

Analisis Sentimen Pengguna Aplikasi Shopee Pada Google Play Store Menggunakan Metode Random Forest

Oleh:

Muhammad Zainal Abidin

Mochammad Alfian Rosid

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

JULI 2025

Pendahuluan

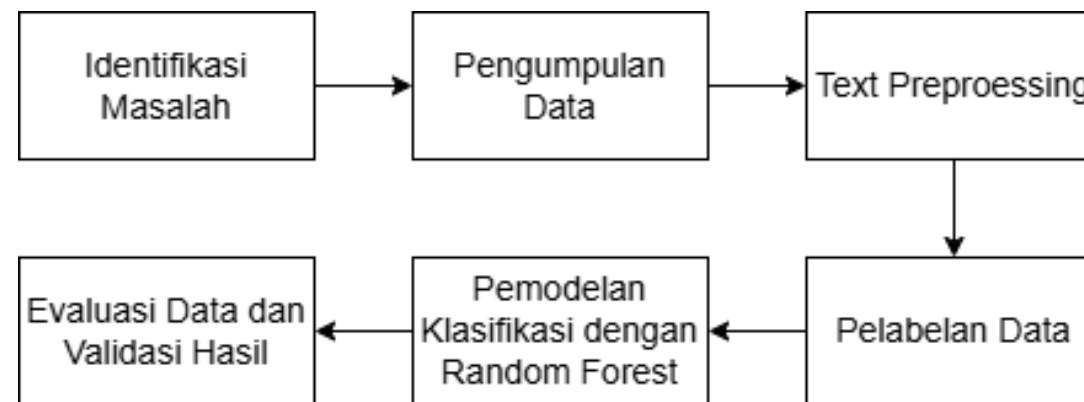
Shopee sebagai aplikasi e-commerce populer memengaruhi perilaku belanja digital. Ulasan pengguna di Google Play mencerminkan kepuasan pelanggan, dipengaruhi fitur, navigasi, dan layanan. Analisis sentimen diperlukan untuk memahami opini publik melalui ulasan tersebut. Penelitian ini menggunakan algoritma Random Forest karena efektivitasnya dalam klasifikasi teks dan penanganan data besar. Studi sebelumnya membuktikan keunggulan Random Forest dalam analisis opini, namun minim perbandingan metode. Penelitian ini bertujuan menganalisis 5.000 ulasan Shopee (Des 2024–Mar 2025), mengevaluasi akurasi sentimen, serta memberikan masukan strategis untuk pengembangan layanan aplikasi secara lebih responsif dan tepat sasaran.

Pertanyaan Penelitian (Rumusan Masalah)

1. Bagaimana menganalisis sentimen pengguna yang diungkapkan dalam ulasan aplikasi Shopee di Google Play Store dari Desember 2024 hingga Maret 2025 sebanyak 5.000 data ulasan dengan Evaluasi yang digunakan yaitu akurasi, recall, presisi dan f1-score
2. Bagaimana menilai dinamika sentimen masyarakat terhadap aplikasi e-commerce ini.

Metode

Penelitian ini akan mengikuti serangkaian langkah sistematis untuk menganalisis sentimen ulasan pengguna aplikasi Shopee di Google Play Store menggunakan metode Random Forest. Tahapan yang dilakukan yaitu identifikasi masalah, pengumpulan data, text preprocessing, pelabelan data, pemodelan klasifikasi dengan Random Forest, dan evaluasi data dan hasil validasi. Diagram alir berikut memberikan gambaran umum prosesnya:



Metode

- **Pengumpulan data** = pada penelitian ini diimplementasikan dengan menggunakan bahasa pemrograman Python dalam lingkungan Google Colab. Metode crawling yang digunakan tidak memanfaatkan API, namun memanfaatkan library Scrapy untuk mengumpulkan hingga 5.000 ulasan pengguna.
- Text Preprocessing = Text Preprocessing dilakukan untuk membentuk kumpulan data yang siap untuk dianalisis. Pada penelitian ini, dilakukan langkah preprocessing dengan empat langkah, yaitu Data Cleansing, Tokenizing, Stopwords Remove dan Stemming.
- Pelabelan data = Ada dua metode untuk pelabelan data yaitu secara manual dan sentiment menggunakan Bahasa pemrograman python.
- Klasifikasi data = Random Forest diukur untuk optimalisasi analisis sentimen ulasan Shopee. Metode klasifikasi supervised learning ini membangun banyak pohon keputusan dari subset acak data dan fitur. Setiap pohon berkontribusi pada hasil klasifikasi akhir melalui mayoritas suara

