



Analisis Sentimen Terhadap Ulasan Aplikasi Tiktok Pada Google Play Store Menggunakan Metode TF-IDF dan Naïve Bayes

Wahyu Riski Maulana^{1*}, Bambang Irawan²

^{1*2}Program Studi Teknik Informatika, Universitas Muhadi Setiabudi, Indonesia

*Email : maulanawahyuriski24@email.com

Abstract: The rapid growth of short-video social media platforms such as TikTok has significantly increased the volume of user reviews that reflect public perceptions of application quality. These reviews constitute electronic word of mouth (e-WOM), which influences brand image, user trust, and adoption decisions. This study aims to analyze user sentiment toward the TikTok application using a text mining approach based on the Naïve Bayes algorithm and Term Frequency–Inverse Document Frequency (TF-IDF) weighting. The dataset consists of user reviews collected from the Google Play Store and categorized into sentiment classes. The data were processed through several preprocessing stages, including text cleaning, tokenization, stopword removal, normalization, and stemming, before feature extraction and classification. The experimental results indicate that the proposed model achieved an accuracy of 67.35% in classifying sentiment. However, analysis of the confusion matrix and prediction distribution reveals a bias toward the majority class due to dataset imbalance. The Naïve Bayes classifier combined with TF-IDF representation demonstrates satisfactory performance in identifying dominant extreme sentiments (positive and negative), yet its effectiveness decreases in multi-class classification scenarios with uneven data distribution. These findings suggest that although probabilistic models remain computationally efficient and relevant, additional strategies such as resampling techniques or class weight adjustments are necessary to improve performance in imbalanced multi-class sentiment classification. This research provides empirical evidence for sentiment analysis of digital application reviews in the Indonesian context and offers a foundation for further model optimization in future studies.

Keywords: sentiment analysis, Naïve Bayes, TF-IDF, TikTok.

1. PENDAHULUAN

Perkembangan teknologi digital dalam satu dekade terakhir telah mentransformasi pola komunikasi, distribusi informasi, serta interaksi sosial masyarakat secara fundamental. Konvergensi teknologi mobile, komputasi awan, dan kecerdasan buatan mendorong pertumbuhan pesat ekosistem media sosial sebagai ruang publik digital. Media sosial tidak lagi sekadar sarana berbagi informasi, tetapi telah berevolusi menjadi medium ekonomi kreatif, personal branding, pemasaran digital, hingga pembentukan opini publik. Dalam konteks ini, TikTok muncul sebagai salah satu platform berbasis video pendek dengan pertumbuhan pengguna paling signifikan secara global maupun di Indonesia. Keunggulan kompetitif platform ini terletak pada sistem rekomendasi berbasis algoritma machine learning yang mampu mempersonalisasi konten secara adaptif berdasarkan perilaku pengguna. (Rahmawati&Putra,2022).

Seiring dengan meningkatnya penetrasi pengguna, intensitas interaksi pengguna terhadap aplikasi juga mengalami peningkatan, termasuk dalam bentuk pemberian ulasan (reviews) pada Google Play Store. Ulasan pengguna tersebut merepresentasikan pengalaman empiris, persepsi terhadap kualitas layanan, evaluasi performa sistem, stabilitas aplikasi, kebijakan privasi, hingga kepuasan terhadap fitur-fitur yang disediakan. Secara konseptual, ulasan digital dapat dipandang sebagai bentuk *electronic word of mouth* (e-WOM) yang memiliki implikasi strategis terhadap reputasi aplikasi dan pengambilan keputusan pengguna lain (Pratama et al., 2023). Namun demikian, karakteristik data ulasan yang bersifat tidak terstruktur (unstructured text), berjumlah besar (high volume), dan dinamis (*velocity*) menyebabkan analisis manual menjadi tidak efisien serta rentan terhadap bias subjektif.

