

ANALISIS SENTIMEN KOMENTAR PADA TIKTOK SHOP MENGUNAKAN METODE TF-IDF DAN KLASIFIKASI NAÏVE BAYES (STUDI KASUS: GLAD2GLOW)

Devy Liani Herawati^{1*}, Rizaldy Khair²,

¹ Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi,
Universitas Muhammadiyah Sumatera Utara, Jl.Kapten Muchtar Basri No.3, Glugur
Darat II, Kec. Medan Timur, Medan, Sumatera Utara, Indonesia, 20238

² Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi,
Universitas Muhammadiyah Sumatera Utara, Jl.Kapten Muchtar Basri No.3, Glugur
Darat II, Kec. Medan Timur, Medan, Sumatera Utara, Indonesia, 20238

*Penulis Korespondensi: devylianiherawati44@gmail.com

Abstract. *The rapid growth of social media and social commerce platforms has generated a large volume of user comments that can be utilized to understand consumer perceptions of a product. TikTok Shop, as one of the leading digital shopping platforms, contains various user comments expressing positive, negative, and neutral sentiments, making an automatic sentiment analysis approach necessary. This study aims to analyze user sentiments toward Glad2Glow products on TikTok Shop by applying the Term Frequency-Inverse Document Frequency (TF-IDF) method for feature weighting and the Naïve Bayes algorithm for sentiment classification. The dataset used in this study consists of 350 user comments collected from TikTok Shop. The research process includes data collection, text preprocessing, feature extraction using TF-IDF, sentiment classification with the Naïve Bayes algorithm, and model evaluation using a confusion matrix. The experimental results show that the classification model achieved an accuracy of 98.06%, with precision values of 1.00 for the negative class, 0.89 for the neutral class, and 1.00 for the positive class. The recall values were 0.92, 1.00, and 1.00, while the F1-scores were 0.96, 0.94, and 1.00, respectively. Furthermore, the web-based sentiment analysis system successfully performed automatic sentiment classification and displayed the classification results effectively. These findings indicate that the TF-IDF method and the Naïve Bayes algorithm provide excellent performance in classifying user sentiments toward Glad2Glow products on the TikTok Shop platform.*

Keywords: *Sentiment Analysis, TF-IDF, Naïve Bayes, TikTok Shop, Glad2Glow.*

Abstrak. Perkembangan media sosial dan platform social commerce telah menghasilkan banyak komentar pengguna yang dapat dimanfaatkan untuk mengetahui persepsi konsumen terhadap suatu produk. TikTok Shop sebagai salah satu platform belanja digital menyediakan berbagai komentar yang mengandung sentimen positif, negatif, dan netral, sehingga diperlukan metode yang mampu menganalisis sentimen secara otomatis. Penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna terhadap produk Glad2Glow pada TikTok Shop dengan menerapkan metode Term Frequency-Inverse Document Frequency (TF-IDF) sebagai pembobotan fitur dan algoritma Naïve Bayes sebagai metode klasifikasi. Data yang digunakan sebanyak 350 komentar pengguna yang diperoleh dari TikTok Shop. Tahapan penelitian meliputi pengumpulan data, preprocessing, pembobotan kata menggunakan TF-IDF, klasifikasi menggunakan Naïve Bayes, serta evaluasi model menggunakan confusion matrix. Hasil penelitian menunjukkan bahwa model klasifikasi memperoleh nilai accuracy sebesar 98,06%, dengan nilai precision sebesar 1,00 untuk kelas negatif, 0,89 untuk kelas netral, dan 1,00 untuk kelas positif. Nilai recall yang diperoleh masing-masing sebesar 0,92, 1,00, dan 1,00, sedangkan nilai F1-Score berturut-turut sebesar 0,96, 0,94, dan 1,00. Selain itu, sistem analisis sentimen berbasis website yang dibangun mampu melakukan proses analisis komentar secara otomatis dan menampilkan hasil klasifikasi dengan baik. Berdasarkan hasil tersebut, metode TF-IDF dan algoritma Naïve Bayes mampu memberikan performa yang sangat baik dalam mengklasifikasikan sentimen komentar pengguna terhadap produk Glad2Glow pada platform TikTok Shop.

Kata kunci: Analisis Sentimen, TF-IDF, Naïve Bayes, TikTok Shop, Glad2Glow.

1. LATAR BELAKANG

Perkembangan e-commerce menunjukkan pertumbuhan yang sangat signifikan selama beberapa tahun terakhir. Jumlah transaksi digital global mencapai triliunan dolar dan masih terus bertambah, didorong oleh penetrasi internet dan penggunaan smartphone yang massif. Di Indonesia, pangsa pasar digital diperkirakan terus tumbuh hingga 2025, dan platform media sosial yang mengintegrasikan fitur belanja seperti TikTok Shop menjadi salah satu tren terbaru dalam perilaku konsumen digital. Data InSight 2025 menunjukkan TikTok Shop masuk dalam peringkat e-commerce dengan jumlah percakapan positif di media sosial, bersaing dengan Shopee, Tokopedia, dan Lazada sebagai indikator kepuasan konsumen di ranah digital (mis. Shopee: 749 ribu percakapan, TikTok Shop: 63 ribu percakapan positif) meskipun persaingannya ketat dan dinamis. (Jurnal 2, n.d.).

