

Algoritma Naive Bayes dalam Analisis Komperatif Sentimen Dompot Digital

Fenty Ariany¹, Agung Tri Prastowo², Imam Ahmad³, Qadhli Jafar Adrian⁴

¹Program Studi Teknologi Informasi, ²Program Studi Teknik Komputer, ^{3,4}Program Studi Sistem Informasi
Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia
Email: fenty@teknokrat.ac.id¹, fenty@teknokrat.ac.id², imamahmad@teknokrat.ac.id³,
qadhliadrian@teknokrat.ac.id⁴

Abstrak—Perkembangan teknologi finansial di Indonesia telah mendorong adopsi dompet digital secara masif, dengan jumlah pengguna aktif mencapai lebih dari 60 juta pada tahun 2024. Namun, tingginya tingkat penggunaan tersebut tidak selalu berbanding lurus dengan tingkat kepuasan pengguna. Oleh karena itu, analisis sentimen menjadi penting untuk memahami persepsi dan opini publik terhadap layanan dompet digital. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna dompet digital menggunakan algoritma Naive Bayes. Dataset yang digunakan terdiri dari 18.869 ulasan pengguna yang dikumpulkan dari Google Play Store pada empat aplikasi dompet digital populer di Indonesia, yaitu DANA, OVO, DOKU, dan LinkAja. Tahapan preprocessing data meliputi case folding, cleaning, tokenization, stopword removal, serta stemming menggunakan library Sastrawi. Proses ekstraksi fitur dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dengan parameter $max_features$ sebesar 5000. Selanjutnya, optimasi parameter model Naive Bayes dilakukan menggunakan GridSearchCV. Hasil penelitian menunjukkan bahwa model Naive Bayes dengan parameter $\alpha = 0,1$ mampu mencapai tingkat akurasi sebesar 88,5% dan nilai AUC-ROC sebesar 0,9237, yang mengindikasikan kinerja klasifikasi sentimen yang sangat baik.

Kata Kunci: Analisis Sentiment ; Digital Wallet; Naive Bayes; TF-IDH; Ulasan Pengguna

Abstract—The rapid development of financial technology in Indonesia has driven the massive adoption of digital wallets, with more than 60 million active users recorded in 2024. However, a high level of usage does not always correspond to user satisfaction. Therefore, sentiment analysis is essential to understand public perceptions and opinions regarding digital wallet services. This study aims to analyze user sentiment toward digital wallets using the Naive Bayes algorithm. The dataset consists of 18,869 user reviews collected from Google Play Store for four popular digital wallet applications in Indonesia, namely DANA, OVO, DOKU, and LinkAja. The data preprocessing stages include case folding, cleaning, tokenization, stopword removal, and stemming using the Sastrawi library. Feature extraction is performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method with a maximum feature parameter of 5,000. Furthermore, parameter optimization for the Naive Bayes model is conducted using GridSearchCV. The experimental results show that the Naive Bayes model with an α parameter of 0.1 achieves an accuracy of 88.5% and an AUC-ROC value of 0.9237, indicating excellent performance in sentiment classification.

Keywords: Sentiment Analysis; Digital Wallet; Naive Bayes; TF-IDH; User Reviews

1. PENDAHULUAN

Transformasi digital di sektor keuangan Indonesia telah mendorong pertumbuhan pesat penggunaan aplikasi dompet digital (*e-wallet*). Berdasarkan data Bank Indonesia, nilai transaksi uang elektronik mencapai Rp516,02 triliun pada tahun 2024, meningkat sebesar 37,3% dibandingkan tahun sebelumnya [1]. Berbagai aplikasi *e-wallet* seperti DANA, OVO, DOKU, dan LinkAja kini telah menjadi bagian integral dalam kehidupan masyarakat Indonesia, khususnya dalam mendukung aktivitas transaksi pembayaran digital sehari-hari.

Perkembangan teknologi finansial di Indonesia telah mendorong peningkatan penggunaan dompet digital (*e-wallet*) secara signifikan khususnya di Indonesia yang mengalami pertumbuhan ekonomi yang sangat pesat khususnya dalam transaksi pembayaran secara digital. Tingginya tingkat adopsi tersebut tidak selalu diikuti dengan kepuasan pengguna terhadap kualitas layanan yang diberikan. Berbagai ulasan pengguna pada platform distribusi aplikasi seperti google play store menunjukkan adanya beragam persepsi, baik positif maupun negatif, terkait performe, keamanan, kemudahan penggunaan maupun kualitas fitur aplikasi *e-wallet* [2]. Informasi ini menjadi penting untuk dianalisis guna mengetahui tingkat kepuasan pengguna serta mengidentifikasi layanan *e-wallet* yang dinilai paling baik oleh pengguna.

