

Analisis Sentimen Pengguna Aplikasi Investasi Saham Ajaib, IPOT, dan Stockbit Menggunakan Machine Learning

Muhammad Vikri Mustafa¹, Dodi Candra Kurniawan², Amanda Putri Cheryl³,
Faris Widhiarta⁴, Muhammad Arifin⁵

^{1,2,3,4,5} Program Studi Sistem Informasi, Universitas Muria Kudus, Indonesia

Received : 7 Mei 2026, Revised : 18 Mei 2026, Published : 26 Mei 2026

Corresponding Author

Nama Penulis: Faris Widhiarta

E-mail: 202353152@std.umk.ac.id

Abstrak

Kajian ini dirancang untuk memetakan dan membandingkan opini pengguna pada tiga platform investasi saham berbasis mobile di Indonesia, yakni Ajaib, IPOT (Indo Premier Online Technology), dan Stockbit, sekaligus mengukur efektivitas sejumlah algoritma machine learning dalam mengotomasi proses klasifikasi sentimen. Rancangan penelitian yang digunakan bersifat kuantitatif komparatif. Sebanyak 15.000 ulasan dikumpulkan secara proporsional—5.000 ulasan per platform—melalui teknik web scraping pada Google Play Store dengan pendekatan purposive sampling. Data mentah kemudian diolah melalui serangkaian tahap pra-pengolahan meliputi normalisasi teks, penghapusan noise, penanganan nilai kosong, serta pembobotan fitur berbasis TF-IDF Vectorizer. Sentimen ulasan ditetapkan sebagai variabel respons, sedangkan jenis algoritma klasifikasi berperan sebagai variabel prediktor. Empat algoritma diuji secara komparatif, yaitu Logistic Regression, Naive Bayes, Support Vector Machine (SGD), dan Random Forest. Logistic Regression mencatat kinerja tertinggi dengan akurasi 87,30%, mengungguli Naive Bayes (86,77%), SVM-SGD (86,23%), dan Random Forest (85,67%). Pendekatan N-gram turut mengungkap keluhan dominan yang khas pada tiap platform: hambatan verifikasi identitas dan perubahan antarmuka di Ajaib, gangguan konektivitas di IPOT, serta kendala sinkronisasi akun keamanan di Stockbit. Secara keseluruhan, temuan ini menegaskan potensi machine learning sebagai fondasi sistem pemantauan sentimen pengguna yang otomatis dan berbasis data secara real-time.

Kata kunci – analisis sentimen, machine learning, aplikasi investasi, klasifikasi teks, ulasan pengguna

Abstract

This study investigates and compares user sentiment across three mobile-based stock investment platforms in Indonesia—Ajaib, IPOT (Indo Premier Online Technology), and Stockbit—while assessing the effectiveness of multiple machine learning algorithms in automating sentiment classification. A quantitative comparative design was employed, with a total sample of 15,000 reviews drawn equally from each platform (5,000 per application) via web scraping on the Google Play Store using purposive sampling. Collected data underwent a multi-stage preprocessing pipeline comprising text normalization, noise removal, missing value handling, tokenization, stopword filtering, stemming, and TF-IDF-based feature extraction. User sentiment served as the response variable, while algorithm type functioned as the predictor variable. Four classifiers were benchmarked: Logistic Regression, Naive Bayes, Support Vector Machine (SGD), and Random Forest, evaluated on accuracy, precision, recall, and F1-score. Logistic Regression delivered the top performance at 87.30% accuracy, outperforming Naive Bayes (86.77%), SVM-SGD (86.23%), and Random Forest (85.67%). N-gram analysis additionally revealed platform-specific user pain points: identity verification bottlenecks and UI inconsistencies in Ajaib, connectivity disruptions in IPOT, and securities account synchronization failures in Stockbit. These findings confirm that

This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license

machine learning offers a scalable, data-driven foundation for real-time sentiment monitoring in digital investment ecosystems.

Keywords - sentiment analysis, machine learning, investment applications, text classification, user reviews

How To Cite : Mustafa, M. V., Kurniawan, D. C., Cheryl, A. P., Widhiarta, F., & Arifin, M. (2026). Analisis Sentimen Pengguna Aplikasi Investasi Saham Ajaib, IPOT, dan Stockbit Menggunakan Machine Learning. *Jurnal Penelitian Multidisiplin Bangsa*, 2(12), 1955–1963. <https://doi.org/10.59837/jpnmb.v2i12.840>

Copyright ©2026 Muhammad Vikri Mustafa, Dodi Candra Kurniawan, Amanda Putri Cheryl, Faris Widhiarta, Muhammad Arifin

