

Perbandingan Metode Naive Bayes, Decision Tree, dan Support Vector Machine dalam Analisis Sentimen Pengguna terhadap Aplikasi E-Wallet Dana

Muhamad Dimas Adityawarman^{1*}, Windi Irmayani², Muhammad Ifan Rifani Ihsan³

¹Fakultas Teknik & Informatika, Informatika, Universitas Bina Sarana Informatika, Pontianak, Indonesia

Email: ^{1*}dimasaditya1631@gmail.com, ²windi.wnr@bsi.ac.id, ³ifan.mii@bsi.ac.id

(*Email Corresponding Author: dimasaditya1631@gmail.com)

Received: 19 Juni 2026 | Revision: 23 Juni 2026 | Accepted: 24 Juni 2026

Abstrak

Penelitian ini dilatarbelakangi oleh pesatnya pertumbuhan *financial technology* di Indonesia, di mana aplikasi DANA menjadi dompet digital (*e-wallet*) paling populer dengan jumlah pengguna terdaftar mencapai lebih dari 200 juta. Tingginya penggunaan aplikasi ini menghasilkan tumpukan ulasan yang melimpah di Google Play Store yang merepresentasikan tingkat kepuasan maupun keluhan pelanggan. Permasalahan yang diangkat dalam penelitian ini adalah bagaimana mengolah tumpukan ulasan teks tersebut secara otomatis serta menentukan metode klasifikasi terbaik dari tiga algoritma *Machine Learning* yang diuji. Tujuan dari penelitian ini adalah untuk menganalisis dan membandingkan performa akurasi algoritma Naive Bayes, Decision Tree, dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen pengguna. Metode yang digunakan dalam penelitian ini adalah pendekatan kuantitatif komputasional dengan teknik pengumpulan data sekunder berupa 50.000 ulasan dari Google Play Store melalui Kaggle. Proses analisis dilakukan dengan menerapkan lima tahapan prapemrosesan teks (*text preprocessing*), pembobotan fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), penanganan ketidakseimbangan kelas (*imbalanced data*) menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE), pembagian data (80% data latih dan 20% data uji), serta pengujian model menggunakan instrumen Confusion Matrix. Hasil penelitian menunjukkan bahwa algoritma Naive Bayes memiliki kinerja paling superior dengan tingkat akurasi mencapai 80%, disusul oleh Decision Tree dan Support Vector Machine (SVM) yang masing-masing memperoleh akurasi sebesar 78%. Dengan demikian, disimpulkan bahwa algoritma Naive Bayes merupakan metode yang paling optimal dan stabil untuk klasifikasi analisis sentimen pada data teks ulasan aplikasi *e-wallet* setelah distribusi kelas disetarakan.

Kata Kunci: Analisis Sentimen, DANA, Naive Bayes, Decision Tree, Support Vector Machine, SMOTE

Abstract

This research is motivated by the rapid growth of financial technology in Indonesia, where the DANA application has become the most popular digital wallet (*e-wallet*) with over 200 million registered users. The high usage of this application results in an abundance of reviews on the Google Play Store, representing both customer satisfaction and complaints. The problem addressed in this research is how to automatically process these textual reviews and determine the best classification method among the three tested Machine Learning algorithms. This research aims to analyze and compare the accuracy performance of Naive Bayes, Decision Tree, and Support Vector Machine (SVM) algorithms in classifying user sentiment. The method used in this research is a computational quantitative approach, utilizing a secondary data collection technique consisting of 50,000 reviews from the Google Play Store via Kaggle. The analysis process was conducted by applying five stages of text preprocessing, feature weighting using Term Frequency-Inverse Document Frequency (TF-IDF), handling data imbalance using the Synthetic Minority Over-sampling Technique (SMOTE), data splitting (80% training data and 20% testing data), and model evaluation using a Confusion Matrix. The results showed that the Naive Bayes algorithm had the most superior performance with an accuracy rate of 80%, followed by Decision Tree and Support Vector Machine (SVM), each obtaining an accuracy of 78%. Therefore, it can be concluded that the Naive Bayes algorithm is the most optimal and stable method for conducting sentiment analysis classification on *e-wallet* application review text data after the class distribution is equalized.