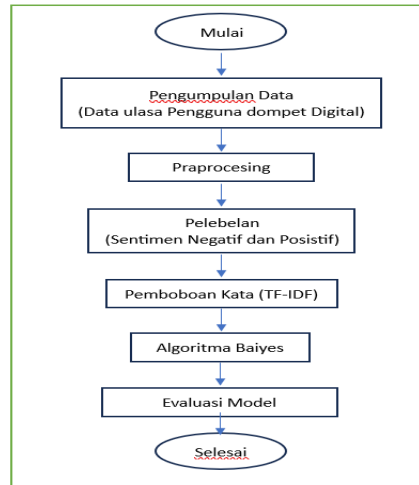
Permasalahan yang muncul adalah jumlah ulasan penggunaan yang sangat besar sehingga analisis secara manual menjadi tidak efisien dan berpotensi menghasilkan subjektivitas. Oleh karena itu diperlukan pendekatan otomatis berbasis machine learning untuk mengklasifikasi sentimen pengguna secara sistematis dan objektif. Analisis sentimen merupakan teknik text mining yang digunakan untuk mengidentifikasi opini, emosi dan evaluasi pengguna terhadap suatu produk atau layanan sehingga dapat mendukung pengambilan keputusan berbasis data [3].

Salah satu metode yang banyak digunakan dalam klasifikasi sentimen adalah algoritma Naive Bayes, dimana algoritma naive bayes memiliki kemampuan dalam mengelola data teks berdimensi tinggi dan efisiensi dalam komputasi [4]. Penelitian sebelumnya menunjukkan bahwa algoritma Naive Bayes mampu memberikan perform yang baik dalam klasifikasi

teks, khususnya pada analisis sentimen ulasan pengguna aplikasi digital[5]. Berdasarkan hal ini, penelitian ini menerapkan metode analisis sentimen berbasis Naive Bayrs terhadap ulasan pengguna e-walet dengan tujuan mengidentifikasi persepsi pengguna serta menentukan layanan e-wallet yang memiliki sentimen terbaik berdasarkan opini pengguna.

2. METODE PENELITIAN

Metodologi penelitian ini disusun untuk memberikan solusi dalam mengklasifikasikan sentimen pengguna terhadap layanan dompet digital di Indonesia. Pendekatan yang digunakan adalah *text mining* dengan algoritma **Naive Bayes**, yang dipilih karena efisien dan efektif dalam menangani data teks berdimensi tinggi. Alur penelitian terdiri dari enam tahapan utama sebagaimana ditunjukkan pada diagram alur penelitian.



Gambar 1 Tahapan Penelitian

Gambar 1 diatas menunjukkan tahapan penelitian analisis sentimen pengguna dompet digital yang dimulai dari proses pengumpulan data, preprocessing, pelabelan data, pembobotan kata menggunakan TF-IDF, klasifikasi sentimen dengan algoritma Naive Bayes, hingga evaluasi kinerja model menggunakan metrik evaluasi. Tahapan dalam melakukan penelitian adalah sebagai berikut :

2.1 Pengumpulan Data

Pengumpulan data dilakukan melalui teknik web scraping dengan memanfaatkan library dari google play scraper yang dijalankan dengan menggunakan bahasa pemrograman Python versi 3.9. Library ini memungkinkan proses ekstrak ulasan pengguna dari google play store secara otomatis dengan berbagai parameter yang dapat disesuaikan sesuai kebutuhan peneliti [13]. Data yang diperoleh berupa ulasan pengguna beserta meta data pendukung seperti rating bintang, tanggal publikasi ulasan serta identitas pengguna yang berasal dari empat aplikasi dompet digital yang ada di Indonesia, yaitu DANA, OVO, DOKU dan LinkAja

Proses pengambilan data dilakukan pada periode januari 2025 dengan batas maksimum 5.000 ulasan setiap aplikasi, keseluruhan jumlag data awal yang berhasil dikumpulkan adalah 23412 ulasan pengguna. Setelah proses pengambilan data maka dilakukan tahap berikutnya yaitu tahapan penyaringan data dengan cara menghapus ulasan yang memiliki rating 3. Ini di karena dianggap bersifat netrall dan tidak menunjukkan kecendrungan sentimen yang jelas. Setelah proses ini dilakukan jumlah data akhir yang digunakan dalam penelitian ini sebanyak 18.869 ulasan pengguna.