PENDAHULUAN

Perkembangan teknologi finansial (*fintech*) di Indonesia telah membuka akses investasi saham kepada masyarakat yang lebih luas, khususnya melalui aplikasi berbasis mobile yang menawarkan kemudahan penggunaan, biaya transaksi yang kompetitif, serta proses registrasi yang serba digital. Kehadiran aplikasi seperti Ajaib, IPOT (*Indo Premier Online Technology*), dan Stockbit mendorong peningkatan jumlah investor ritel, terutama dari kalangan generasi muda yang semakin aktif berpartisipasi di pasar modal. Seiring dengan meningkatnya jumlah pengguna, volume ulasan pada platform distribusi aplikasi seperti Google Play Store juga mengalami peningkatan signifikan. Ulasan tersebut mengandung informasi penting terkait pengalaman pengguna, baik berupa kepuasan maupun keluhan terhadap layanan aplikasi. Namun, besarnya jumlah data ulasan yang bersifat tidak terstruktur (*unstructured data*) menjadi tantangan dalam proses analisis, sehingga diperlukan pendekatan otomatis untuk mengolahnya secara efektif.

Salah satu pendekatan yang terbukti mampu mengotomasi pemahaman teks secara efisien adalah analisis sentimen, sebuah cabang Natural Language Processing (NLP) yang mengekstrak orientasi opini dari data tekstual. Sejumlah studi empiris telah mengonfirmasi kehandalan machine learning dalam domain ini. Dalam konteks perbankan digital, misalnya, SVM terbukti menghasilkan klasifikasi yang presisi pada ulasan berbahasa Indonesia (Puji Astuti et al., 2022). Pada segmen marketplace, Logistic Regression yang dipadukan dengan TF-IDF konsisten mengungguli model lain (Lestari & Hutagalung, 2025). Khusus di ranah aplikasi investasi, SVM dan *Maximum Entropy* telah diterapkan pada ulasan Ajaib (Kavabilla et al., 2023), sementara pendekatan serupa diujicobakan lintas platform saham oleh Lestari dan Saepudin (2021). Yulianto dan Andrianto (2025) memperkuat temuan ini dengan menunjukkan superioritas Logistic Regression dalam perbandingan empat algoritma pada ulasan aplikasi Bibit. Secara kolektif, literatur-literatur ini menegaskan bahwa kombinasi tepat antara teknik representasi fitur dan pemilihan model merupakan kunci keberhasilan klasifikasi sentimen berbasis ulasan digital.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis dan membandingkan sentimen pengguna pada tiga aplikasi investasi saham di Indonesia, yaitu Ajaib, IPOT, dan Stockbit. Pemilihan ketiga aplikasi ini didasarkan pada beberapa pertimbangan: (1) ketiganya termasuk aplikasi investasi saham dengan volume dan frekuensi transaksi investor ritel tertinggi di Bursa Efek Indonesia (Bursa Efek Indonesia, 2024; Otoritas Jasa Keuangan, 2023); (2) ketiganya menawarkan pendekatan layanan yang berbeda satu sama lain, sehingga memungkinkan perbandingan sentimen yang representatif; serta (3) ketiganya memiliki volume ulasan yang memadai untuk mendukung pemodelan machine learning yang valid secara statistik. Penelitian ini juga bertujuan untuk mengevaluasi kinerja beberapa algoritma machine learning dalam melakukan klasifikasi sentimen secara otomatis, serta mengidentifikasi permasalahan utama (*pain points*) yang dialami pengguna berdasarkan analisis ulasan, sehingga hasil penelitian dapat memberikan kontribusi bagi pengembang aplikasi dalam meningkatkan kualitas layanan serta mendukung pengembangan sistem monitoring sentimen berbasis data secara real-time.

TINJAUAN PUSTAKA

Analisis sentimen tergolong dalam bidang *Natural Language Processing* (NLP) yang secara khusus mengkaji orientasi emosional dalam teks—apakah bersifat positif, negatif, atau netral—terhadap suatu entitas produk maupun layanan. Dalam ekosistem layanan digital, ulasan pengguna menjadi sumber data primer yang kaya informasi tentang ekspektasi, kepuasan, dan ketidakpuasan konsumen. Kendati demikian, sifat ulasan yang tidak berstruktur mengharuskan dilakukannya serangkaian tahap pra-pengolahan sebelum data dapat dianalisis, mulai dari normalisasi teks, segmentasi kata (tokenisasi), eliminasi stopwords, hingga penyeteman kata untuk menyeragamkan bentuk morfologisnya. Representasi numerik teks kemudian dilakukan melalui TF-IDF (*Term Frequency–Inverse Document Frequency*)—sebuah skema pembobotan yang mengukur signifikansi suatu token dalam sebuah dokumen secara relatif terhadap seluruh koleksi teks (Septiani & Isabela, 2022).

