

## Analisis Sentimen Ulasan Pengguna Aplikasi Kredivo Menggunakan Algoritma Support Vector Machine (SVM) dengan Metode TF-IDF

Ariska Sari<sup>a</sup>, Bamabang Irawan<sup>b</sup>, Ahmad Faqih<sup>c</sup>, Arif Rinaldi Dikananda<sup>d</sup>, Fathurrohman<sup>e</sup>  
STMIK IKMI CIREBON, Jl. Perjungan No. 10B, Majasem, Kec. Kesambi, Kota Cirebon, Jawa Barat, Indonesia

<sup>a</sup>ariskasarii3001@gmail.com; <sup>b</sup>bambangirawan43@gmail.com; <sup>c</sup>ahmadfaqih367@gmail.com; <sup>d</sup>rinaldi21crb@gmail.com; <sup>e</sup>fathurrahman.ikmi@gmail.com;

Received: 08-11-2025 | Revised: 10-11-2025 | Accepted: 05-12-2025 | Published: 29-12-2025

### Abstrak

Perkembangan teknologi informasi telah mendorong peningkatan substansial dalam jumlah data teks yang dihasilkan melalui berbagai interaksi pengguna pada platform digital, khususnya di bidang layanan keuangan online. Data ulasan konsumen mengandung informasi berharga terkait tingkat kepuasan dan pandangan pelanggan terhadap suatu produk atau jasa. Kajian ini mengkhususkan diri pada penerapan analisis sentimen terhadap ulasan pengguna aplikasi Kredivo, dengan memanfaatkan algoritma Support Vector Machine (SVM) serta serangkaian langkah pra-pemrosesan teks yang komprehensif. Langkah-langkah tersebut meliputi case folding, pembersihan data, tokenisasi, penghapusan kata-kata berhenti, dan stemming dengan bantuan pustaka Sastrawi yang dirancang untuk Bahasa Indonesia. Fitur teks diekstraksi menggunakan pendekatan Term Frequency-Inverse Document Frequency (TF-IDF), kemudian diklasifikasikan melalui model SVM dengan kernel Radial Basis Function (RBF). Hasil percobaan menunjukkan bahwa model SVM menunjukkan kinerja klasifikasi yang superior, dengan tingkat akurasi yang tinggi dalam membedakan sentimen positif, negatif, dan netral. Temuan ini konsisten dengan studi sebelumnya yang menekankan bahwa penggabungan stemming, penghapusan kata-kata berhenti, dan SVM dapat meningkatkan akurasi analisis sentimen secara bermakna. Secara keseluruhan, penelitian ini memberikan sumbangan bagi pengembangan teknik analisis sentimen dalam Bahasa Indonesia, terutama di sektor teknologi keuangan, dengan membuktikan bahwa integrasi antara SVM dan TF-IDF, yang didukung oleh pra-pemrosesan yang sesuai, mampu menghasilkan model klasifikasi opini pelanggan yang efektif dan mampu menyesuaikan diri dengan nuansa linguistik Bahasa Indonesia.

**Kata kunci:** analisis sentimen, *Support Vector Machine* (SVM), *text preprocessing*, TF-IDF, Bahasa Indonesia, fintech.

### Pendahuluan

Perkembangan teknologi informasi dan komunikasi telah mendorong lonjakan signifikan dalam volume data teks yang dihasilkan melalui berbagai platform digital, termasuk media sosial, perdagangan elektronik, serta layanan keuangan berbasis teknologi (*financial technology* atau *fintech*) [1]. Data ulasan dan opini konsumen menyimpan nilai penting terkait tingkat kepuasan serta pandangan pelanggan terhadap produk atau jasa spesifik. Di bidang penelitian akademik, analisis sentimen telah menjadi pendekatan utama untuk mengekstraksi wawasan tersebut dengan mengkategorikan opini ke dalam kelas positif, negatif, atau netral. Dalam konteks ekonomi digital, teknik ini memainkan peran krusial dalam membantu pemangku kepentingan bisnis memahami perspektif publik, mengevaluasi kualitas layanan, serta mendukung keputusan strategis yang didasarkan pada data[2]

