

# Model Analisis Sentimen pada Ulasan Pengguna Mobile Banking Menggunakan Kombinasi K-Means dan Naive Bayes

Nika Nur Cahayana<sup>1</sup>, Ulfi Saidata Aesy<sup>2,\*</sup>, Kharisma<sup>3</sup>

<sup>1,2,3</sup> Program Studi Sistem Informasi, Fakultas Teknik dan Teknologi Informasi, Universitas Jenderal Achmad Yani Yogyakarta

## Article Info

### Article history:

Received August 16, 2024

Accepted January 3, 2025

Published May 20, 2025

### Keywords:

Mobile BSB

Data mining

Clustering

K-means

Clarification

Naive bayes

## ABSTRACT

Application Mobile Banking Sumsel Babel (BSB) faces the problem of low user satisfaction with a System Usability Scale (SUS) score of 31.46, which is included in the "Not Acceptable" category. The main purpose of this study is so that South Sumatra Babel can make improvements to the BSB application. The methods used include K-Means for clustering review data and Naïve Bayes for classifying user review sentiment. The study found four clusters of user reviews, of which sentiment in cluster 0 and cluster 3 were classified as negative, while cluster 1 and cluster 2 were classified as positive. Negative clusters indicate problems in access, account activation, failed transactions, and frequent application errors. In conclusion, users of the BSB application experienced various difficulties and frustrations related to the stability and reliability of the application, indicating an urgent need for improvement, especially in cluster 0 and cluster 3.



## Corresponding Author:

Ulfi Saidata Aesy,

Program Studi Sistem Informasi, Fakultas Teknik dan Teknologi Informasi

Universitas Jenderal Achmad Yani Yogyakarta,

Jl. Siliwangi, Ringroad Barat, Banyuraden, Gamping, Sleman, Daerah Istimewa Yogyakarta 55293

Email: \* ulfiaesy@gmail.com

## 1. PENGANTAR

Perkembangan teknologi telah mengubah aktivitas manusia secara signifikan, termasuk dalam sektor perbankan. Mobile banking telah menjadi elemen esensial dalam era perbankan digital modern, memungkinkan individu untuk melakukan transaksi perbankan menggunakan smartphone dan akses internet kapan saja dan di mana saja [1] Aksesibilitas yang ditawarkan oleh aplikasi mobile banking sangat mempengaruhi keputusan nasabah dalam memilih layanan perbankan [2]. Oleh karena itu, kepuasan nasabah terhadap penggunaan aplikasi mobile banking menjadi faktor penting dalam meningkatkan pengalaman layanan perbankan digital [3].

Meskipun mobile banking menawarkan banyak manfaat dan kemudahan, tidak semua aplikasi berhasil memberikan pengalaman pengguna yang optimal [4]. Studi menunjukkan bahwa beberapa aplikasi mobile banking menghadapi tantangan terkait rendahnya tingkat kepuasan pengguna karena antarmuka yang kompleks, keterbatasan fitur, kinerja aplikasi yang lambat, dan masalah keamanan [5].