Dalam perspektif data science, permasalahan tersebut dapat diatasi melalui pendekatan *text mining* dan *natural language processing* (NLP). Analisis sentimen merupakan salah satu cabang *text mining* yang berfokus pada identifikasi, ekstraksi, dan klasifikasi opini dalam data teks menjadi kategori tertentu, umumnya positif, negatif, dan netral. Pendekatan ini memungkinkan transformasi data kualitatif menjadi informasi kuantitatif yang terukur sehingga dapat digunakan sebagai dasar pengambilan keputusan berbasis data (*data-driven decision making*). Beberapa penelitian di Indonesia menunjukkan bahwa penerapan analisis sentimen pada ulasan aplikasi mobile efektif dalam mengidentifikasi pola kepuasan pengguna serta isu-isu dominan yang perlu ditindaklanjuti pengembang (Sari & Nugroho, 2021; Hidayat et al., 2022).

Secara metodologis, representasi fitur teks memegang peranan krusial dalam menentukan performa model klasifikasi. *Term Frequency-Inverse Document Frequency* (TF-IDF) merupakan metode pembobotan kata yang banyak digunakan karena mampu merepresentasikan tingkat kepentingan suatu term dalam dokumen relatif terhadap keseluruhan korpus. Kombinasi TF-IDF dengan algoritma klasifikasi probabilistik seperti *Naïve Bayes* terbukti memiliki kinerja yang kompetitif, khususnya pada dataset berukuran menengah dengan kompleksitas linguistik moderat (Wibowo et al., 2022). *Naïve Bayes* bekerja berdasarkan prinsip probabilitas bersyarat dengan asumsi independensi antarfitur, sehingga efisien secara komputasi dan relatif stabil terhadap noise pada data teks.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan metode TF-IDF dan algoritma *Naïve Bayes* dalam menganalisis sentimen ulasan pengguna TikTok pada Google Play Store. Penelitian ini tidak hanya berkontribusi pada pengembangan model klasifikasi sentimen berbahasa Indonesia, tetapi juga memberikan implikasi praktis bagi pengembang aplikasi dalam merumuskan strategi peningkatan kualitas layanan berbasis analisis data empiris. Dengan pendekatan yang sistematis dan terukur, hasil penelitian diharapkan mampu memberikan gambaran objektif mengenai persepsi pengguna terhadap performa dan fitur aplikasi TikTok di Indonesia.

2. TINJUAN PUSTAKA

2.1 Analisis sentimen

Analisis sentimen merupakan cabang *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi dan mengklasifikasikan opini dalam teks ke dalam kategori tertentu, seperti positif, negatif, atau netral. Teknik ini digunakan untuk mengubah data teks yang tidak terstruktur menjadi informasi kuantitatif yang dapat dianalisis secara sistematis. Dalam penelitian aplikasi digital, analisis sentimen membantu mengukur persepsi dan tingkat kepuasan pengguna secara objektif.

2.2 Ulasan Pengguna Google pada Play Store

Google Play Store menyediakan fitur rating dan komentar yang memungkinkan pengguna memberikan evaluasi terhadap aplikasi. Ulasan tersebut mencerminkan pengalaman nyata pengguna terhadap performa, fitur, dan stabilitas aplikasi. Karena jumlahnya besar dan tidak terstruktur, diperlukan teknik *text mining* untuk mengekstraksi informasi penting secara otomatis dan efisien.

2.3 TF-IDF

TF-IDF (*Term Frequency–Inverse Document Frequency*) adalah metode pembobotan kata untuk menentukan tingkat kepentingan suatu istilah dalam dokumen dibandingkan dengan seluruh dokumen dalam korpus. Metode ini menekankan kata yang sering muncul dalam satu dokumen namun jarang muncul di dokumen lain.

Visualisasi rumus TF-IDF:

$$TF_{IDF}(t, d) = TF(t, d) \times \log\left(\frac{N}{df(t)}\right)$$

Pendekatan ini efektif dalam merepresentasikan teks ke dalam bentuk numerik sebagai input algoritma klasifikasi.

2.4 Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi probabilistik yang menggunakan Teorema Bayes dengan asumsi independensi antar fitur. Algoritma ini banyak digunakan pada klasifikasi teks karena sederhana, cepat, dan stabil untuk data berdimensi tinggi.

Visualisasi Teorema Bayes:

$$P(C | X) = \frac{P(X | C) \times P(C)}{P(X)}$$

Dalam analisis sentimen, kelas dengan probabilitas tertinggi akan dipilih sebagai hasil prediksi sentimen suatu teks.

