
Penerapan Metode Klasifikasi Support Vector Machine dalam Analisis Sentiment Ulasan Pengguna Layanan Goride Pada Aplikasi Gojek di Google Playstore

M. Adreansyah Pratama Lubis^{1*)}, Ria Eka Sari²⁾

1,2)Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer,
Universitas Potensi Utama

*Corresponding Email: adreansyahlubis@gmail.com

Abstrak

Ulasan pengguna terhadap layanan Goride di Google Playstore merupakan sumber data krusial untuk evaluasi kualitas layanan, namun analisis manual terhadap volume data besar tidak efisien. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem klasifikasi sentimen otomatis menggunakan metode Support Vector Machine (SVM) untuk mengkategorikan ulasan pengguna Goride menjadi sentimen positif atau negatif. Sebanyak 659 ulasan dari periode Januari-Desember 2024 dianalisis melalui serangkaian tahapan pra-pemrosesan data, pelabelan data, dan ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Model SVM dioptimalkan menggunakan GridSearchCV mencapai performa puncak dengan akurasi 87,88%, presisi 66,67%, recall 85,71%, dan F1-Score 75,00% pada konfigurasi terbaik ($C=0.1$, kernel linear). Hasil ini mengonfirmasi bahwa SVM merupakan metode yang efektif dan tangguh untuk analisis sentimen pada ulasan layanan transportasi daring, serta menyediakan alat praktis yang berpotensi mendukung pengambilan keputusan strategis berbasis data untuk peningkatan layanan Gojek.

Kata Kunci: Analisis Sentimen, Support Vector Machine (SVM), TF-IDF, Goride, Streamlit

Abstract

User reviews of the Goride service on the Google Play Store are a crucial data source for evaluating service quality; however, manual analysis of large data volumes is inefficient. This study aims to design and implement an automated sentiment classification system using the Support Vector Machine (SVM) method to categorize Goride user reviews into positive or negative sentiments. A total of 659 reviews from the period January–December 2024 were analyzed through a series of preprocessing stages, data labeling, and feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF). The SVM model, optimized using GridSearchCV, achieved peak performance with an accuracy of 87.88%, precision of 66.67%, recall of 85.71%, and an F1-score of 75.00% under the best configuration ($C=0.1$, linear kernel). These findings confirm that SVM is an effective and robust method for sentiment analysis of online transportation service reviews and provides a practical tool with the potential to support data-driven strategic decision-making for improving Gojek's services.

Keywords: Sentiment Analysis, Support Vector Machine (SVM), TF-IDF, GoRide, Streamlit

PENDAHULUAN

Perkembangan ekonomi digital telah mendorong transformasi layanan transportasi konvensional menjadi ekosistem super app yang mengintegrasikan beragam layanan untuk memenuhi kebutuhan masyarakat. Gojek sebagai salah satu pelopor dalam sektor ini telah membentuk platform terintegrasi yang menyediakan layanan mobilitas seperti GoRide, GoCar, GoSend, dan GoFood untuk menjawab dinamika kebutuhan masyarakat modern (Harimurti & Sofyan, 2022). Berdasarkan survei Institute for Development of Economics and Finance pada tahun 2022, Gojek mendominasi pasar transportasi daring di Indonesia dengan tingkat penggunaan mencapai 82% (Katadata.co.id, 2022). Dalam lanskap yang sangat kompetitif ini, kualitas layanan menjadi faktor diferensiasi krusial untuk mempertahankan loyalitas pengguna. Ulasan pengguna pada platform distribusi aplikasi seperti Google Play Store merupakan sumber data otentik untuk mengevaluasi kepuasan pelanggan (Adhi Putra, 2021). Namun volume data tekstual yang besar dan tidak terstruktur menjadi tantangan utama dalam ekstraksi wawasan yang bermakna sehingga analisis manual menjadi tidak efisien dan rentan terhadap subjektivitas (Rokhman et al., 2021).

Analisis sentimen merupakan cabang ilmu dalam natural language processing yang melakukan proses komputasional untuk mengidentifikasi dan mengkategorikan opini pengguna terhadap suatu entitas ke dalam kategori positif, negatif, atau netral (Mehta & Pandya, 2020). Metode analisis sentimen menggunakan algoritma machine learning telah menjadi pendekatan yang mapan untuk mengklasifikasikan opini dari data tekstual tidak terstruktur. Support Vector Machine (SVM) adalah salah satu metode klasifikasi supervised learning yang bekerja dengan menemukan hyperplane optimal untuk memisahkan data ke dalam kelas-kelas tertentu dengan margin maksimum. Berbagai penelitian telah menunjukkan efektivitas metode SVM dalam domain analisis sentimen. Studi oleh (Alamsyah et al., 2024) mencatat akurasi SVM sebesar 87,24% dalam klasifikasi ulasan aplikasi finansial sementara (Idris et al., 2023) mencapai akurasi 89% pada

ulasan platform e-commerce. Penelitian (Utami et al., 2023) juga membuktikan keunggulan SVM dalam analisis sentimen berbasis aspek pada ulasan aplikasi dengan hasil yang memuaskan. Studi komparatif secara konsisten menempatkan SVM sebagai model dengan performa unggul seperti temuan (Iskandar & Nataliani, 2021) yang menunjukkan akurasi rata-rata SVM mencapai 96,43% dibandingkan Naïve Bayes dan k-NN.

