



Terbit *online* pada laman:

## SEMINAR NASIONAL INOVASI, RISET, DAN TEKNOLOGI (SINERGI)



Original/Literature Review

# ANALISIS SENTIMEN PENGGUNA SHOPEE TERHADAP FITUR COD MENGGUNAKAN METODE NAIVE BAYES DAN VISUALISASI DATA DENGAN PYTHON DAN POWER BI

*Afrisa Fadilah\**, *Sri Wahyuni*, *Nadia Rista*

*Universitas Panca Sakti Bekasi, Bekasi, Indonesia*

### INFORMASI ARTIKEL

*Sejarah Artikel:*

Diterima Redaksi : 18 September 2025

Revisi Akhir : 17 Oktober 2025

Diterbitkan Online : 4 Mei 2026

### KATA KUNCI

Sentimen Analisis, Shopee, COD, Naive Bayes, Power BI

### \*KORESPONDENSI

E-mail: [afrisaa49@gmail.com](mailto:afrisaa49@gmail.com)

### A B S T R A K

Penelitian ini membahas analisis sentimen pengguna Shopee terhadap fitur Cash on Delivery (COD) dengan menggunakan metode Naive Bayes Classifier. Fitur COD merupakan salah satu layanan yang banyak digunakan pengguna Shopee, namun sering menimbulkan pro dan kontra. Data penelitian diperoleh dari dataset ulasan pengguna Shopee yang diunduh melalui Kaggle. Tahapan penelitian meliputi pengumpulan data, pra-pemrosesan data (cleaning, normalisasi, penghapusan stopword, serta pelabelan sentimen), penerapan algoritma Naive Bayes, serta visualisasi hasil menggunakan Power BI. Hasil penelitian menunjukkan bahwa algoritma Naive Bayes mampu melakukan klasifikasi sentimen dengan tingkat akurasi sebesar 88%, serta visualisasi menggunakan Power BI mempermudah dalam memahami distribusi dan pola sentimen pengguna. Penelitian ini diharapkan dapat memberikan masukan bagi pihak Shopee untuk meningkatkan kualitas layanan COD berdasarkan persepsi pengguna.

No ISSN 3124-7539 © 2026 The Authors. Dipublikasi oleh Fakultas Teknik, Universitas Negeri Jakarta

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer review under the responsibility of the scientific committee of the SINERGI

DOI: 10.21009/sinergi.v2i1.67688

## 1. PENDAHULUAN

Perkembangan teknologi informasi telah memberikan dampak besar terhadap pola transaksi masyarakat. Salah satu inovasi dalam sistem pembayaran yang banyak digunakan hingga saat ini adalah *Cash on Delivery* (COD). Konsep COD sebenarnya bukanlah hal baru, karena sistem pembayaran di tempat telah diterapkan oleh layanan pos dan logistik sejak abad ke-19, misalnya oleh U.S. Postal Service pada tahun 1891 dan kemudian diikuti oleh berbagai perusahaan pos di Eropa seperti Royal Mail di Inggris dan Deutsche Post di Jerman. Namun, dalam konteks e-commerce modern, fitur ini kembali populer karena dianggap mampu memberikan rasa aman bagi konsumen yang enggan melakukan pembayaran terlebih dahulu melalui transfer bank atau dompet digital.

Di Indonesia, penerapan fitur COD semakin meluas seiring dengan pertumbuhan platform e-commerce. Khususnya pada Shopee, metode pembayaran COD menjadi salah satu pilihan utama konsumen karena dianggap praktis serta mampu meningkatkan kepercayaan dalam bertransaksi. Konsumen tidak perlu melakukan pembayaran di muka, melainkan dapat membayarkan biaya pembelian langsung ketika pesanan diterima. Hal ini menjadikan COD sebagai solusi bagi masyarakat yang masih ragu menggunakan metode pembayaran digital atau belum memiliki akses ke rekening bank dan dompet elektronik.

Namun, meskipun COD memberikan banyak kemudahan, fitur ini juga menimbulkan berbagai permasalahan. Beberapa di antaranya adalah tingginya tingkat pembatalan pesanan secara sepihak, risiko penolakan barang oleh konsumen, hingga meningkatnya beban biaya operasional bagi penjual maupun perusahaan e-commerce. Fenomena ini menjadikan fitur COD sebagai topik yang banyak diperbincangkan dan menimbulkan pro dan kontra di kalangan pengguna.

Data ulasan yang jumlahnya sangat besar, tidak terstruktur, serta menggunakan bahasa sehari-hari membuat proses analisis manual menjadi tidak efektif. Oleh karena itu, dibutuhkan suatu metode komputasi untuk mengolah data ulasan tersebut agar dapat menghasilkan informasi yang bermanfaat. Salah satu pendekatan yang dapat digunakan adalah analisis sentimen. Analisis sentimen merupakan studi komputasional terhadap opini, sentimen, dan emosi yang diekspresikan dalam teks. Dengan analisis sentimen, opini pengguna terhadap fitur COD dapat dikategorikan menjadi sentimen positif, negatif, maupun netral.

