

Analisis Sentimen Aplikasi *ChatGPT Mobile* Menggunakan Algoritma *Naïve Bayes*

Fahmi Ahmad Tohir^{*1}, Bambang Irawan², Agus Bahtiar³

^{1,2}Program Studi Teknik Informatika, STMIK IKMI Cirebon

³Program Studi Sistem Informasi, STMIK IKMI Cirebon

E-mail: ^{*1}clan7kurcaci@gmail.com, ²bambang_irawan_2000@yahoo.com,

³agusbahtiar038@gmail.com

Abstrak

Aplikasi kecerdasan buatan semakin banyak digunakan, termasuk aplikasi ChatGPT. Aplikasi ini merupakan sebuah model bahasa generatif yang dikembangkan oleh OpenAI. Tujuan utama penelitian adalah untuk memahami bagaimana pengguna merespon aplikasi ini melalui ulasan di Google Play Store. Secara khusus, penelitian ini mencermati kata-kata yang sering muncul dalam ulasan positif dan negatif serta memberikan penilaian terhadap kenyamanan penggunaan, responsivitas antarmuka, dan manfaat yang diperoleh dari interaksi dengan model Bahasa tersebut. Digunakan metode pendekatan KDD dan algoritma Naïve Bayes untuk melangsungkan proses penelitian. Berdasarkan 2.238 ulasan di Google Play Store, ditemukan mayoritas ulasan (87%) adalah positif. Pengguna menyoroti kenyamanan penggunaan, responsivitas antarmuka, dan manfaat yang diperoleh dari interaksi dengan model Bahasa. Namun, ada juga ulasan netral (5%) yang memberikan tanggapan baik dan buruk terhadap aplikasi. Selain itu, terdapat ulasan negatif (8%) yang menyoroti ketidaktepatan jawaban model. Evaluasi algoritma klasifikasi Naïve Bayes menunjukkan performa yang sangat baik. Pada skenario 80:20, diperoleh akurasi sebesar 94%, presisi 94%, recall 99%, dan F1-Score 97%. Temuan penelitian ini menunjukkan bahwa ChatGPT memiliki potensi yang besar untuk menjadi aplikasi kecerdasan buatan yang bermanfaat. Namun, pengembang aplikasi perlu memfokuskan peningkatan akurasi jawaban model untuk mengatasi kritik dari pengguna.

Kata Kunci—*Analisis Sentimen, Chatgpt, Ulasan Pengguna, Algoritma Naïve Bayes*

Abstract

AI applications are increasingly being used, including ChatGPT application. This application is generative language model developed by Open AI. Main objective of research is to understand how users respond this application through reviews on Google Play Store. Specifically, this research looks at words that frequently appear in positive and negative reviews and provides an assessment of comfort use, responsiveness of interface, and benefits obtained from interacting with Language model. KDD approach and Naïve Bayes algorithm were used to carry out research process. Based on 2,238 reviews on Google Play Store, it was found that majority of reviews (87%) were positive. Users highlight convenience of use, responsiveness of interface and benefits obtained from interacting with Language model. However, there are also neutral reviews (5%) that give both good and bad feedback to app. Additionally, there were negative reviews (8%) highlighting inaccuracy of model answers. Evaluation of Naïve Bayes classification algorithm shows excellent performance. In 80:20 scenario, an accuracy 94%, precision 94%, recall 99% and F1-Score 97% were obtained. Findings of this research indicate that ChatGPT has great potential to become a useful AI application. However, application developers need to focus on improving accuracy of model answers to overcome criticism from users.

Keywords—*Sentiment Analysis, Chatgpt, User Reviews, Naïve Bayes Algorithm*

1. PENDAHULUAN

Pada era digitalisasi yang semakin canggih ini, semua hal menjadi semakin mudah. Di dalam berbagai bidang kehidupan tidak akan jauh dari digitalisasi. Baik bidang pertanian, peternakan, pemerintahan, Pendidikan, dan bidang lainnya. Dengan adanya penemuan-penemuan digital membawa dampak baik maupun dampak buruk bagi penggunaannya. Dampak ini tergantung bagaimana pengguna memanfaatkan teknologi.

