

Analisis Sentimen Ulasan Pengguna Aplikasi Bank Jago pada Google Play Store dengan Metode Naïve Bayes

Rena Destiyana Segarani^{*1}, Dwi Hartanti², Mira Erlinawati³

^{1,2,3}Teknik Informatika, Fakultas Ilmu Komputer, Universitas Duta Bangsa Surakarta

Email: ^{*1}210103152@mhs.udb.ac.id, ²dwihartanti@udb.ac.id, ³mira_erlinawati@udb.ac.id

(Naskah masuk: 16 Mei 2025, diterima untuk diterbitkan: 30 Desember 2025)

Abstrak: kemajuan teknologi digital telah mendorong peningkatan pengguna layanan perbankan berbasis aplikasi, termasuk Bank Jago sebagai salah satu innovator di sektor perbankan digital Indonesia. Ulasan pengguna di Google play Store menjadi sumber data yang kaya untuk memahami persepsi, kepuasan, serta keluhan pengguna terhadap aplikasi tersebut. Penelitian ini bertujuan untuk mengkaji sentimen dari ulasan-ulasan tersebut dengan menerapkan metode Naïve Bayes, salah satu algoritma klasifikasi dalam pembelajaran mesin yang dikenal sederhana namun efektif. Data dikumpulkan dari Google Play Store dan diproses melalui tahap pra-pemrosesan yang mencakup pembersihan teks, tokenisasi, penghapusan stopword, serta stemming. Sentiment kemudian dikategorikan ke dalam tiga kelas: positif, negative dan netral. Hasil analisis menunjukkan dominasi sentimen positif, namun juga ditemukan kritik terhadap aspek teknis dan fitur layanan. Model Naïve Bayes berhasil mengklasifikasikan data dengan tingkat akurasi yang baik, menjadikannya alat yang potensial untuk menganalisis opini public secara otomatis. Penelitian ini dapat menjadi masukan bagi pengembang untuk meningkatkan kualitas aplikasi berdasarkan umpan balik pengguna.

Kata Kunci – Analisis Sentimen; Naïve Bayes; Ulasan Pengguna; Bank Jago; Google Play Store

Sentiment Analysis Of User Reviews Of Bank Jago Application On Google Play Store Using The Naïve bayes Method

Abstract: The advancement of digital technology has driven an increase in users of application-based banking services, including Bank Jago as one of the innovators in the Indonesian digital banking sector. User reviews on the Google Play Store are a rich source of data to understand user perceptions, satisfaction, and complaints about the application. This study aims to examine the sentiment of these reviews by applying the Naïve Bayes method, one of the classification algorithms in machine learning that is known to be simple but effective. Data was collected from the Google Play Store and processed through a pre-processing stage that includes text cleaning, tokenization, stopword removal, and stemming. Sentiment was then categorized into three classes: positive, negative and neutral. The results of the analysis showed a dominance of positive sentiment, but criticism of technical aspects and service features was also found. The Naïve Bayes model successfully classified data with a good level of accuracy, making it a potential tool for analyzing public opinion automatically. This study can be input for developers to improve the quality of applications based on user feedback.

Keywords – Sentiment Analysis; Naïve Bayes; User Reviews; Bank Jago; Google Play Store

1. PENDAHULUAN

Sektor perbankan saat ini tengah mengembangkan penggunaan teknologi informasi dengan meningkatkan standar kualitas layanan agar dapat bersaing di pasar era digital yang semakin kompetitif. Nasabah saat ini menggunakan aplikasi *mobile banking* (*m-banking*) yang ditawarkan oleh berbagai organisasi perbankan. Meskipun demikian, bank harus mengumpulkan data tentang layanan yang dibutuhkan nasabahnya [1]. Dengan demikian, aplikasi mobile seperti Bank Jago menjadi semakin relevan dalam mempermudah proses transaksi. Bank Jago merupakan sebuah aplikasi yang dibuat guna membantu penggunanya dalam melakukan transaksi secara online menggunakan smartphone tanpa harus pergi ke bank. Aplikasi ini juga memberikan beberapa fitur seperti tabungan, belanja atau pembayaran tagihan sehingga pengguna lebih bisa mengatur pengelolaan uang.

