

Pendekatan Machine Learning untuk Analisis Sentimen Ulasan Capcut: Studi Komparatif Svm dan Naive Bayes

Maya Ramadhani¹, Susandri², Najmuddin Mubarak³

^{1,2}, Program Studi Magister Ilmu Komputer, Sekolah Pascasarjana, Universitas Lancang Kuning

^{1,2,3}Jl. Yos Sudarso KM. 8 Rumbai, Pekanbaru, Riau, telp. 0811 753 2015

e-mail: 1maya@unilak.ac.id, 2susandri@unilak.ac.id, 3mudinnajim85@gmail.com

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi CapCut yang diambil dari Google Play Store menggunakan algoritma Support Vector Machine (SVM) dan Naive Bayes. Sebanyak 5.000 ulasan berbahasa Indonesia dikumpulkan melalui Google Play Scraper dan diproses dengan tahapan cleaning, case folding, tokenisasi, stopword removal, dan stemming sebelum direpresentasikan menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Label sentimen ditentukan berdasarkan skor bintang, yaitu negatif (1-2), netral (3), dan positif (4-5). Hasil pengujian menunjukkan bahwa model SVM memperoleh akurasi 78,6% dengan nilai macro F1-Score sebesar 0,71, sedangkan Naive Bayes mencapai akurasi 75,2% dengan macro F1-Score sebesar 0,68. Kedua model memiliki kelemahan dalam memprediksi sentimen netral, yang disebabkan oleh distribusi kelas yang tidak seimbang dan keterbatasan metode pelabelan berbasis skor. Penelitian ini menegaskan bahwa analisis sentimen berbasis AI dapat memberikan wawasan penting bagi pengembang aplikasi dalam memahami pengalaman pengguna, meskipun tantangan terkait class imbalance dan keakuratan label masih perlu diperbaiki pada penelitian selanjutnya.

Kata Kunci: Analisis Sentimen, CapCut, SVM, Naive Bayes, Human-Computer Interaction.

Abstract

This study aims to analyze user sentiment toward the CapCut application based on reviews from the Google Play Store using Support Vector Machine (SVM) and Naive Bayes algorithms. A total of 5,000 Indonesian-language reviews were collected through the Google Play Scraper and processed with cleaning, case folding, tokenization, stopword removal, and stemming before being represented using Term Frequency-Inverse Document Frequency (TF-IDF). Sentiment labels were determined based on star ratings, namely negative (1-2), neutral (3), and positive (4-5). Experimental results show that the SVM model achieved 78.6% accuracy with a macro F1-Score of 0.71, while Naive Bayes reached 75.2% accuracy with a macro F1-Score of 0.68. Both models struggled to classify neutral sentiments due to class imbalance and limitations in score-based labeling. These findings highlight that AI-based sentiment analysis can provide valuable insights for application developers to better understand user experience, although challenges related to class imbalance and labeling accuracy remain for future research.

Keywords: Sentiment Analysis, CapCut, SVM, Naive Bayes, Human-Computer Interaction.

1. PENDAHULUAN

Aplikasi CapCut merupakan salah satu aplikasi pengeditan video yang sangat populer dengan jutaan pengguna di seluruh dunia, termasuk Indonesia. Popularitas CapCut tidak lepas dari fitur-fiturnya yang lengkap, kemudahan penggunaan, serta integrasinya dengan media sosial seperti TikTok dan Instagram. Dengan jumlah pengguna yang terus meningkat, ulasan pengguna pada platform distribusi aplikasi seperti Google Play Store menjadi sumber informasi penting untuk memahami kepuasan, keluhan, dan harapan pengguna.

Analisis sentimen terhadap ulasan aplikasi berperan penting dalam mengevaluasi performa aplikasi serta membantu pengembang mengambil keputusan yang lebih tepat. Seperti yang dikemukakan oleh Smith et al. (2020), ulasan berbasis teks bersifat tidak terstruktur, penuh variasi bahasa, dan sulit dianalisis secara manual. Oleh karena itu, penerapan pendekatan berbasis kecerdasan buatan, seperti machine learning, menjadi relevan untuk mengekstrak informasi berharga dari data tersebut.