Salah satu contoh penemuan digital yang mempermudah manusia dalam banyak bidang yaitu adanya *AI (Artificial Intelligence)* atau yang lebih dikenal sebagai kecerdasan buatan, *AI* ditemukan pada tahun 1956 oleh para ilmuwan *OpenAI* yang kemudian mengalami “Musim dingin *AI*” pada 2 tahun berikutnya dimana minat terhadap teknologi ini menurun secara drastis. Perkembangan algoritma dan kemajuan algoritma telah memulai kembali kepopuleran *AI* sekitar tahun 2000-an. *AI* telah membawa banyak manfaat baik dalam bidang pertahanan, kesehatan, Pendidikan dan lainnya., kecerdasan buatan membawa banyak dampak baik untuk proses pembelajaran pelajar ataupun mahasiswa yang semakin dipermudah di awal abad 21. Semua ilmu secara terbuka untuk siapapun dan dapat diakses dengan gratis melalui internet sehingga mempercepat penyampaiannya. Sangat berbeda dengan zaman dahulu yang masih menggunakan mesin ketik manual dan penyebaran informasi hanya melalui mulut ke mulut saja. Kemajuan dan perkembangan kecerdasan buatan (*AI*) secara cepat dan signifikan membentuk suatu kebiasaan baru, masyarakat saat ini menggunakan *AI* dalam mempermudah suatu aktivitas keseharian manusia[1]. *AI* yang paling populer di era ini salah satunya adalah *ChatGPT*, *ChatGPT (Generative Pre-Trained Transformer)* adalah robot atau chatbot yang memanfaatkan artificial intelegent atau kecerdasan buatan yang mampu melakukan interaksi dan membantu manusia dalam mengerjakan berbagai tugas[2], yang dikembangkan oleh *OpenAI*. Berupa model Bahasa generatif yang memprediksi probabilitas kalimat atau kata berikutnya dalam suatu percakapan ataupun perintah teks. *ChatGPT* bekerja dengan mempelajari kumpulan data yang sangat besar, baik teks, seperti artikel, buku, dan kode. *ChatGPT* kemudian menggunakan pengetahuan yang diperoleh dari kumpulan data ini untuk menghasilkan teks baru.

Pada mulanya *ChatGPT* hanya terdapat dalam bentuk antarmuka web, hanya dapat diakses dan berinteraksi melalui *website OpenAI* yang di luncurkan pada Juni 2020. Lalu di tahun 2023 ini *OpenAI* meluncurkan sebuah aplikasi *mobile ChatGPT* yang ditujukan untuk mempermudah akses dan juga memberikan kecepatan respons yang lebih tinggi melalui perangkat seluler. Aplikasi *mobile ChatGPT* dirilis di *Google Play Store* pada 26 juli 2023. Hingga bulan November, aplikasi *mobile* ini telah diunduh sebanyak 10 juta lebih dengan *rating* sebesar 4,7. Pemberian *rating* ini diikuti dengan ulasan dari para pengguna aplikasi mengenai opini mereka terhadap aplikasi *ChatGPT*. Banyaknya ulasan yang diberikan dapat mempengaruhi calon pengguna untuk mengunduh dan menggunakan aplikasi. Sehingga dibutuhkan sebuah teknik untuk mengetahui bagaimana ulasan-ulasan tersebut terhadap aplikasi *ChatGPT*.

Pada penelitian yang dilakukan oleh [3] mengenai aplikasi *MOLA* didapatkan sebuah kesimpulan bahwa aplikasi ini memiliki performa yang baik dengan 312 ulasan positif. Lebih tinggi dibandingkan ulasan negatif yang hanya sebanyak 208. Kendala pada aplikasi ini terdapat pada eror dan macet yang sering terjadi. Penelitian yang dilakukan oleh [4] mengenai aplikasi flip diambil 200 *dataset* yang dibagi menjadi 100 data positif dan 100 data negatif. Menghasilkan pengolahan menggunakan *Naïve Bayes* 82,00%, kemudian dioptimasi dengan *PSO* dan didapatkan hasil 88,24% Penelitian yang dilakukan oleh [5] mengenai sentimen aplikasi *jobstreet* mendapat respon positif 83,2% dibandingkan 16,8% respon negatif. penelitian [6] yang mengungkapkan bahwa *ChatGPT* memiliki kekuatan yang besar untuk memajukan akademisi dengan cara baru. Namun, penting untuk mempertimbangkan bagaimana menggunakan teknologi ini secara bertanggung jawab dan etis sebagai profesional untuk meningkatkan pekerjaan daripada menyalahgunakannya. Selain itu, penelitian [7] mengungkapkan bahwa *ChatGPT* muncul sebagai teknologi yang mampu mengubah interaksi

sosial kita dengan cara baru yang radikal. Kecerdasan buatan memiliki potensi untuk merevolusi cara belajar dan mengajar dan metode dalam proses pembelajaran di dunia Pendidikan.

Dengan melakukan analisis terhadap ulasan dari aplikasi *ChatGPT*, diharapkan sebuah gambaran mengenai ulasan yang dirasakan oleh para pengguna aplikasi. Bagaimana performa aplikasi pada saat digunakan. Apakah aplikasi tersebut berjalan dengan semestinya. Serta ingin mengukur seberapa efektifitas algoritma *Naïve Bayes* untuk mengklasifikasi ulasan pengguna.

