

PERBANDINGAN SENTIMEN ULASAN APLIKASI GEMINI PADA GOOGLE PLAYSTORE MENGGUNAKAN METODE SVM DAN NAIVE BAYES

Dimas Pungky Eprianza¹, Mochamad Agung Laksono², Aria Farhan Qur'ani³, Susliansyah⁴,
Hendro Priyono⁵

¹ 19210719@bsi.ac.id, ²19211127@bsi.ac.id, ³ 19210739@bsi.ac.id, ⁴susliansyah.slx@bsi.ac.id,
⁵Hendro.hop@bsi.ac.id

¹Program Studi Sistem Informasi, ²Fakultas Teknik dan Informatika, ³Universitas Bina Sarana
Informatika

Abstrak

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence) semakin pesat dan telah diterapkan dalam berbagai aspek kehidupan, salah satunya melalui chatbot seperti Gemini AI dari Google. Meskipun menawarkan beragam fitur menarik, beberapa pengguna menyampaikan keluhan terkait performa aplikasi ini. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap Gemini AI berdasarkan ulasan di Google Play Store. Sebanyak 1000 ulasan dikumpulkan melalui proses scraping dan dianalisis menggunakan algoritma Support Vector Machine (SVM) dan Naive Bayes. Tahapan penelitian meliputi preprocessing, pembobotan TF-IDF, pembagian data, penerapan algoritma, serta evaluasi model menggunakan Google Colab. Hasil penelitian menunjukkan bahwa metode Naive Bayes memperoleh akurasi tertinggi sebesar 82,4%, sedangkan SVM mencapai 81,4%. Secara keseluruhan, kedua metode menunjukkan kinerja yang lebih baik dalam mengklasifikasikan sentimen positif dibandingkan negatif.

Kata kunci: *Artificial Intelligence*, Analisis Sentimen, Gemini AI, *Support Vector Machine* (SVM), *Naive Bayes*

Abstract

The development of artificial intelligence (AI) technology is progressing rapidly and has been applied in various aspects of life, one of which is through chatbots like Google's Gemini AI. Although it offers a variety of interesting features, some users have expressed complaints regarding the performance of this application. This study aims to analyze user sentiment towards Gemini AI based on reviews on the Google Play Store. A total of 1,000 reviews were collected through a scraping process and analyzed using Support Vector Machine (SVM) and Naive Bayes algorithms. The research stages included preprocessing, TF-IDF weighting, data splitting, algorithm implementation, and model evaluation using Google Colab. The study results show that the Naive Bayes method achieved the highest accuracy of 82.4%, while SVM reached 81.4%. Overall, both methods demonstrated better performance in classifying positive sentiment compared to negative sentiment.

Keywords: *Artificial Intelligence*, Analisis Sentimen, Gemini AI, *Support Vector Machine* (SVM), *Naive Bayes*

1. Pendahuluan

Artificial Intelligence (AI), juga dikenal dalam bahasa Indonesia sebagai Kecerdasan Buatan, merupakan cabang ilmu komputer yang berfokus pada pengembangan sistem atau mesin yang mampu menjalankan tugas-tugas yang biasanya membutuhkan kecerdasan manusia. AI bekerja dengan memanfaatkan komputer dan sistem berbasis algoritma serta model matematika untuk mempelajari data, mengenali pola, dan mengambil keputusan secara otomatis[1].

Kehadiran Artificial Intelligence (AI) telah membawa perubahan signifikan di berbagai bidang kehidupan. Salah satu inovasi yang menarik dari kecerdasan buatan adalah kemampuan untuk berkomunikasi antara manusia dan mesin yang disebut chatbot. Chatbot adalah program komputer yang dapat berinteraksi dengan manusia melalui internet seperti ChatGPT, Deepseek, ClaudeAI dan Gemini

