun Kedua kebiasaan dilaukan pengukuran berdasarkan sejauh mana orang itu yakin bahwa perilakunya terjadi secara otomatis.

Variabel *Price Value* Dihilangkan

Untuk Variabel *Price Value* tidak digunakan pada penelitian ini, karena Model UTAUT2 umumnya digunakan dalam konteks konsumen, di mana variabel tersebut difokuskan ke sistem maupun teknologi yang dipakai oleh konsumen dan biaya yang terkait. Namun, variabel Price Value kurang relevan dalam konteks penerimaan SIMAKAD karena beberapa penelitian sebelumnya [13] menunjukkan bahwa hubungan antara variabel Price Value dan Behavioral Intention tidak memiliki pengaruh signifikan. Oleh karena itu, dalam penelitian ini, variabel Price Value dihilangkan karena difokuskan pada penerimaan sistem informasi dalam konteks pendidikan akademik, dan aspek biaya tidak dipertimbangkan.

Penambahan Variabel *Perceived Value*

Peneliti memutuskan memakai model UTAUT2 dengan mengganti Variabel Price Value menjadi Perceived Value dalam studi kasus SIMAKAD IKesT MP dengan mengabaikan jenis kelamin responden dan umur responden dari penelitian ini. Pengguna SIMAKAD IKesT MP dalam penelitian ini merupakan mahasiswa berstatus aktif [9]. Hubungan antar variabel yang akan digunakan oleh peneliti dapat dilihat pada Gambar seperti yang dijelaskan berikut ini.



Gambar 1. UTAUT2 dikembangkan dengan menggunakan *Perceived Value*

Penjelasan Mengenai Variabel Tujuan pada UTAUT2 :

1. *Behavioral Intention*

Behavioral Intention ataupun niat perilaku pengguna dapat diartikan suatu kemaunan seseorang dalam menggunakan teknologi informasi sesuai dengan keinginannya. Kemauan dalam memanfaatkan suatu sistem adalah niat pengguna untuk terus memakai sistem tersebut secara terus menerus, dengan dan mereka memiliki pengaksesan ke sistem yang digunakan.

2. *Use Behavior*

Use Behavior atau perilaku suatu penggunaan mengacu pada sejauh mana frekuensi seseorang memakai teknologi informasi. Pengguna akan menggunakan teknologi informasi jika mereka memiliki minat dalam memanfaatkan sistem informasi tersebut, dikarenakan mereka yakin penggunaan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

sistem kinerja pekerjaan mereka akan meningkat. Dengan kata lain, Use Behavior mencerminkan sejauh mana tingkat keterlibatan pengguna dalam menggunakan teknologi informasi dalam konteks pekerjaan atau tugas yang mereka lakukan.

III. METODELOGI PENELITIAN.

A. Jenis Penelitian

Pada penelitian ini, digunakan pendekatan kuantitatif melalui pelaksanaan survei sebagai metode penelitian. Metode survei dalam penelitian kuantitatif digunakan untuk mengumpulkan data dari lokasi yang alami (bukan dibuat-buat), namun peneliti melakukan intervensi dalam proses pengumpulan data, seperti distribusi kuesioner, uji wawancara terstruktur, dan lain-lain.

B. Populasi Responden dan Sampel Responden

Populasi terdiri dari mahasiswa berstatus aktif di Institut Ilmu Kesehatan & Teknologi Muhammadiyah Palembang (IKesT MP) angkatan 2019-2022, yang mencakup dua fakultas yaitu Fakultas Ilmu Kesehatan dan Fakultas Sains dan Teknologi. Adapun jumlah responden sebanyak 150 responden pada penelitian in.

C. Sumber Data

Data primer penelitian ini dikumpulkan dengan menyebarkan kuesioner kepada responden. Kuesioner dikirimkan secara digital melalui link Google Form.

D. Instrumen Penelitian

Instrumen penelitian mengadopsi kuesioner SIA (Sistem Informasi Akademik) yang telah dikembangkan oleh peneliti sebelumnya . Dalam sampel penelitian. Sebelumnya, kuesioner telah menjalani uji validitas dan reliabilitas oleh peneliti.

Hasil dari uji validitas tersebut bahwa 29 item kuesioner, semuanya telah terbukti valid. Selanjutnya, responden juga menjalani uji reliabilitas untuk menilai sejauh mana kuesioner tersebut dapat diandalkan sebelum disebarkan kepada seluruh responden. Hasil dari uji reliabilitas penelitian tersebut nilai alpha Cronbach 0,917, memilik tingkat reliabilitas yang layak untuk digunakan karena nilainya sangat tinggi dalam penelitian ini. Skala pengukuran yang digunakan dalam kuesioner adalah skala Likert dengan 5 poin [14].

E. Hipotesis penelitian

Hipotesis ini dirumuskan berdasarkan prinsip-prinsip yang ada di model UTAUT 2. Adapun hipotesis sebagai berikut [7]. Adapun Uji Hipotesis dapat digambarkan :

H1 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Performance Expectancy dan hasilnya positif

H2 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Effort dan hasilnya positif

H3 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Social Influence dan hasilnya positif

H4 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Facilitating Conditions dan hasilnya positif

H5 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Hedonic Motivation dan hasilnya positif

H6 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Perceived Value dan hasilnya postif.

H7 : Variabel Behavioral Intention akan dipengaruhi oleh Variabel Habit dan hasilnya positif.

H8 : Variabel Use Behavior akan dipengaruhi oleh Variabel Facilitating Condition dan hasilnya positif

H9 : Variabel Use Behavior akan dipengaruhi oleh Variabel Habit dan hasilnya positif

H10 :Variabel Use Behavior akan dipengaruhi oleh Variabel Behavioral Intention dan hasolnya positif

F. Tehnik Analisa Data

Model Persamaan Struktural—Least Square Parsial (SEM-PLS) adalah suatu pendekatan statistik multivariat yang didasarkan pada perbedaan variabel dependen yang berganda dengan variabel independen yang berganda. Penganalisisan dalam SEM-PLS melibatkan dua tahap kunci, yaitu penilaian terhadap model pengukuran (Measurement Model) atau outer model, dan model struktural (Structural Model) atau inner model [15]. Sofwarw analisis data menggunakan *SMART PLS 3.0*.

Tabel 1. Nilai *Outer Loading* Pada Model Pengukuran Awal

Variabel.	Performance Expectancy.	Effort. Expectancy.	Social. Influence.	Facilitating. Conditions.	Hedonic. Motivation.	Perceived. Value.	Habit..	Behavioral. Intention.	Use. Behavior.	Ket
PE1	0.872	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
PE2	0.910	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
PE3	0.902	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
PE4	0.859	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
EE1	0.000	0.919	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
EE2	0.000	0.905	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
EE3	0.000	0.940	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
EE4	0.000	0.929	0.000	0.000	0.000	0.000	0.000	0.000	0.000	V
SI1	0.000	0.000	0.826	0.000	0.000	0.000	0.000	0.000	0.000	V
SI2	0.000	0.000	0.930	0.000	0.000	0.000	0.000	0.000	0.000	V
SI3	0.000	0.000	0.914	0.000	0.000	0.000	0.000	0.000	0.000	V
FC1	0.000	0.000	0.000	0.910	0.000	0.000	0.000	0.000	0.000	V
FC2	0.000	0.000	0.000	0.932	0.000	0.000	0.000	0.000	0.000	V
FC3	0.000	0.000	0.000	0.923	0.000	0.000	0.000	0.000	0.000	V
HM1	0.000	0.000	0.000	0.000	0.914	0.000	0.000	0.000	0.000	V
HM2	0.000	0.000	0.000	0.000	0.957	0.000	0.000	0.000	0.000	V
HM3	0.000	0.000	0.000	0.000	0.841	0.000	0.000	0.000	0.000	V
PV1	0.000	0.000	0.000	0.000	0.000	0.928	0.000	0.000	0.000	V
PV2	0.000	0.000	0.000	0.000	0.000	0.945	0.000	0.000	0.000	V
PV3	0.000	0.000	0.000	0.000	0.000	0.901	0.000	0.000	0.000	V
HI1	0.000	0.000	0.000	0.000	0.000	0.000	0.925	0.000	0.000	V
HI2	0.000	0.000	0.000	0.000	0.000	0.000	0.866	0.000	0.000	V
HI3	0.000	0.000	0.000	0.000	0.000	0.000	0.917	0.000	0.000	V
BI1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.931	0.000	V
BI2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.952	0.000	V
BI3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.910	0.000	V
UB1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.931	V

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

Variabel.	Performance Expectancy.	Effort. Expectancy.	Social. Influence.	Facilitating. Conditions.	Hedonic. Motivation.	Perceived. Value.	Habit..	Behavioral. Intention.	Use. Behavior.	Ket
UB2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.903	V
UB3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.880	V

Hasil uji validitas peneliti menunjukkan *: Outer loading* pada variabel *performance expectancy* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator valid. Nilai *outer loading* pada variabel *effort expectancy* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *effort expectancy* semua valid. Nilai *outer loading* pada variabel *social influence* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada *social influence* semua valid. Nilai *outer loading* pada variabel *facilitating conditions* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *facilitating conditions* semua valid. Nilai *outer loading* pada variabel *hedonic motivation* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *hedonic motivation* semua valid. Nilai *outer loading* pada variabel *perceived value* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *perceived value* semua valid.

Nilai *outer loading* pada variabel *habit* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *habit* semua valid. Nilai *outer loading* pada variabel *behavioral intention* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *behavioral intention* semua valid. Nilai *outer loading* pada variabel *use behavior* semua indikator bernilai lebih dari 0,7, sehingga dapat dinyatakan semua indikator pada variabel *use behavior* semua valid. Jadi semua variable yang ada nilainya lebih dari 0,7 dan kesemuanya variabelnya dinyatakan valid dan tidak perlu dilakukan revisi pengukuran jadi hanya 1x pengukuran

Tabel 2. Nilai Reliabilitas

Variabel	Cronbach's Alpha	Rho A	Composite Reliability	(AVE)	Ket
Performance Expectancy	0.909	0.910	0.936	0.785	Reliabel
Effort Expectancy	0.942	0.943	0.959	0.852	Reliabel
Social Influence	0.871	0.907	0.920	0.794	Reliabel
Facilitating Conditions	0.912	0.914	0.944	0.850	Reliabel
Hedonic Motivation	0.888	0.901	0.931	0.819	Reliabel
Perceived Value	0.915	0.915	0.947	0.856	Reliabel
Habit	0.887	0.903	0.930	0.816	Reliabel
Behavioral Intention	0.923	0.923	0.951	0.867	Reliabel
Use Behavior	0.890	0.894	0.931	0.819	Reliabel

Hasil Uji Reliability : Keseluruhan variable baik , Effort Expectancy, Performance Expectancy, Facilitating Conditions, Social Influence, Hedonic Motivation, Habit, Perceived Value, Use Behavior, Behavioral Intention, dari segi indikator baik itu *Cronbach's Alpha, , Composite Reliability*, *rho_A* menunjukan angka > 0.70 sehingga semua dinyatakan valid. Seluruh variabel pada kolom *Average Variance Extracted (AVE),* dinyatakan Reliabel

IV. HASIL DAN PEMBAHASAN

1. Hasil analisis deskriptif variabel UTAUT2 terhadap perilaku penggunaan SIMAKAD.

Berdasarkan hasil analisis deskriptif secara menyeluruh pada variabel UTAUT2 terhadap perilaku penggunaan SIMAKAD mahasiswa dapat dinyatakan Dimensi yang terukur paling tinggi adalah pada *Use Behavior* (bernilai 89,87%), sementara dimensi yang terendah (bernilai 84,27%) terukur pada *Social Influence*. Hasil rekapitulasi pengukuran dimensi Variabel Penerimaan SIMAKAD pada UTAUT2 berkategori dengan Penelian total termasuk"Tinggi" dan ditampilkan pada tabel dibawah ini

Tabel 2. Rekapitulasi Tanggapan Responden Mahasiswa Terhadap Penggunaan Simakad

Dimensi Variable Penerimaan Teknologi SIMAKAD Pada UTAUT2		Jlh Skor	% Capaian	Kategori
PE	Performance Expectancy	2694	89,80%	Tinggi
EE	Effort Expectancy	2654	88,47%	Tinggi
SI	Social Influence	1896	84,27%	Tinggi
FC	Facilitating Conditions	1970	87,56%	Tinggi
HM	Hedonic Motivation	1954	86,84%	Tinggi
PV	Perceived Value	2008	89,24%	Tinggi
HI	Habit	2694	86,27%	Tinggi
BI	Behavioral Intention	2694	86,27%	Tinggi
UB	Use Behavior	2694	89,87%	Tinggi
Total Pengukuran Pada Variabel UTAUT2		21258	87,62%	Tinggi

2. Nilai *R-Square*

Adapun hasil R-square yang dihasilkan setelah menggunakan *bootstraping* pada Smart PLS ditunjukkan pada tabel 3. sebagai berikut :

Tabel 3. Nilai R – Square.

Variabel	R Square
Behavioral Intention	0,769
Use Behaviour	0,740

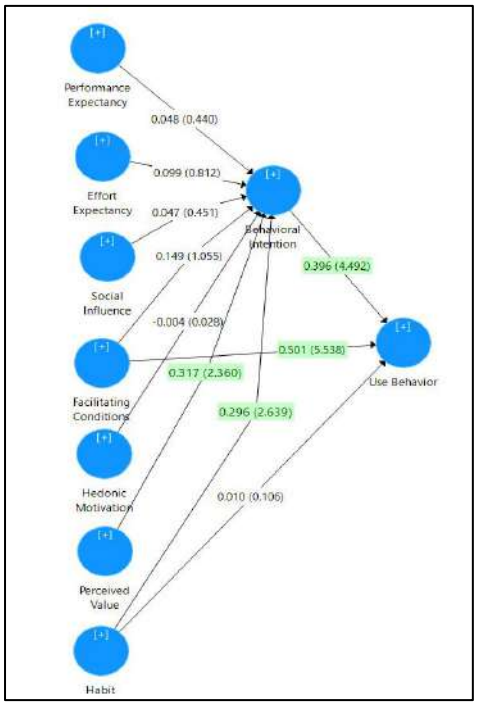
Tabel 3. menjelaskan bahwa nilai hasil *R-square* variabel Behavioral Intention (niat perilaku) 0,769. Berarti hasil presentase besarnya pengaruh *Effort Expectancy, , Facilitating Conditions, Performance Expectancy, Hedonic Motivation Social Influence, Habi, Perceived Value,* terhadap *behavioral Intention* sebesar 76,9 % dan faktor lain sebesar 23,1%.