Aplikasi perbankan digital didistribusikan melalui platform Google Play Store, yang memungkinkan nasabah untuk mengakses proses transaksi hanya dengan telepon pintar, menawarkan kemudahan dan aksesibilitas dari lokasi mana pun. Aplikasi tersebut harus diunduh dari Google Play Store agar setiap pengguna dapat memanfaatkan kemampuannya. Program tersebut dibuat untuk beroperasi pada perangkat yang menggunakan sistem operasi Android. Ulasan pengguna dapat dilihat di Google Play Store sebagai peringkat dan komentar di kolom yang dapat diakses. Penyedia layanan dapat meningkatkan kinerja aplikasi dengan bantuan ulasan pelanggan ini. Perusahaan perbankan dapat menggunakan pemrosesan data ulasan sebagai dasar dalam membuat aplikasi dan memanfaatkan kemampuan yang ditawarkan kepada konsumen. Namun, tantangannya adalah banyaknya volume data ulasan yang perlu ditangani secara manual, yang membutuhkan waktu dan peralatan yang sesuai. Analisis sentimen, yang memastikan nada umpan balik pengguna pada aplikasi perbankan digital, adalah cara terbaik untuk mengatasi masalah ini [2].

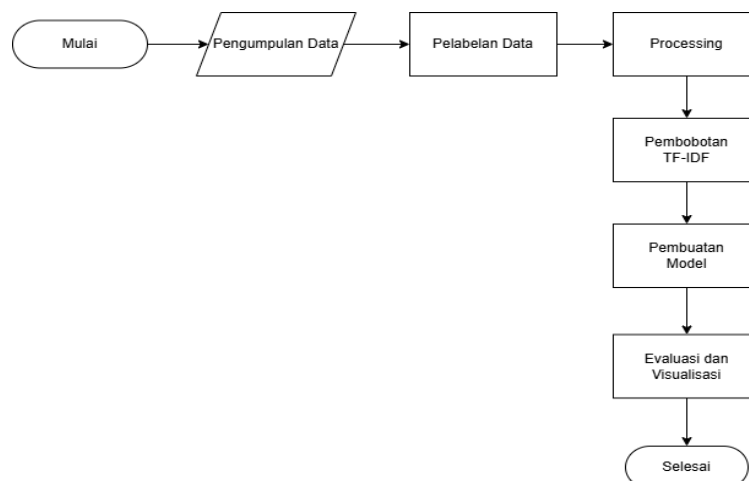
Dalam situasi ini, perusahaan membutuhkan cara untuk memeriksa perasaan dari ulasan aplikasi yang diterima. Supaya perusahaan bisa mengetahui apakah ulasan itu baik atau buruk, analisis perasaan diperlukan. Salah satu cara yang biasa dipakai untuk meneliti data dalam jumlah besar adalah analisis perasaan. Metode untuk mengumpulkan informasi yang dibutuhkan sangat cocok untuk menggunakan teknik pengambilan data. Karena hasil dari proses pengambilan data masih tahap awal, proses pembersihan informasi masih perlu diselesaikan.

Analisis sentiment pada penelitian ini menggunakan Metode *Naive Bayes*. Berdasarkan Teorema Bayes dan didasarkan pada gagasan bahwa fitur atau sifat bersifat independen satu sama lain, teknik *Naive Bayes* merupakan algoritma pembelajaran terbimbing yang mudah dipahami dan terkenal. Penelitian tentang kategorisasi kumpulan data dermatologis telah menunjukkan bahwa metode *Naive Bayes* lebih akurat daripada algoritma lainnya. Kategorisasi kumpulan data tersebut memiliki tingkat akurasi 99,31% dan didasarkan pada banyak kondisi dermatologis. Hasilnya, pendekatan *Naive Bayes* sesuai untuk mengklasifikasikan data dalam berbagai domain [3].

Dengan melihat hasil penelitian terdahulu yang telah mengkonfirmasi kinerja baik metode *Naive Bayes* dengan meneliti temuan penelitian terdahulu yang membuktikan kemampuan analisis sentimen metode *Naive Bayes* yang kuat. Penelitian ini bertujuan untuk lebih jauh memahami sikap pengguna terhadap aplikasi Bank Jago dengan memberikan informasi dan wawasan baru. Untuk meningkatkan pengalaman pengguna dan kemanjuran aplikasi, diyakini bahwa penelitian ini, yang menerapkan pendekatan *Naive Bayes* pada evaluasi pengguna dari Google Play Store, akan menawarkan informasi yang mendalam kepada pengembang aplikasi tentang cara meningkatkan atau menyempurnakan fitur-fitur saat ini. Harapan dari penelitian ini adalah untuk menambah wawasan dan kontribusi tambahan untuk pemahaman sentimen pengguna terhadap aplikasi Bank Jago. Dengan menerapkan metode *Naive Bayes* pada ulasan pengguna dari Google Play Store, diharapkan penelitian ini dapat memberikan wawasan yang berharga bagi pengembang aplikasi untuk meningkatkan atau menyempurnakan fitur-fitur yang ada, sehingga dapat meningkatkan pengalaman pengguna dan efektivitas aplikasi.