Keywords: Sentiment Analysis, DANA, Naive Bayes, Decision Tree, Support Vector Machine, SMOTE

1. PENDAHULUAN

Perkembangan teknologi digital di Indonesia mengalami peningkatan yang sangat pesat, khususnya dalam bidang *financial technology* (fintech). Salah satu bentuk perkembangan fintech yang banyak digunakan masyarakat adalah aplikasi dompet digital atau *e-wallet*. *E-wallet* merupakan layanan pembayaran digital yang memungkinkan pengguna melakukan berbagai transaksi secara praktis, seperti transfer uang, pembayaran tagihan, pembelian produk, hingga transaksi non-tunai lainnya. Berdasarkan survei Jakpat tahun 2025, aplikasi DANA menjadi *e-wallet* yang paling banyak digunakan masyarakat Indonesia dengan persentase penggunaan mencapai 72%, mengungguli beberapa aplikasi dompet digital lainnya seperti GoPay dan OVO [1]. PT Espay Debit Indonesia Koe selaku pengembang DANA juga mencatat jumlah pengguna terdaftar mencapai lebih dari 200 juta pengguna pada akhir tahun 2024 [2]. Tingginya jumlah pengguna

tersebut menghasilkan jumlah ulasan yang sangat masif pada platform Google Play Store, yang memuat spektrum opini pengguna berupa sentimen positif, negatif, maupun netral terkait kualitas layanan [3].

Banyaknya informasi ulasan yang beredar setiap hari menyebabkan proses identifikasi dan pengelompokan keluhan secara manual menjadi kurang efektif karena membutuhkan waktu dan sumber daya yang besar. Seiring berkembangnya teknologi kecerdasan buatan, pendekatan *machine learning* melalui bidang penambangan teks (*text mining*) banyak dimanfaatkan untuk membantu proses klasifikasi sentimen secara otomatis. Dalam konteks klasifikasi teks, beberapa algoritma yang sering direkomendasikan adalah Support Vector Machine (SVM), Decision Tree, dan Naive Bayes [4]. SVM dikenal memiliki kemampuan ekstraksi linear yang tangguh dalam menangani data berdimensi tinggi. Decision Tree menawarkan proses klasifikasi yang transparan karena menghasilkan aturan hierarkis dalam bentuk pohon keputusan. Sementara itu, Naive Bayes merupakan algoritma klasifikasi probabilistik yang memiliki pemrosesan komputasi sangat ringan dan konsisten menghasilkan akurasi tinggi.

Selain pemilihan algoritma, kualitas model klasifikasi sangat dipengaruhi oleh distribusi data yang digunakan. Pada dataset ulasan aplikasi e-wallet, sering ditemukan kondisi ketidakseimbangan kelas (*class imbalance*), di mana ulasan Positif dan Negatif mendominasi secara absolut, sementara ulasan Netral sangat minim. Kondisi ini dapat menyebabkan model memproduksi prediksi yang bias terhadap kelas mayoritas. Untuk mengatasi permasalahan tersebut, diperlukan teknik augmentasi data seperti *Synthetic Minority Over-sampling Technique* (SMOTE) yang bekerja membangkitkan data sintesis pada kelas minoritas agar distribusi label menjadi ekuilibrium [5]. Selain itu, penerapan teknik SMOTE telah terbukti secara akademis mampu memberikan pengaruh signifikan terhadap peningkatan performa model klasifikasi pada kelas minoritas agar distribusi label menjadi ekuilibrium [6].