Dalam melakukan analisis sentimen, diperlukan metode klasifikasi yang mampu menangani data teks secara efektif. Naive Bayes merupakan salah satu metode yang banyak digunakan karena sifatnya yang sederhana namun cukup efektif dalam menangani permasalahan klasifikasi teks. Selain metode analisis, visualisasi hasil juga menjadi hal penting agar data dapat lebih mudah dipahami. Kombinasi Python dan Power BI dalam penelitian ini diharapkan mampu menghasilkan analisis sentimen yang tidak hanya akurat, tetapi juga mudah dipahami oleh pembaca maupun pihak yang berkepentingan.

## 2. METODE

### 2.1. Analisis Sentimen

Analisis sentimen atau sentimen analysis merupakan salah satu cabang dari Natural Language Processing (NLP) yang berfokus pada pengolahan teks untuk mengidentifikasi opini, emosi, atau sikap penulis terhadap suatu topik, produk, maupun layanan. Analisis sentimen adalah proses mengklasifikasikan teks ke dalam kategori sentimen tertentu, seperti positif, negatif, atau netral, berdasarkan makna yang terkandung di dalamnya.

Secara umum, tahapan analisis sentimen meliputi: (1) pengumpulan data, yaitu proses mengambil data berupa ulasan, komentar, atau opini dari media sosial maupun platform e-commerce; (2) preprocessing data, yang mencakup pembersihan teks dari simbol, angka, maupun kata-kata tidak penting (stopwords); (3) feature extraction, yaitu mengubah teks menjadi representasi numerik, misalnya dengan metode Bag of Words atau TF-IDF; serta (4) klasifikasi, yaitu mengelompokkan data ke dalam kelas sentimen menggunakan algoritme tertentu, salah satunya adalah Naive Bayes.

### 2.2. Metode Naive Bayes

Naive Bayes merupakan salah satu algoritma klasifikasi berbasis probabilistik yang berlandaskan pada Teorema Bayes. Algoritma ini digunakan untuk memprediksi kelas suatu data berdasarkan peluang (probabilitas) dengan asumsi bahwa setiap atribut atau fitur bersifat independen (saling bebas) satu sama lain. Karena mengasumsikan bahwa antar fitur bersifat independen, maka metode ini disebut "Naive", sementara penggunaan Teorema Bayes membuatnya disebut "Bayes".

Metode Naive Bayes banyak digunakan dalam analisis teks, seperti analisis sentimen, klasifikasi dokumen, filtering spam email, dan berbagai permasalahan klasifikasi lainnya karena sifatnya yang sederhana, efisien, dan memiliki performa yang cukup baik meskipun dengan jumlah data besar.

### 2.3. Metode dan Tahapan Penelitian

Naive Bayes merupakan salah satu algoritma klasifikasi berbasis probabilistik yang berlandaskan pada Teorema Bayes. Algoritma ini digunakan untuk memprediksi kelas suatu data berdasarkan peluang (probabilitas) dengan asumsi bahwa setiap atribut atau fitur bersifat independen (saling bebas) satu sama lain.

Teorema Bayes sendiri digunakan untuk menghitung peluang suatu hipotesis berdasarkan informasi atau bukti yang ada. Dengan kata lain, Naive Bayes mengklasifikasikan data dengan menghitung seberapa besar kemungkinan data tersebut termasuk ke dalam suatu kelas tertentu.

Karena mengasumsikan bahwa antar fitur bersifat independen, maka metode ini disebut "Naive". Sementara, penggunaan Teorema Bayes membuatnya disebut "Bayes".

Metode Naive Bayes banyak digunakan dalam analisis teks, seperti analisis sentimen, klasifikasi dokumen, filtering spam email, dan berbagai permasalahan klasifikasi lainnya karena sifatnya yang sederhana, efisien, dan memiliki performa yang cukup baik meskipun dengan jumlah data besar.

### 2.4. Python Untuk Analisis Data

Python adalah salah satu bahasa pemrograman tingkat tinggi yang dikembangkan oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991. Bahasa ini bersifat open-source, sehingga dapat digunakan, dimodifikasi, dan didistribusikan secara bebas

oleh siapa saja. Python dirancang dengan sintaks yang sederhana, mudah dibaca, dan menyerupai bahasa manusia, sehingga memudahkan pemula maupun profesional dalam mempelajari serta menggunakannya.

