

# Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM

Nanda Ressq Setiawan<sup>1</sup>, Emil R. Kaburuan<sup>2</sup>

<sup>1,2</sup>Program Studi Teknik Informatika

Universitas Mercu Buana Jakarta

Jl. Meruya Selatan No.1, Meruya Sel., Kembangan, Daerah Khusus Ibukota Jakarta 11650, Indonesia,

[nandaressqsetiawan@gmail.com](mailto:nandaressqsetiawan@gmail.com)<sup>1</sup>, [emil.kaburuan@mercubuana.ac.id](mailto:emil.kaburuan@mercubuana.ac.id)<sup>2</sup>

The SIM renewal service is familiar with a complicated service process. Addressing complaints in the community, Korlantas Polri facilitates fast and easy online services through the digital application for the SINAR Korlantas Polri service. However, the rating, which is accompanied by various negative and positive reviews, shows that the service provided has not fully met the expectations of the application's users. In order to find out how optimal the Digital Korlantas Polri application can be found by analyzing the sentiments of user reviews. The purpose of this study is to implement the sentiment analysis review of the Korlantas Digital Application on the Google Play Store using the SVM method. User review data for the Digital Korlantas Polri application obtained through the Google Play website with a total sample of 1200 review data. Collecting data in this study by crawling data using the google-play-scraper library. The testing method uses a confusion matrix. The results of the study show that the SVM algorithm can perform sentiment analysis on the Korlantas digital application review with the results of 598 positive sentiments and 511 negative sentiments. Based on the test results the SVM model has good performance in the 90:10 data ratio scenario with an accuracy value of 0.82 and the SVM model with the worst performance is in the 80:20 and 60:40 data ratio scenario with an accuracy of 0.74.

**Keywords**—*Digital Korlantas, Google Play Store, Application Reviews, Sentiment Analysis, Support Vector Machine*

Layanan perpanjangan SIM familiar dengan proses pelayanan yang rumit. Mengatasi keluhan di masyarakat, Korlantas Polri memfasilitasi pelayanan online yang cepat dan mudah melalui aplikasi digital layanan SINAR Korlantas Polri. Namun rating yang disertai berbagai ulasan negatif dan positif menunjukkan bahwa pelayanan yang diberikan belum sepenuhnya memenuhi harapan dari pengguna aplikasi tersebut. Agar dapat mengetahui seberapa optimal aplikasi Digital Korlantas Polri dapat diketahui dengan menganalisis sentimen ulasan pengguna. Tujuan penelitian ini adalah mengimplementasikan sentimen analisis review Aplikasi Digital Korlantas pada Google Play Store menggunakan Metode SVM. Data ulasan pengguna aplikasi Digital Korlantas Polri yang didapatkan melalui *website* Google Play dengan jumlah sampel sebanyak 1200 data ulasan. Pengumpulan data dalam penelitian ini dengan cara *crawling data* menggunakan *library google-play-scraper*. Metode pengujian menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa algoritma SVM dapat melakukan analisis sentimen pada ulasan aplikasi digital korlantas dengan hasil 598 sentimen positif dan 511 sentimen negatif. Berdasarkan hasil pengujian model

SVM memiliki kinerja yang baik pada skenario rasio data 90:10 dengan nilai akurasi sebesar 0.82 dan model SVM dengan kinerja terburuk ada pada skenario rasio data 80:20 dan 60:40 yakni dengan akurasi sebesar 0.74.

**Kata Kunci**—*Digital Korlantas, Google Play Store, Review Aplikasi, Sentimen Analisis, Support Vector Machine.*

## I. PENDAHULUAN

Semakin pesat berkembangnya teknologi membuat hampir semua aktivitas hidup dari masyarakat sudah semakin dipermudah dengan adanya kemajuan teknologi digital [1]. Oleh karena itu, tidak menutup kemungkinan jika Indonesia akan gencar dalam melakukan beberapa program serta akselerasi transformasi digital yang bahkan hampir digabungkan sebagai program nasional hampir ke seluruh industri. Transformasi digital ini juga sudah menuju industri perusahaan yang saat ini tidak hanya secara mandiri yang dilakukan oleh pemerintah salah satunya yaitu Polres Gresik dengan pelayanan perpanjangan SIM. Layanan perpanjangan SIM familiar dengan proses pelayanan yang rumit dan berbelit-belit.

Hal tersebut menyebabkan masyarakat yang ingin memperpanjang SIM merasa jenuh dengan proses yang tidak efektif dan efisien. Masyarakat sering mengeluh buruknya pelayanan satpas dan pembuatan SIM. Mengatasi keluhan dan kekurangan di masyarakat, Kepolisian Negara Republik Indonesia Korlantas Polri memfasilitasi pelayanan online yang cepat dan mudah. Pelayanan ini dilaksanakan pada masa penyesuaian kebiasaan baru (*new normal*) atau akan normal baru. Polres Gresik berusaha memberikan pelayanan yang lebih optimal terutama perkembangan teknologi untuk meningkatkan kualitas pelayanan masyarakat melalui aplikasi digital layanan SINAR Korlantas Polri [2] [3].