Machine learning merujuk pada cabang kecerdasan buatan yang membekali sistem dengan kemampuan belajar mandiri dari data guna membuat prediksi atau keputusan tanpa instruksi eksplisit. Dalam konteks klasifikasi teks sentimen, empat algoritma kerap menjadi acuan perbandingan. *Logistic Regression*, meski sederhana secara arsitektur, menunjukkan performa yang tangguh pada data berdimensi tinggi seperti keluaran TF-IDF berkat kestabilan estimasi parameternya. Naive Bayes menawarkan efisiensi komputasi yang menonjol dengan asumsi kondisional independensi antar fitur—suatu penyederhanaan yang jarang berlaku sempurna namun sering menghasilkan kinerja kompetitif. SVM mencari hyperplane pemisah dengan margin optimal, menjadikannya andal pada ruang fitur yang lebar. Adapun *Random Forest*, sebagai teknik ensemble yang membangun hutan pohon keputusan secara paralel, unggul dalam menemukan pola non-linear namun rentan terhadap data yang sangat jarang (*sparse*) seperti *representasi bag-of-words* (Larasati et al., 2022).

Beberapa penelitian terdahulu memperkuat relevansi pendekatan machine learning dalam analisis sentimen. Puji Astuti et al. (2022) menunjukkan bahwa SVM mampu mencapai tingkat akurasi yang tinggi saat diterapkan pada ulasan aplikasi mobile banking di Indonesia. Lestari dan Hutagalung (2025) melaporkan keunggulan *Logistic Regression* yang dikombinasikan dengan TF-IDF dalam mengklasifikasikan sentimen ulasan marketplace berbahasa Indonesia. Di ranah aplikasi investasi, Kavabilla et al. (2023) mengeksplorasi SVM dan *Maximum Entropy* pada ulasan aplikasi Ajaib, sementara Lestari dan Saepudin (2021) menguji SVM pada ulasan lintas platform saham di Google Play Store. Yulianto dan Andrianto (2025) memperluas kajian ini dengan membandingkan empat algoritma supervised learning pada ulasan aplikasi Bibit dan menyimpulkan bahwa *Logistic Regression* konsisten unggul. Normawati dan Prayogi (2021) mengkaji efektivitas Naive Bayes Classifier melalui evaluasi Confusion Matrix pada data Twitter, Tanggraeni dan Sitokdana (2022) mengaplikasikan Naive Bayes pada analisis sentimen ulasan aplikasi e-government di Google Play, sedangkan Maulana et al. (2023) menerapkan pendekatan serupa pada ulasan aplikasi di Google Play Store. Simanjuntak et al. (2023) secara spesifik membandingkan *Logistic Regression* dengan *Random Forest* untuk klasifikasi emosi pada konten media sosial. Keseluruhan temuan tersebut menegaskan bahwa pilihan algoritma serta teknik ekstraksi fitur berkontribusi signifikan terhadap keberhasilan klasifikasi sentimen.

Dalam penelitian ini, keterkaitan antar variabel digambarkan melalui kerangka konsep yang menempatkan data ulasan pengguna sebagai variabel masukan, proses analisis sentimen berbasis machine learning sebagai variabel proses, serta keluaran berupa hasil klasifikasi sentimen dan identifikasi permasalahan pengguna. Data ulasan dari Google Play Store diolah melalui tahap pra-proses, kemudian direpresentasikan secara numerik menggunakan metode TF-IDF dengan pembatasan 2.500 fitur teratas, lalu diklasifikasikan menggunakan beberapa algoritma yang kinerjanya dibandingkan berdasarkan metrik akurasi, precision, recall, dan F1-score.

Berdasarkan kerangka konsep tersebut, hipotesis dalam penelitian ini adalah terdapat perbedaan kinerja antara algoritma *Logistic Regression*, Naive Bayes, Support Vector Machine (SGD), dan *Random Forest* dalam klasifikasi sentimen ulasan pengguna aplikasi investasi saham, serta *Logistic Regression* diduga memiliki kinerja terbaik dalam mengklasifikasikan sentimen berbasis fitur TF-IDF.