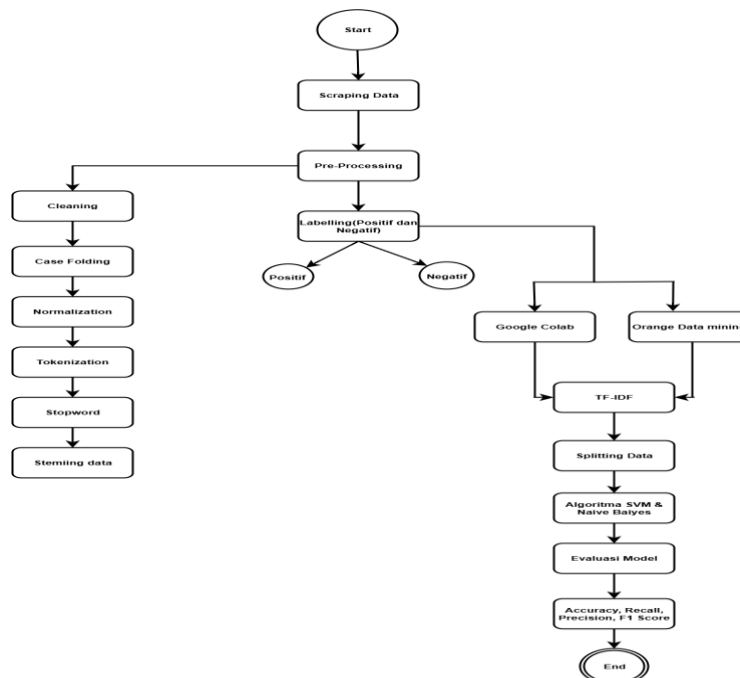
AI[2]. Salah satu chatbot yang banyak dimanfaatkan ialah Gemini AI salah satu kecerdasan buatan yang banyak digunakan saat ini. Yang Dikembangkan oleh Google Deep Mind, aplikasi ini dirancang dengan teknologi multimodal untuk memproses berbagai jenis data seperti teks, gambar, audio, video, dan kode program[3]. Meskipun Gemini AI menawarkan berbagai keunggulan, banyak pengguna mengeluhkan respons yang kurang akurat, khususnya pada pertanyaan teknis seperti pemrograman. Beberapa ulasan juga menyebutkan jawaban yang tidak relevan dengan pertanyaan utama. Hal ini menunjukkan adanya kesenjangan antara ekspektasi dan pengalaman pengguna, sehingga analisis sentimen diperlukan untuk menilai tingkat kepuasan terhadap kinerja Gemini AI.

Penelitian ini berfokus pada penerapan analisis sentimen menggunakan pendekatan machine learning. Analisis sentimen atau opinion mining merupakan cabang ilmu yang fokus pada pengkajian ekspresi sentimen, opini, maupun emosi yang dituangkan dalam bentuk teks[4]. Berdasarkan penelitian sebelumnya Penelitian oleh Pamungkas & Cahyono menunjukkan bahwa algoritma SVM memiliki akurasi 80%, sedangkan algoritma KNN memiliki akurasi 71% [5]. Lalu oleh Najibulloh menunjukkan bahwa metode Neural Network (NN) memiliki akurasi 95,5%, SVM sebesar 95,4%, dan Decision Tree (DT) sebesar 92,1% [6]. Dan penelitian oleh juga menemukan bahwa metode SVM memiliki akurasi lebih tinggi dari Naive Bayes, terutama dalam mendeteksi kedua kelas sentimen yang memiliki akurasi Hasil evaluasi menunjukkan SVM unggul dibandingkan Naive Bayes dalam akurasi 96,54% pada SVM dan pada Naive Bayes 90,32%

Untuk menganalisis sentimen diatas dibutuhkan suatu metode yaitu menggunakan metode Algoritma Support Vector Machine (SVM) dan Naive Bayes. Melalui pendekatan ini pada dua metode ini, Penelitian diharapkan dapat memberikan kontribusi dalam bentuk evaluasi sistematis terhadap pandangan pengguna terhadap Gemini AI, sekaligus menjadi acuan bagi pengembang aplikasi untuk menjadi lebih baik depannya. Dengan membandingkan dua metode klasifikasi tersebut berdasarkan akurasi, presisi, recall, dan F1-score[7] Penelitian ini diharapkan dapat menentukan metode optimal dalam mengukur kepuasan pengguna dan efektivitas Gemini AI, sekaligus memberikan kontribusi akademik serta masukan praktis bagi pengembang untuk meningkatkan kualitas respons dan pengalaman pengguna

2. Metode

Gambar 1 menyajikan gambaran mengenai tahapan penelitian yang dilakukan



Gambar 1
Kerangka Penelitian

2.1. Scrapping Data

Teknik *scraping* digunakan untuk mendapatkan data atau informasi dari situs *web* tertentu. Ini dapat berupa teks, tautan, video, audio, atau dokumen[8]. Untuk *scraping* ini dilakukan dalam batas-batas tertentu. Adapun batasan yang diambil pada Penelitian ini *scraping* data ulasan pada aplikasi Gemini AI dalam *scraping* ini mengambil data berupa Username, Rating, Review Text/Ulasan, Tanggal ulasan.