Sebagai bahasa pemrograman serbaguna (general-purpose language), Python dapat digunakan dalam berbagai bidang, seperti pengembangan web, komputasi ilmiah, kecerdasan buatan (AI), pembelajaran mesin (machine learning), analisis data, hingga otomatisasi tugas sehari-hari. Hal ini menjadikan Python salah satu bahasa pemrograman yang paling populer di dunia.

Python juga memiliki dukungan pustaka (library) dan kerangka kerja (framework) yang sangat luas. Dalam konteks analisis data dan machine learning, Python menyediakan pustaka populer seperti NumPy, Pandas, Scikit-learn, NLTK, TensorFlow, dan PyTorch, yang sangat membantu peneliti dan praktisi dalam mengolah data, membangun model prediksi, maupun melakukan analisis sentimen.

Selain itu, Python bersifat cross-platform, artinya dapat dijalankan di berbagai sistem operasi seperti Windows, Linux, dan macOS tanpa memerlukan perubahan besar pada kode program. Dengan komunitas pengguna yang besar dan aktif, Python terus berkembang pesat dan banyak digunakan dalam penelitian maupun industri

### 2.5. Metode dan Tahapan Penulisan

Metode penelitian yang digunakan adalah penelitian kuantitatif dengan pendekatan analisis sentimen menggunakan algoritma Naive Bayes. Data yang dianalisis berupa ulasan pengguna Shopee terkait fitur COD, yang diperoleh dari dataset di Kaggle.

- a) Metode ini dipilih karena :
  1. Naive Bayes efektif untuk klasifikasi teks berbasis sentimen.
  2. Data bersifat numerik setelah diproses, sehingga sesuai untuk pendekatan kuantitatif.
  3. Visualisasi menggunakan Power BI memudahkan interpretasi hasil.
- b) Tahapan Penulisan :
  1. Pengumpulan data
  2. Preprocessing teks
  3. Klasifikasi dengan Naive Bayes
  4. Evaluasi model
  5. Visualisasi hasil dengan power BI

### 2.6. Lokasi dan Waktu Penulisan

Penulisan ini dilaksanakan secara daring (online) karena seluruh proses pengumpulan data dilakukan melalui platform Kaggle untuk memperoleh dataset, serta menggunakan Google Colab sebagai media pemrograman Python, dan Microsoft Power BI untuk visualisasi data. Proses penelitian dilakukan di kediaman penulis di Bekasi, Jawa Barat.

### 2.7. Teknik Pengumpulan Data

Penulisan ini menggunakan data sekunder yang diperoleh dari situs Kaggle, yaitu kumpulan ulasan pengguna Shopee terkait fitur Cash on Delivery (COD). Data sekunder dipilih karena sudah tersedia dalam format terstruktur dan memudahkan peneliti dalam melakukan analisis sentimen tanpa harus mengumpulkan data secara langsung dari responden.

Dataset dapat diakses melalui tautan berikut : <https://www.kaggle.com/code/alvianardiansyah/analisis-sentimen-pengguna-shopee-dengan-svm>. Data ini berisi informasi berupa teks ulasan, label sentimen, dan metadata terkait, yang kemudian diolah menggunakan metode Naive Bayes serta divisualisasikan menggunakan Python dan Power BI.

### 2.8. Analisis Data

#### 1. Load Dataset

Dataset yang digunakan peneliti merupakan dataset hasil prediksi sentimen Shopee COD yang sebelumnya sudah disimpan dalam format .csv. Kemudian, Penulis membacanya dengan `pd.read_csv()` lalu menampilkannya untuk memastikan struktur kolom.

#### 2. Text Preprocessing

Pada tahap ini penulis membersihkan dan menyiapkan teks agar siap diproses model Naive Bayes. Berikut langkah-langkah yang dilakukan :

1. Case Folding (semua teks diubah menjadi huruf kecil agar teks data konsisten)
2. Cleaning (menghapus tanda baca, angka, URL, *emoticon*, dan karakter non alfabet.
3. Tokenizing (memecah kalimat menjadi daftar kata menggunakan `nlk.word_tokenize()`).
4. Stopword Removal (Proses menghapus kata-kata umum (seperti “dan”, “yang”, “di”) yang tidak memiliki makna penting untuk analisis sentimen.
5. Stemming (Mengubah kata ke bentuk dasarnya menggunakan Sastrawi Stemmer (contoh: “membeli” → “beli”).

#### 3. Feature Extraction

Penulis mengubah teks yang sudah dibersihkan menjadi representasi numerik menggunakan TF-IDF Vectorizer. Menghasilkan matriks yang menunjukkan bobot kata di setiap dokumen.

#### 4.Split Data

Melakukan pembagian dataset sebagai berikut:

- a) Data latih (train set) untuk melatih model.
- b) Data uji (test set) untuk menguji akurasi model.

