

## ANALISIS SENTIMEN APLIKASI HALO BCA DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES, SUPPORT VECTOR MACHINE DAN RANDOM FOREST

**Sebastianus Adi Santoso Mola<sup>1</sup>, Diandra Lyan Bethseba Baun<sup>2</sup>, Ingratcia Oktaviana Nunes<sup>3</sup>,  
 Michelle Maria Angely Rani Sani<sup>4</sup>**

Program Studi Ilmu Komputer, Universitas Nusa Cendana  
 Jl. Adisucipto Penfui, Kupang – Nusa Tenggara Timur, Indonesia  
 Email: [adimola@staf.undana.ac.id](mailto:adimola@staf.undana.ac.id), [2diandrabaun@gmail.com](mailto:2diandrabaun@gmail.com), [3annaanunes17@gmail.com](mailto:3annaanunes17@gmail.com),  
[4michellesani63@gmail.com](mailto:4michellesani63@gmail.com)

### ABSTRAK

Penelitian ini bertujuan mengevaluasi sentimen pengguna terhadap aplikasi Halo BCA berdasarkan ulasan di Google Play Store menggunakan tiga algoritma pembelajaran mesin, yakni Naive Bayes, Support Vector Machine (SVM), dan Random Forest. Data penelitian bersumber dari 6.464 komentar pengguna yang diambil melalui teknik web scraping. Data kemudian mengalami preprocessing meliputi pembersihan, penyeragaman huruf, normalisasi, tokenisasi, penghapusan kata umum, serta stemming. Proses penandaan sentimen dilakukan dengan metode berbasis leksikon, di mana ulasan dikelompokkan menurut jumlah kata positif dan negatif dalam teks. Ulasan diberi label "Positif" jika kata positif lebih banyak, "Negatif" jika kata negatif dominan, dan "Netral" jika keduanya seimbang. Hasil pelabelan menunjukkan dari total data, 2.636 ulasan (40,8%) tergolong positif, 2.425 ulasan (37,5%) negatif, dan 1.403 ulasan (21,7%) netral. Model Random Forest menunjukkan kinerja terbaik dengan akurasi 91,28%, presisi 0,91, recall 0,91, dan F1-score 0,91. Support Vector Machine memiliki akurasi 87,55%, sementara Naive Bayes 81,73%. Analisis wordcloud menunjukkan ulasan positif didominasi kata-kata yang menggambarkan kepuasan terhadap kualitas dan kemudahan aplikasi, sedangkan ulasan negatif berkaitan dengan masalah teknis dan kendala akses. Ulasan netral mencerminkan pengalaman biasa tanpa kesan khusus. Penelitian ini diharapkan memberikan kontribusi nyata bagi pengembangan aplikasi perbankan mobile, khususnya Halo BCA, dengan fokus perbaikan teknis dan aksesibilitas sebagai prioritas utama keluhan pengguna.

Kata kunci: Analisis sentimen, Halo BCA, *Naive Bayes*, *Support Vector Machine*, *Random Forest*

### ABSTRACT

*This study aims to evaluate user sentiment toward the Halo BCA application based on reviews from the Google Play Store using three machine learning algorithms: Naive Bayes, Support Vector Machine (SVM), and Random Forest. The research data consists of 6,464 user comments obtained through web scraping techniques. The data underwent preprocessing, including cleaning, case folding, normalization, tokenization, stopword removal, and stemming. Sentiment labeling was conducted using a lexicon-based method, where reviews were categorized based on the number of positive and negative words in the text. Reviews were labeled as "Positive" if positive words were predominant, "Negative" if negative words dominated, and "Neutral" if both were balanced. The labeling results show that of the total data, 2,636 reviews (40.8%) were positive, 2,425 reviews (37.5%) were negative, and 1,403 reviews (21.7%) were neutral. The Random Forest model demonstrated the best performance, with an accuracy of 91.28%, precision of 0.91, recall of 0.91, and an F1-score of 0.91. Support Vector Machine achieved an accuracy of 87.55%, while Naive Bayes had an accuracy of 81.73%. A word cloud analysis revealed that positive reviews were dominated by words expressing satisfaction with the application's quality and ease of use, while negative reviews were more related to technical issues and access problems. Neutral reviews reflected ordinary experiences without specific impressions. This study is expected to provide meaningful contributions to the development of mobile banking applications, particularly Halo BCA, focusing on technical improvements and accessibility as the main priorities for addressing user complaints.*

*Keywords: Sentiment analysis, Halo BCA, Naive Bayes, Support Vector Machine, Random Forest*



## 1. PENDAHULUAN

### Latar Belakang

Dalam era digital yang terus maju, industri perbankan mengalami perubahan signifikan dalam cara melayani kebutuhan pelanggan. Bank konvensional saat ini perlu beradaptasi dengan cepat agar dapat memenuhi harapan nasabah yang semakin cerdas teknologi, terutama seiring dengan pertumbuhan industri keuangan. Salah satu terobosan yang muncul adalah layanan perbankan mobile, yang menawarkan kemudahan, efisiensi, dan akses yang belum pernah ada sebelumnya. Beragam aplikasi telah dibuat untuk memudahkan nasabah dalam mengakses layanan dan produk perbankan. Salah satu contohnya adalah aplikasi Halo BCA, yang merupakan layanan digital untuk layanan pelanggan dari Bank Central Asia (BCA). Aplikasi ini memungkinkan pemilih untuk berinteraksi secara langsung dengan bank melalui platform digital.

Halo BCA adalah aplikasi mobile yang dirancang untuk mempermudah nasabah maupun non-nasabah BCA dalam mengakses layanan bantuan dan produk perbankan tanpa terbatas oleh waktu dan tempat. Aplikasi ini merupakan inovasi dari layanan call center BCA yang sebelumnya hanya bisa diakses melalui telepon. Dengan adanya aplikasi ini, pengguna dapat berkomunikasi langsung dengan customer service untuk menyelesaikan berbagai masalah perbankan atau mendapatkan informasi terkait produk BCA dengan cara yang cepat dan aman. Fitur-fitur seperti VoIP Call, integrasi media sosial, dan layanan self-service membuat Halo BCA menjadi solusi perbankan yang komprehensif dan efisien. Ketersediaan 24/7 menjadikan Halo BCA sebagai solusi yang praktis, memberikan kenyamanan dan kemudahan bagi penggunanya untuk mendapatkan bantuan dan informasi di era digital yang terus berkembang.