Di tengah perkembangan tersebut, komentar dan ulasan konsumen memainkan peran penting dalam membentuk persepsi produk, pengalaman pengguna, dan keputusan pembelian. Komentar ini sering kali berisi opini eksplisit berupa kritik, saran, dan pujian yang mencerminkan persepsi konsumen terhadap suatu produk atau layanan. Sayangnya, volume besar data komentar ini bersifat tidak terstruktur, sehingga analisis manual tidak efisien serta rentan bias manusia. Kondisi tersebut menjadi tantangan utama bagi peneliti maupun pelaku usaha untuk secara cepat dan akurat mengukur sentiment konsumen. Di sinilah teknik analisis sentiment berbasis text mining dan machine learning menjadi sangat diperlukan untuk mengekstraksi dan mengklasifikasikan opini pengguna secara objektif dan kuantitatif. (Ashari et al., 2025)

Beragam penelitian telah dilakukan untuk menerapkan analisis sentiment dalam konteks e-commerce dan aplikasi lainnya. Sebagai contoh, Indrayuni & Nurhadi (2025) melakukan analisis sentiment ulasan pengguna TikTok Shop Seller Center menggunakan tiga algoritma berbeda untuk menentukan model terbaik, dengan SVM menunjukkan performa tertinggi dibandingkan Naïve Bayes dan Logistic Regression. Temuan lain dari penelitian yang berfokus pada TikTok Shop menemukan pentingnya Teknik analisis sentiment untuk memahami opini pengguna demi meningkatkan platform. Namun demikian, mayoritas studi tersebut masih menggunakan dataset umum atau platform yang berbeda dan belum banyak yang fokus pada produk tertentu seperti Glad2Glow, serta

belum memaksimalkan kombinasi metode pembobotan fitur dan klasifikasi dalam satu konteks spesifik social commerce. Hal ini menunjukkan adanya ruang pengembangan dalam metodologi analisis yang digunakan. (Indrayuni & Acmad Nurhadi, 2025)

Selain itu, sebagian besar penelitian terdahulu masih berfokus pada algoritma tertentu tanpa membandingkan secara sistematis efektivitas model dalam klasifikasi sentiment untuk dataset yang besar dan beragam karakter bahasanya, seperti komentar TikTok Shop. Di sisi lain, metode TF-IDF terbukti efektif dalam berbagai studi text mining karena kemampuannya menilai pentingnya kata dalam dokumen besar, tetapi aplikasinya pada kasus TikTok Shop dengan karakteristik Bahasa Indonesia yang banyak menggunakan slank, singkatan, dan emoji masih relatif kurang dieksplor secara mendalam, sehingga menciptakan kesenjangan penelitian yang perlu diisi melalui pendekatan yang lebih terintegrasi antara Teknik pembobotan fitur dan algoritma klasifikasi yang sesuai.

Penelitian ini hadir untuk mengisisi kesenjangan tersebut dengan melakukan analisis sentiment komentar pengguna pada TikTok Shop produk Glad2Glow dan algoritma Naïve Bayes Classifier. Penelitian terdahulu telah mengaplikasikan metode yang serupa untuk analisis sentiment di media sosial dan e-commerce, misalnya (Utomo, 2021). Penelitian Utomo lebih fokus pada platform umum atau ulasan aplikasi, dan belum banyak yang mengarah pada komentar produk TikTok Shop spesifik satu brand dengan karakter Bahasa yang sangat singkat dan informal.

Memalui pendekatan kuantitatif, penelitian ini bertujuan mengukur distribusi sentiment konsumen secara akurat dan mengevaluasi performa model dengan metrik seperti akurasi, presisi, recall, dan F1-score, sehingga tidak hanya memberikan gambaran opini konsumen secara numerik, tetapi juga bukti empiris penerapan text mining di platform social commerce yang sedang berkembang pesat. Secara teoritis, penelitian ini diharapkan memperluas kajian metodologi analisis sentiment e-commerce di TikTok Shop, sedangkan secara praktis dapat membantu pemilik merek seperti Glad2Glow dan pelaku usaha lain dalam memahami persepsi pelanggan, merancang strategi pemasaran yang lebih tepat, serta meningkatkan kualitas produk dan layanan berdasarkan data opini konsumen nyata.