2.2. Preprocessing Data

Preprocessing adalah tahap awal proses pengolahan teks yang bertujuan untuk mengubah data menjadi format yang sesuai untuk digunakan saat siap untuk analisis[9]. Pada *preprocessing* ini meliputi:

- a. *Cleaning*: sebuah proses untuk menghapus elemen mengganggu agar data lebih bersih dan bisa dianalisis.
- b. *Case folding*: merupakan proses untuk mengubah semua huruf teks menjadi huruf kecil (*lowercase*).
- c. *Normalization*: merupakan suatu proses transformasi data numerik ke dalam skala tertentu
- d. *Tokenization*: merupakan sebuah proses pemisah kalimat dan menjadi perkata di sebuah data.
- e. *Stopword Removal*: merupakan ahapan dalam pemrosesan teks yang bertujuan untuk menghapus kata-kata yang sering muncul namun tidak memiliki nilai penting dalam proses analisis.
- f. *Stemming*: merupakan proses menghapus kata awal dan kata akhir yang nantinya menghasilkan kata dasar.

2.3. Labelling Sentimen

Pada proses Labeling sentimen ini menggunakan Label positif dan label negatif pada proses analisis sentimen yang berguna untuk melatih dan penerapan metode yang digunakan dimana proses ini penting untuk mendapatkan hasil data yang akurat[10]. Pada penelitian ini labelling digunakan untuk memudahkan klasifikasi dan analisis sentimen ulasan pada Aplikasi Gemini AI.

2.4. Term Frequency-Inverse Document Frequency (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) ialah salah satu teknik dalam pemrosesan teks dan pemodelan bahasa alami yang berfungsi untuk mengukur tingkat kepentingan suatu kata (term) dalam sebuah dokumen, relatif terhadap keseluruhan koleksi dokumen yang dianalisis[11]. Adapun rumus Pada Tf-IDF sebagai berikut:

$$TF(t) = \frac{\text{Jumlah Kata (t) dalam komentar}}{\text{Jumlah total kata dalam komentar}}$$
$$IDF = \frac{1 + \text{Total Dokumen Frekuensi (N)}}{1 + \text{Jumlah Total Kata (DF)}} + 1$$

Keterangan:

$$TF-IDF = TF \times IDF$$

TF : Term Frequency

DF : Document Frequency

IDF : Inverse Document Frequency

(t) : Term atau sebuah kata dalam komentar

N : Jumlah total dokumen

2.5. Klasifikasi Support Vector Machine & Naive Bayes

Proses dengan metode Support Vector Machine (SVM) diawali dengan memasukkan data yang telah dipilih. Selanjutnya, data tersebut melalui tahap klasifikasi menggunakan SVM untuk memperoleh

nilai akurasi. Tahapan awal dalam proses ini yaitu mengidentifikasi karakteristik data yang dijadikan parameter klasifikasi, kemudian dilanjutkan dengan perhitungan akurasi dari hasil klasifikasi tersebut[12]. Adapun rumus *Support Vector Machine (Linear)*:

$$f(x)=w.x+b$$

Keterangan:

w: vektor bobot (weight)

x: vektor fitur input (hasil TF-IDF)

b: *bias*

f(x): fungsi keputusan

Naive Bayes adalah sebuah klasifikasi berbasis probabilitas yang sederhana, yang bekerja dengan menghitung peluang berdasarkan frekuensi dan kombinasi nilai pada dataset yang tersedia. Algoritma ini didasarkan pada Teorema Bayes bahwa setiap atribut tidak saling bergantung satu sama lain terhadap nilai pada variabel kelas[13]. Adapun rumus Naive Bayes:

$$P(H|X) = \frac{P(X | H) \cdot P(H)}{P(X)}$$

Keterangan:

X = Data dengan class yang belum diketahui

H = Hipotesis data X dari suatu kelas yang spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi dari X (Posteriori Probality)

P(H) = Probabilitas hipotesis H (Prior Probability)

P(X|H) = Probabilitas X berdasarkan Hipotesis H

P(X) = Probabilitas dari X

2.6. Evaluasi Model

Evaluasi model adalah untuk mengukur baik kinerja atau model dalam menyelesaikan tugas yang telah diterapkan, seperti klasifikasi, regresi, atau analisis sentimen. Evaluasi model dilakukan dengan menggunakan Confusion Matrix, yaitu Confusion Matrix merupakan perbandingan hasil dari prediksi dengan kelas data sebenarnya, lalu menghitung jumlah prediksi yang benar dan salah dari setiap kategori kelas[14]. Confusion Matrix terdiri dari beberapa komponen yaitu:

- a. True Positive (TP): nilai prediksi benar dan nilai sebenarnya benar.
- b. True Negative (TN): nilai prediksi salah dan nilai sebenarnya salah.
- c. False Positive (FP): nilai prediksi benar dan nilai sebenarnya salah.
- d. False Negative (FN): nilai prediksi salah dan nilai sebenarnya benar

Tabel 1. Confusion Matrix

	<i>Predicted Positive</i>	<i>Predicted Negative</i>
<i>Actual Positive</i>	TP	FN
<i>Actual Negative</i>	FP	TN

Confusion matrix bisa dihitung matrix evaluasi seperti tingkat akurasi, presisi, recall, dan F1 score yang nantinya menunjukkan kinerja pada model klasifikasi yang ditentukan.