METODE

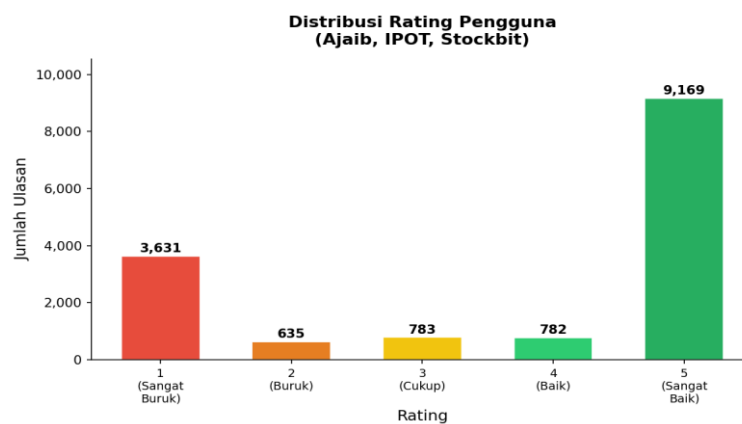
Studi ini dirancang menggunakan pendekatan kuantitatif dengan desain komparatif, berfokus pada pemetaan dan perbandingan sentimen pengguna dari tiga platform investasi saham terkemuka di Indonesia: Ajaib, IPOT (Indo Premier Online Technology), dan Stockbit. Korpus data yang digunakan terdiri atas 15.000 entri ulasan—masing-masing 5.000 ulasan per platform—yang diperoleh dari Google Play Store. Pemilihan sampel dilakukan secara purposif dengan memprioritaskan ulasan terbaru yang relevan secara kontekstual dengan fokus penelitian. Proses akuisisi data dilaksanakan melalui teknik web scraping berbantuan pustaka google-play-scraper pada ekosistem pemrograman Python.

Seluruh data yang berhasil dihimpun kemudian melewati alur pra-pengolahan bertingkat: pembersihan karakter non-alfanumerik dan tanda baca, penanganan entri kosong, normalisasi teks, tokenisasi, penyaringan stopwords, serta stemming untuk menyeragamkan bentuk kata ke akar leksikalnya. Penetapan label sentimen mengacu pada skor rating yang diberikan pengguna secara langsung—rating 4 dan 5 dikodekan sebagai sentimen positif, rating 1 dan 2 sebagai negatif, dan rating 3 sebagai netral. Representasi numerik teks kemudian diperoleh melalui transformasi TF-IDF, menghasilkan matriks fitur yang siap dikonsumsi oleh model machine learning.

Tahap pemodelan melibatkan empat algoritma klasifikasi supervised: Logistic Regression, Naive Bayes, Support Vector Machine berbasis SGD, dan *Random Forest*. Partisi dataset menggunakan rasio 80:20 antara data latih dan data uji melalui stratified sampling, sehingga distribusi kelas terjaga secara proporsional pada kedua subset dan replikabilitas eksperimen terjamin. Kinerja setiap model diukur menggunakan empat metrik evaluasi standar: akurasi, precision, recall, dan F1-score. Di sisi lain, analisis N-gram diterapkan khusus pada segmen ulasan negatif untuk mengekstrak frasa-frasa berulang yang merepresentasikan keluhan dominan pengguna pada masing-masing platform.

PEMBAHASAN

Pola distribusi rating ulasan pada ketiga platform investasi saham—Ajaib, IPOT, dan Stockbit—memperlihatkan kecenderungan bimodal yang kuat, yakni konsentrasi di kutub ekstrem: rating tertinggi (positif) dan rating terendah (negatif), sementara rating tengah (netral) jauh lebih jarang muncul. Fenomena ini lazim dijumpai pada ekosistem ulasan aplikasi digital, di mana pengguna cenderung terdorong untuk memberikan penilaian hanya ketika mengalami kepuasan luar biasa atau kekecewaan mendalam. Yang menarik, lonjakan ulasan negatif pada periode-periode tertentu berkorelasi erat dengan pembaruan versi aplikasi maupun dinamika pasar saham—terutama saat volatilitas harga meningkat tajam sehingga memberikan tekanan ekstra pada infrastruktur sistem. Distribusi rating ketiga aplikasi dapat dilihat secara visual pada Gambar 1 berikut.



