

Perbandingan Algoritma *Machine learning* Untuk Analisis Sentimen Aplikasi Instagram Berbasis *NLP*

Yuridis Seri Rahmat¹, Saskia Renata², Albertus Belo³, Weiskhy Steven Dharmawan⁴, Riski Annisa⁵, Lady Agustin Fitriana^{6*}

^{1,2,3,5} Program Studi Informatika, Universitas Bina Sarana Informatika, Pontianak, Indonesia

⁴ Program Studi Sistem Informasi Akuntansi, Universitas Bina Sarana Informatika, Pontianak, Indonesia

⁶ Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Pontianak, Indonesia

Email: ¹yuridis410@gmail.com, ²y91c.0735@gmail.com, ³albertusbelo964@gmail.com, ⁴weiskhy.wvn@bsi.ac.id, ⁵riski.rnc@bsi.ac.id, ⁶Lady.lag@bsi.ac.id

(*Email Corresponding Author: Lady.lag@bsi.ac.id)

Received: May 20, 2026 | Revision: May 25, 2026 | Accepted: May 30, 2026

Abstrak

Instagram merupakan salah satu platform media sosial terbesar di dunia dengan lebih dari dua miliar pengguna aktif bulanan. Ulasan pengguna di Google Play Store menjadi sumber data yang kaya untuk memahami persepsi dan kepuasan pengguna terhadap fitur serta layanan yang diberikan. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi Instagram menggunakan tiga algoritma klasifikasi *machine learning*, yaitu *Support Vector Machine (SVM)*, *Random Forest (RF)*, dan *Neural Network (NN)*, dengan representasi fitur menggunakan metode *Bag-of-Words (BoW)*. Data yang dikumpulkan sebanyak 10.000 ulasan melalui teknik *web scraping* dari Google Play Store menggunakan *library google-play-scraper*, dan setelah melalui proses *preprocessing* diperoleh 7.236 data bersih yang digunakan untuk pemodelan. Pelabelan data dilakukan secara otomatis menggunakan leksikon *InSet (Indonesia Sentiment Lexicon)*. Tahapan *preprocessing* meliputi pembersihan teks, *case folding*, normalisasi kata tidak baku, tokenisasi, *stopword removal*, dan *stemming* menggunakan algoritma Sastrawi. Hasil eksperimen menunjukkan bahwa *SVM* memperoleh akurasi tertinggi sebesar 90,7%, diikuti *Neural Network* 89,2%, dan *Random Forest* 81,5%. Berdasarkan hasil tersebut, *SVM* terbukti paling unggul dalam mengklasifikasikan sentimen positif, negatif, dan netral pada ulasan pengguna aplikasi Instagram.

Kata Kunci: Analisis Sentimen, Instagram, *Neural Network*, *Random Forest*, *Support Vector Machine*

Abstract

Instagram is one of the largest social media platforms in the world with more than two billion monthly active users. User reviews on Google Play Store represent a rich data source for understanding user perception and satisfaction regarding provided features and services. This study aims to analyze sentiment of Instagram application user reviews using three machine learning classification algorithms, namely *Support Vector Machine (SVM)*, *Random Forest (RF)*, and *Neural Network (NN)*, with *Bag-of-Words (BoW)* feature representation. A total of 10,000 reviews were collected through web scraping techniques from Google Play Store using the *google-play-scraper* library, and after preprocessing, 7,236 clean data records were used for modeling. Data labeling was performed automatically using the *InSet (Indonesia Sentiment Lexicon)*. Preprocessing stages included text cleaning, *case folding*, non-standard word normalization, tokenization, *stopword removal*, and *stemming* using the Sastrawi algorithm. Experimental results show that *SVM* achieved the highest accuracy of 90.7%, followed by *Neural Network* 89.2%, and *Random Forest* 81.5%. Based on these results, *SVM* proved most superior in classifying positive, negative, and neutral sentiments in Instagram application user reviews.

Keywords: Sentiment Analysis, Instagram, *Neural Network*, *Random Forest*, *Support Vector Machine*

1. PENDAHULUAN

Perkembangan platform media sosial dalam satu dekade terakhir telah mengubah cara masyarakat berkomunikasi, berbagi informasi, dan mengekspresikan pendapat secara fundamental. Instagram, yang diluncurkan pada tahun 2010 dan diakuisisi oleh Meta pada tahun 2012, kini menjadi salah satu platform media sosial paling dominan di dunia dengan lebih dari dua miliar pengguna aktif bulanan. Di Indonesia, Instagram menempati posisi sebagai salah satu aplikasi dengan jumlah unduhan terbanyak di *Google Play Store*, yang mencerminkan tingginya adopsi platform ini di berbagai kalangan masyarakat untuk kebutuhan personal maupun bisnis. Seiring dengan masifnya penggunaan tersebut, ulasan pengguna pada platform distribusi aplikasi menjadi sumber informasi yang sangat berharga. Pengguna secara aktif meninggalkan ulasan yang berisi opini, keluhan teknis, dan saran terkait pembaruan fitur serta pengalaman pengguna secara keseluruhan (*user experience*).

Namun, muncul permasalahan utama di mana volume ulasan yang masuk setiap harinya mencapai ribuan data, sehingga proses analisis secara manual tidak lagi praktis dan efisien untuk dilakukan oleh pengembang aplikasi.

Keterbatasan sumber daya manusia dalam membaca dan mengklasifikasikan setiap ulasan menyebabkan banyak umpan balik penting yang terlewatkan. Selain kendala volume, tantangan signifikan muncul dari karakteristik bahasa Indonesia yang digunakan di media sosial, yang sering kali melibatkan penggunaan istilah slang, singkatan unik, serta tata bahasa tidak baku yang sulit diinterpretasikan secara konsisten oleh manusia dalam skala besar. Oleh karena itu, diperlukan sebuah solusi otomatis menggunakan teknik *text mining* dan *machine learning* untuk mengekstraksi wawasan (*insight*) dari data ulasan tersebut secara cepat dan akurat. Analisis sentimen, sebagai salah satu cabang dari *Natural Language Processing (NLP)*, hadir sebagai solusi untuk mengidentifikasi dan mengklasifikasikan polaritas opini pengguna ke dalam kategori positif, negatif, atau netral.