2. METODE PENELITIAN

Metodologi dalam penelitian ini dirancang untuk menjamin ketepatan dan konsistensi selama proses pengumpulan serta analisis data. Rangkaian tahapan metodologis yang digambarkan pada Gambar 1 mencerminkan langkah-langkah utama yang dilakukan guna menghasilkan temuan penelitian yang valid.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Pengumpulan data dilakukan sebagai tahap awal, dengan sumber data berasal dari pengguna aplikasi Bank Jago yang tersedia di Google Play Store. Data tersebut diperoleh melalui teknik *web scraping*, yang memungkinkan pengambilan ulasan pengguna secara otomatis langsung dari halaman aplikasi. Teknik ini dilakukan dengan bantuan salah satu pustaka Python bernama *google-play-scraper*, yang secara otomatis mengekstrak dan mengorganisasi data ulasan ke dalam format CSV [4].

2.2. Aturan Penulisan Numbering secara berurutan

Pelabelan data adalah proses untuk mengelompokkan komentar dari pengguna aplikasi Bank Jago berdasarkan perasaan yang ada di dalamnya. Dalam studi ini, perasaan ditentukan dengan memeriksa nilai emosional di masing-masing komentar. Komentar yang memperoleh skor 1 dan 2 dianggap sebagai "negatif", skor 3 dianggap "netral", sedangkan skor 4 dan 5 dianggap "positif". Ketepatan dan keseragaman dalam proses pelabelan sangat krusial agar data yang diperoleh teratur dan dapat digunakan sesuai dengan tujuan analisis yang telah ditentukan [5].

2.3. Preprocessing

Preprocessing merupakan langkah krusial dalam analisis sentimen, di mana teks terlebih dahulu disiapkan agar dapat ditinjau secara manual maupun digunakan dalam pembangunan model *Naive Bayes*. Proses ini dilakukan di awal untuk memastikan bahwa data yang digunakan bersih dan konsisten, sehingga peneliti dapat menghindari permasalahan yang berkaitan dengan ketidakteraturan data. Tujuan utama dari proses teknis ini adalah untuk menghasilkan klasifikasi yang memiliki tingkat akurasi yang tinggi [6]. Tahapan *preprocessing* dalam penelitian ini meliputi *word normalization*, *case folding*, *cleansing*, *tokenizing*, *stopword removal*, dan *stemming*. Pada tahap normalisasi, kata-kata yang tidak sesuai dengan kaidah bahasa standar diubah menjadi bentuk yang sesuai dengan konvensi linguistik yang berlaku [7]. *Case folding* adalah proses penyamaan bentuk kata dengan mengubah seluruh huruf menjadi huruf kecil atau huruf besar agar arti kata tetap seragam. Sementara itu, *cleansing* bertujuan untuk menghapus elemen-elemen yang tidak relevan dalam teks guna meminimalkan gangguan dalam analisis. Teknik *tokenizing* digunakan untuk memecah teks menjadi unit-unit kecil seperti kata atau frasa berdasarkan pemisah seperti spasi atau simbol tertentu. Proses *stopword removal* menghilangkan kata-kata umum yang sering muncul namun memiliki sedikit makna penting dalam konteks analisis. Terakhir, *stemming* dilakukan untuk mengembalikan kata-kata yang memiliki imbuhan ke bentuk dasarnya atau akar katanya.