Hasil

Dari 5.000 ulasan, ditemukan proporsi tertentu:

- Positif sebanyak 1473 data (29,5%)
- Negatif sebanyak 2546 data (50,9%)
- Netral: 981 data (19,6%)

Ini menunjukkan persepsi umum masyarakat terhadap aplikasi Shopee selama periode yang telah ditentukan

Terdapat **dinamika sentimen**: misalnya, peningkatan sentimen negatif pada Januari 2025 karena isu layanan pengiriman, dan peningkatan sentimen positif pada Maret 2025 karena promo Ramadhan. Hal ini penting untuk melihat **respon masyarakat terhadap kebijakan atau fitur baru** dari Shopee.

Setelah pelatihan model Random Forest, didapatkan hasil evaluasi seperti:

- Akurasi: 89%
- Presisi: 88%
- Recall: 88%
- F1-score: 88%

Ini menunjukkan seberapa baik model dapat mengklasifikasikan sentimen ulasan secara otomatis.

Pembahasan

Dari hasil penelitian ini, dapat disimpulkan bahwa metode klasifikasi Random Forest mampu menghasilkan nilai akurasi yang cukup tinggi, yaitu sebesar 88%. Meskipun begitu, akurasi ini masih bisa ditingkatkan lebih lanjut dengan menggunakan data yang lebih baik atau lebih bersih. Peningkatan akurasi maksimum juga dapat dicapai dengan menambahkan lebih banyak data ke dalam set pelatihan atau dengan mencoba metode klasifikasi lain. Data yang digunakan dalam penelitian ini terdiri dari 5.000 komentar yang dikumpulkan dari aplikasi Shopee di Play Store.

Manfaat Penelitian

Penelitian ini memberikan manfaat yang signifikan dalam berbagai aspek. Bagi perusahaan seperti Shopee, hasil analisis sentimen berbasis ulasan pengguna dapat menjadi sumber wawasan berharga untuk memahami persepsi pelanggan secara langsung. Temuan ini dapat digunakan untuk memperbaiki layanan, meningkatkan fitur aplikasi, serta menyusun strategi promosi dan manajemen reputasi yang lebih efektif. Bagi pengembang aplikasi, penelitian ini membantu mengidentifikasi masalah teknis yang sering dikeluhkan oleh pengguna, sehingga mereka dapat menentukan prioritas dalam proses pengembangan dan pemeliharaan aplikasi. Selain itu, bagi kalangan akademisi atau peneliti di bidang teknologi informasi, penelitian ini menjadi contoh penerapan nyata dari teknologi *Natural Language Processing* (NLP) dalam menganalisis opini masyarakat terhadap layanan digital, khususnya dalam konteks bahasa Indonesia. Penelitian ini juga berguna bagi konsumen dan regulator, karena mampu memberikan gambaran umum tentang tingkat kepuasan masyarakat terhadap aplikasi e-commerce, sehingga dapat menjadi bahan pertimbangan dalam pengambilan keputusan. Terakhir, dari sisi pengembangan teknologi kecerdasan buatan, penelitian ini dapat digunakan untuk mengevaluasi performa algoritma analisis sentimen dan memperkuat penerapan NLP dalam bahasa lokal yang masih relatif kurang dieksplorasi dibandingkan bahasa global seperti Inggris.