2.2 Preprocessing Data

Pada Tahap preprocessing data dilakukan untuk meningkatkan kualitas data teks sebelum proses klasifikasi. Proses ini meliputi *case folding* yang diformalkan sebagai $d' = \text{lower}(d)$, $d' = \text{lower}(d)$, tokenizing untuk memecah teks menjadi token $T = \{t_1, t_2, \dots, t_n\}$, $T = \{t_1, t_2, \dots, t_n\}$, *stopword removal* dengan menghilangkan kata yang tidak bermakna signifikan $T' = T - ST' = T - ST' = T - S$, serta stemming untuk mengubah kata berimbuhan menjadi kata dasar $ti' = \text{stem}(ti)$, $ti' = \text{stem}(ti)$. Tahapan ini bertujuan untuk mengurangi noise dan meningkatkan performa model klasifikasi.

2.3 Perlabelan

Perlabelan data merupakan tahap yang penting dalam analisis sentimen pada algoritma Naive Bayes di mana dalam algoritma ini termasuk metode supervised learning yang membutuhkan data latih berlabel sebagai acuan dalam proses klasifikasi. Label yang digunakan umumnya berupa kategori sentimen seperti positif dan negatif yang merepresentasikan opini pengguna terhadap suatu produk atau layanan.

Penentuan kriteria sentimen pada tahap awal adalah menentukan standar klasifikasi sentimen yang akan digunakan. Dalam penelitian ulasan aplikasi, perlabelan didasarkan pada rating bintang, yaitu :

- a. Rating 4 – 5 sentimen positif
- b. Rating 1 - 2 sentimen negatif
- c. Rating 3 netral

Proses perlabelan dapat dilakukan melalui otomatis yang berdasarkan rating atau aturan tertentu, bisa juga dengan manual, dimana dengan cara membaca isi ulasan untuk meningkatkan akurasi label. dan yang terakhir adalah semi otomatis yang merupakan gabungan kedua kombinasi.

Validasi dan pembersihan label merupakan tahapan berikutnya, dimana tahapan ini memastikan konsistensi label. Ulasan yang ambigu tidak relevan atau berpotensi salah label, pada proses ini label dihapus agar kualitas data latih tetap baik. Label sentimen dikonversikan ke bentuk numerik, dimana positif = 1 dan negatif = 0. Transformasi ini mempermudah proses komputasi dalam tahap pelatihan model.

Dataset berlabel digunakan dalam proses pelatihan pada algoritma Naive Bayes untuk menghitung probabilitas prior setiap kelas sentimen, dan menghitung probabilitas kemunculan kata pada masing masing kelas serta untuk membentuk model klasifikasi yang dapat memprediksi setimen ulasan baru

2.4 Pembobotan

Pembobotan kata dilakukan menggunakan metode **Term Frequency–Inverse Document Frequency (TF-IDF)** untuk mengetahui tingkat kepentingan suatu kata dalam dokumen.

a) Term Frequency (TF)

TF menunjukkan seberapa sering suatu kata muncul dalam satu dokumen

$$TF(t, d) = \frac{f(t, d)}{\sum f(t, d)} \quad (1)$$

Dengan $f(t,d)$ = frekuensi kata t pada dokumen d

b) Inverse Document Frequency (IDF)

IDF menunjukkan seberapa penting suatu kata dalam keseluruhan dokumen.

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

Dengan

N = jumlah seluruh dokumen

df(t) = jumlah dokumen yang mengandung kata t

c) TF-IDF

Nilai TF-IDF diperoleh dari perkalian TF dan IDF.

$$TF - IDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

2.5 Algoritma Naive Bayes

Algoritma **Naive Bayes** digunakan untuk mengklasifikasikan sentimen berdasarkan probabilitas kemunculan kata pada setiap kelas sentimen.

$$P(C|X) = \frac{P(X|C) P(C)}{P(X)} \quad (4)$$

dengan:

C = kelas sentimen (positif atau negatif)

X = dokumen teks

P(C|X) = probabilitas dokumen X termasuk kelas C

Evaluasi model dilakukan menggunakan **confusion matrix** dan metrik evaluasi berikut.

a. Akurasi

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

b. Presisi

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

c. Recall

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

d. F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

Dimana

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

2.6 Evaluasi Kinerja

Evaluasi kinerja model dilakukan menggunakan beberapa metrik, yaitu *accuracy* untuk mengukur proporsi prediksi yang benar terhadap keseluruhan prediksi, *precision* untuk mengukur ketepatan prediksi kelas positif, *recall* untuk mengukur kemampuan model dalam mendeteksi data positif, serta *F1-score* yang merupakan nilai rata-rata harmonik antara *precision* dan *recall*. Pembagian dataset dilakukan menggunakan metode *stratified split* dengan rasio 80:20 guna memastikan proporsi kelas yang seimbang pada data pelatihan dan pengujian