2.4. Pembobotan TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) Term Frequency-Inverse Document Frequency (TF-IDF) merupakan metode untuk mengubah data berbasis teks menjadi data numerik dengan tujuan memberikan bobot pada setiap kata atau fitur. Bobot ini dihitung dengan menggabungkan frekuensi kemunculan suatu kata dalam sebuah dokumen serta kebalikan dari jumlah dokumen yang mengandung kata tersebut [8]. Teknik ini berfungsi untuk menentukan seberapa sering sebuah kata muncul dalam suatu dokumen, sambil mempertimbangkan keberadaannya di seluruh kumpulan dokumen. Persamaan TF-IDF ditunjukkan sebagai berikut:

$$TF - IDF(d, t) = TD(d, t) * IDF(t) \quad (1)$$

di mana:

$$TF(d, t) = \frac{\text{jumlah kata } t \text{ pada dokumen } d}{\text{total kata pada dokumen } d} \quad (2)$$

dan:

$$IDF(t) = \log \frac{\text{total dokumen}}{\text{jumlah dokumen yang mengandung kata } t} \quad (3)$$

Penjelasan:

t = kata

d = dokumen

2.5. Pembuatan Model

Setelah menyelesaikan proses penentuan bobot dengan menggunakan TF-IDF, langkah berikutnya adalah menerapkan model dengan metode Naive Bayes untuk mengklasifikasikan sentimen dari ulasan yang diberikan pengguna. Pembuatan model ini sangat penting dalam analisis sentimen, di mana algoritma Naive Bayes digunakan untuk melatih model agar dapat mengenali sentimen dari ulasan pengguna aplikasi Bank Jago. Implementasi dilakukan dengan menggunakan bahasa pemrograman Python. Klasifikasi Naive Bayes adalah metode yang membuat prediksi berdasarkan prinsip probabilitas. Teknik ini beroperasi berdasarkan Teorema Bayes dan diakui sebagai salah satu metode yang cukup dapat diandalkan dan efisien untuk tugas klasifikasi [9]. Rumus umum dari algoritma klasifikasi Naive Bayes dapat dilihat pada rumus (4).

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)} (H|X) \quad (4)$$

Penjelasan:

H: Membuat asumsi data untuk kategori tertentu.

x: Data yang lapisannya belum diketahui.

P(H); Peluang H.

P(x): Peluang x.

P(H| x): Peluang suatu asumsi sesuai dengan keadaan.

P(x |H): Peluang didasarkan pada keadaan dan asumsi.

2.6. Evaluasi dan visualisasi

Tahap akhir dalam proses ini adalah evaluasi, yang memiliki peran penting dalam menilai kinerja model dengan mengukur metrik seperti *accuracy*, *recall*, *precision*, dan *f1-score* [10]. Selanjutnya, proses visualisasi dilakukan menggunakan diagram lingkaran untuk memberikan pemahaman yang lebih mendalam mengenai kemampuan model dalam mengklasifikasikan setiap kategori label. Untuk memperjelas representasi visual terkait distribusi hasil prediksi serta performa model, digunakan diagram lingkaran yang menampilkan persentase masing-masing label secara grafis.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Dalam studi ini, data dikumpulkan dengan menggunakan Google Collab untuk mengunduh ulasan dari pengguna aplikasi Bank Jago yang ada di Google Play Store. Ulasan tersebut kemudian disimpan dalam format CSV. Dengan cara ini, peneliti berhasil mendapatkan 3.990 ulasan yang dianggap relevan, yang mencakup baik komentar maupun skor penilaian. Data yang telah terkumpul kemudian diekspor dan disimpan sebagai file `jago_digitalBanking_2025.csv`. Gambar 2 di bawah ini memperlihatkan langkah-langkah keseluruhan dalam proses pengambilan data.

```
from google_play_scraper import Sort, reviews_all
import pandas as pd

# Ambil semua ulasan
app_reviews = reviews_all(
    'com.jago.digitalBanking',
    lang='id',
    country='id'
)

# Konversi ke DataFrame
reviews_df = pd.DataFrame(app_reviews)

# Ubah kolom datetime ke format datetime Pandas
reviews_df['at'] = pd.to_datetime(reviews_df['at'])

# Filter hanya ulasan dari tahun 2025
reviews_2025 = reviews_df[reviews_df['at'].dt.year == 2025]

# Simpan ke CSV
reviews_2025.to_csv('jago_digitalBanking_2025.csv', index=False)

print("Ulasan tahun 2025 telah disimpan ke dalam file jago_digitalBanking_2025.csv")
```

Ulasan tahun 2025 telah disimpan ke dalam file jago_digitalBanking_2025.csv

Gambar 2. Scraping Data

Gambar 2 menampilkan program yang digunakan untuk mengumpulkan data ulasan dari aplikasi Bank Jago di Google Play Store menggunakan `google_play_scraper`, kemudian mengubahnya menjadi bentuk DataFrame dengan bantuan pustaka `pandas`, dan menyimpannya ke dalam file CSV bernama `'ulasan_bank_jago.csv'`.