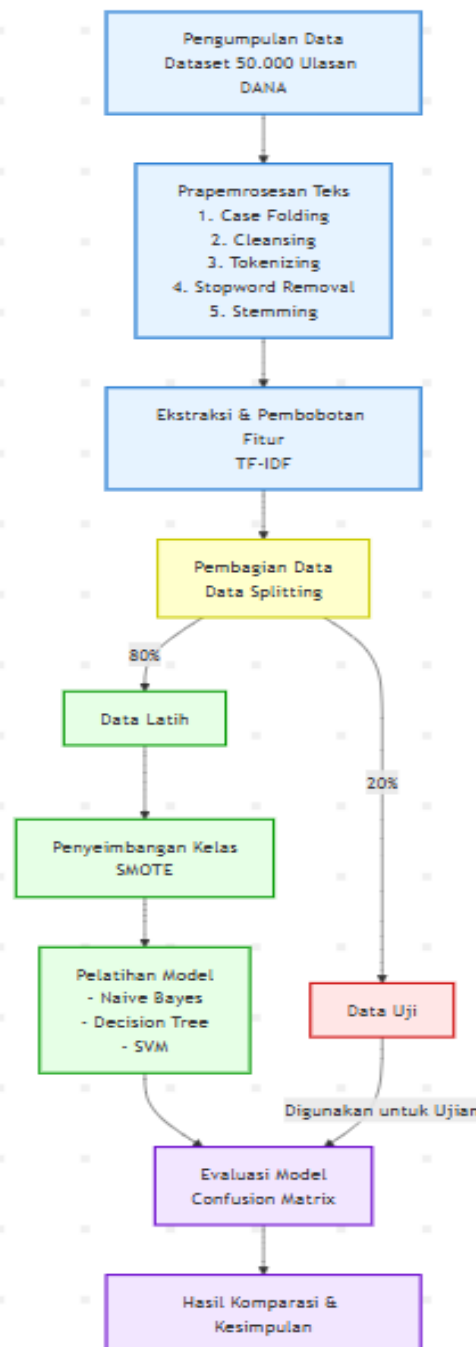
Beberapa penelitian terdahulu telah mengkaji perbandingan metode ini. Analisis opini publik mengenai pembangunan IKN menemukan bahwa algoritma SVM memiliki performa klasifikasi yang lebih tinggi dan unggul dalam memisahkan kelas data dibandingkan dengan algoritma Naive Bayes [7]. Studi komparasi untuk analisis sentimen pada ulasan aplikasi dompet digital OVO juga mengonfirmasi bahwa metode SVM berhasil menghasilkan tingkat akurasi yang lebih superior dan dominan atas algoritma konvensional lainnya [8]. Namun demikian, pendekatan klasifikasi yang menerapkan Naive Bayes yang dikombinasikan dengan metode pembobotan statistik TF-IDF terbukti sanggup mendeteksi opini publik secara efektif dan mendongkrak performa akurasi model secara signifikan [9]. Berdasarkan kajian literatur tersebut, terdapat *gap* penelitian berupa minimnya studi komparatif yang membandingkan performa Naive Bayes, Decision Tree, dan SVM secara simultan dengan menerapkan SMOTE di dalam *pipeline* khusus pada ekosistem dataset ulasan aplikasi DANA. Oleh karena itu, penelitian ini bertujuan untuk membandingkan performa ketiga algoritma tersebut secara *head-to-head* untuk mencari arsitektur mana yang paling stabil, optimal, dan akurat dalam mengekstraksi klasifikasi polaritas pengguna aplikasi *e-wallet* tersebut.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif komputasional berbasis *machine learning*. Dalam ranah analisis teks, pendekatan *supervised machine learning* terbukti sangat andal untuk melatih sistem komputer menggunakan sekumpulan data ulasan digital yang telah diberi label kelas sebelumnya [10], sehingga model yang terbentuk mampu memprediksi dan mengklasifikasikan opini sentimen pada data teks baru dengan tingkat performa yang terukur [11].

Untuk mengklasifikasikan polaritas sentimen pengguna terhadap aplikasi dompet digital DANA secara otomatis. Dataset ulasan bersumber dari Google Play Store yang diperoleh melalui platform Kaggle. Proses penelitian dirancang secara sistematis dan terstruktur melalui beberapa tahapan yang saling berkaitan, yang meliputi pengumpulan data, prapemrosesan teks (*text preprocessing*) berbahasa Indonesia, ekstraksi fitur dan pembobotan menggunakan TF-IDF, pembagian proporsi data latih dan data uji (80:20), penanganan ketidakseimbangan kelas (*imbalanced data*) menggunakan teknik SMOTE, pelatihan tiga algoritma klasifikasi secara paralel (Naive Bayes, Decision Tree, dan SVM), serta komparasi performa model yang dievaluasi menggunakan instrumen *Confusion Matrix*.

Tahapan alur penelitian secara keseluruhan dapat dilihat secara visual pada Gambar 1.



Gambar 1: Alur Tahapan Penelitian

2.1 Pengumpulan Data (*Data Collection*)

Data yang digunakan dalam penelitian ini adalah data sekunder berupa dataset teks ulasan pengguna aplikasi DANA yang diekstraksi dari platform Kaggle (yang bersumber asli dari Google Play Store) [12]. Kumpulan data ini mencakup teks ulasan, *rating*, dan tanggal ulasan yang merepresentasikan opini sentimen positif, negatif, atau netral dari para pengguna aplikasi tersebut. Total dataset yang digunakan berjumlah 50.000 baris data.