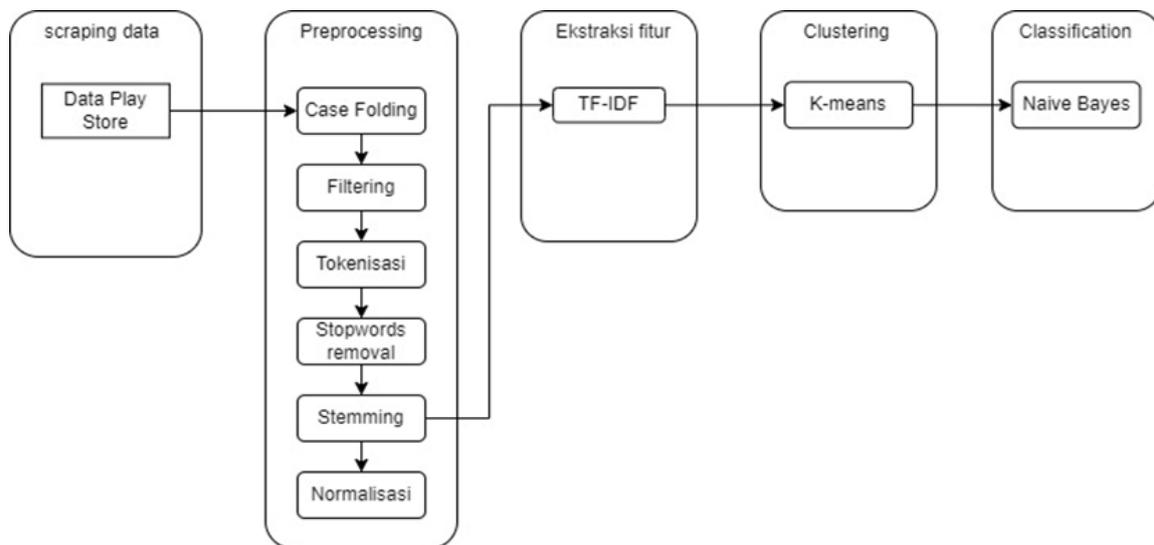
Aplikasi Mobile Banking Sumsel Babel (BSB), yang melayani nasabah di Sumatera Selatan dan Bangka Belitung, menghadapi masalah rendahnya tingkat kepuasan pengguna. Studi sebelumnya menggunakan System Usability Scale (SUS) menunjukkan skor 31,46, yang termasuk dalam kategori "Tidak Dapat Diterima" dan penilaian "Buruk" dalam skala Adjective Rating, menyoroti tantangan besar dalam hal kegunaan [6]. Hal ini menunjukkan perlunya peningkatan substansial untuk memenuhi harapan pengguna. Dengan adanya masalah SUS yang rendah maka perlu dilakukan analisis kembali supaya mendapatkan hasil yang akurat Analisis sentimen pengguna menjadi sangat penting dalam memahami persepsi dan preferensi pengguna terhadap suatu aplikasi. Dengan menggunakan metode K-Means dan Naive Bayes, penelitian ini akan memungkinkan identifikasi pola sentimen yang mungkin tersembunyi dalam data ulasan pengguna

Analisis sentimen pengguna sangat penting untuk memahami persepsi dan preferensi pengguna [7]. Penelitian ini menggunakan metode K-Means dan Naive Bayes untuk mengidentifikasi pola sentimen dalam ulasan pengguna. K-Means digunakan untuk mengelompokkan data ulasan berdasarkan pola atau karakteristik yang serupa [8], sementara Naive Bayes mengklasifikasikan sentimen ulasan tersebut sebagai positif atau negatif [9]. Pemilihan K-Means dan Naive Bayes didasarkan pada keunggulan masing-masing dalam analisis sentimen. K-Means efektif dalam mengelompokkan data tanpa supervisi, mengungkap pola sentimen tersembunyi, sementara Naive Bayes unggul dalam klasifikasi data berdasarkan probabilitas, menawarkan akurasi tinggi dan waktu komputasi cepat [9].

Tujuan utama penelitian ini adalah agar Sumsel Babel dapat melakukan perbaikan terhadap aplikasi BSB. Hasil penelitian diharapkan dapat membantu penyedia layanan perbankan dalam memahami preferensi dan kebutuhan pengguna, sehingga dapat melakukan penyesuaian yang diperlukan untuk meningkatkan kepuasan pengguna dengan aplikasi BSB.

## 2. METODE PENELITIAN

Penelitian ini menggunakan data dari hasil crawling dari Google Play Store dengan menggunakan ulasan pengguna. Tahapan penelitian dapat dilihat pada gambar 1.



Gambar 1. Tahapan penelitian

- a. Pengambilan data  
Pada tahap awal, data akan diambil dari Play Store dengan menggunakan bahasa pemrograman Python dan perangkat lunak Anaconda. Data yang diambil akan berupa ulasan pengguna mengenai aplikasi Mobile Banking Sumsel Babel.
- b. Preprocessing  
Preprocessing merupakan tahap penting pertama dalam memproses data yang bertujuan untuk membersihkan, mengubah format, serta menyiapkan data untuk analisis mendalam [10]. Pada tahap ini, beberapa langkah yang umum dilakukan meliputi (Case folding) merubah huruf kapital menjadi huruf kecil. (Filtering) menghapus tanda baca, angka, dan karakter khusus. (tokenization) membagi teks yang berupa kalimat, dokumen, menjadi bagian bagian tertentu. (stopwords) penghapusan kata-kata umum yang tidak memberikan makna signifikan. (stemming) pengubahan kata-kata ke dalam bentuk dasarnya, dan (normalisasi) kata-kata seperti mengubah kata singkatan menjadi kata baku [11].
- c. Ekstraksi Fitur  
Proses ekstraksi fitur menggunakan metode TF-IDF (Term frequency-Inverse Document Frequency) merupakan langkah penting untuk menganalisis sentimen pengguna terhadap aplikasi Mobile Banking Sumsel Babel di Play Store. Ide dasarnya adalah untuk mengidentifikasi kata-kata kunci yang paling penting dalam setiap ulasan pengguna untuk menentukan tingkat sentimen positif, netral, atau negatif terhadap aplikasi tersebut.
- d. Clustering  
Digunakan untuk mengelompokkan data ulasan pengguna aplikasi Mobile Banking Sumsel Babel ke dalam kelompok-kelompok berdasarkan pola sentimen yang serupa[12].
- e. Klasifikasi