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimental. Proses penelitian dilakukan secara sistematis melalui beberapa tahapan, yaitu pengumpulan data, prapemrosesan teks, ekstraksi fitur, pembangunan model klasifikasi, serta evaluasi performa model. Pendekatan ini dipilih untuk menguji secara empiris kemampuan algoritma dalam mengklasifikasikan sentimen ulasan pengguna secara terukur dan objektif. Berikut Gambar Langkah langkahnya.



Gambar 1. Diagram Alur Penelitian

3.1 Pengumpulan data

Data diperoleh melalui teknik scraping terhadap ulasan aplikasi TikTok pada Google Play Store sebanyak 20.000 ulasan berbahasa Indonesia. Data yang dikumpulkan mencakup tiga atribut utama, yaitu teks ulasan, Review ID, Username, Date rating (skala 1–5), dan tanggal ulasan.

Proses pelabelan sentimen dilakukan secara otomatis berdasarkan rating yang diberikan pengguna, dengan skema sebagai berikut: rating 4–5 dikategorikan sebagai sentimen positif, rating 3 sebagai netral, dan rating 1–2 sebagai negatif. Skema ini digunakan untuk membangun dataset terstruktur yang siap digunakan dalam proses pelatihan dan pengujian model.

3.2 Tahap preprocessing

Tahap preprocessing bertujuan meningkatkan kualitas data teks sebelum dilakukan ekstraksi fitur. Proses ini meliputi:

-
- a. Case Folding – Mengubah seluruh teks menjadi huruf kecil untuk menghindari perbedaan akibat kapitalisasi.
 - b. Tokenisasi – Memecah teks menjadi unit kata (token).
 - c. Stopword Removal – Menghapus kata-kata umum yang tidak memiliki makna signifikan.
 - d. Normalisasi – Mengubah kata tidak baku atau singkatan menjadi bentuk standar.
 - e. Stemming – Mengubah kata ke bentuk dasar untuk mengurangi variasi morfologis.

3.3 Ekstraksi Fitur

Representasi teks dilakukan menggunakan metode TF-IDF untuk mengubah data teks menjadi vektor numerik. Pembobotan ini mempertimbangkan frekuensi kemunculan kata dalam dokumen serta tingkat kelangkaannya dalam seluruh korpus.

Formulasi TF-IDF:

$$TF_{IDF}(t, d) = TF(t, d) \times \log\left(\frac{N}{df(t)}\right)$$

Hasil pembobotan ini menghasilkan matriks fitur yang menjadi input bagi algoritma klasifikasi.

3.4 Kasifikasi

Model klasifikasi yang digunakan adalah Multinomial Naïve Bayes. Algoritma ini dipilih karena sesuai untuk data berbasis frekuensi kata dan memiliki efisiensi komputasi yang tinggi. Multinomial Naïve Bayes menghitung probabilitas posterior setiap kelas berdasarkan distribusi kata dalam dokumen, kemudian memilih kelas dengan probabilitas tertinggi sebagai hasil prediksi.

3.5 Evaluasi Model

Evaluasi performa model dilakukan menggunakan confusion matrix untuk mengukur kualitas prediksi. Berdasarkan confusion matrix, dihitung beberapa metrik evaluasi, yaitu Accuracy, Precision, Recall, dan F1-Score.

Accuracy mengukur proporsi prediksi yang benar terhadap seluruh data. Precision menunjukkan ketepatan model dalam memprediksi suatu kelas, sedangkan Recall mengukur kemampuan model dalam mendeteksi seluruh data yang benar-benar termasuk dalam kelas tersebut. F1-Score merupakan rata-rata harmonik antara Precision dan Recall yang memberikan gambaran keseimbangan performa model secara keseluruhan.