Kemudian variabel *use behavior* (perilaku menggunakan) hasil *R-square* sebesar 0.740. Menunnjukan persentase besarnya tingkat suatu pengaruh variabel *Performance Expectancy, Social Influence, Effort Expectancy, Habit Facilitating Conditions, Perceived Value, Hedonic Motivation,* total nilai 74 % dan faktor lain sebesar 26%.

3. Uji Hipotesis

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

Pada pengujian hipotesis dengan menggunakan uji-t dan koefisien jalur *(path coefficient)*. Pada SmartPLS dilakukan *bootstraping*, maka didapatlah hasil pada Gambar 2. di bawah ini:



Gambar 2. Hasil Pengujian Hipotesis pada Model Struktural

Gambar 2 menunjukkan nilai t-value pada tanda kurung yang merupakan hasil pengujian dan nilai dari koefesien antar variabel. Hasil tersebut menunjukkan bahwa dari sepuluh faktor yang diamati, hanya empat factor yang hasilnya signifikan karena nilai t-valuenya lebih dari 1,96. Berikut ini adalah tabulasi hasil pengujian model persamaan struktural secara keseluruhan (*full model*) berdasarkan koefesien jalur atau hubungan antar variabel laten seperti pada Tabel 4. Sebagai berikut :

Tabel 4.
Nilai *path coeficients, t-statistics significance, p-value*

Hipotesis		*path coefficient*	t-hitung	*path coefficient*	p- Values	Ket
No	variabel					
H1	PE -> BI	0.048	0.440	Positif	0.660	Tidak Sig.
H2	EE -> BI	0.099	0.812	Positif	0.417	Tidak Sig.
H3	SI -> BI	0.047	0.451	Positif	0.652	Tidak Sig.
H4	FC -> BI	0.149	1.055	Positif	0.292	Tidak Sig.
H5	HM -> BH	-0.004	0.028	Negatif	0.978	Tidak Sig.
H6	PV -> BI	0.317	2.360	Positif	0.019	Signifikan
H7	HI -> BI	0.296	2.639	Positif	0.009	Signifikan
H8	FC -> UB	0.501	5.538	Positif	0.000	Signifikan
H9	HI -> UB	0.010	0.106	Positif	0.915	Tidak Sig.
H10	BI -> UB	0.396	4.492	Positif	0.000	Signifikan

Berdasarkan Tabel 4. terlihat bahwa dari sepuluh (10) hipotesis, hanya ada empat hipotesis yang dapat diterima yaitu Pengaruh Positif *Perceived Value* terhadap *Behavioral Intention*, Pengaruh Positif *Habit* terhadap *Behavioral Intention* Pengaruh Positif *Facilitating Conditions* terhadap *Use Behavior* dan terkahir Pengaruh Positif *Behavior Intention* terhadap *Use Behavior*. Adapun analisis suatu pengaruh variabel UTAUT 2 terhadap SIMAKAD :

1. Pengaruh Ekspektasi Kinerja (*Performance Expectancy)* terhadap Niat Perilaku (*Behavioral Intention)*

   Hasil penelitian menunjukkan terdapat pengaruh positif pada variabel *Performance Expectancy* (ekspektasi kinerja) terhadap variabel *Behavioral Intention* (niat perilaku) sebesar 0,048 (4,8%), namun secara statistik tidak signifikan karena diperoleh nilai t-hitung sebesar 0,440 (lebih kecil dari 1,96) atau *p-value* 0,660 (lebih besar dari 5%) dan hipotesis ditolak.

   Implikasi dari hasil tersebut adalah *Performance Expectancy* (ekspektasi kinerja) sistem SIMAKAD di IKest MP dapat membantu pengurusan adminstrasi mahasiswa dalam menyelesaikan administrasi akademik (positif) tetapi tidak signifikan terhadap *Behavioral Intention* (niat perilaku) penggunaan sistem SIMAKAD di IKesT MP karena pada kenyataannya SIMAKAD tidak digunakan mahasiswa secara teratur, hanya pada saat akan menyusun Kartu Rencana Studi (KRS) pada awal perkuliahan dan melihat nilai pada Kartu Hasil Studi (KHS) saja. Penelitian yang dilakukan [16] mengatakan tingkat kepuasan pada sistem informasi akademik dapat diperoleh dari Kinerja (*Performance)* sistem informasi akademik, Informasi Sistem, dan Keamanan Sistem informasi

2. Pengaruh Ekspektasi Usaha (*Effort Expectancy)* terhadap Niat Perilaku *(Behavioral Intention)*

   Hasil penelitian menunjukkan adanya pengaruh positif sebesar 9,9% dari variabel Effort Expectancy terhadap variabel Behavioral Intention. Namun, secara statistik, pengaruh ini tidak signifikan karena nilai t-hitung sebesar 0,812 (lebih kecil dari 1,96), dengan p-value sebesar 0,417 (lebih besar dari 5%), sehingga hipotesis ditolak.

   Hasil ini mengindikasikan bahwa meskipun SIMAKAD telah diberikan fitur-fitur dan kemudahan dalam penggunaannya, namun belum berhasil menaikan niat mahasiswa untuk terus menggunakan sistem tersebut

3. Pengaruh Pengaruh Sosial (*Social Influence)* terhadap Niat Perilaku (*Behavioral Intention)*

   Hasil penelitian menunjukkan terdapat terdapat pengaruh pengaruh positif sebesar 4,7% dari variabel Social Influence (pengaruh sosial) terhadap variabel Behavioral Intention (niat perilaku). Namun, secara statistik, pengaruh ini tidak signifikan karena nilai t-hitung sebesar 0,451 (lebih kecil dari 1,96), dengan p-value sebesar 0,652 (lebih besar dari 5%), sehingga hipotesis ditolak.

   Hasil ini mengindikasikan bahwa lingkungan sekitar mahasiswa di IKesT MP mungkin belum memiliki peran yang signifikan dalam memotivasi individu untuk menggunakan sistem SIMAKAD (Behavioral Intention).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

4. Pengaruh *Facilitating Conditions* (Kondisi Fasilitas) terhadap *Behavioral Intention* (Niat Perilaku)

Hasil penelitian menunjukkan adanya pengaruh positif sebesar 14,9% dari variabel Facilitating Conditions (Kondisi Fasilitas) terhadap variabel Behavioral Intention (niat perilaku). Namun, secara statistik, pengaruh ini tidak signifikan karena nilai t-hitung sebesar 1,055 (lebih kecil dari 1,96), dengan p-value sebesar 0,292 (lebih besar dari 5%), sehingga hipotesis ditolak.

Implikasi dari hasil ini adalah bahwa pengaruh kondisi fasilitas terhadap meningkatnya minat sebesar 14,9% dalam penggunaan SIMAKAD. Namun, dalam konteks SIMAKAD di IKesT MP, pengaruh yang diberikan tidak signifikan terhadap minat penggunaan SIMAKAD. Pada saat ini, ketersediaan akses ke layanan seperti internet serta sumber daya yang memadai seperti smartphone dan jaringan internet (Wi-Fi)[17], serta pengetahuan untuk menggunakannya telah menjadi kebutuhan umum. Oleh karena itu, mahasiswa dapat dengan mudah mengakses sistem, dan hal ini mempengaruhi minat penggunaan SIMAKAD.

5. Pengaruh *Hedonic Motivation* (motivasi hedonis/motivasi kesenangan) terhadap *Behavioral Intention* (niat perilaku)

Hasil penelitian menunjukkan bahwa variabel Hedonic Motivation (motivasi hedonis) tidak memiliki pengaruh negatif yang signifikan terhadap variabel Behavioral Intention (niat perilaku), dengan pengaruh yang sangat kecil sebesar -0,4%. Dari perhitungan statistik tidak ada hubungan yang signifikan antara keduanya karena hasil t-hitung sebesar 0,028 (lebih kecil dari 1,96) dan hasil p-value sebesar 0,978 (lebih besar dari 5%), sehingga hipotesis ditolak.

Dari hasil ini dapat disimpulkan bahwa fitur-fitur yang ada pada tampilan SIMAKAD cenderung lebih berfokus pada administrasi akademik daripada memberikan pengalaman hiburan kepada mahasiswa. Berbeda dengan aplikasi lain yang mungkin menawarkan hiburan seperti musik atau permainan, SIMAKAD dirancang khusus untuk membantu dalam proses administrasi mahasiswa

6. Pengaruh *Perceived value* (nilai keuntungan/manfaat) terhadap *Behavioral Intention* (niat perilaku)

Hasil penelitian menunjukkan terdapat pengaruh positif variabel *Perceived value* (nilai keuntungan/manfaat) terhadap variabel *Behavioral Intention* (niat perilaku) sebesar 0,317 (31,7%) dan secara statistik signifikan karena diperoleh nilai t-hitung sebesar 2,360 (lebih besar dari 1,96) atau *p-value* 0,019 (lebih kecil dari 5%), dan hipotesis diterima

Hal tersebut menunjukan bahwa SIMAKAD memberikan manfaat yang besar bagi mahasiswa dalam proses administrasi akademik Ini menciptakan niat untuk menggunakan SIMAKAD dalam mendukung proses perkuliahan.

7. Pengaruh *Habit* (kebiasaan) terhadap *Behavioral Intention* (niat perilaku)

Hasil penelitian menunjukkan terdapat pengaruh positif variabel *Habit* (kebiasaan) terhadap variabel *Behavioral Intention* (niat perilaku) sebesar 0,296 (29,6%) dan secara statistik signifikan karena diperoleh nilai t-hitung sebesar 2,639 (lebih besar dari 1,96) atau *p-value* 0,009 (lebih kecil dari 5%) dan hipotesis diterima.

Kebiasaan adalah salah satu indikator utama untuk memahami niat penggunaan suatu sistem [18]. Apabila pengguna (mahasiswa) mempunyai pengalaman akan suatu teknologi dan terus menggunakannya maka akan mengubah pengalaman menjadi kebiasaan [17]. Sejalan juga dengan hasil penelitian yang dilakukan oleh [19]. Pengguna yang terbiasa memakai simakad, semakin besar kemungkinan niat perilaku akan meningkatkan untuk terus memakai simakad. Mahasiswa menjadi lebih sering memakai simakad. mahasiswa pada awal semester mereka melakukan penggunaan sistem tersebut untuk mengakses SIMAKAD untuk pengurusan Kartu Rencana Studi (KRS) dan mengetahui perkembangan nilai semester nya melalui Kartu Hasil Studi (KHS) yang dapat dengan mudah di lihat di SIMAKAD melalui ponsel pribadi mahasiswa.

8. Pengaruh *Facilitating Conditions* (kondisi fasilitas) terhadap *Use Behaviour* (perilaku menggunakan)

Hasil penelitian variabel *Facilitating Conditions* (kondisi fasilitas) berpenaruh psotif terhadap variabel *Use Behaviour* (perilaku menggunakan) sebesar 0,501 (50,1%) dan secara statistik signifikan karena diperoleh nilai t-hitung sebesar 5,538 (lebih besar dari 1,96) atau *p-value* 0,000 (lebih kecil dari 5%), dan hipotesis diterima.

Hal tersebut menunjukkan bahwa dengan adanya SIMAKAD dapat digunakan mahasiwa dengan mudah karena adanya fasilitas dan sumber daya yang memadai yang disiapkan oleh institusi, seperti wifi yang dapat mendorong pengguna (mahasiswa) untuk selalu menggunakan SIMAKAD. Sejalan juga dengan penelitian [19] dimana hasil *t-hitung* pada variabel suatu kondisi fasilitas terhadap variabel perilaku dalam menggunakan simakad. Nilainya diatas t hitung 2.22 lebih besar dari 1,98 sehingga ada pengaruh dan terbukti signifikan.

Pengguna sistem informasi akademik telah mendapatkan fasilitas yang memadai sehingga dapat mahasisiwa dapat menggunakan sistem tersebut dengan mudah. Begitu juga dengan penelitian yang dilakukan oleh [20] menyatakan Semakin besar kesiapan pengguna terhadap kondisi yang mendukung, semakin tinggi kemungkinan mereka dapat melakukan hal tersebut, Faktor ini memiliki dampak yang signifikan pada perilaku penggunaan. Faktor-faktor yang memudahkan seperti ketersediaan sarana, layanan internet, dan kondisi lainnya yang memadai. dapat membuat pengguna menggunakan sistem tersebut secara berkelanjutan.

9. Pengaruh *Habit* (kebiasaan) terhadap *Use Behaviour* (perilaku menggunakan)

Hasil penelitian menunjukkan terdapat pengaruh positif variabel *Habit* (kebiasaan) terhadap variabel *Use Behaviour* (perilaku menggunakan) sebesar 0,010 (1%), namun secara statistik tidak signifikan karena diperoleh nilai t-hitung sebesar 0,106 (lebih kecil dari 1,96) atau *p-value* 0,915 (lebih besar dari 5%), dan hipotesis ditolak.

Hasil penelitian ini menunjukkan bahwa walaupun mahasiswa memiliki kebiasaan dalam menggunakan SIMAKAD di tiap semesternya, tetapi tidak menjadikan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

SIMAKAD menjadi hal utama karena setelah pengurusan administrasi berakhir mahasiswa tidak lagi mengakses SIMAKAD kecuali ada kondisi lain yang diperlukan. Hal ini sejalan dengan penelitian [19] yang menunjukkan hasil t hitung pada variabel suatu kebiasaan terhadap variabel perilaku menggunakan suatu sistem diatas t hitung 0.12 lebih kecil dari 1,98, hasilnya pengaruh yang ditunjukkan terbukti tidak signifikan.

Sistem informasi akademik sering digunakan oleh pengguna secara terus menerus, tetapi pengguna merasa bahwa setelah mereka mengambil KRS dan mengakses sistem tersebut, mereka tidak lagi mengautamakan simakad..