Analisis sentimen adalah teknik dalam pemrosesan bahasa alami untuk mengklasifikasikan teks berdasarkan polaritas sentimennya [1]. Dalam hal aplikasi Halo BCA, analisis sentimen berfungsi untuk memahami secara mendalam bagaimana pengguna menilai pengalaman mereka saat menggunakan aplikasi tersebut, termasuk aspek kemudahan akses, kualitas layanan, dan performa teknis. Data ulasan dari pengguna di Google Play Store dapat digunakan untuk mengidentifikasi sentimen positif, negatif, atau netral. Melalui analisis ini, pengembang aplikasi dapat lebih mudah mengidentifikasi keunggulan dan kekurangan Halo BCA secara terstruktur. Di tengah persaingan digital yang semakin ketat, wawasan mendalam tentang sentimen pengguna dapat menjadi elemen penting dalam mempertahankan loyalitas nasabah dan memperkuat posisi BCA dalam industri perbankan digital.

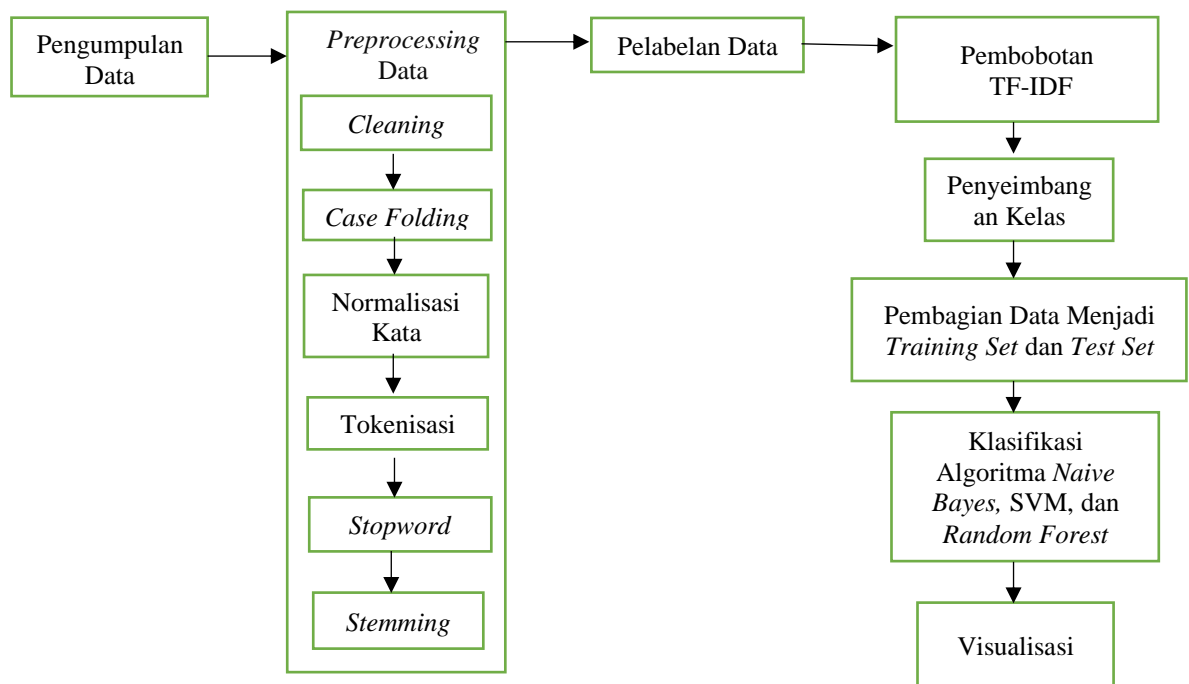
Sejumlah studi telah dilaksanakan mengenai analisis sentimen dengan memanfaatkan metode Naive Bayes, SVM, dan Random Forest. Studi [2], [3], [4] menunjukkan bahwa algoritma Random Forest terbukti paling efektif dalam mengklasifikasikan sentimen pengguna dari aplikasi Ruangguru [2], aplikasi Dana [3], dan aplikasi ShopeePay [4]. Namun, ada pula penelitian yang menunjukkan bahwa metode SVM memiliki kinerja yang lebih baik, seperti yang terdapat dalam [5], [6], [7]. Dalam penelitian [5], algoritma SVM menunjukkan hasil optimal dalam klasifikasi cyberbullying, dengan tingkat akurasi 83% dan waktu proses 4 menit 52 detik. Studi [6] bertujuan untuk menganalisis sentimen tweet mengenai PPKM di wilayah Jawa dan Bali. Temuan menunjukkan bahwa SVM efektif dalam mengklasifikasikan sentimen dan mendukung bahwa PPKM berhasil menurunkan penyebaran Covid-19. Penelitian yang dilakukan selanjutnya [7] menyimpulkan bahwa penerapan Algoritma SVM menunjukkan Pegipegi sebagai aplikasi pemesanan tiket online terbaik dengan tingkat akurasi tertinggi 78,21%, diikuti oleh Agoda (77,00%), Traveloka (75,03%), Mister Aladin (64,00%), dan Tiket.com (58,68%). Selanjutnya ada berbagai penelitian yang menunjukkan bahwa metode Naive Bayes memiliki kinerja yang baik seperti yang tercantum dalam [8], [9]. Studi [8] menyimpulkan bahwa sistem analisis sentimen untuk layanan ojek online di Twitter yang menggunakan Naive Bayes mencapai akurasi 80% pada 800 tweet. Akurasi bisa ditingkatkan dengan menambahkan data pelatihan. Dalam penelitian [9] yang dilakukan untuk menganalisis sentimen terhadap calon presiden 2024 menunjukkan bahwa sentimen positif mendominasi di semua dataset. Melalui Naive Bayes, tingkat akurasi tertinggi dicapai pada dataset Ganjar (73,68%), sedangkan skor sentimen tertinggi ditemukan pada dataset Ridwan (4,878).