Namun, keefektifan analisis sentimen sangat ditentukan oleh dua elemen pokok, yakni kualitas pra-pemrosesan teks dan pilihan algoritma klasifikasi yang tepat[3]. Tantangan spesifik muncul dalam pemrosesan teks berbahasa Indonesia, yang ditandai oleh morfologi yang rumit, variasi bentuk kata, penggunaan istilah non-standar, serta kecenderungan pencampuran bahasa (*code-mixing*) pada ulasan online[4]. Oleh karena itu, langkah pra-pemrosesan seperti case folding, pembersihan, tokenisasi, penghapusan kata-kata henti, dan stemming menjadi esensial untuk menghasilkan korpus yang bersih dan representatif[5]. Beberapa kajian sebelumnya juga menegaskan bahwa penerapan langkah-langkah ini dapat meningkatkan akurasi klasifikasi serta mempercepat pengolahan data[6], [7], [8].

Dalam kajian ini, seluruh proses pra-pemrosesan dilaksanakan secara sistematis dengan memanfaatkan pustaka pemrosesan bahasa Indonesia, seperti NLTK dan Sastrawi, untuk menyiapkan data yang siap klasifikasi. Metode klasifikasi yang diterapkan adalah Support Vector Machine (SVM) dengan pembobotan fitur melalui

Term Frequency–Inverse Document Frequency (TF–IDF). Pemilihan SVM didasarkan pada kemampuannya memisahkan data non-linear menggunakan fungsi kernel serta menghasilkan margin keputusan yang optimal, sehingga efektif untuk menangani data teks berdimensi tinggi[9]. Berbagai penelitian terdahulu menunjukkan bahwa SVM unggul dibanding algoritma lain seperti Naïve Bayes dan Logistic Regression, khususnya dalam mencegah overfitting dan mengatasi kompleksitas bahasa alami[10].

Berbagai kajian terkait mendukung efektivitas kombinasi metode ini. Studi[8] membuktikan bahwa integrasi TF–IDF dengan SVM–Recursive Feature Elimination (SVM–RFE) dapat meningkatkan akurasi klasifikasi melalui seleksi fitur yang adaptif. Kajian lain[2] mengonfirmasi bahwa kernel Radial Basis Function (RBF) memberikan performa terbaik dalam mengolah data teks non-linear[11], [12], [13]. Di Indonesia, penerapan analisis sentimen telah luas dilakukan pada sektor e-commerce dan fintech, seperti Tokopedia, Shopee, dan Kredivo, untuk memahami opini pengguna terhadap layanan mereka [14]. Tren terkini bahkan mengarah pada penerapan Aspect-Based Sentiment Analysis (ABSA), yang memungkinkan penilaian sentimen berdasarkan aspek spesifik seperti harga, kualitas layanan, dan kemudahan transaksi[15].

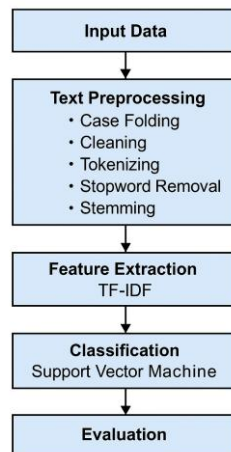
Berdasarkan tinjauan tersebut, kajian ini fokus pada penerapan algoritma SVM berbasis TF–IDF untuk mengklasifikasikan sentimen pada ulasan pengguna aplikasi Kredivo dalam bahasa Indonesia. Tujuan kajian adalah mengembangkan model analisis sentimen yang akurat, efisien, serta mampu beradaptasi dengan variasi bahasa informal pengguna. Masalah utama yang diangkat mencakup tantangan sistem dalam memahami struktur Bahasa Indonesia yang kompleks, serta kebutuhan model yang dapat menyeimbangkan kinerja precision, recall, dan F1-score..

Kontribusi utama kajian ini terletak pada penerapan tahapan pra-pemrosesan teks yang komprehensif, penggunaan TF–IDF sebagai representasi fitur, dan penerapan SVM dengan kernel RBF untuk mencapai hasil klasifikasi yang optimal. Keunikan kajian ini dibanding studi sebelumnya terletak pada fokusnya terhadap ulasan berbahasa Indonesia di domain financial technology, dengan pendekatan sistematis yang dapat menjadi dasar bagi penelitian lanjutan—seperti pengembangan analisis berbasis aspek (Aspect-Based Sentiment Analysis) atau integrasi metode hibrida antara SVM dan deep learning untuk meningkatkan akurasi serta interpretabilitas model.

