

PENERAPAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN ULASAN APLIKASI OVO PADA PLATFORM GOOGLE PLAY STORE

Aisyah Maharani¹, Bambang Irawan², Ahmad Faqih³, Raditya Danar Dana⁴, Ade Rizki Rinaldi⁵.

Program Studi Teknik Informatika¹²³
Program Studi Manajemen Informatika⁴
Program Studi Rekayasa Perangkat Lunak⁵
STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
aisyahmaharani095@gmail.com

(*) Corresponding Author : aisyahmaharani095@gmail.com.

Abstract— This study focuses on examining sentiment patterns expressed by users regarding the OVO digital wallet service. The analysis is conducted using user reviews collected from the Google Play Store. The research adopts a text mining workflow that includes preprocessing steps and sentiment labeling, both manually and through rule-based methods. The feature extraction technique applied is Term Frequency Inverse Document Frequency (TF-IDF), while the classification process is carried out using the Support Vector Machine (SVM) algorithm. Several preprocessing procedures such as text cleaning, normalization, tokenization, stopword removal, and stemming play a crucial role in improving data quality, consistent with previous studies that emphasize the importance of preprocessing for enhancing model performance. The TF-IDF method effectively transforms textual reviews into numerical representations, enabling the model to identify relevant terms more accurately. The experimental results indicate that the SVM algorithm achieves the best performance in classifying sentiments into positive and negative categories. Model evaluation is conducted using accuracy, precision, recall, and F1-score. These findings reinforce the evidence that SVM is highly capable of handling high-dimensional text data. Sentiment visualizations in the form of pie charts, bar charts, and word clouds further illustrate user perceptions of the application. Overall, this study demonstrates that combining effective preprocessing with TF-IDF feature extraction improves the quality of sentiment classification. The outcomes of this research are expected to serve as a reference for developing sentiment analysis systems and to provide recommendations for application developers in enhancing the quality of digital wallet services.

Keywords: Sentiment Analysis, OVO, Support Vector Machine, Preprocessing, TF-IDF, User Reviews.

Abstrak— Penelitian ini difokuskan pada pengkajian pola sentimen yang muncul dari opini pengguna terhadap layanan dompet digital OVO. Analisis dilakukan menggunakan kumpulan ulasan yang dihimpun dari platform Google Play Store. Studi ini menggunakan alur kerja text mining yang mencakup langkah preprocessing serta proses pelabelan sentimen, baik secara manual maupun aturan tertentu. Metode ekstraksi fitur yang diaplikasikan adalah Term Frequency Inverse Document Frequency (TF-IDF), serta Proses klasifikasi dilakukan dengan memanfaatkan algoritma Support Vector Machine (SVM). Beberapa langkah yang dilakukan mencakup pembersihan teks, normalisasi, tokenisasi, penghapusan stopword, serta proses stemming terbukti sangat berpengaruh dalam meningkatkan kualitas data, sejalan dengan penelitian terdahulu yang menegaskan pentingnya preprocessing dalam meningkatkan performa model. Pendekatan TF-IDF terbukti efektif dalam mentransformasikan teks ulasan ke dalam bentuk numerik sebagai masukan bagi model. Sehingga model mampu membedakan kata-kata yang relevan. Hasil eksperimen menunjukkan bahwa algoritma SVM menghasilkan performa terbaik dalam mengklasifikasikan sentimen menjadi kategori positif, dan negatif. Evaluasi performa dilakukan menggunakan indikator akurasi, presisi, recall, dan F1-score. Temuan tersebut memperkuat bukti bahwa SVM memiliki kemampuan yang baik dalam memproses data teks dengan dimensi fitur yang besar. Visualisasi sentimen berupa diagram pie, diagram batang, dan wordcloud juga membantu memperjelas pola persepsi pengguna terhadap aplikasi. Secara keseluruhan, penelitian ini membuktikan bahwa kombinasi preprocessing yang efektif, ekstraksi fitur TF-IDF. Diharapkan hasil penelitian ini dapat dimanfaatkan sebagai acuan pengembangan sistem analisis sentimen serta memberikan rekomendasi bagi pengembang aplikasi dalam meningkatkan mutu layanan dompet digital.