Penelitian-penelitian sebelumnya telah menunjukkan bahwa ketiga metode tersebut memberikan hasil yang efektif dalam analisis sentimen. Berdasarkan hal tersebut, Penelitian ini akan mengevaluasi kinerja tiga algoritma, yaitu Naive Bayes, SVM, dan Random Forest, dalam konteks analisis sentimen terhadap ulasan aplikasi Halo BCA.

Penelitian ini menunjukkan bahwa penerapan tiga metode, yaitu Naive Bayes, SVM, dan Random Forest, sangat penting untuk memahami performa model dalam mengklasifikasikan sentimen pengguna secara menyeluruh. Setiap metode memiliki kelebihan masing-masing dalam pengolahan data teks, tergantung pada karakteristik data yang digunakan. Algoritma Random Forest terbukti efektif dalam menangani dataset yang kompleks dan memiliki ketidakseimbangan kelas yang signifikan, sementara SVM sangat baik untuk data berdimensi tinggi dengan fitur yang spesifik. Sebaliknya, Naive Bayes menawarkan kecepatan pemrosesan yang tinggi dan hasil yang baik untuk dataset sederhana atau dengan distribusi data yang seimbang. Menggunakan ketiga metode ini secara bersamaan memberikan analisis yang lebih mendalam terhadap sentimen pengguna aplikasi Halo BCA. Perbandingan hasil dari algoritma-algoritma tersebut membantu pengembangan memilih pendekatan terbaik berdasarkan kebutuhan, seperti akurasi tinggi, waktu pemrosesan yang cepat, atau kemudahan implementasi. Selain itu, analisis sentimen dengan berbagai metode memungkinkan pemahaman yang lebih luas terhadap preferensi pengguna, sehingga mendukung keputusan untuk meningkatkan pengalaman pengguna aplikasi Halo BCA.

## 2. METODE PENELITIAN

Penelitian ini dimulai dengan pengumpulan data ulasan pengguna dari Google Play Store menggunakan teknik *web scraping*. Data kemudian diproses melalui tahap *preprocessing* yang meliputi pembersihan, penyeragaman huruf, normalisasi, tokenisasi, penghapusan kata umum, dan stemming untuk memastikan kualitas data. Setelah itu, pelabelan sentimen dilakukan menggunakan metode berbasis leksikon, di mana ulasan dikelompokkan menjadi positif, negatif, atau netral. Data yang telah dilabeli diberi bobot menggunakan metode TF-IDF dan diseimbangkan untuk menghindari bias kelas. Selanjutnya, data dibagi menjadi *training set* dan *test set* untuk diuji menggunakan algoritma *Naive Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest*. Hasil analisis kemudian divisualisasikan, memberikan gambaran mengenai sentimen pengguna dan faktor utama yang memengaruhi ulasan mereka.



Gambar 1. Tahap Penelitian

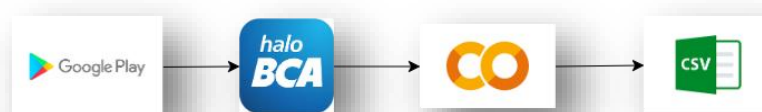
### Sumber dan Jenis Data

Data primer diperoleh melalui teknik *scraping* dari ulasan aplikasi Halo BCA di *Google Play Store*. Ulasan ini digunakan untuk mengevaluasi tingkat kepuasan pengguna terhadap aplikasi tersebut. Data hasil *scraping* akan memberikan gambaran mengenai persepsi pengguna terhadap kualitas layanan, kemudahan penggunaan, serta kinerja aplikasi Halo BCA. Analisis data primer ini diharapkan dapat mengungkap kelebihan dan kekurangan aplikasi, yang nantinya dapat menjadi dasar untuk meningkatkan kualitas layanan serta menjaga kepuasan dan loyalitas nasabah.

### Langkah – Langkah Penelitian

#### a. Pengumpulan Data

Penelitian ini dilakukan dengan mengumpulkan informasi melalui teknik *web scraping* pada tinjauan aplikasi Halo BCA di *Google Play Store*. Proses ini meliputi pengumpulan informasi secara langsung dari halaman ulasan pengguna, seperti tanggapan, penilaian, dan tanggal ulasan.



Gambar 2. Proses Scraping Data

Gambar tersebut menunjukkan alur proses pengolahan data yang dimulai dari aplikasi yang diunduh melalui Google Play. Setelah itu, aplikasi Halo BCA digunakan, yang kemungkinan berfungsi untuk layanan perbankan atau transaksi tertentu. Data yang dihasilkan kemudian diproses lebih lanjut oleh sistem atau aplikasi lain yang dilambangkan dengan logo Collab, sebelum akhirnya diekspor ke dalam format CSV. Format CSV ini biasanya digunakan untuk menyimpan data dalam bentuk tabel sederhana yang dapat diolah lebih lanjut, seperti untuk keperluan pelaporan atau analisis data.

b. *Preprocessing Data*

*Preprocessing* merupakan tahap yang bertujuan untuk membersihkan serta memperbaiki data. Informasi yang dikumpulkan sering kali berupa data tidak terstruktur dan mengandung karakter yang tidak penting. Tujuan dari *preprocessing* ialah membersihkan gangguan atau noise yang ada [10]. Langkah-langkah dalam *preprocessing* data meliputi *Cleaning*, *Case Folding*, Normalisasi kata, Tokenisasi, Penghapusan *Stopword*, dan *Stemming*. *Cleaning* merupakan langkah untuk menghapus atau memperbaiki data yang tidak tepat, tidak lengkap, atau tidak relevan sehingga data menjadi lebih konsisten dan siap untuk diproses. *Case folding* mengubah setiap huruf dalam teks menjadi format yang seragam, biasanya huruf kecil (*lowercase*). Normalisasi kata dimaksudkan untuk meratakan skala nilai dalam dataset, menekan perbedaan signifikan antar fitur, sehingga setiap fitur memberikan kontribusi yang setara dalam analisis dan pemodelan. Tokenisasi merupakan tindakan memisahkan teks menjadi kata-kata terpisah, yang memudahkan dalam menghitung frekuensi kemunculan kata dalam data. Penghapusan *stopword* menghilangkan kata-kata yang tidak berarti signifikan dalam analisis, mempermudah pengolahan teks. *Stemming* merupakan proses penghilangan akhiran dan prefiks pada kata untuk memperoleh bentuk dasar yang lebih sederhana, yang biasanya mencerminkan makna utama dari kata itu.