10. Pengaruh *Behavioral Intention* (niat perilaku) terhadap *use behaviour* (perilaku menggunakan)

Hasil penelitian menunjukkan terdapat pengaruh positif variabel *Behavioral Intention* terhadap *Use Behaviour* sebesar 0,396 (39,6%) dan secara statistik signifikan karena diperoleh nilai t-hitung sebesar 4,492 (lebih besar dari 1,96) atau *p-value* 0,000 (lebih kecil dari 5%), dan hipotesis diterima.

Hasil tersebut menunjukkan bahwa niat mahasiswa akan terus menggunakan dan bergantung pada SIMAKAD dalam pengurusan administrasi akademik. Berbagai fitur yang ada pada SIMAKAD banyak digunakan oleh mahasiswa dan SIMAKAD sering digunakan mahasiswa selama perkuliahan. Teknologi informasi akan digunakan jika pengguna memiliki minat dalam menggunakan sistem informasi tersebut, karena mereka meyakini bahwa penggunaan sistem tersebut dapat meningkatkan kinerja pekerjaan mereka. [7].

V. KESIMPULAN

1. Untuk melihat seberapa besar penerimaan Sistem Informasi Akademik (SIMAKAD) IKesT Muhammadiyah Palembang dengan menggunakan model UTAUT 2, dari nilai *R-square* artinya besar tingkat penerimaan model UTAUT 2, variabel *Use behaviour* (perilaku menggunakan) sebesar 74%.
2. Berdasarkan hasil uji statistic SEM-PLS,pengaruh variabel utama terhadap variabel tujuan (*Behavioral Intention* dan *Use Behaviour*) model UTAUT 2 pada Sistem Informasi Akademik (SIMAKAD) IKesT Muhammadiyah Palembang sebagai berikut :
   a. Untuk variabel *Perceived Value* (*p-Value* : 0,019) dan *Habit* (*p-Value* : 0,009) dimana nilai *p-Value* < 0,05 yang menunjukkan terdapat pengaruh terhadap variabel *Behavioral Intention* yang artinya hipotesis diterima. Sedangkan untuk variabel *Performance Expectancy* (*p-Value* : 0,660), *Effort Expectancy* (*p-Value*: 0,417), *Social Influence* (*p-Value* : 0,652), *Facilitating Conditions* (*p-Value* : 0,292), *Hedonic Motivation* (*p-Value* : 0,978), dimana nilai *p-Value* > 0,05 yang menunjukkan tidak terdapat pengaruh terhadap *Behavioral Intention,* yang artinya hipotesis ditolak

   b. Berdasarkan model UTAUT 2, variabel *Facilitating Conditions* (*p-Value* : 0,000) dan *Behavioral Intention* (*p-Value* : 0,000) terdapat pengaruh positif terhadap variabel *Use Behaviour* yang artinya hipotesis diterima, sedangkan variabel *Habit* (*p-Value* : 0,915) tidak terdapat pengaruh terhadap *Use Behaviour* yang artinya hipotesis ditolak.

Bagi Institusi Pendidikan penelitian ini diharapkan dapat meningkatkan pengembangan Sistem Informasi Akademik mahasiswa dengan meningkatkan performance expectancy dengan menambahkan fitur-fitur yang lebih menarik dan menyenangkan supaya niat perilaku dapat meningkat. Pada penelitian selanjutnya yang sejalan dengan topik dapat menambahkan variabel -variabel baru seperti perceived of risk, self efficacy dan variabel lainnya yang dapat meningkatkan pengembangan sistem informasi mahasiswa dan dosen.

DAFTAR PUSTAKA

[1] E. Fadilah, R., & Surya Negara, "Analisis Faktor Yang Mempengaruhi Penerimaan Pengguna Aplikasi Elektronik Renumerasi Kinerja (E-RK) Menggunakan Metode UTAUT dan SDT (Studi Kasus : Pemerintah Kabupaten Musi Rawas).," *J. Ilm. Matrik, 24(1), 40–50. https//doi.org/10.33557/jurnalmatrik.v24i1.1658*, 2022.

[2] E. Indrayani, "Management of Academic Information System (AIS) at Higher Education in the City of Bandung," *Procedia - Soc. Behav. Sci.*, vol. 103, pp. 628–636, 2013, doi: 10.1016/j.sbspro.2013.10.381.

[3] T. Sutabri, *Analisis Sistem Informasi*. Penerbit Andi, 2012.

[4] T. Sutabri, "Konsep Sistem Informasi," Penerbit Andi, 2012.

[5] T. Sutabri, "Analisis Layanan Tata Kelola Aplikasi Sistem Informasi Akademik dengan Menggunakan Cobit 5 pada STIK Bina Husada Analysis Of Governance Services Of Academic Information System Applications Using Cobit 5 On STIK Bina Husada," *J. Ilm. Bin. STMIK Bina Nusant. Jaya*, vol. 0, no. 01, pp. 2657–2117, 2023.

[6] I. Irawan, "Pengembangan Sistem Informasi Akademik Universitas Pahlawan Tuanku Tambusai Riau," *J. Teknol. Dan Open Source*, vol. 1, no. 2, pp. 55–66, 2018, doi: 10.36378/jtos.v1i2.21.

[7] J. Y. L. T. and X. X. Viswanath Venkatesh, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology.," *Mis Quarterly, Vol. 36, Pp. 157-178., 36(1), 157–178. https//doi.org/doi10.2307/41410412*, 2012.

[8] P. J. B. Tan, "Applying the UTAUT to Understand Factors Affecting the Use of English E-Learning Websites in Taiwan. SAGE Open, pp. 1-12.," 2013.

[9] J. J. C. Christiono, A. T., & Tambotoh, "Analisis Pemanfaatan Teknologi Informasi Menggunakan Pendekatan Unified Theory of Acceptance and Use of Technology 2 (Studi Kasus: Flexible Learning (F-Learn) UKSW).," *Konf. Nas. Sist. Inf. 2014, 2(Utaut 2), 1–7. https//repository.uksw.edu/handle/123456789/12359*, 2014.

[10] N. Shaw and K. Sergueeva, "The non-monetary benefits of mobile commerce: Extending UTAUT2 with perceived value," *Int. J. Inf. Manage.*, vol. 45, no. December 2017, pp. 44–55, 2019, doi: 10.1016/j.ijinfomgt.2018.10.024.

[11] E. al. Limayem, "Electronic Commerce Research and Applications. Electronic Commerce Research and Applications , 50.," 2007.

[12] S. Kim, S. S., Malhotra, N. K., & Narasimhan, "wo Competing Perspectives on Automatic Use: A Theoretical and Empirical Comparison, Information Systems Research 16(4): 418-432," 2005.

[13] and A. S. Tarhini, Ali, Mazen El-Masri, Maged Ali, ""Extending the UTAUT Model to Understand the Customers' Acceptance and Use of Internet Banking in Lebanon." Information Technology & People 29 (4): 830–49. https://doi.org/10.1108/itp-02-2014-0034," 2016.

[14] T. Hadiansyah, D., & Dirgahayu, "Evaluasi Sistem Informasi Akademik Universitas Mercu Buana Yogyakarta Menggunakan UTAUT2 Evaluation of Academic Information System of Universias Mercu Buana Yogyakarta Using UTAUT 2 Tanggal submisi : 04-09-2019 ; Tanggal penerimaan : 27-02-2020.," *J. Multimed. Artif. Intell. 4(1), 1–12. http//jmai.mercubuana-yogya.ac.id/index.php/jmai/article/view/108%0Ahttp//jmai.mercubuana-yogya.ac.id/index.php/jmai/article/download/108/34*, 2020.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

[15] W. A. and J. Hartono, "Partial Least Square (PLS) Alternatif Structural Equation Modeling (SEM) dalam," *Penelit. Bisnis. Yogyakarta Penerbit Andi*, 2015.
[16] R. Klarasati and T. Sutabri, "Analisis Pengukuran Tingkat Kepuasan Dosen Terhadap Sistem Informasi Akademik Menggunakan Metode Usability Pada Universitas Prabumulih," *J. Teknol. Dan Ilmu ...*, vol. 6, no. April, pp. 12–17, 2023.
[17] N. S. Desvira and M. F. Aransyah, "Analisis Faktor-Faktor yang Memengaruhi Minat dan Perilaku Penggunaan Fitur ShopeePay Menggunakan Model Unified Theory of Acceptance and Use of Technology (UTAUT2)," vol. 12, pp. 178–191, 2023.
[18] J. A. Charisma, "Analisis Minat Dan Perilaku Pengguna E-Wallet: Perluasan Utaut 2 Dengan Budaya Sebagai Moderasi (Studi Pada Mahasiswa di Kota Malang)," *http://etheses.uin-malang.ac.id/18709/1/16510088.pdf*, 2020.
[19] S. Anggraini, M. H. Irfani, and S. Rahayu, "Analisis Penerimaan Sistem Informasi Akademik Dengan Menggunakan UTAUT 2 (Studi Kasus: Akademi Keperawatan Pembina Palembang)," *Jusifo*, vol. 6, no. 1, pp. 15–30, 2020, doi: 10.19109/jusifo.v6i1.5616.
[20] M. F. Bhaskara, "Analisis Faktor-faktor yang Mempengaruhi Penerimaan Pengguna Dompet Elektronik Dana Menggunakan Pendekatan Unified Theory Of Acceptance and Use of Technology," 2022.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1978, Copyright ©2023
Submitted : September 28, 2023, Revised : October 9, 2023, Accepted : October 16, 2023, Published : November 7, 2023

Penentuan Penerima Beasiswa di STIT Prabumulih Menggunakan Metode AHP

Andi christian[1]*, Ariansyah[2], Anggie Sri Wahyuni[3]
Program Studi Sistem Informasi. Fakultas Ilmu Komputer[1], [2], [3]
Universitas Prabumulih
Jalan Patra No.50 Kel. Sukaraja Kec. Prabumulih Selatan Kota Prabumulih Sumatera Selatan
andichristian918@gmail.com[1], ayielubai@gmail.com[2], anggiewhyn@gmail.com[3]

***Abstract*— In every educational institution, especially universities, there are lots of scholarships offered to students. Likewise with the Prabumulih College of Engineering (STIT Prabumulih) which has a scholarship program for its students by applying predetermined rules or criteria, for example, parents' income, parents' dependents, student achievement index scores, etc. Due to this, not all scholarship recipients who apply for scholarships will receive a scholarship. The problem faced by the campus today is in the process of winning scholarships. therefore a decision support system is needed that can assist in providing scholarship recipient recommendations. In this study the authors used the AHP method and the Expert Choice application. From the calculation results obtained by the specified criteria, the GPA of 0.389 is the highest priority weight compared to other criteria. Then, from the results of calculating student data or all alternatives, the total value of each student is obtained. It can be concluded that the one who can be recommended to get a UKT scholarship is Student A because it has the highest score, namely 16.6% of the total calculated.**

Keywords*— AHP, *Scholarship, Expert Choice

***Abstrak*—Disetiap lembaga pendidikan khususnya perguruan tinggi banyak sekali beasiswa yang ditawarkan kepada mahasiswa. Demikian halnya dengan Sekolah Tinggi Ilmu Teknik Prabumulih (STIT Prabumulih) yang memiliki program beasiswa untuk mahasiswanya dengan menerapkan aturan-aturan atau kriteria-kriteria yang telah ditetapkan contohnya adalah penghasilan orang tua, tanggungan orang tua, nilai indeks prestasi akademik mahasiswa, dan lain-lain Karena hal ini, tidak semua calon penerima beasiswa yang mengajukan beasiswa akan mendapatkan beasiswa. Permasalahan yang di hadapi oleh pihak kampus saat ini adalah pada proses penentuan penetapan beasiswa. oleh karena itu diperlukannya sistem pendukung keputusan yang dapat membantu dalam memberikan rekomendasi penerima beasiswa. Dalam penelitian ini penulis menggunakan metode AHP dan aplikasi *Expert Choice.* Dari hasil perhitung didapatkan kriteria yang ditentukan, IPK sebesar 0.389 menjadi bobot prioritas tertinggi dibandingkan dengan kriteria lainnyakemudian, dari hasil perhitungan data mahasiswa atau semua alternatif, diperoleh nilai total masing-masing mahasiswa dapat disimpulkan bahwa yang dapat direkomendasikan mendapatkan beasiswa UKT adalah Mahasiswa A karena memiliki nilai yang paling tinggi yaitu 16,6 % dari total seluruh kriteria.**

Kata Kunci*—AHP, Beasiswa, *Expert Choice

I. PENDAHULUAN

Pendidikan adalah untuk mempersiapkan manusia dalam memecahkan problem kehidupan masa kini maupun masa yang akan datang[3]-[14]. Karena itu pentingnya pendidikan tercantum pada UUD 1945 nomor 20 tahun 2003 tentang sistem pendidikan nasional terdapat pada pasal 12 ayat (2) UUD 1945 yang berbunyi "Setiap warga Negara wajib mengikuti pendidikan dasar dan pemerintah wajib membiayai nya". Dari pasal tersebut dapat ditarik kesimpulan bahwa pembentukan memperoleh pendidikan yang layak pemerintah wajib membiayai warga Negaranya agar mendapatkan kehidupan yang layak dan lebih baik. Salah satu di antaranya melalui Program beasiswa[4].

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi ke berlangsung pendidikan yang ditempuh [5]-[8]. Di setiap lembaga pendidikan khususnya perguruan tinggi banyak sekali beasiswa yang diberikan untuk mahasiswa. Demikian halnya dengan Sekolah Tinggi Ilmu Teknik Prabumulih (STIT Prabumulih) yang memiliki program beasiswa untuk mahasiswa nya dengan menerapkan aturan-aturan atau kriteria-kriteria yang telah ditetapkan contohnya adalah penghasilan orang tua, tanggungan orang tua, nilai indeks prestasi akademik, dan lain-lain. Karena hal ini, calon penerima beasiswa yang mengajukan beasiswa tidak semua mendapatkan beasiswa. Permasalahan yang di hadapi oleh pihak kampus saat ini adalah pada proses penetapan beasiswa karena dalam proses pengumpulan datanya masih menggunakan cara konvensional sehingga sering terjadi kesalahan karena tidak adanya kriteria yang jelas untuk calon mahasiswa yang mendapatkan beasiswa. Sehingga dalam melakukan seleksi beasiswa tersebut tentu akan mengalami kesulitan karena banyak calon pelamar beasiswa dan banyaknya kriteria yang digunakan dalam menentukan keputusan penerima beasiswa yang sesuai dengan harapan, oleh karena itu diperlukannya sistem pendukung keputusan yang dapat membantu dalam memberikan rekomendasi penerima beasiswa.