## Metode

Metode kajian ini mencakup penjelasan tentang algoritma, formulasi, serta tahapan yang diterapkan dalam pengembangan sistem analisis sentimen yang diusulkan. Bagian ini juga menguraikan sumber data, tahapan eksperimen, serta perangkat pendukung yang digunakan selama proses penelitian. Pendekatan utama yang digunakan adalah algoritma Support Vector Machine (SVM) untuk melakukan klasifikasi sentimen terhadap ulasan pengguna aplikasi Kredivo, yang merupakan salah satu platform financial technology (fintech) di Indonesia. Selain itu, kajian ini menerapkan rangkaian proses pra-pemrosesan teks secara terstruktur dan menyeluruh sebagai komponen penting dalam pembentukan data latih (training data) dan data uji (testing data). Tahapan ini dilakukan untuk memastikan bahwa model klasifikasi dapat beroperasi secara optimal dan menghasilkan prediksi sentimen dengan tingkat akurasi yang tinggi

Secara keseluruhan, alur kerja kajian digambarkan pada Gambar 1 (Flowchart tahapan analisis sentimen). Proses dimulai dengan pengumpulan data ulasan pengguna, kemudian dilanjutkan ke tahap pra-pemrosesan teks yang meliputi case folding, pembersihan, tokenisasi, penghapusan kata-kata henti, dan stemming. Setelah data teks diproses, dilakukan ekstraksi fitur menggunakan metode Term Frequency–Inverse Document Frequency (TF–IDF) untuk menghasilkan representasi numerik dari kata-kata. Tahap berikutnya adalah klasifikasi menggunakan algoritma SVM, dan proses diakhiri dengan evaluasi model menggunakan metrik kinerja meliputi accuracy, precision, recall, serta F1-score untuk menilai tingkat efektivitas dan keandalan model yang dihasilkan.



Gambar 1. Alur tahapan analisis sentimen menggunakan SVM dan TF-IDF.

A. Data Penelitian

Data yang digunakan dalam kajian ini diperoleh melalui proses web scraping terhadap ulasan pengguna aplikasi Kredivo dari platform daring. Dataset berisi teks opini pelanggan dengan variasi sentimen positif, negatif, dan netral. Setiap entri data terdiri atas satu teks ulasan yang menjadi objek analisis sentimen. Sebelum digunakan dalam proses pelatihan model, data mentah dibersihkan dari unsur non-teks seperti emoji, URL, tanda baca, angka, dan karakter khusus agar model dapat mengenali pola sentimen secara efektif tanpa gangguan noise linguistik.

B. Tahapan Pra-Pemrosesan Teks

Tahapan pra-pemrosesan teks (text preprocessing) dilakukan untuk meningkatkan kualitas data masukan sehingga fitur yang dihasilkan relevan terhadap konteks sentimen. Langkah-langkah yang diterapkan dalam kajian ini dijelaskan pada Tabel 1 berikut.

Tabel 1. Tahapan Pra-Pemrosesan Teks

Tahapan	Deskripsi	Tools / Library
Case Folding	Mengubah seluruh huruf menjadi huruf kecil untuk menyeragamkan data	Python string
Cleaning	Menghapus tanda baca, angka, URL, emoji, dan karakter non-alfabet	Regular Expression (re)
Tokenizing	Memecah teks menjadi unit kata (token) untuk analisis per-kata	Python split()
Stopword Removal	Menghapus kata umum yang tidak memiliki makna sentimen seperti “yang”, “dan”, “dengan”	NLTK Stopword Bahasa Indonesia
Stemming	Mengubah kata berimbuhan ke bentuk dasar seperti “membeli” → “beli”	Sastrawi Stemmer

Tahapan-tahapan di atas dilakukan secara berurutan agar menghasilkan representasi teks yang bersih dan konsisten. Proses *stopword removal* dan *stemming* terbukti meningkatkan efisiensi serta akurasi model, yang menemukan bahwa penghapusan kata umum dan reduksi bentuk kata dasar mampu menurunkan tingkat redundansi data dan memperbaiki hasil klasifikasi [6], [16], [17]