Kata Kunci: Analisis Sentimen, OVO, Support Vector Machine, Preprocessing, TF-IDF, Ulasan Pengguna.

INTRODUCTION

Perkembangan teknologi finansial di Indonesia dalam beberapa tahun terakhir, perkembangan yang terjadi menunjukkan peningkatan yang sangat signifikan, khususnya dalam penggunaan layanan pembayaran digital. Dompot digital atau e-wallet telah menjadi solusi praktis bagi masyarakat untuk memfasilitasi transaksi nontunai yang berlangsung lebih cepat, aman, dan memiliki efisiensi yang lebih tinggi. Salah satu aplikasi yang menempati posisi penting adalah OVO, yang digunakan secara luas oleh berbagai kalangan masyarakat. Popularitas OVO menunjukkan adanya tingkat adopsi teknologi yang tinggi, namun di sisi lain juga menghadirkan tantangan dalam hal pemanfaatan fitur secara optimal serta munculnya beragam ulasan dari pemakai di platform *Google Play Store*.

Ulasan pengguna menjadi sumber data penting dalam mengevaluasi kualitas aplikasi. Tidak hanya dalam bentuk rating bintang yang bersifat kuantitatif, ulasan juga berupa komentar teks yang bersifat kualitatif, mencerminkan kepuasan, keluhan, maupun harapan pengguna. Analisis ulasan ini sangat bermanfaat bagi pengembang aplikasi untuk memahami persepsi pengguna secara lebih komprehensif. Akan tetapi, jumlah ulasan yang sangat besar dan sifat data yang tidak terstruktur membuat analisis secara manual menjadi tidak tepat. Oleh karena itu, diperlukan metode otomatis yang mampu mengolah ulasan secara sistematis dan akurat.

Analisis sentimen merupakan metode yang umum digunakan dalam bidang text mining dan Natural Language Processing (NLP) untuk mengelompokkan suatu pendapat ke dalam kategori sentimen seperti positif maupun negatif. Melalui analisis ini, pengembang dapat memperoleh gambaran proporsi kepuasan pengguna sekaligus mengidentifikasi aspek layanan yang sering mendapat perhatian. Bidang Teknik Informatika berperan penting dalam pengembangan metode ini, khususnya melalui penerapan machine learning untuk mengolah data teks dalam jumlah besar. Penelitian [1] berhasil meningkatkan kinerja SVM dengan optimasi Particle Swarm Optimization (PSO) untuk klasifikasi ulasan dompet digital. Sementara itu, [2] menemukan bahwa SVM tetap konsisten menghasilkan klasifikasi yang baik pada aplikasi keuangan lain, seperti Bank Syariah Indonesia. Identifikasi persoalan yang muncul dari fenomena ini meliputi:

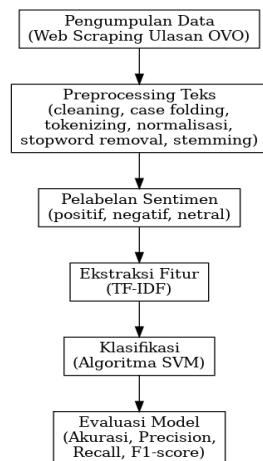
1. besarnya jumlah ulasan yang belum dimanfaatkan secara optimal;
2. bentuk teks ulasan yang tidak terstruktur sehingga menyulitkan analisis manual; dan
3. belum adanya sistem otomatis yang mampu mengklasifikasikan opini pengguna secara cepat, akurat, dan konsisten. Dengan demikian, pendekatan berbasis machine learning menjadi solusi untuk mengatasi keterbatasan analisis konvensional.