2.2 Prapemrosesan Data Teks (*Text Preprocessing*)

Data ulasan yang diperoleh dari platform digital merupakan data tekstual tidak terstruktur yang memiliki tingkat *noise* tinggi, seperti penggunaan huruf kapital yang tidak konsisten, bahasa gaul (*slang*), angka, singkatan, maupun tanda

baca yang tidak relevan bagi model komputer [13]. Oleh karena itu, dilakukan pembersihan data teks melalui lima tahapan berurutan.

- Case Folding*: Mengubah seluruh huruf kapital dalam teks ulasan menjadi huruf kecil (*lowercase*), agar sistem menganggap kata yang sama dengan variasi kapitalisasi sebagai satu kata yang identik.
- Cleansing*: Menghapus angka, tanda baca, simbol khusus, *emoticon*, tautan URL, dan karakter khusus lainnya yang mengganggu proses analisis.
- Tokenizing*: Memotong atau memecah sebaris kalimat ulasan panjang menjadi kumpulan potongan kata-kata tunggal (*token*).
- Stopword Removal*: Menyaring dan membuang kata-kata umum (*stopwords*) yang sering muncul namun tidak memiliki muatan informasi sentimen penting, seperti kata hubung "yang", "dan", "di", serta "ke".
- Stemming*: Mengubah kata-kata yang memiliki imbuhan menjadi bentuk kata dasarnya yang baku sesuai dengan struktur linguistik bahasa Indonesia.

2.3 Pembobotan Kata (TF-IDF)

Setelah teks ulasan bersih dari tahap prapemrosesan, data tekstual tersebut ditranslasikan ke dalam bentuk representasi matriks angka numerik agar dapat dipahami dan diproses oleh algoritma *machine learning* [14]. Metode *Term Frequency–Inverse Document Frequency* (TF-IDF) digunakan dalam tahap ini untuk menghitung frekuensi kemunculan suatu kata kunci di dalam dokumen ulasan dan memberikan bobot signifikansi yang proporsional pada kata-kata yang bermuatan sentimen. Rumus matematis TF-IDF secara umum diformulasikan sebagai berikut:

$$W_{dt} = t f_{dt} \times \log \left(\frac{N}{df_t} \right)$$

Keterangan:

- W_{dt} = Nilai bobot kata
- $t f_{dt}$ = Frekuensi kemunculan kata t dalam dokumen d
- N = Total seluruh dokumen/ulasan dalam *dataset*
- df_t = Jumlah dokumen yang mengandung kata t .

2.4 Pembagian Data dan Penanganan Ketidakseimbangan (SMOTE)

Matriks pembobotan fitur numerik dari hasil proses TF-IDF selanjutnya dibagi menjadi dua bagian utama (*data splitting*) secara acak dengan proporsi parameter 80% untuk data latih (*training data*) dan 20% untuk data uji (*testing data*). Sebelum data dilatih ke dalam model, dilakukan penyeimbangan proporsi kelas sentimen menggunakan algoritma *Synthetic Minority Over-sampling Technique* (SMOTE) [5]. Algoritma ini mensintesis data baru pada kelas minoritas (kelas Netral) berdasarkan kedekatan jarak antar data (*K-Nearest Neighbors*), sehingga proporsi data pada kelas minoritas menjadi setara dengan kelas mayoritas pada fase pelatihan dan mencegah model menghasilkan prediksi yang bias [15].

2.5 Pemodelan dan Evaluasi Performa

Data latih yang telah diseimbangkan dimasukkan ke dalam lingkungan komputasi untuk membentuk model klasifikasi sentimen menggunakan tiga jenis algoritma berbeda secara terpisah:

- Naive Bayes*: Memanfaatkan pendekatan probabilitas statistik bersyarat untuk mengelompokkan teks berdasarkan kemunculan kata kunci.
- Decision Tree*: Membentuk aturan keputusan hierarkis berbasis struktur pohon logika untuk memisahkan kategori sentimen.
- Support Vector Machine* (SVM): Mencari bidang pemisah linier optimal (*hyperplane*) dengan jarak (*margin*) terbesar di antara kelas-kelas sentimen yang berbeda.