3.2. Pelabelan Dataset

Tahap pelabelan data mencakup proses mengelompokkan setiap ulasan pengguna aplikasi Bank Jago ke dalam kategori sentimen positif, negatif, atau netral, yang selanjutnya digunakan sebagai data latih untuk membangun model *Naive Bayes* dalam analisis sentimen.

```
content score label
0 bagus 4 positif
1 Nice App, membantu banget buat alokasi dana, t... 5 positif
2 mohon agar sediakan menu setor tarik tunai di ... 4 positif
3 makasih banyak yang sudah membuat aplikasi ini... 5 positif
4 apps apa ni suruh daftar udh d masukin no dan ... 1 negatif
... ..
3985 lama bener prosesnya, terus aja stack dalam pr... 1 negatif
3986 aku mau login pakai aplikasi ini, tapi kok gab... 1 negatif
3987 Mana whatsapp buat konsultasi 5 positif
3988 Gak bisa masuk akun Padahal mo transfer uang m... 1 negatif
3989 Bagus . 5 positif
```

[3990 rows x 3 columns]

Gambar 3. Pelabelan Data

3.3. Preprocessing

Tahap preprocessing memegang peranan penting dalam proses kerja data science karena dapat sangat memengaruhi hasil analisis secara keseluruhan. Dengan memahami proses ini secara menyeluruh, peneliti akan lebih siap dalam menangani proyek data yang lebih rumit. Pada langkah

ini, dilakukan normalisasi teks dalam data frame untuk menyamakan penggunaan kata-kata yang disingkat, misalnya mengganti "yg" menjadi "yang", "dgn" menjadi "dengan", "udh" menjadi "udah", dan "tp" menjadi "tetapi", serta berbagai bentuk singkatan lainnya. Selain itu, beberapa kata seperti "udah" dan "sudah", atau "tidak" dan "nggak", diperlakukan sebagai sinonim karena memiliki arti yang sama. [11].

```
➡ Sebelum mengganti singkatan:
0                                     bagus
1   Nice App, membantu banget buat alokasi dana, t...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar udh d masukin no dan ...
Name: content, dtype: object

Sesudah mengganti singkatan:
0                                     bagus
1   Nice App, membantu banget buat alokasi dana, t...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar sudah d masukin no da...
Name: content, dtype: object
```

Gambar 4. Normalisasi Teks

Setelah teks dinormalisasi, langkah selanjutnya adalah case folding, yaitu mengubah seluruh huruf dalam ulasan pengguna menjadi huruf kecil. Setelah itu, dilakukan proses cleansing atau pembersihan teks menggunakan berbagai teknik seperti menghapus data yang duplikat, teks kosong, karakter dan tanda baca, URL, tag HTML, simbol, serta referensi yang tidak relevan [12].

```
➡ Sebelum Case Folding:
0                                     bagus
1   Nice App, membantu banget buat alokasi dana, t...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar sudah d masukin no da...
Name: content, dtype: object

Sesudah Case Folding:
0                                     bagus
1   nice app, membantu banget buat alokasi dana, t...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar sudah d masukin no da...
Name: cleaned_content, dtype: object
```

Gambar 5. Case Folding

```
➡ Sebelum membersihkan:
0                                     bagus
1   nice app membantu banget buat alokasi dana tan...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar sudah d masukin no da...
Name: cleaned_content, dtype: object

Sesudah membersihkan:
0                                     bagus
1   nice app membantu banget buat alokasi dana tan...
2   mohon agar sediakan menu setor tarik tunai di ...
3   makasih banyak yang sudah membuat aplikasi ini...
4   apps apa ni suruh daftar sudah d masukin no da...
Name: cleaned_content, dtype: object
```