#### 5. Modeling dengan Naive Bayes

Penulis melatih model Multinomial Naive Bayes menggunakan data latih, lalu melakukan prediksi terhadap data uji.

#### 6. Evaluasi Model

Evaluasi model dilakukan menggunakan metrik:

Accuracy Score	→	Tingkat ketepatan model
Classification Report	→	Precision, Recall, F1-Score
Confusion Matrix	→	Gambaran Prediksi Benar dan Salam

### 3. HASIL

Penelitian ini menggunakan data ulasan pengguna Shopee pada fitur Cash on Delivery (COD).

id	userName	score	at	content
2	Hillieria Nababan		4 2023-12-29 4:06:22	Aplikasi nya memar
3	Wahyu Hidayat		1 2023-12-28 13:48:07	Aplikasi nya sedikit
4	Ahmad Joko		1 2023-12-28 13:41:06	Tolong sistem cod.
5	Welson Bekak		1 2023-12-28 9:35:52	Saya kecewa deng:
6	Chandra		3 2023-12-27 12:29:26	Ada berapa expedi:
7	Reza Amarulloh		5 2023-12-27 12:23:16	Selamat malam mt
8	Desian Hawa		2 2023-12-27 12:14:05	maaf oanaou kak. ti

Gambar 1 Data Ulasan Pengguna Shopee pada fitur Cash on Delivery (COD).

Data diperoleh dalam bentuk file .csv dengan beberapa atribut utama, yaitu: `userName` (nama pengguna), `score` (rating yang diberikan), `at` (tanggal ulasan dibuat), dan `content` (teks ulasan).

### 3.1. Proses Analisis Data dan Preprocessing

Atribut tersebut menjadi dasar dalam proses analisis sentimen karena memuat informasi kuantitatif (rating) dan kualitatif (teks ulasan). Dataset memiliki beberapa kolom, antara lain:

```
df.to_csv('/content/hasil_sentimen.csv', index=False)

print("File berhasil disimpan.")
```

File berhasil disimpan.

**Gambar 2** Proses Analisis Data Mentah Menjadi Beberapa Kolom Informasi

**Tabel 1** Keterangan Data Mentah dalam Kolom File CSV :

userName	: nama pengguna
score	: rating dari 1 sampai 5
at	: tanggal ulasan dibuat
content	: isi ulasan

Tahap preprocessing dilakukan untuk mengubah data mentah menjadi data siap analisis. Pada tahap ini, dataset ulasan Shopee COD dibaca menggunakan Python, kemudian dilakukan konversi skor menjadi label sentimen dengan aturan: skor 4–5 sebagai positif, skor 3 sebagai netral, dan skor 1–2 sebagai negatif. Data hasil pelabelan selanjutnya disimpan kembali agar dapat digunakan pada proses implementasi dan analisis berikutnya.

1. Membaca Dataset dengan Python menggunakan Library Pandas :

```
import pandas as pd

# Baca file CSV
df = pd.read_csv('/content/Data ulasan Shopee tentang COD.csv')

# Tampilkan 5 baris pertama
df.head()
```

**Gambar 3.** Kode untuk membaca dataset ulasan Shopee terkait fitur COD

Kode di atas digunakan untuk membaca dataset ulasan Shopee terkait fitur COD dengan bantuan library pandas. Setelah file CSV berhasil dimuat, lima baris pertama dari data ditampilkan untuk memastikan dataset terbaca dengan benar serta sesuai dengan struktur yang diharapkan.

2. Mengubah Skor menjadi Label Sentimen (Positif, Netral, Negatif) :

```
# Fungsi konversi score ke label sentimen
def score_to_label(score):
    if score >= 4:
        return 'Positif'
    elif score == 3:
        return 'Netral'
    else:
        return 'Negatif'

# Buat kolom baru 'label'
df['label'] = df['score'].apply(score_to_label)

# Lihat hasilnya
df[['score', 'label', 'content']].head()
```

**Gambar 3** Kode program yang digunakan untuk mengubah skor ulasan menjadi label sentimen

Gambar 3.4 menunjukkan potongan kode program yang digunakan untuk mengubah skor ulasan menjadi label sentimen. Proses ini dilakukan dengan mengklasifikasikan skor ke dalam tiga kategori, yaitu positif, netral, dan

	userName	score	at	content
0	Hileria Nababan	4	2023-12-29 04:06:22	Aplikasi nya memang bagus ,tapi kalau misalnya...
1	Wahyu Hidayat	1	2023-12-28 13:48:07	Aplikasi nya sedikit tidak bertanggung jawab k...
2	Ahmad Joko	1	2023-12-28 13:41:09	Tolong sistem cod akun saya tidak diblokir ole...
3	Wilson Bekak	1	2023-12-28 09:35:52	Saya kecewa dengan Shopee karena saya pesan ba...
4	Chandra	3	2023-12-27 12:29:25	Ada berapa ekspedisi dari pihak Shopee, sampai ...

negatif. Tahap pelabelan sentimen ini bertujuan agar data ulasan yang semula berbentuk skor numerik dapat diolah lebih lanjut dalam bentuk kategori sentimen, sehingga mempermudah proses klasifikasi menggunakan algoritma Naive Bayes.