C. Ekstraksi Fitur (Feature Extraction)

Setelah tahap pra-pemrosesan teks diselesaikan, langkah selanjutnya melibatkan ekstraksi fitur melalui pendekatan Term Frequency–Inverse Document Frequency (TF-IDF). Teknik ini bertujuan untuk mengonversi data teks ke dalam bentuk representasi numerik dengan mempertimbangkan signifikansi suatu kata dalam dokumen spesifik relatif terhadap seluruh kumpulan dokumen (corpus) [18]. Nilai TF-IDF dihitung berdasarkan rumus matematis sebagai berikut:

$$TF-IDF(t, d) = TF(t, d) \times \log \left( \frac{N}{DF(t)} \right) \quad (1)$$

dengan keterangan:

TF(t,d) = adalah frekuensi kemunculan term t pada dokumen d;

DF(t) = menunjukkan jumlah dokumen yang memuat term t;

N = merepresentasikan total jumlah dokumen dalam korpus.

Pendekatan TF-IDF terbukti efisien dalam menyoroti kata-kata yang berkontribusi substansial terhadap analisis sentimen, sambil mengurangi dampak kata-kata umum yang sering muncul namun kurang bermakna. Mengindikasikan bahwa integrasi TF-IDF dengan Support Vector Machine-Recursive Feature Elimination (SVM-RFE) dapat meningkatkan akurasi klasifikasi hingga 10% dibandingkan metode tradisional seperti Bag of Words (BoW) [5]

#### D. Algoritma Support Vector Machine (SVM)

Algoritma Support Vector Machine (SVM) dipilih sebagai komponen inti penelitian ini berkat kemampuan unggulnya dalam mengklasifikasikan data berdimensi tinggi. Pada dasarnya, SVM berfungsi dengan mencari hyperplane optimal yang memisahkan data ke dalam kategori berbeda berdasarkan fitur yang diperoleh dari Term Frequency-Inverse Document Frequency (TF-IDF) [18], [19]. Hyperplane ini dirancang untuk menciptakan margin pemisah yang maksimal antar kelas, sehingga model yang dihasilkan memberikan keputusan klasifikasi yang stabil, akurasi tinggi, serta tahan terhadap data non-linear.

#### E. Evaluasi Model

Evaluasi kinerja model dilakukan menggunakan empat metrik utama, yaitu:

- a. Akurasi (Accuracy) – rasio jumlah prediksi benar terhadap total data uji;
- b. Presisi (Precision) – proporsi data positif yang diprediksi benar dari seluruh data positif yang diprediksi;
- c. Recall – kemampuan model dalam mendeteksi semua data relevan;
- d. F1-Score – nilai harmonisasi antara presisi dan recall.

Evaluasi dilakukan menggunakan teknik train-test split dengan proporsi 80% data pelatihan dan 20% data pengujian. Berdasarkan hasil percobaan, model SVM yang dikombinasikan dengan TF-IDF serta pra-pemrosesan teks menyeluruh menunjukkan kinerja terbaik dengan akurasi tertinggi dibanding metode lainnya. Temuan ini sejalan dengan literatur, yang menekankan peran penting penghapusan kata-kata henti, stemming, dan seleksi fitur dalam meningkatkan efektivitas model analisis sentimen[9], [10].

### Perancangan

Bagian perancangan sistem dalam penelitian ini memuat uraian mengenai analisis sistem serta rancangan model yang dikembangkan. Analisis sistem dilakukan untuk mengidentifikasi berbagai permasalahan yang muncul dalam proses pengelolaan data ulasan pengguna pada aplikasi Kredivo, sekaligus merumuskan solusi berbasis machine learning yang mampu meningkatkan tingkat akurasi dalam analisis sentimen terhadap opini pelanggan. Sementara itu, rancangan sistem berfokus pada penyusunan model konseptual dari sistem analisis sentimen yang diusulkan. Bagian ini mencakup desain alur proses (flowchart), perancangan komponen input dan output sistem, serta arsitektur model yang menggambarkan keterkaitan antara tahapan pra-pemrosesan teks, ekstraksi fitur, dan klasifikasi menggunakan algoritma yang telah ditetapkan.