Gambar 6. Cleansing

Setelah teks dibersihkan, proses dilanjutkan dengan tokenizing, yaitu memecah teks menjadi bagian-bagian yang lebih kecil agar lebih mudah dianalisis. Tahapan berikutnya adalah stopwords removal, yaitu menghapus kata-kata umum yang tidak memiliki makna penting dalam analisis teks.



```
Sebelum tokenisasi:
0                                     bagus
1 nice app membantu banget buat alokasi dana tan...
2 mohon agar sediakan menu setor tarik tunai di ...
3 makasih banyak yang sudah membuat aplikasi ini...
4 apps apa ni suruh daftar sudah d masukin no da...
Name: cleaned_content, dtype: object

Sesudah tokenisasi:
0                                     [bagus]
1 [nice, app, membantu, banget, buat, alokasi, d...
2 [mohon, agar, sediakan, menu, setor, tarik, tu...
3 [makasih, banyak, yang, sudah, membuat, aplika...
4 [apps, apa, ni, suruh, daftar, sudah, d, masuk...
Name: tokenized_content, dtype: object
```

Gambar 7. Tokenizing



```
Sebelum menghapus stopwords:
0                                     [bagus]
1 [nice, app, membantu, banget, buat, alokasi, d...
2 [mohon, agar, sediakan, menu, setor, tarik, tu...
3 [makasih, banyak, yang, sudah, membuat, aplika...
4 [apps, apa, ni, suruh, daftar, sudah, d, masuk...
Name: tokenized_content, dtype: object

Sesudah menghapus stopwords:
0                                     [bagus]
1 [nice, app, membantu, banget, alokasi, dana, r...
2 [mohon, sediakan, menu, setor, tarik, tunai, m...
3 [makasih, aplikasi, memudahkan, transfer, mena...
4 [apps, ni, suruh, daftar, d, masukin, no, sand...
Name: filtered_content, dtype: object
```

Gambar 8. Stopword Removal

Tahap akhir dari preprocessing adalah stemming, yaitu mengubah kata yang memiliki imbuhan ke dalam bentuk dasarnya. Proses ini penting agar semua bentuk variasi dari suatu kata dapat dikenali sebagai satu entitas yang sama oleh model [13]. Hasil dari proses preprocessing seperti pada gambar 4 dan 5.



```
Sebelum stemming:
0                                     [bagus]
1 [nice, app, membantu, banget, alokasi, dana, r...
2 [mohon, sediakan, menu, setor, tarik, tunai, m...
3 [makasih, aplikasi, memudahkan, transfer, mena...
4 [apps, ni, suruh, daftar, d, masukin, no, sand...
Name: filtered_content, dtype: object

Sesudah stemming:
0                                     [bagus]
1 [nice, app, bantu, banget, alokasi, dana, ribe...
2 [mohon, sedia, menu, setor, tarik, tunai, mini...
3 [makasih, aplikasi, mudah, transfer, tabung, h...
4 [apps, ni, suruh, daftar, d, masukin, no, sand...
Name: stemmed_content, dtype: object
```

Gambar 9. Stemming

3.4. Pembobotan TF-IDF

Setelah data melewati seluruh tahapan preprocessing, yang mencakup normalisasi teks, case folding, pembersihan teks, tokenisasi, penghapusan stopwords, dan stemming, tahap berikutnya adalah perhitungan bobot TF-IDF.

```
# Inisialisasi TF-IDF Vectorizer
tfidf_vectorizer = TfidfVectorizer(max_features=5000)

# Transformasi teks
X = tfidf_vectorizer.fit_transform(reviews_df['final_content'])

# Label sebagai target
y = reviews_df['label']
```

Gambar 10. Pembobotan TF-IDF

3.5. Pembuatan model

Langkah berikutnya dalam proses analisis adalah membangun model, yaitu dengan menerapkan algoritma machine learning untuk menganalisis sentimen dari data ulasan pengguna. Dalam penelitian ini, algoritma Naive Bayes dipilih sebagai metode klasifikasi teks. Model tersebut dilatih menggunakan data yang telah diproses sebelumnya dan diberi bobot menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), sehingga siap digunakan untuk menentukan label sentimen.

```
[ ] # Bagi data menjadi training dan testing set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Terapkan SMOTE pada data training
smote = SMOTE(random_state=42)
X_train_resampled, y_train_resampled = smote.fit_resample(X_train, y_train)

# Inisialisasi dan latih model Naive Bayes
nb_model = MultinomialNB()
nb_model.fit(X_train_resampled, y_train_resampled)
```

