

Analisis Sentimen Ulasan Pengguna Game Honor of Kings Menggunakan Support Vector Machine (SVM)

Khoer Fadilah Fathurokhman*, Irving Vitra Paputungan

Universitas Islam Indonesia, Indonesia

Email: 21523101@students.uii.ac.id*, irving@uii.ac.id

Keywords:

sentiment analysis; support vector machine; TF-IDF; text classification; mobile game reviews

Abstract

User reviews in mobile gaming platforms provide valuable insights into user satisfaction but are difficult to analyze manually due to their large volume and unstructured nature. This study aims to develop and evaluate a sentiment classification model for Honor of Kings user reviews using a machine learning approach. The research employs a quantitative experimental method with data collected from Google Play Store, consisting of 3,500 Indonesian-language reviews. The preprocessing stages include cleaning, normalization, stemming, stopword removal, and tokenization. The data are then translated into English and labeled using a lexicon-based approach with TextBlob. Feature extraction is performed using Term Frequency-Inverse Document Frequency (TF-IDF), followed by classification using Support Vector Machine (SVM) with a linear kernel. The evaluation results show that the proposed model achieves an accuracy of 86.86%, indicating good performance in distinguishing positive and negative sentiments. The findings confirm that TF-IDF combined with SVM remains effective for sentiment classification on medium-sized Indonesian text datasets. However, limitations were observed in handling semantic context and translation bias during labeling. Future work may consider using deep learning-based models and manual annotation to improve classification performance and contextual understanding.

Kata Kunci:

analisis sentimen; support vector machine; TF-IDF; klasifikasi teks; ulasan game mobile

Abstrak

Ulasan pengguna di platform game seluler memberikan wawasan berharga mengenai kepuasan pengguna, namun sulit dianalisis secara manual karena volumenya yang besar dan sifatnya yang tidak terstruktur. Penelitian ini bertujuan untuk membangun dan mengevaluasi model klasifikasi sentimen pada ulasan pengguna Honor of Kings menggunakan pendekatan machine learning. Metode penelitian yang digunakan adalah eksperimen kuantitatif dengan data sebanyak 3.500 ulasan berbahasa Indonesia yang diperoleh dari Google Play Store. Tahapan preprocessing meliputi pembersihan data, normalisasi, stemming, stopword removal, dan tokenisasi. Data kemudian diterjemahkan ke dalam Bahasa Inggris dan dilabeli menggunakan pendekatan berbasis leksikon dengan TextBlob. Ekstraksi fitur dilakukan menggunakan metode TF-IDF, dan klasifikasi dilakukan menggunakan Support Vector Machine (SVM) dengan kernel linear. Hasil penelitian menunjukkan bahwa model mencapai akurasi sebesar 86,86%, yang menunjukkan kemampuan yang baik dalam mengklasifikasikan sentimen positif dan negatif. Temuan ini menunjukkan bahwa kombinasi TF-IDF dan SVM masih efektif untuk analisis sentimen pada dataset teks berbahasa Indonesia. Namun, terdapat keterbatasan dalam memahami konteks semantik serta potensi bias akibat proses translasi. Penelitian selanjutnya disarankan menggunakan pendekatan berbasis deep learning dan pelabelan manual untuk meningkatkan akurasi model.

PENDAHULUAN

Industri mobile gaming dalam beberapa tahun terakhir mengalami pertumbuhan yang sangat pesat dan menjadi kontributor utama dalam industri game global (Basri, 2024; Rohmah et al., 2024). Salah satu game yang memiliki dominasi signifikan adalah Honor of Kings, yang dikenal dengan basis pengguna yang besar serta tingkat interaksi yang tinggi. Interaksi tersebut salah satunya tercermin melalui ulasan pengguna pada platform distribusi aplikasi seperti Google Play Store (Irawan et al., 2024). Ulasan ini mengandung informasi penting terkait pengalaman bermain, kualitas sistem, serta tingkat kepuasan pengguna. Namun, volume data ulasan yang sangat besar dan terus bertambah menjadikan analisis secara manual tidak lagi efisien dan sulit dilakukan secara konsisten (Zahoor & Rohilla, 2020).