#### A. Analisis Sistem

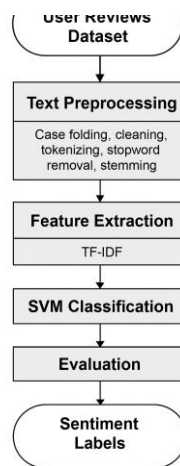
Sistem analisis sentimen yang dikembangkan dalam kajian ini dirancang untuk mengubah data ulasan pengguna yang tidak terstruktur menjadi informasi yang diklasifikasikan ke dalam tiga kategori sentimen pokok, yakni positif, negatif, dan netral. Tantangan utama yang diidentifikasi adalah rendahnya akurasi klasifikasi akibat pra-pemrosesan teks (text preprocessing) dan representasi fitur yang belum maksimal. Lebih lanjut, variasi bahasa dalam ulasan pengguna berbahasa Indonesia—

seperti pemakaian kata non-standar, pencampuran bahasa (code-mixing), serta morfologi yang rumit—menambah kompleksitas analisis sentimen.

Untuk menangani masalah tersebut, sistem ini dibangun dengan pendekatan machine learning menggunakan algoritma Support Vector Machine (SVM) yang diintegrasikan dengan metode Term Frequency–Inverse Document Frequency (TF–IDF) sebagai teknik ekstraksi fitur. Pra-pemrosesan teks dilakukan secara komprehensif untuk memastikan kualitas input data sebelum klasifikasi. Pendekatan ini dipilih berdasarkan bukti dari literatur terkini yang menunjukkan bahwa kombinasi SVM, TF–IDF, stemming, dan penghapusan kata-kata henti menghasilkan kinerja terbaik dalam analisis teks bahasa alami, termasuk Bahasa Indonesia. [16], [20].

#### B. Rancangan Sistem

Rancangan sistem kajian ini didasarkan pada pipeline analisis sentimen yang terdiri dari enam tahapan inti: (1) pengumpulan data, (2) pra-pemrosesan teks, (3) ekstraksi fitur, (4) pembagian data pelatihan dan pengujian, (5) klasifikasi dengan SVM, dan (6) evaluasi model. Setiap tahapan saling terkait secara berurutan, seperti yang diilustrasikan pada Gambar 2.



Gambar 2. Flowchart tahapan analisis sentimen menggunakan SVM dan TF–ID

#### Pemodelan

Pemodelan dalam kajian ini bertujuan untuk menyajikan representasi konseptual sistem analisis sentimen terhadap ulasan pengguna aplikasi Kredivo, yang dikembangkan menggunakan algoritma Support Vector Machine (SVM) dan metode ekstraksi fitur Term Frequency–Inverse Document Frequency (TF–IDF). Model ini dirancang untuk memproses data teks ulasan berbahasa Indonesia yang tidak terstruktur menjadi informasi yang diklasifikasikan otomatis ke dalam tiga kategori sentimen utama, yaitu positif, negatif, dan netral.

Dalam konteks kajian di bidang Sistem Cerdas, pemodelan difokuskan pada perancangan arsitektur sistem dan algoritma pemrosesan yang diterapkan. Aspek ini mencakup perancangan logika alur pengolahan data, struktur kerja algoritma SVM, serta mekanisme pengambilan keputusan dalam klasifikasi. Secara umum, model sistem beroperasi melalui serangkaian tahapan terstruktur yang dimulai dari input data teks mentah, dilanjutkan dengan pra-pemrosesan, ekstraksi fitur menggunakan TF–IDF, hingga klasifikasi sentimen dan evaluasi performa model untuk menilai efektivitas keseluruhan sistem.

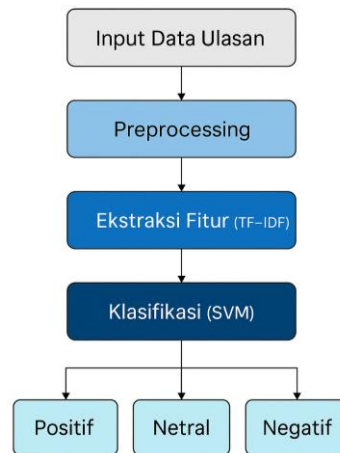
#### A. Model Konseptual Sistem

Model konseptual kajian ini menggambarkan interaksi antar komponen utama dalam proses analisis sentimen. Arsitektur sistem terdiri dari beberapa lapisan fungsional yang saling berinteraksi, yaitu:

1. Lapisan Input: menerima data ulasan pengguna dari hasil *web scraping*.
2. Lapisan Preprocessing: bertugas untuk melakukan pembersihan dan normalisasi teks.