Bayesian classification adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayesian classification didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan kedalam database dengan data yang besar. Metode Bayes merupakan pendekatan statistic untuk melakukan inferensi induksi pada persoalan klasifikasi. Teorema Bayes digunakan untuk melakukan klasifikasi dalam Data Mining[8].

Penelitian ini melibatkan algoritma *Naïve Bayes* sebagai metode untuk mengklasifikasikan ulasan. Dengan proses pendekatan *Knowledge Discovery in Database (KDD)* digunakan untuk menemukan pengetahuan yang berguna dan bermakna dari kumpulan data besar. Langkah-langkah *Knowledge Discovery in Database (KDD)* meliputi metode *Data Selection, Preprocessing, Transformation, data mining* dan *Evaluation*. Data akan dibagi ke dalam 2 bagian, data latih dan data uji dengan perbandingan 80:20.

2. METODE PENELITIAN

Pengolahan data pada penelitian ini menggunakan proses *data mining*, melibatkan pendekatan *KDD* dan algoritma *Naïve Bayes*. Dimana metode *KDD* meliputi 5 tahapan berikut [9]:

1. *Data Selection*: Tahapan awal pada penelitian ini diawali dengan tahap *data selection*, dimana dilakukan proses penambangan data dan pemilihan terhadap data yang dibutuhkan untuk keberlangsungan penelitian.
2. *Preprocessing*: Menyiapkan data mentah yang baru didapatkan sehingga menjadi *dataset* yang matang.
3. *Transformation*: Proses konversi data atau informasi dari suatu bentuk ke bentuk lain.
4. *Data Mining*: Dilakukannya proses pemodelan menggunakan algoritma tertentu sesuai dengan kebutuhan dan menyesuaikan jenis data yang ada.
5. *Evaluation*: Pada tahapan ini mengukur kinerja model pada data pengujian untuk menilai sejauh mana model dapat menggeneralisasi ke data yang tidak terlihat. Proses tersebut menggunakan metrik evaluasi seperti *accuracy*, *presisi*, *recall*, dan *F1-score*.

3. HASIL DAN PEMBAHASAN

3.1. Data Selection

Pemilihan dan penambangan data dilakukan dengan menggunakan teknik *Scrapping* untuk memperoleh data ulasan dari *Google Playstore*. Didapatkan sebanyak 2.238 yang kemudian diseleksi menjadi hanya atribut *content* dan *score* yang dibutuhkan untuk penelitian.

```
# Menampilkan jumlah keseluruhan data dalam dataset
print('Jumlah Keseluruhan Data:', df.shape[0])

# Menampilkan 15 data teratas
print(df.head(15))
```

	content	score
1119	Sangat bagus 🍌	5
1582	Keren	5
1461	luar biasa	5
1400	Nice application	5
1459	mantap bnget	5
1458	bgusssss bngatttt	5
1457	canggih bgttt	5
1456	MANTAP 🍌	5
1455	shortcut	5
1454	siip	5
1453	mantab	5
1452	Mantap	5
1451	interactive	5
1450	Top	5
1449	Wawawww	5

Gambar 1. Proses *Scrapping Data*

Pada Gambar 1 memperlihatkan hasil dari pengambilan data dengan scrapping data, hanya menampilkan 15 data dari keseluruhan data yang ada.

3.2. Preprocessing

3.2.1. Cleaning Data

Dilakukannya penghapusan karakter-karakter yang tidak penting seperti tanda baca karakter khusus, maupun *URL*. Seperti terlihat pada Gambar 2, data pada atribut content sudah tidak memiliki tanda baca karakter khusus, maupun *URL*.

```
# Mengubah nilai kosong menjadi NaN
df.replace('', pd.np.nan, inplace=True)

# Menghapus kolom yang kosong
df.dropna(how='all', axis=1, inplace=True)

# Menghapus baris dengan nilai kosong pada kolom 'text_culum'
df.dropna(subset=['content'], inplace=True)

# Simpan DataFrame yang telah diubah ke dalam file CSV
df.to_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_cleaning.csv', index=False)
```

	content	score
0	Sangat bagus	5
1	Keren	5
2	luar biasa	5
3	Nice application	5
4	mantap bnget	5
5	bgusssss bngatttt	5
6	canggih bgttt	5
7	MANTAP	5
8	shortcut	5
9	siip	5
10	mantab	5
11	Mantap	5
12	interactive	5
13	Top	5
14	Wawawww	5

Gambar 2. Proses *Cleaning Data*

Pada Gambar 2 memperlihatkan hasil dari cleaning data, terlihat perbedaan hasil datanya dari Gambar 4.1 dimana data tersebut memperlihatkan karakter khusus.