3. HASIL DAN PEMBAHASAN

Data ulasan berhasil dikumpulkan dari empat aplikasi dompet digital (*e-wallet*) dengan jumlah review 4.526 review yang menjadi objek penelitian. Tabel 1 menyajikan distribusi jumlah data untuk setiap aplikasi setelah melalui tahap *preprocessing* serta penghapusan data dengan sentimen netral (rating 3). Hasil distribusi menunjukkan bahwa sentimen negatif mendominasi dengan persentase 57,5%, sedangkan sentimen positif sebesar 42,5%.

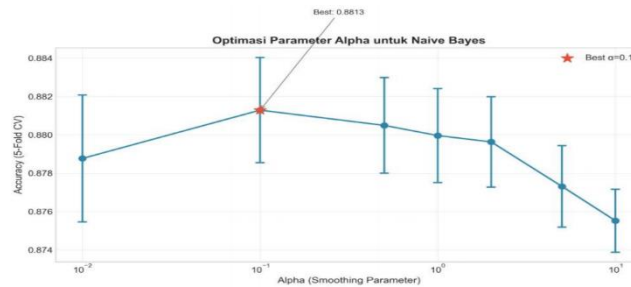
Tabel 1. Distribusi Data

Aplikasi	Positif	Negatif	% Positif	% Negatif
DANA	1.529	2.997	33.8	66.2
OVO	622	4.083	13.2	86.8
DOKU	2.547	2.262	53	47
LinkAja	3.314	1.515	68.8	31.4
TOTAL	8.012	10.857	42.5	57.5

Analisis lebih mendalam pada tingkat aplikasi individual mengungkapkan variasi yang sangat signifikan dalam distribusi sentimen antar aplikasi dompet digital. OVO menunjukkan proporsi sentimen negatif tertinggi yaitu mencapai 86,78% (4.083 ulasan negatif dari 4.705 total) dengan rata-rata rating hanya 1,59 dari skala 5, mengindikasikan tingkat ketidakpuasan pengguna yang sangat tinggi. Sebaliknya, LinkAja menunjukkan performa persepsi pengguna yang jauh lebih baik dengan proporsi sentimen positif tertinggi mencapai 68,63% (3.314 ulasan positif) dan rata-rata rating 3,76. DANA menunjukkan proporsi sentimen negatif 66,22% (2.997 ulasan) dengan rata-rata rating 2,40, sementara DOKU memiliki distribusi yang relatif seimbang dengan 47,04% sentimen negatif (2.262 ulasan) dan rata-rata rating 3,15.

Sebaliknya, nilai alpha yang lebih besar (>1.0) menyebabkan penurunan performa secara gradual hingga mencapai sekitar 87,55% pada $\alpha=10.0$.

Parameter alpha pada algoritma Multinomial Naive Bayes berfungsi sebagai Laplace smoothing factor yang sangat penting untuk menangani masalah zero probability. Nilai alpha yang optimal (0.1) memberikan keseimbangan yang baik antara model yang terlalu fit terhadap training data (alpha terlalu kecil) dan model yang terlalu general sehingga kehilangan informasi discriminative (alpha terlalu besar). Temuan ini konsisten dengan rekomendasi dari penelitian terdahulu yang merekomendasikan nilai alpha dalam rentang 0.1 hingga 1.0 untuk tugas klasifikasi teks.



Gambar 4. Optimasi Parameter

3.1. Klasifikasi

Pada table 2 menunjukkan hasil dari klasifikasi yang dilakukan menggunakan algoritma Naïve Baiyes dengan $\alpha=0,1$ pada data testing adalah

Tabel 2. Hasil Clasifikasi

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,8573	0,9600	0,9058	2.172
Positif	0,9352	0,7831	0,8523	1.602
Accuracy			0,8850	3.774

3.2. Hasil 10-Fold Cross Validation

Hasil cross validation menunjukkan algoritma Naïve Bayes konsisten dan stabilitas model sangat baik

Tabel 3. Cross Validation

Matrik	NB Mean ± STD
Accuracy	0.8808 ± 0,0070
Precision	0.8905 ± 0,0065
Recall	0.8808 ± 0,0070
F1-Score	0.8792 ± 0,0074

4. KESIMPULAN

Hasil penelitian, algoritma Naive Bayes dengan parameter optimal $\alpha = 0,1$ menunjukkan performa yang lebih baik dalam klasifikasi sentimen ulasan pengguna. Hasil analisis distribusi sentimen memperlihatkan bahwa sentimen negatif mendominasi sebesar 57,5%. Di antara aplikasi yang dianalisis, OVO memiliki persentase sentimen negatif tertinggi, yaitu sebesar 86,8%, sedangkan LinkAja menunjukkan persentase sentimen positif tertinggi sebesar 68,6%.