4. HASIL dan PEMBAHASAN

4.1 Deskripsi Dataset

Dataset penelitian terdiri dari 20.000 ulasan aplikasi TikTok yang diambil dari Google Play Store. Setiap data memiliki atribut:

- Review ID
- Username
- Rating (1–5)
- Review Text
- Date

Contoh data yang digunakan dapat dilihat pada Tabel berikut (berdasarkan gambar):

index	Review ID	Username	Rating	Review Text	Date
0	de4d9ffd-8ab2-4c99-956f-068ee2142ce7	Pengguna Google	5	bagus aku suka banget	2025-02-19 07:36:29
1	e791207c-29aa-40b1-9c17-3b93771e2539	Pengguna Google	5	Mantap	2025-02-19 07:35:50
2	0cbe1fa8-25d8-4082-a48f-8fad62a3dde4	Pengguna Google	1	Bisa mendownload tapi tidak bisa masuk,dan jika saya masuk aplikasinya ada kata tiktok terhenti tolong ini kenapa saya tidak bisa masuk ke aplikasi tiktoknya,apakah ada masalah dengan tiktok saya atau memang masalahnya dari tiktoknya,tolong diperbaiki masalah saya akan memberikan bintang 5!	2025-02-19 07:35:06
3	8eabce5c-b4cb-47e3-ae55-b33c93f325f1	Pengguna Google	1	Aku suka tiktok tetapi tidak bisa membuat jeda untuk patah patah aku gk suka	2025-02-19 07:34:08
4	876c3be7-7f10-4358-891f-6c27262ccda7	Pengguna Google	1	Tiktok sekarang udah jelek kebanyakan di update jadi lemot ditambah HP ku mati setiap buka tiktok	2025-02-19 07:33:29

Gambar 2. dataset

Distribusi rating menunjukkan bahwa kelas 5 (bintang lima) memiliki jumlah data paling banyak, sedangkan rating 2–4 relatif lebih sedikit. Hal ini menyebabkan ketidakseimbangan kelas (*class imbalance*) yang memengaruhi performa model.

4.2 Hasil Evaluasi Model

Akurasi: 0.6735

Classification Report:

	precision	recall	f1-score	support
1	0.54	0.85	0.66	5149
2	0.26	0.01	0.01	1525
3	0.27	0.04	0.07	1639
4	0.30	0.05	0.09	1702
5	0.79	0.89	0.84	9985
accuracy			0.67	20000
macro avg	0.43	0.37	0.33	20000
weighted avg	0.60	0.67	0.60	20000

Confusion Matrix:

```
[[4379  9  43  28 690]
 [1200  9  35  27 254]
 [1076  4  65  62 432]
 [ 595  5  56  87 959]
 [ 926  7  38  84 8930]]
```

Gambar 3. Hasil Evaluasi Model

4.2.1 Akurasi

Hasil pengujian menunjukkan:

Akurasi = 0.6735 (67,35%)

Artinya, dari 20.000 data uji, sekitar 67% berhasil diklasifikasikan dengan benar.

4.3 Analisis Hasil

4.3.1 Performa Terbaik

Model menunjukkan performa terbaik pada:

Rating 5 → F1-score = 0.84

Rating 1 → F1-score = 0.66

Hal ini menunjukkan model cukup baik dalam membedakan sentimen ekstrem (sangat positif dan sangat negatif).

4.3.2 Performa Terlemah

Kelas rating 2, 3, dan 4 memiliki recall yang sangat rendah:

Rating 2 → Recall 0.01

Rating 3 → Recall 0.04

Rating 4 → Recall 0.05

Artinya model hampir tidak mampu mengenali kelas menengah dan cenderung mengklasifikasikan ulasan tersebut ke kelas mayoritas.

4.4 Pembahasan

Berdasarkan hasil evaluasi, dapat disimpulkan bahwa:

1. Model Naïve Bayes efektif untuk klasifikasi sentimen polar (positif vs negatif ekstrem).
2. Ketidakseimbangan data menyebabkan rendahnya performa pada kelas menengah.
3. Akurasi 67% tergolong moderat, namun Macro F1-score rendah (0.33) menunjukkan distribusi prediksi tidak merata.

5 KESIMPULAN

Model klasifikasi sentimen yang dibangun menggunakan algoritma Naïve Bayes dengan representasi fitur TF-IDF mampu mencapai tingkat akurasi sebesar 67,35%. Secara umum, hasil ini menunjukkan bahwa model telah mampu menangkap pola linguistik dasar pada ulasan pengguna TikTok, khususnya dalam membedakan sentimen yang bersifat ekstrem, yaitu sangat positif (rating 5) dan sangat negatif (rating 1). Hal ini terlihat dari nilai recall dan f1-score yang relatif tinggi pada kedua kelas tersebut dibandingkan kelas lainnya.