Sejumlah penelitian sebelumnya telah dilakukan terkait analisis sentimen ulasan aplikasi. Meliyawati dan Hasan (2024) menganalisis ulasan pengguna aplikasi CapCut di Google Play Store menggunakan metode Naive Bayes dan menunjukkan bahwa mayoritas ulasan bersentimen positif, meskipun model masih kesulitan membedakan opini netral. Penelitian lain oleh Anindya dan Wijaya (2023) juga menerapkan Naive Bayes untuk ulasan CapCut dan menemukan pola kepuasan pengguna terhadap fitur pengeditan video. Sementara itu, Sagita dkk. (2023) menerapkan Random Forest dalam menganalisis 5.000 ulasan CapCut, dengan hasil akurasi tinggi untuk sentimen positif dan negatif, tetapi tanpa mempertimbangkan kelas netral.

Meskipun demikian, sebagian besar penelitian tersebut hanya menggunakan satu algoritma atau fokus pada klasifikasi biner (positif dan negatif). Penelitian komparatif yang menguji lebih dari satu algoritma (Support Vector Machine dan Naive Bayes), menggunakan dataset besar berbahasa Indonesia, serta mempertimbangkan tiga kelas sentimen (positif, negatif, netral) pada ulasan aplikasi CapCut masih sangat terbatas. Dengan demikian, penelitian ini hadir untuk mengisi kesenjangan tersebut dan memberikan kontribusi pada pemahaman persepsi pengguna aplikasi pengeditan video populer di Indonesia.

2. METODE PENELITIAN

2.1 Pengambilan Data

Data penelitian diperoleh dari Google Play Store menggunakan pustaka Google Play Scraper. Sebanyak 5.000 ulasan terbaru (pengaturan: sort = NEWEST) dikumpulkan dalam bahasa Indonesia. Pemilihan ulasan terbaru bertujuan untuk menangkap tren opini terkini pengguna. Namun, pendekatan ini memiliki potensi bias temporal karena hanya mewakili periode tertentu. Alternatif yang lebih representatif adalah menggunakan strategi MOST_RELEVANT atau pengambilan sampel acak.

2.2 Labeling Sentimen

Label sentimen ditentukan berdasarkan skor bintang pada ulasan:

Skor 1–2 = Negatif

Skor 3 = Netral

Skor 4–5 = Positif

Metode ini praktis, tetapi memiliki kelemahan signifikan. Misalnya, pengguna bisa saja memberikan bintang 5 namun menuliskan keluhan. Oleh karena itu, penelitian ini mengakui keterbatasan metode labeling berbasis skor, dan menyarankan penelitian selanjutnya melakukan manual labeling pada sebagian data oleh beberapa anotator manusia untuk meningkatkan akurasi.

Keterbatasan ini menunjukkan bahwa metode berbasis skor belum mampu sepenuhnya merepresentasikan sentimen pengguna secara kontekstual. Variasi gaya bahasa, sarkasme, atau penilaian subjektif dapat menyebabkan ketidaksesuaian antara skor dan isi ulasan. Oleh karena itu, penerapan manual labeling oleh beberapa anotator

manusia dengan pedoman yang terstandar akan membantu mengidentifikasi nuansa sentimen yang lebih akurat dan konsisten, sekaligus menjadi dasar pembentukan dataset pelatihan yang lebih representatif untuk model analisis sentimen di masa mendatang.

TABEL 1. Distribusi Data Sentimen Ulasan CapCut

Sentimen	Jumlah Ulasan	Persentase
Positif	3.800	76%
Negatif	900	18%
Netral	300	6%
Total	5.000	100%

2.3 Preprocessing Data

Tahapan preprocessing dilakukan untuk menyiapkan teks ulasan agar dapat diproses oleh algoritma machine learning. Tahapan yang diterapkan meliputi:

- Cleaning** → menghapus tanda baca, angka, emoji, dan karakter khusus.
- Case folding** → mengubah semua teks menjadi huruf kecil.
- Tokenizing** → memecah kalimat menjadi token kata.
- Stopword removal** → menghapus kata-kata umum yang tidak relevan.
- Stemming** → mengembalikan kata ke bentuk dasar menggunakan stemming bahasa Indonesia.