Dalam memberikan pelayanan terhadap *customer* mereka, Korlantas Polri banyak membuka masukan untuk pelayanan yang mereka berikan, mulai dari akun media sosial sampai aplikasi yang disediakan aplikasi Digital Korlantas Polri sebagai tempat aktifitas *customer* dalam melakukan transaksi. Saat ini pada Google Play Store aplikasi Digital Korlantas Polri sudah diunduh dengan rating 4,1 dan 71,4 ribu ulasan. Rating yang disertai berbagai ulasan negatif dan positif menunjukkan

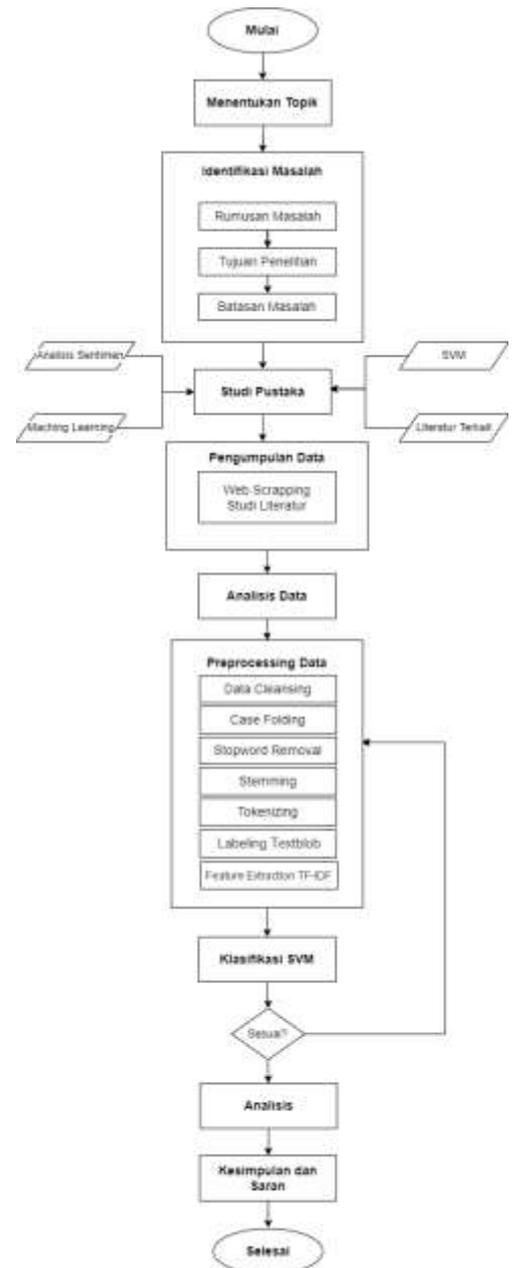
bahwa pelayanan yang diberikan oleh Korlantas Polri belum sepenuhnya memenuhi harapan dari pengguna aplikasi tersebut.

Oleh karena hal tersebut, agar dapat mengetahui seberapa optimal aplikasi Digital Korlantas Polri dapat diketahui dengan menganalisis sentimen ulasan pengguna aplikasi Digital Korlantas Polri. Analisis sentimen merupakan suatu proses pengolahan atau klasifikasi sebuah kata atau text yang bersumber dari internet atau media sosial dengan tujuan mendapatkan respon atau opini masyarakat terkait media sosial [4]. Pada penelitian ini digunakan metode SVM untuk analisis sentimen. SVM adalah algoritma yang dikenal untuk menghasilkan solusi optimal pada kasus klasifikasi. Vapnik memperkenalkan model *machine learning* SVM berbasis kernel. SVM dapat melakukan generalisasi dengan baik, meskipun menggunakan data pelatihan yang kecil [4].

Pada penelitian terdahulu yang ditulis oleh [5] berjudul "Sentiment Analysis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine (SVM)". Penelitian tersebut menganalisis rating aplikasi Maxim di Google Play Store karena pengguna sering membaca ulasan yang ditemukan pengguna lain di Google Play Store sebelum memilih dan menginstal aplikasi. Analisis sentimen memiliki daya tarik yang tinggi bagi pemilik aplikasi dalam membuat keputusan, karena dapat dijadikan bahan evaluasi pengembangan dari aplikasi tersebut. Maxim merupakan aplikasi baru transportasi online yang menawarkan layanan berbeda dari aplikasi sejenis. Peneliti bertujuan menganalisis sentimen pada review pengguna terhadap aplikasi Maxim di Google Play store Analisis rating pengguna ini menggunakan metode (SVM) yang mencapai akurasi sebesar 79%.

## II. METODE PENELITIAN

### 2.1 Jenis Penelitian



Gambar 1. Alur Penelitian

Berdasarkan alur penelitian yang ditampilkan pada gambar 1, penjelasan rinci tentang langkah-langkah penelitian:

1. Menentukan topik

Dalam penelitian ini, awal langkah yang dilakukan dengan pencarian topik yang akan diimplementasikan. Topik yang dipilih adalah Analisis Sentimen Review Aplikasi Korlantas Digital di Google Play Store dengan menggunakan metode SVM

## 2. Identifikasi Masalah

Pada tahap ini identifikasi masalah penulis memanfaatkan situs Google Play Store. Aplikasi Korlantas Digital Polri merupakan salah satu aplikasi layanan perpanjangan SIM yang banyak digunakan khususnya oleh masyarakat Indonesia. Oleh karena itu, analisis sentimen diperlukan untuk mengetahui sentimen pada opini pengguna terhadap aplikasi Digital Korlantas Polri.

## 3. Rumusan Masalah dan Tujuan Penelitian

Langkah ini dilakukan dengan merumuskan masalah terkait penelitian yakni opini pengguna aplikasi Digital Korlantas Polri, hasil penerapan klasifikasi ulasan aplikasi Digital Korlantas Polri menggunakan metode SVM menjadi kelas positif dan negatif, bagaimana informasi yang dihasilkan pada setiap kategori diimplementasikan.

Pada tahap ini, tujuan penelitian ditentukan, yang dihasilkan dari rumusan tujuan yang diinginkan dari masalah yang telah ditetapkan sebelumnya. Penulis bertujuan untuk mendapatkan tanggapan pengguna aplikasi Korlantas Digital Polri dan mengetahui hasil penerapan metode *Support Vector Machine* dalam klasifikasi data ulasan pengguna aplikasi Digital Korlantas Polri.