Gambar 1.
Distribusi Rating per Aplikasi

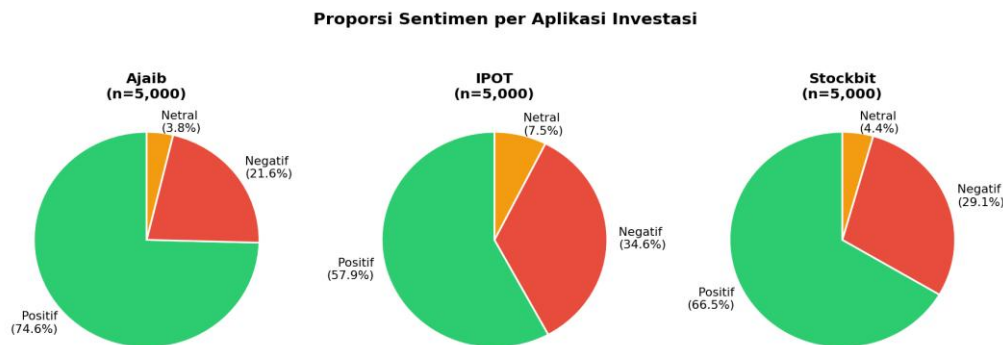
Gambar 1 memperlihatkan distribusi rating gabungan dari ketiga aplikasi (Ajaib, IPOT, dan Stockbit) dengan total 15.000 ulasan. Terlihat pola bimodal yang dominan, di mana rating 5 (Sangat Baik) mendominasi dengan 9.169 ulasan (61,1%), diikuti rating 1 (Sangat Buruk) sebagai kelompok terbesar kedua dengan 3.631 ulasan (24,2%). Sebaliknya, rating 2, 3, dan 4 masing-masing hanya mencapai 635, 783, dan 782 ulasan. Pola ini mencerminkan kecenderungan pengguna untuk memberikan ulasan hanya ketika mengalami kepuasan luar biasa atau kekecewaan mendalam, sementara pengalaman biasa jarang mendorong pengguna untuk menulis ulasan.

Untuk memberikan gambaran distribusi data, berikut disajikan ringkasan kategori sentimen berdasarkan rating pengguna:

Tabel 1.
Distribusi Kategori Sentimen Berdasarkan Rating

Kategori Sentimen	Rentang Rating	Keterangan
Positif	4 – 5	Pengguna puas terhadap layanan
Netral	3	Pengalaman biasa/tidak signifikan
Negatif	1 – 2	Pengguna tidak puas

Perbandingan proporsi sentimen antar aplikasi secara keseluruhan disajikan pada Gambar 2.



Gambar 2.
Proporsi Sentimen per Aplikasi (%)

Gambar 2 menunjukkan perbandingan proporsi sentimen antar aplikasi. Ajaib memiliki sentimen positif tertinggi (74,6%), diikuti Stockbit (66,5%) dan IPOT (57,9%). Sebaliknya, IPOT mencatat proporsi sentimen negatif terbesar (34,6%), yang mengindikasikan adanya tingkat ketidakpuasan pengguna yang lebih tinggi dibandingkan dua platform lainnya. Pola ini konsisten dengan temuan distribusi rating yang menunjukkan Ajaib sebagai platform dengan kepuasan pengguna tertinggi secara keseluruhan.

Tabel 2.
Distribusi Sentimen per Aplikasi

Aplikasi	Positif	Negatif	Netral	Total
Ajaib	3.729 (74,6%)	1.082 (21,6%)	189 (3,8%)	5.000
IPOT	2.896 (57,9%)	1.731 (34,6%)	373 (7,5%)	5.000
Stockbit	3.326 (66,5%)	1.453 (29,1%)	221 (4,4%)	5.000
Total	9.951 (66,3%)	4.266 (28,4%)	783 (5,2%)	15.000

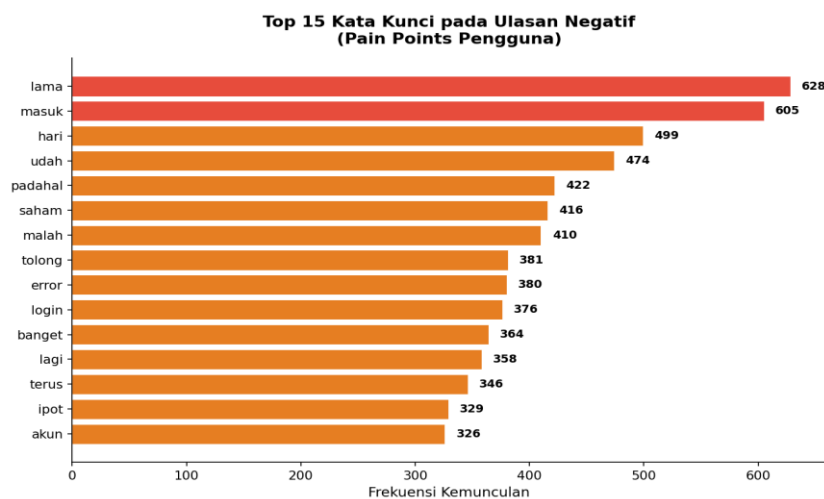
Berdasarkan analisis lebih lanjut menggunakan metode N-gram, diperoleh informasi terkait permasalahan utama (pain points) yang sering muncul dalam ulasan negatif pada masing-masing aplikasi sebagai berikut:

Tabel 3.
Identifikasi Permasalahan Utama Berdasarkan Analisis N-gram

Aplikasi	Permasalahan Utama
Ajaib	Dana masuk rekening lambat, saldo RDN tidak sinkron, proses verifikasi berulang (data diri), perubahan antarmuka membingungkan, respons CS lambat. Bigram dominan: masuk rekening (26), saldo rdn (23), data diri (22).
IPOT	Masalah autentikasi (kode OTP tidak terkirim), log out paksa setelah update, fitur lupa password tidak berfungsi, koneksi terputus saat jam pasar. Bigram dominan: kode otp (44), update malah (31), lupa password (28).
Stockbit	Error saat open market (jam buka perdagangan), sinkronisasi portofolio lambat, fitur komunitas sering error/eror. Bigram dominan: open market (58), sering error/eror (104), error mulu (45).