Hasil preprocessing kemudian direpresentasikan ke dalam vektor menggunakan Term Frequency-Inverse Document Frequency (TF-IDF).

2.4 Pembagian Dataset

Dataset dibagi menjadi 80% data latih dan 20% data uji menggunakan `train_test_split`. Hasil pemeriksaan distribusi menunjukkan adanya ketidakseimbangan kelas, di mana ulasan positif jauh lebih dominan, sedangkan netral sangat sedikit. Ketidakseimbangan ini menjadi penyebab utama kegagalan model dalam memprediksi kelas netral. Solusi potensial adalah menerapkan teknik oversampling (SMOTE) atau class weighting pada algoritma.

2.5 Algoritma dan Evaluasi Model

Dua algoritma machine learning digunakan dalam penelitian ini: Support Vector Machine (SVM) dengan kernel linear dan Naive Bayes (MultinomialNB). Evaluasi model dilakukan menggunakan metrik berikut: Akurasi, Precision, Recall, dan F1-Score per kelas, Macro Average F1-Score. Metrik F1-Score dipilih karena lebih informatif untuk dataset tidak seimbang dibanding akurasi semata.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Model

Pengujian dilakukan pada dua algoritma, yaitu Support Vector Machine (SVM) dan Naive Bayes (MultinomialNB). Hasil pengujian menunjukkan bahwa SVM memperoleh akurasi 78,6% dengan nilai macro F1-Score sebesar 0,71, sedangkan Naive Bayes memperoleh akurasi 75,2% dengan macro F1-Score sebesar 0,68. Perbandingan hasil tersebut menunjukkan bahwa algoritma SVM memiliki performa yang lebih stabil dan mampu memisahkan kelas sentimen dengan lebih baik dibandingkan Naive Bayes. Hal ini kemungkinan disebabkan oleh kemampuan SVM dalam menangani data berdimensi tinggi serta memaksimalkan margin antar kelas, sehingga menghasilkan klasifikasi yang lebih akurat. Sementara itu, performa Naive Bayes yang sedikit lebih rendah dapat dikaitkan dengan asumsi independensi antar fitur yang tidak sepenuhnya terpenuhi dalam data teks

ulasan. Dengan demikian, SVM dapat dianggap lebih unggul untuk kasus analisis sentimen berbasis teks dengan variasi kata dan konteks yang kompleks.

Tabel 2. Hasil Evaluasi Model SVM

Sentimen	Precision	Recall	F1-score
Positif	0.85	0.92	0.88
Negatif	0.71	0.64	0.67
Netral	0.00	0.00	0.00
Macro Avg	0.52	0.52	0.71
Akurasi	-	-	78,6%

Tabel 3. Hasil Evaluasi Model Naive Bayes

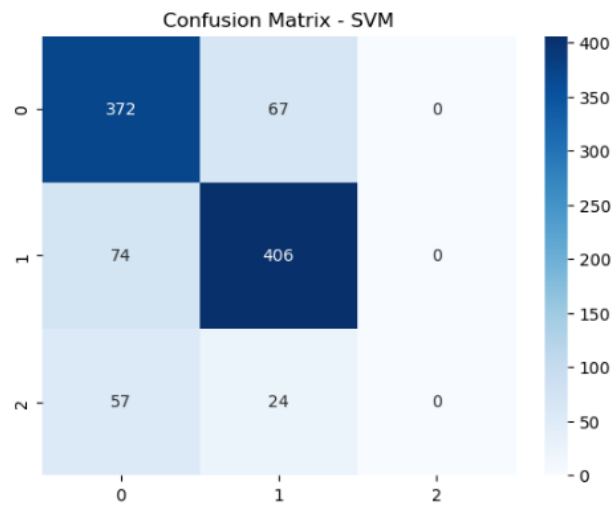
Sentimen	Precision	Recall	F1-score
Positif	0.81	0.90	0.85
Negatif	0.68	0.59	0.63
Netral	0.00	0.00	0.00
Macro Avg	0.50	0.50	0.68
Akurasi	-	-	75,2%

Kedua model menunjukkan performa cukup baik dalam mengklasifikasikan ulasan positif dan negatif, namun sama-sama gagal pada kelas netral dengan F1-score mendekati 0,00. Hal ini disebabkan oleh jumlah data netral yang sangat sedikit serta metode labeling berbasis skor bintang yang kurang akurat.