## 4. Menentukan Batasan dan Metode Penelitian

Pada tahap ini menentukan beberapa batasan penelitian agar penelitian dapat berjalan sesuai tujuan dan tidak melebar ke topik lainnya. Batasan penelitian ini seperti data yang diambil dari website Google Play Store berupa ulasan menggunakan Bahasa Indonesia. Penelitian yang digunakan yaitu metode SVM.

## 5. Studi Pustaka

Studi pustaka dilakukan untuk menunjang penulis pada penelitian yang akan dilakukan terkait literatur dan topik yang sudah ditentukan. studi pustaka yang digunakan sebagai referensi pada penelitian ini terkait dengan teori analisis sentimen, penggunaan machine learning, teori penggunaan metode SVM dan literatur lainnya terkait dengan pengolahan teks.

## 6. Pengumpulan Data

Tahap ini mengumpulkan data ulasan aplikasi Digital Korlantas Polri menggunakan teknik *crawling* atau *scrapping* data menggunakan *library google-play-scraper* yang dimiliki oleh *Google Chrome* pada website Google Play Store.

## 7. Analisis Data

Pada tahap ini dilakukan analisis data. Tahapan ini bertujuan untuk mendeskripsikan persepsi pengguna aplikasi Korlantas Digital Polri.

## 8. Pre-processing Data

Pada tahap ini dilakukan proses preprocessing data, dimana pada proses ini menggunakan software Google Colab dengan bahasa pemrograman python. Berikut merupakan beberapa proses yang terjadi:

- a. *Data Cleansing*
- b. *Case Folding*
- c. *Tokenizing*
- d. *Stopword Removal*
- e. *Stemming*
- f. Pelabelan Data

Pada tahap ini dilakukan pelabelan menggunakan *library textblob*. Pada proses labelling data menggunakan *textblob* data ulasan harus diterjemahkan ke Bahasa Inggris, karena *library textblob* hanya mendeteksi teks berbahasa Inggris.

- g. Feature Extraction menggunakan TF-IDF untuk pembobotan kata

## 9. Klasifikasi

Setelah melalui pre-processing dilanjutkan dengan klasifikasi sentimen menggunakan metode SVM.

## 10. Kesimpulan dan Saran

Setelah melalui seluruh tahap pengolahan data dilakukannya penarikan kesimpulan dari serangkaian proses yang telah dilakukan oleh penulis. Kesimpulan dituliskan berdasarkan hasil penelitian yang didapatkan. Selain itu penulis membuat saran berdasarkan penelitian yang telah dilakukan

## 2.2 Tahapan Penelitian

### 2.2.1 Alat

Perangkat lunak yang digunakan dalam pembuatan adalah sebagai berikut:

- Laptop Processor Intel® Core™ i5 8250U Processor (6M Cache, up to 3.40 GHz), RAM 4GB / 8GB DDR4 2133MHz SDRAM, VGA Discrete Graphics Nvidia GT 930MX 2GB / Nvidia GT 940MX 2GB, Hard Disk 1TB SATA HDD 5400RPM / 246 SSD Sandisk
- Google Collaboratory, sebagai alat bantu untuk menuliskan kode ataupun melakukan proses debugging program Python.
- Microsoft Office, sebagai alat bantu untuk penulisan laporan dan proposal selama pengembangan system

2.2.2 Bahan

Bahan yang terdapat pada penelitian ini merupakan semua data ulasan pengguna aplikasi Digital Korlantas Polri yang didapatkan melalui *website* Google Play dengan jumlah sampel yang digunakan sebanyak 1200 data ulasan

2.3 Metode Pengumpulan Data

Berikut ini merupakan metode yang digunakan dalam penelitian:

1. Web Scraping

Pengumpulan data dalam penelitian ini dengan cara *crawling data* menggunakan library *google-play-scraper*. Teknik pengambilan data ini dapat menemukan data ulasan aplikasi Digital Korlantas secara otomatis.

2. Studi Pustaka

Penulis menggunakan literatur yang bersumber dari jurnal dan e-book yang berkaitan dengan penelitian, kemudian tujuan serta rumusan masalah digunakan pada penelitian ini.

2.4 Metode Analisis Data

Penulis menganalisis data dengan metode *SVM* untuk mengklasifikasi ulasan menjadi sentimen positif dan negatif.

2.5 Rancangan Pengujian

Metode pengujian menggunakan *confusion matrix* untuk mengetahui kinerja dari model klasifikasi yang dibuat. Model *SVM* terbaik akan digunakan untuk diimplementasikan pada ulasan, untuk mengetahui sentimen yang dihasilkan pada tiap ulasan

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini, Dataset yang dibutuhkan penulis untuk menyelesaikan Tugas akhir meliputi konfigurasi Keamanan Jaringan Enterprise dengan metode Intrusion Detection & Prevention System (IDPS), yang mana penulis menyiapkan semua konfigurasi baik itu Network, System Security, End point maupun Penetration testing sebagai Intruder. Berikut detail perangkat yang dilampirkan dalam bentuk tabel.

A. Dataset

Penulis mendapatkan dataset dengan melakukan *crawling* pada *website google play store*. Library yang digunakan untuk melakukan *crawling data* adalah *google-play-scraper*. Didapatkannya 1200 data dengan 12 macam atribut dari proses *crawling data*.