2. KAJIAN TEORITIS

Perkembangan *e-commerce* menunjukkan pertumbuhan yang sangat signifikan selama beberapa tahun terakhir. Jumlah transaksi digital global mencapai triliunan dolar dan terus bertambah, didorong oleh penetrasi internet dan penggunaan *smartphone* yang masif. Di Indonesia, pangsa pasar digital diperkirakan terus tumbuh hingga tahun 2025, dan platform media sosial yang mengintegrasikan fitur belanja seperti TikTok Shop menjadi salah satu tren terbaru dalam perilaku konsumen digital (*social commerce*). Data Insight 2025 menunjukkan bahwa TikTok Shop masuk dalam peringkat *e-commerce* dengan jumlah percakapan positif yang tinggi di media sosial, bersaing secara ketat dan dinamis dengan platform lain seperti *Shopee*, *Tokopedia*, dan *Lazada* sebagai indikator kepuasan konsumen di ranah digital.

Dalam ekosistem *social commerce* tersebut, komentar dan ulasan konsumen memainkan peran krusial dalam membentuk persepsi produk, pengalaman pengguna, dan keputusan pembelian. Komentar ini sering kali berisi opini eksplisit berupa kritik, saran, maupun pujian. Namun, volume data komentar yang besar bersifat tidak terstruktur, sehingga analisis manual menjadi tidak efisien dan *machine learning* menjadi pendekatan teoritis yang sangat diperlukan untuk mengekstraksi dan mengklasifikasikan opini pengguna secara objektif dan kuantitatif (Ashari et al., 2025).

Berbagai penelitian terdahulu telah mengaplikasikan analisis sentimen dalam konteks *e-commerce* dan media sosial. Indrayuni & Nuhadi (2025) melakukan analisis sentimen terhadap ulasan pengguna *TikTok Shop Seller Center* dengan membandingkan tiga algoritma, dimana *Support Vector Machine (SVM)* menunjukkan performa tertinggi dibandingkan *Naive Bayes* dan *Logistic Regression*. Temuan tersebut menegaskan pentingnya teknik analisis sentimen untuk memahami opini pengguna guna pengembangan platform. Selain itu, Utomo (2021) juga mengaplikasikan metode serupa untuk analisis sentimen di media sosial dan *e-commerce*. Namun, mayoritas studi terdahulu masih berfokus pada dataset umum atau platform yang berbeda, dan belum banyak yang berfokus pada satu produk spesifik, seperti Glad2Glow.

Disisi lain, metode pembobotan kata seperti *Term Frequency-Inverse Document Frequency (TF-IDF)* secara teoretis terbukti efektif dalam berbagai studi text mining karena kemampuannya menilai urgensi kata dalam korpus dokumen yang besar. Akan

tetapi, penerapannya pada kasus *TikTok Shop* dengan karakteristik bahasa Indonesia yang didominasi bahasa gaul (slang), singkatan, dan emoji masih kurang dieksplorasi secara mendalam. Terdapat kesenjangan penelitian (research gap) dalam memaksimalkan kombinasi metode pembobotan fitur dan algoritma klasifikasi dalam satu konteks spesifik social commerce yang memiliki karakter bahasa sangat singkat dan informal.

Melalui kerangka teoretis tersebut, penelitian ini mengisi kesenjangan yang ada dengan mengukur distribusi sentimen konsumen secara akurat pada produk *Glad2 Glow* di *TikTok Shop* menggunakan metode *TF-IDF* dan *Naive Bayes Classifier*. Evaluasi performa model melalui metrik akurasi, presisi, *recall*, dan *F1-score* secara implisit diproyeksikan mampu memberikan hasil klasifikasi yang optimal terhadap bahasa informal *social commerce*. Secara teoretis, penelitian ini diharapkan memperluas kajian metodologi analisis sentimen *e-commerce* di *TikTok Shop*. Secara praktis, pendekatan ini dapat menjadi landasan bagi pemilik merek maupun pelaku usaha lain dalam memahami persepsi pelanggan, merancang strategi pemasaran berbasis data empiris, serta meningkatkan kualitas produk berdasarkan opini konsumen nyata.

3. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah text mining, yaitu teknik pengolahan data teks yang bersifat tidak terstruktur untuk menghasilkan informasi yang bermakna. Menurut (Fauzi & Heri Yunial, 2024) text mining adalah teknologi yang digunakan untuk menganalisis data teks yang bersifat semi-terstruktur atau tidak terstruktur dengan memanfaatkan berbagai teknik seperti data mining, machine learning, dan natural language processing (NLP) untuk mengatasi masalah informasi yang berlebihan. Prosesnya melibatkan tahapan preprocessing, ekstraksi fitur, dan analisis representasi mengeah seperti pengelompokkan dan analisis tren. Proses klasifikasi sentiment dilakukan menggunakan algoritma Naïve Bayes sebagai metode supervised learning.