Dalam konteks tersebut, analisis sentimen berbasis komputasi menjadi pendekatan yang relevan untuk mengekstraksi opini pengguna secara otomatis. Analisis sentimen merupakan bagian dari bidang Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan polaritas suatu teks ke dalam kategori tertentu, umumnya positif dan negative (Ariq, 2025; Ertansyah et al., 2025). Pendekatan ini telah banyak digunakan dalam berbagai domain, seperti evaluasi layanan digital, kebijakan publik, hingga analisis ulasan produk, dan terbukti mampu memberikan gambaran umum mengenai persepsi pengguna dalam skala besar (Pamungkas & Cahyono, 2024; Witanti et al., 2022).

Sejumlah penelitian terdahulu menunjukkan bahwa pendekatan machine learning seperti Naïve Bayes, K-Nearest Neighbors, Decision Tree, dan Support Vector Machine dapat digunakan untuk melakukan klasifikasi sentimen. Di antara berbagai algoritma tersebut, Support Vector Machine (SVM) sering dilaporkan memiliki performa yang lebih stabil dalam menangani data teks berdimensi tinggi (Hasibuan & Serdano, 2022; Seno et al., 2024). Beberapa studi yang memanfaatkan data dari platform seperti Twitter, Google Maps, dan Google Play Store menunjukkan bahwa SVM mampu mencapai tingkat akurasi yang relatif tinggi, berkisar antara 73% hingga 87%, bergantung pada kualitas data dan tahapan preprocessing yang digunakan (Rhomaningtias et al., 2025; Surya et al., 2024).

Meskipun demikian, terdapat sejumlah keterbatasan dalam penelitian-penelitian sebelumnya. Sebagian besar penelitian masih berfokus pada domain umum, seperti media sosial atau aplikasi layanan, sementara kajian pada domain game mobile, khususnya genre MOBA, masih relatif terbatas (Maulana et al., 2023; Ramlan et al., 2023). Selain itu, penggunaan bahasa Indonesia dalam analisis sentimen menghadirkan tantangan tersendiri, terutama karena tingginya penggunaan bahasa informal, slang, serta istilah teknis yang spesifik pada komunitas game. Di sisi lain, beberapa penelitian juga menunjukkan adanya permasalahan pada ketidakseimbangan data serta penggunaan metode pelabelan otomatis yang berpotensi memengaruhi akurasi hasil klasifikasi (Kumar & Dalal, 2025).

Dalam menghadapi permasalahan tersebut, pendekatan berbasis kecerdasan buatan, khususnya Natural Language Processing (NLP), menjadi salah satu solusi yang banyak dikembangkan untuk memahami opini pengguna secara otomatis. Analisis sentimen merupakan salah satu cabang NLP yang bertujuan mengidentifikasi polaritas suatu teks berdasarkan opini atau emosi yang terkandung di dalamnya, seperti sentimen positif, negatif, maupun netral. Pemanfaatan analisis sentimen telah berkembang luas dalam berbagai bidang, mulai dari evaluasi produk digital, layanan publik, aplikasi finansial, hingga media sosial. Penelitian terdahulu menunjukkan bahwa algoritma machine learning mampu memberikan

hasil yang efektif dalam melakukan klasifikasi teks dengan jumlah data besar. Beberapa penelitian menggunakan algoritma seperti Naïve Bayes, K-Nearest Neighbor (KNN), Decision Tree, Logistic Regression, hingga Support Vector Machine (SVM) untuk mengidentifikasi pola sentimen pengguna. Di antara berbagai algoritma tersebut, SVM banyak digunakan karena memiliki kemampuan dalam menangani data berdimensi tinggi serta menghasilkan performa yang stabil pada klasifikasi teks berbasis fitur numerik. Penelitian mengenai analisis sentimen menggunakan SVM pada berbagai objek menunjukkan tingkat akurasi yang kompetitif, seperti penelitian Ipmawati et al. (2024) yang memperoleh akurasi tinggi dalam analisis ulasan berbasis Google Maps, serta penelitian lain yang menunjukkan efektivitas SVM dalam mengklasifikasikan opini pengguna pada berbagai platform digital.