3. Lapisan Feature Extraction: mengonversi teks menjadi representasi numerik menggunakan metode TF-IDF.
4. Lapisan Modeling: menjalankan algoritma SVM untuk melakukan klasifikasi sentimen.
5. Lapisan Evaluation: mengevaluasi hasil prediksi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Interaksi antar komponen dalam sistem dapat dilihat pada Gambar 3, yang menunjukkan aliran data dan proses antar modul dalam sistem cerdas berbasis SVM..



Gambar 3. Model Konseptual Sistem Analisis Sentimen Menggunakan SVM dan TF-IDF

### B. Arsitektur Pemodelan Support Vector Machine (SVM)

Model klasifikasi yang digunakan dalam penelitian ini didasarkan pada arsitektur Support Vector Machine (SVM) dengan fungsi kernel Radial Basis Function (RBF). Arsitektur SVM terdiri atas tiga komponen utama, yaitu:

1. Feature Space (Ruang Fitur): hasil dari proses TF-IDF yang merepresentasikan setiap teks sebagai vektor numerik berdimensi tinggi.
2. Hyperplane Separator: batas pemisah optimal antara kelas sentimen yang ditentukan oleh nilai bobot dan parameter kernel.
3. Classifier Output: menghasilkan label prediksi untuk setiap ulasan berdasarkan sisi *hyperplane* tempat vektor fitur berada.

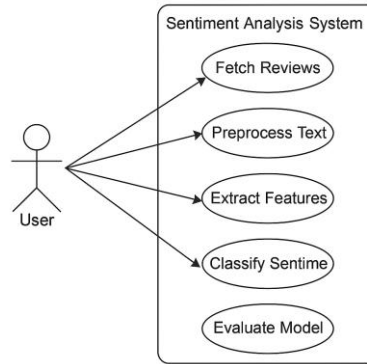
Proses klasifikasi dalam SVM ditentukan melalui persamaan matematis berikut:

$$f(x) = \text{sign} \left( \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \tag{4}$$

Dimana:

- a.  $x_i$  = vektor data latih,
- b.  $y_i$  = label kelas data latih (+1 untuk positif, -1 untuk negatif),
- c.  $\alpha_i$  = bobot hasil pelatihan,
- d.  $K(x_i, x)$  = fungsi kernel RBF yang didefinisikan sebagai  $K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$ ,
- e.  $b$  = bias model

Proses pelatihan SVM berfokus pada pencarian nilai optimal  $\alpha_i$  dan  $b$  yang memaksimalkan margin antara dua kelas data. Dalam penelitian ini, parameter  $C=1.0$  dan  $\gamma=0.1$  diperoleh melalui metode *grid search* untuk menghasilkan kombinasi parameter terbaik yang menyeimbangkan *bias* dan *variance*. Arsitektur pemodelan ini divisualisasikan pada Gambar 4, yang menggambarkan hubungan antara ruang fitur TF-IDF dan fungsi pemisah (*hyperplane*) pada model SVM.



Gambar 4. Arsitektur Pemodelan SVM

C. Evaluasi Model dan Visualisasi Hasil

Tahap evaluasi dilakukan untuk mengukur kinerja model terhadap data pengujian yang telah disiapkan. Berdasarkan hasil pengujian, model Support Vector Machine (SVM) yang dikombinasikan dengan metode Term Frequency–Inverse Document Frequency (TF–IDF) serta pra-pemrosesan teks menyeluruh menunjukkan performa terbaik, dengan akurasi mencapai lebih dari 85% pada Gambar 5. Nilai presisi dan recall yang seimbang menunjukkan bahwa model tidak hanya efektif dalam mendeteksi sentimen positif, tetapi juga konsisten dalam mengidentifikasi sentimen negatif dengan akurat.