- a. Accuracy adalah rasio prediksi yang benar terhadap seluruh jumlah data yang diuji.

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

b. *Recall* mengukur seberapa baik model dalam mendeteksi kelas positif yang sebenarnya.

$$\text{Recall (Kelas Positif)} = \frac{TP}{TP+FN}$$

$$\text{Recall (Kelas Negatif)} = \frac{TN}{TN+FP}$$

c. *Precision* mengukur seberapa banyak dari prediksi positif yang benar-benar positif.

$$\text{Precision (Kelas Positif)} = \frac{TP}{TP+FP}$$

$$\text{Precision (Kelas Negatif)} = \frac{TN}{TN+FN}$$

d. *F1 Score* adalah rata-rata harmonis dari *precision* dan *recall*. Metrik ini digunakan untuk menyeimbangkan keduanya.

$$F1 \text{ Score (Positif)} = 2 \times \frac{\text{Precision(Positif)} \times \text{Recall(Positif)}}{\text{Precision (Positif)} + \text{Recall(Positif)}}$$

$$F1 \text{ Score (Negatif)} = 2 \times \frac{\text{Precision(Negatif)} \times \text{Recall(Negatif)}}{\text{Precision(Negatif)} + \text{Recall(Negatif)}}$$

3. Hasil dan Pembahasan (size 11)

Penelitian ini diawali dengan pengambilan data ulasan aplikasi Gemini AI melalui proses scraping di Google Colab. Data kemudian diproses melalui tahapan preprocessing (cleaning, case folding, normalisasi, tokenisasi, penghapusan stopwords, dan stemming) serta dilakukan pelabelan sentimen positif dan negatif. Dataset hasil pelabelan diimpor ke Orange Data Mining untuk penerapan pembobotan kata dengan metode TF-IDF, lalu dibagi menjadi data latih dan uji. Tahap utama penelitian menggunakan algoritma SVM dan Naive Bayes untuk klasifikasi sentimen, dengan evaluasi kinerja melalui metrik accuracy, recall, precision, dan F1-score.

3.1. Pengumpulan Data

Untuk pengumpulan data ini dilakukan dengan teknik scraping yang menggunakan bahasa python dengan menggunakan library googleplayscraper. Untuk pengambilan data set ini diambil pada tanggal 01 Mei 2024 – 11 Mei 2025 sebanyak 1000 data. Dari data tersebut akan menampilkan hasil dari scraping yang berisi berbagai ulasan yang diberikan oleh user dari aplikasi Gemini AI yang ada di playstore yang ditunjukkan pada tabel 2.

Table 2. Hasil Scrapping Data

No	Nama	Rating	Ulasan	Tanggal Ulasan
1	Hahan Dirdja Subagdja	5	sajauh ini bagus, semoga terus bagus biar bayak penggunanya	11/06/2025 17:34:19
2.	Arman William	5	Sangat Berguna dan bermanfaat. mantap. Dua	24/05/2025 22:37:56

jempol untuk developer,

3.	Galuh Ika Permadi	4	mudah di pakai dan sangat membantu.	11/06/2025 13:42:29
⋮	⋮	⋮	⋮	⋮
1000	Muh Amar	5	sangat membantu dan tentunya bikin bangga.	15/04/2025 23:59:57

Dari Tabel 2 ini dapat Hasil dari scraping data yaitu ulasan mentah yang diambil dari Google Play Store. Data tersebut masih belum bersih ditunjukkan pada Tabel 2. Dalam ini menunjukkan bahwa hasil scraping diperlukan proses preprocessing agar data tersebut bisa diolah.

