

Analisis Sentimen Ulasan Dan Komentar Pengguna Pada Aplikasi Toco Menggunakan Algoritma Support Vector Machine

Sentiment Analysis of User Reviews and Comments on the Toco Application Using the Support Vector Machine Algorithm

Nuraulia Octaviana*¹, Eva Yumami², Desi Wahana³

^{1, 2, 3}Program Studi Rekayasa Perangkat Lunak, Politeknik Negeri Bengkalis, Sungai Alam, Bengkalis, Riau, +62822 3572 0044

E-mail : nurauliaaa12@gmail.com*¹, evayumami@polbeng.ac.id², desiwahana@polbeng.ac.id³

*Corresponding author

Received 26 March 2026; Revised 21 April 2026; Accepted 24 April 2026

Abstrak - Penelitian ini bertujuan untuk menganalisis sentimen ulasan dan komentar pengguna terhadap aplikasi Toco menggunakan algoritma Support Vector Machine (SVM). Seiring dengan meningkatnya jumlah pengguna aplikasi, volume ulasan yang dihasilkan juga semakin besar sehingga diperlukan metode otomatis untuk mengklasifikasikan opini pengguna secara efektif. Metode yang digunakan dalam penelitian ini adalah Knowledge Discovery in Database (KDD) yang meliputi tahapan seleksi data, preprocessing, transformasi data menggunakan TF-IDF, pembagian data, klasifikasi, dan evaluasi. Dataset yang digunakan berjumlah 3000 ulasan yang diperoleh dari Google Play Store, TikTok, dan Instagram, kemudian dilabeli secara manual menjadi sentimen positif dan negatif. Hasil pengujian menunjukkan bahwa model SVM mampu menghasilkan akurasi sebesar 93,17%, precision 97,89%, recall 83,41%, dan F1-score 90,07%. Hasil tersebut menunjukkan bahwa metode SVM efektif dalam melakukan analisis sentimen pada data teks berbahasa Indonesia. Penelitian ini diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap aplikasi Toco serta menjadi bahan evaluasi bagi pengembang dalam meningkatkan kualitas layanan aplikasi.

Kata kunci - Analisis Sentimen, Support Vector Machine, TF-IDF, Ulasan Pengguna, Aplikasi Toco

Abstract - This study aims to analyze the sentiment of user reviews and comments on the Toco application using the Support Vector Machine (SVM) algorithm. As the number of application users increases, the volume of generated reviews also grows, making it necessary to implement an automated method for effectively classifying user opinions. The research method used is Knowledge Discovery in Databases (KDD), which consists of data selection, preprocessing, data transformation using TF-IDF, data splitting, classification, and evaluation. The dataset consists of 3,000 reviews collected from Google Play Store, TikTok, and Instagram, which were manually labeled into positive and negative sentiments. The experimental results show that the SVM model achieved an accuracy of 93.17%, precision of 97.89%, recall of 83.41%, and F1-score of 90.07%. These results indicate that the SVM method is effective in performing sentiment analysis on Indonesian text data. This research is expected to provide insights into user perceptions of the Toco application and serve as evaluation material for developers to improve application service quality.

Keywords - Sentiment Analysis, Support Vector Machine, TF-IDF, User Reviews, Toco App