c. Pelabelan Data

Tahap selanjutnya dalam analisis sentimen adalah pelabelan, yang bertujuan untuk mengklasifikasikan setiap ulasan di Google Play Store berdasarkan sentimen yang terkandung di dalamnya. Proses pelabelan ini dilakukan dengan dua pendekatan utama: berdasarkan rating pengguna dan menggunakan metode berbasis leksikon (*lexicon-based*). Metode berbasis leksikon menganalisis teks ulasan menggunakan dua kumpulan leksikon, yaitu positif dan negatif. Kumpulan leksikon ini diambil dari file eksternal (*positive.tsv* dan *negative.tsv*) dan diubah menjadi struktur data set, yang mempercepat dan mempermudah proses pencocokan kata dalam teks. Pada proses pelabelan ini, jumlah kata positif dan negatif dalam setiap ulasan dihitung. Jika kata positif lebih dominan, ulasan diberi label "Positif". Jika kata negatif lebih banyak, ulasan dikategorikan sebagai "Negatif". Jika keduanya seimbang, ulasan diberi label "Netral". Proses pelabelan ini bertujuan untuk membantu model memahami sentimen yang diungkapkan pengguna dalam ulasan mereka. Dengan adanya kategori yang jelas, analisis sentimen dapat dilakukan dengan lebih terarah dan menghasilkan wawasan yang lebih mendalam tentang pengalaman pengguna terhadap aplikasi tersebut.

d. Pembobotan TF-IDF

Pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan metode yang diterapkan untuk menilai seberapa sering suatu kata muncul dalam sebuah dokumen atau artikel. Tujuan dari perhitungan ini adalah untuk mengevaluasi seberapa signifikan kata tersebut dalam konteks dokumen yang sedang ditelaah [11] TF-IDF merupakan nilai yang diperoleh dari penggabungan dua komponen, yaitu TF (*Frekuensi Istilah*) dan IDF (*Frekuensi Dokumen Invers*). TF (*Frekuensi Term*) mengevaluasi seberapa sering suatu istilah muncul dalam dokumen tertentu, sedangkan IDF (*Frekuensi Dokumen Invers*) menilai seberapa signifikan istilah tersebut dalam seluruh koleksi dokumen [12].

e. Penyeimbangan kelas dengan *Synthetic Minority Oversampling Technique (SMOTE)*

SMOTE dirancang khusus untuk menangani dataset yang tidak seimbang, di mana jumlah data pada satu kelas jauh lebih besar dibandingkan dengan kelas lainnya. Ketidakseimbangan ini dapat membuat model cenderung lebih fokus pada kelas mayoritas, sehingga mengurangi kemampuannya dalam memprediksi kelas yang lebih sedikit [13].

f. Pembagian Data Menjadi *Training set* dan *Test set*

Pembagian data menjadi set pelatihan dan set pengujian adalah langkah penting dalam penelitian *machine learning* untuk mengevaluasi kinerja model pada data yang belum dikenal sebelumnya. Data disiapkan terlebih dahulu, setelah itu proporsi pembagiannya ditentukan. Pembagian dilakukan secara acak untuk menjamin bahwa data yang digunakan bersifat representatif. Set pelatihan digunakan untuk mengajarkan model, sedangkan set pengujian disimpan untuk mengevaluasi kinerja model setelah proses pelatihan. Langkah ini bertujuan untuk memastikan bahwa model tidak hanya terlampau sesuai dengan data latih, tetapi juga dapat berfungsi dengan baik pada data yang belum diperkenalkan sebelumnya.

g. Klasifikasi dengan Algoritma *Naive Bayes*, *Support Vector Machine*, dan *Random Forest Naive Bayes*

<p>Input: Dataset D dengan fitur <math>X_1, X_2, \dots, X_n</math> dan label Y  Output: Model untuk prediksi kelas</p> <p>1. Hitung probabilitas prior untuk setiap kelas (<math>P(Y)</math>):  <math>P(Y=c) = \text{Jumlah sampel dengan } Y=c / \text{Total sampel}</math></p> <p>2. Hitung probabilitas likelihood untuk setiap fitur (<math>P(X Y)</math>):  Untuk setiap fitur <math>X_i</math> dan kelas <math>Y=c</math>:  <math>P(X_i Y=c) = \text{Frekuensi } X_i \text{ diberikan } Y=c / \text{Total sampel dengan } Y=c</math></p> <p>3. Untuk setiap sampel baru dengan fitur <math>X_1', X_2', \dots, X_n'</math>:  a. Hitung posterior untuk setiap kelas:  <math>P(Y=c X) = P(Y=c) * P(X_1 Y=c) * P(X_2 Y=c) * \dots * P(X_n Y=c)</math>  b. Prediksi kelas dengan nilai posterior terbesar.</p> <p>Return: Kelas dengan nilai posterior terbesar</p>
---

Gambar 3. Pseudo Code Algoritma Naive Bayes

Algoritma Naive Bayes adalah metode pembelajaran yang menggunakan statistik dan probabilitas untuk memprediksi kemungkinan suatu peristiwa di masa depan, berdasarkan data atau pengalaman yang telah dikumpulkan sebelumnya [14]. Teknik ini berfokus pada perkiraan probabilitas dan memiliki keunggulan dalam menghasilkan tingkat kesalahan yang lebih rendah, terutama saat digunakan pada dataset berukuran besar.