Meskipun berbagai penelitian telah membuktikan efektivitas metode SVM dalam analisis sentimen, masih terdapat beberapa keterbatasan yang menjadi ruang pengembangan penelitian lebih lanjut. Sebagian besar penelitian terdahulu masih berfokus pada objek umum seperti media sosial, aplikasi layanan digital, aplikasi transportasi, maupun ulasan produk, sementara penelitian yang secara khusus membahas komunitas pengguna game mobile, terutama pada genre MOBA, masih relatif terbatas. Karakteristik ulasan dalam industri game memiliki kompleksitas tersendiri karena banyak menggunakan bahasa informal, istilah komunitas, singkatan, slang, serta ekspresi emosional yang berbeda dibandingkan domain lainnya. Selain itu, beberapa penelitian sebelumnya masih menghadapi tantangan dalam proses pelabelan data otomatis, ketidakseimbangan kelas sentimen, serta keterbatasan metode berbasis frekuensi kata seperti TF-IDF dalam memahami konteks semantik suatu kalimat. Kondisi tersebut menunjukkan adanya kesenjangan penelitian (research gap) terkait bagaimana membangun model analisis sentimen yang mampu menangani karakteristik khusus ulasan pengguna game berbahasa Indonesia secara lebih akurat dan relevan.

Urgensi penelitian ini semakin meningkat karena ulasan pengguna pada platform digital saat ini tidak hanya berfungsi sebagai bentuk ekspresi individu, tetapi juga menjadi sumber data strategis bagi pengembang aplikasi dalam mengambil keputusan berbasis pengguna (user-driven decision making). Pada industri game kompetitif seperti Honor of Kings, pemahaman terhadap sentimen pemain dapat membantu pengembang mengidentifikasi kelemahan sistem, mengevaluasi pembaruan permainan, memahami tingkat kepuasan pengguna, serta mempertahankan loyalitas pemain. Tanpa metode analisis otomatis, informasi yang terkandung dalam ribuan ulasan akan sulit dimanfaatkan secara optimal karena keterbatasan waktu dan sumber daya manusia. Oleh karena itu, penerapan model klasifikasi sentimen berbasis machine learning menjadi penting untuk menyediakan mekanisme analisis yang cepat, objektif, dan terukur. Penelitian ini menjadi relevan karena berupaya memberikan pendekatan komputasional yang dapat membantu mengubah data teks tidak terstruktur menjadi informasi yang bernilai bagi pengambilan keputusan strategis.

Kebaruan penelitian ini terletak pada penerapan kombinasi metode preprocessing teks, ekstraksi fitur TF-IDF, dan algoritma SVM pada domain spesifik berupa ulasan pengguna Honor of Kings berbahasa Indonesia. Berbeda dengan penelitian sebelumnya yang banyak menggunakan data dari media sosial atau aplikasi umum, penelitian ini secara khusus mengeksplorasi karakteristik opini pengguna dalam lingkungan game mobile yang memiliki pola komunikasi dan kosakata unik. Selain itu, penelitian ini memberikan evaluasi empiris terhadap kemampuan SVM dalam mengklasifikasikan sentimen pada dataset berukuran

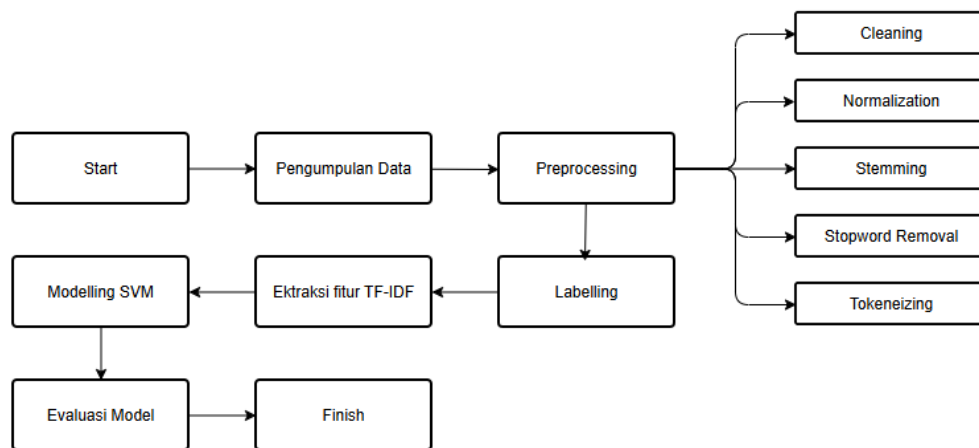
menengah dengan karakteristik bahasa informal. Penggunaan tahapan preprocessing yang meliputi cleaning, normalisasi, stemming, stopword removal, dan tokenisasi menjadi strategi penting untuk meningkatkan kualitas representasi data sebelum dilakukan proses klasifikasi. Dengan demikian, penelitian ini memberikan kontribusi terhadap pengembangan kajian analisis sentimen pada bidang gaming analytics serta memperluas penerapan metode NLP dalam memahami perilaku pengguna digital.

