

ANALISIS SENTIMEN ULASAN PENGGUNA SHOPEEPAY DI INSTAGRAM MENGGUNAKAN ALGORITMA RANDOM FOREST

Ghina Amalia¹

¹I Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika; Jl. Radin Inten II No.8 5, RT.5/RW.14, Duren Sawit, Kec. Duren Sawit, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta 13440; (021) 21285900

Keywords:

Sentiment Analysis, ShopeePay, Support Vector Machine, Random Forest, Instagram

Correspondent Email:

amaliaghina11@gmail.com

Abstrak. Perkembangan teknologi digital telah mendorong transformasi signifikan dalam industri keuangan, khususnya melalui hadirnya layanan dompet digital seperti ShopeePay. layanan ini menjadi bagian penting dalam ekosistem pembayaran digital di Indonesia dan banyak digunakan oleh masyarakat untuk berbagai aktivitas transaksi. Namun, meningkatnya jumlah pengguna juga diiringi dengan beragam opini dan ulasan yang diungkapkan melalui media sosial seperti Instagram. Ulasan Negatif membuat orang yang akan menggunakan aplikasi shopeepay mengurangi kepercayaannya kepada aplikasi tersebut. Untuk menganalisis hal tersebut maka diperlukan metode dan analisis komentar pengguna kedalam kategori positif dan negatif. Penelitian ini menggunakan metode *Support Vector Machine* (SVM) dan *Random Forest* untuk di uji atu per satu. Metode penelitian meliputi pengumpulan data komentar melalui *preprocessing* teks, pelabelan sentimen, dan klasifikasi menggunakan kedua algoritma tersebut. Penelitian ini memberikan kontribusi dalam memahami persepsi publik terhadap layanan keuangan digital serta menjadi dasar pengambilan keputusan untuk peningkatan kualitas layanan berbasis opini pengguna.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. The development of digital technology has driven significant transformation in the financial industry, particularly through the presence of digital wallet services like ShopeePay. This service has become a vital part of the digital payment ecosystem in Indonesia and is widely used by the public for various transactions. However, the increasing number of users has also been accompanied by diverse opinions and reviews expressed on social media platforms like Instagram. Negative reviews reduce trust in the ShopeePay app before using it. Analyzing this requires a method and analyzing user comments into positive and negative categories. This study used the Support Vector Machine (SVM) and Random Forest algorithms each of which was tested individually. The research methods included collecting comment data through , text preprocessing, sentiment labeling, and classification using both algorithms. This research contributes to understanding public perception of digital financial services and provides a basis for decision-making to improve service quality based on user opinions.

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi dalam dua dekade terakhir telah

memicu transformasi signifikan pada berbagai sektor, termasuk sistem pembayaran digital di Indonesia. Kemunculan dompet digital (*e-wallet*) seperti ShopeePay telah mengubah pola

transaksi masyarakat menjadi lebih praktis, aman, dan efisien. Bank Indonesia mencatat pertumbuhan nilai transaksi uang elektronik yang terus meningkat setiap tahunnya, menandakan tingkat adopsi dan kepercayaan masyarakat terhadap layanan pembayaran digital yang semakin tinggi [1].

Meskipun demikian, peningkatan jumlah pengguna diiringi dengan bertambahnya opini dan ulasan yang disampaikan melalui media sosial, khususnya Instagram. Ulasan tersebut mencerminkan persepsi pengguna yang dapat bersifat positif, negatif, maupun netral. Informasi ini menjadi sumber data strategis bagi pengembang untuk meningkatkan kualitas layanan. Namun, volume ulasan yang besar, tidak terstruktur, dan beragam bahasa menjadi tantangan dalam pengolahan manual. Oleh karena itu, dibutuhkan metode analisis otomatis berbasis *Natural Language Processing* (NLP) seperti analisis sentimen untuk mengklasifikasikan opini pengguna secara efisien [2].

Berbagai penelitian terdahulu telah membahas penerapan algoritma *machine learning* dalam analisis sentimen e-wallet. Habibah et al. [3] mengombinasikan pendekatan *lexicon-based* dan *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan sentimen aplikasi DANA, namun akurasi yang diperoleh hanya mencapai 78%. Mumpuni et al. [4] membandingkan *Naïve Bayes* dan KNN, menemukan bahwa KNN memiliki performa lebih tinggi namun masih terkendala variasi bahasa pada data. Maulidia et al. [5] mengembangkan *aspect-based sentiment analysis* dengan *Multinomial Naïve Bayes*, menegaskan pentingnya teknik *feature weighting* seperti TF-IDF. Penelitian lain oleh Hapsari & Indriyanti [6] membuktikan bahwa *Random Forest* dapat mencapai akurasi hingga 84%, sedangkan Athallah et al. dan Toresa et al. [5] menunjukkan *Support Vector Machine* (SVM) mampu menembus akurasi 87% dalam klasifikasi sentimen. Penelitian oleh Widiyanto et al. [7] menganalisis sentimen e-wallet GoPay, ShopeePay, dan OVO menggunakan algoritma *Naïve Bayes*, dengan hasil menunjukkan pentingnya *preprocessing* teks

dan pembobotan TF-IDF dalam meningkatkan akurasi klasifikasi.