Model klasifikasi yang telah terbentuk kemudian diuji kemampuannya menggunakan data uji (*testing data*). Kinerja prediktif dari masing-masing algoritma dievaluasi secara komprehensif menggunakan instrumen matematis *Confusion Matrix* (Valerian et al., 2025). Parameter evaluasi yang dihitung meliputi *Accuracy*, *Precision*, *Recall*, dan *F1-Score* (Hidayah, 2024), dengan rumusan sebagai berikut:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Pada bagian ini diuraikan hasil dari tahapan pengumpulan data, pengolahan, pemodelan data teks ulasan pengguna aplikasi DANA, hingga komparasi evaluasi dari tiga algoritma *machine learning* yang diuji secara paralel.

3.1 Hasil Eksplorasi dan Pembagian Data

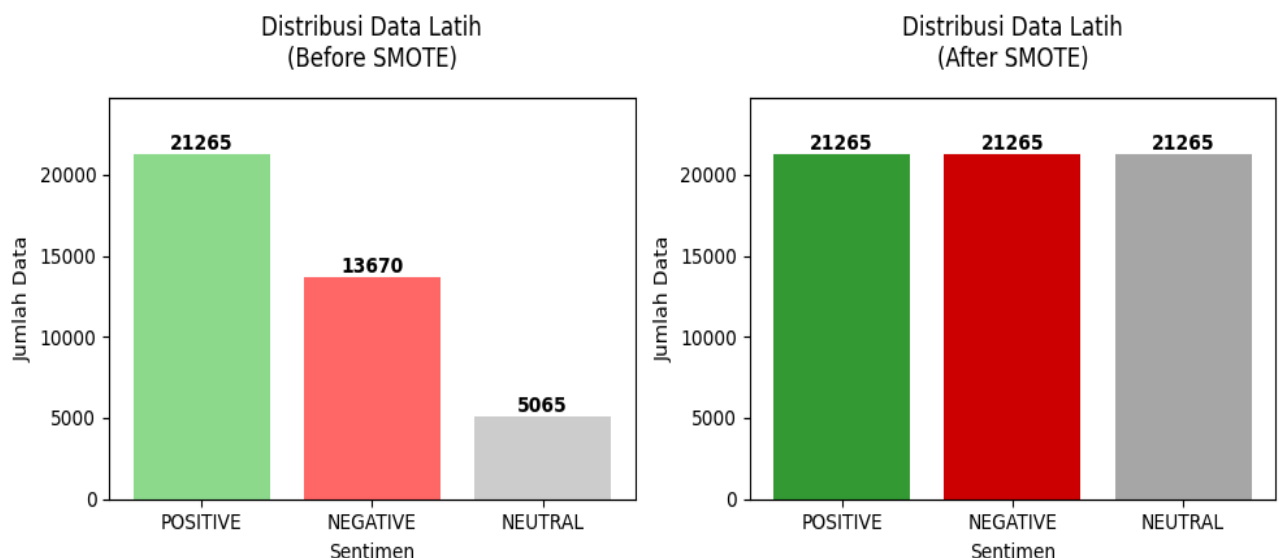
Berdasarkan hasil ekstraksi dataset sekunder dari platform Google Play Store, diperoleh total data ulasan sebanyak 50.000 baris data yang terbagi ke dalam tiga kelas polaritas sentimen. Rincian distribusi kelas sentimen pada keseluruhan dataset awal tersebut adalah Sentimen Positif sebanyak 26.555 ulasan (53,11%), diikuti oleh Sentimen Negatif sebanyak 17.073 ulasan (34,15%), dan Sentimen Netral sebanyak 6.372 ulasan (12,74%). Dataset berjumlah 50.000 ulasan tersebut kemudian dibersihkan dengan menghapus nilai kosong (*missing value*) dan ditransformasikan menggunakan pembobotan matriks TF-IDF.

Matriks data selanjutnya dipartisi menggunakan parameter rasio 80:20, yang memecah korpus menjadi 10.000 ulasan sebagai target Data Uji (*Testing Data*) independen, dan 40.000 ulasan dialokasikan sebagai Data Latih (*Training Data*) untuk melatih algoritma model. Melalui tahapan *Exploratory Data Analysis* (EDA) yang difokuskan pada 40.000 Data Latih tersebut, teridentifikasi adanya rasio ketidakseimbangan kelas (*imbalanced data*) yang ekstrem. Rincian hierarki distribusinya adalah: Sentimen Positif mendominasi dengan 21.265 ulasan, Sentimen Negatif membuntuti dengan 13.670 ulasan, dan Sentimen Netral menjadi kelas minoritas dengan 5.065 ulasan.