Berdasarkan permasalahan dan kesenjangan penelitian tersebut, tujuan utama penelitian ini adalah mengembangkan dan mengevaluasi model klasifikasi sentimen terhadap ulasan pengguna Honor of Kings menggunakan algoritma Support Vector Machine dengan pendekatan ekstraksi fitur TF-IDF. Penelitian ini bertujuan untuk melakukan pengumpulan data ulasan pengguna, melakukan tahapan pengolahan teks, membangun model klasifikasi sentimen, serta mengukur performa model berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Melalui pendekatan eksperimen kuantitatif, penelitian ini diharapkan mampu memberikan gambaran empiris mengenai efektivitas SVM dalam mengolah data opini pengguna pada platform game digital. Selain itu, penelitian ini juga berupaya menjawab pertanyaan mengenai sejauh mana metode berbasis machine learning konvensional masih mampu digunakan dalam menghadapi kompleksitas bahasa pengguna pada lingkungan digital modern.

Hasil penelitian ini diharapkan memberikan manfaat secara teoritis maupun praktis. Secara teoritis, penelitian ini dapat memperkaya literatur mengenai penerapan NLP dan machine learning dalam bidang analisis sentimen, khususnya pada konteks industri game mobile berbahasa Indonesia. Temuan penelitian dapat menjadi referensi bagi penelitian selanjutnya yang ingin mengembangkan model analisis opini berbasis deep learning, transformer model, maupun pendekatan analisis sentimen berbasis aspek (aspect-based sentiment analysis). Secara praktis, penelitian ini memberikan manfaat bagi pengembang game dan pengelola platform digital dalam memahami persepsi pengguna secara lebih cepat dan sistematis. Informasi hasil klasifikasi sentimen dapat digunakan sebagai dasar evaluasi kualitas layanan, peningkatan fitur permainan, serta penyusunan strategi pengembangan produk yang lebih berorientasi pada kebutuhan pengguna. Dengan demikian, penelitian ini tidak hanya berkontribusi terhadap pengembangan teknologi analisis teks, tetapi juga mendukung pengambilan keputusan berbasis data dalam industri digital yang semakin kompetitif.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun dan mengevaluasi model klasifikasi sentimen berbasis machine learning (Murhadi, 2025; Sugiyono, 2022). Objek penelitian berupa ulasan pengguna game Honor of Kings yang diperoleh dari platform Google Play Store Indonesia. Pengumpulan data dilakukan melalui teknik web scraping menggunakan bahasa pemrograman Python dengan bantuan pustaka google-play-scraper. Dataset yang digunakan terdiri dari 3.500 ulasan berbahasa Indonesia yang dikumpulkan dalam periode 24 September hingga 23 Oktober 2025. Tahapan penelitian dalam studi ini dirancang secara sistematis mulai dari pengumpulan data hingga evaluasi model. Alur metodologi penelitian ditunjukkan pada Gambar 1.