1. PENDAHULUAN

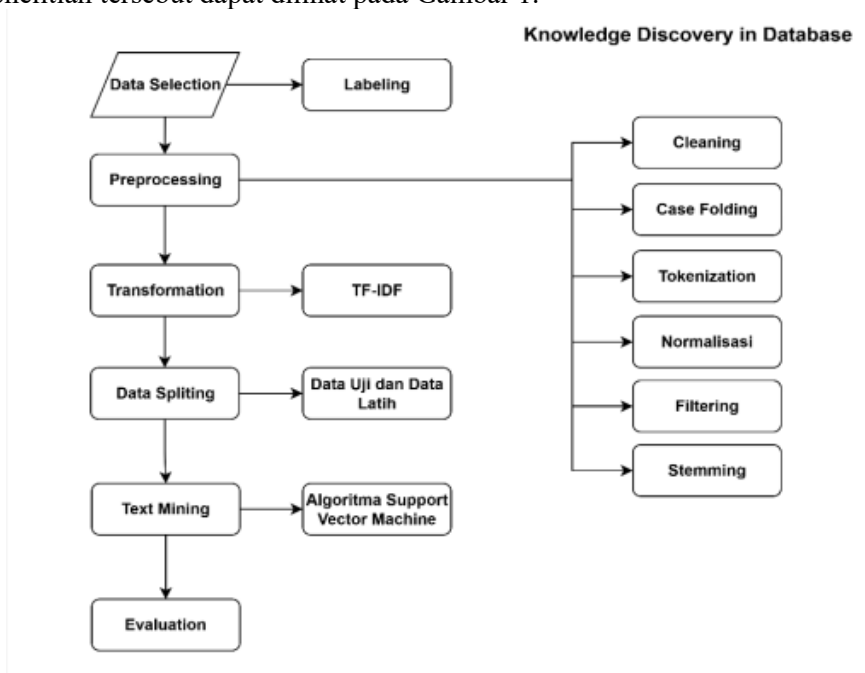
Perkembangan teknologi digital telah mendorong peningkatan penggunaan aplikasi dalam berbagai aspek kehidupan masyarakat. Aplikasi kini dimanfaatkan untuk berbagai kebutuhan, seperti transaksi digital, layanan publik, hingga aktivitas perdagangan. Kemudahan akses serta kecepatan dalam memperoleh informasi menjadikan aplikasi sebagai sarana yang semakin penting dalam menunjang aktivitas sehari-hari. Kondisi ini menuntut pengembang aplikasi untuk tidak hanya memperhatikan fungsionalitas sistem, tetapi juga kualitas pengalaman pengguna yang berpengaruh terhadap tingkat kepuasan dan loyalitas pengguna[1, 2]. Salah satu sumber informasi yang dapat digunakan untuk mengetahui pengalaman pengguna adalah ulasan yang diberikan melalui berbagai platform digital, seperti Google Play Store maupun media sosial. Ulasan tersebut berisi berbagai opini pengguna yang mencerminkan tingkat kepuasan, kritik, maupun saran terhadap layanan yang digunakan. Informasi ini dapat menjadi sumber data yang penting bagi pengembang dalam mengevaluasi kualitas layanan aplikasi[3]. Namun, jumlah ulasan yang terus meningkat serta variasi bahasa yang digunakan oleh pengguna menyebabkan proses analisis secara manual menjadi kurang efektif dan berpotensi mengabaikan informasi penting[4]. Oleh karena itu, diperlukan pendekatan berbasis teknologi yang mampu mengelompokkan sentimen pengguna secara otomatis untuk mengetahui kecenderungan opini yang bersifat positif maupun negatif. Toco merupakan salah satu platform digital yang menyediakan berbagai layanan dalam satu aplikasi terpadu. Aplikasi ini dikembangkan sejak tahun 2022 dengan tujuan mempermudah aktivitas pengguna melalui integrasi layanan belanja, transaksi, serta konsultasi digital. Toco menyediakan dua layanan utama, yaitu Toco Marketplace untuk transaksi jual beli produk umum dan Toco Classified sebagai platform iklan baris untuk berbagai kategori produk seperti properti dan kendaraan. Seiring dengan meningkatnya jumlah pengguna aplikasi, jumlah ulasan yang diberikan oleh pengguna juga terus bertambah sehingga diperlukan metode yang mampu mengelola dan menganalisis data ulasan tersebut secara efektif[5]. Salah satu metode yang dapat digunakan dalam proses klasifikasi sentimen adalah Support Vector Machine (SVM). Algoritma ini merupakan salah satu metode machine learning yang banyak digunakan dalam permasalahan klasifikasi karena memiliki kemampuan yang baik dalam memisahkan data ke dalam dua kelas dengan tingkat akurasi yang tinggi. SVM bekerja dengan mencari hyperplane optimal yang mampu memisahkan data dari kelas yang berbeda dan terbukti efektif dalam pengolahan data teks berdimensi besar[6, 7].

Beberapa penelitian sebelumnya telah menunjukkan keberhasilan penggunaan metode SVM dalam analisis sentimen. Penelitian yang dilakukan oleh Fitriyana dkk. pada ulasan aplikasi Jamsostek Mobile (JMO) menunjukkan bahwa algoritma SVM mampu mengklasifikasikan lebih dari 17.000 ulasan pengguna secara optimal dengan tingkat akurasi mencapai 96% dan F1-score 94%. Hasil penelitian tersebut menunjukkan bahwa metode SVM efektif digunakan dalam menganalisis sentimen pengguna terhadap layanan digital berbasis aplikasi[8]. Penelitian lain yang dilakukan oleh Yolanda dkk. mengenai analisis sentimen ulasan pengguna aplikasi Sayurbox menggunakan metode SVM dengan pembobotan TF-IDF menghasilkan tingkat akurasi sebesar 89,29%, yang menunjukkan bahwa metode tersebut mampu mengelola data teks berbahasa Indonesia secara akurat pada aplikasi berbasis e-commerce[9]. Penelitian oleh Azzahra W dkk. juga menunjukkan bahwa penerapan SVM dan TF-IDF dalam analisis sentimen layanan RSUD Salatiga mampu menghasilkan akurasi 84% dengan nilai precision dan recall yang seimbang[10]. Selain itu, penelitian oleh Afandi dkk. pada aplikasi MPStore-Super App UMKM menunjukkan bahwa metode SVM mampu mencapai akurasi 90%, sehingga dinilai cukup tangguh dalam melakukan klasifikasi sentimen pada aplikasi berbasis layanan digital[11]. Penelitian lain juga menunjukkan performa tinggi metode SVM pada analisis ulasan pengguna. Penelitian oleh Idris dkk. pada ulasan aplikasi Shopee menghasilkan akurasi sebesar 98% dengan penerapan preprocessing dan pembobotan fitur TF-IDF[6]. Selain itu, penelitian oleh Pratama dkk. pada ulasan Google Review hotel menunjukkan bahwa SVM mampu mencapai akurasi