*Support Vector Machine*

<p>Input: Dataset D dengan fitur X dan label Y  Output: Model SVM dengan hyperplane terbaik</p> <p>1. Inisialisasi parameter model (misalnya, bobot w dan bias b).</p> <p>2. Optimasi:  a. Cari hyperplane yang memaksimalkan margin:  - Solve: Minimize <math>\ w\ ^2</math> (panjang vektor w)  - Dengan constraint: <math>Y * (w \cdot X + b) \geq 1</math> untuk semua sampel</p> <p>3. Jika data tidak terpisah secara linear:  a. Terapkan kernel trick untuk memetakan data ke ruang berdimensi lebih tinggi.  - Contoh kernel: Linear, Polynomial, Gaussian (RBF).</p> <p>4. Untuk prediksi sampel baru <math>X'</math>:  a. Hitung nilai decision function:  <math>f(X') = w \cdot X' + b</math>  b. Jika <math>f(X') &gt; 0</math>, prediksi kelas positif; jika tidak, prediksi kelas negatif.</p> <p>Return: Model dengan hyperplane terbaik</p>
--

Gambar 4. Pseudo Code Algoritma Support Vector Machine

Support Vector Machine (SVM) adalah metode pembelajaran mesin yang bertujuan menemukan hyperplane terbaik untuk memisahkan dua kategori. SVM bekerja dengan memisahkan elemen-elemen dari dua kategori, yakni kategori 1 dan kategori -1 [15]. *Hyperplane* adalah garis atau permukaan (tergantung pada dimensi data) yang digunakan untuk memisahkan kelas-kelas dengan jarak margin terbesar.

*Random Forest*

<p>Input: Dataset D, jumlah pohon N, jumlah fitur m yang dipilih secara acak  Output: Model Random Forest</p>
---

```

1. Untuk i = 1 hingga N:
  a. Ambil sampel bootstrap dari dataset D.
  b. Bangun decision tree:
     - Pada setiap node, pilih m fitur secara acak.
     - Pilih fitur terbaik berdasarkan metrik tertentu (misalnya, Gini Index atau Entropy).
     - Bagi node menggunakan fitur terbaik hingga mencapai kondisi berhenti (misalnya, jumlah sampel minimum di node atau kedalaman maksimum).
  c. Simpan decision tree.

2. Untuk prediksi sampel baru X:
  a. Kumpulkan prediksi dari semua pohon.
  b. Gunakan voting mayoritas (untuk klasifikasi) atau rata-rata (untuk regresi).

Return: Prediksi akhir

```

Gambar 5. Pseudo Code Algoritma Random Forest

*Random Forest* merupakan algoritma *ensemble learning* yang menghasilkan keputusan akhir melalui voting mayoritas. Prinsip kerjanya adalah menggabungkan beberapa *Decision Tree* untuk proses klasifikasi. Setiap pohon keputusan menerima sampel data acak dari *training set* menggunakan metode *bagging* (Bootstrap Aggregating), di mana data dipilih secara acak dalam jumlah tertentu. Dari sampel tersebut, sejumlah *Decision Tree* dibangun, dan hasil dari setiap pohon disimpan sebagai model yang digunakan saat pengujian [16].

#### h. Pemodelan dan Evaluasi

Evaluasi model merupakan proses untuk mengukur sejauh mana model memenuhi tujuan penelitian. Salah satu metode untuk mengevaluasi kinerja model klasifikasi adalah dengan memanfaatkan *Confusion Matrix*. *Confusion Matrix* merupakan metrik evaluasi yang digunakan untuk mengukur kinerja model klasifikasi. Matriks ini memberikan representasi visual dari kinerja model klasifikasi, dengan memetakan prediksi yang benar dan salah untuk setiap kelas. Tabel *Confusion Matrix* dapat ditemukan pada Tabel 1.

Tabel 1. *Confusion Matrix*

<i>Actual</i>	<i>Predicted</i>		
	<i>Prediksi Positif</i>	<i>Prediksi Netral</i>	<i>Prediksi Negatif</i>
<i>Positive</i>	True positive (TP)	False Negative (FN)	False Negative (FN)
<i>Netral</i>	False Positive (FP)	True Neutral (TN)	False Negative (FN)
<i>Negative</i>	False positive (FP)	False Positive (FP)	True Negative (TN)

Tabel matriks ini memberikan gambaran rinci tentang akurasi model dalam mengklasifikasikan sentimen pengguna ke dalam kategori positif, netral, atau negatif. Baris "*Actual*" merepresentasikan label asli data, sedangkan kolom "*Predicted*" menunjukkan hasil prediksi dari model. Matriks ini menggambarkan hubungan antara data asli dan prediksi, seperti *True Positive* (TP), di mana data sebenarnya positif dan diprediksi positif, serta *True Neutral* (TN) dan *True Negative* (TN), yang menunjukkan prediksi model sesuai dengan label netral dan negatif. Selain itu, matriks ini juga mencatat kesalahan prediksi, seperti *False Positive* (FP), ketika data sebenarnya bukan kelas tertentu tetapi diprediksi sebagai kelas tersebut, dan *False Negative* (FN), ketika model gagal memprediksi kelas asli. Informasi yang terdapat dalam matriks ini merupakan dasar perhitungan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*, yang digunakan untuk mengukur kinerja model secara komprehensif. Berdasarkan *confusion matrix* dapat dihitung adalah akurasi, presisi, *recall*, *F1-score*, dengan persamaan (1) sampai (4).

$$\text{Akurasi} = \frac{TP_{\text{positif}} + TP_{\text{netral}} + TP_{\text{negatif}}}{TP_{\text{positif}} + TP_{\text{netral}} + TP_{\text{negatif}} + FP_{\text{positif}} + FP_{\text{netral}} + FP_{\text{negatif}} + FN_{\text{positif}} + FN_{\text{netral}} + FN_{\text{negatif}}} \quad (1)$$

$$\text{Presisi}_{\text{Kelas}} = \frac{TP_{\text{Kelas}}}{TP_{\text{Kelas}} + FP_{\text{Kelas}}} \quad (2)$$

$$\text{Recall}_{\text{Kelas}} = \frac{TP_{\text{Kelas}}}{TP_{\text{Kelas}} + FN_{\text{Kelas}}} \quad (3)$$

$$F1_{\text{Kelas}} = 2 \times \frac{\text{Precision}_{\text{Kelas}} \times \text{Recall}_{\text{Kelas}}}{\text{Precision}_{\text{Kelas}} + \text{Recall}_{\text{Kelas}}} \quad (4)$$

#### i. Visualisasi

Visualisasi *word cloud* merupakan representasi grafis dari frekuensi kata dalam suatu teks, di mana ukuran font kata mencerminkan frekuensi kemunculannya. Visualisasi *word cloud* memungkinkan kita mengidentifikasi kata-kata kunci yang paling sering muncul dalam masing-masing kategori sentimen (positif, netral, negatif).