Gambar 1. Metode Penelitian

Gambar 1 menunjukkan bahwa penelitian dimulai dari tahap pengumpulan data melalui proses scraping ulasan pengguna, yang kemudian dilanjutkan dengan preprocessing data yang meliputi pembersihan teks, normalisasi, stemming, penghapusan stopwords, dan tokenisasi. Data yang telah diproses selanjutnya melalui tahap pelabelan menggunakan pendekatan berbasis leksikon dengan bantuan TextBlob setelah dilakukan translasi ke Bahasa Inggris (Laia et al., 2024). Translasi ke Bahasa Inggris hanya dilakukan untuk keperluan pelabelan otomatis menggunakan TextBlob. Sementara itu, data teks yang digunakan pada tahap ekstraksi fitur TF-IDF dan pelatihan model SVM tetap menggunakan teks asli berbahasa Indonesia hasil preprocessing. Tahap berikutnya adalah ekstraksi fitur menggunakan metode TF-IDF untuk mengubah data teks menjadi representasi numerik (Hartatik, 2024; Kurniasari et al., 2025). Data tersebut kemudian digunakan dalam proses pemodelan menggunakan algoritma Support Vector Machine (SVM), dan hasil klasifikasi dievaluasi menggunakan confusion matrix.

Preprocessing Data

Data ulasan yang diperoleh masih dalam bentuk teks tidak terstruktur sehingga perlu dilakukan tahapan preprocessing untuk meningkatkan kualitas data. Proses ini meliputi pembersihan teks (cleaning), normalisasi, stemming, stopwords removal, dan tokenisasi.

Pembersihan data dilakukan dengan menghapus karakter yang tidak relevan seperti simbol, angka, dan tanda baca, serta mengubah seluruh teks menjadi huruf kecil. Normalisasi dilakukan dengan mengonversi kata tidak baku menjadi bentuk baku menggunakan kamus bahasa Indonesia. Selanjutnya, proses stemming dilakukan menggunakan pustaka Sastrawi untuk mengubah kata menjadi bentuk dasar, diikuti dengan stopwords removal untuk menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap makna sentimen. Tahap akhir adalah tokenisasi, yaitu pemecahan teks menjadi unit kata.

Pelabelan Data

Pelabelan dilakukan secara otomatis menggunakan pendekatan berbasis leksikon melalui pustaka TextBlob. Karena keterbatasan TextBlob dalam memproses Bahasa Indonesia, teks hasil preprocessing diterjemahkan terlebih dahulu ke dalam Bahasa Inggris menggunakan pustaka deep-translator. Proses translasi ini hanya digunakan untuk menghasilkan label sentimen otomatis (ground truth annotation), sedangkan data teks yang digunakan dalam proses ekstraksi fitur dan pelatihan model tetap menggunakan korpus asli berbahasa Indonesia. Nilai polaritas digunakan sebagai dasar klasifikasi, di mana nilai lebih dari nol dikategorikan sebagai

sentimen positif dan nilai kurang dari atau sama dengan nol dikategorikan sebagai sentimen negatif.

Ekstraksi Fitur

Representasi numerik teks dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) [23], [24]. Ekstraksi fitur dilakukan terhadap teks hasil preprocessing dalam Bahasa Indonesia yang tersimpan pada korpus penelitian. Metode ini mengukur pentingnya suatu kata dalam dokumen relatif terhadap seluruh korpus.

$$TF(t, d) = \frac{f_{t,d}}{\sum_k f_{k,d}}$$

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

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

di mana $f_{t,d}$ adalah frekuensi kemunculan kata t dalam dokumen d , N adalah jumlah total dokumen, dan $df(t)$ adalah jumlah dokumen yang mengandung kata t .

Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan teknik *train-test split*.

Model klasifikasi dibangun menggunakan algoritma Support Vector Machine (SVM) dengan kernel linear. SVM bekerja dengan mencari hyperplane optimal yang memisahkan dua kelas dengan margin maksimum.

$$w \cdot x + b = 0$$

Untuk data berlabel, fungsi keputusan SVM dinyatakan sebagai:

$$y_i(wx_i + b) \geq 1$$

Margin antara dua kelas didefinisikan sebagai:

$$Margin = \frac{1}{\|w\|}$$

Model dilatih menggunakan pustaka *scikit-learn* dengan parameter kernel linear dan nilai regulasi $C = 1.0$ untuk mengontrol kompleksitas model.