Klasifikasi sentimen dilakukan menggunakan metode Naïve Bayes yang mana dalam penelitian ini digunakan untuk menentukan sentimen negative, positif, dan netral pada setiap cluster [13].

### 3. HASIL DAN ANALISIS

Pada bagian ini adalah pembahasan hasil penelitian berdasarkan tahap penelitian yang ditunjukkan gambar 1. Pengumpulan Data dilakukan dengan mengambil data ulasan pengguna di google play store. Data diambil pada 30 April 2024. Data ulasan yang diperoleh sebanyak 2985 data ulasan pengguna. Data sebanyak 2985 berasal dari komentar di google play store oleh pengguna dengan metode pengumpulan secara acak. Jumlah ini mencakup berbagai pengguna untuk memastikan keberagaman opini. Sample data yang diperoleh ditunjukkan pada gambar 2.

1	score	content
2	2	Aplikasi sering keluar sendiri. Baru mau sudah ada peringatan 184 . Harus aktivasi terus aktivasi tidak bisa sendiri harus ke kantor truss.. Mohon untuk di perbaiki sekelas mobile banking aja lemot
3	1	Tolong diperbaiki lagi aplikasinya, hampir setiap login sulit, diminta aktivasi lagi nanti ujung-ujungnya diblokir padahal nomer yg dimasukkan sudah benar. Update terbaru malah tidak bisa digunakan untuk
4	1	Ribet, Aktivasi harus ke bank padahal BNI mobile, BRimo sama Livin mandiri udah bisa aktivasi tanpa ke Bank
5	1	2024, transfer rempong, kalau lagi tf pembayarannya sering gagal tapi saldo kesedot. Tf ke Dana tidak bisa, virtual akun jg terbatas.. Sedangkan biaya admin transaksi mahal... Aaaaaghhhhh
6	3	M-banking sering eror minta aktivasi ulang dan proses sangat mempersulit nasabah.
7	3	Sering gangguan pas mau tf
8	5	Dengan aplikasi ini memudahkan kita untuk bertransaksi,...

Gambar 2. Sample hasil crawling data

Tahap selanjutnya adalah *preprocessing* data. Pada tahap ini digunakan untuk membersihkan dan pengolahan data yang diperlukan untuk diolah pada tahap selanjutnya. Pada gambar 3, ditunjukkan data hasil *preprocessing*.

1	score	Preprocessing
2	2	2 aplikasi sering keluar sendiri baru mau ingat aktivasi terus aktivasi sendiri kantor truss mohon baik kelas mobile banking aja lot
3	1	1 baik aplikasi hampir login sulit minta aktivasi ujung ujung blokir padahal nomer yang masuk benar updete baru malah guna transfer baik atmnya kadaluwarsa stok kantor cabang
4	1	1 ribet aktivasi bank padahal bni mobile brimo sama livin mandiri udah aktivasi bank
5	1	1 transfer rempong kalau transfer bayar sering gagal saldo sedot transfer dana bisa virtual akun jg bata biaya admin transaksi mahal
6	3	3 banking sering eror minta aktivasi ulang prose sangat sulit nasabah
7	3	3 sering gangguan pa mau transfer
8	5	5 aplikasi mudah transaksi