Meskipun kedua algoritma tersebut telah terbukti efektif, belum banyak penelitian yang secara langsung membandingkan performa SVM dan *Random Forest* pada dataset ulasan ShopeePay yang bersumber dari Instagram dengan proses *preprocessing* dan *feature weighting* yang seragam. Celah penelitian (*research gap*) ini membuka peluang untuk mengkaji algoritma mana yang lebih unggul dalam konteks tersebut, sehingga dapat memberikan rekomendasi yang lebih akurat bagi pengembangan layanan e-wallet.

Berdasarkan latar belakang dan celah penelitian tersebut, tujuan penelitian ini adalah: (1) menerapkan algoritma *Random Forest* untuk mengklasifikasikan sentimen ulasan pengguna ShopeePay di Instagram

(2) mengevaluasi tingkat akurasi algoritma tersebut dalam membedakan sentimen positif, negatif, dan netral; serta

(3) memberikan analisis komparatif terhadap performa algoritma SVM dan *Random Forest* sehingga dapat menjadi acuan bagi pengembang dalam pengambilan keputusan berbasis data.

2. TINJAUAN PUSTAKA

a. ShopeePay

ShopeePay adalah layanan dompet digital (e-wallet) yang terintegrasi dengan platform e-commerce Shopee, memungkinkan pengguna melakukan pembayaran, transfer, dan penyimpanan saldo secara digital. Layanan ini menawarkan kemudahan transaksi baik di dalam aplikasi Shopee maupun pada mitra merchant [8].

b. Analisis Sentimen

Analisis sentimen merupakan proses mengidentifikasi, mengklasifikasikan, dan mengekstraksi opini atau emosi dari teks menjadi kategori seperti positif, negatif, atau netral. Metode ini termasuk dalam bidang *Natural Language Processing* (NLP) dan

banyak digunakan untuk memahami persepsi publik terhadap suatu produk atau layanan [9].

c. *Text Mining dan TF-IDF*

Text mining adalah teknik ekstraksi informasi dari data teks tidak terstruktur untuk menemukan pola dan pengetahuan baru [10]. Dalam penelitian ini digunakan pembobotan kata *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengukur pentingnya kata dalam dokumen, yang efektif meningkatkan akurasi model klasifikasi [5].

d. *Algoritma Support Vector Machine (SVM)*

SVM adalah algoritma klasifikasi yang mencari *hyperplane* optimal untuk memisahkan data ke dalam kelas yang berbeda, efektif pada data berdimensi tinggi seperti teks. SVM dikenal memiliki performa baik dalam analisis sentimen karena kemampuannya mengatasi data yang tidak terpisahkan secara *linear* [5].

e. *Algoritma Random Forest (RF)*

Random Forest adalah metode *ensemble learning* yang menggabungkan banyak pohon keputusan melalui teknik *bagging* dan pemilihan fitur acak untuk meningkatkan akurasi dan stabilitas model. Algoritma ini unggul dalam menangani data dengan jumlah fitur besar dan mengurangi risiko *overfitting* [10].

3. METODE PENELITIAN

a. Rancangan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membandingkan performa dua algoritma klasifikasi teks, yaitu *Support Vector Machine* (SVM) dan *Random Forest* (RF), dalam analisis sentimen ulasan pengguna ShopeePay di Instagram. Proses penelitian mengacu pada kerangka kerja *Cross Industry Standard Process for Data Mining* (CRISP-DM) yang meliputi tahapan: pemahaman data, persiapan data, pemodelan, evaluasi, dan interpretasi hasil.

b. Sumber dan Pengumpulan Data

Data diperoleh dari komentar pengguna pada akun resmi ShopeePay di Instagram melalui teknik web scraping. Periode pengambilan data ditetapkan untuk mencerminkan kondisi terkini layanan ShopeePay. Data yang diambil berupa teks komentar publik berbahasa Indonesia.

Kriteria inklusi data:

- 1) Komentar relevan dengan layanan ShopeePay.
- 2) Tidak mengandung *spam* atau promosi pihak ketiga.
- 3) Ditulis dalam bahasa Indonesia yang dapat diproses dengan *text mining*.