Menurut (Setiawan et al., 2023) algoritma Naïve Bayes merupakan alogritma supervised learning yang menggunakan peluang dengan pemelihan hasil akhir peluang terbesar untuk memasukkan data bar uke dalam kelasnya. Keuntungan algoritma Naïve Bayes adalah tingkat nilai error yang didapat lebih rendah ketika data set berjumlah besar, akurasi Naïve Bayes dan kecepatannya lebih tinggi pada saat diaplikasikan ke dalam data

set yang jumlahnya lebih besar. Sehingga, penelitian ini menggunakan pendekatan kuantitatif dengan metode text mining dan algoritma Naïve Bayes untuk menganalisis dan mengklasifikasikan sentiment komentar pengguna secara sistematis, terukur, dan berbasis data numerik. Pendekatan tersebut dipilih karena sesuai dengan tujuan penelitian ini, yaitu untuk membangun sistem analisis sentiment yang mampu menghasilkan informasi objektif mengenai persepsi konsumen terhadap produk yang sedang diteliti.

Tahapan Penelitian terdiri atas:

1. Pengumpulan Data
2. Preprocessing Text
3. Ekstraksi Fitur (TF-IDF)
4. Pembagian Dataset
5. Klasifikasi Sentimen
6. Evaluasi Model
7. Hasil analisis Sentimen

4. HASIL DAN PEMBAHASAN

Pada tahap ini dilakukan proses pengumpulan data berupa komentar pengguna terhadap produk Glad2Glow yang diperoleh dari platform TikTok Shop. Pengumpulan data dilakukan secara manual dengan mengamati dan mencatat komentar yang terdapat pada halaman produk Glad2Glow di TikTok Shop.

Data komentar yang berhasil dikumpulkan berjumlah 350 komentar yang terdiri dari berbagai opini pengguna, baik yang menunjukkan sentimen positif, netral, maupun negatif. Seluruh data yang telah dikumpulkan kemudian disusun ke dalam sebuah file Comma Separated Values (CSV) agar lebih mudah diproses pada tahapan analisis selanjutnya menggunakan bahasa pemrograman Python..

A. Implementasi Pengambilan Data

Setelah proses pengumpulan data selesai dilakukan, tahap berikutnya adalah membaca dataset yang telah disimpan dalam format CSV menggunakan bahasa pemrograman Python. Dataset dibaca menggunakan library Pandas sehingga data dapat dimuat ke dalam bentuk DataFrame untuk memudahkan proses pengolahan pada tahapan preprocessing dan klasifikasi sentimen

```
# 01. MEMBACA DATASET GLAD2GLOW
import pandas as pd
import numpy as np

df = pd.read_csv("Data ulasan Glad2Glow.csv")
print("="*50)
display(df.head())
```

Gambar 4.1 Implementasi Pembacaan Dataset Glad2Glow

Kode program pada Gambar 4.1 digunakan untuk membaca dataset Data ulasan Glad2Glow.csv menggunakan library Pandas. Fungsi `read_csv()` digunakan untuk memuat data ke dalam DataFrame, sedangkan fungsi `display(df.head())` digunakan untuk menampilkan lima data pertama sebagai bentuk verifikasi bahwa dataset berhasil dibaca dengan baik. Setelah dataset berhasil dimuat, data siap digunakan pada proses preprocessing, pembobotan kata menggunakan metode TF-IDF, serta proses klasifikasi menggunakan algoritma Naïve Bayes.

B. Hasil Pengumpulan Data

Hasil dari proses pengumpulan data berupa kumpulan komentar pengguna TikTok Shop terhadap produk Glad2Glow. Dataset yang digunakan dalam penelitian ini terdiri dari dua atribut utama, yaitu rating dan ulasan. Atribut rating menunjukkan penilaian pengguna terhadap produk dengan skala 1 sampai 5, sedangkan atribut ulasan berisi komentar atau pendapat pengguna mengenai pengalaman penggunaan produk Glad2Glow. Berikut merupakan sebagian data hasil pengumpulan komentar yang digunakan dalam penelitian.

Tabel 4.1 Contoh Data Komentar Pengguna Produk Glad2Glow

No	Tanggal	Rating	Ulasan
1	7-04-2026	3	Masih dicoba dulu, teksturnya cukup ringan, standar seperti serum lain.
2	29-05-2026	5	Kulit terasa lebih halus setelah rutin pemakaian!!!
3	18-04-2026	1	Jujur kecewa, menurut saya overhype, hasil biasa saja.

Berdasarkan data yang telah dikumpulkan, terlihat bahwa komentar pengguna memiliki beragam pendapat mengenai produk Glad2Glow. Sebagian pengguna memberikan komentar positif karena merasa produk mampu memberikan hasil yang sesuai harapan, sementara sebagian lainnya menyampaikan komentar netral maupun negatif berdasarkan pengalaman penggunaan yang berbeda. Variasi sentimen tersebut menjadi dasar dalam proses analisis sentimen pada penelitian ini. Selanjutnya, seluruh data akan melalui tahap preprocessing untuk membersihkan dan menyiapkan data sebelum dilakukan pembobotan menggunakan metode TF-IDF dan proses klasifikasi menggunakan algoritma Naïve Bayes.