Penelitian ini bertujuan merancang model klasifikasi sentimen terhadap ulasan pengguna aplikasi OVO pada platform *Google Play Store* dengan memanfaatkan algoritma Support Vector Machine. Selain mengukur performa model melalui metrik akurasi, presisi, recall, dan F1-score, studi ini juga bertujuan menyajikan pemahaman yang lebih jelas mengenai kecenderungan sikap serta persepsi pengguna terhadap layanan OVO.

Kebaruan penelitian ini terletak pada penggunaan ulasan *Google Play Store* sebagai sumber data utama, yang lebih relevan karena berasal langsung dari pengalaman pengguna aplikasi, berbeda dengan sebagian besar penelitian terdahulu yang berfokus pada media sosial. Selain itu, penelitian ini menggabungkan metode preprocessing teks, pelabelan berbasis rating dan kamus kata, serta pembobotan TF-IDF untuk meningkatkan kualitas klasifikasi. Dengan pendekatan ini, penelitian diharapkan memberikan kontribusi praktis bagi pengembang aplikasi dalam meningkatkan kualitas layanan sekaligus memperkaya literatur analisis sentimen berbasis machine learning pada domain fintech di Indonesia.

MATERIALS AND METHODS

Penelitian ini menerapkan pendekatan kuantitatif melalui desain eksperimen yang memanfaatkan teknik machine learning. Fokus utama studi ini adalah menghasilkan model klasifikasi sentimen untuk ulasan pengguna aplikasi OVO pada *Google Play Store* dengan menggunakan algoritma Support Vector Machine (SVM). Pemilihan SVM didasari oleh keampuannya dalam memproses data teks yang memiliki format tinggi serta rekam jejak performa yang stabil pada berbagai penelitian terkait analisis sentimen.



Gambar 1 Metode penelitian

Tahapan penelitian dilakukan secara sistematis, dimulai dari pengumpulan data hingga evaluasi model. Alur penelitian dapat dijelaskan sebagai berikut:

A. Pengumpulan Data

Data diperoleh melalui web scraping dari Google Play Store yang berisi teks ulasan dan rating bintang, kemudian disimpan dalam format Excel.

B. Preprocessing

Tahap ini bertujuan menyiapkan teks agar layak digunakan dalam proses klasifikasi. Proses meliputi:

1. Cleaning
Pada tahap *cleaning*, teks dibersihkan dari elemen-elemen yang tidak diperlukan, seperti tag HTML, tautan, serta berbagai simbol yang berlebihan dan tidak relevan dengan isi utama dokumen [3].
2. Case folding
menyeragamkan huruf menjadi kecil semua.
3. Tokenizing
Tokenizing merupakan proses untuk memisahkan kalimat menjadi beberapa kata. Biasanya setiap kata akan dipisahkan oleh spasi sehingga pada kasus ini akan digunakan space delimiter [4].
4. Normalisasi kata
mengubah kata tidak baku menjadi baku.
5. Stopword removal
Menghapus kata-kata yang tidak penting, seperti "apakah", "ke", "luar", dan "biasanya" (Farhan & Setiaji, 023).
6. Stemming
mengubah kata berimbuhan menjadi kata dasar.

C. Pelabelan Sentimen

Setiap ulasan diberi label sentimen berdasarkan kombinasi rating bintang dan kamus kata sentimen. Label dikelompokkan ke dalam tiga kategori: positif, negatif, dan netral.

D. Ekstraksi Fitur

Teks yang telah diproses direpresentasikan secara numerik menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk mengukur sejauh mana sebuah kata dianggap penting dalam suatu dokumen. Setelah setiap kata direpresentasikan dalam bentuk numerik, nilai TF-IDF tersebut kemudian dimasukkan ke dalam model klasifikasi berbasis supervised learning [5]. Tujuan utama dari TF-IDF adalah memberikan nilai lebih besar pada kata-kata yang kemunculannya dominan dalam satu dokumen, namun relatif jarang ditemukan pada dokumen lainnya dalam korpus kumpulan tersebut. Dengan demikian, jika sebuah kata muncul kali dalam suatu teks, maka teks tersebut akan memiliki skor [6].