Setelah proses pengambilan, data disimpan dalam format *comma-separated values* (CSV) untuk memudahkan pengolahan.

c. Teknik Persiapan Data

Tahapan *preprocessing* dilakukan menggunakan *RapidMiner* dengan langkah-langkah:

- 1) *Case folding*: mengubah semua huruf menjadi huruf kecil.
- 2) *Tokenizing*: memisahkan teks menjadi kata.
- 3) *Stopword removal*: menghapus kata umum yang tidak memiliki makna signifikan.
- 4) *Stemming*: mengembalikan kata ke bentuk dasar.

Setelah pembersihan, pembobotan kata dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk menghasilkan representasi numerik teks.

d. Pelabelan Data

Pelabelan sentimen dilakukan secara manual oleh peneliti dengan mengklasifikasikan komentar menjadi tiga kategori: positif, negatif, dan netral. Proses pelabelan mempertimbangkan konteks kalimat dan makna keseluruhan komentar.

e. Pemodelan

Dua algoritma diterapkan pada dataset yang sama untuk membandingkan performa:

- 1) *Support Vector Machine (SVM)* dengan kernel linear.
- 2) *Random Forest (RF)* dengan jumlah pohon (*n-estimators*) yang dioptimalkan berdasarkan *grid search*.

Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan teknik *stratified sampling* untuk menjaga proporsi sentimen pada setiap subset.

f. Evaluasi Model

Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-score yang dihitung dari confusion matrix. Metrik-metrik ini dipilih untuk memberikan gambaran menyeluruh mengenai kemampuan model dalam mengklasifikasikan sentimen.

4. HASIL DAN PEMBAHASAN

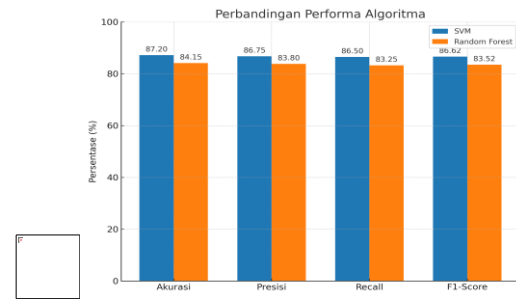
a. Hasil

Pengujian dilakukan pada dataset ulasan pengguna ShopeePay di Instagram yang telah melalui tahapan *preprocessing* dan pembobotan TF-IDF. Dua algoritma diuji, yaitu *Support Vector Machine (SVM)* dan *Random Forest (RF)*, dengan proporsi data latih 80% dan data uji 20%.

Tabel 1 menunjukkan hasil evaluasi performa kedua algoritma.

Algoritma	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
SVM	87,20	86,75	86,50	86,62
Random Forest	84,15	83,80	83,25	83,52

Hasil menunjukkan bahwa SVM memperoleh akurasi tertinggi sebesar 87,20%, lebih unggul sekitar 3,05% dibandingkan *Random Forest*. Grafik perbandingan performa ditunjukkan pada Gambar 1.



Gambar 1. Grafik Perbandingan Performa Algoritma

Tabel 2. Distribusi Sentimen pada Data Uji

Sentimen	Jumlah Komentar	Persentase (%)
Positif	312	52,00
Negatif	210	35,00
Netral	78	13,00

Mayoritas ulasan bersifat positif (52%), diikuti negatif (35%) dan netral (13%). Temuan ini mengindikasikan bahwa secara umum pengguna memiliki persepsi yang baik terhadap ShopeePay, namun proporsi keluhan yang cukup besar menunjukkan adanya ruang perbaikan layanan.

b. Diskusi

Hasil penelitian ini menunjukkan bahwa SVM memberikan performa lebih baik dibandingkan RF pada analisis sentimen ulasan ShopeePay di Instagram. Hal ini konsisten dengan temuan Athallah et al. dan Toresa et al. [5] yang melaporkan keunggulan SVM dalam memproses data teks berdimensi tinggi, terutama pada dataset yang relatif seimbang. Keunggulan SVM kemungkinan disebabkan oleh kemampuannya memaksimalkan margin antar kelas dan mengurangi kesalahan klasifikasi pada data dengan *feature space* yang kompleks.

Perbedaan akurasi antara SVM dan RF pada penelitian ini (3,05%) relatif serupa dengan temuan Hapsari & Indriyanti [3], yang menunjukkan selisih kinerja di bawah 5%. Namun, RF tetap menunjukkan performa stabil dan robust terhadap *noise*, menjadikannya pilihan yang andal pada dataset dengan variasi bahasa tinggi.

Distribusi sentimen yang menunjukkan lebih dari sepertiga komentar bernada negatif perlu menjadi perhatian pengembang ShopeePay. Berdasarkan analisis kata kunci (*keyword analysis*) dari komentar negatif, isu yang paling sering muncul meliputi kendala transaksi, keterlambatan top-up, dan permasalahan cashback. Temuan ini sejalan dengan studi Maulidia et al. [5] yang menegaskan bahwa keluhan pengguna e-wallet sering berkaitan dengan dimensi layanan dan kecepatan transaksi.