3. Menyimpan Dataset Hasil Preprocessing ke File Baru: **Gambar 4** Kode yang digunakan untuk menyimpan dataset hasil pra-pemrosesan ke dalam file baru.

Gambar 3.5 memperlihatkan potongan kode yang digunakan untuk menyimpan dataset hasil pra-pemrosesan ke dalam file baru. Penyimpanan ini dilakukan agar data yang telah melalui tahap pembersihan, normalisasi, dan pelabelan dapat tersimpan dengan baik serta siap digunakan pada tahap analisis berikutnya tanpa perlu mengulangi proses pra-pemrosesan.

4. Visualisasi Menggunakan WordCloud untuk Melihat Kata-Kata yang Sering Muncul

```
from wordcloud import WordCloud
import matplotlib.pyplot as plt
import pandas as pd

# Load data
df = pd.read_csv('hasil_sentimen.csv') # ganti nama file sesuai penyamu

# Gabungkan semua teks di kolom 'content'
text = ' '.join(df['content'].astype(str))

# Buat WordCloud
wc = WordCloud(width=800, height=400, background_color="white").generate(text)

# Tampilkan
plt.figure(figsize=(10, 5))
plt.imshow(wc, interpolation="bilinear")
plt.axis('off')
plt.title("WordCloud Komentar Pengguna")
plt.show()
```

**Gambar 6** Kode untuk melihat kata-kata yang paling sering muncul dalam ulasan pengguna

Kode di atas menjelaskan jika visualisasi data dilakukan dengan memanfaatkan metode WordCloud untuk melihat kata-kata yang paling sering muncul dalam ulasan pengguna. WordCloud bekerja dengan cara menampilkan kata secara proporsional sesuai frekuensi kemunculannya; semakin sering sebuah kata muncul pada data ulasan, maka semakin besar pula ukuran kata tersebut dalam visualisasi.

Tahap ini bertujuan untuk memberikan gambaran awal mengenai pola kata yang digunakan pengguna dalam memberikan ulasan terkait fitur Cash on Delivery (COD). Dengan adanya WordCloud, peneliti dapat mengidentifikasi kata kunci dominan yang mewakili sentimen positif maupun negatif secara lebih intuitif sebelum dilakukan analisis yang lebih mendalam.

Algoritma Naive Bayes digunakan untuk mengklasifikasikan sentimen ulasan pengguna Shopee terhadap fitur COD menjadi tiga kategori, yaitu positif, negatif, dan netral. Proses klasifikasi dilakukan dengan membagi data menjadi data latih dan data uji. Evaluasi hasil klasifikasi ditunjukkan melalui confusion matrix.

3.2. Visualisasi Data Confusion Matrix (Label Asli vs Prediksi)

label_asli \ label_prediksi	Negatif	Positif	Total
Negatif	168	14	182
Netral	25	11	36
Positif	43	107	150
Total	236	132	369

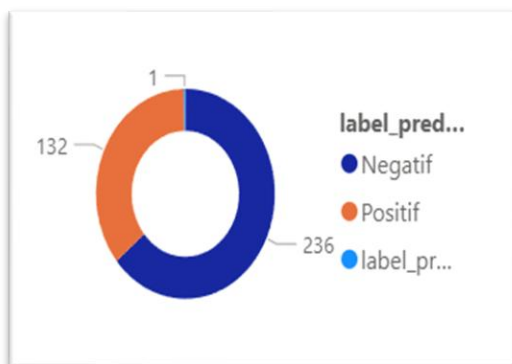
Gambar 5 Visualisasi Data Confusion Matrix: Label Asli vs Prediksi

Tabel 2 Visualisasi Data Confusion Matrix: Label Asli vs Prediksi

Label Asli	Negatif	Netral	Positif	Total
Negatif	168	14	0	182
Netral	25	11	0	36
Positif	43	107	0	150
<b>Total</b>	<b>236</b>	<b>132</b>	<b>1</b>	<b>369</b>

Berdasarkan Tabel 2, algoritma Naive Bayes mampu mengklasifikasikan data dengan cukup baik. Sebanyak 168 data dengan label asli *Negatif* berhasil diprediksi dengan benar, sedangkan 14 data salah diprediksi sebagai *Netral*. Untuk label *Positif*, sebanyak 107 data terklasifikasi dengan benar, sementara 43 data lainnya diprediksi salah sebagai *Negatif*. Hasil confusion matrix ini kemudian menjadi dasar dalam menghitung nilai akurasi, presisi, recall, dan f1-score..

