

# Systematic Literature Review: Metode Machine Learning dalam Klasifikasi Emosi pada Data Tekstual

Putu Widyantara Artanta Wibawa<sup>[1]</sup>, Cokorda Pramarta<sup>[2]\*</sup>

Program Studi Informatika<sup>[1]</sup>

Net-Centric Computing Research Lab<sup>[2]</sup>

Fakultas MIPA, Universitas Udayana

Jimbaran, Bali, Indonesia

[putuwaw973@gmail.com](mailto:putuwaw973@gmail.com)<sup>[1]</sup>, [cokorda@unud.ac.id](mailto:cokorda@unud.ac.id)<sup>[2]</sup>

**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

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.

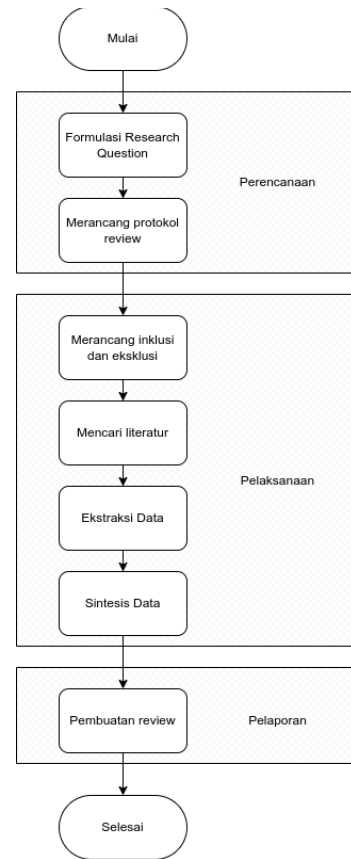
## II. METODE PENELITIAN

### 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?

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.

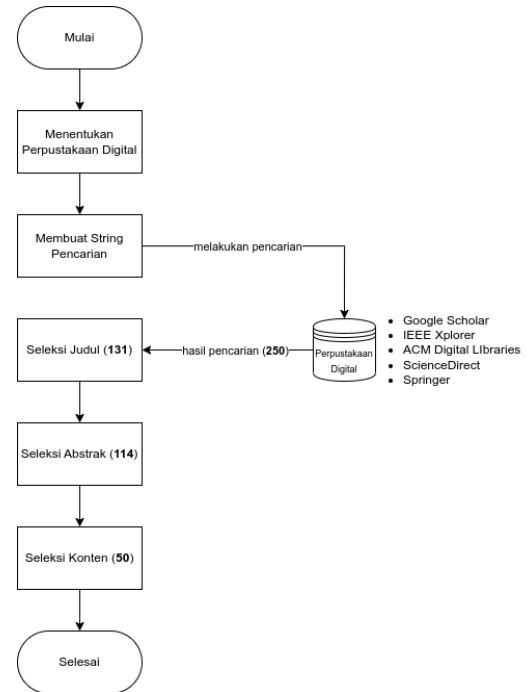
TABEL IV TABEL KRITERIA INKLUSI DAN EKSKLUSI

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.



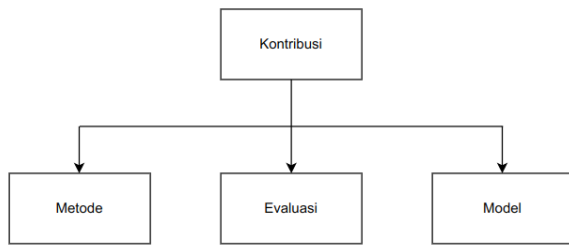
Gambar 2. Tahapan Pencarian dan Seleksi Literatur

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



Gambar 3 Kerangka Kontribusi dari Literatur

F. Sintesis Data

Sintesis data melibatkan 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 Teksual?

Penelitian ini menghasilkan tiga kontribusi, yaitu kontribusi terhadap metode, evaluasi, dan model dalam klasifikasi emosi pada data teksual. 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].