sebesar 95% dalam mengklasifikasikan sentimen pengguna[12]. Di sisi lain, beberapa penelitian juga mengembangkan pendekatan berbasis deep learning seperti BERT dan Bi-LSTM yang memiliki kemampuan lebih baik dalam memahami konteks bahasa yang kompleks[13, 14]. Meskipun demikian, metode SVM masih banyak digunakan karena memiliki performa yang stabil, efisien, serta mampu memberikan hasil klasifikasi yang baik pada data teks berdimensi besar[5, 15]. Berdasarkan berbagai penelitian terdahulu tersebut, penelitian ini menerapkan metode Support Vector Machine (SVM) untuk melakukan analisis sentimen terhadap ulasan dan komentar pengguna aplikasi Toco. Hasil penelitian diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap layanan aplikasi serta menjadi bahan evaluasi bagi pengembang dalam meningkatkan kualitas layanan. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan dan komentar pengguna aplikasi Toco menggunakan algoritma Support Vector Machine (SVM). Melalui pendekatan ini diharapkan dapat diperoleh gambaran mengenai kecenderungan opini pengguna terhadap aplikasi Toco, baik yang bersifat positif maupun negatif. Hasil analisis diharapkan dapat menjadi masukan bagi pengembang dalam meningkatkan kualitas layanan aplikasi serta mendukung pengambilan keputusan berbasis data dalam pengembangan sistem.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini mengikuti alur Knowledge Discovery in Database (KDD) yang terdiri dari beberapa proses utama, yaitu seleksi data, preprocessing, transformasi data, pembagian data, proses klasifikasi, serta evaluasi hasil. Alur tahapan penelitian tersebut dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian Menggunakan Metode KDD

2.1. Data Selection

Berdasarkan Gambar 1, tahapan penelitian dimulai dari data selection, yaitu proses pengumpulan data ulasan dan komentar pengguna aplikasi Toco dari beberapa platform digital seperti Google Play Store, TikTok, dan Instagram menggunakan teknik web scraping. Data yang dikumpulkan berjumlah 3000 komentar dan ulasan pengguna yang berkaitan dengan pengalaman penggunaan aplikasi. Seluruh data kemudian melalui proses pelabelan manual untuk menentukan

kategori sentimen positif dan negatif. Proses pelabelan dilakukan secara manual dan divalidasi oleh ahli bahasa untuk menjaga konsistensi klasifikasi.

2.2. Preprocessing

Tahap selanjutnya adalah preprocessing, yang bertujuan untuk membersihkan dan menyiapkan data teks agar siap dianalisis. Proses ini meliputi beberapa langkah, yaitu cleaning, case folding, tokenization, normalisasi, filtering, dan stemming. Tahapan ini dilakukan untuk menghilangkan karakter yang tidak relevan, menyeragamkan bentuk kata, serta mengurangi variasi kata sehingga data menjadi lebih terstruktur.

2.3. Transformasi

Setelah preprocessing, data kemudian memasuki tahap transformasi dengan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi.

2.4. Data Splitting

Selanjutnya dilakukan data splitting, yaitu pembagian dataset menjadi data latih sebesar 80% dan data uji sebesar 20%. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengukur performa model yang dihasilkan.

2.5. Text Mining

Tahap modelling dilakukan dengan menerapkan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan sentimen ulasan pengguna. Model ini dilatih menggunakan data latih dan kemudian diuji menggunakan data uji untuk melihat kemampuan model dalam mengenali pola sentimen.

2.6. Evaluation

Tahap terakhir adalah evaluation, yaitu proses pengukuran kinerja model menggunakan confusion matrix yang menghasilkan nilai accuracy, precision, recall, dan F1-score. Evaluasi ini bertujuan untuk mengetahui tingkat efektivitas algoritma SVM dalam mengklasifikasikan sentimen ulasan dan komentar pengguna aplikasi Toco.

Selain tahapan pengolahan data, penelitian ini juga merancang sebuah sistem sederhana untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Toco. Sistem ini dirancang untuk memproses teks ulasan yang dimasukkan oleh pengguna dan mengklasifikasikan sentimen menggunakan algoritma Support Vector Machine (SVM). Perancangan sistem bertujuan untuk mempermudah proses analisis sentimen secara otomatis serta memberikan gambaran mengenai persepsi pengguna terhadap aplikasi Toco.