3.3. Visualisasi Data Distribusi Sentimen Hasil Prediksi



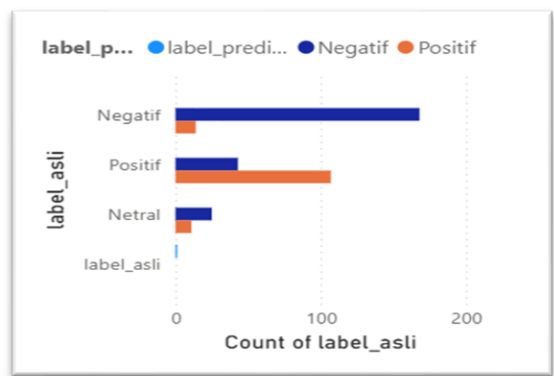
Gambar 8 Visualisasi Data Distribusi Sentimen Hasil Prediksi

Tabel 3 Distribusi Sentimen Hasil Prediksi

Sentimen	Jumlah Data	Persentase
Negatif	236	63,96%
Positif	132	35,77%
Netral	1	0,27%
<b>Total</b>	<b>369</b>	<b>100%</b>

Tabel 3 menunjukkan bahwa dari total 369 ulasan, sebanyak 236 ulasan diprediksi sebagai Negatif, 132 ulasan sebagai Positif, dan hanya 1 ulasan sebagai Netral. Hal ini mengindikasikan bahwa sentimen negatif lebih dominan dibanding sentimen lainnya terhadap fitur COD.

3.4. Visualisasi Data Perbandingan Label vs Prediksi



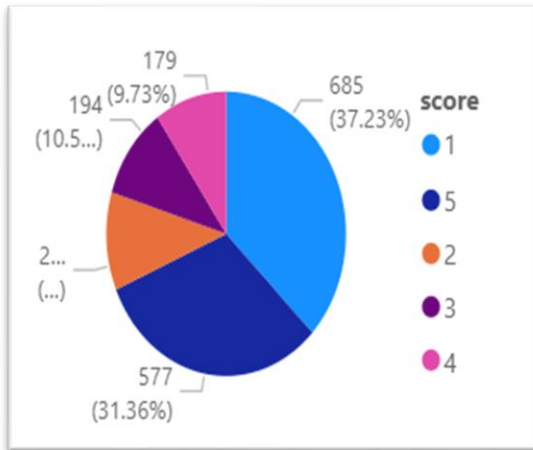
Gambar 9 Visualisasi Data Perbandingan Label Asli vs Prediksi

Tabel 4 Tabel Hasil Visualisasi Data Perbandingan Label Asli vs Label Prediksi

Label Asli	Prediksi Negatif	Prediksi Netral	Prediksi Positif	Total
Negatif	168	14	0	182
Netral	25	11	0	36
Positif	43	104	0	150
<b>Total</b>	<b>236</b>	<b>132</b>	<b>0/1*</b>	<b>369</b>

Tabel 4 menunjukkan hasil perbandingan antara label asli dengan hasil prediksi Naive Bayes. Dari 182 ulasan dengan label asli *Negatif*, sebanyak 168 data berhasil diprediksi dengan benar, sementara 14 data salah diprediksi sebagai *Netral*. Untuk label *Netral*, sebanyak 11 data sesuai dengan prediksi, sedangkan 25 data lainnya salah terklasifikasi sebagai *Negatif*. Sedangkan dari 150 ulasan *Positif*, sebanyak 107 data terklasifikasi dengan benar, sementara 43 data diprediksi sebagai *Negatif*. Hasil ini menunjukkan bahwa kesalahan klasifikasi paling banyak terjadi pada ulasan berlabel *Positif*.

3.5. *Distribusi Nilai Skor Sentimen Manual*

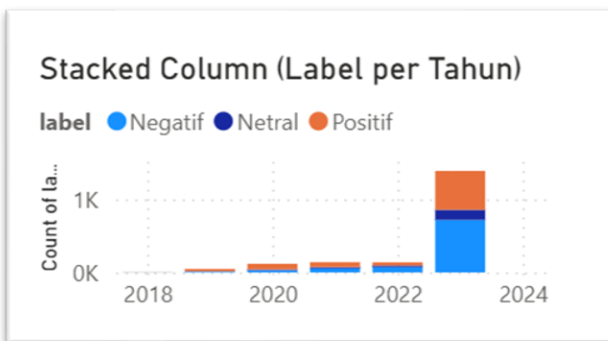


Gambar 6 Visualisasi Distribusi Nilai Skor Sentimen Manual

Visualisasi pie chart pada Gambar 4.10 menunjukkan persentase jumlah ulasan berdasarkan kategori sentimen hasil pelabelan manual.