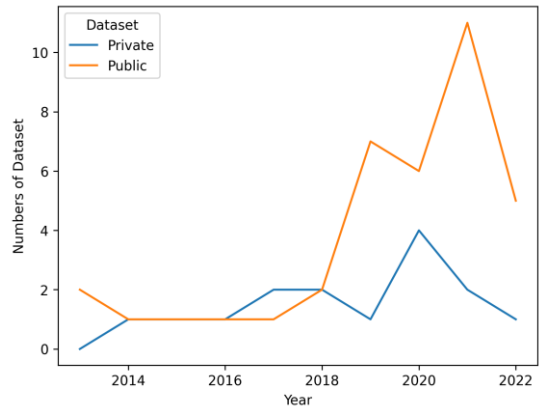
Penelitian yang dilakukan oleh Batbaatar, et al mengusulkan sebuah metode baru dalam klasifikasi emosi, yaitu SENN (Semantic Emotion Neural Network) [20]. Metode yang berbeda juga diusulkan pada penelitian yang dilakukan oleh Chatterjee et. al. Pada penelitian tersebut peneliti menggunakan SS-BED yang 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

<b>Id</b>	<b>Referensi</b>	<b>Tahun</b>	<b>Sumber</b>	<b>Kontribusi</b>
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

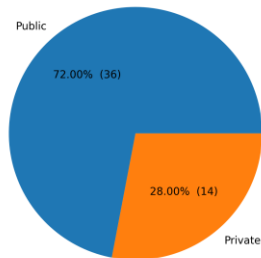
222	[63]	2022	Springer	Model
225	[64]	2022	Springer	Model
229	[65]	2019	Springer	Model
238	[66]	2021	Springer	Metode
240	[67]	2021	Springer	Model
248	[68]	2021	Springer	Evaluasi
249	[69]	2022	Springer	Metode



Gambar 5. Grafik Tren Penggunaan Dataset

**B. RQ2: Apakah Jenis Dataset yang Digunakan dalam Klasifikasi Emosi pada Data Teksual?**

Dataset yang digunakan dalam *systematic literature review* ini adalah dataset teksual 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.

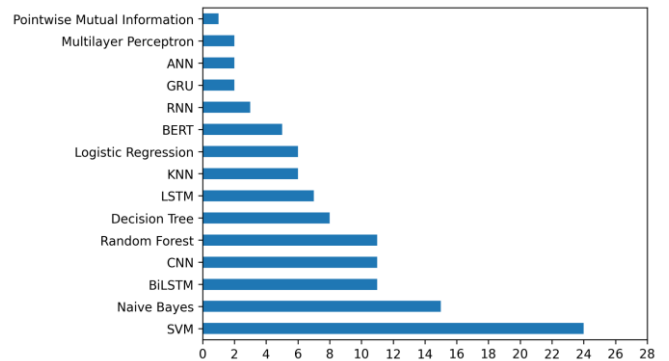


Gambar 3. Grafik Distribusi Dataset

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 bahwa 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.

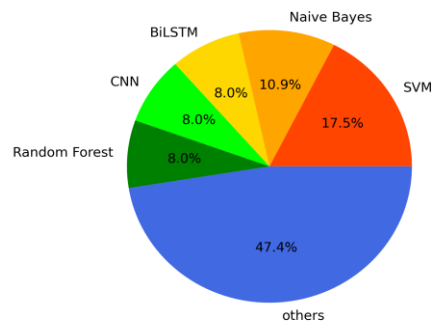
**C. RQ3: Apakah Metode Yang Digunakan Dalam Klasifikasi Emosi Pada Data Teksual?**

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

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 yangimbang 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].

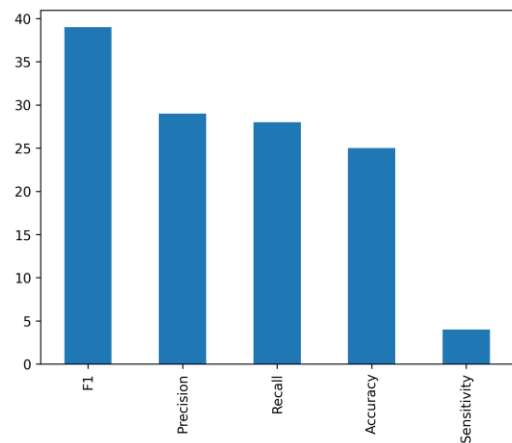
Neural Network atau Deep Neural Network adalah model yang meniru bagaimana otak manusia bekerja [16]. Arsitektur dari Neural Network terbentuk dari lapisan yang saling terhubung yang terdiri dari tiga tingkatan yang berbeda, yaitu *input layer*, *hidden layer*, dan *output layer* [17]. Secara umum terdapat 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

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 (AIoT), 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. In *2017 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-Classifer 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

- 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.knsys.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.knsys.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.knsys.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., Moseleh, 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., Houbay, 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., Mimalinee, 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>



[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>