Hasil ini menunjukkan bahwa masing-masing aplikasi memiliki karakteristik permasalahan yang berbeda, yang dipengaruhi oleh fitur utama dan arsitektur sistem yang digunakan. Permasalahan teknis seperti stabilitas sistem dan integrasi data menjadi faktor dominan yang mempengaruhi kepuasan pengguna.

Visualisasi kata kunci permasalahan utama berdasarkan analisis N-gram pada ulasan negatif ditampilkan pada Gambar 3.



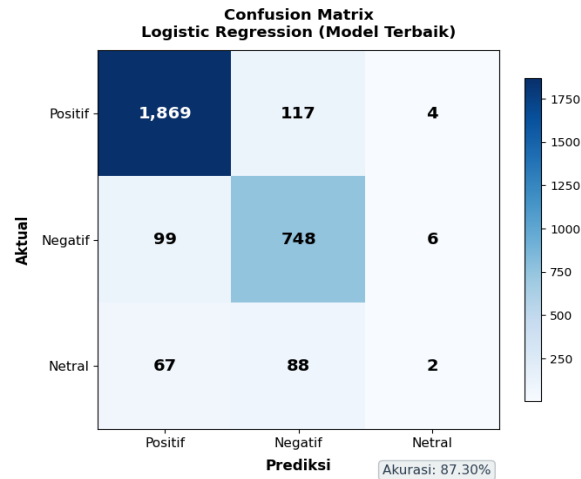
Gambar 3.
Visualisasi Kata Kunci Pain Points per Aplikasi

Selanjutnya, hasil pengujian model machine learning dilakukan untuk mengetahui algoritma terbaik dalam klasifikasi sentimen. Hasil evaluasi disajikan pada tabel berikut:

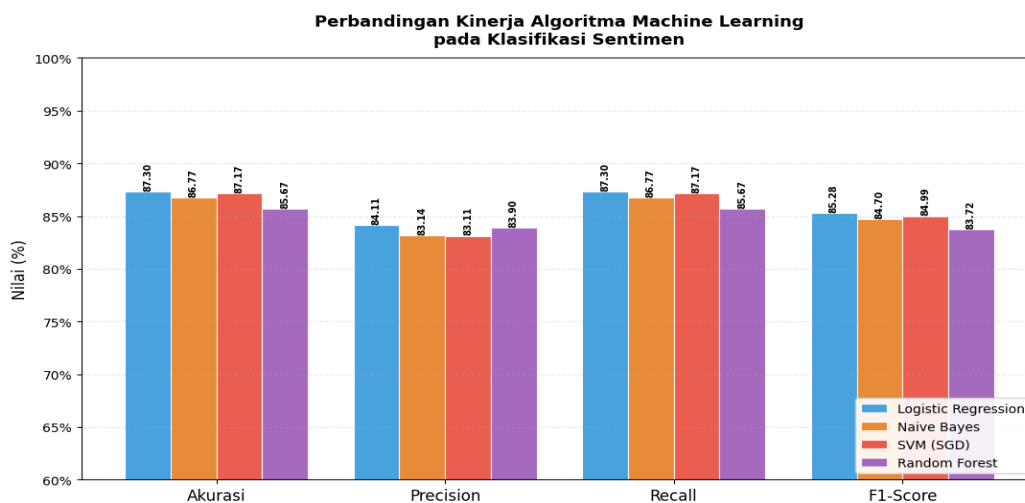
Tabel 4.
Perbandingan Kinerja Algoritma Machine Learning

Algoritma	Akurasi	Precision	Recall	F1-Score
Logistic Regression	87,30%	84,11%	87,30%	85,28%
Naive Bayes	86,77%	83,14%	86,77%	84,70%
SVM (SGD)	86,23%	81,89%	86,23%	83,99%
Random Forest	85,67%	83,90%	85,67%	83,72%

Hasil prediksi model terbaik (*Logistic Regression*) divisualisasikan melalui confusion matrix pada Gambar 4. Dari 3.000 data uji, model Logistic Regression berhasil mengklasifikasikan dengan benar 1.869 dari 1.990 ulasan positif (93,92%), 748 dari 853 ulasan negatif (87,69%), namun hanya 2 dari 157 ulasan netral (1,27%). Rendahnya recall kelas netral menunjukkan bahwa ulasan bernada ambigu sulit diklasifikasikan oleh model berbasis fitur leksikal, terutama akibat ketidakseimbangan distribusi kelas netral yang hanya mencakup 5,22% dari total data. Perbandingan akurasi keempat algoritma selanjutnya disajikan pada Gambar 5.