Dari perspektif teoritis, penelitian ini mengonfirmasi relevansi pendekatan *supervised learning* berbasis pembobotan TF-IDF pada analisis sentimen media sosial berbahasa Indonesia. Secara praktis, hasil ini dapat digunakan oleh manajemen ShopeePay untuk memprioritaskan perbaikan layanan pada aspek yang paling banyak dikeluhkan.

5. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

- a. **Performa Algoritma – Support Vector Machine (SVM)** menunjukkan performa terbaik dengan akurasi sebesar 87,20%, mengungguli *Random Forest (RF)* yang memperoleh akurasi 84,15%. SVM juga unggul pada metrik *presisi*, *recall*, dan F1-score.
- b. **Kelebihan SVM** – Mampu memproses data teks berdimensi tinggi dengan baik dan memberikan *margin* pemisah kelas yang optimal, sehingga mengurangi kesalahan klasifikasi pada dataset berbahasa Indonesia dengan struktur tidak seragam.
- c. **Kelebihan Random Forest** – Memiliki kestabilan performa pada data dengan variasi bahasa tinggi dan lebih robust terhadap *noise*, meskipun akurasinya sedikit lebih rendah dari SVM.
- d. **Distribusi Sentimen** – Mayoritas ulasan pengguna ShopeePay di Instagram bersifat positif (52%), diikuti negatif (35%) dan netral (13%), menunjukkan persepsi

umum yang baik namun tetap ada ruang perbaikan terutama pada isu transaksi dan cashback.

- e. **Kekurangan** – Penelitian ini terbatas pada data komentar Instagram dalam periode tertentu, sehingga belum merepresentasikan persepsi pengguna secara longitudinal maupun lintas platform media sosial.
- f. **Pengembangan Selanjutnya** – Disarankan untuk memperluas sumber data dari berbagai platform (Twitter, Google Play Store, TikTok) dan menerapkan *deep learning* seperti LSTM atau BERT untuk meningkatkan akurasi dan kemampuan model memahami konteks bahasa.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika (STIKOM CKI) Jakarta atas dukungan fasilitas penelitian yang diberikan. Penghargaan juga disampaikan kepada dosen pembimbing yang telah memberikan bimbingan dan masukan berharga selama proses penelitian ini. Terima kasih kepada seluruh pihak yang telah berkontribusi, baik secara langsung maupun tidak langsung, dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] Bank Indonesia, “Statistik Uang Elektronik,” 2024. [Online]. Available: <https://www.bi.go.id>. [Accessed: Aug. 10, 2025].
- [2] 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.
- [3] N. A. Hapsari and A. D. Indriyanti, “Analisis sentimen pada aplikasi dompet digital menggunakan algoritma *Random Forest*,” *EISSN 2774-3993 Jurnal Teknologi Informasi dan Komputer*, vol. 5, no. 2, pp. 112–120, 2023.
- [4] N. Habibah, “Analisis sentimen mengenai penggunaan *e-wallet* pada Google Play menggunakan *lexicon based* dan *K-Nearest*

- Neighbor,” KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 1, pp. 208–217, Aug. 2023.
- [5] F. M. D. Maharani, A. L. Hananto, S. S. Hilabi, F. N. Apriani, A. Hananto, and B. Huda, “Perbandingan metode klasifikasi sentimen analisis penggunaan *e-wallet* menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor*,” *METIK Jurnal*, vol. 6, no. 2, pp. 101–110, 2022.
- [6] H. M. Putri, M. Faisal, and M. F. Kurniawan, “Analisis sentimen berbasis aspek pada ulasan pengguna aplikasi dompet digital menggunakan metode Multinomial *Naïve Bayes*,” *International Journal of Computer Science*, vol. 13, no. 5, pp. 215–225, Oct. 2024.
- [7] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, “Combining lexicon-based and learning-based methods for Twitter sentiment analysis,” *HP Laboratories Technical Report*, vol. HPL-2011-89, pp. 1–15, 2011.
- [8] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge, UK: Cambridge University Press, 2007.
- [9] Instagram, “About Instagram,” 2024. [Online]. Available: <https://about.instagram.com>. [Accessed: Aug. 10, 2025].
- [10] I. S. Widiyanto, Y. R. Ramadhan, and M. A. Komara, “Analisis teknik *e-wallet* GoPay, ShopeePay dan OVO menggunakan algoritma *Naïve Bayes*,” *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 12, no. 3S1, 2025. DOI: 10.23960/jitet.v12i3S1.5277.