

ANALISIS SENTIMEN ULASAN *GAME* DELTA FORCE MOBILE MENGGUNAKAN *NAIVE BAYES* DAN *SUPPORT VECTOR MACHINE*

Faiz Najwan Zaky¹⁾, Muhamad Rendi Gibran²⁾, Sumanto³⁾, Jefina Tri Kumalasari⁴⁾, Ghofar Taufiq⁵⁾, Ade Christian⁶⁾

^{1,2}Prodi Informatika, Fakultas Teknik & Informatika, Universitas Bina Sarana Informatika

Correspondence author: M.R.Gibran, rendigibran1@gmail.com, Bekasi, Indonesia

Abstract

This study aims to analyze sentiment in Delta Force Mobile reviews on the Google Play Store using Naive Bayes and Support Vector Machines. Sentiment classification in reviews helps game developers understand player sentiment and improve user satisfaction. The method used is the KDD framework, which focuses on web scraping, text cleaning, tokenization, stemming, and TF-IDF for data weighting. The experimental data consisted of 29,454 reviews, with a peak accuracy rate of 86.17%. The study found a significant improvement in SVM's ability to classify negative sentiment accurately. SVM had a score of 0.69, while Naive Bayes had a score of 0.56. This indicates that SVM is more accurate at classifying reviews that contain technical complaints, such as lag and bugs.

Keywords: sentiment analysis, naive bayes, Support Vector Machine, delta force

Abstrak

Penelitian ini bertujuan menganalisis sentimen pada ulasan Delta Force Mobile di Google Play Store menggunakan *Naive Bayes* dan *Support Vector Machine*. Klasifikasi sentimen dalam ulasan digunakan untuk membantu pengembang *game* memahami sentimen pemain dan membantu meningkatkan kepuasan pengguna. Metode yang digunakan yaitu kerangka kerja KDD yang berfokus pada *web scraping*, pembersihan teks, tokenisasi, *stemming*, dan TF-IDF untuk memberi bobot pada data. Data yang digunakan dalam eksperimen berjumlah 29.454 ulasan, dengan tingkat akurasi tertinggi sebesar 86.17%. Hasil penelitian mendapatkan temuan signifikan bahwa SVM memiliki kemampuan yang lebih baik dalam mengklasifikasikan sentimen negatif secara akurat. SVM memiliki nilai 0,69 dan *Naive Bayes* memiliki nilai 0,56. Hal ini menunjukkan bahwa SVM lebih akurat dalam mengklasifikasikan ulasan yang mengandung keluhan teknis seperti *lag* dan *bug*.

Kata Kunci: analisis sentimen, *Naive Bayes*, *Support Vector Machine*, delta force

A. PENDAHULUAN

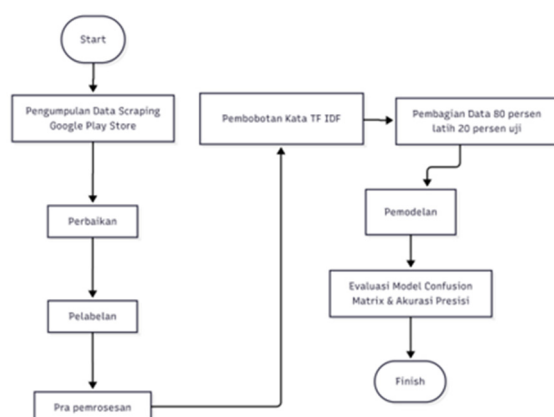
Selama sepuluh tahun terakhir ini industri *game* seluler mengalami pertumbuhan secara eksponensial. Hal ini dipicu oleh *smartphone*

yang semakin terjangkau, walaupun masih terdapat pengguna aktif yang terdaftar di berbagai platform aplikasi, yang mengakibatkan bergesernya pola umpan balik pengguna dalam jumlah besar. *Feedback* dari

pengguna di Google Play Store menjadi pertimbangan bagi pengembang untuk mengatasi masalah dari *bug*, dan mengenali fitur yang diinginkan pengguna. Begitupun, dari ribuan data yang terdaftar, *feedback* satu per satu secara random yang dianalisis sentimen secara komputasi (Nugraha et al., 2023). Pada aplikasi “Delta Force Mobile”, pengembangan masih kesulitan untuk memprioritaskan area perbaikan, meski secara manual mendapatkan data *feedback* yang melimpah. Hal ini tidak berbeda dari penelitian sebelumnya pada sasaran *feedback* data *game (Mobile Apps)* PUBG MOBILE (Sari et al., 2024). Dimana analisis sentimen menjadi salah satu fokus untuk permasalahan tersebut. Dua diantaranya, *Naive Bayes* dan *Support Vector Machine (SVM)* (Alyandi, 2025) yang sering dicontohkan. *Naive Bayes* terkenal cepat dan sederhana, sedangkan SVM lebih baik untuk menangani data yang berukuran besar (*high dimensional data*) dengan memisahkan data dan mencari *hyperplane* pemisah (Widyanto et al., 2023). Pengujian SVM pada dataset yang tidak seimbang oleh (Yani, 2025) pada perbandingan Mobile Legends, menunjukkan hasil yang lebih stabil. Penelitian ini bertujuan membandingkan kinerja kedua metode tersebut untuk memberikan masukan dalam mengembangkan Delta Force Mobile Game.