Kondisi ketimpangan distribusi ini merupakan fenomena empiris yang sangat wajar dalam data ulasan aplikasi digital. Secara psikologis, probabilitas pengguna untuk meluangkan waktu memberikan ulasan akan memuncak pada dua kondisi ekstrem, yakni ketika mereka merasa sangat terpuaskan oleh layanan (membentuk kelas positif yang besar) atau ketika mereka merasa sangat dirugikan (membentuk kelas negatif). Sementara itu, pengguna dengan pengalaman standar jarang memberikan ulasan, sehingga kelas netral selalu menyusut menjadi minoritas.

3.2 Hasil Penanganan Ketidakseimbangan Kelas

Untuk menangani masalah ketimpangan kelas yang signifikan tersebut, teknik augmentasi data *Synthetic Minority Over-sampling Technique* (SMOTE) diterapkan secara eksklusif pada matriks 40.000 Data Latih. Algoritma SMOTE secara efektif berhasil mensintesis sampel-sampel teks buatan pada titik koordinat kelas minoritas, sehingga proporsi kelas Sentimen Negatif dan Netral mengembang dan disetarakan secara matematis agar sejajar dengan dominasi kelas Positif, yakni masing-masing ditarik mencapai 21.265 sampel data latih. Perbandingan distribusi kelas sentimen pada data latih sebelum dan sesudah penerapan SMOTE divisualisasikan pada Gambar 2.



Gambar 2: Distribusi Data Latih Sebelum dan Sesudah SMOTE

Ekuilibrum proporsi data ini mutlak diperlukan sebelum matriks diserap oleh model klasifikasi agar sistem tidak menghasilkan prediksi yang bias dan terhindar dari kecenderungan menghafal (*overfitting*) pada kelas mayoritas.

3.3 Hasil Evaluasi Model dan Analisis Komparasi

Setelah matriks data latih berhasil diseimbangkan, pengujian klasifikasi sentimen dieksekusi menggunakan ketiga algoritma secara paralel terhadap 10.000 data uji independen yang berada dalam kondisi asli (tanpa efek augmentasi SMOTE). Rincian performa dari masing-masing model dievaluasi melalui *Confusion Matrix* dan dijabarkan sebagai berikut:

- Naive Bayes:** Berdasarkan pengujian menggunakan algoritma probabilitas *MultinomialNB*, model ini meraih skor akurasi (*Accuracy*) global tertinggi sebesar 80,00%. Metrik pendukung lainnya menunjukkan tingkat kepresisian (*Precision* rata-rata makro) sebesar 73,00%, kepekaan klasifikasi (*Recall* rata-rata makro) 74,00%, dan indeks stabilitas performa (*F1-Score* rata-rata makro) sebesar 73,00%.
- Decision Tree:** Berdasarkan pengujian menggunakan arsitektur pohon keputusan (*DecisionTreeClassifier*), hasil metrik evaluasi menunjukkan capaian *Accuracy* yang tertahan di level 78,00%. Metrik pendukung lainnya mencatatkan *Precision* rata-rata makro 70,00%, *Recall* rata-rata makro 71,00%, dan *F1-Score* rata-rata makro 70,00%.
- Support Vector Machine (SVM):** Pengujian yang dilakukan menggunakan algoritma batas linier *LinearSVC* mendapati lonjakan metrik yang juga terhenti pada level *Accuracy* 78,00%. Untuk metrik pendukungnya, SVM memperoleh *Precision* rata-rata makro 71,00%, *Recall* rata-rata makro 74,00%, dan *F1-Score* rata-rata makro 72,00%.

Berdasarkan komparasi performa tersebut, dapat diamati bahwa implementasi teknik augmentasi SMOTE memberikan dampak komputasional yang sangat krusial dan transformatif terhadap fondasi performa algoritma Naive Bayes, khususnya dalam mengenali sentimen kelas minoritas (Netral). Terjadi lonjakan rasio *recall* yang sangat signifikan, di mana model probabilitas tersebut mampu mengklasifikasikan sebaran data kelas minoritas dengan tingkat kepekaan analitik yang jauh lebih tajam. Hal ini sukses mengukuhkan algoritma berbasis probabilitas bersyarat ini sebagai metode dengan tingkat akurasi prediksi tertinggi, yakni menyentuh angka 80,00%.