## Analisis Sentimen

Analisis sentimen adalah proses otomatis untuk mengidentifikasi dan mengekstrak opini atau emosi yang terkandung dalam teks [17]. Dalam penelitian ini digunakan metode Naive Baayes, SVM dan Random Forest.

### 3. HASIL DAN PEMBAHASAN

Ulasan tentang aplikasi Halo BCA dikelompokkan menjadi tiga kategori utama: positif, netral, dan negatif. Proses scraping yang dilakukan di halaman Google Play Store menghasilkan total 6.464 ulasan, yang kemudian diberi label sesuai dengan kategori yang telah ditentukan. Tabel Kelas Ulasan memberikan gambaran tentang distribusi sentimen dalam ulasan pengguna, yang terdiri dari kategori Positif, Negatif, dan Netral. Dari total ulasan tersebut, mayoritas—sebanyak 2.745—termasuk dalam kategori positif, menandakan bahwa sebagian besar pengguna memberikan tanggapan yang baik terhadap aplikasi. Ulasan dengan sentimen negatif mencapai 2.416, yang menunjukkan bahwa ada sejumlah pengguna yang merasa tidak puas atau memberikan kritik terhadap aplikasi. Sementara itu, ulasan netral hanya berjumlah 1303, mencerminkan bahwa sedikit pengguna yang memberikan tanggapan yang tidak condong ke arah positif atau negatif. Informasi ini sangat berguna untuk memahami pola sentimen yang ada dan juga dapat menjadi pertimbangan dalam mengatasi ketidakseimbangan data (*class imbalance*) selama proses analisis sentimen.

Tabel 2. Tabel Kelas Ulasan

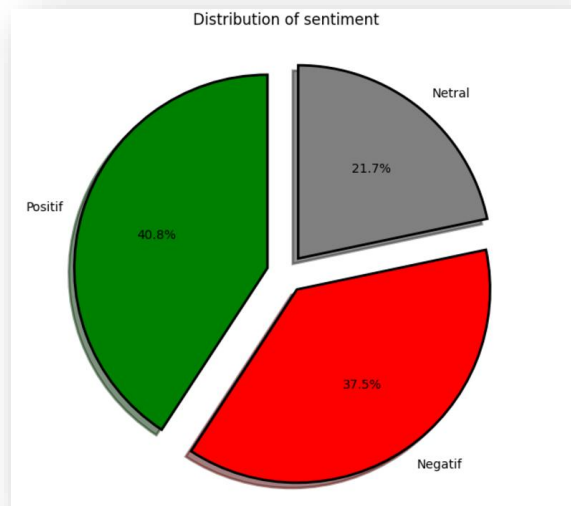
Kelas Ulasan	Jumlah
Positif	2745
Negatif	2416
Netral	1303
Total	6464

Tabel Hasil Preprocessing dan Pelabelan menggambarkan langkah-langkah pengolahan data ulasan pengguna sebelum dianalisis lebih lanjut. Kolom Ulasan Awal berisi teks asli dari pengguna yang masih mengandung elemen-elemen seperti tanda baca, kata-kata yang tidak relevan, dan struktur kalimat yang rumit. Setelah melalui proses *preprocessing*, teks tersebut disederhanakan, seperti yang terlihat pada kolom Ulasan Bersih, di mana tanda baca dan kata-kata berlebihan dihapus agar teks menjadi lebih jelas dan siap untuk dianalisis. Kolom Label menunjukkan sentimen dari setiap ulasan, apakah itu Positif (untuk ulasan yang mengungkapkan kepuasan) atau Negatif (untuk ulasan yang berisi kritik atau ketidakpuasan). Proses ini sangat penting agar data yang digunakan dalam analisis sentimen terstruktur dengan baik, memudahkan pemrosesan lebih lanjut oleh model, dan menghasilkan analisis yang lebih akurat.

Tabel 3 menyajikan hasil akhir dari tahap preprocessing data. Data yang telah dibersihkan dan diberi label sentimen ditampilkan secara rinci.

Tabel 3. Hasil *Preprocessing* dan Pelabelan

Ulasan awal	Ulasan bersih	Label
Dengan menggunakan aplikasi halo bca ini menjadi sangat gampang dan nyaman menggunakan mobile banking bca sangat terbantu dan cs nya juga dengan bgitu cepat dalam menangani masalah pemblokiran dan memberikan solusi yang tepat thanks a lot	aplikasi halo gampang nyaman mobile banking bantu cepat tangan blokir solusi thanks	Positif
Selalu gagal masuk aplikasi error. Mau buat laporan tarik tunai uangnya tidak keluar tp saldo sudah motong harus kemana kalo aplikasinya error?	gagal masuk aplikasi error lapor tarik tunai uang saldo motong mana aplikasi error	Negatif



Gambar 6. Perbandingan Pelabelan Ulasan

Setelah tahap preprocessing data, diperoleh ulasan mengenai aplikasi Halo BCA, dengan rincian 40,8% atau 2.636 ulasan yang memiliki sentimen positif, 37,5% atau 2.425 ulasan yang memiliki sentimen negatif, dan 21,7% atau 1403 ulasan yang memiliki sentimen netral, seperti yang ditampilkan pada Gambar 3.

Tahap berikutnya adalah menerapkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk memberikan bobot pada setiap istilah. Tode ini mengukur tingkat kepentingan sebuah kata dalam sebuah dokumen dengan mempertimbangkan frekuensi kemunculan kata tersebut dan frekuensi invers dokumen (IDF). Dengan cara ini, TF-IDF memberikan bobot lebih besar pada kata-kata yang lebih spesifik, sehingga membantu model untuk fokus pada istilah yang relevan dalam membedakan dokumen satu dengan yang lainnya.