Gambar 4. Confusion Matrix Prediksi Sentimen



Gambar 5. Perbandingan Akurasi Antar 4 Algoritma

Berdasarkan Tabel 4, algoritma Logistic Regression menunjukkan performa terbaik dengan akurasi sebesar 87,30%. Hal ini disebabkan oleh karakteristik data teks yang telah direpresentasikan menggunakan TF-IDF, di mana distribusi fitur bersifat linier sehingga dapat dipisahkan dengan baik oleh model linear. Selain itu, perbedaan performa antar algoritma relatif kecil, dengan selisih maksimal sekitar 1,63%, yang menunjukkan bahwa kualitas praproses dan ekstraksi fitur memiliki pengaruh yang signifikan terhadap hasil klasifikasi.

Untuk memperjelas perbandingan, berikut disajikan kelebihan dan kelemahan masing-masing algoritma:

Tabel 5.
Analisis Kelebihan dan Kelemahan Algoritma

Algoritma	Kelebihan	Kelemahan
Logistic Regression	Akurat pada data teks, cepat, stabil	Kurang menangani hubungan non-linear
Naive Bayes	Cepat dan efisien	Asumsi independensi fitur kurang realistis
SVM (SGD)	Baik untuk dataset besar	Sensitif terhadap parameter
Random Forest	Menangani non-linearitas	Kurang optimal pada data sparse

Dikaitkan dengan landasan teoritis, capaian Logistic Regression ini selaras dengan argumentasi bahwa fitur TF-IDF menghasilkan ruang vektor yang secara inheren linier terpisah, sehingga model linear memiliki keunggulan alamiah. Hal ini dikuatkan oleh sejumlah studi yang secara konsisten melaporkan superioritas Logistic Regression pada klasifikasi teks berbahasa Indonesia (Lestari & Hutagalung, 2025; Yulianto & Andrianto, 2025; Simanjuntak et al., 2023). Sebaliknya, *Random Forest* yang bertumpu pada partisi rekursif ruang fitur kurang adaptif terhadap data berdimensi sangat tinggi namun sangat jarang seperti matriks TF-IDF.

Selain itu, hasil penelitian ini menunjukkan bahwa analisis sentimen tidak hanya mampu mengklasifikasikan opini pengguna, tetapi juga memberikan wawasan mendalam terkait permasalahan spesifik yang dialami pengguna. Informasi ini sangat penting bagi pengembang aplikasi dalam menentukan prioritas perbaikan sistem, seperti peningkatan stabilitas server, penyederhanaan proses verifikasi, serta optimalisasi integrasi data antar platform.

Secara keseluruhan, studi ini menegaskan bahwa kualitas tahap pra-pengolahan dan strategi representasi fitur memiliki bobot yang jauh lebih determinan terhadap akurasi klasifikasi dibanding sekadar memilih algoritma yang lebih kompleks. Ini berarti investasi pada pipeline data yang solid – bukan semata-mata pada kerumitan model – adalah kunci untuk membangun sistem analisis sentimen yang andal, cepat, dan dapat beroperasi secara otomatis dalam skala besar pada domain ulasan aplikasi digital.

KESIMPULAN

Studi ini berhasil memetakan dan membandingkan sentimen pengguna pada tiga platform investasi saham berbasis mobile di Indonesia – Ajaib, IPOT, dan Stockbit – serta mengukur performa komparatif empat algoritma machine learning dalam klasifikasi sentimen otomatis. Distribusi ulasan memperlihatkan polarisasi yang jelas antara sentimen positif dan negatif, mencerminkan sikap pengguna yang cenderung ekspresif ketika merasakan kepuasan ataupun kekecewaan yang signifikan. Melalui analisis N-gram, ditemukan pola keluhan yang khas per platform: hambatan verifikasi dan inkonsistensi antarmuka di Ajaib, gangguan koneksi di IPOT, serta sinkronisasi akun keamanan yang bermasalah di Stockbit.

Dari sisi pemodelan, Logistic Regression tampil sebagai algoritma terbaik yang secara konsisten mengungguli Naive Bayes, SVM-SGD, dan *Random Forest* pada seluruh metrik evaluasi. Hasil ini memperkuat tesis bahwa kombinasi TF-IDF dengan model linear merupakan pasangan yang efisien untuk klasifikasi teks tidak berstruktur berbahasa Indonesia. Meski demikian, kelas netral masih menjadi titik lemah dengan recall yang sangat rendah akibat distribusi data yang tidak seimbang – penanganannya imbalanced class menjadi agenda prioritas untuk iterasi pengembangan berikutnya.