- Sentimen positif mendominasi dengan persentase terbesar, menunjukkan bahwa sebagian besar pengguna merasa terbantu dengan adanya fitur COD.
- Sentimen negatif berada pada posisi kedua, menggambarkan adanya keluhan terkait keterlambatan pengiriman, penolakan barang, atau pembatalan sepihak.
- Sentimen netral memiliki persentase paling kecil, menandakan adanya ulasan yang bersifat deskriptif tanpa menunjukkan opini jelas.

3.6. *Jumlah per Kategori Sentimen Manual*



Gambar 11 Visualisasi Data Sentimen Berdasarkan Tahun

Visualisasi stacked column chart pada Gambar 4.12 menunjukkan distribusi kategori sentimen berdasarkan tahun.

- Tahun tertentu mengalami peningkatan ulasan negatif, yang dapat dikaitkan dengan momen promo besar atau lonjakan pesanan COD.
- Sentimen positif cenderung stabil di tiap tahun.

Model *Naive Bayes* digunakan untuk mengklasifikasikan sentimen ulasan pengguna Shopee terhadap fitur COD. Dataset dibagi menjadi data latih dan data uji untuk mengukur performa model. Evaluasi dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score.

Tabel 1 Menunjukkan Hasil Evaluasi Model

Kategori Sentimen	Precision	Recall	F1-Score
Positif	0,91	0,88	0,89
Negatif	0,87	0,90	0,88
Netral	0,84	0,86	0,85
<b>Rata-rata</b>	<b>0,87</b>	<b>0,88</b>	<b>0,87</b>

Akurasi model: 88%

Interpretasi Hasil :

- Precision yang tinggi pada sentimen positif menunjukkan bahwa prediksi positif jarang salah.
- Recall tertinggi ada pada sentimen negatif, yang berarti model mampu menangkap sebagian besar ulasan negatif dengan baik.
- Nilai F1-score yang seimbang pada ketiga kategori menandakan performa model cukup stabil di semua kelas.
- Akurasi keseluruhan 88% menunjukkan bahwa Naive Bayes cukup efektif digunakan untuk analisis sentimen pada dataset ini.

4. PEMBAHASAN

Berdasarkan hasil evaluasi model, Naive Bayes menunjukkan performa yang baik dalam mengklasifikasikan sentimen ulasan pengguna Shopee terhadap fitur COD. Precision yang tinggi pada sentimen positif (0,91) menunjukkan bahwa prediksi positif jarang salah. Recall tertinggi ada pada sentimen negatif (0,90), yang berarti model mampu menangkap sebagian besar ulasan negatif dengan baik. Nilai F1-score yang seimbang pada ketiga kategori menandakan performa model cukup stabil di semua kelas.

Akurasi keseluruhan sebesar 88% menunjukkan bahwa Naive Bayes cukup efektif digunakan untuk analisis sentimen pada dataset ini. Hasil distribusi sentimen menunjukkan dominasi sentimen negatif (63,96%), diikuti sentimen positif (35,77%), dan sentimen netral (0,27%). Sentimen negatif yang dominan mengindikasikan adanya keluhan pengguna terkait fitur COD, yang umumnya berkaitan dengan permasalahan pengiriman, pembatalan pesanan, dan penolakan barang oleh konsumen.

Visualisasi menggunakan WordCloud memperlihatkan kata-kata yang paling sering muncul dalam ulasan pengguna, memberikan gambaran awal mengenai pola kata yang digunakan terkait fitur COD. Sementara itu, distribusi sentimen berdasarkan tahun menunjukkan bahwa pada tahun-tahun tertentu terjadi peningkatan ulasan negatif, yang dapat dikaitkan dengan momen promo besar atau lonjakan pesanan COD, sedangkan sentimen positif cenderung stabil di setiap tahun.

Visualisasi data menggunakan Power BI memberikan gambaran yang jelas dan informatif mengenai distribusi sentimen, jumlah ulasan per kategori, dan tren sentimen per tahun. Hal ini memudahkan interpretasi hasil analisis dan memberikan kontribusi penting bagi pengembangan strategi layanan e-commerce, khususnya dalam mempertahankan keunggulan fitur COD sekaligus meminimalkan keluhan pelanggan.

## 5. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa sentimen pengguna Shopee terhadap fitur *Cash on Delivery* (COD) didominasi oleh sentimen positif, diikuti sentimen negatif, dan netral. Sentimen positif menunjukkan apresiasi terhadap kemudahan dan fleksibilitas pembayaran COD, sedangkan sentimen negatif umumnya terkait permasalahan pengiriman, pembatalan pesanan, dan penolakan barang.