B. METODE PENELITIAN

Penelitian ini bersifat eksperimen dan berbasis pada komputer. Ilustrasi Gambar 1 menjadi acuan untuk mengikuti langkah-langkah yang dilakukan penelitian.



Gambar 1. Alur Metodologi Penelitian

Data Scraping

Data ulasan dari pengguna telah diperoleh dari Google Play Store untuk permainan “Delta Force Mobile”. Pengambilan data (*scraping*) dilakukan menggunakan Pustaka Python *google-play-scraper* (Rismansyah et al., 2025). Parameter pengumpulan data diatur untuk negara Indonesia (*country='id'*) dan Bahasa Indonesia (*lang='id'*), dengan pengurutan berdasarkan ulasan yang relevan. Proses ini berhasil mengumpulkan 40.588 data ulasan mentah

Data Reparation

Sebelum memasuki tahap pemrosesan teks, dilakukan proses pembersihan data untuk memastikan kualitas dataset tetap terjaga. Pada tahap ini, data yang memiliki nilai kosong (*missing values*) dihapus, begitu pula ulasan yang terdeteksi duplikat. Dari total 40.588 data awal setelah proses pembersihan diperoleh 31.631 data ulasan yang unik dan siap diproses pada tahap selanjutnya

Data Labeling

Pelabelan data dilakukan dengan mengacu pada skor rating yang diberikan oleh pengguna. Ulasan dengan rating 4 dan 5 diklasifikasikan sebagai sentimen positif, sedangkan ulasan dengan bintang 1 dan 2 dikategorikan sebagai sentimen negatif. Sementara itu, ulasan dengan rating bintang 3 dianggap bersifat netral sehingga dikeluarkan dari dataset karena tidak menunjukkan kecenderungan sentimen yang jelas. Setelah proses pelabelan, diperoleh distribusi data sebanyak 22.933 ulasan bersentimen positif dan 6.521 ulasan bersentimen negatif.

Text Preprocessing

Data teks diproses melalui serangkaian tahapan agar dapat dibaca oleh algoritma:

1. Cleaning

Pada tahap ini, teks dibersihkan dari berbagai elemen yang tidak diperlukan, seperti URL, karakter selain huruf, angka,

serta spasi yang berlebihan. Proses ini dilakukan dengan bantuan *Regular Expression* “Regex” agar teks menjadi lebih bersih, konsisten, dan mudah diolah pada tahap berikutnya.

2. *Case Folding*

Tahap ini dilakukan dengan mengubah seluruh huruf dan teks menjadi huruf kecil (*lower case*). Tujuan adalah untuk menyeragamkan penulisan kata sehingga perbedaan huruf besar dan kecil tidak memengaruhi proses analisis selanjutnya.

3. *Tokenizing*

Pada tahap ini, teks dipecah menjadi unit-unit kata atau potongan kata (token). Proses tokenisasi bertujuan untuk mengubah kalimat menjadi daftar kata sehingga lebih mudah dianalisis pada tahap pemrosesan berikutnya.

4. *Stemming*

Tahap *stemming* dilakukan untuk mengubah kata-kata yang memiliki imbuhan menjadi bentuk kata dasarnya. Proses ini menggunakan pustaka Sastrawi, yang dirancang khusus untuk menangani proses *stemming* dalam Bahasa Indonesia, sehingga kata-kata menjadi lebih seragam dan mudah dianalisis

5. *Stopword Removal*

Pada tahap ini, kata-kata umum atau kata hubung yang tidak memiliki pengaruh signifikan terhadap sentimen dihapus dari teks. Proses penghapusan *stopword* dilakukan menggunakan daftar *stopword* dari Sastrawi, sehingga hanya kata-kata yang relevan dengan analisis sentimen yang dipertahankan..

Ekstraksi Fitur dan Pemodelan

Data teks yang telah melalui tahap pembersihan selanjutnya direpresentasikan dalam bentuk vektor numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Untuk menjaga efisiensi komputasi, jumlah fitur dibatasi hingga maksimal 5.000 kata. Selanjutnya, dataset dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan perbandingan 80:20.