Sebelum menerapkan SMOTE, data menunjukkan ketidakseimbangan kelas yang signifikan: kelas Positif memiliki 2.636 sampel, kelas Negatif 2.425 sampel, dan kelas Netral hanya 1.403 sampel. Ketidakseimbangan ini menyebabkan model lebih cenderung memprediksi kelas dengan jumlah sampel lebih banyak (Positif), sehingga kinerja prediksi untuk kelas minoritas (Negatif dan Netral) menjadi rendah. Setelah menggunakan SMOTE, jumlah sampel di setiap kelas menjadi seimbang, dengan 2.636 sampel per kelas. SMOTE berhasil mencapainya dengan membuat sampel sintetik baru untuk kelas minoritas (Negatif dan Netral). Sebagai hasilnya, model dapat mempelajari setiap kelas dengan lebih baik tanpa adanya bias, yang diharapkan akan meningkatkan akurasi dan kestabilan prediksi untuk seluruh kelas.

Tabel 4. Hasil Penyeimbangan Kelas dengan SMOTE

Label	Sebelum SMOTE	Sesudah SMOTE
Positif	2636	2636
Negatif	2425	2636
Netral	1403	2636

Dataset ini dipartisi menjadi dua set, yaitu set pelatihan (80%) dan set pengujian (20%). Pembagian ini adalah praktik umum dalam pemodelan untuk memastikan bahwa model mendapatkan cukup data untuk dilatih sebelum diuji. Sebanyak 5.050 data digunakan sebagai data pelatihan untuk melatih model, sedangkan 1.263 data sisanya berfungsi sebagai data pengujian untuk mengevaluasi kinerja model. Pembagian data ini bertujuan untuk mencegah overfitting dan meningkatkan kemampuan generalisasi model pada data yang belum pernah ditemui sebelumnya.

### Hasil Klasifikasi

Berdasarkan Tabel 5, Metode Naïve Bayes berhasil mengklasifikasikan dengan benar 348 ulasan positif, 519 ulasan netral, dan 426 ulasan negatif. Berdasarkan confusion matrix yang ditampilkan, dapat diperoleh nilai-nilai evaluasi model. Proses klasifikasi yang menerapkan metode Naïve Bayes menghasilkan matriks evaluasi dengan tingkat akurasi 81,73%, presisi 0,83, *recall* 0,82, dan *F1-score* 0,82.

Tabel 5. *Confusion Matrix* Metode *Naive Bayes*

Aktual	Prediksi Negatif	Prediksi Netral	Prediksi Positif
Negatif	426	75	28
Netral	7	519	22
Positif	58	99	348



Berdasarkan Tabel 6, Metode Support Vector Machine (SVM) berhasil mengklasifikasikan dengan benar 415 ulasan positif, 521 ulasan negatif, dan 449 ulasan netral. Berdasarkan confusion matrix yang ditampilkan, diperoleh nilai-nilai evaluasi model. Proses klasifikasi dengan metode SVM menghasilkan matriks evaluasi yang menunjukkan akurasi sebesar 87,55%, presisi 0,88, *recall* 0,88, dan *F1-score* 0,88.

Tabel 6. Confusion Matrix Metode Support Vector Machine

Aktual	Prediksi Negatif	Prediksi Netral	Prediksi Positif
Negatif	449	58	22
Netral	8	521	19
Positif	33	57	415

Berdasarkan Tabel 7, metode *Random Forest* berhasil mengklasifikasikan dengan benar 433 ulasan positif, 490 ulasan negatif, dan 521 ulasan netral. Dari confusion matrix yang ditampilkan, dapat dilihat nilai-nilai matriks evaluasi. Proses klasifikasi dengan algoritma *Random Forest* menghasilkan matriks evaluasi yang menunjukkan tingkat akurasi 91,28%, presisi 0,91, *recall* 0,91, dan *F1-score* 0,91.

Tabel 7. Confusion Matrix Metode Random Forest

Aktual	Prediksi Negatif	Prediksi Netral	Prediksi Positif
Negatif	490	4	35
Netral	10	521	17
Positif	42	30	433

Tabel 8 menyajikan perbandingan matriks evaluasi dari masing-masing metode. Metode *Naïve Bayes* menunjukkan tingkat presisi 0,83, yang lebih rendah jika dibandingkan dengan metode *Support Vector Machine* dan *Random Forest*, yang masing-masing memiliki presisi 0,88 dan 0,91. Berdasarkan penilaian lain, metode *Random Forest* menunjukkan hasil terbaik. Akurasi, *recall*, dan *F1-score* pada metode *Random Forest* berhasil mencapai 0,91, yang lebih tinggi dibanding hasil yang didapat dari *Naïve Bayes* dan SVM. Hal ini menunjukkan bahwa metode *Random Forest* memberikan performa terbaik di antara ketiga metode yang diuji.

Tabel 8. Perbandingan Metode *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest*

Matrik Evaluasi	Naive Bayes	Support Vector Machine	Random Forest
Akurasi	0.81	0.87	0.91
Presisi	0.83	0.88	0.91
Recall	0.82	0.88	0.91
F1-score	0.82	0.88	0.91

Frekuensi kemunculan kata dalam *review* aplikasi Halo BCA divisualisasikan dalam bentuk *wordcloud*. Ukuran kata dalam *wordcloud* berfungsi sebagai indikator visual dari frekuensi kemunculannya. Kata yang paling sering muncul akan memiliki ukuran font yang paling besar, sehingga mudah diidentifikasi.



Gambar 7. Wordcloud Sentimen Positif

Opini positif pengguna terhadap aplikasi Halo BCA mencakup kepuasan pengguna karena kualitas aplikasi yang baik, kemudahan penggunaan, serta dukungan aplikasi dalam memfasilitasi transaksi. Gambar 7 menunjukkan wordcloud yang mewakili sentimen positif dari pengguna aplikasi Halo BCA.



Gambar 8. Wordcloud Sentimen Negatif

Opini negatif pengguna terhadap aplikasi Halo BCA mencakup ketidakpuasan pengguna terkait masalah teknis, kesulitan dalam mengakses layanan, dan kendala yang dihadapi saat bertransaksi. Gambar 8 menampilkan visualisasi sentimen negatif pengguna terhadap aplikasi Halo BCA dalam bentuk *wordcloud*.