Beberapa penelitian terdahulu telah mengeksplorasi penggunaan *machine learning* untuk analisis sentimen pada media sosial. Penelitian oleh Hayuningtyas et al. 2025 [1] menerapkan metode *Support Vector Machine (SVM)* untuk menganalisis sentimen pengguna Instagram di *Google Play Store* dan berhasil membuktikan bahwa *SVM* efektif dalam menangani data teks media sosial. Selanjutnya, Surya Sayogo et al. 2023 [2] melakukan penelitian serupa menggunakan algoritma *Naïve Bayes* pada 1.100 data ulasan Instagram dan memperoleh akurasi sebesar 89%. Di sisi lain, Zaiem Praghakusma et al. 2021 [3] membandingkan berbagai fungsi *kernel* pada *SVM* untuk analisis sentimen Instagram dan Twitter, di mana *kernel linear* ditemukan sebagai yang paling optimal dengan akurasi 83,06%. Pendekatan berbasis *deep learning* juga diuji oleh Cahyadi et al. 2020 [4] menggunakan *Recurrent Neural Network (RNN)* dengan *Long Short-Term Memory (LSTM)* yang menunjukkan keunggulan dalam memahami konteks teks sekuensial pada data Instagram. Terakhir, penelitian oleh Latif et al. 2026 [5] memberikan landasan komparatif dengan membandingkan performa *SVM*, *Random Forest*, dan *Neural Network* pada ulasan aplikasi di *Google Play Store*, yang menyimpulkan bahwa *SVM* cenderung memberikan hasil yang paling stabil dibandingkan algoritma lainnya. Penelitian Agustin Fitriana et al. 2026 [6] membandingkan CNN dan LSTM pada 5.000 ulasan aplikasi Roblox dari *Google Play Store* berbahasa Indonesia menggunakan pelabelan berbasis leksikon InSet dan *stemming* Sastrawi, dengan hasil CNN unggul pada akurasi 89%. Namun, penelitian tersebut hanya membandingkan dua arsitektur *deep learning* tanpa menyertakan algoritma klasik seperti *SVM* dan *Random Forest* sebagai pembanding, sehingga belum diketahui bagaimana performa relatif pendekatan *machine learning* tradisional pada domain dan bahasa yang sama.

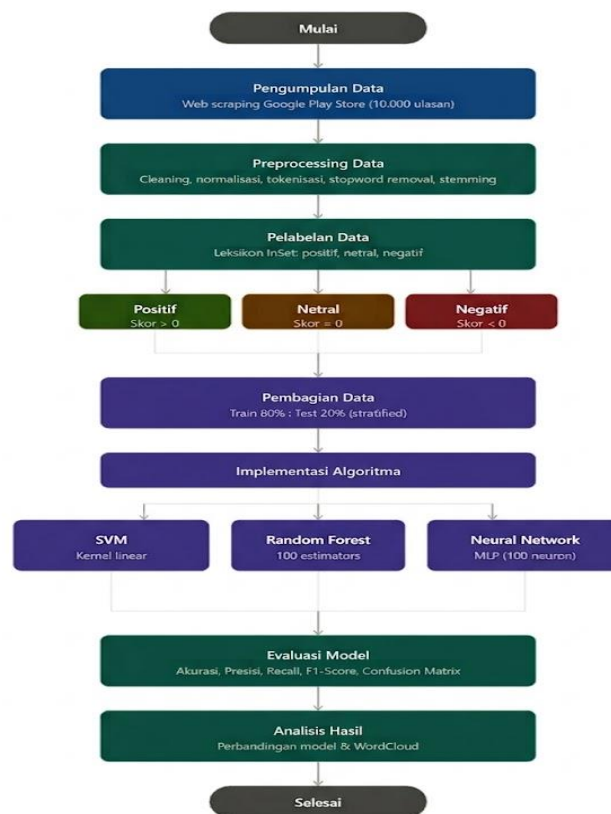
Dari kajian penelitian-penelitian tersebut, terdapat beberapa *research gap* yang belum terjawab dalam literatur yang ada. Belum ada penelitian yang secara khusus membandingkan tiga algoritma *SVM*, *Random Forest*, dan *Neural Network* secara bersamaan pada dataset ulasan aplikasi Instagram berbahasa Indonesia dari *Google Play Store*. Sebagian besar penelitian tidak menerapkan *pipeline preprocessing NLP* Bahasa Indonesia yang menyeluruh, khususnya normalisasi kata tidak baku dan *stemming* berbasis Sastrawi secara terintegrasi. Penggunaan kamus InSet sebagai metode pelabelan sentimen otomatis tiga kelas pada ulasan aplikasi Instagram belum banyak dieksplorasi. Perbandingan representasi fitur *Bag of Words (CountVectorizer)* pada ketiga algoritma tersebut di atas dalam konteks ulasan aplikasi media sosial berbahasa Indonesia masih sangat terbatas.

Berdasarkan gap tersebut, penelitian ini hadir dengan *novelty* pada kombinasi yang belum pernah dilakukan secara bersamaan, yaitu penggunaan data ulasan aplikasi Instagram dari *Google Play Store* berbahasa Indonesia dalam jumlah besar hingga 10.000 data. *Pipeline preprocessing NLP* Bahasa Indonesia yang lengkap mencakup normalisasi kata tidak baku, *stopword removal* NLTK, dan *stemming* Sastrawi. Pelabelan sentimen otomatis tiga kelas menggunakan InSet, serta evaluasi komparatif yang sistematis antara *SVM*, *Random Forest*, dan *Neural Network* dengan representasi fitur *CountVectorizer*.