Pada tahap klasifikasi, digunakan dua algoritma pembelajaran mesin sebagai berikut:

1. *Naive Bayes* (MultinomialNB): Algoritma *Naive Bayes* dengan pendekatan Multinomial dipilih karena kemampuan yang efisien dalam menangani data teks dengan dimensi tinggi, serta performanya yang baik pada tugas klasifikasi.
2. *Support Vector Machine* (SVM): Algoritma *Support Vector Machine* digunakan dengan kernel linear (SVC kernel= ‘linear’) karena mampu memisahkan dua kelas sentimen secara efektif, khususnya pada data teks yang memiliki karakteristik berdimensi tinggi.

Evaluasi Model

Kinerja kedua model dievaluasi menggunakan *Confusion Matrix* untuk memperoleh nilai akurasi, presisi (*precision*), *recall*, dan *F1-score*. Proses evaluasi ini dilakukan untuk membandingkan performa masing-masing algoritma serta menentukan model yang paling akurat dalam mengklasifikasikan sentimen pengguna.

C. HASIL DAN PEMBAHASAN

Hasil Pra-pemrosesan Data

Berdasarkan hasil proses *scraping*, diperoleh sebanyak 40.588 data ulasan. Setelah dilakukan pembersihan dengan menghapus data duplikat (*duplicate removal*), jumlah ulasan berkurang menjadi 31.631 data yang unik. Tabel 1 menyajikan contoh perubahan data, mulai dari teks ulasan asli sehingga teks yang telah melalui tahap pembersihan dan siap untuk diolah lebih lanjut.

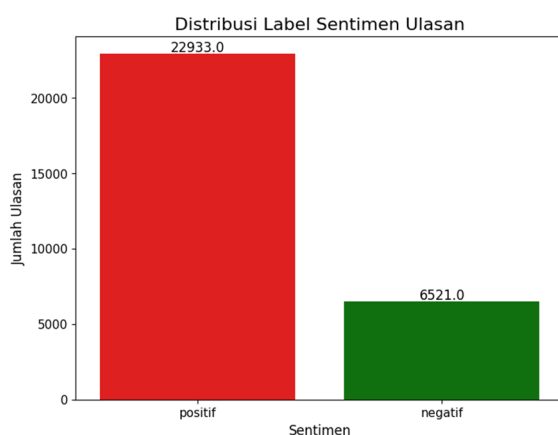
Tabel 1. Hasil Pra-Pemrosesan Data

Tahapan	Input (Data Mentah)	Output (Hasil Proses)
Ulasan Asli	"Game-nya SERU banget!! Tapi sayang,	-

Tahapan	Input (Data Mentah)	Output (Hasil Proses)
	banyak cheater di rank :("	
<i>Cleaning</i>	"Game-nya SERU banget!! Tapi sayang, banyak cheater di rank :("	"gamenya seru banget tapi sayang banyak cheater di rank"
<i>Case Folding</i>	"gamenya seru banget tapi sayang banyak cheater di rank"	"gamenya seru banget tapi sayang banyak cheater di rank"
<i>Stopword Removal</i>	"gamenya seru banget tapi sayang banyak cheater di rank"	"gamenya seru sayang banyak cheater rank"
<i>Stemming</i>	"gamenya seru sayang banyak cheater rank"	"game seru sayang banyak cheater rank"

Distribusi Sentimen

Setelah melalui proses pelabelan otomatis serta penghapusan ulasan dengan rating netral, diperoleh dataset akhir sebanyak 29.454 ulasan. Distribusi data menunjukkan adanya ketidakseimbangan (*imbalanced dataset*), di mana sentimen positif mendominasi dengan 22.933 ulasan (77,86%), sementara sentimen negatif berjumlah 6.521 ulasan (22,14%). Perbandingan distribusi sentimen tersebut ditampilkan pada Gambar 2.



Gambar 2. Distribusi Label Sentimen

Temuan ini menunjukkan adanya masalah teknis signifikan yang perlu dianalisis lebih lanjut.

Word Cloud

Berdasarkan hasil analisis *word cloud* yang ditampilkan pada gambar 3, terlihat bahwa sentimen positif didominasi oleh kata-kata seperti "bagus", "mantap", "seru" dan "grafik". Kata-kata tersebut mencerminkan kepuasan pengguna terhadap kualitas dan pengalaman bermain yang ditawarkan.