3.2.2. Case Folding

Tahapan ini dilakukan untuk mengubah semua huruf dalam teks menjadi huruf-huruf kecil agar tidak ada perbedaan antara huruf besar dan kecil yang dianggap sebagai hal yang berbeda.

```
# Case folding
df['content'] = df['content'].apply(lambda x: x.lower() if isinstance(x, str) else x)
df.to_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_casefolding.csv', index=False)
print(df.head(15))
```

	content	score
0	sanget bagus	5
1	keren	5
2	luar biasa	5
3	nice application	5
4	mantap bnget	5
5	bgusssss bngatttt	5
6	canggih bgttt	5
7	mantap	5
8	shortcut	5
9	slip	5
10	mantab	5
11	mantap	5
12	interactive	5
13	top	5
14	wawawaw	5

Gambar 3. Proses Case Folding

Pada Gambar 3, melakukan case folding dan menampilkan data yang tadinya perbedaan dari huruf besar dan kecil sehingga sekarang berupa huruf kecil semua.

3.2.3. Tokenizing

Tokenizing adalah tahap dimana teks kalimat dipecah menjadi bagian-bagian kecil atau dipisahkan perkata. Langkah ini penting untuk dilakukan dalam pengolahan teks mining karena mempermudah menganalisis dan memproses teks.

```
# Tokenizing
# Baca data dari CSV
df = pd.read_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_casefolding.csv')

df['content'] = df['content'].apply(lambda x: x.split() if isinstance(x, str) else np.nan)
df.to_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_tokenizing.csv', index=False)
print(df.head(15))
```

	content	score
0	[sang,at, bagus]	5
1	[keren]	5
2	[luar, biasa]	5
3	[nice, applicat,ion]	5
4	[mantap, bnget]	5
5	[bgusssss, bngatttt]	5
6	[canggih, bgttt]	5
7	[mantap]	5
8	[shortcut]	5
9	[slip]	5
10	[mantab]	5
11	[mantap]	5
12	[interact,ive]	5
13	[top]	5
14	[wawawaw]	5

Gambar 4. Proses Tokenizing

Pada Gambar 4 melakukan tokenizing unuk membagi/memecah teks menjadi bagian-bagian kecil yang digunakan untuk analisis pengolahan kata.

3.2.4. Stopword

Di dalam setiap ulasan pastilah terdapat kata yang umum digunakan tetapi tidak memiliki makna spesifik, maka dari itu dilakukan proses *Stopword* untuk menghilangkan kata-kata tersebut agar pengolahan data dapat terfokuskan pada kata-kata yang relevan untuk menentukan sentimen.

```

# Simpan ke file CSV
df.to_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_stopword.csv', index=False)
print(df.head(15))

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!

```

	content	score
0	['sangat ', 'bagus ']	5
1	['keren ']	5
2	['luar ', 'biasa ']	5
3	['nice ', 'application ']	5
4	['mantap ', ' ']	5
5	[' ', ' ']	5
6	['canggih ', ' ']	5
7	['mantap ']	5
8	[' ']	5
9	[' ']	5
10	['mantab ']	5
11	['mantap ']	5
12	[' ']	5
13	[' ']	5
14	[' ']	5

Gambar 5. Proses Stopword

Pada Gambar 5, stopwords menghapus kata yang kurang relevan atau tidak memiliki makna lebih untuk analisis sentimen.

3.2.5. Stemming

Setiap kata berimbuhan dikembalikan ke bentuk kata dasar atau akarnya. Langkah ini dilakukan dengan menggunakan *Stemming* yaitu pemisahan kata dasar dari imbuhan, seperti imbuhan 'me', 'di', 'kan', dan lain-lain.

```

# Mendapatkan semua kata unik dari kolom 'content'
all_text = ' '.join(df['content'].dropna().astype(str))
words = set(all_text.split())

# Inisialisasi StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()

# Melakukan stemming pada setiap kata
stemmed_words = [stemmer.stem(word) for word in words]

# Mengambil 300 kata yang telah di-stemming
top_300_stemmed_words = stemmed_words[:300]

# Menggabungkan kata-kata yang telah di-stemming menjadi string dengan format yang diminta
stemmed_words_string = ', '.join(f'{word}' for word in top_300_stemmed_words)
df.to_csv('/content/drive/MyDrive/Analisis Sentimen ChatGPT/data_stemming.csv', index=False)

# Menampilkan string dengan 300 kata yang telah di-stemming
print(stemmed_words_string)

```

'gua', 'good', 'sih', 'sumpah', 'kurang', 'play', 'cakap', 'ke', 'malah', 'ikan', 'doang', 'ban

Gambar 6. Proses Stemming

Pada Gambar 6, melakukan proses stemming data untuk mendapatkan kata dasar dalam kalimat, serta menampilkan kata kata dasar yang diambil dari kolom content pada data.