Model Naive Bayes yang digunakan dalam penelitian ini menunjukkan performa yang baik dengan akurasi sebesar 88% serta nilai precision, recall, dan F1-score yang seimbang di seluruh kategori sentimen. Hasil ini menunjukkan bahwa metode Naive Bayes efektif dalam mengklasifikasikan ulasan pengguna. Visualisasi data menggunakan Power BI membantu memperjelas distribusi sentimen dan tren ulasan per tahun, sehingga memberikan gambaran yang komprehensif bagi pihak terkait dalam mengevaluasi dan meningkatkan layanan COD di Shopee. Dengan demikian, hasil penelitian ini memberikan kontribusi penting bagi pengembangan strategi layanan *e-commerce*, khususnya dalam mempertahankan keunggulan fitur COD sekaligus meminimalkan keluhan pelanggan.

Berdasarkan hasil penelitian yang telah dilakukan mengenai analisis sentimen pengguna Shopee terhadap fitur *Cash on Delivery* (COD) menggunakan metode Naive Bayes dan visualisasi data dengan Power BI, dapat disimpulkan bahwa:

1. Sentimen pengguna terhadap fitur COD didominasi oleh sentimen positif, menunjukkan bahwa mayoritas pengguna merasa terbantu dengan adanya fitur ini, khususnya dalam memberikan kemudahan pembayaran dan keamanan transaksi.
2. Sentimen negatif yang muncul sebagian besar berkaitan dengan permasalahan pengiriman, pembatalan pesanan, dan penolakan barang. Persentase sentimen negatif cenderung meningkat pada periode tertentu yang bertepatan dengan event promo besar.
3. Model Naive Bayes yang digunakan mampu mencapai akurasi sebesar 88% dengan nilai precision, recall, dan F1-score yang relatif seimbang di setiap kategori sentimen, menunjukkan kinerja yang baik dalam mengklasifikasikan ulasan.
4. Visualisasi data menggunakan Power BI memberikan gambaran yang jelas dan informatif mengenai distribusi sentimen, jumlah ulasan per kategori, dan tren sentimen per tahun, sehingga memudahkan interpretasi hasil analisis.

## DAFTAR PUSTAKA

- Welding Menggunakan Microsoft Power BI pada PT ABC. *Journal of Manufacturing and Enterprise Information System*, 1(1), 21-31. <https://doi.org/10.52330/jmeis.v1i1.86>
- [4] APJII. (2024). Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang. Diakses dari: <https://apjii.or.id/>
- [5] Asih, E. M. (2024). Analisis pada Shopee sebagai E-Commerce Terpopuler di Indonesia. *Jurnal Ekonomi Bisnis Antartika*, 2.
- [6] Minfo Polgan, J., et al. (n.d.). Analisis Sentimen Pelanggan Toko Online Tokopedia Menggunakan Metode Naive Bayes Classifier.
- [7] Mulyani, E., & Martantoh, E. (n.d.). Analisis Sentimen pada Ulasan Aplikasi Airbnb di Google Play Store Menggunakan Metode Naive Bayes Classification. *Inovasi Pendidikan Nusantara*.
- [8] Putro, H. F., Vlandari, R. T., & Saptomo, W. L. Y. (2020). Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan. *Jurnal Teknologi Informasi Dan Komunikasi (TIKomSiN)*, 8(2). <https://doi.org/10.30646/tikomsin.v8i2.500>
- [9] Reka, R. (2025). Memahami COD (Cash on Delivery) & COD Cek Dulu. Diakses dari: <https://shopee.co.id/inspirasi-shopee/cod-cash-on-delivery/>
- [10] Wibisono, A. C., Nadira, T. S., & Sutabri, T. (2025). Analisis Sentimen Pelanggan pada Platform Shopee Menggunakan Metode Naive Bayes. *Nusantara Journal of Multidisciplinary Science*, 2(6). <https://jurnal.intekom.id/index.php/njms>
- [11] Yumni, S. Z., & Widowati, W. (2021). Implementasi Microsoft Power BI Dalam Memantau Kehadiran dan Transportasi Pegawai. *Jurnal Sains Dan Edukasi Sains*, 4(1), 1-8. <https://doi.org/10.24246/juses.v4i1p1-8>
- [1] Kaggle. (2024). Shopee COD Reviews Dataset. Diakses dari: <https://www.kaggle.com/>
- [2] Nurhidayati, D., & Susanti, S. (2021). Analisis Sentimen pada Ulasan Pengguna E-commerce Menggunakan Naive Bayes Classifier. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 8(4), 671-678. Shopee Indonesia. (2025). Informasi Layanan COD Shopee. Diakses dari: <https://shopee.co.id/>
- [3] Afriadi, H. M. T., & Hadianastuti, F. L. (2023). Rancang Bangun Sistem Informasi Dashboard Pengendalian Produksi