Di luar dimensi teknis, studi ini memberikan nilai praktis yang konkret: temuan berbasis N-gram dapat langsung dimanfaatkan oleh tim pengembang sebagai umpan balik sistematis untuk menentukan prioritas perbaikan layanan. Di sisi yang lebih luas, kerangka analisis sentimen yang dibangun dalam penelitian ini berpotensi dikembangkan menjadi sistem pemantauan umpan balik pengguna secara real-time yang adaptif.

Untuk penelitian selanjutnya, disarankan untuk: (1) mengeksplorasi model bahasa berbasis transformer seperti IndoBERT atau XLM-RoBERTa yang lebih memahami konteks bahasa Indonesia; (2) menerapkan aspect-based sentiment analysis (ABSA) agar dapat mengidentifikasi sentimen pada aspek spesifik seperti UI, performa, atau layanan pelanggan; (3) memperluas sumber data dari App Store (iOS) dan media sosial seperti Twitter/X untuk perspektif pengguna yang lebih beragam; dan (4) menerapkan teknik oversampling seperti SMOTE untuk mengatasi ketidakseimbangan kelas netral. Selain itu, pengembangan sistem dashboard visualisasi sentimen secara real-time juga menjadi peluang penelitian lanjutan yang dapat memberikan manfaat lebih luas bagi pengembang dan pengguna aplikasi.

DAFTAR PUSTAKA

- Bursa Efek Indonesia. (2024). Laporan Tahunan Jumlah Investor Ritel. *BEI*.
- Kavabilla, F. E., Widiariyah, T., & Warsito, B. (2023). Analisis sentimen pada ulasan aplikasi investasi online Ajaib pada Google Play menggunakan metode Support Vector Machine dan Maximum Entropy. *Jurnal Gaussian*, 11(4), 542–553. <https://doi.org/10.14710/j.gauss.11.4.542-553>
- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis sentimen ulasan aplikasi Dana dengan metode *Random Forest*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(9), 4305–4313. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11583>
- Lestari, D., & Hutagalung, J. (2025). Analisis sentimen ulasan marketplace berbahasa Indonesia menggunakan Logistic Regression dan TF-IDF. *Jurnal Sistem dan Teknologi Informasi (JustIN)*, 13(2), 85–93. <https://doi.org/10.26418/justin.v13i2.67891>
- Lestari, S., & Saepudin, S. (2021). Support Vector Machine: Analisis sentimen aplikasi saham di Google Play Store. *JUSIFO (Jurnal Sistem Informasi)*, 7(2), 81–90. <https://doi.org/10.19109/jusifo.v7i2.9825>
- Maulana, R., Voutama, A., & Ridwan, T. (2023). Analisis sentimen ulasan aplikasi MyPertamina pada Google Play Store menggunakan algoritma NBC. *Jurnal Teknologi Terpadu*, 9(1), 42–48. <https://doi.org/10.54914/jtt.v9i1.609>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier dan Confusion Matrix pada analisis sentimen berbasis teks pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Otoritas Jasa Keuangan. (2023). Statistik Pasar Modal Indonesia. *OJK*.
- Puji Astuti, N., Purbowati, R., & Lia Kurniawati. (2022). Analisis sentimen ulasan aplikasi BRImo menggunakan Support Vector Machine dan Naive Bayes. *Jurnal Informatika Upgris*, 8(2), 120–128. <https://doi.org/10.26877/jiu.v8i2.13021>
- Septiani, D., & Isabela, I. (2022). Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam temu kembali informasi pada dokumen teks. *SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia*, 1(2), 81–88. <https://journal.unj.ac.id/unj/index.php/SINTESIA/article/view/39364>
- Simanjuntak, W. O., Bijaksana, A., Negara, P., & Septriana, R. (2023). Perbandingan algoritma Logistic Regression dan *Random Forest* (Studi Kasus: Klasifikasi Emosi Tweet). *Jurnal Aplikasi dan Riset Informatika*, 2(1), 160–164. <https://doi.org/10.26418/juara.v2i1.69682>
- Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis sentimen aplikasi e-government pada Google Play menggunakan algoritma Naïve Bayes. *JATISI: Jurnal Teknik Informatika dan Sistem Informasi*, 9(2), 785–795. <https://doi.org/10.35957/jatisi.v9i2.1835>
- Yulianto, M. A., & Andrianto, R. (2025). Analisa kinerja algoritma supervised learning pada sentimen ulasan aplikasi investasi online Bibit. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 4(3), 1486–1496.