3.2.6. Labeling

Proses dimana setiap data diklasifikasikan menjadi 'positif', 'negatif' dan 'netral' berdasarkan sentiment yang terkandung di dalam data tersebut. Label diberikan berdasarkan *score* dari setiap data *content*[10].



Gambar 9. Word Cloud Sentimen Negatif

Pada Gambar 8, Untuk sentimen negatif kata yang sering muncul adalah “gak”, “login”, “tidak”, dan “salah”.



Gambar 10. Word Cloud Sentimen Netral

Pada gambar 10, sentimen netral kata yang sering muncul adalah kata “tidak”, “akurat”, “bagus” dan “bisa”.

3.3. Transformation

Transformation adalah proses konversi data atau pengubahan informasi dari suatu bentuk ke bentuk lainnya. Langkah ini bertujuan untuk mengubah teks ulasan menjadi representasi numerik dan prosesnya menggunakan metode *TF-IDF* (*Term Frequency-Inverse Document Frequency*) untuk mengkonversikan teks menjadi vektor numerik sehingga menghasilkan pembobotan untuk memberikan nilai pada frekuensi dari munculnya kata (*TF*) dan frekuensi kemunculan kata atau *term* pada seluruh dataset (*IDF*).

Hasil Pembobotan TF-IDF:

	ada	adalah	add	agak	agar	ai	ainya	aja	akan	akses
0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
1	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
2	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
3	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
4	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
...
2206	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
2207	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
2208	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0
2209	0.0	0.0	0.0	0.0	0.0	0.100397	0.0	0.0	0.134455	0.0
2210	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0

	web	website	when	will	with	wow	ya	yang	yg	you
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
...
2206	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
2207	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
2208	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0
2209	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.251879	0.0	0.0
2210	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0

[2211 rows x 441 columns]

Gambar 11. Proses Transformation TF-IDF

Pada Gambar 11, tahapan transformation diperlihatkan hasil pembobotan yang menggunakan metode TF-IDF pada teks dengan mengkonversikan kedalam vektor numerik.

3.4. Data Mining

3.4.1. Pembagian Data Uji dan Data Latih

Membagi data menjadi 2 subset yaitu data latih dan data uji, metodenya menggunakan *train-test split* untuk menguji kinerja model pada data yang tidak terlihat. Pembagian data latih dan data uji dilakukan dengan rasio 80:20

```

import pandas as pd
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, classification_report
from sklearn.utils import shuffle

# Mem baca file CSV
df = pd.read_csv('/content/drive/MyDrive/Analisis Sentimen chatGPT/data_transformed.csv')

# variabel dependen (y) dan variabel independen (X)
X = df_transformed.drop('sentiment', axis=1)
y = df_transformed['sentiment']

# Pembagian data latih dan data uji dengan rasio 80:20
X_train_80, X_test_20, y_train_80, y_test_20 = train_test_split(Tfidf_matrix, df['sentiment'], test_size=0.2, random_state=42)
    
```

Gambar 12. Proses Pembagian Data

Pada Gambar 12, tahapan ini merupakan proses pembagian data menjadi data uji dan data latih.

3.4.2. Melatih Model Sentimen

Tahap ini dilakukan dengan melibatkan algoritma *Naive Bayes* untuk melatih model analisis sentimen sehingga dapat dipahami korelasi antara fitur dan sentimen. Dapat dilihat dalam Gambar 12 bagaimana *syntax* algoritma *Naive Bayes* untuk melakukan pelatihan model pada data uji dan mengetahui evaluasi performa dari model yang dihasilkan.

```

from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, classification_report
from sklearn.model_selection import cross_val_score

# Fungsi untuk melatih dan mengevaluasi model
def train_and_evaluate(model, X_train, y_train, X_test, y_test, ratio_name):
    print(f'\nModel dengan rasio {ratio_name}')

    # Melatih model
    model.fit(X_train, y_train)

    # Mengevaluasi model pada data uji
    y_pred = model.predict(X_test)

    # Mengevaluasi performa model
    accuracy = accuracy_score(y_test, y_pred)
    print(f'Accuracy: {accuracy:.2f}')

    print('Classification Report:')
    print(classification_report(y_test, y_pred))

# Inisialisasi model Naive Bayes
model = MultinomialNB(alpha=0.12)

# Train dan evaluasi model untuk rasio 80:20
train_and_evaluate(model, X_train_80, y_train_80, X_test_80, y_test_80, '80:20')
    
```

Gambar 13. Pelatihan Data

Pada Gambar 13, tahapan ini merupakan tahapan pelatihan data untuk menentukan nilai akurasi dari klasifikasi.