TABEL 1. DAFTAR ARTIBUT DATASET

Nama Atribut	Keterangan
--------------	------------

<i>reviewId</i>	Berisi no Id review
<i>username</i>	Berisi username pengguna aplikasi cek bansos
<i>userImage</i>	Berisi link gambar profil pengguna
<i>content</i>	Berisi komentar pengguna terkait aplikasi digital korlantas
<i>score</i>	Berisi angka penilaian yang diberikan untuk aplikasi cek bansos
<i>thumbsUpCount</i>	Berisi jumlah penyuka komentar
<i>reviewCreatedVersion</i>	Berisi seri atau versi aplikasi yang digunakan oleh pengguna
<i>At</i>	Berisi waktu pengguna mengunggah komentar
<i>replyContent</i>	Berisi balasan komentar antar pengguna
<i>repliedAt</i>	Berisi waktu balasan komentar antar pengguna
<i>sortOrder</i>	Berisi kategori komentar diurutkan berdasarkan most relevant dan newest
<i>appld</i>	Berisi Id aplikasi digital korlantas

Tidak semua atribut digunakan pada penelitian ini. Penulis hanya menggunakan atribut *content*, karena hanya pada atribut *content* yang berisi ulasan pengguna.

TABEL 2. DATA ULASAN PENGGUNA

Content
Aplikasi tidak bisa dipakai misalkan klik e-psi tidak bisa, klik e-rikkes juga tidak bisa, jika belum sempurna aplikasinya lebih baik tidak usah di launching. Jika begini kita tetap harus ke satpas didaerah rumah kita.
Regnya gagal. Masukkan no HP -> RPC time out. Setelah itu baru dpt massage OTP. Sepertinya timer nya terlalu cepat sehingga tidak nyambung dengan OTPnya. Atau OTPnya telat masuk. Tolong di perbaiki bugnya
Aplikasinya mungkin belum sepenuhnya siap, mau verifikasi KTP dengan foto wajah, sudah mengikuti sesuai intruksi dan hasilnya diterima tapi tanda silang setelah dilihat di profil tapi belum terverifikasi KTP nya, sudah dicoba beberapa kali tetap belum terverifikasi. Tolong diperbaiki aplikasinya, 108alua belum siap harusnya jangan dulu dilaunching.
Ini di implementasikan nasional tapi sepertinya servernya seadanya ya? Susah banget masuknya time out terus. Belum bisa review lebih lanjut belum lihat semua. Ini salah satu wujud POLRI go digital untuk menjawab kebutuhan masa depan dan mengurangi potensi pungli tapi mohon dengan sangat komitmennya, servernya juga di kembangkan skala nasional dong.
Tidak bisa masukkan nomor Handphone, selalu muncul "Terjadi

Kesalahan, Coba lagi beberapa saat" atau biasa muncul " RPC Timeout", sebenarnya aplikasi ini bisa di pakek apa tidak sih,, kalau belum bisa di pakek,, jangan di luncurkan dulu
Apanya yg simpel perpanjang sim cukup dari HP. Proses ribet berkas pendukung harus instal apk lagi atau daftar lg via web disitus yg berbeda. Total jadinya ada 3 apk/situs. Belum lagi tetep harus datang untuk cek kondisi. Gk cocok sama apa yg digembor2kan dimedia.
Saya sudah berhasil memasukan untuk nomer SIM A. Tapi ketika saya input untuk SIM C, selalu saja gagal. Sudah cara input manual & scan barcode. Ketika simpan selalu tidak berhasil. Verifikasi akun sudah validasi lengkap semua. Sudah coba Logout dan login ulang. Tapi belum bisa. Server nya apakah penuh ? Efek louncing trial barusan aplikasinya.

beberapa kali tetap gagal. .. Mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus, mungkin bisa diperbaiki lagi baik dari server dan aplikasinya. Jika aplikasi ini sukses secara Nasional, maka akan bagus. Karena akan menghilangkan pungli2 yg ada di daerah2 yg jauh dri pusat.	beberapa kali tetap gagal. .. Mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus, mungkin bisa diperbaiki lagi baik dari server dan aplikasinya. Jika aplikasi ini sukses secara Nasional, maka akan bagus. Karena akan menghilangkan pungli yg ada di daerah yg jauh dri pusat.
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

B. Pre-Processing

Pre-processing terdiri dari beberapa tahapan yakni *data cleansing*, *case folding*, *stopword removal*, *stemming*, *tokenizing*, *labeling*, dan *feature extraction*. Hasil pre-processing dituliskan sebagai berikut :

a. Data Cleansing

Pada tahap *data cleansing* hal pertama yang dilakukan oleh penulis adalah menghapus kolom atau atribut yang tidak digunakan untuk analisis sentimen.



Gambar 2. Hasil Penghapusan Kolom

Setelah melakukan penghapusan kolom, dilakukannya penghapusan angka perubahan hasil ulasan setelah penghapusan angka dituliskan pada tabel berikut.

TABEL 3. HASIL PENGHAPUSAN ANGKA

Sebelum Penghapusan Angka	Setelah Penghapusan Angka
Bintang 1 dulu, Konfirmasi kode OTP beberapa kali salah, disuruh cek jaringan terus padahal jaringan bagus. Pas sudah bisa masuk, mau update profil,, sudah di isi. Gagal simpan terus. Dicoba	Bintang dulu, Konfirmasi kode OTP beberapa kali salah, disuruh cek jaringan terus padahal jaringan bagus. Pas sudah bisa masuk, mau update profil,, sudah di isi. Gagal simpan terus. Dicoba

Kemudian dilanjutkan dengan menghapus simbol dan tanda baca pada teks. Hasil dari proses penghapusan simbol dan tanda baca dituliskan pada tabel 4.