C. Preprocessing Data

Preprocessing data merupakan tahapan yang dilakukan untuk membersihkan dan menyiapkan data komentar sebelum dilakukan proses pembobotan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan klasifikasi menggunakan algoritma Naïve Bayes. Tahapan ini bertujuan untuk menghilangkan karakter atau informasi yang tidak diperlukan sehingga data menjadi lebih terstruktur dan mudah diproses. Pada penelitian ini, preprocessing dilakukan melalui beberapa tahapan, yaitu cleaning, penghapusan emoji (emoji removal), pengurangan karakter yang berulang (reduce character), dan case folding. Hasil dari preprocessing akan menjadi data masukan pada proses pembobotan TF-IDF.

```
# --- 02. PREPROCESSING ---
import re
import nltk
nltk.download('punkt')

def cleaningulasan(content):
    ulasan = re.sub(r'[A-Za-z0-9]+', ' ', str(content))
    ulasan = re.sub(r'[A-Za-z0-9]+', ' ', ulasan)
    ulasan = re.sub(r'http\S+', ' ', ulasan)
    ulasan = re.sub(r'[0-9]+', ' ', ulasan)
    ulasan = re.sub(r'[\s\S]*?@.*?$', '', ulasan)
    ulasan = ulasan.strip(' ')
    return ulasan

def clearEmoji(content):
    return content.encode('ascii', 'ignore').decode('ascii')

def replaceTOM(content):
    pola = re.compile(r'(\w)\2+', re.DOTALL)
    return pola.sub(r'\1', content)

def casefoldingText(content):
    return content.lower()

def preprocess(df):
    df = df.copy()

    # kolom ulasan Glad2Glow
    df['cleaning'] = df['content'].apply(cleaningulasan)

    df['hapusEmoji'] = df['cleaning'].apply(clearEmoji)
    df['reduceChar'] = df['hapusEmoji'].apply(replaceTOM)
    df['casefolding'] = df['reduceChar'].apply(casefoldingText)

    return df

df_preprocessed = preprocess(df)
display(df_preprocessed.head())
```

Gambar 4.2 Implementasi Preprocessing Data

Kode program pada Gambar 4.2 digunakan untuk melakukan preprocessing terhadap data komentar pengguna produk Glad2Glow yang telah dibaca dari file Data ulasan Glad2Glow.csv. Proses diawali dengan cleaning untuk menghapus karakter yang tidak diperlukan, seperti tanda baca, angka, tautan (URL), mention, dan hashtag. Selanjutnya dilakukan penghapusan emoji, pengurangan karakter yang berulang, serta proses case folding yang mengubah seluruh huruf menjadi huruf kecil (lowercase). Hasil dari setiap tahapan preprocessing disimpan ke dalam kolom baru sehingga perubahan data dapat diamati sebelum digunakan pada proses pembobotan TF-IDF.

Tabel 4.2 Hasil Preprocessing Data

No	Ulasan	Cleaning	Case Folding
1	Masih dicoba dulu, Teksturnya cukup ringan, standar seperti serum lain	Masih dicoba dulu Teksturnya cukup ringan standar seperti serum lain	masih dicoba dulu teksturnya cukup ringan standar seperti serum lain
2	Kulit terasa lebih halus setelah rutin pemakaian!!!	Kulit terasa lebih halus setelah rutin pemakaian	kulit terasa lebih halus setelah rutin pemakaian
3	Jujur kecewa, Menurut saya overhype, hasil biasa saja	Jujur kecewa Menurut saya overhype hasil biasa saja	jujur kecewa menurut saya overhype hasil biasa saja

D. Pelabelan Data

Tahap sentiment labeling merupakan proses pemberian label sentimen pada setiap komentar pengguna yang telah melalui seluruh tahapan preprocessing. Label sentimen digunakan sebagai kelas target pada proses klasifikasi menggunakan algoritma Naïve Bayes. Pada penelitian ini, pelabelan sentimen dilakukan secara otomatis berdasarkan nilai rating yang diberikan oleh pengguna terhadap produk Glad2Glow di TikTok Shop. Label sentimen yang digunakan terdiri dari tiga kategori, yaitu Positif, Netral, dan Negatif. Proses pelabelan ini bertujuan untuk menghasilkan data yang telah memiliki kelas sehingga dapat digunakan pada tahap pembobotan TF-IDF dan proses klasifikasi.