Evaluasi Model

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

Metrik evaluasi yang digunakan meliputi akurasi, presisi, dan recall yang dirumuskan sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

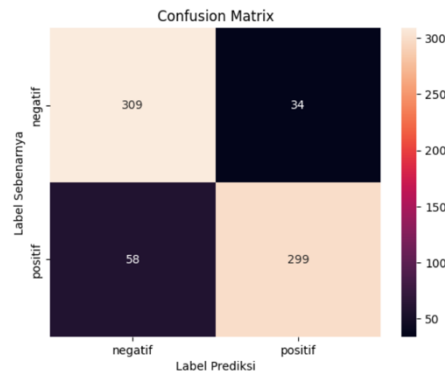
$$Recall = \frac{TP}{TP + FN}$$

Confusion matrix digunakan untuk menggambarkan distribusi hasil prediksi terhadap label aktual, sedangkan metrik evaluasi digunakan untuk mengukur performa model dalam mengklasifikasikan sentimen secara kuantitatif.

HASIL DAN PEMBAHASAN

Hasil Evaluasi Model

Hasil pengujian model Support Vector Machine (SVM) menunjukkan bahwa model mampu mengklasifikasikan sentimen dengan tingkat akurasi sebesar **86,86%**. Untuk memahami performa model secara lebih komprehensif, digunakan confusion matrix yang menggambarkan distribusi hasil prediksi terhadap label aktual.



Gambar 2. Confusion matrix hasil klasifikasi sentimen menggunakan SVM

Confusion matrix menunjukkan bahwa model menghasilkan distribusi klasifikasi sebagai berikut:

Tabel 1. Hasil Confusion Matrix

	Prediksi Negatif	Prediksi Positif
Aktual Negatif	309 (TN)	34 (FN)
Aktual Positif	58 (FP)	299 (TP)

Berdasarkan tabel 1, dapat diamati bahwa model memiliki kemampuan yang baik dalam mengidentifikasi kedua kelas, ditunjukkan oleh nilai *true negative* dan *true positive* yang relatif tinggi.

Metrik Evaluasi Model

Untuk memberikan gambaran kuantitatif terhadap performa model, dilakukan perhitungan metrik evaluasi sebagai berikut:

Tabel 2. Hasil Evaluasi Model

Metrik	Nilai
Accuracy	86.86%
Precision	83.75%
Recall	89.76%
F1-Score	86.64%

Nilai precision menunjukkan tingkat ketepatan prediksi positif, sedangkan recall menggambarkan kemampuan model dalam menangkap seluruh data positif. F1-score sebagai rata-rata harmonik dari precision dan recall menunjukkan keseimbangan performa model dalam klasifikasi.

Analisis Temuan Ilmiah

Temuan utama dalam penelitian ini menunjukkan bahwa kombinasi TF-IDF dan SVM mampu memberikan performa klasifikasi yang tinggi pada data ulasan berbahasa Indonesia. Secara teoritis, hal ini disebabkan oleh kemampuan SVM dalam memaksimalkan margin antar kelas pada ruang fitur berdimensi tinggi, sehingga mampu memisahkan data dengan karakteristik kompleks.

Namun, distribusi kesalahan menunjukkan adanya kecenderungan nilai *false negative* yang lebih tinggi dibandingkan *false positive*. Hal ini mengindikasikan bahwa model lebih sering mengklasifikasikan ulasan positif sebagai negatif. Secara ilmiah, fenomena ini dapat dijelaskan oleh keterbatasan TF-IDF yang tidak mempertimbangkan konteks semantik, serta adanya distorsi makna akibat proses translasi pada tahap pelabelan.

Selain itu, penggunaan bahasa informal dalam ulasan game, termasuk slang dan ekspresi campuran, menyebabkan ambiguitas dalam representasi fitur. Kondisi ini memperkuat asumsi bahwa pendekatan berbasis frekuensi kata belum sepenuhnya mampu menangkap kompleksitas bahasa alami dalam domain spesifik seperti game.

Perbandingan dengan Penelitian Terdahulu

Hasil penelitian ini konsisten dengan penelitian sebelumnya yang menunjukkan bahwa SVM memiliki performa yang baik dalam analisis sentimen. Penelitian oleh Adelia (2024) melaporkan akurasi sebesar 73,6%, sedangkan penelitian oleh Ipmawati et al. (2024) mencapai akurasi hingga 87%. Dengan hasil akurasi sebesar 86,86%, penelitian ini berada dalam rentang performa yang sama, sehingga memperkuat validitas metode yang digunakan.