E. Klasifikasi dengan SVM

Model SVM digunakan untuk mengklasifikasikan ulasan ke dalam kelas sentimen. Algoritma ini dipilih karena efektif untuk data linier dan non-linier serta memiliki generalisasi yang baik. Dalam proses klasifikasi, setiap data masukan diasumsikan termasuk ke dalam salah satu dari dua kelas yang tersedia. Kondisi ini menjadikan Support Vector Machine (SVM) berperan sebagai algoritma pengklasifikasi linier biner yang bersifat non-probabilistik. SVM bekerja dengan memanfaatkan himpunan data pelatihan yang telah diberi label, di mana masing-masing data sudah ditentukan keanggotaannya terhadap salah satu kategori. Berdasarkan informasi tersebut, SVM kemudian membangun sebuah batas pemisah yang optimal untuk membedakan kedua kelas secara sistematis. [7].

F. Evaluasi Model

Model dievaluasi menggunakan confusion matrix dengan metrik akurasi, presisi, recall, dan F1-score. Uji dilakukan dengan pembagian data menjadi training set dan testing set, untuk mengukur performa model dalam mengenali sentimen pengguna terhadap aplikasi OVO.

1. Data Penelitian

Setelah data dibersihkan, tahap selanjutnya adalah memberikan label pada setiap ulasan. Pelabelan dilakukan dengan mengacu pada kombinasi rating bintang dan kamus kata sentimen. Pendekatan ini membantu membangun dataset terlabel yang diperlukan untuk proses klasifikasi menggunakan machine learning.

2. Pengujian / Eksperimen

Proses pengujian dilakukan dengan memisahkan dataset ke dalam dua bagian, yakni data pelatihan (training set) dan data pengujian (testing set). Model SVM dibangun menggunakan data pelatihan, kemudian dievaluasi melalui data pengujian untuk menilai kinerjanya. Hasil evaluasi tersebut dianalisis lebih lanjut untuk mengidentifikasi kecenderungan sentimen yang paling dominan serta mengukur tingkat ketepatan model dalam mengklasifikasikan opini pengguna terhadap aplikasi OVO.

RESULTS AND DISCUSSION

A. Scraping Data

Tahap Tahap awal penelitian dimulai dengan pengumpulan ulasan pengguna aplikasi OVO melalui Google Play Store menggunakan pustaka *google-play-scrapers*. Ulasan yang berhasil diambil tersebut kemudian menjadi dasar bagi proses analisis selanjutnya. mencakup rating bintang, teks ulasan, dan label sentimen, lalu disimpan dalam format .xlsx untuk analisis lanjutan. Ulasan yang beragam, baik positif maupun negatif, menjadikannya sesuai sebagai dataset analisis sentimen.

	ulasan	Labelin
0		memuaskan
1	ini kenapa yaa , saya transfer dari ovo ke ban...	negatif
2	maaf kenapa top up saya ga ada perihal laporan...	negatif
3		bgus
4	beli pulsa tapi dalam proses trus udah lebih d...	negatif
5	Makin lama potonganya makin naik dulu 1000 tru...	negatif
6	saya suka sekali dengan aplikasi ovo, transfer...	positif
7	sy sangat kecewa dgn aplikasi ovo in, skrng GK...	negatif
8	maaf saya beri bintang 1, kecewa banget... isi...	negatif
9		mudahan bisa aman

Gambar 2 Hasil Scraping data

B. Preprocessing

Proses preprocessing dilakukan untuk menyiapkan data teks agar siap dianalisis, meliputi tahap pembersihan teks, normalisasi, tokenisasi, penghapusan stopwords, dan stemming guna menghasilkan data yang bersih dan terstandar.