Tabel 3. Hasil Drop Duplicates

No	Nama	Rating	Ulasan	Tanggal Ulasan
1	Hahan Dirdja Subagdja	5	sajauh ini bagus, semoga terus bagus biar bayak penggunanya	11/06/2025 17:34:19
2.	Yans Lombok	3	berlangganan ke advance dapet 3 generate video per 24 jam wkwk	11/06/2025 14:53:03
3.	Galuh Ika Permadi	4	mudah di pakai dan sangat membantu.	11/06/2025 13:42:29
⋮	⋮	⋮	⋮	⋮
999	Metta Yana	5	sangat membantu buat ngerjain tugas.	11/04/2025 12:18:06

Drop duplicates dilakukan untuk menghilangkan data ulasan yang semula 1.000 menjadi 999 agar data ulasan tidak menjadi duplicate yang ditunjukkan pada tabel 3

3.2. Preprocessing Data

Pada tahap preprocessing data ini untuk mempersiapkan proses agar siap digunakan untuk menganalisis data ataupun pemodelan dengan tujuan agar data lebih terstruktur. Maka dari itu memerlukan beberapa tahapan proses seperti Cleaning membersihkan teks dari karakter yang tidak diperlukan seperti angka, tanda baca, URL, dan simbol agar data lebih bersih dan terstruktur, Case Folding mengubah seluruh huruf menjadi huruf kecil (*lowercase*) untuk menyamakan format penulisan kata, Normalization mengubah kata tidak baku atau singkatan ke bentuk baku agar konsisten tokenization Memecah teks menjadi potongan kecil (token) seperti kata atau frasa untuk memudahkan analisis, Stopword Removal menghapus kata umum yang tidak berpengaruh terhadap makna, Stemming mengubah kata ke bentuk dasar dengan menghilangkan imbuhan.

Tabel 4. Hasil Preprocessing Text

Ulasan	Cleaning	Case folding	Normalizati on	Tokenization	Stopword Removal	Stemmin g
sajauh ini bagus, semoga terus bagus biar banyak penggunanya	sajauh ini bagus semoga terus bagus biar banyak penggunaan ya	sajauh ini bagus semoga terus bagus biar banyak penggunaan ya	sajauh ini bagus semoga terus bagus biar banyak penggunaan ya	['sajauh', 'ini', 'bagus', 'semoga', 'terus', 'bagus', 'biar', 'banyak', 'penggunanya']	['sajauh', 'bagus', 'semoga', 'bagus', 'biar', 'banyak', 'penggunanya']	sajauh bagus moga bagus biar banyak guna
Sangat Berguna dan bermanfaat. Dua jempol untuk developer,	Sangat Berguna dan bermanfaat. Dua jempol untuk developer	sangat berguna dan bermanfaat. dua jempol untuk developer	sangat berguna dan bermanfaat. dua jempol untuk developer	['sangat', 'berguna', 'dan', 'bermanfaat', 'jempol', 'dua', 'untuk', 'developer']	['berguna', 'bermanfaat', 'jempol', 'developer']	guna manfaat mantap jempol develope r
Gemini payah tidak bisa menjawab pertanyaan atau soal mata pelajaran	Gemini payah tidak bisa menjawab pertanyaan atau soal mata pelajaran	gemini payah tidak bisa menjawab pertanyaan atau soal mata pelajaran	gemini payah tidak bisa menjawab pertanyaan atau soal mata pelajaran	['gemini', 'payah', 'tidak', 'bisa', 'menjawab', 'pertanyaan', 'atau', 'soal', 'mata', 'pelajaran']	['gemini', 'payah', 'mata', 'pelajaran']	gemini payah mata ajar

3.3. Hasil Labelling Sentimen

pada proses Labelling Sentimen ini dilakukan berdasarkan rating dimana rating dua dikategorikan sebagai label “Negatif” dan rating selain daripada dua maka akan dikategorikan sebagai label “Positif” dimana proses labelling ini sangat penting untuk menentukan isi dari review text berdasarkan rating yang ditunjukkan pada tabel 5

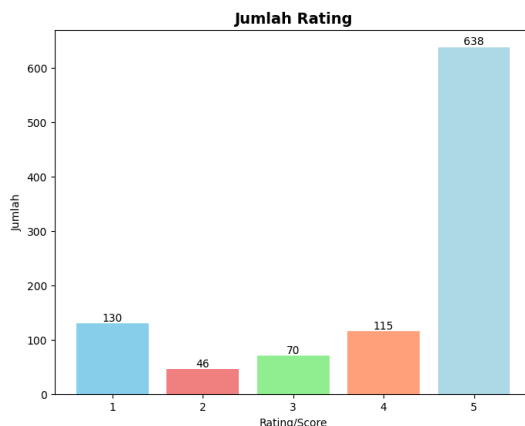
Tabel 5. Hasil Labelling Sentimen

Ulasan	Rating	Sentimen
gua suka bug bug ya bikin kesal woi	2	Negatif
wow aplikasi bantu tugas request apa pokok keren	5	Positif
responsif relevan sesuai	5	Positif

```
Sentiment
Positif    818
Negatif    176
Name: count, dtype: int64
```

Gambar 2. Jumlah Labelling Sentimen

Berikut ini adalah hasil dari jumlah data positif dan negatif yang dimana data positif berjumlah 818 dan data negatif berjumlah 176, langkah ini penting dalam analisis sentimen selanjutnya.