Sistem pendukung keputusan adalah sistem berbasis komputer yang interaktif, yang membantu pengambil keputusan dalam menggunakan data dan model untuk menyeleksi masalah yang tidak ter struktur [1]-[12]. Sistem

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023

pendukung keputusan ini dapat membantu dalam memecahkan masalah yang ada, termasuk pada permasalahan yang ada pada STIT Prabumulih dalam seleksi penerima beasiswa dengan kriteria-kriteria yang telah ditetapkan, salah satu metode yang dapat membantu dalam sistem pendukung keputusan seleksi penerima beasiswa adalah metode *Analytical Hierarchy Process* (*AHP*) untuk membantu dalam penetapan kriteria-kriteria yang dapat digunakan dalam pengambilan keputusan.

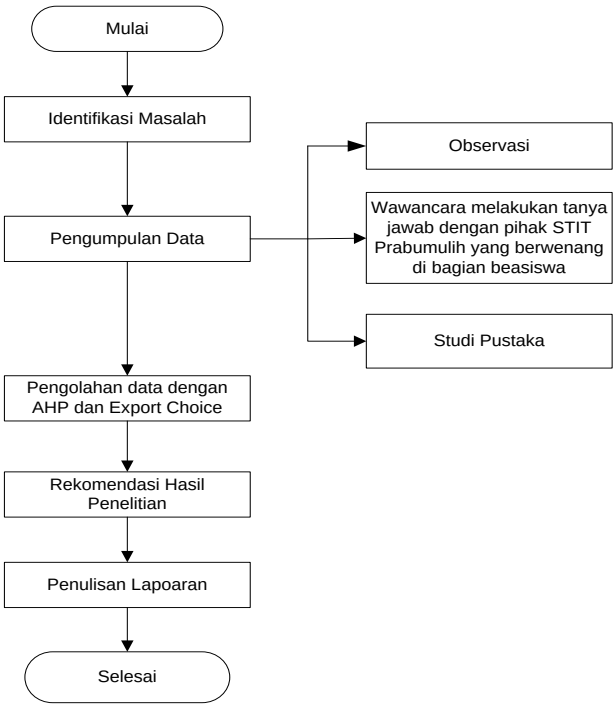
Metode (*AHP*) adalah suatu model pendukung keputusan untuk penyelesaian masalah *multi* kriteria yang dikembangkan oleh Thomas L. Saaty. Pada metode *Analytical Hierarchy Process* (*AHP*) dapat membantu dalam pembobotan terhadap kriteria-kriteria yang digunakan [6]-[9]. Penyelesaian yang dilakukan oleh metode *Analytical Hierarchy Process* (*AHP*) dalam seleksi penerima beasiswa di STIT Prabumulih adalah sebatas bobot tiap kriteria atau pun sub-kriteria. Bobot tersebut diperoleh dengan melakukan perbandingan berpasangan antar kriteria atau pun antar sub kriteria yang telah ditentukan oleh pihak STIT Prabumulih.

Aplikasi *Expert Choice* sangat bagus digunakan untuk penganalisis permasalahan dalam pengambilan keputusan dengan menggunakan metode *Analytical Hierarchy Process* (*AHP*) dapat menghasilkan alternatif yang banyak dan hierarki yang besar atau hierarki yang mempunyai banyak level, karena tidak perlu menghitung bobot secara manual, hingga tingkat kesalahan dalam perhitungan bobotnya sangat kecil, namun tergantung ketelitian kita dalam menginputkan data[7].

II. METODE PENELITIAN

Metode penelitian adalah cara ilmiah untuk mendapatkan data dengan tujuan dapat mendeskripsikan, dibuktikan, dikembangkan dan ditemukan pengetahuan, teori, untuk memahami, memecahkan, dan mengantisipasi masalah dalam kehidupan manusia[2]. Metode penelitian kualitatif juga merupakan metode penelitian yang juga menekankan pada aspek pemahaman secara mendalam terhadap suatu masalah dari pada melihat permasalahan untuk penelitian generalisasi. Metode ini lebih suka menggunakan teknik analisis mendalam, yaitu mengkaji masalah secara kasus perkasus karena metodologi kualitatif yakin bahwa sifat suatu masalah akan berbeda dengan sifat masalah lainnya.

Kerangaka pemikiran dibawah ini menggambarkan langkah-langkah pemikiran dalam penentuan penerimaan beasiswa di STIT Prabumulih



Gambar 1. Diagram Tahapan Penelitian

Uraian Penjelasan dari pada diagram tahap penelitian atas adalah sebagai berikut:

A. Indentifikasi Masalah

Mengidentifikasi permasalahan yang menjadi penentu pengambil keputusan bagi calon penerima beasiswa di STIT Prabumulih.

B. Teknik Pengumpulan Data

1. Observasi

Melakukan pengamatan (Observasi) secara langsung ke STIT Prabumulih dengan mencatat hal-hal yang berhubungan dengan penelitian ini.

2. Wawancara

Melakukan pengumpulan data dengan melakukan tanya jawab dengan pembantu ketua III dan bagian terkait pada STIT Prabumulih dengan menanyakan hal-hal yang berhubungan dengan penelitian ini.

3. Studi Pustaka

kegiatan pengumpulan data yang bersumber dari buku-buku, jurnal sebagai bahan referensi yang digunakan sebagai bahan acuan yang bertujuan untuk mendapatkan panduan yang di perlukan untuk menyelesaikan penelitian ini.

C. Pengolahan Data dengan metode AHP dan Expert Choice

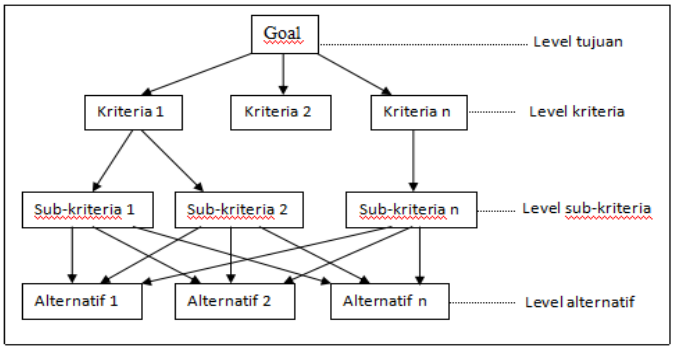
1. Metode Analytical Hierarchy Process (AHP)

Penyelesaian yang dilakukan oleh metode Metode Analytical Hierarchy Process (AHP) adalah sebatas mencari bobot tiap kriteria atau pun sub-kriteria. Bobot tersebut diperoleh dengan melakukan perbandingan berpasangan atar kriteria antar sub kriteria[14]-[15]. Melakukan perbandingan kriteria secara berpasangan atau hanya 2 kriteria pada waktu

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023

yang bersamaan jauh lebih mudah bila dibandingkan dengan melakukan perbandingan semua kriteria sekaligus. Umumnya, bentuk kuesioner untuk perbandingan berpasangan antar kriteria menggunakan skala 1 – 9 untuk menunjukkan tingkat kepentingan tiap kriteria [1]-[11].

Model pendukung keputusan ini akan menguraikan masalah multi kriteria yang kompleks dalam suatu struktur multi level di mana level pertama menjadi tujuan, yang diikuti level kriteria, sub kriteria, dan seterusnya ke bawah hingga level terakhir dari alternatif. Struktur hierarki tampak seperti pada Gambar 2



Gambar 2. Arsitektur Hierarki Pada AHP

2. Penggunaan Aplikasi *Expert Choice*

Dalam melakukan penelitian ini, peneliti menggunakan sofware sebagai alat bantu dalam pengambilan keputusan untuk melakukan perbandingan kriteria-kriteria penerima beasiswa. Sofware atau aplikasi yang digunakan dalam penelitian ini adalah aplikasi *Expert Choice*. Tujuannya untuk membandingkan dan membuktikan analisis perhitungan penulis sesuai dengan aplikasi *Expert Choice* yang sudah teruji kehandalannya [10].

D. Rekomendasi Hasil Penelitian

Tahap selanjutnya adalah merekomendasikan hasil penelitian agar dapat menjadi bahan pengambilan keputusan dalam penetuan penerimaan Beasiswa pada STIT Prabumulih.

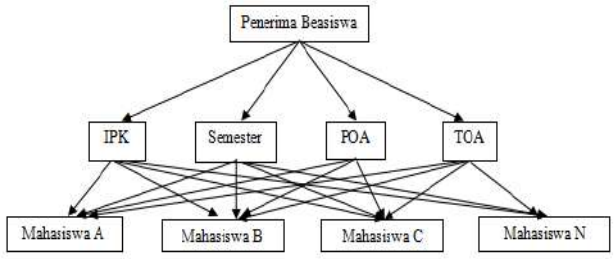
E. Penulisan Laporan

Tahap akhir dari metodologi penelitian ini adalah membuatkan laporan hasil penelitian.

III. HASIL DAN PEMBAHASAN

A. Struktur *Analytical Hierarchy Process*

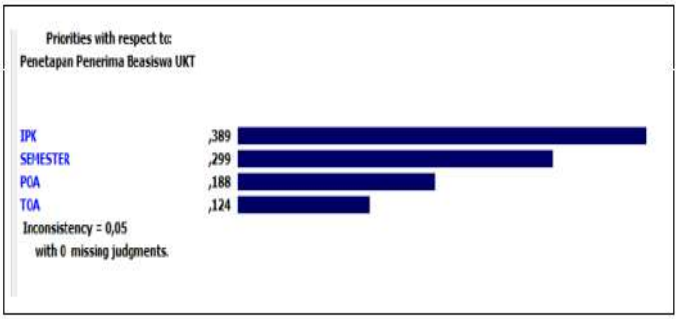
Proses analisis ini dinamakan hierarki, yang terdiri dari Goal, kriteria, dan alternatif. Goal atau tujuan pada hierarki ini adalah sistem pendukung keputusan penerimaan beasiswa UKT, sedangkan kriteria-nya adalah, nilai indeks prestasi akademik, jumlah semester, penghasilan orang tua, dan tanggungan orang tua. Dana alternatif terdiri dari 12 mahasiswa. Kriteria dan alternatif didapat dari hasil wawancara peneliti dengan pembantu ketua III, Dosen dan Mahasiswa di STIT Prabumulih sehingga menghasilkan informasi yang diperlukan dalam penelitian ini. Berikut struktur hierarki AHP sistem pendukung keputusan penerima beasiswa



Gambar 3. Struktur Hierarki Penerimaan Beasiswa UKT

B. Hasil Pengolahan Data *Analytical Hierarchy Process*

Hasil pengolahan dan analisis data penelitian ini menghasilkan bahwa dari empat faktor (kriteria) yang menjadi kriteria seleksi penerima beasiswa di STIT Prabumulih masih dilakukan secara konvensional, kriteria yang memiliki pengaruh paling besar (bobot prioritas tertinggi) berdasarkan analisis dengan metode AHP yaitu kriteria Indeks Prestasi Akademik (IPK) mempunyai nilai tertinggi yaitu sebesar 0.389, kemudian diurutkan kedua ada kriteria semester sebesar 0.299, pada urutan ketiga dengan bobot 0.188 pada kriteria penghasilan orang tua, dan yang terakhir pada kriteria tanggungan orang tua sebesar 0,124. Hasil analisis dapat dilihat pada gambar 3 di bawah ini:



Gambar 4. Grafik Normalisasi Matriks Antar Kriteria

Grafik Normalisasi Matriks Antar Kriteria Dalam gambar 3, hasil pengolahan dan analisis data penelitian ini menghasilkan bahwa dari empat faktor (kriteria) yang menjadi kriteria seleksi penerima beasiswa di STIT Prabumulih masih dilakukan secara konvensional, kriteria yang memiliki pengaruh paling besar (bobot prioritas tertinggi) berdasarkan analisis dengan metode AHP yaitu kriteria Indeks Prestasi Akademik (IPK) mempunyai nilai tertinggi yaitu sebesar 0.389, kemudian diurutkan kedua ada kriteria semester sebesar 0.299, pada urutan ketiga dengan bobot 0.188 pada kriteria penghasilan orang tua, dan yang terakhir pada kriteria tanggungan orang tua sebesar 0,124.

Tabel 1. Matriks Perbandingan Berpasangan

Kriteria	IPK	Semester	Poa	Toa
IPK	1	2	2	2
Semester	0,5	1	2	3
Poa	0,5	0,5	1	2
Toa	0,5	0,33	0,5	1

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023

Tabel grafik normalisasi matriks antar kriteria merupakan hasil dari tabel 1 Matriks Perbandingan Berpasangan, *Pairwase comparison* matriks antar kriteria. hal ini dibuktikan dengan hitungan manual di bawah ini.