3. HASIL DAN PEMBAHASAN

3.1. Data Selection

Pada tahap ini dilakukan pengambilan data ulasan dan komentar pengguna aplikasi Toco dari beberapa platform digital. Sebagian kecil data digunakan sebagai contoh proses analisis manual sebelum dilakukan pengolahan menggunakan sistem. Data tersebut kemudian diberi label sentimen positif dan negatif berdasarkan isi ulasan pengguna. Contoh data ulasan yang telah diberi label dapat dilihat pada Tabel 1.

Tabel 1. Pelabelan data ulasan

No	Dokumen	Label
1	aplikasinya bgus dan SANGAT bermanfaat cari barang jdi lebih mudah anti ribett	Positif
2	top, aplikasi nya ringan dan mudah utk digunakan, CS nya juga fast respon, rekomended	Positif
3	Website jualannya lengkap, barangnya mudah untuk di cari, aplikasinya mudah dimengerti, aplikasi nya rekomendasi banget	Positif

4	Bintang 4 dlu,aplikasi lemot dan banyak error nya	Negatif
5	Betulan dulu aplikasinya,msh byk yg ngeluh lemot dan mah sepi penjualannya.jgn klo nanti udh rame jd byk tingkah naiikin biaya dan nambahin biaya seenak2nya ya	Negatif

Data tersebut kemudian diproses lebih lanjut melalui tahapan preprocessing sebelum dilakukan proses klasifikasi.

3.2. Preprocessing

Tahap preprocessing bertujuan untuk membersihkan dan menormalisasi data teks agar siap digunakan dalam proses analisis. Tahapan preprocessing yang dilakukan meliputi cleaning, case folding, normalisasi, tokenization, filtering (stopword removal), dan stemming. Sebagai contoh, proses preprocessing pada salah satu ulasan dapat dilihat pada Tabel 2.

Tabel 2. Preprocessing

Tahap	Hasil
Teks Awal	aplikasinya bgus dan SANGAT bermanfaat cari barang jdi lebih mudah anti ribett
Case Folding	aplikasinya bgus dan sangat bermanfaat cari barang jdi lebih mudah anti ribett
Normalisasi	aplikasinya bagus dan sangat bermanfaat cari barang jadi lebih mudah anti ribet
Tokenization	[aplikasinya, bagus, sangat, bermanfaat, cari, barang, mudah]
Filtering	[aplikasi, bagus, manfaat, cari, barang, mudah]
Stemming	[aplikasi, bagus, manfaat, cari, barang, mudah]

Melalui proses ini, teks ulasan yang awalnya tidak terstruktur dapat diubah menjadi data yang lebih bersih dan terstandarisasi sehingga memudahkan proses analisis selanjutnya.

3.3. Transformasi Data (TF-IDF)

Setelah melalui tahap preprocessing, data teks selanjutnya ditransformasikan ke dalam bentuk numerik menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya dalam dokumen serta tingkat keunikannya pada keseluruhan dataset. Nilai Term Frequency (TF) dihitung menggunakan persamaan (1) berikut:

$$TF = \frac{\text{Jumlah kemunculan kata dalam dokumen}}{\text{Total kata dalam dokumen}} \tag{1}$$

Sedangkan nilai Inverse Document Frequency (IDF) dihitung dengan persamaan (2) :

$$IDF = \log \left(\frac{N}{df} \right) \tag{2}$$

di mana N merupakan jumlah total dokumen dan df merupakan jumlah dokumen yang mengandung kata tersebut. Bobot akhir setiap kata diperoleh dengan mengalikan nilai TF dan IDF menggunakan persamaan (3) berikut:

$$TF - IDF = TF \times IDF \tag{3}$$

Contoh hasil pembobotan TF-IDF pada beberapa kata dapat dilihat pada Tabel 3 berikut.

Tabel 3. Hasil pembobotan TF-IDF

Token	D1	D2	D3
aplikasi	0.0369	0.0369	0.0369
bagus	0.0871	0	0
barang	0.0871	0	0.0871
mudah	0.0871	0	0
rekomendasi	0	0.0871	0.0871

Hasil pembobotan tersebut kemudian menghasilkan vektor fitur TF-IDF yang digunakan sebagai representasi numerik dari setiap dokumen. Vektor ini selanjutnya digunakan sebagai input pada proses klasifikasi menggunakan algoritma Support Vector Machine (SVM).