Referensi

- Rizky, I. and Mahfudz, A. (2022) 'Sebagai Variabel Intervening (Studi Pada Pengguna Shopee Di Kota Semarang)', *Diponegoro J. Manag.*, 11(1), pp. 1–13. Tersedia pada: <https://ejournal3.undip.ac.id/index.php/djom/index>.
- Putra, R.F. and Sumarno (2022) 'Aplikasi Sistem Pakar Perencanaan Investasi Pasar Modal Menggunakan Metode Forward Chaining Berbasis Website', *Procedia Eng. Life Sci.*, 3(December).
- Arsi, P., Wahyudi, R. and Waluyo, R. (2021) 'Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia', *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, 5(2), pp. 231–237. doi: 10.29207/resti.v5i2.2698.
- Amaliah, S., Nusrang, M. and Aswi, A. (2022) 'Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng', *VARIANSI J. Stat. Its Appl. Teach. Res.*, 4(3), pp. 121–127. doi: 10.35580/variansium31.
- Gifari, O.I. et al. (2022) 'Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine', *J. Inf. Technol.*, 2(1), pp. 36–40. doi: 10.46229/jifotech.v2i1.330.
- Mahmuda, S. (2024) 'Implementasi Metode Random Forest pada Kategori Konten Kanal Youtube', *J. Jendela Mat.*, 2(01), pp. 21–31. doi: 10.57008/jjm.v2i01.633.
- Vonega, D.A., Fadila, A. and Kurniawan, D.E. (2022) 'Analisis Sentimen Twitter Terhadap Opini Publik Atas Isu Pencalonan Puan Maharani dalam PILPRES 2024', *J. Appl. Informatics Comput.*, 6(2), pp. 129–135. doi: 10.30871/jaic.v6i2.4300.
- Wardani, N.S., Prahutama, A. and Kartikasari, P. (2020) 'Analisis Sentimen Pemindahan Ibu Kota Negara Dengan Klasifikasi Naïve Bayes Untuk Model Bernoulli Dan Multinomial', *J. Gaussian*, 9(3), pp. 237–246. doi: 10.14710/j.gauss.v9i3.27963.
- Kurniawan, I. and Susanto, A. (2019) 'Implementasi Metode K-Means dan Naïve Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019', *Eksplora Inform.*, 9(1), pp. 1–10. doi: 10.30864/eksplora.v9i1.237.
- Alvanof, M.M. and Dinata, R.K. (2024) 'Penerapan Algoritma Random Forest dalam Deteksi dan Klasifikasi Ransomware', 5(2).

Referensi

- Siregar, A.P. et al. (2023) 'Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke', J. Penelit. Rumpun Ilmu Tek., 2(4), pp. 155–164. doi: 10.55606/juprit.v2i4.3039.
- Erkamim, M. et al. (2023) 'Komparasi Algoritme Random Forest dan XGBoosting dalam Klasifikasi Performa UMKM', J. Sist. Inf. Bisnis, 13(2), pp. 127–134. doi: 10.21456/vol13iss2pp127-134.
- Puspitasari, R., Findawati, Y. and Rosid, M.A. (2023) 'Sentiment Analysis of Post-Covid-19 Inflation Based on Twitter Using the K-Nearest Neighbor and Support Vector Machine Classification Methods', J. Tek. Inform., 4(4), pp. 669–679. doi: 10.52436/1.jutif.2023.4.4.801.
- Fauzan, A. et al. (2021) 'Pengembangan Aplikasi Virtual Tour sebagai Media Pengenalan Lingkungan Kampus PENS berbasis Website', J. Teknol. Terpadu, 7(1), pp. 23–30. doi: 10.54914/jtt.v7i1.341.
- Wahyuningtias, P. et al. (2022) 'COMPARISON OF RANDOM FOREST AND SUPPORT VECTOR MACHINE METHODS ON TWITTER SENTIMENT ANALYSIS (CASE STUDY : INTERNET SELEBGRAM RACHEL VENNYA ESCAPE FROM QUARANTINE) PERBANDINGAN METODE RANDOM FOREST DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TWITT', J. Tek. Inform., 3(1), pp. 141–145.
- Mukarramah, R., Atmajaya, D. and Ilmawan, L.B. (2021) 'Performance comparison of support vector machine (SVM) with linear kernel and polynomial kernel for multiclass sentiment analysis on twitter', Ilk. J. Ilm., 13(2), pp. 168–174. doi: 10.33096/ilkom.v13i2.851.168-174.
- Warjiyono et al. (2024) 'Analisa Prediksi Harga Jual Rumah Menggunakan Algoritma Random Forest Machine Learning', JURSIKSTEKNI (Jurnal Sist. Inf. dan Teknol. Informasi), 6(2), pp. 416–423.
- Aditya, M.F.R., Lutvi, N. and Indahyanti, U. (2024) 'Prediksi Penyakit Hipertensi Menggunakan Metode Decison Tree dan Random Forest', J. Ilm. Komputasi, 23(1), pp. 9–16. doi: 10.32409/jikstik.23.1.3503.
- Hidayat, H., Sunyoto, A. and Al Fatta, H. (2023) 'Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier', J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan), 7(1), pp. 31–40. doi: 10.47970/siskom-kb.v7i1.464.
- Bin, A., Alkatiri, M. and Nasution, A.N.S. (2020) 'OPINI PUBLIK TERHADAP PENERAPAN NEW NORMAL DI MEDIA SOSIAL TWITTER', J. Strateg. Commun., 11(1).