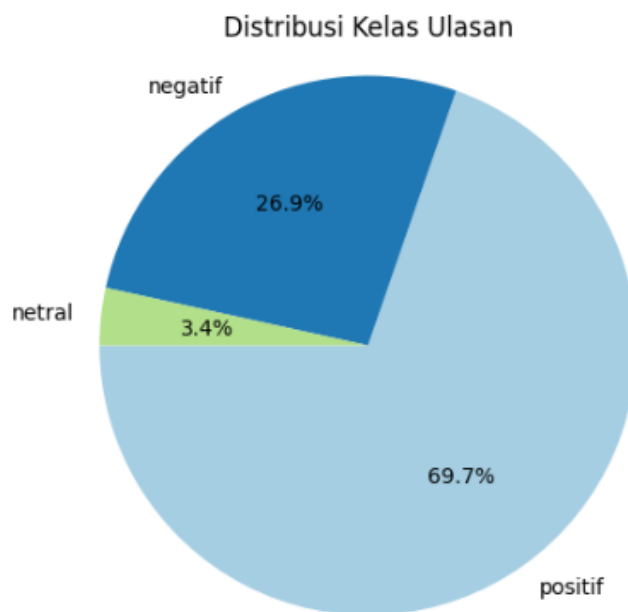
Gambar 11. Pembuatan Model

3.6. Evaluasi dan Visualisasi

Pada tahap ini, proses evaluasi dilakukan dengan memanfaatkan *confusion matrix* dan *pie chart* sebagai alat visualisasi. Selama evaluasi berlangsung, algoritma Naive Bayes digunakan untuk menghitung metrik kinerja seperti *accuracy*, *recall*, *precision*, dan *f1-score*.

Table 1. Hasil Evaluasi

	Precision	Recall	F1-score	Support
Negatif	0.74	0.76	0.75	224
Netral	0.15	0.40	0.22	30
Positif	0.95	0.85	0.90	544
Accuracy	-	-	0.81	798
Macro avg	0.62	0.58	0.62	798
Weighted avg	0.87	0.81	0.83	798



Gambar 12. Diagram Lingkaran

Hasil klasifikasi menunjukkan bahwa model Naive Bayes mampu mencapai akurasi sebesar 81%. Jika ditinjau lebih dalam per kategori, model ini menunjukkan performa yang cukup baik dalam mengenali ulasan dengan sentimen negatif, dengan precision 0,74, recall 0,76, dan f1-score 0,75. Untuk kategori netral, kinerjanya masih rendah, dengan precision hanya 0,15, recall 0,40, dan f1-score 0,22. Sementara itu, pada kelas positif, model bekerja sangat baik, dengan precision 0,95, recall 0,85, dan f1-score 0,90. Jika dilihat dari nilai rata-rata makro (macro average), model mencatatkan precision sebesar 0,62, recall 0,58, dan f1-score 0,62. Sedangkan pada rata-rata tertimbang (weighted average), performa model tetap tinggi dengan precision 0,87, recall 0,81, dan f1-score 0,83. Hal ini disebabkan oleh dominasi data pada kelas positif yang jumlahnya jauh lebih banyak. Selain itu, confusion matrix juga memperkuat temuan ini, menunjukkan bahwa model sangat tepat dalam mengenali ulasan yang bersentimen positif.

4. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma klasifikasi Naive Bayes mampu bekerja secara efektif dalam mengidentifikasi sentimen pengguna terhadap aplikasi Bank Jago, berdasarkan data ulasan terbaru. Sebanyak 3.990 ulasan berhasil dikumpulkan dalam tahap pengambilan data, yang terdiri dari komentar pengguna dan rating yang mereka berikan. Ulasan dikategorikan berdasarkan skala penilaian: rating 1 dan 2 dianggap sebagai sentimen negatif, rating 3 sebagai netral, dan rating 4 serta 5 sebagai sentimen positif. Untuk mengubah teks menjadi bentuk numerik yang dapat diolah oleh model, digunakan metode TF-IDF. Salah satu kendala yang dihadapi adalah jumlah data yang tidak seimbang di setiap kelas sentimen. Untuk mengatasi hal ini, diterapkan teknik SMOTE (Synthetic Minority Over-sampling Technique) pada data pelatihan.