Pengembangan penelitian selanjutnya, model klasifikasi dapat ditingkatkan dengan menerapkan pendekatan *deep learning* seperti IndoBERT maupun metode *Aspect-Based Sentiment Analysis* guna memperoleh pemahaman sentimen yang lebih mendalam. Dari sisi *computational cost*, algoritma Naive Bayes menunjukkan keunggulan yang signifikan dengan waktu prediksi yang sekitar 129 kali lebih cepat dibandingkan model pembanding (0,0105 s vs 1,3571 s), sehingga sangat sesuai untuk diimplementasikan pada aplikasi berbasis *real-time*.

5. REFERENCES

- [1] Bank Indonesia, 2024, "Statistik Sistem Pembayaran dan Infrastruktur Pasar Keuangan," Laporan Tahunan Bank Indonesia
- [2] T. Dahlberg, J. Guo, and J. Ondrus, "A critical review of mobile payment research," *Electronic Commerce Research and Applications*, vol. 14, no. 5, pp. 265–284, 2015.
- [3] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, 2012.
- [4] A. McCallum and K. Nigam, "A comparison of event models for Naive Bayes text classification," *Proc. AAAI Workshop Learning for Text Categorization*, 1998.
- [5] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
- [6] A. Hidayat dan S. Rahmawati, 2023, "Implementasi Machine Learning untuk Analisis Sentimen Media Sosial," *Jurnal Teknik Informatika*, vol. 18, no. 2, pp. 145-158
- [7] A. Pratama dan B. Wijaya, 2023, "Analisis Sentimen Ulasan Aplikasi Mobile Banking Menggunakan Naive Bayes," *Jurnal Teknologi Informasi*, vol. 15, no. 2, pp. 112-125
- [8] A. F. Hidayatullah and M. R. Ma'arif, "Analysis of Indonesian text preprocessing techniques for sentiment classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 512–520, 2023
- [9] A. Alsaeedi and M. Z. Khan, "A study on sentiment analysis techniques of Twitter data," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 1–10, 2022.
- [10] D. P. Putra and L. E. Nugroho, "Sentiment analysis using TF-IDF and machine learning classifiers," *Journal of Big Data*, vol. 10, no. 1, pp. 1–18, 2023.
- [11] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC and informedness," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2022.
- [12] F. Koto, E. Rahmanningtyas, and J. H. Lau, "IndoBERTweet: A pretrained language model for Indonesian Twitter sentiment analysis," in *Proc. ACL*, Toronto, Canada, 2023, pp. 1–12.
- [13] H. Setiawan dan I. Kurniawan, 2021, "Sentiment Labeling Strategies for Indonesian Social Media Data," *Procedia Computer Science*, vol. 195, pp. 234-243
- [14] J. Hermawan et al., 2022, "Feature Extraction Methods for Indonesian Text Classification," *Journal of Information Technology*, vol. 14, no. 2, pp. 89-102
- [15] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," *Information*, vol. 13, no. 2, pp. 1–68, 2022.
- [16] R. A. Pratama and Suyanto, "Performance comparison of Naive Bayes and support vector machine for sentiment analysis," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 21, no. 4, pp. 1021–1029, 2023.
- [17] R. Saputra and A. Nugraha, "Sentiment analysis of Indonesian e-wallet services using Naive Bayes classifier," *Journal of Information Systems Engineering and Business Intelligence*, vol. 10, no. 2, pp. 85–95, 2024.
- [18] K. Santoso dan L. Putri, 2023, "Optimasi Algoritma Naive Bayes untuk Klasifikasi Teks Bahasa Indonesia," *Jurnal Informatika*, vol. 19, no. 3, pp. 178-191.
- [19] Y. A. Sari and A. Wibowo, "Text preprocessing impact on sentiment analysis using machine learning," *Journal of Information and Communication Technology*, vol. 21, no. 3, pp. 295–310, 2022.
- [20] T. Joachims, "Text categorization with machine learning: Learning with many relevant features," *Machine Learning: ECML*, pp. 137–142, updated survey reference, 2022.