3.5. Evaluation

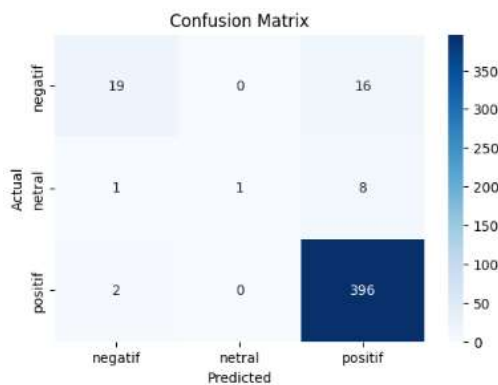
3.5.1. Confusion Matrix

Model dengan rasio 80:20
 Accuracy: 0.94
 Classification Report:

	precision	recall	f1-score	support
negatif	0.86	0.54	0.67	35
netral	1.00	0.10	0.18	10
positif	0.94	0.99	0.97	398
accuracy			0.94	443
macro avg	0.94	0.55	0.61	443
weighted avg	0.94	0.94	0.93	443

Gambar 14. Performa Model Algoritma Naive Bayes

Pada gambar 14, diperlihatkan hasil evaluasi keakuratan algoritma Naive Bayes dengan rasio model 80:20.



Gambar 15. Confusion Matrix dari Rasio 80:20

Pada gambar 15, confusion matrix dengan rasio 80:20 menunjukkan akurasi dan prediksi antara positif dan negatif. Dan berikut merupakan tabel seluruh hasil yang diperoleh semua rasio yang dilatih.

Evaluasi model dilakukan dengan menggunakan *confusion matrix*, dimana confusion matrix sendiri merupakan tabel yang digunakan dalam *machine learning* untuk mengevaluasi kinerja model klasifikasi. Bertujuan untuk membandingkan prediksi model dengan nilai sebenarnya yang membantu mengukur nilai akurasi, presisi, *recall* dan *f-1 score*. Dari keenam (3x3) hasil *confusion matrix* rasio didapatkan bahwa rasio dengan perbandingan nilai 80:20 memperoleh nilai performa paling tinggi, yaitu 94% untuk nilai akurasi, 94% nilai presisi, 99% nilai *recall*, dan 97% untuk nilai *f-1 score*. berikut penghitungan nilai performa menggunakan rumus.

Table 1. Confusion Matrix

Kelas Sebelumnya	Kelas Prediksi			
		Positif	Negatif	Netral
	Positif	(396) TPP	(2) PFNeg	(0) PFNet
	Negatif	(16) NegFP	(19) TNegNeg	(0) NegFNet
	Netral	(8) NetFP	(1) NetFNeg	(1) TNetNet

Keterangan [14]:

1. TPP (True Positif Positif) = Jumlah prediksi yang positif dari data kelas yang positif.
2. TNegNeg (True Negatif Negatif) = Jumlah prediksi yang negatif dari data kelas yang negatif.
3. TNetNet (True Netral Netral) = Jumlah prediksi yang netral dari data kelas yang netral.
4. PFNeg (Positif False Negatif) = Jumlah prediksi yang positif dari data kelas yang negatif.
5. NegFP (Negatif False positif) = Jumlah prediksi yang negatif dari data kelas yang positif.
6. NegFP (Negatif False Positif) = Jumlah prediksi yang negatif dari data kelas yang positif.
7. PFNet (Positif False Netral) = Jumlah prediksi yang positif dari data kelas yang netral.
8. NetFNeg (Netral False Negatif) = Jumlah prediksi yang netral dari data kelas yang negatif.
9. NegFNet (Negatif False Netral) = Jumlah prediksi yang negatif dari data kelas yang netral [15].

$$accuracy = \frac{TPP + TNegNeg + TNetNet}{Total} = \frac{396 + 19 + 1}{396 + 2 + 0 + 16 + 19 + 0 + 8 + 1 + 1} = \frac{416}{443} \times 100\% = 94\%$$

$$precision = \frac{TPP}{TPP + NegFP + NetFP} = \frac{396}{396 + 16 + 8} = \frac{396}{420} \times 100\% = 94\%$$

$$recall = \frac{TPP}{TPP + PFNeg} = \frac{396}{396 + 2 + 0} = \frac{396}{398} \times 100\% = 99\%$$