TABEL 4. HASIL PENGHAPUSAN SIMBOL DAN TANDA BACA

Sebelum Penghapusan Simbol	Setelah Penghapusan Simbol
Bintang dulu, Konfirmasi kode OTP beberapa kali salah, disuruh cek jaringan terus padahal jaringan bagus. Pas sudah bisa masuk, mau update profil,, sudah di isi. Gagal simpan terus. Dicoba beberapa kali tetap gagal. .. Mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus, mungkin bisa diperbaiki lagi baik dari server dan aplikasinya. Jika aplikasi ini sukses secara Nasional, maka akan bagus. Karena akan menghilangkan pungli yg ada di daerah yg jauh dri pusat.	Bintang dulu Konfirmasi kode OTP beberapa kali salah disuruh cek jaringan terus padahal jaringan bagus Pas sudah bisa masuk mau update profil sudah di isi Gagal simpan terus Dicoba beberapa kali tetap gagal Mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus mungkin bisa diperbaiki lagi baik dari server dan aplikasinya. Jika aplikasi ini sukses secara Nasional maka akan bagus Nasional maka akan bagus Karena akan menghilangkan pungli yg ada di daerah yg jauh dri pusat

b. Case Folding

Setelah melakukan tahap *data cleansing*, penulis melalui tahap *case folding*. Pada tahap ini penulis mengubah seluruh teks pada dokumen menjadi *lowercase*. Hasil dari tahap *case folding* di tuliskan pada tabel 5.

TABEL 5. HASIL TAHAP CASE FOLDING

Sebelum Case Folding	Setelah Case Folding
----------------------	----------------------

Bintang dulu Konfirmasi kode OTP beberapa kali salah disuruh cek jaringan terus padahal jaringan bagus Pas sudah bisa masuk mau update profil sudah di isi Gagal simpan terus Dicoba beberapa kali tetap gagal Mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus mungkin bisa diperbaiki lagi baik dari server dan aplikasinya Jika aplikasi ini sukses secara Nasional maka akan bagus Karena akan menghilangkan pungli yg ada di daerah yg jauh dri pusat	bintang dulu konfirmasi kode otp beberapa kali salah disuruh cek jaringan terus padahal jaringan bagus pas sudah bisa masuk mau update profil sudah di isi gagal simpan terus dicoba beberapa kali tetap gagal mungkin servernya belum terlalu kuat untuk diakses oleh banyak pengguna sekaligus mungkin bisa diperbaiki lagi baik dari server dan aplikasinya jika aplikasi ini sukses secara nasional maka akan bagus karena akan menghilangkan pungli yg ada di daerah yg jauh dri pusat
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Pada tahap ini penulis juga melakukan penghapusan data teks yang duplikat dari data semula yang didapatkan sejumlah 1200 berubah menjadi 1109 data.

c. *Stopword Removal*

Setelah melalui tahap *case folding*, penulis melalui tahap *stopword removal* yang bertujuan untuk menghilangkan kata kata yang tidak bermakna. Penulis menggunakan library NLTK untuk menyediakan kata kata *stopword* dalam bahasa Indonesia dan juga penulis menentukan *stopword* tambahan. List dari *stopword* tambahan ditunjukkan pada tabel 6.

TABEL 6. STOPWORD TAMBAHAN

List <i>stopword</i> tambahan
['sim','yg','lhaaa','gimana','gmn','bs','mw','mau','klo','kalo','klw','t','tdk','ndak','nggak','jd','jdi','jadi','a','b','c','ktp','psikotes','lg','ektp','lagi','lgi','jg','juga','jga','bgt','stnk','dgn','dengan','x','nya','krn','mgkn','soal','soalnya','sangat','sgt','bagi','sekali','untuk','hadeuuh','pkb','utk','udh','sdh','sudah','ribu','rb','hnya','hanya','hny','dlu','dulu','dl','iy','trs','trus','norek','bni','bca','bank','polda','polri','eh','stnk','ppkm','terus','ga','ada','gk','tp','pdhl','sih','ya','sy','saya','min','lgsg','terooooos','pke','va','cm','cuma','ema','erikkes','eppi','korlantas','tsb','ggk']

d. *Stemming dan Tokenizing*

Setelah melalui tahap *stopword removal*, penulis melakukan tahap *stemming* dan *tokenizing*. *Stemming* digunakan untuk merubah kata menjadi kata dasar menggunakan bantuan library Sastrawi dan *tokenizing*

digunakan untuk menjadikan kalimat dipisah per kata menggunakan library NLTK. Hasil dari proses ini ditunjukkan pada tabel 7.

TABEL 7. HASIL STEMMING DAN TOKENIZING

Sebelum Stemming&Tokenizing	Setelah Stemming&Tokenizing
aplikasi tidak bisa dipakai misalkan klik eppi tidak bisa klik erikkes juga tidak bisa jika belum sempurna aplikasi nya lebih baik tidak usah di launching jika begini kita tetep harus ke satpas didaerah rumah kita	['aplikasi', 'pakai', 'misal', 'klik', 'klik', 'sempurna', 'aplikasi', 'lebih', 'baik', 'usah', 'launching', 'begini', 'tetep', 'satpas', 'daerah', 'rumah']

Setelah melalui tahap tersebut data akan di *untokenize* menjadi kalimat seutuhnya hasil ditunjukkan pada tabel 8.

TABEL 8. HASIL UNTOKENIZE

Hasil Untokenize
aplikasi pakai misal klik klik sempurna aplikasi lebih baik usah launching begini tetep satpas daerah rumah

e. *Labeling*

Setelah melalui tahap data *cleansing*, *case folding*, *stopword removal*, *stemming* dan *tokenizing*, data dipersiapkan untuk diberi label atau *labeling*. *Labeling* menggunakan library *TextBlob*, namun sebelum data diberi label data perlu diterjemahkan ke bahasa Inggris. Hal ini dikarenakan library *TextBlob* hanya dapat berfungsi pada teks berbahasa Inggris. Penulis menerjemahkan teks menggunakan library *googlettranslate*. Hasil teks terjemahan ditunjukkan pada tabel 9.