Di sisi lain, injeksi sampel data sintesis dari SMOTE justru mendorong algoritma Decision Tree untuk mengonstruksi cabang aturan keputusan (*rule-based trees*) yang terlampau kompleks dan spesifik terhadap data latih. Ekstraksi fitur yang terlampau kaku (*overfitting*) ini menyebabkan berkurangnya fleksibilitas generalisasi model ketika berhadapan dengan data uji yang baru. Fenomena komputasi serupa yang anomali menimpa algoritma Support Vector Machine (SVM), di mana penambahan sampel data sintesis secara masif pada kelas Netral menyebabkan terjadinya kepadatan dan tumpang tindih (*overlapping*) pada lanskap ruang fitur klasifikasi. Kondisi tabrakan tersebut mendesak posisi koordinat batas pemisah (*hyperplane*) linier milik SVM agar bergeser menjauh dari titik optimalnya, sehingga mereduksi tingkat presisi absolut model dalam mendefrensiasi antarkelas, dan secara instan mendegradasi kinerja puncaknya menjadi 78,00%.

4. KESIMPULAN

Penelitian ini berhasil menyimpulkan bahwa algoritma pembelajaran mesin Naive Bayes, Decision Tree, dan Support Vector Machine (SVM) dapat diimplementasikan secara komprehensif untuk mengklasifikasikan polaritas sentimen ulasan pengguna aplikasi *e-wallet* DANA ke dalam kategori positif, negatif, dan netral. Berdasarkan pengujian komparatif terhadap 10.000 data uji independen yang telah diseimbangkan proporsinya menggunakan algoritma SMOTE, Naive Bayes terbukti menjadi arsitektur klasifikasi yang paling superior dan optimal dengan capaian akurasi puncak sebesar 80,00%. Performa probabilitas ini secara mutlak berhasil mengungguli arsitektur Decision Tree dan SVM yang keduanya tertahan pada capaian akurasi maksimal sebesar 78,00%. Penerapan teknik augmentasi SMOTE terbukti sangat esensial dalam memecahkan masalah ketidakseimbangan kelas—di mana kelas netral awalnya menderita defisit ekstrem di angka 12,74%—sehingga sukses mendongkrak kepekaan prediksi (*recall*) secara signifikan pada tubuh algoritma Naive Bayes. Meskipun demikian, tumpukan sampel data sintesis tersebut justru memicu pergeseran pada ruang keputusan batas linier (*hyperplane*) milik SVM, yang mengakibatkan pelemahan akurasi dan memvalidasi keunggulan Naive Bayes dalam ekosistem data hasil *oversampling* ini. Temuan empiris ini memberikan rekomendasi kuat bahwa model probabilitas Naive Bayes merupakan instrumen analitik yang paling tepat untuk diadopsi oleh pihak pengembang aplikasi dalam mengevaluasi *feedback* pelanggan secara otomatis. Lebih lanjut, tingginya persentase keluhan bersentimen negatif yang menyentuh angka 34,15% seyogianya menjadi landasan evaluasi operasional bagi *developer* aplikasi DANA, secara spesifik untuk memprioritaskan perbaikan stabilitas *server* pada fitur transfer data (tf) serta mekanisme penarikan dana tunai (wd) yang teridentifikasi mendominasi keluhan utama pengguna. Untuk pengembangan riset komputasional di masa mendatang, peneliti selanjutnya sangat direkomendasikan untuk mengeksplorasi dan membandingkan performa arsitektur *Deep Learning* tingkat lanjut, seperti *Long Short-Term Memory* (LSTM) maupun model representasi bahasa BERT, guna memperluas dimensi komparasi akurasi pada pengolahan teks ulasan berbahasa Indonesia.