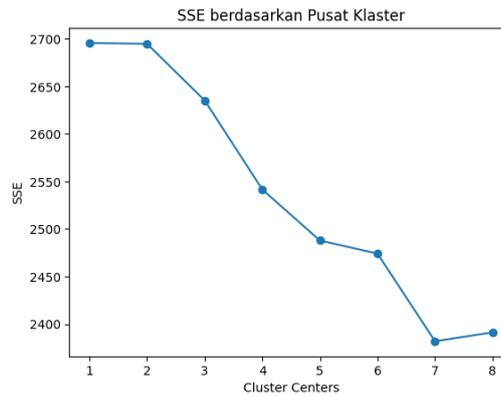
Gambar 3. Data hasil preprocessing

(0, 625)	0.22158232667943062
(0, 606)	0.19099012426733497
(0, 538)	0.1703299314224059
(0, 537)	0.471650009138615
(0, 392)	0.20107280072047187
(0, 390)	0.18242744165110117
(0, 374)	0.14453248447823358
(0, 345)	0.23055721941303517
(0, 280)	0.22495936278715387
(0, 279)	0.2768056321214634
(0, 269)	0.24309357936575984
(0, 232)	0.282777543718692
(0, 60)	0.1897758103690501
(0, 55)	0.2896720635437971
(0, 50)	0.13669397955478485
(0, 33)	0.10390172427323349

Gambar 4. TF-IDF

Perubahan data dari gambar 2 dan gambar 3 terlihat cukup jelas. Pada gambar 2, data masih asli dan masih banyak *noise*. Pada gambar 3, data sudah bersih setelah melewati proses *case folding*, *filtering*, *tokenizing stopword*, *stemming*, dan *normalisasi* yaitu tidak ada *noise*, tidak ada kata imbuhan (hanya kata dasar).

Setelah data bersih, maka data akan diolah dengan menghitung TF-IDF seperti pada gambar 4 setiap katanya. Konversi TF-IDF ke vektor ini digunakan untuk mengubah TF-IDF menjadi bentuk bilangan vektor dan hasil konversi tersebut akan digunakan dalam menentukan jumlah kluster dengan menggunakan metode elbow [14]. Pada gambar 5 menunjukkan penerapan elbow menggunakan SSE.



Gambar 5. Grafik Elbow

Elbow method diterapkan menggunakan perhitungan Sum of Squared Error (SSE) untuk menentukan jumlah kluster yang optimal [15]. Berdasarkan grafik ini, titik elbow tidak terlihat jelas pada K=4, tetapi pemilihan K=4 didasarkan pada pertimbangan praktis dan relevansi domain, seperti kebutuhan untuk membagi data menjadi empat kategori yang dapat diinterpretasikan dengan baik oleh pengguna. Meskipun K=7 menunjukkan penurunan yang lebih signifikan dalam SSE, pemilihan K=4 lebih seimbang antara kompleksitas model dan interpretasi hasil, sehingga cocok untuk tujuan analisis ini.

Setelah K ditentukan, maka data akan dikelompokkan ke dalam *cluster* yang mana terdapat 4 *cluster*. WordCloud digunakan untuk menggambarkan kata-kata yang paling sering muncul dalam teks atau dokumen yang termasuk dalam *cluster* tersebut. Kata-kata yang muncul dengan frekuensi yang lebih tinggi akan ditampilkan dengan ukuran yang lebih besar dalam WordCloud, sementara kata-kata yang jarang muncul akan lebih kecil atau bahkan tidak muncul sama sekali [16]. Pada gambar 5 wordcloud ditunjukkan untuk memahami tema yang dibahas dalam setiap *cluster*.



Gambar 6. Wordcloud Cluster

Dari gambar 5 dan 6, dapat dilihat pada kluster 0 adalah kelompok menurut masyarakat aplikasi bank sumsel babel banyak kekurangan. Sehingga pada kelompok kata kluster 0 berisi keluhan masyarakat tentang aplikasi sumsel babel, misal aktivasi, transaksi, buka, eror dan lain-lain. Dan pada kluster 1 lebih kepada pembahasan terkait kebermanfaatannya aplikasi sumsel babel. Kluster 2 membahas mengenai layanan aplikasi dan peningkatan pada aplikasi sumsel babel. Sedangkan pada kluster 3 lebih membahas masalah operasional dan pengalaman pengguna.

Tabel 1. Analisis Pembahasan

Cluster	Kata Yang Sering Muncul	Hasil Analisis
Cluster 0	aplikasi, aktivasi, transaksi, buka, terus, tidak, blokir, eror, ribet	Pengguna sering menghadapi kesulitan dalam mengakses aplikasi, melakukan aktivasi akun, dan proses transaksi. Keluhan tentang aplikasi yang sering mengalami error, ribet, atau tidak berfungsi dengan baik juga umum.