Namun demikian, dibandingkan dengan pendekatan berbasis *deep learning*, metode ini masih memiliki keterbatasan dalam memahami konteks kalimat secara lebih mendalam. Oleh karena itu, peningkatan performa dapat dilakukan dengan mengintegrasikan model berbasis *transformer* atau melakukan pelabelan manual untuk meningkatkan kualitas data.

KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi sentimen ulasan pengguna Honor of Kings menggunakan pendekatan Support Vector Machine (SVM) dengan representasi fitur TF-IDF. Berdasarkan hasil evaluasi, model menunjukkan kemampuan yang baik dalam mengklasifikasikan sentimen positif dan negatif, sehingga tujuan penelitian untuk membangun dan mengevaluasi model analisis sentimen berbasis machine learning dapat tercapai.

Temuan ilmiah utama dalam penelitian ini menunjukkan bahwa pendekatan berbasis TF-IDF dan SVM masih efektif dalam menangani data teks berbahasa Indonesia yang tidak terstruktur, khususnya pada domain ulasan game mobile. Hal ini mengindikasikan bahwa metode klasifikasi klasik tetap relevan untuk dataset berukuran menengah dengan kompleksitas linguistik tertentu.

Namun demikian, penelitian ini juga mengidentifikasi adanya keterbatasan dalam proses klasifikasi, terutama terkait ketidakmampuan model dalam memahami konteks semantik secara mendalam serta potensi bias yang muncul akibat proses translasi pada tahap pelabelan. Kondisi ini berdampak pada munculnya kesalahan klasifikasi, khususnya pada data dengan ekspresi bahasa yang ambigu atau mengandung campuran sentimen.

Berdasarkan temuan tersebut, penelitian selanjutnya disarankan untuk mengembangkan pendekatan yang lebih adaptif terhadap bahasa Indonesia, seperti penggunaan model berbasis deep learning atau transformer, serta mempertimbangkan pelabelan data secara manual untuk meningkatkan kualitas ground truth. Selain itu, pengembangan analisis sentimen berbasis aspek (aspect-based sentiment analysis) juga dapat menjadi arah penelitian lanjutan untuk memperoleh pemahaman yang lebih spesifik terhadap faktor-faktor yang memengaruhi pengalaman pengguna.

REFERENSI

- Adelia Irawan, F., Rialdy Atmadja, A., Wahana, A., & Universitas Islam Negeri Sunan Gunung Djati Bandung. (2024). *Analisis sentimen ulasan aplikasi bank digital menggunakan algoritma Naive Bayes*.
- Ariq, H. (2025). Analisis sentimen ulasan produk untuk mengukur kepuasan pelanggan e-commerce menggunakan metode natural language processing (NLP) dan support vector machine (SVM). *JPNM: Jurnal Pustaka Nusantara Multidisiplin*, 3. <https://doi.org/10.59945/jpnm.v3i3.739>
- Basri, F. E. (2024). E-sport dalam industri game menjadi strategi soft diplomacy untuk memperkuat pertumbuhan perekonomian China. *Al-Mutsala*, 6(1), 268–280. <https://doi.org/10.46870/jstain.v6i1.1042>
- Ertansyah, G., Kusuma, R., & Sari, A. (2025). Analisis sentimen pada media sosial menggunakan teknik natural language processing. *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, 183–189. <https://doi.org/10.47701/qgcey104>
- Hartatik, H. (2024). Enhancing customer experience in e-commerce through lexicon and TextBlob sentiment analysis. *West Science Social and Humanities Studies*, 2(7), 1237–1245. <https://doi.org/10.58812/wsshs.v2i07.1747>
- Hasibuan, M. S., & Serdano, A. (2022). Analisis sentimen kebijakan pembelajaran tatap muka menggunakan support vector machine dan Naive Bayes. *JRST (Jurnal Riset Sains dan Teknologi)*, 6(2), 199. <https://doi.org/10.30595/jrst.v6i2.15145>
- Ipmawati, J., Saifulloh, S., & Kusnawi, K. (2024). Analisis sentimen tempat wisata berdasarkan ulasan pada Google Maps menggunakan algoritma support vector machine. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 247–256. <https://doi.org/10.57152/malcom.v4i1.1066>
- Kumar, V., & Dalal, M. (2025). Constitutional morality in the digital age: AI surveillance and privacy concerns. *Journal of Multidisciplinary Knowledge*, 5(2), 212–220. <https://doi.org/10.36676/jmk.v5.i2.147>
- Kurniasari, D., Hdiana, Y. Z., Lumbanraja, F. R., Warsono, W., & Hadi, N. A. (2025). Understanding consumer sentiments: A TextBlob-based sentiment analysis study. *Integra: Journal of Integrated Mathematics and Computer Science*, 2(3), 81–88. <https://doi.org/10.26554/integrajimcs.20252340>
- Laia, Y., Berutu, S. S., Sumihar, Y. P., & Budiati, H. (2024). Implementasi library TextBlob dan metode support vector machine pada analisis sentimen pelanggan terhadap jasa transportasi online. *Building of Informatics, Technology and Science (BITS)*, 6(1). <https://doi.org/10.47065/bits.v6i1.5090>
- Maulana, I., Apriandari, W., & Pambudi, A. (2023). *Analisis sentimen berbasis aspek terhadap ulasan aplikasi MyPertamina menggunakan support vector machine*. <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index>
- Murhadi, W. (2025). *Metode penelitian: Pendekatan kuantitatif dan kualitatif*. <https://repository.ubaya.ac.id/48121/>