Namun demikian, evaluasi yang lebih mendalam melalui precision, recall, f1-score, serta confusion matrix menunjukkan adanya bias yang signifikan terhadap kelas mayoritas. Kelas dengan jumlah data besar, terutama rating 5, mendominasi prediksi model, sementara kelas menengah (rating 2, 3, dan 4) memiliki performa yang rendah, khususnya pada metrik recall. Kondisi ini mengindikasikan bahwa model cenderung mengalami kesulitan dalam membedakan nuansa sentimen yang tidak bersifat ekstrem, sehingga banyak ulasan dengan polaritas moderat salah diklasifikasikan ke dalam kelas dominan.

Rendahnya macro-average f1-score dibandingkan weighted-average f1-score semakin memperkuat indikasi ketidakseimbangan performa antar kelas. Dengan kata lain, meskipun akurasi keseluruhan terlihat cukup baik, kemampuan generalisasi model terhadap seluruh kategori rating belum merata. Hal ini menunjukkan bahwa pendekatan Naïve Bayes berbasis TF-IDF cukup efektif untuk klasifikasi sentimen polar, tetapi kurang optimal ketika diterapkan pada skema multi-kelas dengan distribusi data yang tidak seimbang.

Secara keseluruhan, hasil penelitian ini menegaskan bahwa pemilihan algoritma dan strategi penanganan data sangat berpengaruh terhadap performa klasifikasi sentimen. Untuk meningkatkan kualitas model pada konteks multi-kelas, diperlukan pendekatan tambahan seperti penyeimbangan data, rekayasa fitur lanjutan, atau penggunaan algoritma yang lebih kompleks agar representasi semantik ulasan dapat ditangkap secara lebih komprehensif.

REFERENSI

- Hidayat, R., Firmansyah, M., & Kusuma, A. (2022). Analisis sentimen ulasan aplikasi mobile menggunakan metode Naïve Bayes dan TF-IDF. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(3), 450–458.
- Pratama, D., Saputra, R., & Wicaksono, A. (2023). Electronic word of mouth dan pengaruhnya terhadap persepsi kualitas aplikasi digital di Indonesia. *Jurnal Ilmu Komunikasi*, 21(1), 55–67.
- Rahmawati, L., & Putra, Y. (2022). Algoritma rekomendasi dan personalisasi konten pada platform media sosial berbasis artificial intelligence. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 9(4), 789–798.
- Sari, N., & Nugroho, E. (2021). Klasifikasi sentimen ulasan aplikasi pada Google Play Store menggunakan pendekatan text mining. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNETI)*, 10(2), 120–127.
- Wibowo, A., Santoso, H., & Prasetyo, B. (2022). Perbandingan performa algoritma klasifikasi pada analisis sentimen teks bahasa Indonesia. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)*, 6(8), 3775–3783.
- Putri, M. A., Kurniawan, D., & Ramadhan, F. (2021). Implementasi metode TF-IDF dan Naïve Bayes untuk analisis sentimen ulasan aplikasi berbasis Android. *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, 2(3), 321–330.
- Saputra, I. G., Mahendra, G. S., & Wijaya, I. M. (2022). Analisis sentimen pengguna aplikasi mobile menggunakan algoritma klasifikasi machine learning. *Jurnal Media Informatika Budidarma*, 6(4), 2101–2109.
- Kusnadi, A., Lestari, S., & Hidayanto, A. (2023). Evaluasi performa model klasifikasi teks bahasa Indonesia pada data tidak seimbang. *Jurnal Teknologi dan Sistem Komputer*, 11(2), 95–103.
- Anggraini, D., Prakoso, B., & Utami, E. (2021). Penerapan text mining untuk klasifikasi opini pengguna pada platform digital. *Jurnal Ilmiah Teknologi Informasi Asia*, 15(2), 85–94.

Firmanto, Y., Setiawan, R., & Wahyudi, T. (2023). Perbandingan algoritma Naïve Bayes dan Support Vector Machine pada analisis sentimen multi-kelas. *Jurnal Pengembangan IT (JPIT)*, 8(1), 44–53.