REFERENCES

- [1] A. Z. Yonatan, "5 Platform E-Wallet Terpopuler di Indonesia," *5 Platf. E-Wallet Terpopuler di Indones.*, 2025, [Online]. Available: <https://goodstats.id/article/5-platform-e-wallet-terpopuler-di-indonesia-2025-R1GNX>
- [2] E. Kure, "Pengguna Platform Dana Tembus 200 Juta," *Pengguna Platf. Dana Tembus 200 Juta*, 2025, [Online]. Available: <https://investor.id/finance/387968/2024-pengguna-platform-dana-tembus-200-juta>
- [3] M. A. Sriani; Armansyah; Rambe, "Analisis Sentimen Pengguna Aplikasi Dana Menggunakan Metode Naïve Bayes," *Anal. Sentimen Pengguna Apl. Dana Menggunakan Metod. Naïve Bayes*, 2024, [Online]. Available: https://www.researchgate.net/publication/385869493_Analisis_Sentimen_Pengguna_Aplikasi_Dana_Menggunakan_Metode_Naive_Bayes
- [4] R. D. Cahyani and P. T. Prasetyaningrum, "Sentiment Analysis of User Reviews for AI Applications : Evaluating SVM , Logistic Regression , and Random Forest," vol. 8, no. 1, pp. 1–27, 2026, doi: 10.63158/journalisi.v8i1.1366.
- [5] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE : Synthetic Minority Over-sampling Technique," vol. 16, pp. 321–357, 2022.
- [6] C. Haryawan and Y. M. K. Ardhana, "Analisa Perbandingan Teknik Oversampling SMOTE," *JIRE (Jurnal Inform. Rekayasa Elektron.*, vol. 6, no. 1, pp. 73–78, 2023.
- [7] N. Hadi and D. Sugiarto, "Analisis Sentimen Pembangunan IKN pada Media Sosial X Menggunakan Algoritma SVM , Logistic Regression dan Naïve Bayes," vol. 10, no. 1, pp. 37–49, 2025, doi: 10.30591/jpit.v10i1.7106.
- [8] A. Lowell, A. Lowell, K. Candra, and E. Indra, "Perbandingan Metode Support Vector Machine (SVM) Dan Naive Bayes Pada Analisis Sentimen Ulasan Aplikasi OVO JURNAL MEDIA INFORMATIKA [JUMIN]," vol. 6, no. 2, pp. 896–905, 2025.
- [9] D. S. Nugroho, I. F. Hanif, and M. A. Hasbi, "Sentiment Analysis of Alleged 2024 Election Fraud Based on Tweets Using the Naïve Bayes Classifier Algorithm Analisis Sentimen Dugaan Pelanggaran Pemilu 2024 Berdasarkan Tweet Menggunakan Algoritma Naïve Bayes Classifier," vol. 4, no. July, pp. 1169–1176, 2024.
- [10] R. Artikel, N. Luh, P. Chandra, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, "Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning," vol. 7, no. April, pp. 47–58, 2021.
- [11] A. M. Wahid, Turino, K. A. Nugroho, D. Titi Safitri4, and F. S. Utomo, "Optimasi Logistic Regression dan Random Forest untuk Deteksi Berita Hoax Optimasi Logistic Regression dan Random Forest untuk Deteksi Berita Hoax Berbasis Hyperparameter Optimization of Logistic Regression and Random Forest for Hoax News Detection Using T," *J. Pendidik. dan Teknol. Indones.*, vol. 4, no. January, pp. 381–392, 2025.
- [12] B. Irawan, A. Bahtiar, N. Bayes, G. P. Store, and N. Bayes, "PENERAPAN METODE NAIVE BAYES PADA ANALISIS SENTIMEN APLIKASI," vol. 8, no. 1, pp. 759–766, 2024.
- [13] B. Hakim, "ANALISA SENTIMEN DATA TEXT PREPROCESSING PADA DATA MINING DENGAN MENGGUNAKAN MACHINE LEARNING DATA TEXT PRE-PROCESSING SENTIMENT ANALYSIS IN DATA MINING USING MACHINE LEARNING School of Computer Science and Technology , Harbin Institute of Technology," vol. 4, no. 2, pp. 16–22, 2021.
- [14] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes Pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH TF-IDF Weighting Using Naïve Bayes on Public Sentiment on The Issue of Rising BIPIH," vol. 13, no. April, pp. 84–93, 2023.
- [15] M. A. Hermawan, A. Faqih, G. Dwilestari, T. Informatika, and S. Informasi, "IMPLEMENTASI AKURASI MODEL NAIVE BAYES MENGGUNAKAN SMOTE DALAM ANALISIS SENTIMEN PENGGUNA APLIKASI BRIMO," vol. 13, no. 1, 2025.