Penelitian ini hadir untuk mengisi celah tersebut dengan melakukan pengujian menyeluruh terhadap tiga arsitektur algoritma yang berbeda: *SVM* yang berbasis *hyperplane*, *Random Forest* yang berbasis *ensemble tree*, dan *Neural Network* yang berbasis neuron berlapis. Dengan menggunakan representasi fitur *Bag-of-Words (BoW)* dan dataset yang lebih besar, penelitian ini bertujuan untuk memberikan gambaran yang lebih akurat mengenai algoritma mana yang paling tangguh dalam menangani data teks media sosial yang bersifat *sparse* dan memiliki dimensi fitur tinggi. Selain itu, integrasi visualisasi *WordCloud* dan analisis frekuensi kata dilakukan untuk memberikan dimensi analisis yang lebih dalam, tidak hanya dari sisi akurasi teknis, tetapi juga substansi pesan yang disampaikan oleh pengguna Indonesia. Diharapkan, hasil dari studi ini dapat memberikan kontribusi praktis bagi pengembang perangkat lunak dalam memetakan tren kepuasan pengguna secara ilmiah, serta memberikan rekomendasi metodologis yang solid bagi penelitian analisis sentimen di masa depan pada domain media sosial Indonesia.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode komparatif antara tiga algoritma klasifikasi *machine learning*. Alur penelitian terdiri dari beberapa tahapan, yaitu pengumpulan data, *preprocessing* teks, pelabelan teks, pembagian data, implementasi algoritma, evaluasi model, dan analisis hasil. Berikut penjelasan masing-masing tahapan seperti gambar 1.



Gambar 1. Alur penelitian

2.1 Pengumpulan Data

Penelitian ini menggunakan *library google-play-scrapers* untuk mengumpulkan total 10.000 ulasan pengguna berbahasa Indonesia, Proses scraping dilakukan pada halaman ulasan aplikasi Instagram (*'com.instagram.android'*) di Google Play Store dengan parameter *Sort.NEWEST* untuk memastikan data yang dikumpulkan merupakan ulasan terbaru. Data yang dikumpulkan meliputi teks ulasan (*content*), *rating* bintang (*score*), nama pengguna (*userName*), dan tanggal ulasan (*at*).

2.2 Preprocessing

Tahap *preprocessing* merupakan langkah krusial untuk meningkatkan kualitas representasi data teks sebelum proses ekstraksi fitur. *Preprocessing* diawali dengan pembersihan teks (*text cleaning*) menggunakan ekspresi reguler untuk menghapus elemen-elemen yang tidak berkontribusi terhadap analisis, seperti *URL*, *username (@mention)*, *tag HTML*, emoji, simbol, serta angka[7].

Selanjutnya dilakukan *case folding* dengan mengubah seluruh teks menjadi huruf kecil (*lowercase*) untuk menyeragamkan representasi kata[8]. Tahap berikutnya adalah normalisasi kata tidak baku, yaitu mengganti kata-kata slang dan singkatan dengan bentuk bakunya menggunakan kamus kata baku yang diperoleh dari repositori GitHub. Setelah normalisasi, dilakukan tokenisasi untuk memecah teks menjadi unit kata individual.

Proses dilanjutkan dengan *stopword removal*, yaitu penghapusan kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen. Daftar *stopword* mengacu pada *library NLTK* untuk bahasa Indonesia yang diperkaya dengan *custom stopwords* tambahan. Tahap terakhir adalah *stemming* menggunakan algoritma Sastrawi untuk mengubah kata menjadi bentuk dasarnya[9]. Seluruh rangkaian *preprocessing* ini menghasilkan kolom *stemming_data* yang digunakan sebagai input pada proses ekstraksi fitur.

2.3 Pelabelan Data

Pelabelan data dilakukan secara otomatis menggunakan pendekatan berbasis leksikon *InSet (Indonesia Sentiment Lexicon)*, yaitu kamus sentimen bahasa Indonesia yang mengelompokkan kata ke dalam kategori positif dan negatif beserta skor polaritasnya. Pendekatan ini menghitung skor sentimen setiap ulasan berdasarkan selisih antara jumlah kata positif dan kata negatif yang ditemukan dalam teks ulasan.

Ulasan dengan skor sentimen positif (> 0) dikategorikan sebagai sentimen positif, skor negatif (< 0) sebagai sentimen negatif, dan skor nol sebagai sentimen netral. Dari 10.000 data mentah yang terkumpul, setelah melalui proses penghapusan duplikat dan pembersihan nilai kosong pada tahap *preprocessing*, diperoleh 7.236 data bersih yang digunakan untuk pelabelan dan pemodelan.

2.4 Pembagian Data

Dataset yang telah diberi label selanjutnya dibagi menjadi data latih (*training data*) dan data uji (*testing data*) menggunakan fungsi *train_test_split* dari *library scikit-learn* dengan rasio 80:20 dan parameter *random_state* untuk memastikan konsistensi dan reproduktibilitas hasil pembagian pada setiap pengujian.

Dengan proporsi tersebut, dari total 7.236 data bersih diperoleh 5.788 data latih (80%) yang digunakan oleh algoritma untuk mempelajari pola sentimen, serta 1.448 data uji (20%) yang dialokasikan sebagai data baru yang belum pernah dilihat model untuk mengukur tingkat akurasi prediksi. Proporsi 80:20 dipilih karena dianggap ideal dalam menjaga keseimbangan antara kecukupan data untuk proses pembelajaran model dan validitas hasil evaluasi.

2.5 Implementasi Algoritma

Representasi fitur teks dilakukan menggunakan metode *Bag-of-Words (BoW)* yang mengubah setiap ulasan menjadi vektor numerik berdasarkan frekuensi kemunculan kata. Metode ini dipilih karena kesederhanaannya dan efektivitasnya yang telah terbukti pada berbagai tugas klasifikasi teks. Setelah representasi fitur terbentuk, tiga algoritma klasifikasi *machine learning* diimplementasikan secara terpisah dan dibandingkan performanya.

2.5.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi yang bekerja dengan mencari *hyperplane* optimal yang memisahkan kelas-kelas data dengan margin maksimal menurut Tommy Rustandi et al. 2023[10]. Pada penelitian ini digunakan *SVM* dengan *kernel linear* karena representasi *Bag-of-Words* menghasilkan ruang fitur berdimensi tinggi yang bersifat *linearly separable*. Parameter *C* (*regularization*) digunakan untuk mengontrol *trade-off* antara margin maksimal dan kesalahan klasifikasi pada data latih.

2.5.2 Random Forest (RF)

Random Forest adalah algoritma *ensemble* yang membangun sejumlah pohon keputusan secara paralel menggunakan subset data dan fitur yang dipilih secara acak, kemudian mengagregasi prediksi seluruh pohon melalui mekanisme voting mayoritas menurut Qomariyati et al. 2024[11]. Pendekatan *ensemble* ini memberikan ketahanan terhadap *overfitting* dibandingkan pohon keputusan tunggal. Pada penelitian ini digunakan parameter *n_estimators* untuk menentukan jumlah pohon yang dibangun.