Berikut merupakan implementasi kode program yang digunakan untuk melakukan proses sentiment labeling.

```
# --- 05. SENTIMENT LABELING ---

def label_sentiment(score):
    if score == 1 or score == 2:
        return 'Negatif'
    elif score == 3:
        return 'Netral'
    else:
        return 'Positif'

df_preprocessed['sentimen'] = df_preprocessed['score'].apply(label_sentiment)
print(df_preprocessed['sentimen'].value_counts())

display(df_preprocessed.head())
```

Gambar 4.3 Implementasi Sentiment Labeling

Kode program pada Gambar 4.3 digunakan untuk memberikan label sentimen berdasarkan nilai rating pada setiap komentar pengguna. Rating 1 dan 2 dikategorikan sebagai sentimen Negatif, rating 3 dikategorikan sebagai sentimen Netral, sedangkan rating 4 dan 5 dikategorikan sebagai sentimen Positif. Proses pelabelan dilakukan menggunakan fungsi `label_sentiment()`, kemudian diterapkan pada seluruh data sehingga menghasilkan kolom sentimen yang akan digunakan sebagai kelas target pada proses klasifikasi menggunakan algoritma Naïve Bayes.

Tabel 4.3 Hasil Sentiment Labeling

No	Rating	Hasil Preprocessing	Labeling
1	3	coba tekstur cukup ringan standar serum	Netral
2	5	kulit rasa halus rutin pakai	Positif
3	1	kulit rasa halus rutin pakai	Negatif

E. Visualisasi Distribusi Sentimen (BarChart)

Visualisasi distribusi sentimen dilakukan menggunakan diagram batang (bar chart) untuk menampilkan jumlah komentar pada masing-masing kategori sentimen, yaitu Positif, Negatif, dan Netral. Diagram batang memberikan gambaran yang jelas mengenai persebaran data hasil pelabelan sentimen sehingga memudahkan dalam melakukan

analisis terhadap kecenderungan pendapat pengguna mengenai produk Glad2Glow di TikTok Shop.

```
# --- 07. SENTIMENT BAR CHART ---

import matplotlib.pyplot as plt

sentimen_data = df_preprocessed['sentimen'].value_counts()

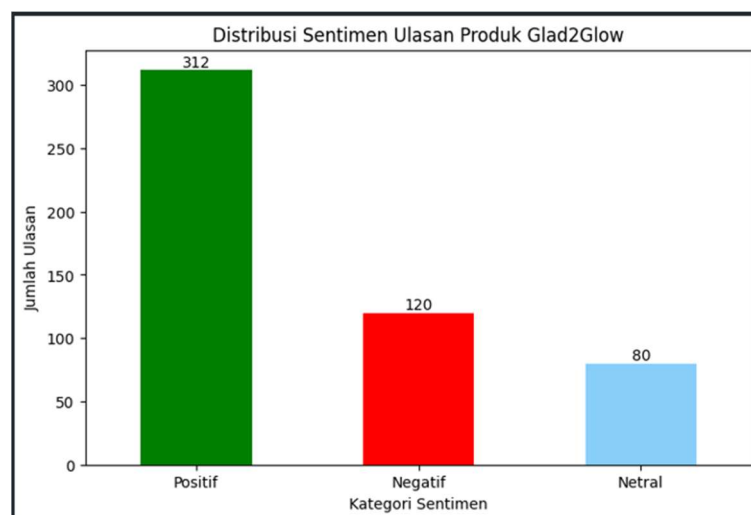
ax = sentimen_data.plot(
    kind='bar',
    color=['green', 'red', 'lightskyblue'],
    figsize=(8,5)
)

# Menampilkan jumlah pada setiap batang
for p in ax.patches:
    ax.annotate(
        str(int(p.get_height())),
        (p.get_x() + p.get_width()/2, p.get_height()),
        ha='center',
        va='bottom'
    )

plt.title("Distribusi Sentimen Ulasan Produk Glad2Glow")
plt.xlabel("Kategori Sentimen")
plt.ylabel("Jumlah Ulasan")
plt.xticks(rotation=0)
plt.show()
```

Gambar 4.4 Implementasi Visualisasi Distribusi Sentimen (Bar Chart)

Kode program pada Gambar 4.4 digunakan untuk menghitung jumlah data pada setiap kategori sentimen menggunakan fungsi `value_counts()`. Selanjutnya, hasil perhitungan divisualisasikan dalam bentuk diagram batang (bar chart) sehingga distribusi sentimen positif, negatif, dan netral dapat ditampilkan secara lebih informatif. Setiap batang mewakili jumlah komentar pada masing-masing kategori sentimen sehingga memudahkan dalam membandingkan banyaknya komentar pada setiap kelas.



Gambar 4.5 Visualisasi Distribusi Sentimen Menggunakan Bar Chart

Berdasarkan hasil visualisasi distribusi sentimen pada Gambar 4.5, diperoleh bahwa komentar dengan sentimen positif berjumlah 312 data, sentimen negatif sebanyak 120

data, dan sentimen netral sebanyak 80 data. Hasil tersebut menunjukkan bahwa sebagian besar pengguna memberikan tanggapan positif terhadap produk Glad2Glow di TikTok Shop. Hal ini mengindikasikan bahwa produk memperoleh respons yang cukup baik dari pengguna berdasarkan pengalaman mereka setelah menggunakan produk tersebut. Meskipun demikian, masih terdapat sejumlah komentar yang termasuk ke dalam kategori negatif dan netral, sehingga hasil analisis sentimen ini dapat menjadi bahan evaluasi bagi pihak terkait dalam meningkatkan kualitas produk maupun pelayanan kepada konsumen.