TABEL 9. STOPWORD TAMBAHAN

Hasil Terjemahan
Translated(src=id, dest=en, text=Using the application, for example, clicking on the perfect click on the application, it's better not to launch like this, but keep the security guard at home, pronunciation=Using the application, for example, clicking on the perfect click on the application, it's better not to launch like this, but keep the security guard at home, extra_data="{ 'translat...")

Hasil dari terjemahan menggunakan *google-trans* menghasilkan data terjemahan yang tidak bersih. Terdapat kalimat yang tidak dibutuhkan seperti *Translated* (src=id,

dest=en, text=. Selain itu juga terdapat data yang duplikat dan adanya tanda baca. Sehingga diperlukan proses pembersihan data kembali. Penulis menggunakan Microsoft excel untuk melakukan pembersihan manual pada data. Sehingga didapatkannya hasil terjemahan data yang bersih yang dituliskan pada tabel 10.

TABEL 10. HASIL DATA TERJEMAHAN

Hasil Clean Translate
using the application for example clicking on the perfect click on the application its better not to launch like this but keep the security guard at home

Setelah melalui proses penerjemahan, data siap untuk diberi label menggunakan library TextBlob. TextBlob memberikan label berdasarkan nilai polaritas pada suatu teks. Apabila nilai polaritas lebih dari 0, maka TextBlob akan melabelkan teks tersebut menjadi positif. Sedangkan jika nilai polaritas kurang dari 0, maka TextBlob akan melabelkan teks menjadi negatif. Hasil dari pemberian label menggunakan library TextBlob dituliskan pada tabel 11.

TABEL 11. HASIL LABELING MENGGUNAKAN TEXTBLOB

Teks	Nilai Polaritas	Label
using the application for example clicking on the perfect click on the application its better not to launch like this but keep the security guard at home	0.75	Positif
reg failed to enter no hp rpc time out just got massage otp timer too fast to connect otpnya otp late entry good bug	0.025	Positif
the application may be fully ready if you want to verify the face photo you have followed the instructions for the results receive a cross see the verification profile have tried several times verify that the application is good if its ready dont launch it	0.22	Positif
the implementation of the national server is very difficult to enter the time out and continue to review further see all one of the forms of go digital answer the need for a future lack of potential for extortion please commit to a	-0.13	Negatif

national scale development server please		
is it simple long enough for a cell phone complicated process file support install the application register via the web the site is different the total application the site still comes checks the conditions match what the media hype	- 0.09166666667	Negatif

Berdasarkan labeling menggunakan TextBlob didapatkannya 589 label positif dan 520 label negative.

```
32] print(data_trans["Label"].value_counts())

Positif    589
Negatif    520
Name: Label, dtype: int64
```

Gambar 3. Jumlah Hasil Labeling

f. Feature Extraction

Setelah melakukan labeling pada data, dilakukannya feature extraction menggunakan TF-IDF. Pada tahap ini menggunakan library scikit-learn, untuk melakukan feature extraction dengan TF-IDF. TF-IDF diterapkan pada data agar diperoleh data yang relevan. TF-IDF memiliki parameter min\_df yang berfungsi memfilter kata, pada penelitian ini penulis menggunakan min\_df=3, sehingga kata yang akan diproses adalah kata yang muncul minimal pada tiga dokumen. Hasil dari TF IDF ditunjukkan pada gambar 4.

Gambar 4. Hasil TF-IDF

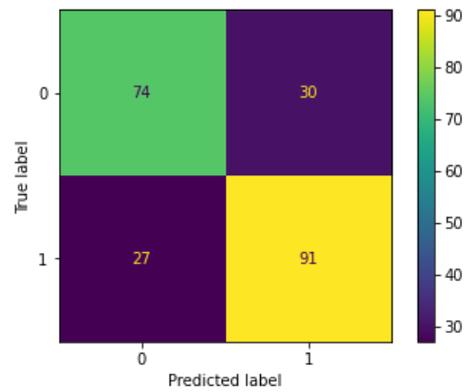
C. Pembuatan Model

Setelah melalui beberapa tahap dari pre-processing, penulis melanjutkan ke tahap pembuatan model untuk klasifikasi sentimen. Namun dalam pembuatan model, penulis merubah label pada data menjadi angka (label encoding) untuk





Gambar 8. Wordcloud Ulasan Negatif.



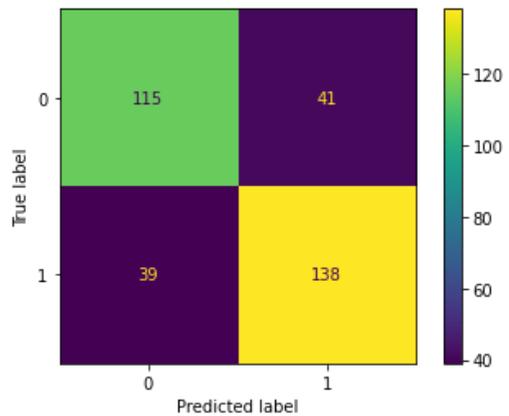
Gambar 10. Confusion Matrix 80:20

Hasil pengujian *confusion matrix* pada skenario dengan rasio data 80% data latih: 20 % data uji ditunjukkan pada gambar 4.9. Dari *confusion matrix* tersebut didapatkannya nilai precision, nilai recall, dan nilai f1-score yang ditunjukkan pada tabel 15. dan nilai accuracy yang dihasilkan adalah 0.74.