Tabel 2. Matriks Normalisasi Antar Kriteria

Kriteria	IPK	Semester	Poa	Toa	Jumlah	Prioritas	%
IPK	0,4	0,25	0,364	0,25	1,534	0,389	38%
Semester	0,2	0,26	0,364	0,375	1,199	0,299	29%
Poa	0,2	0,13	0,182	0,25	0,762	0,188	19%
Toa	0,2	0,09	0,09	0,125	0,505	0,124	12%
Jumlah	1	1	1	1	4	1	100%

Total nilai prioritas digunakan untuk mendapatkan nilai konsistensitasnya seperti tabel 2 sehingga mendapatkan nilai :

Tabel 3. Nilai Prioritas

Kriteria	Prioritas
Ipk	0,389
Semester	0,299
Poa	0,188
Toa	0,124

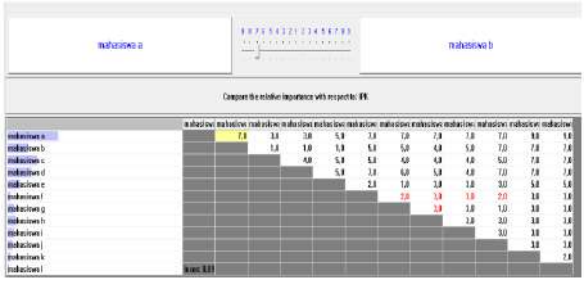
Untuk mendapatkan nilai konsistensi-nya dilakukan perhitungan ini digunakan untuk memastikan bahwa nilai rasio konsistensi (CR) <= 0.1. Jika ternyata nilai CR lebih dari 0.1, maka matriks perbandingan berpasangan harus diperbaiki. Untuk menghitung rasio konsistensi menggunakan rumus sebagai berikut :

$$\lambda maks = \text{Jumlah Matriks Perbandingan Berpasangan} \times \text{Prioritas}$$
$$= (2{,}5 \times 0{,}389) + (3{,}83 \times 0{,}299) + (5{,}5 \times 0{,}188) + (8 \times 0{,}124)$$
$$= 4{,}14367$$
$$CI = \frac{(\lambda maks - n)}{(n-1)} = \frac{4{,}14367-4}{4-1} = \frac{0{,}14367}{3} = 0{,}04789$$
$$IR = 0{,}90$$
$$CR = \frac{CI}{IR} = \frac{0{,}04789}{0{,}90} = 0{,}053 \longrightarrow 0{,}05 \text{ (konsisten)}$$

Oleh karena CR < 0.1, maka rasio konsistensi dari perhitungan tersebut bisa diterima.

1. Analisis Pengolahan Data Alternatif IPK.

Setiap alternatif mendapatkan nilai-nilai pembobotan berdasarkan kriteria indeks prestasi akademik (IPK). Data yang sudah diperoleh kemudian dimasukkan ke dalam aplikasi *expert choice* guna untuk menghitung data yang sudah masuk sehingga mendapatkan *inconsistensi* yang diinginkan dan sesuai dengan aturan perhitungan AHP.



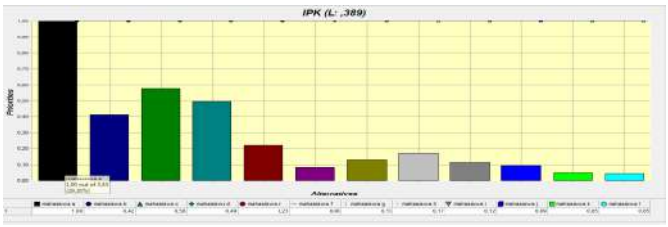
Gambar 5. Pairwise Numerical Comparisons

Dari data di atas hasil yang diperoleh dapat dilihat pada tabel 4 dibawah ini :

Tabel 4. Pairwise Numerical Comparisons

Mahasiswa	Bobot	Persentase
A	0,292	29%
B	0,122	12%
C	0,168	16%
D	0,144	14%
E	0,066	6%
F	0,024	2%
G	0,039	3%
H	0,051	5%
I	0,035	3%
J	0,027	2%
K	0,016	1%
L	0,014	1%

Hasil analisis dapat juga dilihat pada gambar 5



Gambar 6. Grafik Normalisasi Matriks Alternatif IPK

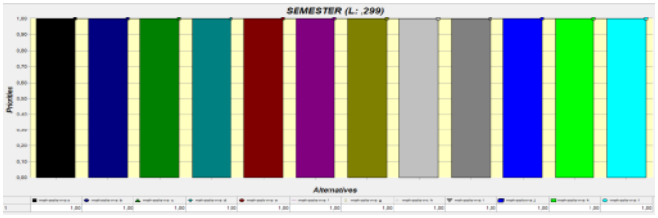
2. Analisis Pengolahan Data Alternatif Semester.

Setiap alternatif mendapatkan nilai-nilai pembobotan berdasarkan kriteria semester. Data yang sudah diperoleh kemudian dimasukkan ke dalam aplikasi expert choice guna untuk menghitung data yang sudah masuk sehingga mendapatkan inconsistensi yang diinginkan dan sesuai dengan aturan perhitungan AHP.

Gambar 7. Pairwise Numerical Comparisons

Pada kriteria semester data yang didapat oleh penulis adalah setiap mahasiswa yang mengikuti seleksi penerimaan beasiswa memiliki nilai yang sama dikarenakan semua mahasiswa yang mengikuti seleksi penerima beasiswa UKT berada pada semester yang sama yaitu semester 4. Oleh karena itu data yang didapat dan diolah memiliki bobot yang sama antar mahasiswa yaitu 0,083 atau sebesar 8%. Hasil analisis dapat dilihat pada gambar 7.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023

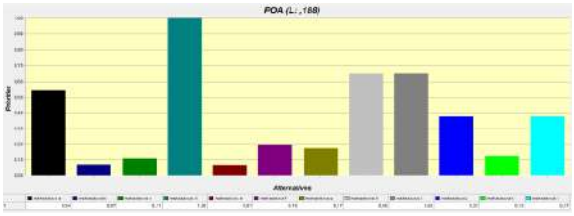


Gambar 8. Grafik Normalisasi Matriks Alternatif Semester

3. Analisis Pengolahan Data Alternatif Penghasilan Orang Tua. Setiap alternatif mendapatkan nilai-nilai pembobotan berdasarkan kriteria Penghasilan Orang Tua. Data yang sudah diperoleh kemudian dimasukkan ke dalam aplikas*i expert choice* guna untuk menghitung data yang sudah masuk sehingga mendapatkan inconsistensi yang diinginkan dan sesuai dengan aturan perhitungan AHP.

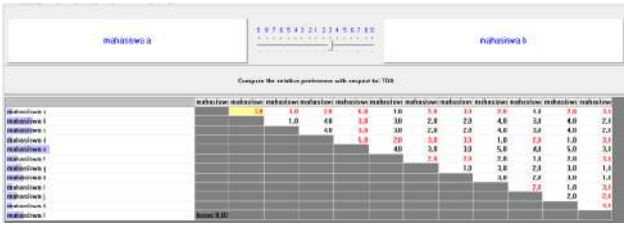
Gambar 9. Pairwise Numerical Comparisons

Dari data tersebut dapat disimpulkan bahwa Mahasiswa A mendapatkan bobot 0,125 atau sebesar 12%, Mahasiswa B mendapatkan bobot 0,016 atau sebesar 1%, Mahasiswa C mendapatkan bobot 0,025 atau sebesar 2%, Mahasiswa D mendapatkan bobot 0,231 atau sebesar 23%, Mahasiswa E mendapatkan bobot 0,016 atau sebesar 1%, Mahasiswa F mendapatkan bobot 0,044 atau sebesar 4%, Mahasiswa G mendapatkan bobot 0,040 atau sebesar 4%, Mahasiswa H mendapatkan bobot 0,150 atau sebesar 15%, Mahasiswa I mendapatkan 0,150 atau sebesar 15%, Mahasiswa J mendapatkan bobot 0,086 atau sebesar 8%, Mahasiswa K mendapatkan bobot 0,030 atau sebesar 3%, Mahasiswa L mendapatkan bobot 0,086 atau sebesar 8%. Hasil analisis dapat dilihat pada gambar 9.



Gambar 10. Grafik Normalisasi Matriks Alternatif Penghasilan Orang Tua

4. Analisis Pengolahan Data Alternatif Tanggungan Orang Tua. Data yang sudah diperoleh kemudian dimasukkan ke dalam aplikasi *expert choice* guna untuk menghitung data yang sudah masuk sehingga mendapatkan *inconsistensi* yang diinginkan dan sesuai dengan aturan perhitungan AHP.



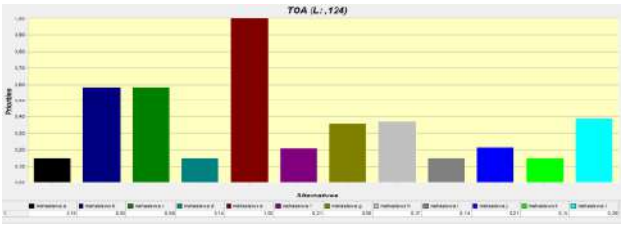
Gambar 11. Pairwise Numerical Comparisons

Dari data di atas hasil yang diperoleh dapat dilihat pada tabel 5 dibawah ini :

Tabel 5. Pairwise Numerical Comparisons

Mahasiswa	Bobot	Persentase
A	0,034	3%
B	0,135	13%
C	0,135	13%
D	0,034	3%
E	0,234	23%
F	0,048	4%
G	0,084	8%
H	0,087	8%
I	0,034	3%
J	0,050	5%
K	0,034	3%
L	0,091	9%

Hasil analisis dapat dilihat pada gambar 11.

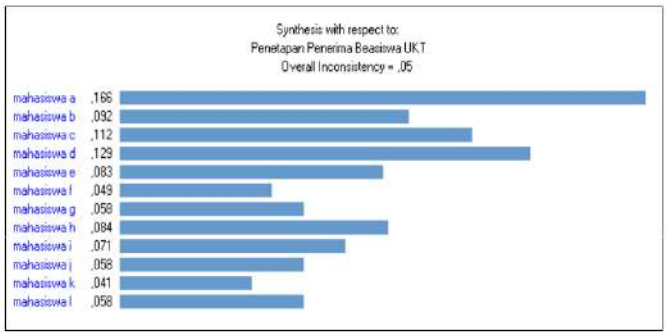


Gambar 12. Grafik Normalisasi Matriks Antar Alternatif Berdasarkan Tanggungan Orang Tua

3.1. Perhitungan Hasil Pengolahan Data *Analytical Hierarchy Process.*

Setelah mendapatkan nilai masing-masing dari setiap pembobotan kriteria dan setiap alternatif berdasarkan kriteria. Hasil penjumlahan ini merupakan hasil akhir dari penilaian kinerja mahasiswa. Mahasiswa A mendapatkan nilai 0,166, Mahasiswa B mendapatkan nilai 0,092, Mahasiswa C mendapatkan nilai 0,112, Mahasiswa D mendapatkan nilai 0,129, Mahasiswa E mendapatkan nilai 0,083, Mahasiswa F mendapatkan nilai 0,049, Mahasiswa G mendapatkan nilai 0,053, Mahasiswa H mendapatkan nilai 0,084, Mahasiswa I mendapatkan 0,071, Mahasiswa J mendapatkan nilai 0,058, Mahasiswa K mendapatkan bobot 0,041, Mahasiswa L mendapatkan nilai 0,058. Hasil seperti gambar 12. berikut ini:

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023



Gambar 13. Hasil Synthesis With Respect

Hal ini menujukan bahwa Mahasiswa A lebih unggul dari alternatif lainnya dalam seleksi penerima beasiswa UKT. Hasil perhitungan ini menujukan juga bahwa Mahasiswa A lebih memenuhi kriteria yang telah ditentukan.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan untuk mengukur prioritas kriteria dari empat perspektif/ kriteria yaitu indeks prestasi akademik, semester, penghasilan orang tua, tanggungan orang tua, bobot prioritas tertinggi berdasarkan hasil analisis dengan metode *AHP* yaitu didapatkan kriteria yang ditentukan, IPK sebesar 0.389 menjadi bobot prioritas tertinggi dibandingkan dengan kriteria lainnya. kemudian, dari hasil perhitungan data mahasiswa atau semua alternatif, diperoleh nilai total masing-masing mahasiswa dapat disimpulkan bahwa yang dapat direkomendasi-kan mendapatkan beasiswa UKT adalah Mahasiswa A karena memiliki nilai yang paling tinggi yaitu 16,6 % dari total seluruh kriteria.

REFERENCES

[1] Ari Basuki, S.T, M.T, Andharini Dwi Cahyani, S. Kom, M.Kom., *Sistem pendukung keputusan*. Yogyakata: Deepublish, 2016.

[2] Sugiyono, Metode Penlitian Kuantitatif, Kualitatif R & D, Sutopo. Bandung: ALFABETA, cv, 2019.

[3] Djumali dkk., Landasan Pendidikan. Yogyakara: Gava Media, 2014.

[4] Depdiknas, Udang-Undang RI Nomor 20 Tahun 2003, tentang Sistem Pendidikan Nasional di unduh dari https://peraturan.bpk.go.id/Home/Details/43920/uu-no-20-tahun-2003

[5] Izhar Salim, Fatmawati Fatmawati Sablan Tusti, "Analisis Implementasi Program Beasiswa *Credit Union* Keling Kumang Pada Anggota Tingkat SMA," *JPPK: Journal of Equatorial Education and Learning*, vol. 8, no. 6, pp. 1-9, 2019.

[6] Frieyadie, "Penerapan Metode AHP Sebagai Pendukungkeputusan Penetapan Beasiswa," *Jurnal Pilar Nusa Mandiri*, vol. 13, no. 1, pp. 49-58, 2017.

[7] Dian Hartanti, Karina Djunaidi Rahma Farah Ningrum, "Sistem Pendukung Keputusan Promosi Jabatan Struktural Dosen Menggunakan AHP (Analytical Hierarchy Process)," *Jurnal Teknik: Universitas Muhammadiyah Tangerang*, vol. 6, no. 2, pp. 62-71, 2017.

[8] Johannes Kristoffel Santie, "Implementasi Kebijakan Program Bantuan Besiswa Bidikmisidi Politeknik Negeri Manado," MAP (Jurnal Manajemen dan Administrsi Publik), vol. 1, no. 2, pp. 183-192, 2018.

[9] Yustina Meisella Kristania, Rousyati, Dany Pratmanto, Sopian Aji, "Sistem Penunjang Keputusan Pemilihan Penerima Beasiswa DenganMetode Analytical Hierarchy Process (AHP) Di SMKEra Informatika Tangerang Selatan," *Indonesian Journal on Software Engineering (IJSE)*, vol. 7, no. 2, pp. 212-219, 2021.

[10] Handayani, Rani Irma, "Pemanfaatan Aplikasi Expert Choice Sebagai Alat Bantu Dalam Pengambilan Keputusan (Studi Kasus: PT. BIT Teknologi Nusantara)," Journal Of Computing And Information System, Vol. Xi, No. 1, Pp. 53-59, 2015.