F. Visualisasi WordCloud

Visualisasi word cloud digunakan untuk menampilkan kata-kata yang paling sering muncul pada komentar pengguna produk Glad2Glow di TikTok Shop. Visualisasi ini bertujuan untuk memberikan gambaran mengenai kata-kata yang dominan pada setiap kategori sentimen sehingga karakteristik masing-masing sentimen dapat lebih mudah dianalisis. Pada penelitian ini, word cloud dibagi menjadi tiga kategori berdasarkan hasil pelabelan sentimen, yaitu Positif, Negatif, dan Netral.

```
# --- 06. WORDCLOUD VISUALIZATION ---

from wordcloud import WordCloud
import matplotlib.pyplot as plt

def draw_wordcloud(df, label, color_map):
    data = df[df['sentimen'] == label]

    all_text = ' '.join(
        [' '.join(tokens) for tokens in data['Stemming']]
    )

    wordcloud = WordCloud(
        width=1000,
        height=500,
        background_color='white',
        colormap=color_map
    ).generate(all_text)

    plt.figure(figsize=(12,6))
    plt.imshow(wordcloud, interpolation='bilinear')
    plt.axis('off')
    plt.title(f'WordCloud Sentimen {label} - Glad2Glow')
    plt.show()

# WordCloud setiap sentimen
draw_wordcloud(df_preprocessed, 'Negatif', 'Reds')
draw_wordcloud(df_preprocessed, 'Netral', 'Blues')
draw_wordcloud(df_preprocessed, 'Positif', 'Greens')
```

Gambar 4.6 Implementasi Visualisasi Word Cloud

Kode program pada Gambar 4.6 digunakan untuk menghasilkan visualisasi word cloud berdasarkan kategori sentimen. Proses diawali dengan memisahkan data sesuai label sentimen, yaitu Positif, Negatif, dan Netral. Selanjutnya, seluruh kata hasil proses

stemming pada masing-masing kategori digabungkan menjadi satu teks. Library WordCloud kemudian digunakan untuk menghasilkan visualisasi berdasarkan frekuensi kemunculan kata. Semakin sering suatu kata muncul, maka ukuran kata pada word cloud akan semakin besar sehingga memudahkan dalam mengidentifikasi kata-kata yang paling dominan pada setiap kategori sentimen.

G. Ekstraksi Fitur (TF-IDF)

Data komentar pengguna produk Glad2Glow yang telah melalui seluruh tahapan preprocessing masih berbentuk data teks sehingga belum dapat digunakan secara langsung pada proses klasifikasi menggunakan algoritma Naïve Bayes. Oleh karena itu, setiap kata perlu dikonversi ke dalam bentuk numerik melalui proses pembobotan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode TF-IDF digunakan untuk menghitung tingkat kepentingan setiap kata berdasarkan frekuensi kemunculannya pada suatu dokumen dan keseluruhan dokumen dalam dataset. Hasil dari proses pembobotan ini berupa nilai numerik yang merepresentasikan bobot masing-masing kata sehingga dapat digunakan sebagai fitur masukan pada proses klasifikasi sentimen menggunakan algoritma Naïve Bayes. Berikut merupakan implementasi kode program yang digunakan untuk melakukan proses pembobotan TF-IDF, sedangkan hasil pembobotannya dapat dilihat pada Tabel 4.4.

```
# --- 08. TF-IDF VECTOR & NAIVE BAYES MODELING ---

from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.preprocessing import LabelEncoder
from sklearn.naive_bayes import MultinomialNB

# Ambil fitur dan label
X = df_preprocessed['Stemming']
Y = df_preprocessed['sentimen']

# Gabungkan token menjadi kalimat
x_text = [' '.join(tokens) for tokens in X]

# Split data
x_train, x_test, y_train, y_test = train_test_split(
    x_text,
    Y,
    test_size=0.2,
    random_state=42,
    stratify=Y
)
```

Gambar 4.7 Ekstraksi Fitur (TF-IDF)

Tabel 4.4 Pembobotan Kata TF-IDF

<i>Term</i>	<i>Weigh</i>
nan	124.00
pakai	24.83

kulit	22.91
banget	22.22
suka	20.61
serum	18.96
hasil	17.28
cocok	15.78
wajah	15.12
kecewa	14.91