TABEL 15. HASIL SKENARIO 80:20

Skenario	Label	Precision	Recall	F1-score
80:20	Negatif	0.73	0.71	0.72
	Positif	0.75	0.77	0.76

c. Hasil *Confusion Matrix* Skenario 70:30



Gambar 11. Confusion Matrix 70:30

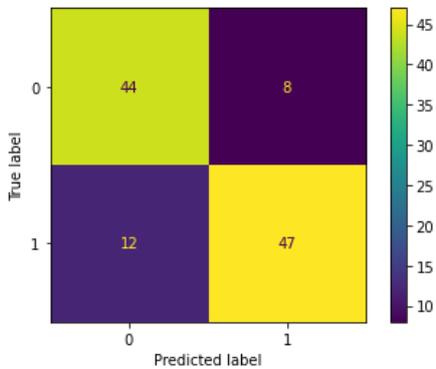
Hasil pengujian *confusion matrix* pada skenario dengan rasio data 70% data latih: 30 % data uji ditunjukkan pada gambar 4.9. Dari *confusion matrix* tersebut didapatkannya nilai precision, nilai recall, dan nilai f1-score yang ditunjukkan pada tabel 16. dan nilai accuracy yang dihasilkan adalah 0.76.

TABEL 16. HASIL SKENARIO 70:30

Skenario	Label	Precision	Recall	F1-score
70:30	Negatif	0.73	0.71	0.72
70:30	Positif	0.75	0.77	0.76

E. Pengujian

a. Hasil *Confusion Matrix* 90:10



Gambar 9. *Confusion Matrix* 90:10

Hasil pengujian *confusion matrix* pada skenario dengan rasio data 90% data latih : 10 % data uji ditunjukkan pada gambar 4.8. Dari *confusion matrix* tersebut didapatkannya nilai precision, nilai recall, dan nilai f1-score yang ditunjukkan pada tabel 4.14 dan nilai *accuracy* yang dihasilkan adalah 0.82

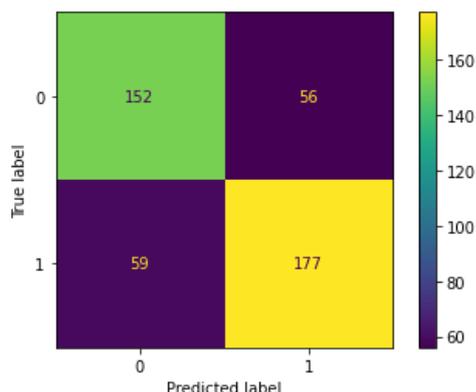
TABEL 14. HASIL SKENARIO 90:10

Skenario	Label	precision	recall	F1-score
90:10	Negatif	0.79	0.85	0.81
	Positif	0.85	0.80	0.82

b. Hasil *Confusion Matrix* 80:20

70:30	Negatif	0.75	0.74	0.74
	Positif	0.77	0.78	0.78

d. Hasil Confusion Matrix Skenario 60:40



Gambar 11. Confusion Matrix 60:40

Hasil pengujian pada skenario dengan rasio data 60% data latih: 40 % data uji ditunjukkan pada gambar 3.9. Dari confusion matrix tersebut didapatkannya nilai *precision*, *recall*, dan *f1-score* yang ditunjukkan pada tabel 4.17. dan nilai *accuracy* yang dihasilkan adalah 0.74.

TABEL 17. HASIL SKENARIO 60:40

Skenario	Label	Precision	Recall	F1-score
60:40	Negatif	0.72	0.73	0.73
	Positif	0.76	0.75	0.75

F. Analisis Hasil

Berdasarkan hasil pengujian model SVM dengan parameter *probability* = true, *kernel* = linear , *C* = 10, *gamma* = auto. Model SVM yang memiliki kinerja paling baik pada skenario pertama yakni dengan rasio data 90 : 10. Pada skenario pertama memiliki hasil evaluasi paling tinggi. Hasil evaluasi pada label negatif menghasilkan nilai *precision* sebesar 0.79, nilai *recall* sebesar 0.85, nilai *f1-score* sebesar 0.81. Sedangkan hasil evaluasi pada label positif adalah nilai *precision* sebesar 0.85, nilai *recall* sebesar 0.80, dan nilai *f1 score* sebesar 0.82. Pada skenario rasio data 90:10 juga menghasilkan nilai *accuracy* tertinggi dari pada skenario lainnya yakni 0.82.

Sedangkan model dengan kinerja terburuk ada pada model SVM dengan skenario rasio data 80 : 20 dan skenario rasio data 60 : 40. Hasil Akurasi yang dihasilkan pada dua skenario tersebut adalah 0,74. Nilai *precision*, *recall*, dan *F1-Score* yang dihasilkan skenario rasio data 80:20 pada label negatif adalah 0.73, 0.71, 0.72 dan pada label positif adalah 0.75, 0.77, 0.76. Sedangkan nilai *precision*, *recall*, dan *F1-Score* yang dihasilkan skenario 60 : 40 pada label negatif adalah 0.72, 0.73, 0.73 dan pada label positif adalah 0.76, 0.75 , 0.75.

Hasil evaluasi model SVM terbaik digunakan oleh penulis

untuk klasifikasi sentimen pada teks ulasan aplikasi digital korlantas. Berdasarkan hasil labeling menggunakan TextBlob dan klasifikasi menggunakan SVM memiliki perbedaan pada jumlah sentimen. Pada proses labeling menggunakan TextBlob menghasilkan 589 sentimen positif dan 520 sentimen negatif. Sedangkan hasil klasifikasi sentimen menggunakan SVM menghasilkan 598 sentimen positif dan 511 sentimen negatif. Hasil dari klasifikasi sentimen pada ulasan dituliskan pada table 18.