Tabel 1 Hasil preprocessing

Teks Asli	Aku nggak suka banget sama produk ini, harganya mahal tp kualitasnya jelek!
Cleaning	aku nggak suka banget sama produk ini harganya mahal tp kualitasnya jelek
Normalisasi Kata	aku tidak suka banget sama produk ini harganya mahal tapi kualitasnya jelek
Tokenisasi	['aku', 'tidak', 'suka', 'banget', 'sama', 'produk', 'ini', 'harganya', 'mahal', 'tapi', 'kualitasnya', 'jelek']

Stopword Removal	['aku', 'suka', 'banget', 'sama', 'produk', 'harganya', 'mahal', 'kualitasnya', 'jelek']
Stemming	['aku', 'suka', 'banget', 'sama', 'produk', 'harga', 'mahal', 'kualitas', 'jelek']

C. Pelabelan Sentimen

Proses pelabelan manual dilakukan dengan validasi oleh ahli bahasa guna memastikan ketepatan klasifikasi sentimen setiap ulasan, sehingga label yang diberikan sesuai dengan makna dan konteks bahasa yang digunakan.

	ulasan	Labelin
0		memuaskan
1	ini kenapa yaa , saya transfer dari ovo ke ban...	negatif
2	maaf kenapa top up saya ga ada perihal laporan...	negatif
3		bgus
4	beli pulsa tapi dalam proses trus udah lebih d...	negatif
5	Makin lama potonganya makin naik dulu 1000 tru...	negatif
6	saya suka sekali dengan aplikasi ovo, transfer...	positif
7	sy sangat kecewa dgn aplikasi ovo in, skrng GK...	negatif
8	maaf saya beri bintang 1, kecewa banget... isi...	negatif
9		mudahan bisa aman

Gambar 3 Hasil Pelabelan sentimen

D. Ekstraksi Fitur (TF-IDF Vectorizer)

Proses pembobotan tutur kata dalam studi ini dilakukan menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF). Pada tahap ini, setiap kata dalam tinjauan data praproses diubah menjadi vektor numerik berdasarkan tingkat kepentingannya dalam dokumen. Nilai term frequency digunakan untuk menghitung seberapa

```
Shape dari matriks TF-IDF: (2984, 3915)
Beberapa fitur TF-IDF pertama: ['aada' 'aahhh' 'aaldo' 'aargghhh' 'abal' 'abang' 'abe' 'abis' 'acara' 'acc']
```

Gambar 4 Hasil Ekstraksi Fitur TF-IDF

sering suatu kata muncul dalam tinjauan, sementara inverse document frequency mengukur keunikan kata di seluruh korpus. Gabungan kedua elemen ini mendatangkan bobot TF-IDF yang dapat merepresentasikan relevansi setiap kata secara lebih akurat. Hasil pembobotan ini kemudian menjadi dasar cara training model Support Vector Machine (SVM) untuk klasifikasi sentimen.

E. Klasifikasi

Pada tahap klasifikasi, model yang dikembangkan diuji untuk menentukan kemampuannya dalam mengelompokkan data ulasan ke dalam kategori sentimen yang telah ditentukan. Melalui serangkaian evaluasi, seperti laporan klasifikasi, matriks konfusi, dan validasi silang, kinerja model dapat dianalisis secara komprehensif untuk memastikan kualitas klasifikasi yang dihasilkan.