2.5.3 Neural Network (NN)

Neural Network yang digunakan adalah *Multi-Layer Perceptron (MLP)* dengan satu lapisan tersembunyi. MLP mampu mempelajari representasi fitur non-linear dari data teks melalui mekanisme propagasi balik (*backpropagation*). Meskipun penelitian Arasy et al. 2025[12] menggunakan ukuran lapisan yang lebih kecil yaitu konfigurasi (10, 20), (10, 8), atau (5, 3). pada studi ini menggunakan Konfigurasi arsitektur meliputi 100 neuron pada lapisan tersembunyi dengan fungsi aktivasi *ReLU*, serta lapisan keluaran dengan fungsi aktivasi *softmax* untuk menghasilkan probabilitas pada tiga kelas sentimen.

2.6 Evaluasi Model

Evaluasi performa ketiga model dilakukan pada data uji (1.448 sampel) menggunakan empat metrik standar klasifikasi teks, yaitu akurasi, presisi, *recall*, dan *F1-Score*. Akurasi mengukur proporsi prediksi yang benar dari keseluruhan data uji. Presisi mengukur ketepatan prediksi positif, sedangkan *recall* mengukur kemampuan model dalam menemukan seluruh sampel yang benar-benar positif. *F1-Score* merupakan rata-rata harmonik dari presisi dan *recall* yang digunakan untuk menilai keseimbangan antara keduanya.

Selain metrik per kelas, digunakan juga *macro average* dan *weighted average* untuk mengevaluasi performa model secara keseluruhan. *Macro average* menghitung rata-rata metrik tanpa memperhitungkan distribusi kelas, sementara *weighted average* memperhitungkan proporsi setiap kelas dalam dataset. Perbandingan ketiga model dilakukan berdasarkan nilai akurasi dan *macro F1-Score* sebagai metrik utama.

2.7 Analisis Hasil

Analisis hasil dilakukan melalui tiga pendekatan komplementer. Pertama, perbandingan performa model secara kuantitatif berdasarkan metrik evaluasi untuk menentukan algoritma yang paling optimal pada dataset ulasan Instagram berbahasa Indonesia. Kedua, analisis visualisasi *WordCloud* untuk mengidentifikasi pola kata dominan pada setiap kelas sentimen sebelum dan sesudah proses *preprocessing*.

Ketiga, analisis frekuensi kata dilakukan untuk memahami topik dan isu yang paling sering disampaikan pengguna Instagram dalam ulasan mereka. Kombinasi ketiga pendekatan analisis ini diharapkan memberikan wawasan yang komprehensif, baik dari sisi teknis perbandingan algoritma maupun dari sisi substansi persepsi pengguna terhadap aplikasi Instagram.

Tahap berikutnya adalah normalisasi kata tidak baku, yaitu proses mengganti kata-kata slang, singkatan, atau variasi penulisan informal dengan bentuk bakunya. Normalisasi dilakukan menggunakan kamus kata baku yang diperoleh dari repositori GitHub, sehingga mampu meningkatkan konsistensi data teks dan meminimalkan ambiguitas makna, Seperti gambar ke 5.

case_folding	normalisasi
aplikasinya bagus gak ada minusnya	aplikasinya bagus tidak ada minusnya
itu ig kalo mau perbarui profil sorotan lama b...	itu ig kalau mau perbarui profil sorotan lama ...
bisa di privasi dari semua akun	bisa di privasi dari semua akun
sangatt menarik untuk menemani hari yang membo...	sangatt menarik untuk menemani hari yang membo...
adakan siaran langsung lagi!!!!	adakan siaran langsung lagi!!!!

Gambar 5. Normalisasi

Setelah proses normalisasi, teks diproses melalui tiga tahap *preprocessing* secara berurutan sebagaimana ditunjukkan pada Gambar 6. Pertama, tokenisasi dilakukan untuk memecah teks menjadi unit kata individual guna memudahkan analisis fitur. Selanjutnya, *stopword removal* diterapkan menggunakan *library* NLTK bahasa Indonesia yang diperkaya dengan *custom stopwords* seperti "mau", "yang", "itu", "hari", dan "kalau" untuk menyaring kata-kata yang tidak bermakna secara analitis. Terakhir, *stemming* menggunakan algoritma Sastrawi dilakukan untuk mengubah setiap kata menjadi bentuk dasarnya, sehingga variasi morfologis dapat diseragamkan[9]. Hasil keseluruhan tahap *preprocessing* ini menghasilkan data teks yang lebih bersih dan terstruktur, seperti yang terlihat pada kolom *stemming data* di Gambar 6.

tokenize	stopword removal	stemming_data
[aplikasinya, bagus, tidak, ada, minusnya]	[aplikasinya, bagus, minusnya]	aplikasi bagus minus
[itu, ig, kalau, mau, perbarui, profil, sorota...]	[ig, perbarui, profil, sorotan, bolak, wifi, d...]	ig baru profil sorot bolak wifi data wifi data...
[bisa, di, privasi, dari, semua, akun]	[privasi, akun]	privasi akun
[sangatt, menarik, untuk, menemani, hari, yang...]	[sangatt, menarik, menemani, membosankan]	sangatt tarik tani bosan
[adakan, siaran, langsung, lagi!!!!]	[adakan, siaran, langsung, lagi!!!!]	adakan siar langsung lagi!!!!