[11] Nurina Yasin, "Pemanfaatan Aplikasi Expert Choice Penanganan Gangguan Operasioal Kereta Api Akibat Genangan Air," Ug Jurnal, Vol. 14, No. 5, Pp. 04-07, 2020.

[12] Heni Ayu Septilia, Styawati, "Sistem Pendukung Keputusan Pemberian Dana Bantuan Menggunakan Metode Ahp," Jurnal Teknologi Dan Sistem Informasi (Jtsi), Vol. 1, No. 2, Pp. 34-41, 2020.

[13] Erliyan Redy Susanto, Ajeng Savitri Puspaningrum, Neneng, "Rancang Bangunrekomendasi Penerima Bantuan Sosial Berdasarkan Data Kesejahteraan Rakyat," *Jurnal Tekno Kompak*, Vol. 15, No. 1, Pp. 1-12, 2021.

[14] Jadiaman Parhusip, "Penerapan Metode Analytical Hierarchy Process (Ahp) Pada Desain Sistem Pendukung Keputusan Pemilihan Calon Penerima Bantuan Pangan Non Tunai (Bpnt) Di Kota Palangka Raya," *Jurnal Teknologi Informasi*, Vol. 13, No. 2, Pp. 18-29, 2019.

[15] Hermanda Ihut Tua Simamora, "Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Metode Analytical Hierarchy Process (AHP) Pada SMA Pencawan Medan," *Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD*, vol. 2, no. 1, pp. 19-25, 2019.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1717, Copyright ©2023
Submitted : May 14, 2023, Revised : June 24, 2023, Accepted : September 8, 2023, Published : November 8, 2023

Identifikasi PenipuanKartu Kredit Pada Transaksi Ilegal Menggunakan Algoritma Random Forest dan Decision Tree

Indah Werdiningsih[1], Endah Purwanti[2], Gede Rangga Wira Aditya[3], Auliya Rakhman Hidayat[4], R. Sulthan Rafi Athallah[5], Virda Adisty Sahar[6], Tio Satrio Wibisono[7], Darren Febriand Nura Somba[8]
Sistem Informasi, Fakultas Sains dan Teknologi Universitas Airlangga
indah-w@fst.unair.ac.id[1], endahpurwanti@fst.unair.ac.id[2], gede.rangga.wira-2020@fst.unair.ac.id[3], auliya.rakhman.hidayat-2020@fst.unair.ac.id[4] , raden.sulthan.rafi-2020@fst.unair.ac.id[5] , virda.adisty.sahar-2020@fst.unair.ac.id[6], tio.satrio.wibisono-2020@fst.unair.ac.id[7] , darren.febriand.nura-2020@fst.unair.ac.id[8]

Abstract*— The use of credit cards is increasing in today's digital era. This increase has resulted in many cases of fraud which have had a negative impact on credit card owners. To overcome this, many financial institutions have developed credit card fraud detection systems that can identify suspicious transactions. This study uses a classification method, namely random forest and decision tree to identify illegal transactions using a credit card, which then compares the results and attempts to create a model that can be useful for detecting fraud using a credit card that is more accurate and effective. The result of this study is that the accuracy provided by the Decision Tree Classifier is 0.98, while the accuracy provided by the Random Forest Classification is also 0.975. The conclusion obtained that the decision tree has a higher level of accuracy compared to the Random Forest Classification Algorithm, which is 98%. On the other hand, the Random Forest classification algorithm has a slightly lower level of accuracy compared to the Decision Tree classification algorithm, with an accuracy rate of 97.5%

***Keywords*—Credit Card, Classification, Decision Tree, RandomForest**

***Abstrak*— Penggunaan kartu kredit semakin meningkat dalam era digital saat ini. Peningkatan tersebut berdampak padabanyaknya kasus penipuan yang berdampak buruk bagi pemilik kartu kredit. Untuk mengatasinya, banyak lembaga keuangan telah mengembangkan sistem deteksi penipuan menggunakan kartu kredit yang dapat mengidentifikasi transaksi yang mencurigakan. Penelitian ini menggunakan metode klasifikasiyaitu *random forest* dan *decision tree* dalam mengidentifikasi transaksi ilegal menggunakan kartu kredit yang kemudian dibandingkan hasilnya dan berusaha untuk membuat model untuk mendeteksi penipuan menggunakan kartu kredit yang lebih akurat dan efektif. Hasil dari penelitian ini adalah akurasi yang diberikan oleh *Decision Tree Classifier* sebesar 0.98, sementara akurasi yang diberikan oleh *Random Forest Classification* juga sebesar 0.975. Kesimpulan yang didapat bahwa decision tree memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan Algoritma Klasifikasi Random Forest yaitu 98%. Kemudian, algoritma klasifikasi Random Forest memiliki tingkat akurasi yang sedikit lebih rendah dengan algoritma klasifikasi Decision Tree dengan tingkat akurasi 97,5%.**

Kata Kunci*—Kartu Kredit, Klasifikasi, *Decision Tree*, *Random Forest

I. PENDAHULUAN

Dalam era digital saat ini, penggunaan kartu kredit semakin meluas dan telah menjadi salah satu metode pembayaran yang paling populer di seluruh dunia. Kredit biasanya digunakan untuk merujuk pada transaksi keuangan elektronik yang dilakukan tanpa penggunaan kas fisik [1]. Namun, peningkatan penggunaan kartu kredit juga diikuti dengan peningkatan kasus penipuan yang menggunakan kartu kredit. Penipuan ini dapat berdampak buruk bagi pemilik kartu kredit dan lembaga keuangan yang terkait sehingga perlu adanya upaya untuk mengidentifikasi dan mencegah transaksi ilegal yang menggunakan kartu kredit [2].

Untuk mengatasi masalah ini, banyak lembaga keuangan telah mengembangkan sistem deteksi penipuan menggunakan kartu kredit yang dapat membantu mengidentifikasi transaksi yang mencurigakan atau ilegal. Namun, sebuah transaksi tidak dapat secara murni diklasifikasikan sebagai penipuan atau secara asli pada sistem, mereka hanya mencari kemiripan dan kemungkinan transaksi menjadi penipuan berdasarkan studi ekstensif perilaku pelanggan, kebiasaan belanja mereka dan juga menganalisis penipuan yang dilakukan sebelumnya, yang kemudian akan diamati pola mereka [3]. Selain itu, masalah yang sering muncul adalah kurangnya akurasi dalam mengklasifikasikan transaksi sebagai legal atau ilegal sehingga memungkinkan terjadinya kesalahan dalam menolak transaksi yang sebenarnya legal atau menerima transaksi yang seharusnya ditolak.

Penyelesaian masalah terkait klasifikasi/prediksi menggunakan *machine learning* saat ini menjadi acuan yang

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

dapat diandalkan karena dirasa dapat memberikan akurasi terkait permasalahan [4]. *Machine learning* adalah subjek yang mempelajari cara menggunakan komputer untuk mensimulasikan peningkatan diri komputer untuk memperoleh pengetahuan dan keterampilan baru, mengidentifikasi pengetahuan yang ada, dan terus meningkatkan kinerja dan pencapaian [5].

Penerapan yang telah diimplementasikan dalam berbagai aplikasi, machine learning atau pembelajaran mesin telah mencakup hampir semua aspek dan wilayah ilmiah, yang dimana telah membawa dampak besar pada ilmu pengetahuan dan masyarakat [6]. Ada hubungan yang erat antara istilah machine learning dan data mining, jadi seringkali dalam literatur ilmiah, metode machine learning adalah disebut metode data mining [7].

Ada beberapa metode pengklasifikasian yang digunakan dalam machine learning untuk melakukan prediksi, diantaranya adalah Decision Tree dan Random Forest. Decision tree learning adalah algoritma induktif tipikal berdasarkan contoh, yang berfokus pada aturan klasifikasi yang ditampilkan sebagai pohon keputusan yang disimpulkan dari sekelompok gangguan dan contoh tidak teratur [8]. Di antara metode-metode data mining lainnya, klasifikasi Decision Tree memiliki berbagai keunggulan diantaranya sederhana untuk dipahami, mudah untuk diterapkan, membutuhkan sedikit pengetahuan, mampu menangani data numerik dan kategorikal, tangguh, dan dapat menangani dataset yang besar [9]. Sementara itu, klasifikasi Random Forest adalah pengklasifikasi ansambel yang menghasilkan banyak pohon keputusan, menggunakan subset sampel dan variabel pelatihan yang dipilih secara acak [10]. Klasifikasi Random Forest memiliki beberapa keunggulan. Pertama, algoritma ini dapat meningkatkan akurasi ketika terdapat data yang hilang dan juga dapat mengatasi pencilan (outliers). Selain itu, klasifikasi Random Forest efisien dalam penyimpanan data. Selain itu, algoritma ini juga melibatkan proses seleksi fitur, yang memungkinkan untuk memilih fitur-fitur terbaik dan meningkatkan performa model klasifikasi. Dengan adanya fitur seleksi ini, klasifikasi Random Forest dapat bekerja secara efektif pada data yang besar dengan parameter yang kompleks. Selain itu, Random Forest juga mampu bekerja secara paralel melalui metode multiple random forest. Namun, terkadang klasifikasi Random Forest dapat menghasilkan nilai yang tidak diharapkan dan tidak memprediksi rentang nilai respons pada data latih [11].

Dalam konteks ini, penelitian ini bertujuan untuk menentukan metode klasifikasi yang dapat membantu dalam mengidentifikasi transaksi ilegal yang menggunakan kartu kredit secara akurat. Metode klasifikasi yang akan dikembangkan akan didasarkan pada algoritma Machine Learning. Penelitian ini akan dilakukan dengan menggunakan dataset credit card fraud bersumber pada Kaggle yang mencakup informasi seperti jarak dari rumah tempat transaksi, jarak dari transaksi terakhir, rasio transaksi rata-rata harga beli terhadap harga pembelian, repeat retailer, transaksi menggunakan chip (kartu kredit), transaksi menggunakan nomor PIN, online order, dan fraud transaksi.

Penelitian sebelumnya oleh Dhwanir Shah dan Lokesh Kumar Sharma [12], dilakukan pendekatan untuk mendeteksi penggunaan kartu kredit palsu menggunakan algoritma Decision Tree dan Random Forest. Penelitian ini menggunakan metode Parameter Tuning sebelum dan sesudah dari model yang dipilih, yang dimana model yang dimaksud disini adalah Decision Tree dan Random Forest. Akurasi yang didapatkan oleh Decision Tree sebelum Parameter Tuning sebesar 99.05% untuk Train Data dan 57,64% untuk Test Data, sedangkan setelah dijalankan Parameter Tuning didapatkan akurasi Train Data sebesar 96,12% dan 74,76% untuk Test Data. Kemudian pada model Random Forest, didapatkan train data sebesar 87,87% dan 63,27% sebelum dijalankannya parameter tuning. Setelah dijalankan parameter tuning, didapatkan hasil sebesar 53,61% pada train data dan 26,16% pada test data. Dengan begitu, didapatkan kesimpulan bahwa akurasi yang didapatkan oleh Decision Tree lebih besar daripada Random Forest.

Kemudian terdapat juga penelitian lainnya oleh T. Tulasi Bhavani, M. Kameswara Rao, dan A. Manohar Reddy [21], dilakukan pendekatan untuk mendeteksi kejadian serangan jaringan menggunakan algoritma Decision Tree dan Random Forest. Dari penelitian ini diperoleh hasil bahwa dengan menggunakan algoritma Decision Tree, akurasi yang didapatkan menunjukkan angka 81,87%. Sedangkan dengan menggunakan algoritma Random Forest, didapatkan akurasi senilai 95,32%. Dengan ini dapat disimpulkan bahwa pada penelitian ini, algoritma Random Forest lebih efektif dalam menjalankan tugas regresi dan klasifikasi.

Kedua penelitian sebelum tidak terdapat seleksi fitur. Oleh karena itu, penelitian ini akan menggunakan attribute selection dan algoritma decision tree dan random forest untuk mengidentifikasi transaksi ilegal menggunakan kartu kredit. Attribute Selection menggunakan feature importance dari Decision tree dan Random Forest. Disamping itu, dataset yang digunakan pada penelitian ini berbeda dengan penelitian sebelumnya. Hasil akurasi dari kedua algoritma tersebut akan dibandingkan akurasinya dan berusaha untuk membuat model yang dapat berguna untuk mendeteksi penipuan menggunakan kartu kredit yang lebih akurat dan efektif. Hasil penelitian diharapkan dapat memberikan kontribusi pada perkembangan teknologi Machine Learning dalam bidang keamanan finansial menggunakan algoritma klasifikasi terbaik yang didapatkan dari penelitian.

II. METODOLOGI PENELITIAN

Penelitian ini terdapat 6 tahapan, yaitu pengumpulan data, attribute selection, data preprocessing, split data, klasifikasi, dan evaluasi. Klasifikasi menggunakan Decision Tree dan Random Forest. Tahapan penelitian dapat dilihat pada Figure 1.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

A. *Pengumpulan Data*

Data yang digunakan dalam penelitian ini adalah data public yang diambil dari dari website

https://www.kaggle.com/datasets/dhanushnarayananr/credit-card-fraud

Dataset tersebut berupa data transaksi kartu kredit illegal. Dataset terdiri dari 8 atribut, terdiri dari 7 atribut input dan 1 atribut class. 7 atribut input adalah Distance_ from_home, distance_from_last_transaction, ratio_to_median_purchase _price, repeat_retailer, used_chip, used_pin_number, online_order. 1 atribut kelas adalah fraud. Deskripsi atribut ditunjukkan pada Table 2.