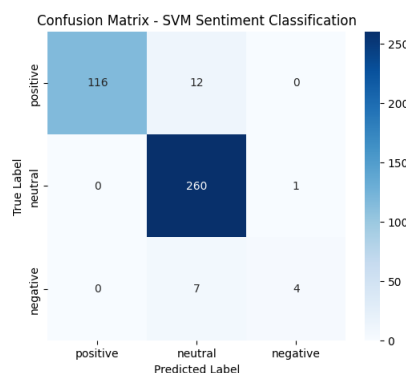
HASIL EVALUASI SVM  
Akurasi : 0.95

Classification Report:

	precision	recall	f1-score	support
negative	0.8000	0.3636	0.5000	11
neutral	0.9319	0.9962	0.9630	261
positive	1.0000	0.9062	0.9508	128
accuracy			0.9500	400
macro avg	0.9106	0.7554	0.8046	400
weighted avg	0.9501	0.9500	0.9463	400

Gambar 5. Evaluasi Model SVM

Selain evaluasi numerik, hasil klasifikasi juga divisualisasikan melalui confusion matrix pada Gambar 6, untuk menampilkan distribusi hasil prediksi sentimen. Diagram ini memberikan gambaran tentang jumlah data dari setiap kategori sentimen yang diklasifikasikan dengan benar, serta data yang mengalami kesalahan klasifikasi, sehingga memfasilitasi analisis mendalam terhadap kinerja model secara keseluruhan.



Gambar 6. Confusion Matrix Hasil SVM

## Kesimpulan

Penelitian ini mengungkapkan bahwa penggunaan algoritma Support Vector Machine (SVM) yang dikombinasikan dengan teknik Term Frequency–Inverse Document Frequency (TF–IDF) serta serangkaian langkah pra-pemrosesan teks yang mendalam berhasil menciptakan model klasifikasi sentimen yang handal untuk menilai ulasan pengguna pada aplikasi Kredivo, yang ditulis dalam Bahasa Indonesia. Tahapan pra-pemrosesan seperti case folding, pembersihan data, tokenisasi, penghapusan kata-kata umum, dan stemming ternyata krusial dalam memperbaiki mutu data serta efisiensi model secara menyeluruh. Dari hasil percobaan, algoritma SVM dengan kernel Radial Basis Function (RBF) menonjol sebagai yang paling unggul, menghasilkan nilai precision, recall, dan F1-score yang harmonis. Temuan ini menegaskan bahwa integrasi TF–IDF dan SVM merupakan strategi yang ampuh untuk mengatasi kerumitan bahasa Indonesia, termasuk variasi kosakata dan pola kalimat yang tidak standar. Pada dasarnya, model yang dibangun dapat mendeteksi sentimen positif, negatif, dan netral dengan tingkat konsistensi yang tinggi. Lebih lanjut, hasil studi ini berpotensi besar untuk diimplementasikan dalam analisis opini pelanggan di bidang teknologi finansial (fintech), sekaligus menjadi fondasi bagi riset mendatang, seperti pengembangan Aspect-Based Sentiment Analysis (ABSA) atau penerapan pendekatan hibrid berbasis deep learning untuk meningkatkan ketepatan dan kemampuan interpretasi model.