Gambar 3. Jumlah Rating Sentimen

Berdasarkan grafik di atas, mayoritas pengguna memberikan rating 3-5, yang mengindikasikan kepuasan terhadap aplikasi Gemini AI. Meskipun demikian, masih terdapat sebagian pengguna yang memberikan rating 1-2, menunjukkan adanya keluhan atau ketidakpuasan. Secara umum, ulasan dengan sentimen positif lebih mendominasi dibandingkan dengan sentimen negatif.

3.4. Hasil Term Frequency-Inverse Document Frequency (TF-IDF)

Pada titik ini, pembobotan kata dilakukan secara manual menggunakan metode Term Frequency (TF) dan Term Frequency Inverse Document Frequency (TF-IDF). Tujuannya untuk menilai seberapa penting sebuah kata dalam sebuah ulasan, dibandingkan dengan kumpulan ulasan lain. salah satu contoh ulasan dari review text yang akan digunakan untuk pembobotan kata “aplikasi bagus banget update ulang ya”.

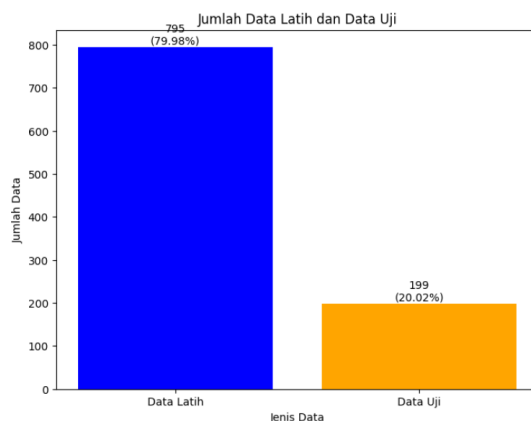
Tabel 6. Hasil Perhitungan TF-IDF

Term	TF	DF	IDF	TF-IDF
aplikasi	$1/6 = 0,16$	193	5,231	0,836
bagus	$1/6 = 0,16$	285	3,870	0,619
banget	$1/6 = 0,16$	118	7,899	1,263
update	$1/6 = 0,16$	39	21.52	3,443
ulang	$1/6 = 0,16$	18	44.21	7,073
ya	$1/6 = 0,16$	167	5,886	0,941

Hasil pembobotan TF-IDF berdasarla memperlihatkan bahwa kata “ulang” memperoleh nilai tertinggi sebesar 7,073. Hal ini menunjukkan bahwa meskipun jarang muncul pada keseluruhan korpus,

kata tersebut memiliki tingkat kepentingan yang tinggi dalam 1 ulasan tertentu. Sebaliknya, kata “bagus” memiliki nilai lebih rendah (0,619) karena frekuensinya yang tinggi pada berbagai dokumen, sehingga tingkat kekhususannya menjadi rendah. Adapun kata “update” (3,443) dan “banget” (1,263) juga menunjukkan nilai TF-IDF yang relatif tinggi, menandakan kemunculan yang lebih jarang namun signifikan terhadap makna ulasan.