- Pamungkas, A. S., & Cahyono, N. (2024). Analisis sentimen review ChatGPT di Play Store menggunakan support vector machine dan K-nearest neighbor. *Edumatic: Jurnal Pendidikan Informatika*, 8(1), 1–10. <https://doi.org/10.29408/edumatic.v8i1.24114>
- Ramdan Adi Surya, M., Martanto, M., & Hayati, U. (2024). Analisis sentimen ulasan pengguna OVO menggunakan algoritma Naive Bayes pada Google Play Store. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 2780–2786. <https://doi.org/10.36040/jati.v8i3.8739>
- Ramlan, R., Satyahadewi, N., & Andani, W. (2023). Analisis sentimen pengguna Twitter menggunakan support vector machine pada kasus kenaikan harga BBM. *Jambura Journal of Mathematics*, 5(2), 431–445. <https://doi.org/10.34312/jjom.v5i2.20860>
- Rhomaningtias, L., Khairunisa, A., Wara, S. S. M., & Hindrayani, K. M. (2025). Analisis sentimen ulasan aplikasi Smile Indonesia menggunakan metode Naive Bayes dan support vector machine (SVM). *HOAQ (High Education of Organization Archive Quality): Jurnal Teknologi Informasi*, 16(1), 79–91. <https://doi.org/10.52972/hoa.vol16no1.p79-91>
- Rohmah, M., Ismail, K., Rahmadani, R., Masitoh, G., & Putri, D. A. P. (2024). Inovasi dan transformasi industri dalam mendorong pertumbuhan ekonomi Indonesia. *Jurnal Neraca: Jurnal Pendidikan dan Ilmu Ekonomi Akuntansi*, 8(1), 43–52. <https://doi.org/10.31851/neraca.v8i1.14391>
- Seno, B. A., Widodo, & Adhi, B. P. (2024). Penerapan algoritma support vector machine untuk mendeteksi emosi dari teks bahasa Indonesia. *PINTER: Jurnal Pendidikan Teknik Informatika dan Komputer*, 8(1), 72–80. <https://doi.org/10.21009/pinter.8.1.8>
- Sugiyono. (2022). *Metode penelitian kuantitatif, kualitatif, dan R&D* (27th ed.). Alfabeta.
- Witanti, A., Yogyakarta Jl Raya Wates-Jogjakarta, B., Sedayu, K., Bantul, K., & Istimewa Yogyakarta, D. (2022). Analisis sentimen masyarakat terhadap vaksinasi COVID-19 pada media sosial Twitter menggunakan algoritma support vector machine (SVM). *Jurnal Sistem Informasi dan Informatika (SIMIKA)*, 5.
- Zahoor, S., & Rohilla, R. (2020). Twitter sentiment analysis using lexical or rule based approach: A case study. In *Proceedings of the 2020 IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)* (pp. 537–542). IEEE. <https://doi.org/10.1109/ICRITO48877.2020.9197910>