Gambar 6. Tokenisasi, *Stopword removal*, *Stemming Data*

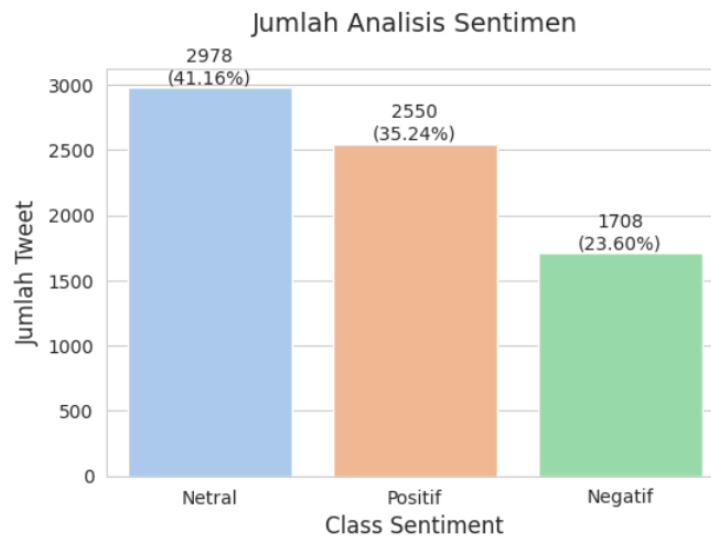
3.3 Distribusi Label Sentimen

Pelabelan data dilakukan secara otomatis menggunakan pendekatan berbasis leksikon *InSet (Indonesia Sentiment Lexicon)*, yaitu kamus sentimen bahasa Indonesia yang mengelompokkan kata ke dalam kategori positif dan negatif beserta skor polaritasnya. Pendekatan ini menghitung skor sentimen setiap ulasan berdasarkan selisih antara jumlah kata positif dan kata negatif yang terdapat dalam kamus leksikon. Ulasan dengan skor sentimen positif (> 0) dikategorikan sebagai sentimen positif; skor negatif (< 0) sebagai sentimen negatif; dan skor nol sebagai sentimen netral. Dari 10.000 data mentah yang terkumpul, setelah melalui proses penghapusan duplikat dan pembersihan nilai kosong pada tahap *preprocessing*, diperoleh 7.236 data bersih yang digunakan untuk pelabelan dan pemodelan. Distribusi hasil pelabelan menggunakan InSet Lexicon dapat dilihat pada Gambar 7, yang menunjukkan perbandingan jumlah ulasan positif dan negatif secara signifikan. Setelah melewati tahap pelabelan menggunakan kamus InSet Lexicon, seluruh data ulasan dikelompokkan ke dalam tiga kategori sentimen, yaitu netral, positif, dan negatif. Berdasarkan hasil klasifikasi yang disajikan pada Gambar 8, kategori sentimen netral mendominasi dengan sebanyak 2.978 ulasan atau sekitar 41,16%

dari total data. Selanjutnya, sentimen positif menempati posisi kedua dengan total 2.550 ulasan atau sebesar 35,24%, sementara sentimen negatif tercatat sebanyak 1.708 ulasan dengan persentase 23,60%.

Rating	steming_data	Score	Sentiment
0	5	aplikasi bagus minus	-2 Negatif
1	1	ig baru profil sorot bolak wifi data wifi data...	0 Netral
2	5	privasi akun	0 Netral
3	5	sangatt tarik tani bosan	1 Positif
4	4	adakan siar langsung lagi!!!!	1 Positif

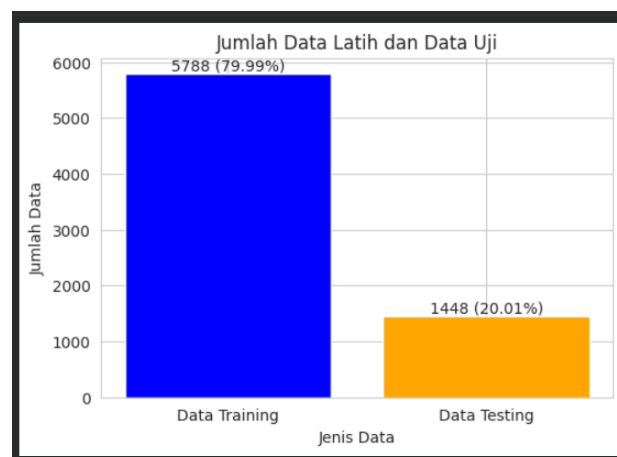
Gambar 7. Labeling Sentiment Leksikon InSet



Gambar 8. Distribusi kelas sentimen hasil pelabelan InSet Lexicon

3.4 Pembagian Data

Untuk mengevaluasi kinerja model secara objektif, dataset yang telah diberi label dibagi menjadi data latih (*training data*) dan data uji (*testing data*) dengan rasio 80:20. Melalui pembagian ini, 80% dari total data digunakan oleh algoritma untuk mempelajari pola sentimen, sementara 20% sisanya dialokasikan sebagai data baru yang belum pernah dilihat oleh model untuk mengukur tingkat akurasi prediksi. Proses ini dilakukan menggunakan fungsi *train_test_split* dengan parameter *random_state* guna memastikan konsistensi hasil pembagian data pada setiap pengujian. Proporsi 80:20 dipilih karena dianggap ideal dalam menjaga keseimbangan antara kecukupan data untuk proses pembelajaran model dan validitas hasil pengujian, seperti pada gambar ke 9.



Gambar 9. Data Training & Data Testing

3.5 Implementasi Algoritma

3.5.1 Support Vector Machine (SVM)

Algoritma *Support Vector Machine (SVM)* mencapai akurasi tertinggi sebesar 90,7% dengan performa terbaik dalam klasifikasi sentimen dibandingkan kedua algoritma lainnya. Berdasarkan gambar 10, *SVM* menunjukkan dominasi akurasi yang signifikan dengan berhasil memprediksi 286 data negatif, 545 data netral, dan 481 data positif secara tepat. *SVM* menunjukkan keseimbangan yang baik pada seluruh kategori kelas, dengan nilai *macro precision* sebesar 0,909, *macro recall* sebesar 0,902, dan *macro F1-Score* sebesar 0,905. Pada kelas positif, *SVM* mencatatkan performa paling optimal dengan *precision* 0,932, *recall* 0,943, dan *F1-Score* 0,938. Minimnya angka *misklasifikasi* pada matriks, seperti hanya 3 data positif yang terprediksi sebagai negatif, membuktikan bahwa *SVM* sangat akurat, stabil, dan merupakan model paling andal untuk analisis sentimen pada dataset ini.