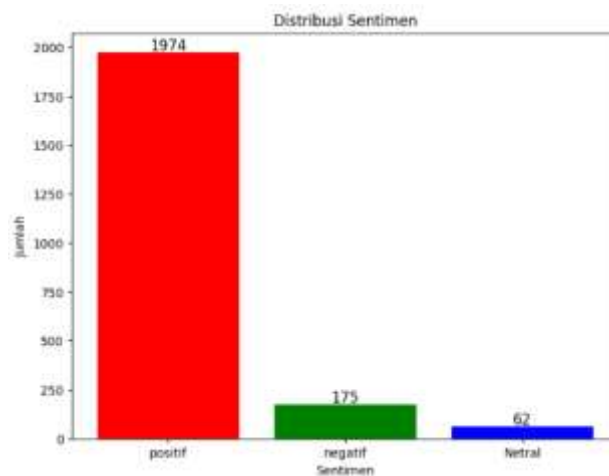
$$f1\ score = 2 \times \frac{presisi \times recall}{presisi + recall} = 2 \times \frac{99 \times 94}{99 + 94} = 2 \times \frac{9306}{192} \times 100\% = 97\%$$

Dari hasil evaluasi tersebut, dapat disimpulkan bahwa model klasifikasi sentimen yang telah diperoleh dapat digunakan untuk memprediksi sentimen dari teks dengan akurasi yang tinggi. Model ini dapat digunakan untuk berbagai aplikasi, seperti analisis sentimen media sosial, analisis sentimen produk, dan analisis sentimen pelanggan.

3.5.2. Visualisasi Data

Menyajikan data secara visual untuk memahami pola dan tren secara lebih intuitif. Dengan menggunakan grafik untuk menampilkan distribusi sentimen. Berdasarkan Gambar 15. Ditemukan sebanyak 1.974 ulasan yang memberikan sentimen positif terhadap aplikasi ChatGPT, seperti pada ulasan kata-kata yang sering muncul yaitu “mantap”, “membantu”, “bagus”, dan “keren”. Secara umum, pengguna menyoroti kenyamanan dalam penggunaan, responsivitas antarmuka dan manfaat yang diperoleh dari interaksi dengan model Bahasa tersebut. Ulasan positif ini memberikan gambaran yang bagus dan menunjukkan aplikasi diterima dengan baik di kalangan pengguna.

Namun disamping itu, sebanyak 62 dan 175 ulasan yang merepresentasikan sentimen netral dan negatif terhadap aplikasi ChatGPT seperti ulasan kata-kata yang sering muncul yaitu netral “tidak”, “akurat”, “bagus” dan “bisa” dan negatif “gak”, “login”, “tidak”, dan “salah”. Kata-kata tersebut merujuk pada kritik pada aplikasi mengenai kurang tepatnya jawaban yang diberikan oleh model Bahasa AI.



Gambar 16. Diagram Sentimen

Pada gambar 16, memvisualisasikan jumlah keseluruhan sentimen positif dan negatif terlihat dominan sentimen positif dengan 1974, sentimen netral 62 dan untuk sentimen negatif 175.

4. KESIMPULAN

Metode KDD meliputi 5 tahapan, dimana tahap pertama adalah data selection yaitu mengambil data dari ulasan platform Google Play Store mengenai aplikasi ChatGPT. Pengambilan data tersebut diambil menggunakan teknik scrapping API dengan mengirim permintaan untuk mendapatkan data yang diperlukan demi keberlangsungan penelitian. Setelah 2.238 data berhasil didapatkan, kemudian dilakukan penyeleksian terhadap atribut sehingga hanya menampilkan atribut content dan score. Tahap kedua yaitu preprocessing data dilakukan dengan data cleaning, membersihkan karakter khusus selain huruf dan spasi seperti hashtag, tanda petik, mention, dan lain-lain. Kemudian dilakukannya case folding yaitu membuat huruf kapital menjadi huruf kecil setelah huruf menjadi sama, setiap kalimat dipisahkan perkata menggunakan tokenizing. Kata-kata yang telah dipisahkan disaring berdasarkan makna, kata yang tidak memiliki makna akan dihapus di langkah stopwords. Stemming, yaitu melakukan penghapusan imbuhan dari setiap kata yang sudah memiliki makna. Setelah dataset selesai dipersiapkan maka dilakukan pelabelan menggunakan score dari setiap ulasan.

Dataset telah menjadi data akhir yang matang tetapi masih belum bisa dilakukan tahap data mining karena harus dilakukan tahap transformation terlebih dahulu, yaitu pembobotan terhadap setiap kata yang ada. Baru tahap data mining menggunakan algoritma Naïve Bayes

dapat dilakukan sehingga menghasilkan sebuah model yang akan dievaluasi menggunakan confusion matrix.