TABEL 18. HASIL KLASIFIKASI PADA TEKS ULASAN

No	Teks Ulasan	Hasil Label TextBlob	Hasil Label SVM
1	using the application for example clicking on the perfect click on the application its better not to launch like this but keep the security guard at home	Positif	Positif
2	reg failed to enter no hp rpc time out just got massage otp timer too fast to connect otpnya otp late entry good bug	Positif	Positif
3	the application may be fully ready if you want to verify the face photo you have followed the instructions for the results receive a cross see the verification profile have tried several times verify that the application is good if its ready dont launch it	Positif	Positif
4	the implementation of the national server is very difficult to enter the time out and continue to review further see all one of the forms of go digital answer the need for a future lack of potential for extortion please commit to a national scale development server please	Negatif	Negatif
5	entering the cellphone number always appears so its wrong try a few it usually appears the rpc timeout is correct what application do you use dont launch it	Negatif	Negatif
6	is it simple long enough for a	Negatif	Negatif

	cell phone complicated process file support install the application register via the web the site is different the total application the site still comes checks the conditions match what the media hype		
7	the result of entering the input number always fails manual input method scan barcode save always account verification results complete validation all try to log out re login server full effect launching trial just now application	Negatif	Negatif
8	even though the beta mode will definitely launch too many obstacles smart people when the application loses neighbors please pull the application sir its good the main bug is minimized send the otp complete with the new launch break through good ready responsibility spirit sorry star	Positif	Positif
9	just installed the cellphone register then sent the otp code and sent it but still doesnt respond the otp code comes in give it one star fix it dont be too hasty the perfect application the perfect launch a lot of disappointment the application has a lot of perfect errors just launched	Positif	Positif
10	difficult procedure even taking photos when wearing glasses is clearly difficult many stages of the application difficult access lot d perfect new system launching starting trying afternoon hours success going easy instead of hassle	Positif	Positif

#### IV. KESIMPULAN

Setelah melalui banyak tahapan untuk melakukan analisis sentimen pada ulasan pengguna aplikasi digital

korlantas, didapatkannya kesimpulan sebagai berikut:

1. Algoritma SVM dapat melakukan analisis sentimen pada ulasan aplikasi digital korlantas dengan hasil 598 sentimen positif dan 511 sentimen negatif.
2. Berdasarkan hasil pengujian model SVM memiliki kinerja yang baik pada skenario rasio data 90:10 dengan nilai akurasi sebesar 0.82 dan model SVM dengan kinerja terburuk ada pada skenario rasio data 80:20 dan 60:40 yakni dengan akurasi sebesar 0.74.

#### DAFTAR PUSTAKA

- [1] S. Wahyu, "Era Digital dan Tantangannya," *Seminar Nasional Pendidikan*, pp. pp. ISBN.978-602-50088-0-1, 2017.
- [2] S. A. Robbianyah, "Penerapan Metode EUCS Terhadap Keuasan Pengguna Layanan SINAR pada Aplikasi Digital Korlantas POLRI," *Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya*, 2022.
- [3] T. Annisa, "Mengenal Peran Sentiment Analysis Beserta Cara Kerjanya," 2021.
- [4] P. A. N. A. a. R. M. T. Meisya, "Perbandingan Kernel Support Vector Machine (SVM) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19," *SINTECH J*, pp. 139-145, 2021.
- [5] Akbar, "Sentiment Analysis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine (SVM)," *Journal of Artificial Intelligence & Data Science*, pp. Vol. 2 No. 2 p-ISSN: 2746-9204. e-ISSN: 2746-9190, 2022.
- [6] M. R. Fahlevvi, "Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine," *Jurnal Teknologi dan Komunikasi*, 2022.
- [7] R. Wahyudi and G. Kusumawardhana, "Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *JURNAL INFORMATIKA*, Vol. 8 No. 2, pp. Halaman 200-207 ISSN: 2355-6579 | E-ISSN: 2528-224, 2021.
- [8] A. F. Rozi and A. S. Purnomo, "Analisis Sentimen Untuk Respon Masyarakat Terhadap Universitas (Studi Kasus : Universitas Mercu Buana Yogyakarta)," *Jurnal Of Information System And Artificial Intelligence (JISAI)*, 2020.
- [9] M. Abdulrahman, "The Impact of Electronic Word of Mouth on Consumers Purchasing Decisions," *International Journal of Computer Applications*, 2018.
- [10] M. S. Utomo, "Implementasi PHP Sebagai Penghasil Konten Otomatis Pada Halaman Situs," *Jurnal Teknologi Informasi Dinamik*, pp. 147-153, 2022.
- [11] A. Josi, "Penerapan Teknik Web Scraping Pada Mesin Pencari Artikel Ilmiah," 2018.
- [12] S. Arthur, "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal*, 2017.
- [13] T. A. Mitchell, "Machine Learning," *McGraw-Hill*, 2017.
- [14] Samuel, "Implementasi metode K-Nearest Neighbor dengan decision rule untuk klasifikasi subtopik berita," *Jurnal Informatika*, 2018.
- [15] D. Pratama, "Analisis Permasalahan Perangkat Jaringan Menggunakan Metode Algoritma K-Means Dan Kmedoids," *Program Studi Teknik Informatika Fakultas Ilmu Komputer. Universitas Mercu Buana*, 2021.
- [16] F. E. d. K. Prabowo, "Analisis Prediksi Masa Studi Mahasiswa Menggunakan Algoritma Naïve Bayes," *Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana*, 2019.
- [17] Y. E. Alamarani, "Sentiment Analysis Using Hybrid Method Of

- Support Vector Machine And Decision Tree," *Journal of Theoretical & Applied Information*, 2018.
- [18] A. Kowalczyk, "Support Vector Machines Succinctly," *Syncfusion Inc*, 2017.
- [19] R. A. Tiniges, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *Jurnal Media Informatika Budidarma*, 2020.
- [20] T. Wahyono, *Fundamental of Python for Machine Learning: Dasar-Dasar Pemrograman Python untuk Machine Learning dan Kecerdasan Buatan*, Yogyakarta: Gava Media, 2018.
- [21] E. R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insani Ict Journal*, pp. 156-165, 2020.