Cluster 1	mantap, sangat, bantu, transaksi, bagus, mudah, good, manfaat, aman	Pengguna merasa puas dengan kemudahan penggunaan aplikasi dalam melakukan transaksi. Mereka menyoroti keamanan yang baik dan manfaat yang diperoleh dari penggunaan aplikasi.
Cluster 2	baik, aplikasi, sangat, layanan, mohon, tingkat, susah, aman, bantu	Pengguna menganggap layanan aplikasi sudah baik, tetapi ada harapan untuk peningkatan lebih lanjut. Permintaan untuk bantuan lebih intens dalam mengatasi masalah dan memastikan keamanan aplikasi.
Cluster 3	transfer, transaksi, Aktivasi, sering, gagal, masuk, error, ribet, update	Mencakup masalah operasional dan pengalaman pengguna. Pengguna menghadapi masalah teknis seperti kegagalan dalam transfer dan transaksi, serta proses aktivasi yang sering tidak berhasil. Masalah aplikasi yang sering error dan memerlukan update teratur untuk meningkatkan kinerja.

Tahap selanjutnya untuk memperoleh hasil klasifikasi akan di lakukan pelabelan pada setiap cluster, label ini berupa kategori seperti "Positif", "Netral", atau "Negatif" berdasarkan skor sentimen yang diberikan. Pembagian data (split data) yang mana proses membagi dataset menjadi dua subset: data latih (training data) dan data uji (testing data). Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk mengevaluasi kinerja model. Yang terakhir yaitu Klasifikasi dengan *Naive Bayes* model *Naive Bayes* dilatih menggunakan data latih, kemudian digunakan untuk memprediksi label data uji. Hasil prediksi dievaluasi menggunakan metrik evaluasi yang relevan.

Tabel 2. Hasil Klasifikasi

No	Cluster	Hasil Klasifikasi
1	Cluster 0	Negatif
2	Cluster 1	Positif
3	Cluster 2	Positif
4	Cluster 3	Negatif

Berdasarkan gambar tabel 2 analisis pada empat cluster yang telah dilakukan, ditemukan bahwa cluster 0 terkait masalah teknis dalam penggunaan aplikasi hasil sentimennya yaitu 'negatif' (-). Cluster 1 terkait Kepuasan dan Kemudahan Penggunaan hasil sentimennya yaitu 'positif' (+). Cluster 2 terkait Layanan dan Keamanan hasil sentimennya yaitu 'positif' (+). Cluster 3 terkait operasional dan pengalaman pengguna hasil sentimennya 'negatif' (-) [17].

Berdasarkan penelitian ini, rekomendasi untuk BSB adalah fokus pada perbaikan pada Cluster 0 dan Cluster 3. Dimana cluster 0 rata-rata Pengguna sering menghadapi kesulitan dalam mengakses aplikasi, melakukan aktivasi akun, dan proses transaksi. Keluhan tentang aplikasi yang sering mengalami error, ribet, atau tidak berfungsi dengan baik juga umum. dan untuk cluster 3 kegagalan dalam transfer dan transaksi, serta proses aktivasi yang sering tidak berhasil. Masalah aplikasi yang sering error dan memerlukan update teratur untuk meningkatkan kinerja.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian ini, ditemukan berbagai keluhan dari pengguna aplikasi BSB. Pengguna sering menghadapi kesulitan dalam mengakses aplikasi, melakukan aktivasi akun, dan proses transaksi. Ada banyak keluhan tentang aplikasi yang sering mengalami error, terasa ribet, atau tidak berfungsi dengan baik. Pengguna juga melaporkan kegagalan dalam transfer dan transaksi serta masalah dalam proses aktivasi yang sering tidak berhasil. Selain itu, ada keluhan mengenai aplikasi yang membutuhkan update terus menerus untuk meningkatkan kinerja. Semua ini menunjukkan bahwa pengguna mengalami frustrasi dengan stabilitas dan keandalan aplikasi, yang memerlukan perhatian segera untuk perbaikan.