Gambar 3. Word Cloud Sentimen Positif

Sebaliknya, *word cloud* pada gambar 4 menunjukkan bahwa sentimen negatif banyak didominasi oleh kata "lag", "login", "banyak", dan "berat". Temuan ini mengindikasikan bahwa permasalahan terkait stabilitas koneksi serta beban aplikasi merupakan faktor utama yang menyebabkan ketidakpuasan atau kekecewaan para pemain.



Gambar 4. Word Cloud Sentimen Negatif

Evaluasi dan Perbandingan Model

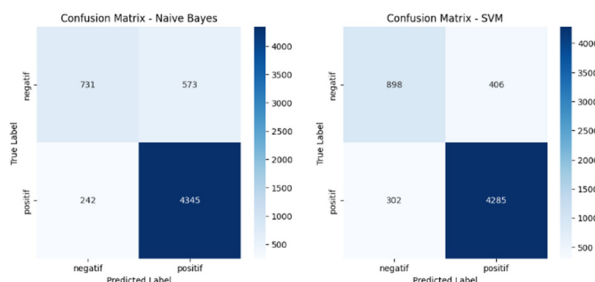
Pengujian model dilakukan menggunakan data uji sebesar 20% dari total dataset, yaitu sebanyak 5.891 data. Hasil perbandingan kinerja kedua algoritma ditampilkan pada Tabel 2, sedangkan visualisasi hasil klasifikasi disajikan dalam bentuk *Confusion Matrix* pada Gambar 5.

Berdasarkan hasil yang disajikan pada Tabel 2, dapat disimpulkan bahwa algoritma SVM menunjukkan kinerja yang lebih unggul dibandingkan dengan *Naive Bayes*. Perbedaan yang paling signifikan terlihat pada nilai recall untuk kelas negatif. SVM mampu mengidentifikasi sekitar 69% ulasan negatif atau keluhan pengguna, sedangkan *Naive Bayes* hanya dapat mengenali sekitar 56%.

Tabel 2. Perbandingan performa Algoritma

Model Algoritma	Akurasi (Accuracy)	Presisi (Negatif)	Recall (Negatif)	F1-Score (Negatif)
Naive Bayes	86,17%	0,75%	0,56%	0,64%
SVM	87,98%	0,75%	0,69%	0,72%

Hasil ini menunjukkan bahwa SVM memiliki tingkat kepekaan yang lebih tinggi dalam mendeteksi umpan balik negatif. Kemampuan tersebut sangat penting karena ulasan negatif dapat menjadi sumber informasi utama bagi pengembangan dalam melakukan evaluasi serta perbaikan terhadap produk atau layanan yang dikembangkan



Gambar 5. Confusion Matrix antara Naïve Bayes (Kiri) dan SVM (Kanan)

Berdasarkan *Confusion Matrix* pada gambar 5, dapat dilihat bahwa model *Naive Bayes* masih melakukan kesalahan yang cukup besar, yaitu mengklasifikasikan 573 ulasan negatif sebagai sentimen positif. Sebaliknya model SVM menunjukkan kinerja yang lebih baik dengan menurunkan jumlah kesalahan tersebut menjadi 406 ulasan. Hal ini mengindikasikan bahwa SVM lebih efektif dalam menangani kelas minoritas, khususnya sentimen negatif. Temuan ini juga sejalan

dengan hasil penelitian yang di laporkan oleh (Alga et al., 2024) yang menyebutkan bahwa SVM memiliki kemampuan yang lebih banyak dalam mendeteksi sentimen negatif

Perhitungan Manual Kinerja Model

Dengan tujuan mengamati akurasi dan evaluasi yang diperoleh, dilakukan perhitungannya secara manual dengan menggunakan nilai konfusif dari model SVM, dengan TP=4.285, FP=406, dan FN=302. Dalam melakukan evaluasi, dilakukan perhitungan menggunakan rumus evaluasi standar, dengan tujuan melakukan validasi dan verifikasi bahwa perhitungan model yang dipercaya.

a. Accuracy

Metrik ini untuk mengukur korelasi dari jumlah nilai perkiraan yang tepat, baik pada kelas positif maupun negatif, dalam satuan persentase terhadap nilai seluruhnya dari data yang di uji coba. Perhitungan adalah nilai yang dijumlahkan atau penjumlahan nilai TP dan TN, kemudian dibagi dengan total seluruhnya, merupakan penjumlahan nilai TP, TN, FP, dan FN

$$\begin{aligned}
 &= \frac{TP+TN}{TP+TN+FP+FN} \\
 &= \frac{4285+898}{4285+898+406+302} \\
 &= \frac{5183}{5891} \\
 &= 0,8798(87,98\%)
 \end{aligned}$$

b. Precision

Metrik ini digunakan untuk nilai sejauh mana ketepatan model dalam memprediksi ulasan yang termasuk ke dalam kelas negatif

$$\begin{aligned}
 &= \frac{TN}{TN + FN} \\
 &= \frac{898}{898 + 302} \\
 &= \frac{898}{1200}
 \end{aligned}$$

$$= 0,748 (75\%)$$

c. *Recall*

Metrik ini digunakan untuk mengukur sejauh mana model mampu mengenali ulasan negatif yang sebenarnya dari seluruh ulasan negatif yang ada dalam data.