3.5. Splitting Data

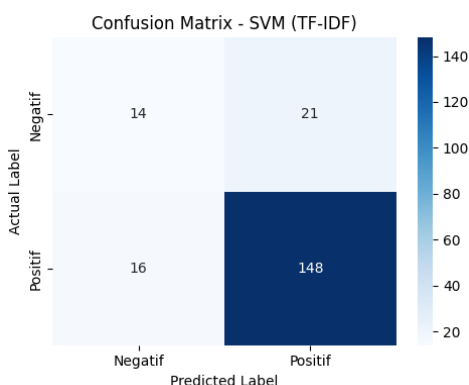


Gambar 4. Hasil Splitting Data

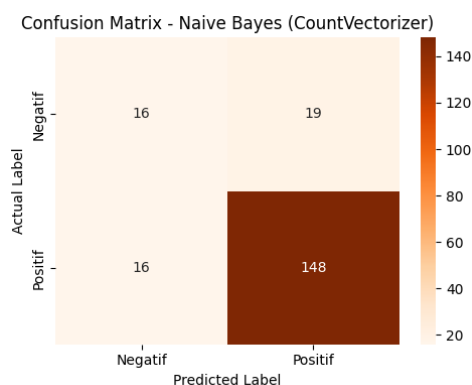
Grafik di atas menunjukkan hasil pembagian dataset menjadi dua bagian, yaitu data latih sebanyak 795 data (79,98%) yang ditampilkan dengan warna biru, dan data uji sebanyak 199 data (20,02%) yang ditampilkan dengan warna oranye. Visualisasi ini bertujuan untuk memastikan proporsi pembagian data sesuai dengan parameter yang telah ditentukan, yakni 80% untuk pelatihan dan 20% untuk pengujian.

3.6. Hasil Klasifikasi SVM & Naive Bayes

Evaluasi model dilakukan untuk mengukur performa klasifikasi sentimen yang dihasilkan oleh algoritma SVM & Nave Bayes. Metrik yang digunakan meliputi accuracy, precision, recall, dan f1-score sebagai indikator utama. Confusion matrix disajikan untuk melihat perbandingan antara hasil prediksi dan data sebenarnya.



Gambar 5. Confusion Matrix SVM



Gambar 6. Confusion Matrix Naive Bayes

Perbandingan antara *Confusion Matrix* SVM dan Naive Bayes yang ditunjukkan pada gambar 5 dan 6 menunjukkan bahwa SVM Dari 199 data uji, sebanyak 148 data positif berhasil diklasifikasikan dengan benar, sedangkan 16 data positif salah diklasifikasikan sebagai negatif. Untuk data negatif, hanya 14 yang diprediksi dengan benar, sementara 21 data negatif salah diklasifikasikan sebagai positif. Hal ini

menunjukkan bahwa model lebih baik dalam mengenali sentimen positif dibandingkan sentimen negatif. Meskipun akurasi model mencapai sekitar 81,41%. Sedangkan pada Naive Bayes Model berhasil mengidentifikasi 148 data positif dengan tepat (True Positive) dan 16 data negatif dengan benar (True Negative). Tetapi, terdapat kesalahan klasifikasi pada 19 data negatif yang terdeteksi sebagai positif (False Positive) dan 16 data positif yang terbaca sebagai negatif (False Negative).

```
Confusion Matrix:
[[ 14 21]
 [ 16 148]]

Classification Report:
      precision    recall  f1-score   support

   Negatif      0.47      0.40      0.43         35
   Positif      0.88      0.90      0.89        164

 accuracy          0.81         199
 macro avg          0.67      0.65      0.66         199
 weighted avg       0.80      0.81      0.81         199

Accuracy: 0.8140703517587939
```

Gambar 7. Evaluasi Kinerja Model SVM

```
Confusion Matrix:
[[ 16 19]
 [ 16 148]]

Classification Report:
      precision    recall  f1-score   support

   Negatif      0.50      0.46      0.48         35
   Positif      0.89      0.90      0.89        164

 accuracy          0.82         199
 macro avg          0.69      0.68      0.69         199
 weighted avg       0.82      0.82      0.82         199

Accuracy: 0.8241206030150754
```

Gambar 8. Evaluasi Kinerja Model Naive Bayes

Perbandingan evaluasi kinerja pada model SVM dan Naive Bayes yang ditunjukkan pada gambar 7 dan 8 menunjukkan bahwa hasil evaluasi model SVM menunjukkan akurasi model sebesar 81,4%. Model memiliki kinerja sangat baik dalam mendeteksi ulasan positif precision 88%, recall 90,2%, F1-score 0,90, tetapi kurang optimal pada ulasan negatif precision 47%, recall 40%, F1-score 0,43. Hal ini mengindikasikan ketidakseimbangan performa, di mana model lebih dominan mengenali pola sentimen positif dibandingkan negatif. Sedangkan hasil evaluasi pada model Naive Bayes menghasilkan akurasi 82,4%. Pada kelas positif, performanya tinggi dengan precision 88%, recall 90,2%, dan F1-score 0,89, menunjukkan konsistensi dalam mengenali ulasan positif. Namun, pada kelas negatif, precision 50%, recall 45%, dan F1-score 0,47 menandakan kesulitan dalam mengidentifikasi ulasan negatif. Secara keseluruhan, model lebih optimal menangani sentimen positif dibandingkan negatif.