3.4. Klasifikasi Menggunakan Support Vector Machine (SVM)

Pada tahap ini dilakukan proses klasifikasi sentimen menggunakan algoritma Support Vector Machine (SVM). Dataset yang telah melalui tahap preprocessing dan pembobotan TF-IDF kemudian dibagi menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%. Model SVM dibangun menggunakan data latih untuk menemukan hyperplane terbaik yang dapat memisahkan kelas sentimen positif dan negatif. Selanjutnya model yang telah dilatih diuji menggunakan data uji untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Toco. Proses klasifikasi dilakukan dengan memanfaatkan nilai vektor TF-IDF dari setiap dokumen sebagai fitur input. Model kemudian menghasilkan prediksi berupa kelas sentimen positif atau negatif berdasarkan fungsi keputusan dari algoritma SVM. Hasil klasifikasi selanjutnya dievaluasi menggunakan beberapa metrik evaluasi seperti accuracy, precision, recall, dan F1-score untuk mengetahui performa model dalam mengidentifikasi sentimen pengguna secara otomatis.

3.5. Dataset dan Skenario Pengujian

```

Preview data:
                                     content
0  Wah mantap sekali aplikasi ini... Sangat mudah...
1  """"Terus berkembang yaa btw cara hapus history...
2  Toco benar-benar memudahkan saya! Aplikasi ini...
3  Aplikasi Toco sangat membantu untuk cari baran...
4  Cuma bisa nyundul produk terus walau ga ada yg...
-----
Info dataset:
Shape: (3000, 1)
Columns: ['content']
Sample content:
Wah mantap sekali aplikasi ini... Sangat mudah diakses dan sangat ringan
    
```

Gambar 2. Hasil dan Jumlah Dataset

Berdasarkan Gambar 4, dataset dalam penelitian ini terdiri dari 3000 data ulasan dan komentar pengguna aplikasi Toco yang dikumpulkan dari beberapa platform digital, yaitu Google Play Store, TikTok, dan Instagram. Data yang diperoleh mencerminkan opini pengguna terhadap berbagai aspek aplikasi, seperti kemudahan penggunaan, performa sistem, fitur yang tersedia, serta pengalaman pengguna secara keseluruhan. Sebelum digunakan dalam proses analisis, data terlebih dahulu melalui tahap preprocessing yang mencakup cleaning, case folding, tokenization, normalisasi, filtering, dan stemming. Tahapan ini bertujuan untuk meningkatkan kualitas data dengan menghilangkan noise serta menyeragamkan struktur teks agar lebih mudah diproses oleh algoritma klasifikasi. Dari total 3000 data ulasan yang digunakan dalam penelitian ini, diperoleh distribusi sentimen sebagai berikut:

- a. Sentimen positif sebanyak 1115 data
- b. Sentimen negatif sebanyak 1885 data

Distribusi ini menunjukkan bahwa dataset cenderung tidak seimbang, di mana jumlah sentimen negatif lebih dominan dibandingkan sentimen positif. Kondisi ini dapat mempengaruhi performa model dalam proses klasifikasi, khususnya dalam mengenali pola pada masing-masing kelas sentimen. Selanjutnya, dataset dibagi menjadi dua bagian menggunakan metode train-test split dengan komposisi:

- a. 80% data latih (2400 data)
- b. 20% data uji (600 data)

Pembagian ini bertujuan untuk melatih model sekaligus menguji kemampuan generalisasi model terhadap data baru yang belum pernah dipelajari sebelumnya.

3.6. Transformasi Data Menggunakan TF-IDF

Data teks hasil preprocessing kemudian ditransformasikan ke dalam bentuk numerik menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculan dalam dokumen serta tingkat keunikannya dalam keseluruhan dataset.

Dengan demikian, proses transformasi dilakukan dengan membangun model TF-IDF pada data latih, kemudian model tersebut digunakan untuk mentransformasikan data uji ke dalam representasi vektor numerik. Setiap dokumen direpresentasikan sebagai vektor berdimensi tinggi yang mencerminkan bobot kata-kata penting dalam dokumen tersebut.

Hasil dari proses ini berupa matriks TF-IDF yang berisi nilai bobot setiap kata terhadap dokumen. Representasi ini memungkinkan algoritma klasifikasi untuk mengenali pola penting dalam data teks. Dengan demikian, metode TF-IDF berperan dalam meningkatkan kualitas fitur yang digunakan dalam proses klasifikasi sentimen menggunakan Support Vector Machine (SVM).

3.7. Hasil Klasifikasi Menggunakan Support Vector Machine (SVM)

Proses klasifikasi dilakukan menggunakan algoritma Support Vector Machine (SVM) dengan kernel linear. Algoritma ini dipilih karena memiliki kemampuan yang baik dalam menangani data berdimensi tinggi seperti hasil transformasi TF-IDF. Model dilatih menggunakan data latih sebanyak 2400 data dan diuji menggunakan 600 data uji. Hasil klasifikasi ditampilkan dalam bentuk confusion matrix sebagai berikut:

Tabel 4. Hasil klasifikasi

	Prediksi Positif	Prediksi Negatif
Aktual Positif	186	37
Aktual Negatif	4	373

Berdasarkan hasil tersebut, diperoleh nilai evaluasi sebagai berikut:

- a. Accuracy : 93,17%
- b. Precision : 97,89%
- c. Recall : 83,41%
- d. F1-Score : 90,07%

Nilai-nilai tersebut menunjukkan bahwa model memiliki performa yang cukup baik dalam mengklasifikasikan sentimen ulasan pengguna.