Confusion Matrix for SVM:

		SVM Confusion Matrix		
		Negatif	Netral	Positif
Actual	Negatif	293	46	3
	Netral	24	540	32
	Positif	5	24	481
		Predicted		
		Negatif	Netral	Positif

Gambar 10. Confusion Matrix SVM

3.5.2 Random Forest

Algoritma *Random Forest* mencapai akurasi sebesar 81,5% dan menempati posisi terendah di antara ketiga model yang diuji. Berdasarkan gambar 11, model ini berhasil memprediksi dengan benar sebanyak 250 data negatif, 486 data netral, dan 444 data positif. *Random Forest* mengklasifikasikan sentimen berdasarkan voting mayoritas dari sejumlah pohon keputusan, dengan nilai *macro precision* sebesar 0,817, *macro recall* sebesar 0,806, dan *macro F1-Score* sebesar 0,810. Performa terendah dijumpai pada kelas negatif, di mana terdapat 68 data negatif yang salah diprediksi sebagai netral dan 24 data lainnya sebagai positif. Hal ini menghasilkan nilai *recall* sebesar 0,731, yang mengindikasikan bahwa model ini kurang optimal dalam mendeteksi ulasan bersentimen negatif dibandingkan model lainnya. Meskipun demikian, *Random Forest* tetap dapat digunakan sebagai model pembanding karena memiliki ketahanan terhadap *overfitting*, sebagaimana ditunjukkan pada pola distribusi prediksi.

Confusion Matrix for Random Forest:

		Random Forest Confusion Matrix		
		Negatif	Netral	Positif
Actual	Negatif	250	68	24
	Netral	35	486	75
	Positif	17	49	444
		Predicted		
		Negatif	Netral	Positif

Gambar 11. Confusion Matrix Random Forest

3.5.3 Neural Network

Algoritma *Neural Network* memperoleh akurasi sebesar 89,2%, yang menempatkannya sebagai model kompetitif dalam klasifikasi sentimen ini. Berdasarkan Gambar 12, *Confusion Matrix* menunjukkan bahwa *Neural Network* memiliki kemampuan yang sangat baik dalam mengklasifikasikan data, terutama pada kelas Netral dengan 534 prediksi benar dan kelas Positif dengan 473 prediksi benar. Meskipun terdapat sedikit kerancuan (*misklasifikasi*) di mana 52 data negatif diprediksi sebagai netral, model ini tetap mencatatkan performa yang solid dengan nilai *macro precision* sebesar 0,898, *macro recall* sebesar 0,885, dan *macro F1-Score* sebesar 0,890. Pada kelas positif, *Neural Network* mencatatkan *precision* 0,910 dan *recall* 0,927, sementara kelas negatif berhasil meraih *precision* tertinggi di antara semua kelas yaitu 0,922. Secara keseluruhan, hasil ini mengonfirmasi bahwa *Neural Network* mampu mempelajari representasi fitur non-linear dari teks ulasan pengguna dengan sangat efektif, menjadikannya alternatif utama selain *SVM*.

Confusion Matrix for Neural Network:

		Neural Network Confusion Matrix		
		Negatif	Netral	Positif
Actual	Negatif	284	52	6
	Netral	21	534	41
	Positif	3	34	473
		Predicted		
		Negatif	Netral	Positif

Gambar 12. Confusion Matrix Neural Network

3.6 Evaluasi Model

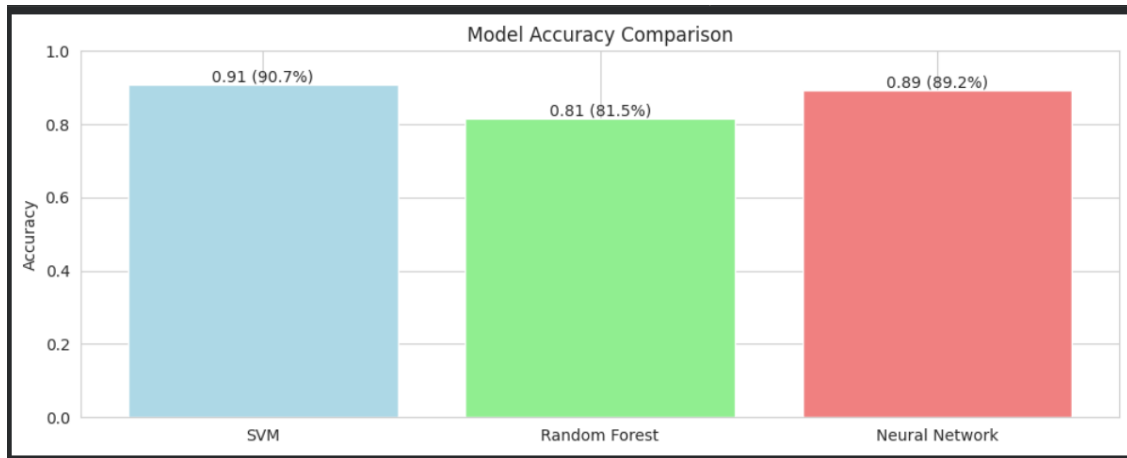
Hasil evaluasi pada *Support Vector Machine (SVM)* menunjukkan tingkat akurasi paling tinggi sebesar 90,7%, menandakan kemampuan yang unggul dalam mengklasifikasikan data sentimen dengan konsistensi tinggi. *Neural Network* memperoleh akurasi 89,2%, cukup kompetitif namun berada di bawah *SVM*, sedangkan *Random Forest* mencatat akurasi terendah sebesar 81,5%, yang meskipun memiliki ketahanan terhadap *overfitting*, masih kalah akurat dibandingkan kedua model lain. Berdasarkan perbandingan tersebut, *SVM* direkomendasikan sebagai pilihan utama untuk analisis sentimen ulasan pengguna aplikasi Instagram, sebagaimana ditunjukkan pada Tabel 1.

Capaian akurasi *SVM* 90,7% dalam penelitian ini melampaui hasil yang dilaporkan oleh Ulhaq et al. 2025 [13] (88%) pada ulasan Tokopedia dan Rismanah et al. 2024 [14] (80,31%) pada layanan ShopeeFood. Tingginya performa ini menunjukkan bahwa penggunaan model *SVM* dengan *kernel linear* sangat tangguh dalam menghadapi fitur *Bag-of-Words* yang bersifat *sparse* dan berdimensi tinggi. Jika dibandingkan dengan studi Faisal et al. 2024 [15] yang mencatat *recall* rendah pada kelas negatif, model dalam penelitian ini menunjukkan stabilitas yang lebih baik di seluruh metrik evaluasi.