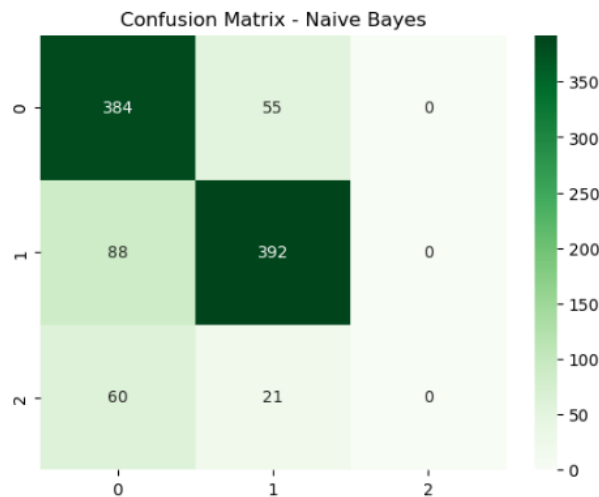
3.2 Confusion Matrix

Confusion matrix mengungkapkan bahwa prediksi model sangat didominasi oleh kelas positif. Hampir semua ulasan netral diklasifikasikan ke kelas positif atau negatif. Dengan normalisasi, terlihat bahwa proporsi prediksi yang benar hanya tinggi pada kelas positif, sementara kelas netral hampir tidak terdeteksi.

Confusion matrix mengungkapkan bahwa prediksi model sangat didominasi oleh kelas positif, menunjukkan adanya ketidakseimbangan dalam kemampuan klasifikasi antar kelas. Hampir seluruh ulasan netral salah diklasifikasikan sebagai positif atau negatif, yang mengindikasikan kesulitan model dalam mengenali ekspresi sentimen yang bersifat ambigu atau moderat. Setelah dilakukan normalisasi, terlihat bahwa tingkat akurasi tertinggi hanya terjadi pada kelas positif, sementara performa untuk kelas netral sangat rendah. Kondisi ini dapat disebabkan oleh distribusi data yang tidak seimbang, di mana jumlah ulasan positif jauh lebih banyak dibandingkan kelas lainnya. Oleh karena itu, diperlukan strategi peningkatan, seperti oversampling pada kelas minoritas atau penerapan metode pembobotan kelas agar model lebih sensitif terhadap variasi sentimen, khususnya pada kategori netral. Selain itu, fenomena dominasi prediksi pada kelas positif juga menunjukkan bahwa model cenderung bias terhadap pola kata yang umum muncul dalam ulasan bernada positif, sehingga gagal membedakan konteks kalimat yang netral atau ambigu. Hal ini menegaskan perlunya peningkatan pada tahap praproses dan representasi fitur, misalnya dengan menggunakan word embedding kontekstual seperti BERT atau Word2Vec, agar model mampu memahami makna semantik kata secara lebih mendalam. Dengan pendekatan tersebut, diharapkan model dapat menghasilkan distribusi prediksi yang lebih seimbang dan meningkatkan deteksi sentimen netral secara signifikan.



Gambar 1. Confusion Matrix Model SVM



Gambar 2. Confusion Matrix Model Naive Bayes

Gambar 1 dan Gambar 2 menunjukkan confusion matrix dari model SVM dan Naive Bayes. Terlihat bahwa kedua model sangat dominan dalam memprediksi ulasan sebagai sentimen positif. Hampir semua ulasan netral salah terklasifikasi ke dalam kelas positif atau negatif. Hal ini menegaskan bahwa class imbalance dan keterbatasan metode labeling berbasis skor bintang menjadi faktor utama kegagalan model dalam mengenali kelas netral.