3.8. Analisis Performa Model

Nilai accuracy sebesar 93,17% menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Hal ini mengindikasikan bahwa kombinasi antara preprocessing, TF-IDF, dan SVM efektif dalam menangani data teks berbahasa Indonesia. Nilai precision yang tinggi (97,89%) menunjukkan bahwa model sangat baik dalam mengidentifikasi sentimen positif. Artinya, sebagian besar prediksi positif yang dihasilkan oleh model benar-benar sesuai dengan kondisi aktual. Namun, nilai recall sebesar 83,41% menunjukkan bahwa masih terdapat data positif yang tidak berhasil dikenali oleh model. Hal ini dapat disebabkan oleh variasi bahasa yang digunakan oleh pengguna, seperti penggunaan singkatan, bahasa tidak baku, serta konteks kalimat yang ambigu. Nilai F1-Score sebesar 90,07% menunjukkan bahwa model

memiliki keseimbangan yang baik antara precision dan recall, sehingga dapat dikatakan memiliki performa yang stabil.

3.9. Analisis Kesalahan Klasifikasi

Meskipun performa model tergolong tinggi, masih terdapat kesalahan klasifikasi yang perlu dianalisis. Berdasarkan confusion matrix, terdapat:

- a. 37 data false negative
- b. 4 data false positive

Kesalahan ini umumnya disebabkan oleh beberapa faktor, antara lain:

a. Ambiguitas Makna Kata

Beberapa ulasan memiliki makna yang bergantung pada konteks sehingga sulit dipahami oleh model.

b. Bahasa Tidak Baku

Penggunaan bahasa sehari-hari, singkatan, atau slang menyebabkan model kesulitan dalam mengenali pola.

c. Keterbatasan TF-IDF

TF-IDF hanya mempertimbangkan frekuensi kata tanpa memahami konteks kalimat secara mendalam.

d. Ketidakseimbangan Pola Data

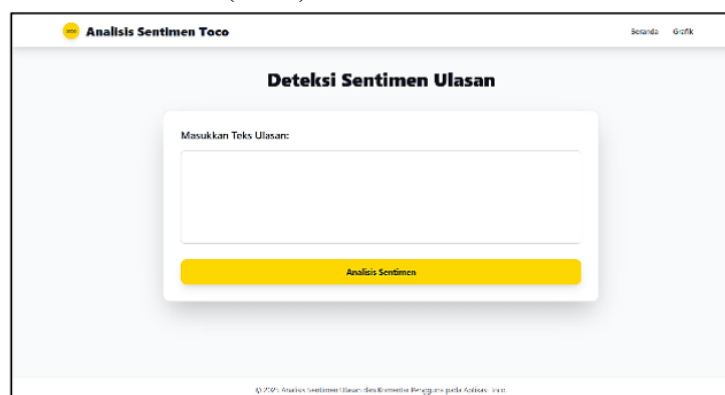
Model cenderung lebih optimal dalam mengenali sentimen positif dibandingkan negatif.

3.10. Perbandingan dengan Penelitian Sebelumnya

Hasil penelitian ini menunjukkan bahwa algoritma SVM mampu menghasilkan performa yang kompetitif dalam analisis sentimen. Dengan akurasi sebesar 93,17%, hasil ini sejalan dengan berbagai penelitian sebelumnya yang juga menunjukkan performa tinggi dari metode SVM dalam klasifikasi teks. Dibandingkan dengan metode lain, SVM memiliki keunggulan dalam hal kestabilan dan efisiensi, terutama pada dataset berukuran sedang. Namun demikian, pendekatan berbasis deep learning seperti BERT memiliki keunggulan dalam memahami konteks bahasa yang lebih kompleks. Meskipun demikian, dari segi implementasi dan kebutuhan komputasi, SVM tetap menjadi metode yang relevan dan efektif untuk analisis sentimen, khususnya dalam konteks penelitian akademik.

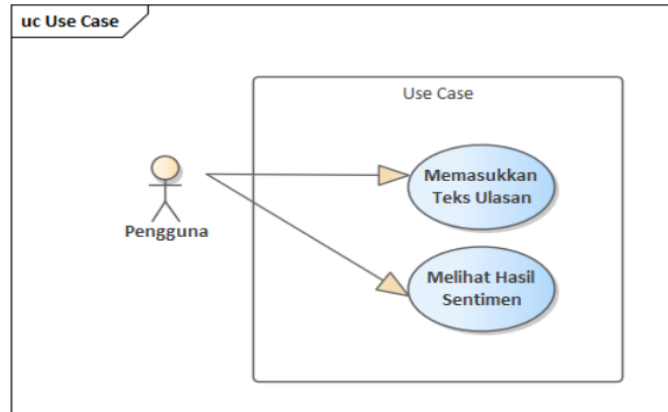
3.11. Implementasi Sistem

Pada tahap implementasi, sistem dirancang dalam bentuk antarmuka berbasis web yang memungkinkan pengguna untuk melakukan analisis sentimen secara otomatis. Pengguna dapat memasukkan teks ulasan melalui halaman input yang tersedia, kemudian sistem akan memproses teks tersebut melalui tahapan preprocessing, transformasi TF-IDF, dan klasifikasi menggunakan algoritma Support Vector Machine (SVM).