4. Kesimpulan dan Saran

Penelitian ini menganalisis sentimen pengguna terhadap aplikasi Gemini AI berdasarkan 1000 ulasan Google Play Store menggunakan metode SVM dan Naive Bayes. Proses dilakukan melalui tahapan preprocessing, pembobotan TF-IDF, serta pembagian data latih dan uji. Hasil evaluasi menunjukkan Naive Bayes memperoleh akurasi tertinggi 82,4%, sedangkan SVM mencapai 81,4%. Kedua metode menunjukkan kinerja yang lebih baik dalam mengenali sentimen positif dibandingkan negatif. Secara keseluruhan, Naive Bayes lebih unggul dalam klasifikasi sentimen pada penelitian ini.

Untuk penelitian selanjutnya, diharapkan menambahkan algoritma lain seperti Random Forest, Logistic Regression, LSTM, atau BERT untuk perbandingan. Selain itu, penerapan teknik balancing data seperti SMOTE atau undersampling diperlukan agar performa model terhadap kelas minoritas, khususnya sentimen negatif, dapat ditingkatkan.

Daftar Pustaka

- [1] Eriana, E. S., & Zein, D. A. (2023). Artificial Intelligence. *Angewandte Chemie International Edition*, 6(11), 1.
- [2] A. D. Pratama and H. Hendry, "Analisa Sentimen Masyarakat Terhadap Penggunaan Chatgpt Menggunakan Metode Support Vector Machine (Svm)," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 1, pp. 327–338, 2024, doi: 10.29100/jupi.v9i1.4285.
- [3] R. Maulida, "Komparasi Respons ChatGPT dan Gemini terhadap Command Pattern Identik dengan Metode Black Box," vol. X, no. 2, pp. 2442–2445, 2024.
- [4] O. Manullang and C. Prianto, "Analisis Sentimen dalam Memprediksi Hasil Pemilu Presiden dan

-
- Wakil Presiden : Systematic Literature Review,” *J. Inform. Dan Teknol. Komput.*, vol. 4, no. 2, pp. 104–113, 2023, [Online]. Available: <https://ejurnalunsam.id/index.php/jicom/>
- [5] A. S. Pamungkas and N. Cahyono, “Analisis Sentimen Review ChatGPT di Play Store menggunakan Support Vector Machine dan K-Nearest Neighbor,” *Edumatic J. Pendidik. Inform.*, vol. 8, no. 1, pp. 1–10, 2024, doi: 10.29408/edumatic.v8i1.24114.
- [6] V. No, I. K. Najibulloh, D. Intan, and S. Saputra, “Edumatic : Jurnal Pendidikan Informatika Analisis Sentimen Ulasan Co-Pilot Google Play dengan SVM , Neural Network , dan Decision Tree,” vol. 9, no. 1, pp. 275–283, 2025, doi: 10.29408/edumatic.v9i1.29673.
- [7] A. Of, H. Disease, P. Using, A. C. Study, and M. L. Algorithms, “ANALISIS PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN PERBANDINGAN ALGORITMA MACHINE LEARNING ANALYSIS OF HEART DISEASE PREDICTION USING A COMPARATIVE STUDY OF,” vol. 4, no. 2, pp. 262–269, 2025.
- [8] S. Kusumo, “Penerapan Web Scraping Deskripsi Produk Menggunakan Selenium Python Dan Framework Laravel,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 4, pp. 3426–3435, 2022, doi: 10.35957/jatisi.v9i4.2727.
- [9] H. H. Mubaroroh, H. Yasin, and A. Rusgiyono, “Analisis Sentimen Data Ulasan Aplikasi Ruangguru Pada Situs Google Play Menggunakan Algoritma Naive Bayes Classifier Dengan Normalisasi Kata Levenshtein Distance,” *J. Gaussian*, vol. 11, no. 2, pp. 248–257, 2022, doi: 10.14710/j.gauss.v11i2.35472.
- [10] D. Moldovan, “A majority voting framework for reliable sentiment analysis of product reviews,” *PeerJ Comput. Sci.*, vol. 11, pp. 1–31, 2025, doi: 10.7717/peerj-cs.2738.
- [11] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *SINTESIA J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 2, pp. 81–88, 2023.
- [12] O. : Muhammad *et al.*, “Analisa Algoritma Support Vector Machine Pada Data Bunga Iris,” vol. 30, no. 1, pp. 477–487, 2022.
- [13] M. Iqbal, “PENERAPAN METODE NAIVE BAYES PADA ULASAN APLIKASI REELSHORT,” vol. 10, no. 1, pp. 49–59, 2025.
- [14] T. Gori, A. Sunyoto, and H. Al Fatta, “Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 1, pp. 215–224, 2024, doi: 10.25126/jtiik.20241118074.