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian analisis sentimen komentar TikTok Shop terhadap produk Glad2Glow menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan algoritma Naïve Bayes, dapat disimpulkan bahwa metode tersebut mampu mengklasifikasikan 350 data komentar ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Hasil pengujian menunjukkan nilai akurasi sebesar 98,06%, dengan nilai precision, recall, dan F1-score yang tinggi pada setiap kelas sentimen. Selain itu, sistem analisis sentimen berbasis website berhasil diimplementasikan dan seluruh fungsinya berjalan sesuai dengan yang diharapkan berdasarkan pengujian black box. Hasil penelitian ini menunjukkan bahwa kombinasi TF-IDF dan Naïve Bayes dapat digunakan untuk menganalisis persepsi konsumen terhadap produk Glad2Glow berdasarkan komentar pengguna TikTok Shop. Untuk pengembangan selanjutnya, disarankan agar penelitian menggunakan jumlah data komentar yang lebih besar, lebih beragam, dan lebih terbaru agar model dapat mengenali pola sentimen secara lebih menyeluruh. Penelitian juga dapat dikembangkan dengan memanfaatkan data dari platform digital lain serta membandingkan algoritma Naïve Bayes dengan metode klasifikasi lainnya, seperti Support Vector Machine (SVM), Random Forest, atau Deep Learning, guna memperoleh perbandingan performa yang lebih komprehensif dan meningkatkan kualitas hasil klasifikasi sentimen..

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumateta Utara atas dukungan fasilitas penelitian.

DAFTAR REFERENSI

- Ahmad Fauzi, & Agus Heri Yunial. (2024). Analisis Sentimen (Sentiment Analysis). CV. Bintang Semesta Media.
- Ashari, A., Dwi, G., Harahap, F., & Rosnelly, R. (2025). Tokopedia Menggunakan Mesin Learning. 132–140.
- Aulia Fitri Maharan, Christian Wiradendi Wolor, & Roni Faslah. (2024). Analisis Sistem Informasi Pengolahan Data Gugus Tugas Pemenangan Caleg X. *Jurnal Multidisiplin Ilmu Akademik*, 1(4), 629–636. <https://doi.org/10.61722/jmia.v1i4.2347>
- I Kadek Dwi Gandika Supartha, Elly, Nofri Yudi Arifin, Achmad Ridwan, Tubagus Riko Rivanthio, Putu Praba Santika, Hita, & Dewi Primasari. (2023). Buku Ajar Analisis Perancangan Sistem. PT. Sonpedia Publishing Indonesia.
- Iman Subasman, Ansar CS, Ade Suhara, Nur Sehang Thamrin, Nanny Mayasari, & Achmad Harristhana Mauldfi Sastraatmadja. (2025). Metode dan Teknik Penelitian Kuantitatif, Kualitatif, dan Pengembangan untuk Mahasiswa. Penerbit Widina.
- Indrayuni, E., & Acmad Nurhadi. (2025). Analisis Sentimen Aplikasi Tiktok Shop Seller Center Menggunakan Naive Bayes, Svm Dan Logistic Regression. In *INTI Nusa Mandiri* (Vol. 20, Number 1, pp. 26–34).
- Jl, A., No, W., Semarang, K., & Tengah, J. (2025). Strategi Engagement Konsumen dan Brand di Instagram : Studi Kasus tindakan . Selain mengandalkan konten publik , Glad2Glow juga perlu mengembangkan strategi konsisten , mempertahankan gaya visual yang kohesif , dan menyisipkan momen interaktif. 32–45. *Jurnal 2*. (n.d.).
- Khairunnisa, G., & Voutama, A. (2024). Peminjaman Inventaris Berbasis Web Di Bem Fasilkom Unsika. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 48–55.
- Pudya Ardhana, V. Y. (2021). Perancangan Sistem Informasi Apotek Qamarul Huda Menggunakan Unified Modeling Language (UML). *Jurnal Kesehatan Qamarul Huda*, 9(2), 115–119.
- Septiani, A., Voutama, A., Siska, S., Andri Hendriadi, A., & Heryana, N. (2024). Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Regulasi Tiktok Shop Pada Media Sosial X (Twitter). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(4), 5729–5735. <https://doi.org/10.36040/jati.v8i4.10040>
- Zaenab Kurnia, Amalina Maryam Zakiyyah, Nur Qodariyah Fitriyah, & Agus Milu Susetyo. (2024). Analisis Sentimen Masyarakat Berdasarkan Komentar Kerja Sama Tiktok Shop dan Tokopedia di Instagram Menggunakan Metode Naïve Bayes Classifier. *Jurnal Penelitian Teknologi Informasi Dan Sains*, 2(2), 115–125. <https://doi.org/10.54066/jptis.v2i2.1978>
- Zhafira, D. F., Rahayudi, B., & Indriati. (2021). Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naïve Bayes dan Pembobotan TF-IDF. *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi (JUST-SI)*, 2(1), 55–63.
- Zunan Setiawan, Muhammad Fajar, Arif Mudi Priyatno, & Anggi Yhurinda Perdana Putri. (2023). Buku Ajar Data Mining. PT. Sonpedia Publishing Indonesia.