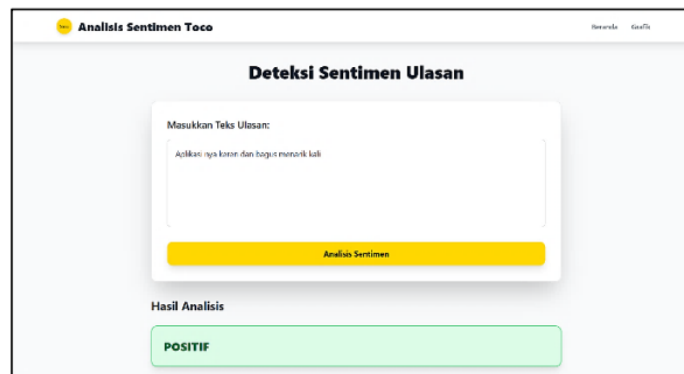


Gambar 3. Tampilan Awal Website

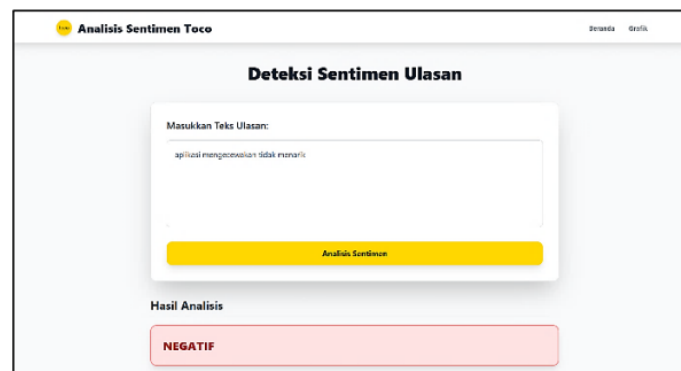
Berdasarkan Gambar 3, sebagai bentuk implementasi dari penelitian ini, dikembangkan sebuah sistem berbasis web sederhana untuk melakukan analisis sentimen secara otomatis. Sistem ini memungkinkan pengguna untuk memasukkan teks ulasan, yang kemudian diproses melalui tahapan preprocessing, transformasi TF-IDF, dan klasifikasi menggunakan SVM.



Gambar 4. Use Case Pengguna



Gambar 5. Analisis Positif



Gambar 6. Analisis Negatif

Berdasarkan Gambar 5 dan Gambar 6, hasil klasifikasi ditampilkan secara langsung dalam bentuk kategori sentimen, yaitu positif atau negatif. Implementasi ini menunjukkan bahwa model yang dikembangkan tidak hanya bersifat teoritis, tetapi juga dapat diterapkan secara praktis dalam membantu memahami opini pengguna.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma Support Vector Machine (SVM) dalam analisis sentimen ulasan pengguna aplikasi Toco mampu memberikan hasil yang baik dan akurat. Dengan menggunakan tahapan metode KDD yang meliputi data selection, preprocessing, transformasi TF-IDF, serta proses klasifikasi, sistem berhasil mengolah data teks yang tidak terstruktur menjadi informasi yang bernilai. Hasil evaluasi menunjukkan bahwa model SVM memperoleh nilai akurasi sebesar 93,17%, precision 97,89%, recall 83,41%, dan F1-score 90,07%. Nilai tersebut menunjukkan bahwa model memiliki performa yang tinggi dan stabil dalam mengklasifikasikan sentimen positif dan negatif. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi yang disebabkan oleh faktor seperti penggunaan bahasa tidak baku, ambiguitas makna, serta keterbatasan metode TF-IDF dalam memahami konteks kalimat.

Secara keseluruhan, penelitian ini berhasil mencapai tujuan yaitu mengidentifikasi kecenderungan sentimen pengguna terhadap aplikasi Toco secara otomatis. Hasil analisis dapat digunakan sebagai bahan evaluasi bagi pengembang untuk meningkatkan kualitas layanan aplikasi. Untuk pengembangan selanjutnya, penelitian ini dapat ditingkatkan dengan menggunakan metode berbasis deep learning seperti BERT atau LSTM agar mampu memahami konteks bahasa secara lebih mendalam serta meningkatkan akurasi model.