## Daftar Pustaka

- [1] M. R. Amalsyah, D. Kurniawan, A. Rifai, and P. Sari, “Sentiment Analysis of Fintech Application User Reviews using the CRISP-DM Framework for Product Development Prioritization,” *SISTEMASI*, vol. 14, no. 2, p. 813, Mar. 2025, doi: 10.32520/stmsi.v14i2.5064.
- [2] H. Huang, A. Asemi, and M. B. Mustafa, “Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions,” *Ieee Access*, vol. 11, pp. 90367–90382, 2023, doi: 10.1109/access.2023.3307308.
- [3] M. Roman, A. Shahid, S. Khan, L. Yu, M. Asif, and Y. Y. Ghadi, “Investigating Maps of Science Using Contextual Proximity of Citations Based on Deep Contextualized Word Representation,” *Ieee Access*, vol. 10, pp. 31397–31419, 2022, doi: 10.1109/access.2022.3159980.
- [4] N. Mughal, G. Mujtaba, S. Shaikh, A. H. kumar, and S. M. Daudpota, “Comparative Analysis of Deep Natural Networks and Large Language Models for Aspect-Based Sentiment Analysis,” *Ieee Access*, vol. 12, pp. 60943–60959, 2024, doi: 10.1109/access.2024.3386969.
- [5] A. Erkan and T. Güngör, “Analysis of Deep Learning Model Combinations and Tokenization Approaches in Sentiment Classification,” *Ieee Access*, vol. 11, pp. 134951–134968, 2023, doi: 10.1109/access.2023.3337354.
- [6] N. Jacob and V. M. Viswanatham, “Sentiment Analysis Using Improved Atom Search Optimizer With a Simulated Annealing and ReLU Based Gated Recurrent Unit,” *Ieee Access*, vol. 12, pp. 38944–38956, 2024, doi: 10.1109/access.2024.3375119.
- [7] H. J. Alshahrani, A. Q. A. Hassan, N. S. Almalki, M. M. Alnfai, A. S. Salama, and M. A. Hamza, “Applied Linguistics With Red-Tailed Hawk Optimizer-Based Ensemble Learning Strategy in Natural Language Processing,” *Ieee Access*, vol. 11, pp. 132448–132456, 2023, doi: 10.1109/access.2023.3331061.
- [8] N. S. M. Nafis and S. Awang, “An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification,” *Ieee Access*, vol. 9, pp. 52177–52192, 2021, doi: 10.1109/access.2021.3069001.
- [9] J. Khan, A. Alam, and Y. Lee, “Intelligent Hybrid Feature Selection for Textual Sentiment Classification,” *Ieee Access*, vol. 9, pp. 140590–140608, 2021, doi: 10.1109/access.2021.3118982.
- [10] F. Feizian and B. Amiri, “Cryptocurrency Price Prediction Model Based on Sentiment Analysis and Social Influence,” *Ieee Access*, vol. 11, pp. 142177–142195, 2023, doi: 10.1109/access.2023.3342688.
- [11] P. Durga and D. Godavarthi, “Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-Rnn),” *Ieee Access*, vol. 11, pp. 108433–108447, 2023, doi: 10.1109/access.2023.3320738.
- [12] A. Chugh *et al.*, “Spider Monkey Crow Optimization Algorithm With Deep Learning for Sentiment Classification and Information Retrieval,” *Ieee Access*, vol. 9, pp. 24249–24262, 2021, doi: 10.1109/access.2021.3055507.
- [13] F. A. Ramadhan, R. Ridwan, P. Ruslan, and A. Zahra, “SENTIMENT ANALYSIS OF E-COMMERCE PRODUCT REVIEWS FOR CONTENT INTERACTION USING MACHINE LEARNING,” vol. 6, no. 1, p. 2023, 2023.

- [14] M. Fadillah and P. Nugroho, "Human-in-the-Loop Approaches for Sentiment Classification in Indonesian Contexts," *Computational Social Science Review*, vol. 11, no. 4, pp. 233–249, 2024, doi: 10.1016/cssr.2024.09.010.
- [15] G. Mumtaz *et al.*, "Classification and Prediction of Significant Cyber Incidents (SCI) Using Data Mining and Machine Learning (DM-ML)," *Ieee Access*, vol. 11, pp. 94486–94496, 2023, doi: 10.1109/access.2023.3249663.
- [16] O. M. Alyasiri, Y. Cheah, A. K. Abasi, and O. M. Al-Janabi, "Wrapper and Hybrid Feature Selection Methods Using Metaheuristic Algorithms for English Text Classification: A Systematic Review," *Ieee Access*, vol. 10, pp. 39833–39852, 2022, doi: 10.1109/access.2022.3165814.
- [17] A. Kurnia, H. Harlinda, and H. Darwis, "Analisis Sentimen Pengguna Gojek Berdasarkan Ulasan pada App Store dengan Metode KNN, Naive Bayes, dan SVM," *LINIER: Literatur Informatika dan Komputer*, vol. 2, no. 2, pp. 234–245, Jun. 2025, doi: 10.33096/linier.v2i2.3132.
- [18] M. Salman Al Markas, S. Anraeni, and L. Budiman Ilmuwan, "Implementasi Fitur Vector Bag Of Word Dan TF IDF untuk Analisis Sentiment," *LINIER: Literatur Informatika dan Komputer*, vol. 2, no. 2, pp. 136–146, Jun. 2025, doi: 10.33096/linier.v2i2.3104.
- [19] N. Rahmah, P. Purnawansyah, and F. Umar, "Metode Support Vector Machine Untuk Klasifikasi Data Penyakit Hati Yang Imbalance," *Buletin Sistem Informasi dan Teknologi Islam*, vol. 5, no. 1, pp. 55–64, Apr. 2024, doi: 10.33096/busiti.v5i1.2189.
- [20] P. Durga and D. Godavarthi, "Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-Rnn)," *Ieee Access*, vol. 11, pp. 108433–108447, 2023, doi: 10.1109/access.2023.3320738.