Referensi

- Dikiyanti, T.D., Rukmi, A.M. and Irawan, M.I. (2021) 'Sentiment analysis and topic modeling of BPJS Kesehatan based on twitter crawling data using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation algorithm', J. Phys. Conf. Ser., 1821(1). doi: 10.1088/1742-6596/1821/1/012054.
- Ardhani, B.A., Chamidah, N. and Saifudin, T. (2021) 'Sentiment Analysis Towards Kartu Prakerja Using Text Mining with Support Vector Machine and Radial Basis Function Kernel', J. Inf. Syst. Eng. Bus. Intell., 7(2), p. 119. doi: 10.20473/jisebi.7.2.119-128.
- Khairunnisa, S., Adiwijaya, A. and Al Faraby, S. (2021) 'Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)', J. Media Inform. Budidarma, 5(2), p. 406. doi: 10.30865/mib.v5i2.2835.
- Dwiki, A. et al. (2021) 'Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN', J. Tek. Inform. dan Sist. Inf., 8(2), pp. 636–646.
- Hidayati, N., Suntoro, J. and Setiaji, G.G. (2021) 'Perbandingan Algoritma Klasifikasi untuk Prediksi Cacat Software dengan Pendekatan CRISP-DM', J. Sains dan Inform., 7(2), pp. 117–126. doi: 10.34128/jsi.v7i2.313.
- Afdal, M. and Elita, L.R. (2022) 'Penerapan Text Mining Pada Aplikasi Tokopedia Menggunakan Algoritma K-Nearest Neighbor', J. Ilm. Rekayasa dan Manaj. Sist. Inf., 8(1), pp. 78–87.
- Laurensz, B. and Sedyono, E. (2021) 'Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19', J. Nas. Tek. Elektro dan Teknol. Inf., 10(2), pp. 118–123. doi: 10.22146/jnteti.v10i2.1421.
- Pamungkas, B., Purbaya, M.E. and K, D.J.A. (2021) 'Analisis Sentimen Twitter Menggunakan Metode Support Vector Machine (SVM) pada', J. Informatics, Inf. Syst. Softw. Eng. Appl., 3(2), pp. 10–20.
- Kurniasih, U. and Suseno, A.T. (2022) 'Analisis Sentimen Terhadap Bantuan Subsidi Upah (BSU) pada Kenaikan Harga Bahan Bakar Minyak (BBM)', J. Media Inform. Budidarma, 6(4), pp. 2335–2340. doi: 10.30865/mib.v6i4.4958.
- Herdiyani, T.C. and Zailani, A.U. (2022) 'Sentiment Analysis Terkait Pemindahan Ibu Kota Indonesia Menggunakan Metode Random Forest Berdasarkan Tweet Warga Negara Indonesia', J. Teknol. Sist. Inf., 3(2), pp. 154–165. doi: 10.35957/jtsi.v3i2.2920.
- Rania, N.Z. and Syah, R.D. (2024) 'Analisis Sentimen Terhadap Aplikasi Gojek Pada Play Store Menggunakan Metode Random Forest Classifier', J. Ilm. Inform. Komput., 29(2), pp. 144–153. doi: 10.35760/ik.2024.v29i2.11877.
- Alita, D. and Isnain, A.R. (2020) 'Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier', J. Komputasi, 8(2), pp. 50–58. doi: 10.23960/komputasi.v8i2.2615.