Algoritma *Support Vector Machine (SVM)* mencapai akurasi tertinggi sebesar 90,7% dengan performa terbaik dalam klasifikasi sentimen dibandingkan kedua algoritma lainnya. *SVM* menunjukkan keseimbangan yang baik pada seluruh kategori kelas, dengan nilai *macro precision* sebesar 0,909, *macro recall* sebesar 0,902, dan *macro F1-Score* sebesar 0,905. Pada kelas positif, *SVM* mencatatkan performa paling optimal dengan *precision* 0,932, *recall* 0,943, dan *F1-Score* 0,938. Hasil evaluasi membuktikan bahwa *SVM* akurat, stabil, dan merupakan model paling andal untuk analisis sentimen pada dataset ini.

Algoritma *Random Forest* mencapai akurasi sebesar 81,5% dan menempati posisi terendah di antara ketiga model yang diuji. *Random Forest* mengklasifikasikan sentimen berdasarkan voting mayoritas dari sejumlah pohon keputusan, dengan nilai *macro precision* sebesar 0,817, *macro recall* sebesar 0,806, dan *macro F1-Score* sebesar 0,810. Performa terendah dijumpai pada kelas negatif dengan *recall* 0,731, yang mengindikasikan bahwa model ini kurang optimal dalam mendeteksi ulasan bersentimen negatif. Meskipun demikian, *Random Forest* tetap dapat digunakan sebagai model pembandingan karena memiliki ketahanan terhadap *overfitting*.

Algoritma *Neural Network* memperoleh akurasi sebesar 89,2%, menempatkannya pada posisi kedua di bawah *SVM*. *Neural Network* mampu mempelajari representasi fitur non-linear dari data teks ulasan pengguna dengan baik, yang tercermin dari nilai *macro precision* sebesar 0,898, *macro recall* sebesar 0,885, dan *macro F1-Score* sebesar 0,890. Pada kelas positif, *Neural Network* mencatatkan *precision* 0,910, *recall* 0,927, dan *F1-Score* 0,918, sementara kelas negatif memperoleh *precision* tertinggi di antara semua kelas yaitu 0,922. Hasil ini menunjukkan bahwa *Neural Network* merupakan alternatif yang kompetitif terhadap *SVM* dalam analisis sentimen pada dataset ini.



Gambar 13. Perbandingan Accuracy Model

Berdasarkan hasil perbandingan ketiga model yang ditunjukkan pada Gambar 13, *Support Vector Machine* (*SVM*) terbukti sebagai algoritma paling efektif dengan akurasi tertinggi sebesar 90,7%, diikuti *Neural Network* sebesar 89,2%, dan *Random Forest* sebesar 81,5%. Keunggulan *SVM* disebabkan oleh kemampuannya dalam mencari *hyperplane* optimal yang memisahkan kelas sentimen dengan margin maksimal, sehingga lebih andal dalam menangani data teks berdimensi tinggi seperti representasi *Bag-of-Words*. *Neural Network* menempati posisi kedua karena kemampuannya mempelajari pola non-linear meskipun hanya menggunakan satu lapisan tersembunyi dengan 100 neuron. Sementara itu, *Random Forest* mencatat akurasi terendah karena sifat algoritma tree-based yang kurang optimal pada ruang fitur yang sangat *sparse* dan berdimensi tinggi.

3.7 Analisis Hasil

Analisis hasil visualisasi *WordCloud* dilakukan untuk mengidentifikasi pola kata dominan sebelum dan sesudah proses *preprocessing*, sebagaimana ditampilkan pada Gambar 14. Sebelum *preprocessing*, teks ulasan didominasi oleh konjungsi dan kata ganti yang kurang informatif bagi analisis sentimen, seperti "dan", "saya", "yang", "di", serta penggunaan imbuhan atau partikel seperti "nya". Kata-kata tersebut muncul dengan ukuran yang sangat besar karena frekuensinya yang tinggi dalam bahasa sehari-hari, namun tidak memiliki kontribusi makna yang kuat terhadap polaritas sentimen.

Setelah melalui seluruh tahapan *preprocessing* yang meliputi *cleaning*, *case folding*, normalisasi, tokenisasi, *stopword removal*, dan *stemming*, distribusi kata berubah secara signifikan. Kata-kata yang bermakna dan merepresentasikan substansi ulasan menjadi jauh lebih menonjol. Kata "instagram", "akun", dan "aplikasi" muncul sebagai topik utama pembicaraan. Selain itu, kata-kata yang membawa muatan sentimen dan pengalaman pengguna seperti "bagus", "bug", "tolong", "banget", dan "suka" terlihat lebih jelas. Fokus fitur juga terekam melalui kemunculan kata "ig", "fitur", "foto", dan "video". Dengan demikian, *WordCloud* sesudah *preprocessing* menjadi lebih representatif dalam mencerminkan topik utama dan emosi yang disampaikan pengguna Instagram dalam ulasan mereka di Google Play Store.



Gambar 14. *WordCloud* Before and After

4. KESIMPULAN

Berdasarkan seluruh rangkaian tahapan eksperimen yang telah dilakukan, dapat disimpulkan bahwa penerapan teknik *Natural Language Processing* (*NLP*) yang dikombinasikan dengan algoritma pembelajaran mesin terbukti sangat efektif dalam membedah pola opini pengguna terhadap aplikasi Instagram secara sistematis. Dari 10.000 sampel ulasan

yang diperoleh pada April 2026, penelitian ini berhasil mengidentifikasi struktur sentimen pengguna di Google Play Store yang didominasi oleh kategori netral (41,16%), sentimen positif (35,24%), dan sentimen negatif (23,60%). Hasil distribusi ini menunjukkan bahwa tingkat kepercayaan dan kepuasan pengguna terhadap layanan Instagram masih berada pada level yang terjaga, meskipun terdapat porsi ulasan netral yang cukup tinggi. Hal ini menandakan bahwa banyak pengguna yang memanfaatkan kolom ulasan sebagai media penyampaian informasi fungsional atau saran pembaruan fitur daripada sekadar mengekspresikan kekecewaan atau pujian yang ekstrem. Selain itu, analisis komparatif antara tiga algoritma klasifikasi menunjukkan temuan yang sangat berarti bagi pengembangan sistem cerdas di masa mendatang. Algoritma *Support Vector Machine (SVM)* terbukti memiliki stabilitas dan akurasi yang lebih unggul dibandingkan dengan *Random Forest (RF)* dan *Neural Network (NN)*. Keunggulan performa *SVM* ini didukung oleh kapabilitas algoritma dalam memetakan fitur teks jarang (*sparse*) ke dalam dimensi yang lebih tinggi melalui penggunaan *hyperplane* optimal dalam metode *Bag-of-Words*. Temuan penelitian ini memberikan rekomendasi praktis bagi pengembang perangkat lunak dan manajemen media sosial untuk mengadopsi model *SVM* sebagai instrumen pemantauan reputasi digital secara *real-time*. Dengan demikian, integrasi antara pengolahan bahasa alami dan algoritma pembelajaran mesin tidak hanya mampu memetakan tren opini publik secara luas, tetapi juga menyediakan landasan data yang ilmiah dan terukur untuk pengambilan keputusan strategis dalam pengembangan teknologi informasi di masa depan.