Gambar 9. Wordcloud Sentimen Netral

Opini netral pengguna terhadap aplikasi Halo BCA mencakup pandangan yang tidak terlalu positif maupun negatif, seperti pengalaman yang biasa saja dalam penggunaan aplikasi, fitur yang cukup membantu tetapi masih perlu peningkatan, serta respon yang sesuai harapan tanpa kesan yang istimewa. Gambar 9 menampilkan visualisasi sentimen netral pengguna aplikasi Halo BCA dalam bentuk *wordcloud*.

#### 4. SIMPULAN

Evaluasi sentimen terhadap aplikasi Halo BCA menggunakan algoritma *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest* menghasilkan beberapa temuan signifikan. Dari 6.464 ulasan yang dianalisis, sebagian besar ulasan bersifat positif dengan persentase 40,8% (2.636 ulasan), diikuti ulasan negatif sebanyak 37,5% (2.425 ulasan), dan sisanya 21,7% (1.403 ulasan) bersifat netral. Dalam hal performa metode machine learning, *Random Forest* menunjukkan hasil terbaik dengan akurasi 91,28%, presisi 0,91, *recall* 0,91, dan *F1-score* 0,91, diikuti oleh *Support Vector Machine* dengan akurasi 87,55%, sementara *Naïve Bayes* memperoleh akurasi 81,73%. Analisis *wordcloud* mengungkap bahwa sentimen positif didominasi oleh kepuasan pengguna terhadap kualitas dan kemudahan penggunaan aplikasi, sedangkan sentimen negatif lebih banyak berfokus pada masalah teknis dan kesulitan akses. Ulasan netral cenderung mencerminkan pengalaman yang biasa tanpa kesan tertentu. Hasil analisis sentimen ini dapat dijadikan sebagai dasar untuk melakukan optimasi pada aplikasi Halo BCA, terutama pada aspek-aspek yang sering menjadi sumber ketidakpuasan pengguna.

Penelitian selanjutnya dapat mengkaji penggunaan metode analisis sentimen yang lebih canggih, seperti *deep learning*, untuk meningkatkan akurasi dalam mengklasifikasikan sentimen pengguna terhadap aplikasi Halo BCA. Analisis sentimen secara berkala terhadap ulasan pengguna memungkinkan kita untuk mengidentifikasi tren perubahan persepsi pengguna terhadap aplikasi, terutama setelah adanya pembaruan fitur.

#### DAFTAR PUSTAKA

- [1] I. D. Onantya and P. P. Adikara, "Analisis Sentimen Pada Ulasan Aplikasi BCA Mobile Menggunakan BM25 dan Improved K-Nearest Neighbor," *J-Ptiik.Ub.Ac.Id*, vol. 3, no. 3, pp. 2575–2580, 2019.
- [2] E. Fitri, "Sentiment Analysis of the Ruangguru Application Using Naive Bayes, Random Forest and Support Vector Machine Algorithms," *J. Transform.*, vol. 18, no. 1, p. 71, 2020.
- [3] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 9, pp. 4305–4313, 2022.
- [4] T. Fadiyah Basar, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen Pengguna Twitter terhadap Pembayaran Cashless menggunakan ShopeePay dengan Algoritma Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 3, pp. 1426–1433, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [5] M. F. Naufal, T. Arifin, and H. Wirjawan, "Analisis Perbandingan Tingkat Performa Algoritma SVM, Random Forest, dan Naïve Bayes untuk Klasifikasi Cyberbullying pada Media Sosial," *J. Ris. Sist. Inf. dan Tek. Inform.*, vol. 8, p. 82, 2023, [Online]. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>
- [6] R. Wati and S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," *J. Tek. Inform. UNIKA St. Thomas*, vol. 06, pp. 240–247, 2021, doi: 10.54367/jtiust.v6i2.1465.
- [7] F. Bei and S. Sudin, "Analisis Sentimen Aplikasi Tiket Online Di Play Store Menggunakan Metode Support Vector Machine (Svm)," *Sismatik*, vol. 01, no. 01, pp. 91–97, 2021.
- [8] D. G. Nugroho, Y. H. Chrisnanto, and A. Wahana, "Analisis Sentimen pada Jasa Ojek Online. (Nugroho dkk.)," pp. 156–161, 2015.
- [9] M. R. Fais Sya' bani, U. Enri, and T. N. Padilah, "Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritme Naïve Bayes," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 265, 2022, doi: 10.30865/jurikom.v9i2.3989.
- [10] M. K. Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo pada Ulasan Pengguna Di," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 478–483, 2023.
- [11] P. H. Saputro, M. Aristin, and Dy. L. Tyas, "Klasifikasi Lagu Daerah Indonesia Berdasarkan Lirik Menggunakan Metode Tfidf Dan Naïve Bayes," *J. Teknol. Inform. dan Terap.*, vol. 4, no. 1, pp. 45–50, 2017.
- [12] Z. Zhu, J. Liang, D. Li, H. Yu, and G. Liu, "Hot Topic Detection Based on a Refined TF-IDF Algorithm," *IEEE Access*, vol. 7, pp. 26996–27007, 2019, doi: 10.1109/ACCESS.2019.2893980.
- [13] R. Siringoringo, "Klasifikasi Data Tidak Seimbang Menggunakan Algoritma Smote dan K-Nearest Neighbor," *J. ISD*, vol. 3, no. 1, pp. 44–49, 2018.
- [14] A. J. Tobing and A. Febriandirza, "Analisis Sentimen Aplikasi Mobile Banking Bca Pada Ulasan Pengguna di Google Play Store Menggunakan Metode Naive Bayes," *J. Inf. Syst. Res.*, vol. 5, no. 4, pp. 998–1005, 2024, doi: 10.47065/josh.v5i4.5485.
- [15] R. W. Pratiwi, S. F. H, D. Dairoh, D. I. Af'idah, Q. R. A, and A. G. F, "Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM)," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 4, no. 1, pp. 40–46, 2021, doi: 10.20895/inista.v4i1.387.
- [16] T. B. Rohman, D. Dwi Purwanto, and J. Santoso, "Sentiment Analysis Terhadap Review Rumah Makan di Surabaya Memanfaatkan Algoritma Random Forest," *Fak. Sist. Inf.*, vol. 60284, 2018.
- [17] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, "Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.