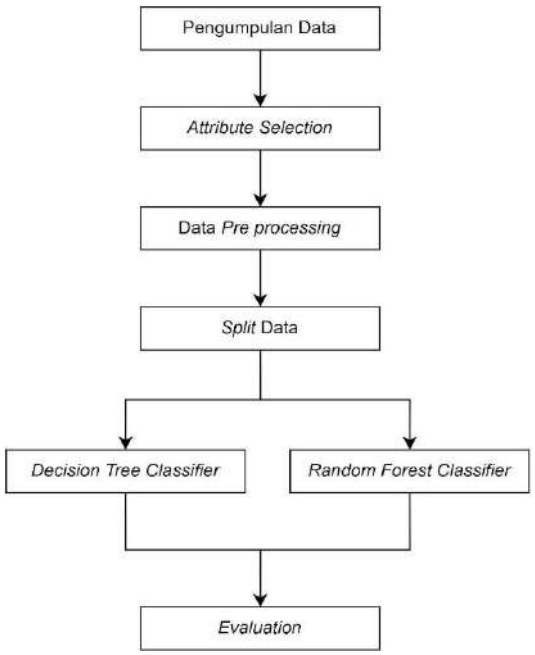


Fig 1. Tahapan Penelitian

TABLE 1. Dataset yang digunakan

	Distance_from_home	distance_from_last_transaction	ratio_to_median_purchase_price	repeat_retailer	used_chip	used_pin_number	online_order	fraud
0	57.877857	0.311140	1.945940	1.0	1.0	0.0	0.0	0.0
1	10.829943	0.175592	1.294219	1.0	0.0	0.0	0.0	0.0
2	5.091049	0.805153	0.427715	1.0	0.0	0.0	1.0	0.0
…	…	…	…	…	…	…	…	…
999997	2.914857	1.472687	0.218075	1.0	1.0	0.0	1.0	0.0
999998	5.258729	0.242023	0.475822	1.0	0.0	0.0	1.0	0.0
999999	52.108125	0.318110	0.386920	1.0	1.0	0.0	1.0	0.0

B. Attribute Selection

Salah satu teknik dalam data mining yang sering digunakan pada tahap preprocessing adalah seleksi atribut/feature, yang bertujuan untuk mengurangi kompleksitas atribut. Proses seleksi atribut membantu dalam memilih atribut yang memiliki dampak signifikan (atribut optimal) dan mengabaikan atribut yang tidak berpengaruh sehingga mempercepat proses pemodelan.

Attribute Selection untuk menentukan variabel-variabel yang berpengaruh pada hasil atau output data yang digunakan yaitu fraud. Attribute Selection menggunakan features importance. Hasil Feature importance menunjukan bahwa repeat_retailer memiliki features importanace yang terkecil sehingga kedua atribut tersebut tidak dipakai. Atribut yang akan digunakan pada klasifikasi adalah Distance_from_home, distance_from_last_ transaction, ratio_to_median_purchase_price, used_chip, used_pin_ number, dan online order. Atribut ratio_to_median_purchase_price memiliki korelasi paling besar dengan variabel fraud, maka atribut tersebut dijadikan sebagai atribut root node Hasil feature importance dapat dilihat pada Figure 2.

TABLE 2. Deskripsi Atribut

Attribute	Deskripsi
distance_from_home	Jarak dari rumah tempat terjadinyatransaksi.
distance_from_last_transaction	Jarak dari transaksi terakhir terjadi.
ratio_to_median_purchase_price	Rasio harga pembelian transaksi terhadap harga pembelian median.
repeat_retailer	Transaksi terjadi dari pengecer yang sama (1 = Ya, 0 = Tidak).
used_chip	Transaksi melalui chip (kartukredit) (1 = Ya, 0 = Tidak).
used_pin_number	Transaksi dilakukan d menggunakan nomor PIN (1 = Ya, 0 = Tidak).
online_order	Transaksi order online (1 = Ya, 0= Tidak).
fraud	Transaksi tersebut penipuan (1 =Ya, 0 = Tidak).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

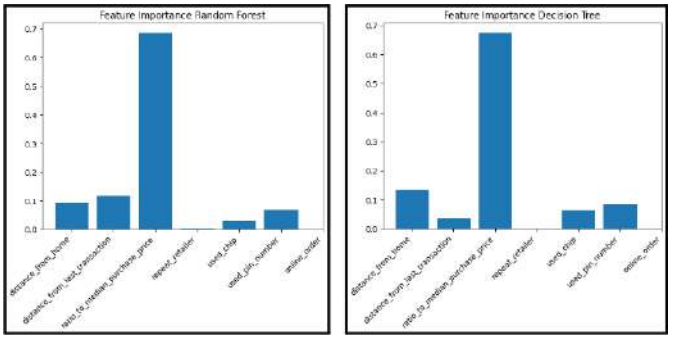


Fig 2. Features Importance Random Forest dan Decision Tree

A. *DATA PRE-PROCESSING*

Data Preprocessing memegang peran utama dalam *data mining*. Dalam langkah pertama melakukan *data mining* adalah melakukan *training* pada *data* supaya dalam penemuan ilmu tidak akansulit memahaminya dengan tidak adanya data yang tidak diinginkan, tidak relevan atau berantakan dandata yang tidak dapat diandalkan [13]. *Preprocessing* dianggap sebagai langkah yang diperlukan untuk mendapatkan model klasifikasi yang memiliki kinerja yang tinggi dalam memprediksi [14]. Tujuan dari *data pre-processing* adalah untuk meminimalkan efek dari data mentah yang tidak berkualitas pada hasil analisis atau pemodelan.

Proses *pre-processing* pada penelitian ini meliputi *data cleaning* terhadap data yang memiliki *missing value* dan *outlier* kemudian dilakukan normalisasi menggunakan metode Z-Score.Pada tahap *data cleaning*, *dataset* yang terdiri dari satu juta record ini tidak terdapat *missing values*, yang dapat dilihat pada Figure 2. Outlier dapat dilihat pada Table 3.

Setelah data *cleaning*, dilakukan normalisasi dengan menggunakan metode Z-Score. Hasil setelah dilakukan normalisasi dapat dilihat pada Table 4.

distance_from_home	0
distance_from_last_transaction	0
ratio_to_median_purchase_price	0
repeat_retailer	0
used_chip	0
used_pin_number	0
online_order	0
fraud	0

Fig 3. Hasil Missing Value

B. *Split data*

Split data adalah proses pembagian *dataset* menjadi dua bagian, yaitu *training set* dan *testing set*, dengan menggunakan proporsi 7:3 yang dilakukan secara acak. Tujuan dari *split data* adalah untuk melatih model dengan *data training* dan menguji model tersebut dengan *data test* yang belum pernah dilihat sebelumnya untuk mengevaluasi performa model. Pembagian data ini sangat penting dalam pengembangan model *machine learning* karena dapat membantu menghindari *overfitting* dan memastikan bahwa model yang dibangun dapat menggeneralisasi data dengan baik [15].

C. Klasifikasi

Klasifikasi menggunakan Random Forest dan Decision Tree. *Input* dari klasifikasi adalah atribut yang diperoleh dari *attribute selection*, yaitu *Distance_from_home, distance_from_last_transaction, ratio_to_median_purchase _price, used_chip, used_pin_number, dan online order. Output* dari klasifikasi adalah penipuan atau tidak.

1. *Decision Tree Classifier*

Decision Tree merupakan sebuah algoritma klasifikasi yang dapat dijelaskan sebagai pemisahan rekursif dari ruang sampel. *Decision Tree* terdiri dari simpul-simpul yang membentuk struktur pohon, di mana pohon tersebut memiliki simpul awal yang disebut akar dan terarah. Pohon keputusan didasarkan pada teknik data mining yang secara rekursif mempartisi sekumpulan data menggunakan metode *depth-first greedy* atau pendekatan *breadth-first* [16]. Semua *node* akan terhubung dengan garis. Dalam pohon keputusan, setiap simpul internal akan membagi ruang sampel menjadi subruang yang lebih kecil, biasanya dua atau lebih subruang, berdasarkan suatu fungsi diskrit yang terkait dengan nilai atribut.

Klasifikasi pohon keputusan memecahkan masalah kompleks dengan memisahkannya menjadi yang sederhana dan menyelesaikannya dengan membangun pohon keputusan berdasarkan pengetahuan yang diperoleh melalui teknik penambangan data. Dasar model pohon keputusan adalah membangun pohon keputusan dengan presisi tinggi dan skala kecil [17].

2. *Random Forest Classifier*

Random Forest (RF) merupakan pengembangan dari metode *Classification and Regression Tree* (CART) dengan menerapkan teknik *bootstrap aggregating* (*bagging*) dan seleksi fitur acak [10]. Metode ini mudah digunakan dan diakui keakuratannya dalam mengatasi sampel kecil dan fitur dengan dimensi tinggi. Selain itu, metode ini dapat dengan mudah diparalelkan, sehingga cocok untuk sistem yang kompleks dalamkehidupan nyata [18]. Kelebihan yang lain adalah pada algoritma RF tidak terdapat *pruning* atau pemangkasan variabel seperti pada algoritma *decision tree.*

Random Forest (RF) mengkombinasikan beberapa pohon (*tree*), berbedadengan pohon tunggal yang hanya terdiri dari satu pohon, untuk melakukan klasifikasi dan prediksi kelas. Pada RF, pembentukan pohon dilakukan dengan melatih sampel data. Pengambilan sampel dilakukan dengan penggantian (*sampling with replacement*). Pemilihan variabel untuk melakukan pemisahan (*split*) dilakukan secara acak. Setelah semua pohon terbentuk, klasifikasi dilakukan. Keputusan

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

klasifikasi akhir diambil dengan cara menghitung rata-rata (menggunakan *mean* aritmatika) probabilitas penugasan kelas yang dihitung oleh semua pohon yang dihasilkan. Setiap pohon memberikan suara untuk keanggotaan kelas.Kelas keanggotaan dengan suara terbanyak akan menjadi yang akhirnya dipilih [10].

D. Evaluation

Tahap ini dilakukan dengan melihat tingkat performa dari pola yang dihasilkan oleh model yang digunakan. Evaluasi perbandingan algoritma menggunakan parameter confusion matrix dengan mengacu pada akurasi, presisi, dan recall. Confusion matrix tetap menjadi metode yang efektif hingga saat ini untuk mengukur dan mengevaluasi kinerja model klasifikasi [19]. Persamaan yang digunakan untuk dijadikan sebagai bahan evaluasi dalam mengukur akurasi model adalah sebagai berikut.

Persamaan yang digunakan untuk sebagaibahan evaluasi dalam mengukur akurasi model adalah Eq. (1).

$$\text{Akurasi} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \text{ x } 100\ \% \qquad (1)$$

TABLE 3. Hasil Outlier

	Distance_from home	Distance_from_last_transaction	Ratio_to_median_purchase_price	Repeat_retailer	Used_chip	Used_pin_number	Online_order	Fraud
12	765,282559	0,371562	0,551245	1.0	1.0	0.0	0.0	0.0
13	2,131956	56,372401	6,358667	1.0	0.0	0.0	1.0	1.0
24	3.803057	67,241081	1,87295	1.0	0.0	0.0	1.0	1.0
…….								
999949	15.724799	1.875906	11.009366	1.0	1.0	0.0	1.0	1.0
999961	337.291331	4.606990	0.181799	1.0	1.0	0.0	1.0	0.0
999973	10.148074	4.465290	12.022734	1.0	0.0	1.0	0.0	0.0

TABLE 4. Hasil Normalisasi Z-Score

	Distance_from_home	Distance_from_last_transaction	Ratio_to_median_purchase_price	Repeat_retailer	Used_chip	Used_pin_number	Online_order	Fraud
0	0,477882	-0,182849	0,043491	0,366584	1,361576	-0,334458	-1.364425	-0,309474
1	-0,241607	-0,188094	-0,189300	0,366584	-0,734443	-0,334458	-1.364425	-0,309474
2	-0,329369	-0,163733	-0,498812	0,366584	-0,734443	-0,334458	0.732909	-0,309474
…								
999997	-0,362650	-0,137903	-0,573694	0,366584	1,361576	-0,334458	0.732909	-0,309474
999998	-0,342098	-0,185523	-0,481628	0,366584	-0,734443	-0,334458	0.732909	-0,309474
999999	0,481403	-0,182849	0,513384	0,366584	1,361576	-0,334458	-1.364425	-0,309474

III. HASIL DAN PEMBAHASAN

a. Heatmap

Heatmap merupakan suatu representasi grafis dari data dengan nilai individu yang terkandung dalam suatu matriks [20]. Visualisasi *heatmap* digunakan untuk melihat korelasi masing-masing variabel dalam kontribusinya dengan *output* data. Pada Gambar 3, menunjukkan grafik *heatmap* yang berisi nilai korelasi antar variabel yang terdapat pada data mengenai kontribusinya dalam menentukan *output* data. Setelah menjalankan fitur *heatmap*, didapatkan kesimpulan bahwa variabel *ratio_to_median_purchase_price*, variabel yang mewakili rasio harga pembelian transaksi terhadap harga pembelian median, memiliki korelasi paling besar dengan variabel *fraud*, variabel yang mewakili apakah transaksi tersebut penipuan, dengan nilai sebesar 0.46. Hasil *Heatmap* dapat dilihat pada Figure 4.

b. Decision Tree

Figure 5 yang menunjukkan hasil *Decision Tree*, didapatkan hasil mengenai klasifikasi dari *Decision Tree* yang dijalankan tanpa adanya atribut*repeat_retailer*. Pada awalnya, dilakukan klasifikasi apakah *ratio_to_median_purchase _price* kurang dari sama dengan 0.777 yang diambil mnggunakan 1000000 sampel merupakan *true* atau *false.* Pada klasifikasi *true*, diklasifikasi apakah *distance_from_home* kurang dari sama dengan 1.122 menggunakan 896842 sampel. Jika true, maka kelas *output*-nya adalah *fraud* dengan sampel berjumlah

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

852163. Jika false maka dilakukan klasifikasi apakah *online_order* kurang dari sama dengan -0.316. Jika *true* maka kelas *output*-nya adalah *fraud* dengan sampel sejumlah 15628, dan jika *false* maka kelas *output*-nya adalah *not fraud* dengan sampel sejumlah 29053.

Klasifikasi *false* merupakan hasil dariklasifikas penentua apakah *ratio_to_median_purchase_price* kurang dari sama dengan 0.777, dilakukan klasifikasi penentuan apakan *online_order* kurang dari sama dengan -0.316 menggunakan sampel sejumlah 103158. Jika*true* maka dilakukan klasifikasi penentuan apakah *distance_from_home* kurang dari sama dengan 1.123 menggunakan 36190 sampel yang jika *true* maka akan menghasilkan kelas *output fraud* dengansampel sejumlah 34415 dan jika false maka akan -0.316 adalah false maka menghasilkan kelas output not fraud sejumlah1775 sampel.