REFERENSI

- [1] R. Y. Hayuningtyas dan R. Sari, “Analisis Sentimen Terhadap Pengguna Aplikasi Instagram di Google Play Store Menggunakan *Support Vector Machine*,” *Journal of Accounting Information System*, vol. 05, no. 1, Jun 2025, doi: 10.31294/jais.v5i01.8823.
- [2] D. Surya Sayogo, B. Irawan, dan A. Bahtiar, “Analisis Sentimen Ulasan Instagram Di Google Play Store Menggunakan Algoritma *Naïve Bayes*,” *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, 2023, doi: 10.36040/jati.v7i6.8178.
- [3] A. Zaiem Praghakusma dan N. Charibaldi, “Komparasi Fungsi *Kernel* Metode *Support Vector Machine* untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi),” vol. 9, no. 2, hlm. 33–42, 2021, doi: 10.12928/jstie.v8i3.xxx.
- [4] R. Cahyadi, A. Damayanti, dan D. Aryadani, “*Recurrent Neural Network (Rnn)* Dengan *Long Short Term Memory (Lstm)* Untuk Analisis Sentimen Data Instagram,” *Jurnal Informatika dan Komputer*, vol. 5, no. 1, Feb 2020, doi: 10.26798/jiko.v5i1.407.
- [5] A. Latif, M. Yugi, F. S. Utomo, dan T. Hariguna, “Comparison Of The Performance Of *Svm*, *Random Forest*, And *Neural Network* Algorithms In Sentiment Analysis Of Openai Application Reviews On The Google Play Store,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 11, no. 1, hlm. 478–489, Feb 2026, doi: 10.29100/jupi.v11i1.7793.
- [6] Lady Agustin Fitriana, I. Sugiyarto, dan U. Faddillah, “Penerapan Model *LSTM* dan *CNN* Untuk Klasifikasi Sentimen Pada Ulasan Aplikasi Roblox,” *Infotek: Jurnal Informatika dan Teknologi*, vol. 9, no. 1, hlm. 71–82, Jan 2026, doi: 10.29408/jit.v9i1.32848.
- [7] R. Mei Sasmita, A. Meiriza, dan H. Novianti, “An *Ensemble Learning* Approach for Sentiment Analysis of Maxim Application Reviews Using *SVM*, *KNN*, and *Random Forest*,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 6, hlm. 3697, 2025, doi: 10.30871/jaic.v9i6.11447.
- [8] J. O. Leandro dan M. I. Fianty, “Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of *Support Vector Machines* and *Naïve Bayes*,” *International Journal on Informatics Visualization*, vol. 9, no. 2, hlm. 796–807, Mar 2025, [Daring]. Tersedia pada: www.joiv.org/index.php/joiv
- [9] R. Lubis dan S. Anggraeni, “Klasifikasi Sentiment Analysis Terhadap Usulan KB Vasektomi Syarat Penerima BANSOS dengan Metode *Naive Bayes*,” *Jurnal Komputer Teknologi Informasi Sistem Informasi (JUKTISI)*, vol. 4, no. 3, hlm. 1384–1394, Okt 2025, doi: 10.62712/juktisi.v4i3.631.

- [10] S. R. K. W. Tommy Rustandi, D. Suhaedi, dan Y. Pemasari, “Pemetaan *Hyperplane* Pada *Support Vector Machine*,” *Bandung Conference Series: Mathematics*, vol. 3, no. 2, hlm. 109–119, Agu 2023, doi: 10.29313/bcsm.v3i2.8187.
- [11] L. N. Qomariyati, S. Nurpadillah, N. F. Rosyidin, F. Sofyan, dan L. R. Mubarak, “Perbandingan Akurasi *Random Forest* Classifier dalam Memprediksi Kelayakan Kendaraan berdasarkan Jumlah *Decision tree* dan *Selection Feature*,” *Jurnal FUSE-Teknik Elektro*, vol. 4, hlm. 21–30, Jun 2024.
- [12] A. Arasy, S. Agustian, L. Handayani, dan I. Iskandar, “Klasifikasi Sentimen Menggunakan Metode *Multilayer Perceptron* dengan Fitur *TF-IDF*,” *MALCOM: Indonesian Journal of Machine learning and Computer Science*, vol. 5, no. 3, hlm. 908–919, Jun 2025, doi: 10.57152/malcom.v5i3.2052.
- [13] A. L. D. Ulhaq dan Suprayogi, “Comparing *Machine learning* Models for Sentiment Analysis of Tokopedia Reviews,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 6, hlm. 3642–3647, Des 2025, doi: 10.30871/jaic.v9i6.11239.
- [14] S. N. Rismanah, R. Astuti, dan F. M. Basysyar, “Penerapan Algoritma *Support Vector Machine* Dalam Menganalisis Sentimen Ulasan Pelanggan *Shopeefood* Berdasarkan *Twitter*,” *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 1, Feb 2024, doi: 10.36040/jati.v8i1.8401.
- [15] H. Faisal, A. Febriandirza, dan F. N. Hasan, “Analisis Sentimen Terkait Ulasan Pada Aplikasi *PLN Mobile* Menggunakan Metode *Support Vector Machine*,” *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, vol. 5, no. 1, hlm. 303–312, Jan 2024, [Daring]. Tersedia pada: <http://repository.uhamka.ac.id/id/eprint/32133/5/4.%20%20Artikel.pdf>