Dari hasil evaluasi Confusion Matrix didapatkan performa dari model yang diperoleh yaitu nilai akurasi sebesar 94%, presisi 94%, recall 99% dan f1-score 97%. Dengan memperoleh hasil tersebut membuktikan bahwa algoritma naïve bayes mampu mendapatkan hasil yang baik..

5. SARAN

Adapun saran untuk penelitian selanjutnya antara lain:

1. Mendapatkan dataset yang seimbang,
2. Disarankan untuk penelitian selanjutnya agar menggunakan algoritma lain agar bisa dijadikan sebagai perbandingan dengan algoritma yang telah digunakan. Seperti contohnya algoritma Decision Tree, Support Vector Machine, dan algoritma lainnya.

DAFTAR PUSTAKA

- [1] M. J. Maulana, C. Darmawan, and ..., "PENGUNAAN CHATGPT DALAM TINJAUAN PENDIDIKAN BERDASARKAN PERSPEKTIF ETIKA AKADEMIK," *Bhineka Tunggal Ika* ..., 2023, [Online]. Available: <https://ejournal.unsri.ac.id/index.php/jbti/article/view/21090>
- [2] A. Faiz and I. Kurniawaty, "Tantangan Penggunaan ChatGPT dalam Pendidikan Ditinjau dari Sudut Pandang Moral," *Edukatif J. Ilmu Pendidik.*, vol. 5, no. 1, pp. 456–463, 2023, doi: 10.31004/edukatif.v5i1.4779.
- [3] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Mola Application Reviews on Google Play Store Using Support Vector Machine Algorithm," *J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, pp. 1–7, 2022.
- [4] O. Irnawati and K. Solecha, "Analisis Sentimen Ulasan Aplikasi Flip Menggunakan Naïve Bayes dengan Seleksi Fitur PSO," *J. Ilm. Intech Inf.* ..., 2022, [Online]. Available: <http://jurnal.umus.ac.id/index.php/intech/article/view/868>
- [5] B. K. Widodo, N. H. Matondang, and D. S. Prasvita, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Penggunaan Aplikasi Jobstreet," *Techno. Com*, 2022, [Online]. Available: <http://publikasi.dinus.ac.id/index.php/technoc/article/view/6361>
- [6] B. D. Lund and T. Wang, "Chatting about ChatGPT: how may AI and GPT impact academia and libraries?," *Libr. Hi Tech News*, vol. 40, no. 3, pp. 26–29, Jan. 2023, doi: 10.1108/LHTN-01-2023-0009.
- [7] S. T. Rudolph J, Samson Tan, "Journal of Applied Learning & Teaching ChatGPT : Bullshit spewer or the end of traditional assessments in higher education ?," vol. 6, no. 1, 2023.
- [8] A. Haditsah, "Klasifikasi Masyarakat Miskin menggunakan Metode Naïve Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018.
- [9] M. K. Khoirul Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 478–483, 2023, doi: 10.36040/jati.v7i1.6373.
- [10] N. R. Siahaan, R. Y. Tiffany, S. R. E. Sinaga, and ..., "ANALISIS SENTIMEN ULASAN APLIKASI MEDIA SOSIAL WHATSAPP MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," *J. Ilm.* ..., 2023, [Online]. Available: <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/104>
- [11] S. Masturoh, R. L. Pratiwi, and ..., "APPLICATION OF THE K-NEAREST NEIGHBOR (KNN) ALGORITHM IN SENTIMENT ANALYSIS OF THE OVO E-WALLET APPLICATION," *JITK (Jurnal Ilmu* ..., 2023, [Online]. Available: <https://ejournal.nusamandiri.ac.id/index.php/jitk/article/view/3997>

- [12] Z. Annisa and B. S. S. Ulama, “Analisis Sentimen Data Ulasan Pengguna Aplikasi ‘PeduliLindungi’ pada Google Play Store Menggunakan Metode Naïve Bayes Classifier Model Multinomial,” *J. Sains dan Seni ITS*, vol. 11, no. 6, 2023, doi: 10.12962/j23373520.v11i6.94064.
- [13] Raksaka Indra Alhaqq, I Made Kurniawan Putra, and Yova Ruldeviyani, “Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 105–113, 2022, doi: 10.22146/jnteti.v11i2.3528.
- [14] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <https://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369/348>
- [15] A. Herdhianto, *Sentiment Analysis Menggunakan Naïve Bayes Classifier (NBC) Pada Tweet Tentang Zakat*. 2020.