1. Report

Pada fase evaluasi ini, model dianalisis menggunakan laporan klasifikasi yang berisi pengukuran presisi, recall, skor F1, dan akurasi total. Laporan ini memberikan gambaran komprehensif tentang kemampuan model untuk membedakan setiap kelas sentimen. Presisi menggambarkan akurasi model dalam menghasilkan prediksi yang tepat untuk suatu kelas, sementara recall menunjukkan seberapa baik model menangkap semua data yang seharusnya termasuk dalam kelas tersebut. Kedua metrik ini kemudian dirangkum dalam skor F1 untuk menilai keseimbangan kinerja model. Laporan ini memungkinkan penilaian yang lebih objektif terhadap kualitas hasil klasifikasi.

```

=== Classification Report ===
              precision    recall  f1-score   support

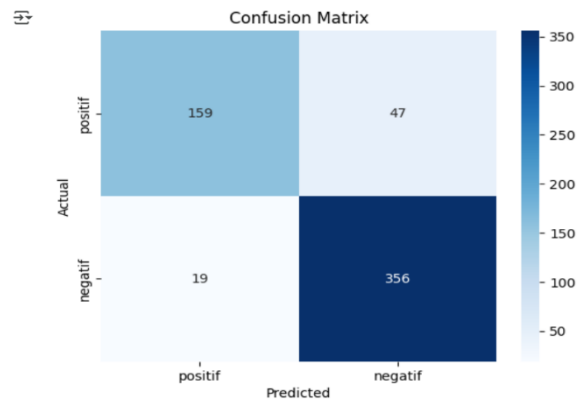
   negatif      0.88      0.97      0.92       307
   positif      0.93      0.76      0.83       165

 accuracy              0.89       472
 macro avg              0.90       472
 weighted avg          0.90       472
    
```

Gambar 5 Hasil klasifikasi report

2. Confusin Matrix

Confusin Matrix digunakan untuk memeriksa pola prediksi model secara lebih rinci untuk setiap kategori sentimen. Matriks ini menampilkan jumlah prediksi yang sesuai dengan data aktual (positif benar dan negatif benar), serta jumlah prediksi yang salah (positif salah dan negatif salah). Dengan memahami distribusi kesalahan ini, peneliti dapat mengidentifikasi kecenderungan model untuk salah mengklasifikasikan data. Analisis ini membantu dalam menilai akurasi model secara lebih mendalam dan mengidentifikasi area yang perlu ditingkatkan.



Gambar 6 Hasil klasifikasi Confusin matrix

3. Cross Validation

Untuk memastikan model berkinerja konsisten pada set data yang berbeda, teknik Cross Validation digunakan. Teknik ini membelah set data menjadi beberapa partikel (fold), akhirnya digunakan secara bergantian sebagai data pelatihan dan uji. Setiap lipatan memiliki kesempatan untuk berfungsi sebagai data uji sekali, memastikan hasil evaluasi yang lebih stabil dan representatif. Metode ini menguji kemampuan generalisasi model SVM dan memastikan bahwa model tidak berkinerja baik hanya pada satu set data tertentu.

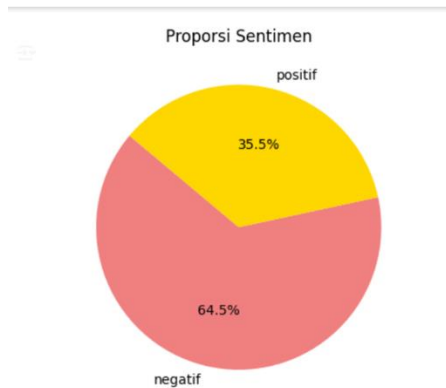
```

=== Cross Validation Score (SVM) ===
Akurasi rata-rata: 0.8736711143166247
    
```

Gambar 7 Hasil klasifikasi cross valiation

F. Visualisasi

Hasil visualisasi dalam penelitian ini menyajikan pemahaman yang lebih terstruktur mengenai pola serta karakteristik sentimen yang muncul pada ulasan pengguna. Diagram pie memperlihatkan pembagian proporsi sentimen positif, dan negatif, sehingga dapat menggambarkan kecenderungan persepsi pengguna terhadap aplikasi secara umum.