$$= \frac{TN}{TN + FP}$$

$$= \frac{898}{898 + 406}$$

$$= \frac{898}{1304}$$

$$= 0,688 (69\%)$$

d. *F1-Score*

Metrik ini merupakan rata-rata harmonis antara nilai presisi dan *recall*, yang digunakan untuk menilai keseimbangan kinerja model dalam melakukan klasifikasi

$$= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$= 2 \times \frac{0,75 \times 0,69}{0,75 + 0,69}$$

$$= 2 \times \frac{0,5175}{1,44}$$

$$= 0,718 (72\%)$$

D. PENUTUP

Berdasarkan hasil analisis sentimen terhadap ulasan *game* Delta Force Mobile, dapat disimpulkan mayoritas pengguna memberikan tanggapan positif, khususnya terhadap sejumlah keluhan yang cukup dominan, terutama pada aspek teknis seperti masalah *lag*, gangguan jaringan, serta kendala saat proses login, yang menjadi penyumbang utama sentimen negatif.

Dari sisi perbandingan kinerja algoritma *Support Vector Machine* (SVM) menunjukkan performa yang lebih unggul dibandingkan dengan *Naive Bayes*, SVM berhasil mencapai tingkat akurasi sebesar 87,98% serta memiliki kemampuan yang lebih baik dalam mendeteksi sentimen negatif, yang tercermin dari nilai *recall* sebesar 0,69, lebih tinggi dibandingkan *Naive Bayes* yang hanya mencapai 0,56. Temuan tidak seimbang, sehingga lebih andal untuk mengidentifikasi umpan balik negatif dari pengguna tersebut.

E. DAFTAR PUSTAKA

- Alga, J., Wulandari, C., & Intan, B. (2024). Analisis Sentimen Aplikasi Youtube di Google Play Store Menggunakan Machine Learning. *Resolusi : Rekayasa Teknik Informatika Dan Informasi*, 4(4), 408–416.
<https://doi.org/10.30865/resolusi.v4i4.1750>
- Alyandi, L. O. (2025). Perbandingan Kinerja Naive Bayes dan SVM Dalam Analisis Sentimen Ulasan Free Fire di Play Store. *JATI : Jurnal Mahasiswa Teknik Informatika*, 9(3), 5127–5135.
<https://doi.org/10.36040/jati.v9i3.13599>
- Nugraha, M. T., Sulistiyowati, N., & Enri, U. (2023). Analisis Sentimen Ulasan Aplikasi Satu Sehat Pada Google Play Store Menggunakan Naive Bayes Classifier. *JATI : Jurnal Mahasiswa Teknik Informatika*, 7(5), 3593–3601.
<https://doi.org/10.36040/jati.v7i5.7753>
- Rismansyah, R. R., Sudiarjo, A., & Mufizar, T. (2025). Analisis Sentimen Ulasan Shopee Pada Google Play Store Menggunakan Algoritma Naive Bayes. *JEIS: Jurnal Elektro Dan Informatika Swadharma*, 5(1), 109–120.
<https://doi.org/10.56486/jeis.vol5no1.661>
- Sari, P. R., Indah, D. R., Rasywir, E., Firdaus, M. A., & Athalina, G. (2024). Comparison of Naive Bayes and SVM Algorithms for Sentiment Analysis of PUBG Mobile on

Google Play Store. *Sistemasi: Jurnal Sistem Informasi*, 13(6), 2767–2779.
<https://doi.org/10.32520/stmsi.v13i6.4814>

Widyanto, T., Ristiana, I., & Wibowo, A. (2023). Komparasi Naïve Bayes dan SVM Analisis Sentimen RUU Kesehatan di Twitter. *SINTECH: Science and Information Technology Journal*, 6(3), 147–161.
<https://doi.org/10.31598/sintechjournal.v6i3.1433>

Yani, F. (2025). Analisis Sentiment Pengguna Aplikasi Mobile Legend di Playstore Dengan Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM). *Djtechno: Jurnal Teknologi Informasi*, 6(2), 810–819.
<https://doi.org/10.46576/djtechno.v6i2.7313>