REFERENSI

- [1] H. W. Raff and F. Ratnawati, “Klasifikasi Sentimen Ulasan Pengguna terhadap Aplikasi Video Editing Play Store Menggunakan Random Forest,” *Techno.Com*, vol. 24, no. 3, pp. 646–657, Aug. 2025, doi: 10.62411/tc.v24i3.12873.
- [2] M. I. Burhan, “Analisis Sentimen Publik Terhadap Aplikasi MyPertamina Pada Google Playstore,” *Jurnal Manajemen Perbankan Keuangan Nitro*, vol. 8, no. 1, pp. 1–21, 2021, doi: 10.56858/jmpkn.v8i1.390.
- [3] F. Rahmasari, N. Rahaningsih, R. D. Dana, and C. L. Rohmat, “Optimasi Analisis Sentimen Aplikasi Glints Menggunakan Algoritma Support Vector Machine (SVM),” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, pp. 558–570, Jan. 2025, doi: 10.23960/jitet.v13i1.5681.
- [4] D. N. I. Huda, C. Prianto, and R. M. Awangga, “Analisis Sentimen Perbandingan Layanan Jasa Pengiriman Kurir Pada Ulasan Play Store Menggunakan Metode Random Forest dan Decision Tree,” *Jurnal Ilmiah Informatika (JIF)*, vol. 11, pp. 150–158, Sep. 2023.
- [5] M. Maulidah, A. Ardiansyah, S. Suleman, L. P. Gemilang, and N. F. Indriarti, “Analisis Sentimen Pada Ulasan Aplikasi Superbank Dengan Metode Support Vector Machine Dan Naive Bayes,” *Indonesian Journal on Software Engineering (IJSE)*, vol. 10, no. 2, pp. 147–155, Dec. 2024, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijse>
- [6] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, “Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM),” *Jambura Journal of Electrical and Electronics Engineering*, vol. 5, no. 1, pp. 32–35, Jan. 2023.
- [7] R. Raudhah and F. P. Putra, “Klasifikasi Sentimen Ulasan Turis Terhadap Objek Wisata di Bali Menggunakan Support Vector Machine,” *Techno.Com*, vol. 24, no. 3, pp. 720–732, Aug. 2025, doi: 10.62411/tc.v24i3.12868.
- [8] V. Fitriyana, L. Hakim, D. C. R. Novitasari, and A. H. Asyhar, “Analisis Sentimen Ulasan Aplikasi Jamsostek Mobile Menggunakan Metode Support Vector Machine,” *Jurnal Buana Informatika*, vol. 14, no. 1, pp. 40–49, Apr. 2023.
- [9] A. M. Yolanda and R. T. Mulya, “Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store,” *VARIANSI*:

- Journal of Statistics and Its Application on Teaching and Research*, vol. 6, no. 2, pp. 76–83, 2024, doi: 10.35580/variasiunm258.
- [10] W. L. Azzahra and E. Mailoa, “Analisis Sentimen terhadap RSUD Salatiga Menggunakan SVM dan TF-IDF,” *Jurnal Indonesia : Manajemen Informatika dan Komunikasi (JIMIK)*, vol. 6, no. 1, pp. 478–489, Jan. 2025, doi: doi.org/10.35870/jimik.v6i1.1208.
- [11] R. B. Afandi, T. F. Nurdiansyah, A. N. Ramadhani, and A. P. Sari, “Implementasi Support Vector Machine Untuk Analisis Sentimen Aplikasi ‘MPStore-Super App UMKM,’” *JIP (Jurnal Informatika Polinema)*, vol. 10, no. 4, pp. 565–569, Aug. 2024.
- [12] H. I. Pratama and P. T. Prasetyaningrum, “Penerapan Support Vector Machine untuk Analisis Sentimen pada Google Review Hotel,” *Journal of Information System Research (JOSH)*, vol. 6, no. 2, pp. 1244–1252, Jan. 2025, doi: 10.47065/josh.v6i2.6645.
- [13] K. S. Nugroho, A. Y. Sukmadewa, H. W. DW, F. A. Bachtiar, and N. Yudistira, “BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews,” Jul. 2021. doi: doi.org/10.48550/arXiv.2107.06802.
- [14] A. Purwarianti, I. A. Putu, and A. Crisdayanti, “Improving Bi-LSTM Performance for Indonesian Sentiment Analysis Using Paragraph Vector,” Sep. 2020. doi: doi.org/10.1109/ICAICTA.2019.8904199.
- [15] R. Satria Buana, W. Gata, A. Zevana, P. Widodo, H. Setiawan, and K. Hilyati, “Analisis Sentimen pada Komen Twitter Pawang Hujan Mandalika dengan Support Vector Machine (SVM) dan Naïve Bayes,” *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 7, no. 2, pp. 194–200, Feb. 2023, doi: 10.35870/jti.