## DAFTAR PUSTAKA

- [1] E. Khoiroh, Farida; Latifah, "Fintech ( Financial Technology ) Dalam Sistem Pembayaran Non Tunai Di Perbankan Syariah ( Studi Kasus Pada Bank Muamalat KCP Lamongan )," *Public Serv. Gov. J.*, vol. 4, 2023.
- [2] D. Anggraeni, "Pengaruh Kualitas Sistem Informasi Akuntansi Terhadap Kepuasan Pengguna Aplikasi Berbasis Mobile Banking Pada Bank BCA KCP Banda," *Rev. Account. Bus.*, vol. 3, no. 2, pp. 170–181, 2023, doi: 10.52250/reas.v3i2.600.
- [3] V. Andriyati, S., Hidayah, N., & Rismayani, "Pengaruh Kualitas Pelayanan, Kemudahan Transaksi dan Fitur Produk Terhadap Kepuasan Nasabah Dalam Penggunaan Bsi Mobile Banking," *J. Ekon. Syariah*, vol. 4, pp. 139–159, 2022.
- [4] D. P. Utami, "Aplikasi Mobile Banking Dalam Transaksi Pada PT . Bank Sumsel Babel Cabang Pembantu Tulung Selapan," *J. Ilm. Mhs. Perbank. Syariah*, vol. 2, pp. 183–196, 2022.
- [5] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018 / 2019," *J. Media Inform. Budidarma*, vol. 4, pp. 51–58, 2020, doi: 10.30865/mib.v4i1.1784.
- [6] Aziz Saputra, "Analisis Pengguna Aplikasi Mobile Banking Sumsel Babel Menggunakan Metode System Usability Scale," *Indones. J. Comput. Sci.*, vol. 12, no. 5, pp. 3115–3126, 2023, doi: 10.33022/ijcs.v12i5.3455.
- [7] I. T. Julianto, "Analisis Sentimen Terhadap Sistem Informasi Akademik Mahasiswa Institut Teknologi Garut," *J. Algoritm.*, vol. 19, pp. 458–465, 2022.
- [8] N. Amalia, N. I. Royanti, and B. Ismanto, "Analisa Sentimen Pelanggan pada Review Belanja Online Berbasis Text Mining Menggunakan Metode K-Means," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1441–1447, 2023, doi: 10.47065/josh.v4i4.3781.
- [9] T. A. P. Pradani Ayu Widya Purnama, "Klasifikasi Penjualan Produk Menggunakan Algoritma Naive Bayes pada Konter HP Bayu Cell," *J. Manaj. Inform. Komput.*, vol. 8, pp. 286–292, 2024.
- [10] A. Agung, A. Daniswara, I. Kadek, and D. Nuryana, "Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru," *J. Informatics Comput. Sci.*, vol. 05, pp. 97–100, 2023.
- [11] R. Rina, T. Hidayat, and D. Uki Eka Saputri, "Analisis Percepatan Pemulihan Ekonomi Indonesia Pasca Pandemi Dengan Big Data Dan Deep Learning," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 3244–3252, 2024, doi: 10.36040/jati.v8i3.9420.
- [12] K. Pratama Simanjuntak and U. Khaira, "Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 7–16, 2021.
- [13] P. W. Cahyo, U. S. Aesy, and B. D. Santosa, "Topic Sentiment Using Logistic Regression and Latent Dirichlet Allocation as a Customer Satisfaction Analysis Model," *J. Infotel*, vol. 16, no. 1, pp. 1–16, 2024, doi: 10.20895/infotel.v16i1.1081.
- [14] R. A. Husen, R. Astuti, L. Marlia, R. Rahmadden, and L. Efrizoni, "Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 211–218, 2023, doi: 10.57152/malcom.v3i2.901.
- [15] V. A. Ekasetya and A. Jananto, "Klusterisasi Optimal Dengan Elbow Method Untuk Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Semarang," *J. Din. Inform.*, vol. 12, no. 1, pp. 20–28, 2020, doi: 10.35315/informatika.v12i1.8159.
- [16] S. M. Fani, R. Santoso, and S. Suparti, "Penerapan Text Mining Untuk Melakukan Clustering Data Tweet Akun Blibli Pada Media Sosial Twitter Menggunakan K-Means Clustering," *J. Gaussian*, vol. 10, no. 4, pp. 583–593, 2021, doi: 10.14710/j.gauss.v10i4.30409.
- [17] H. Hairani, K. E. Saputro, and S. Fadli, "K-means-SMOTE for handling class imbalance in the classification of diabetes with C4.5, SVM, and naive Bayes," *J. Teknol. dan Sist. Komput.*, vol. 8, no. 2, pp. 89–93, 2020, doi: 10.14710/jtsiskom.8.2.2020.89-93.