Model ini berhasil mencapai akurasi sebesar 81% dalam melakukan klasifikasi. Kinerja terbaik ditunjukkan oleh kelas sentimen positif, dengan precision sebesar 0,95, recall 0,85, dan f1-score 0,90. Sebaliknya, sentimen netral memiliki performa yang jauh lebih rendah, dengan precision 0,15, recall 0,40, dan f1-score 0,22. Sementara itu, untuk sentimen negatif, precision tercatat sebesar 0,74, recall 0,76, dan f1-score 0,75. Secara keseluruhan, temuan ini mengindikasikan bahwa mayoritas pengguna memberikan tanggapan positif terhadap aplikasi Bank Jago.

DAFTAR PUSTAKA

- [1] Irawan, F. A., Atmadja, A. R., & Wahana, A. (2024). Analisis Sentimen Ulasan Aplikasi Bank Digital Menggunakan Algoritma Naïve Bayes. *Journal of Computer Science and Information Technology*, 4(2), 60-68. <https://doi.org/10.47065/explorer.v4i2.1181>
- [2] Astuti, K. C., Firmansyah, A., & Riyadi, A. (2024). Implementasi Text Mining untuk Analisis Sentimen Masyarakat terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store. *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, 8(1), 383-394. <https://doi.org/10.33395/remik.v8i1.13421>
- [3] Eina, M. F., Chrisnanto, Y. H., & Melina, M. (2024). Klasifikasi Telemarketing Menggunakan Naïve Bayes Classification Dan Wrapper Sequential Feature Selection. *INTECOMS: Journal of Information Technology and Computer Science*, 7(4), 1189-1198. <https://doi.org/10.31539/intecom.v7i4.10846>
- [4] Rahma Firmansyah, D., & Lestariningsih, E. (2024). Analisis Sentimen Ulasan Aplikasi Smart Campus Unisbank di Google Playstore Menggunakan Algoritma Naive Bayes. *Jurnal Teknologi Informasi Dan Komunikasi*, 8(2), 2024. <https://doi.org/10.35870/jti>
- [5] Wijaya, D., Saputra, R. A., & Irwiensyah, F. (2024). Analisis Sentimen Ulasan Aplikasi Samsat Digital Nasional Pada Google Playstore Menggunakan Algoritma Naïve Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(4). <https://doi.org/10.30865/klik.v4i4.1738>
- [6] Anwar, K. (2022). Analisa sentimen Pengguna Instagram Di Indonesia Pada Review Smartphone Menggunakan Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 2(4), 148-155. <https://doi.org/10.30865/klik.v2i4.315>
- [7] Putri, N. A., Srirahayu, A., & Sudibyo, N. A. (2025). Analisis Sentimen Terhadap Aplikasi KitaLulus Menggunakan Metode Naive Bayes dari Ulasan Google Play Store. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 14(2), 269-279. <https://doi.org/10.30591/smartcomp.v14i2.7230>
- [8] Fahrezi, I. A., & Verdikha, N. A. (2024). Analisis Sentimen Twitter Atas Isu Hak Angket Menggunakan Pembobotan TF-IDF dan Algoritma SVM. *Sci-tech Journal*, 3(2), 179-192. <https://doi.org/10.56709/stj.v3i2.526>
- [9] Zhafran Muflih, H., Rizki Abdillah, A., & Noor Hasan, F. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(3), 1613-1621. <https://doi.org/10.30865/klik.v4i3.1303>
- [10] Hamdani, I. M., Nurhidayat, N., Karman, A., & Julyaningsih, A. H. (2024). Edukasi dan Pelatihan Data Science dan Data Preprocessing. *Intisari: Jurnal Inovasi Pengabdian Masyarakat*, 2(1), 19-26. <https://doi.org/10.58227/intisari.v2i1.125>
- [11] Nehe, P. H., Berutu, S. S., & Budiati, H. (2024). Analisis Sentimen Opini Masyarakat Terhadap Presiden Jokowi Sebelum Dan Sesudah Pilpres 2024 Menggunakan Metode Naive Bayes Classification. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 13(1), 451. <https://doi.org/10.35889/jutisi.v13i1.1841>
- [12] Ramadhani, B., & Suryono, R. R. (2024). Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse. *Jurnal Media Informatika Budidarma*, 8(2). <https://doi.org/10.30865/mib.v8i2.7458>
- [13] Darmawan, G., Alam, S., & Imam Sulisty, M. (2023). Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi Mypertamina Pada Google Playstore Menggunakan Metode Naïve Bayes. *Storage - Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 100-108. <https://doi.org/10.55123>