Gambar 8 Hasil Visualisasi

Visualisasi yang dihasilkan pada penelitian ini memiliki peran penting dalam memberikan gambaran yang lebih mendalam terkait pola sentimen yang muncul pada ulasan pengguna. Melalui penyajian diagram pie, dapat dilihat bahwa sentimen pengguna terbagi menjadi tiga kelompok utama, yakni positif, dan negatif. Persentase setiap kategori tersebut menggambarkan kecenderungan umum penilaian pengguna terhadap aplikasi, sehingga dapat menjadi tolok ukur awal dalam memahami kualitas pengalaman yang mereka rasakan.

Secara keseluruhan, visualisasi ini tidak hanya menyajikan data dalam format yang lebih mudah dipahami, tetapi juga menjadi landasan penting dalam menginterpretasikan arah kecenderungan sentimen. Hasil visualisasi memungkinkan peneliti memperoleh penafsiran yang lebih m mengenai persepsi pengguna serta dampaknya terhadap evaluasi performa aplikasi. Temuan ini juga dapat dijadikan acuan dalam merancang strategi peningkatan kualitas atau pengembangan fitur aplikasi pada tahap selanjutnya.

CONCLUSION

Penelitian ini dilakukan untuk menganalisis sentimen ulasan pengguna aplikasi dengan menerapkan serangkaian tahapan pengolahan data, mulai dari pra-pemrosesan, pelabelan sentimen, pembobotan kata menggunakan TF-IDF, hingga pemodelan menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM). Berdasarkan rangkaian proses dan hasil

evaluasi yang diperoleh, beberapa kesimpulan dapat ditarik sebagai berikut.

Pertama, proses pra-pemrosesan terbukti memiliki peran yang sangat penting dalam meningkatkan kualitas data yang digunakan. Tahapan seperti cleaning, normalisasi, tokenisasi, stopword removal, dan stemming mampu mengurangi noise dan menghasilkan teks yang lebih konsisten. Temuan ini sejalan dengan penelitian [1] serta [8] yang menegaskan bahwa preprocessing berpengaruh besar terhadap peningkatan akurasi model NLP, khususnya pada bahasa dengan karakteristik informal.

Kedua, metode pembobotan TF-IDF efektif dalam merepresentasikan teks ulasan ke dalam bentuk numerik yang informatif. Pembobotan ini membantu menonjolkan kata-kata penting yang berpengaruh terhadap sentimen. Hasil ini konsisten dengan penelitian [9] yang menunjukkan bahwa TF-IDF mampu meningkatkan kinerja model klasifikasi berbasis teks.

Ketiga, hasil eksperimen menunjukkan bahwa algoritma SVM memberikan performa lebih baik dibandingkan Naive Bayes dalam mengklasifikasikan sentimen. Nilai akurasi, presisi, recall, dan F1-score pada SVM lebih tinggi, serta menunjukkan kestabilan performa yang lebih baik. Kesimpulan ini sejalan dengan hasil penelitian [10], [4], [11], [12], serta [2], yang sama-sama melaporkan bahwa SVM unggul dalam tugas klasifikasi teks karena kemampuannya menghadapi data berdimensi tinggi. Selain itu, temuan ini juga menguatkan penelitian [13] yang menempatkan SVM sebagai model dengan kinerja terbaik dalam analisis sentimen aplikasi digital.

Keempat, visualisasi data seperti diagram pie, diagram batang, wordcloud, dan confusion matrix memberikan gambaran menyeluruh tentang distribusi sentimen serta kualitas hasil klasifikasi. Visualisasi ini mempermudah interpretasi pola ulasan pengguna dan membantu mengidentifikasi kecenderungan tertentu, sebagaimana direkomendasikan oleh penelitian-penelitian terdahulu terkait analisis sentimen pada platform ulasan dan sosial media.