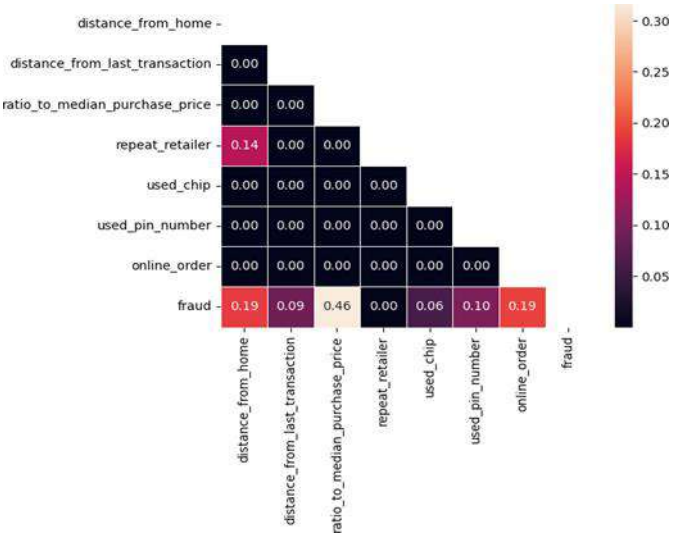


Fig 4. Hasil Heatmap

Jika penentuan apakah *online_order* kurang dari sama dengan dilakukan klasifikasi penentuan apakah *used_pin_number* lebih kecil dari sama dengan 1.328. Jika *true* maka akan menghasilkan kelas *output fraud* sejumlah 6765 sampel, sedangkan jika*false* maka akan menghasilkan kelas *output not fraud* sejumlah 60203 sampel. Hasil tree yang dihasilkan oleh Decision Tree dapat dilihat pada Figure 5.

TABLE 5. *Confusion Matrix Decision Tree*

		Actual Values	
		Positive	*Negative*
Predicted Values	*Positive*	270403	*3537*
	Negative	*2518*	*23542*

TABLE 6. Akurasi Decision Tree

	precision	**recall**	**f1-score**	**support**
0	0.99	0.99	0.99	273940
1	0.87	0.90	0.89	26060
accuracy			0.98	300000
macro avg	0.93	0.95	0.94	300000
weighted avg	0.98	0.98	0.98	300000

Tabel 5 menjelaskan hasil *confusion matrix.* Hasil yang didapatkan berupa nilai *True Positive* (TP) sebesar 270403, nilai *False Positive* (FP) sebesar 3537, nilai *True Negative* (TN) sebesar 23542, dan nilai *False Negative* (FN) sebesar 2518. Kemudian pada Tabel 6, didapatkan tingkat akurasi dari dijalankannya *Decision Tree*. Tingkat akurasi yang didapatkan dengan menggunakan Persamaan (1) sebesar 0.99 atau 99%. Hasil akurasi yang diperoleh dari Decision Tree ditunjukkan pada Table 6.

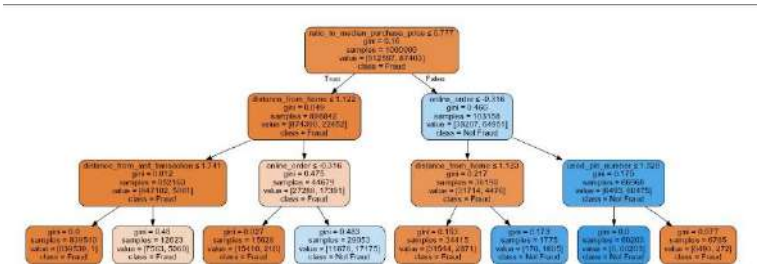


Fig 5. Hasil Tree yang dihasilkan Decision Tree

a. Random Forest

Figure 6 didapatkan hasil mengenai random forest yang telah dijalankan tanpa adanya atribut repeat_retailer. Pada klasifikasi pertama, yaitu apakah ratio_to_median_purchase_price kurang dari sama dengan 0.777. Jika true, maka akan dilakukan klasifikasi lanjutan apakah online_order kurang dari sama dengan minus 0.316 dengan sampel 396948. Dengan hasil true, dilakukan klasifikasi tambahan distance_from_last_transaction apakah kurang dari sama dengan 1.751 dengan sampel 138535, yang dimana jika true maka output akan menunjukkan fraud dengan sampel 136514, begitu pula jika false yang menunjukkan kelas fraud dengan sampel 2021.

Klasifikasi variabel online_order kurang dari sama dengan minus 0.316 false, akan dilanjutkan klasifikasi tambahan menggunakan variabel used_pin_number kurang dari 1.328 dengan sampel 258413. Jika true, maka output akan menghasilkan kelas fraud dengan sampel 232194 dan gini

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

0.081, begitu pula dengan false dengan sampel 26219 dan gini 0.0.

Klasifikasi variabel ratio_to_median_purchase_price false, dilakukan klasifikasi lebih lanjut menggunakan variabel used_pin_number kurang dari 1.328. Jika true, maka akan dilanjutkan klasifikasi lebih lanjut menggunakan variabel distance_from_home kurang dari 1.121 dengan sampel 40976 yang dimana true dan false menghasilkan output yang sama yaitu not fraud dengan true bersampel 38924 dengan gini0.435 dan false bersampel 2052 dengan gini 0.0.

Hasil false used_pin_number kurang dari 1.328, dilakukan klasifikasi lebih lanjut menggunakan variabel distance_from_last_transaction kurang dari 1.761. Jika true, maka akan menghasilkan output kelas fraud dengan sampel 4589 dan gini 0.038. Jika false, maka akan menghasilkan output not fraud dengan sampel 77 dan gini 0.484.

Table 7 didapatkan hasil setelah dijalankannya confusion matrix. Hasil yang didapatkan berupa nilai True Positive (TP) sebesar 273940, nilai False Positive (FP) sebesar 0, nilai True Negative (TN) sebesar 18601, dan nilai False Negative (FN) sebesar 7459. Kemudian pada Table 8, didapatkan tingkat

akurasi dari dijalankannya Random Forest. Tingkat akurasi yang didapatkan dengan menggunakan Eq. (1) sebesar 0.975 atau 97,5%.

TABLE 7. Confusion Matrix Random Forest

	precision	**recall**	**f1-score**	**support**
0	0.97	1.00	0.99	273940
3	1.00	0.71	0.83	26060
accuracy			0.98	300000
macro avg	0.99	0.86	0.91	300000
weighted avg	0.98	0.98	0.97	300000

TABLE 8. Akurasi Random Forest

		Actuaal Values	
		Positive	Negative
Predictd Values	Positive	273940	0
	Negative	7459	18601

Berdasarkan akurasi yang diperoleh menunjukkan bahwa Random forest dan decision tree dengan attribute selection memiliki akurasi yang lebih tinggi dibandingkan dengan penelitian sebelumnya[12] yang memperoleh akurasi sebesar 74,76% untuk Decision Tree sebesar dan Random Forest sebesar 63,27% . Hal ini menunjukkan bahwa attribute selection dapat meningkatkan akurasi.

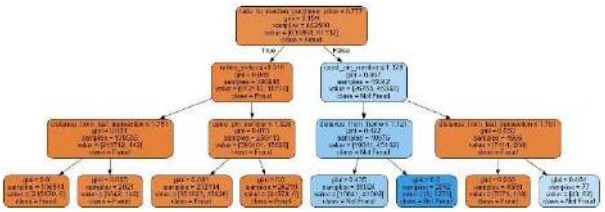


Fig. 6 Hasil Tree yang dihasilkan Random Forest

IV. KESIMPULAN

Kartu kredit semakin populer dalam era digital, tetapi penipuan menggunakan kartu kredit juga semakin meningkat. Penelitian ini bertujuan untuk menentukan algoritma klasifikasi yang dapat membantu dalam mengidentifikasi kartu kredit pada transaksi ilegal secara akurat. Dataset yang digunakan dalam penelitian ini adalah data transaksi menggunakan kartu kredit yang diambil dari Kaggle. Data tersebut sebanyak satu juta records.

Dataset terdapat variabel fraud yang mewakili apakah transaksi tersebut penipuan atau tidak. Oleh karena itu variabel ini akan digunakan sebagai target. Setelah dilakukan visualisasi dengan heatmap didapatkan hasil bahwa atribut ratio_to_median_purchase_price memiliki korelasi paling besar dengan atribut fraud, dengan nilai sebesar 0,46. Dengan demikian, atribut ratio_to_median_purchase_price akan dijadikan sebagai atribut root node yaitu atribut yang menjadi keputusan besar pertama yang akan menentukan apakah record tersebut termasuk penipuan atau tidak di akhir.

Klasifikasi yang telah dilakukan menggunakan atribut Distance_from_home, distance_from_last_ transaction, ratio_to_median_purchase_price, used_chip, used_pin_ number, dan online order.

Hasil akurasi yang diperoleh adalah menggunakan Decision Tree memiliki tingkat akurasi sebesar 98%d Random Forest memiliki tingkat akurasi sebesar 97,5%. Berdasarkan hasil akurasi, disimpulkan bahwa Decision Tree memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan Random Forest untuk identifikasi penipuan kartu kredit menggunakan attribute selection serta attribute selection dapat meningkatkan akurasi.

REFERENCES

[1] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit Card Fraud Detection Using AdaBoost and Majority Voting," IEEE Access, vol. 6, pp. 14277–14284, 2018, doi: 10.1109/ACCESS.2018.2806420.

[2] A. De Sá, A. Pereira, and G. Pappa, "A Customized Classification Algorithm for Credit-Card Fraud Detection," Eng Appl Artif Intell, vol. 72, May 2018, doi: 10.1016/j.engappai.2018.03.011.

[3] Y. Jain, N. Tiwari, S. Dubey, and S. Jain, "A comparative analysis of various credit card fraud detection techniques," International Journal of Recent Technology and Engineering, vol. 7, pp. 402–407, May 2019.

[4] J. K. Afriyie et al., "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," Decision Analytics Journal, vol. 6, p. 100163, 2023, doi: https://doi.org/10.1016/j.dajour.2023.100163.

[5] H. Wang, C. Ma, and L. Zhou, "A Brief Review of Machine Learning and Its Application," in 2009 International Conference on Information

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023

Engineering and Computer Science, 2009, pp. 1–4. doi: 10.1109/ICIECS.2009.5362936.

[6] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," EURASIP J Adv Signal Process, vol. 2016, no. 1, p. 67, 2016, doi: 10.1186/s13634-016-0355-x.

[7] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine Learning and Data Mining Methods in Diabetes Research," Comput Struct Biotechnol J, vol. 15, pp. 104–116, 2017, doi: https://doi.org/10.1016/j.csbj.2016.12.005.

[8] Q. Dai, C. Zhang, and H. Wu, "Research of Decision Tree Classification Algorithm in Data Mining," International journal of database theory and application, vol. 9, pp. 1–8, 2016.

[9] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques third edition." Morgan Kaufmann Publishers, Waltham, Mass., 2012. [Online]. Available:http://www.amazon.de/Data-Mining-Concepts-TechniquesManagement/dp/0123814790/ref=tmm_hrd_title_0?ie=UTF8&qid=136603903 &sr=1-1

[10] M. Belgiu and L. Drăguţ, "Random Forest in remote sensing: A review of applications and future directions," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 114, pp. 24–31, 2016, doi: https://doi.org/10.1016/j.isprsjprs.2016.01.011.

[11] P. Kashyap, Machine Learning for Decision Makers: Cognitive Computing Fundamentals for Better Decision Making, 1st ed. USA: Apress, 2018.

[12] Shah, D and Sharma, LK. "Credit Card Fraud Detection using Decision Tree and Random Forest." ITM Web of Conferences, 2023, search.proquest.com,

[13] M. Durairaj and N. Ramasamy, "A comparison of the perceptive approaches for preprocessing the data set for predicting fertility success rate," International Journal of Control Theory and Applications, vol. 9, no. 27, pp. 255–260, 2016.

[14] E. Alshdaifat, D. Alshdaifat, A. Alsarhan, F. Hussein, and S. M. F. S. El-Salhi, "The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms' Performance," Data (Basel), vol. 6, no. 2, 2021, doi: 10.3390/data6020011

[15] S. Cho et al., "A Hybrid Machine Learning Approach for Predictive Maintenance in Smart Factories of the Future," in Advances in Production Management Systems. Smart Manufacturing for Industry 4.0, I. Moon, G. M. Lee, J. Park, D. Kiritsis, and G. von Cieminski, Eds., Cham: Springer International Publishing, 2018, pp. 311–317.

[16] J. R. Gaikwad, A. B. Deshmane, H. V Somavanshi, S. V Patil, and R. A. Badgujar, "Credit card fraud detection using decision tree induction algorithm," International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 4, no. 6, pp. 2278–3075, 2014.

[17] D. L. Talekar and K. P. Adhiya, "Credit Card Fraud Detection System: A Survey," International journal of modern engineering research (IJMER), vol. 4, no. 9, 2014.

[18] G. Biau and E. Scornet, "A random forest guided tour," TEST, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.

[19] A. Nurmasani and Y. Pristyanto, "Algoritme Stacking Untuk Klasifikasi Penyakit Jantung Pada Dataset Imbalanced Class," Pseudocode, vol. 8, no. 1, pp. 21–26, Mar. 2021, doi: 10.33369/pseudocode.8.1.21-26.

[20] C.-S. Yu et al., "Clustering Heatmap for Visualizing and Exploring Complex and High-dimensional Data Related to Chronic Kidney Disease," J Clin Med, vol. 9, no. 2, 2020, doi: 10.3390/jcm9020403.

[21] Bhavani, T. T., Rao, M. K., & Reddy, A. M. (2019, November). Network intrusion detection system using random forest and decision tree machine learning techniques. In First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019 (pp. 637-643). Singapore: Springer Singapore.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v12i3.1730, Copyright ©2023
Submitted : May 17, 2023, Revised : July 5, 2023, Accepted : October 17, 2023, Published : November 9, 2023