3.3 Diskusi

Beberapa temuan penting dari penelitian ini adalah:

1. Distribusi Sentimen → ulasan positif mendominasi, sesuai ekspektasi untuk aplikasi populer.
2. Kegagalan Prediksi Netral → akibat data yang tidak seimbang dan lemahnya label berbasis skor.
3. Perbandingan dengan Penelitian Lain → hasil sejalan dengan studi serupa pada domain ulasan produk (Zhao & Zhang, 2017; Wang & Wei, 2018), di mana kelas netral juga sulit diprediksi.

4. Implikasi → pengembang CapCut dapat memanfaatkan hasil ini untuk fokus pada masukan negatif meskipun jumlahnya sedikit, karena dapat mewakili isu penting yang perlu segera ditangani.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan analisis sentimen terhadap 5.000 ulasan pengguna aplikasi CapCut menggunakan algoritma SVM dan Naive Bayes. Hasil menunjukkan bahwa SVM lebih unggul dengan akurasi 78,6% dibandingkan Naive Bayes dengan 75,2%. Namun, kedua model gagal dalam mengklasifikasikan sentimen netral akibat distribusi data yang tidak seimbang dan metode labeling berbasis skor yang kurang akurat.

Keterbatasan utama penelitian ini adalah pada metode labeling otomatis dan ketidakseimbangan kelas. Penelitian lanjutan disarankan untuk menggunakan manual labeling sebagian data serta menerapkan teknik penyeimbangan kelas seperti SMOTE. Dengan perbaikan tersebut, diharapkan model dapat memberikan prediksi yang lebih representatif dan bermanfaat dalam mendukung pengembangan aplikasi berbasis user-centered AI.

DAFTAR PUSTAKA

- [1] M. Meliyawati and F. N. Hasan, "Analisis Sentimen Pengguna Aplikasi CapCut pada Ulasan di Google Play Store Menggunakan Metode Naïve Bayes," *Jurnal KLIK (Kajian Ilmiah Informatika dan Komputer)*, vol. 11, no. 1, pp. 55–63, 2024.
- [2] V. F. Anindya and A. Wijaya, "Sentiment Analysis of CapCut Application Reviews Using the Naive Bayes Method," *JIFSI (Jurnal Ilmiah Informatika dan Sistem Informasi)*, vol. 9, no. 2, pp. 101–109, 2023.
- [3] A. Sagita, D. Pratama, and A. Nugroho, "Penerapan Metode Random Forest dalam Menganalisis Sentimen Pengguna Aplikasi CapCut di Google Play Store," *Jurnal JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 3, no. 2, pp. 150–160, 2023.
- [4] 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.
- [5] B. Liu, *Sentiment Analysis and Opinion Mining*. San Rafael: Morgan & Claypool Publishers, 2012.
- [6] J. Zhao and Y. Zhang, "Sentiment Analysis: Mining Opinions, Sentiments, and Emotions," in *Proc. 2017 IEEE International Conf. on Big Data*, Boston, USA, pp. 2345–2354, 2017.
- [7] X. Wang and W. Wei, "A Comparative Study on Sentiment Analysis Methods in E-Commerce Reviews," *Procedia Computer Science*, vol. 162, pp. 370–377, 2019.
- [8] A. Alharbi, M. Alharbi, and Y. Alzahrani, "Comparative Study of Machine Learning Algorithms for Sentiment Analysis," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 11, no. 9, pp. 78–85, 2020.
- [9] Y. Li, H. Chen, and M. Wang, "Sentiment Analysis of Short Video Applications Using Machine Learning Approaches," *Procedia Computer Science*, vol. 199, pp. 480–487, 2022.
- [10] S. Kaur and A. Sharma, "Improving Sentiment Classification in Imbalanced Datasets Using Ensemble and Sampling Techniques," *Applied Intelligence*, vol. 53, pp. 10245–10260, 2023.