Secara keseluruhan, penelitian ini menunjukkan bahwa kombinasi preprocessing yang baik, pelabelan yang valid, TF-IDF sebagai ekstraksi fitur, serta penggunaan SVM sebagai algoritma utama memberikan hasil klasifikasi sentimen yang optimal. Temuan ini menguatkan bukti dari berbagai penelitian sebelumnya bahwa SVM merupakan salah satu metode paling efektif untuk tugas klasifikasi teks. Penelitian ini diharapkan dapat menjadi rujukan bagi pengembangan sistem analisis sentimen berbasis ulasan pengguna serta memberikan dasar bagi pengembang aplikasi dalam

meningkatkan kualitas layanan berdasarkan masukan pengguna.

REFERENCE

- [1] Y. Albalawi, J. Buckley, and N. S. Nikolov, "Investigating the impact of pre-processing techniques and pre-trained word embeddings in detecting Arabic health information on social media," *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00488-w.
- [2] A. Noviriandini, H. Hermanto, and Y. Yudhistira, "Klasifikasi Support Vector Machine Berbasis Particle Swarm Optimization Untuk Analisa Sentimen Pengguna Aplikasi Pedulilindungi," *JIKA (Jurnal Inform.)*, vol. 6, no. 1, p. 50, 2022, doi: 10.31000/jika.v6i1.5681.
- [3] A. S. Agil Rafsanjani, D. L. Fithri, and S. Supriyono, "Sentiment Analysis of User Reviews of the KitaLulus Application on Google Play Store using the Support Vector Machine (SVM) Algorithm," *Sistemasi*, vol. 14, no. 5, p. 2519, 2025, doi: 10.32520/stmsi.v14i5.5519.
- [4] E. Hokijuliandy, H. Napitupulu, and F. Firdaniza, "Analisis Sentimen Menggunakan Metode Klasifikasi Support Vector Machine (SVM) dan Seleksi Fitur Chi-Square," *SisInfo J. Sist. Inf. dan Inform.*, vol. 5, no. 2, pp. 40–49, 2023, doi: 10.37278/sisinfo.v5i2.670.
- [5] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Comparative Evaluation of SVM Kernels for Sentiment Classification in Fuel Price Increase Analysis," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023.
- [6] D. Toresa, S. Rico Francisco Sitorus, I. Muzdalifah, F. Wiza, and R. Syelly, "Analisis Sentimen Terhadap Ulasan Penggunaan Dompot Digital Dana Menggunakan Metode Klasifikasi Support Vector Machine," *Technologica*, vol. 3, no. 2, pp. 64–74, 2024, doi: 10.55043/technologica.v3i2.163.
- [7] A. A. Munandar, F. Farikhin, and C. E. Widodo, "Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 8, no. 2, p. 77, 2023, doi: 10.31328/jointecs.v8i2.4747.
- [8] Rianto, A. B. Mutiara, E. P. Wibowo, and P. I. Santosa, "Improving the accuracy of text classification using stemming method, a case of non-formal Indonesian conversation," *J. Big Data*, vol. 8, no. 1, pp. 1–16, 2021, doi: 10.1186/s40537-021-00413-1.
- [9] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 422, 2021, doi: 10.30865/mib.v5i2.2845.
- [10] C. Anilkumar, S. V E., S. Kanchana, and S. B. Kumar, "Sentimental Analysis on Product Reviews Using Support Vector Machine and Nave Bayes," *Appl. Comput. Eng.*, vol. 2, no. 1, pp. 1067–1073, 2023, doi: 10.54254/2755-2721/2/20220586.
- [11] P. Arsi and R. Waluyo, "Sentiment Analysis of Discourse on Moving the Indonesian Capital City Using the Support Vector Machine (SVM) Algorithm," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.202183944.
- [12] E. R. Kaburuan and N. R. Setiawan, "Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 1, pp. 105–116, 2023, doi: 10.32736/sisfokom.v12i1.1614.
- [13] M. Iqbal, M. Afdal, and R. Novita, "Implementation of Support Vector Machine Algorithm for Sentiment Analysis of Online Loan Application Review Data on Google Play Store Implementasi Algoritma Support Vector Machine untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online Di Google Play Store," vol. 4, no. October, pp. 1244–1252, 2024.