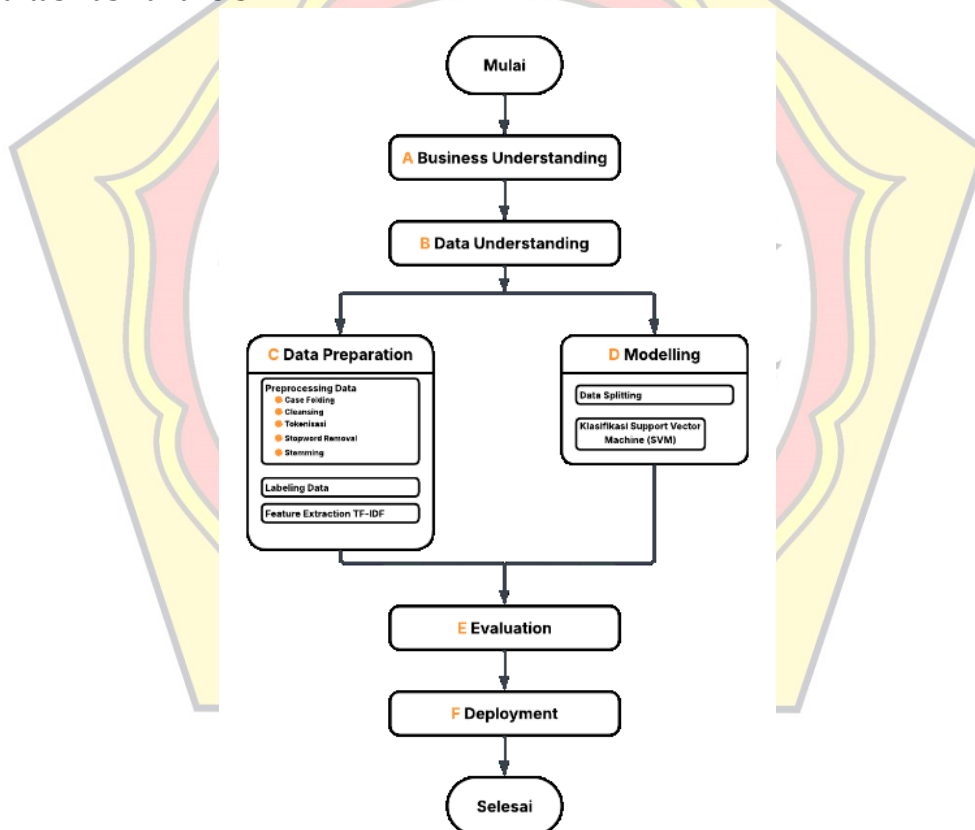
Dalam konteks layanan transportasi daring dan media sosial, penerapan SVM juga menunjukkan hasil yang memuaskan. (Dwianto & Sadikin, 2021) menerapkan SVM untuk analisis sentimen transportasi online pada Twitter dan menemukan bahwa SVM memberikan performa klasifikasi yang lebih stabil. (Pranata et al., 2022) menggunakan algoritma SVM untuk klasifikasi sentimen terhadap layanan transportasi online dengan hasil yang menjanjikan. (Rizki et al., 2021) juga membuktikan efektivitas SVM dalam analisis sentimen pada media sosial. Untuk mengoptimalkan proses klasifikasi, teknik ekstraksi fitur seperti Term Frequency-Inverse Document Frequency (TF-IDF) telah terbukti efektif dalam menentukan bobot kata kunci dari data ulasan untuk keperluan analisis lanjutan (Joergensen E Munthe et al., 2022). Meskipun berbagai penelitian telah mengaplikasikan analisis sentimen pada platform media sosial dan e-commerce secara umum, kajian yang secara spesifik menargetkan layanan inti dalam ekosistem super app seperti GoRide pada aplikasi Gojek masih terbatas. Selain itu sebagian besar penelitian berhenti pada pelaporan metrik akurasi model tanpa mengembangkan sistem aplikatif yang dapat memvisualisasikan hasil analisis untuk mendukung pengambilan keputusan strategis oleh penyedia layanan.

Penelitian ini bertujuan merancang dan mengevaluasi sistem analisis sentimen terhadap ulasan pengguna layanan GoRide di Google Play Store menggunakan metode klasifikasi Support Vector Machine untuk mengkategorikan ulasan ke dalam sentimen positif atau negatif. Kontribusi utama penelitian ini terletak pada fokus analisis terhadap domain spesifik layanan GoRide sebagai studi kasus yang relevan dengan pasar transportasi daring di Indonesia serta pengembangan purwarupa sistem visualisasi data interaktif menggunakan

framework Streamlit. Sistem ini dirancang untuk menyajikan hasil analisis secara informatif sehingga dapat memberikan landasan data yang kuat bagi penyedia layanan dalam upaya peningkatan kualitas layanan secara berkelanjutan.

METODE PENELITIAN

Penelitian ini menerapkan metodologi Cross Industry Standard Process for Data Mining (CRISP-DM) yang merupakan framework standar untuk proyek data mining dan telah banyak digunakan dalam penelitian analisis sentimen berbasis machine learning (Hakim et al., 2023). Pendekatan sistematis ini mencakup enam tahapan utama yang dikembangkan lebih lanjut untuk memastikan kualitas dan validitas hasil analisis.



Gambar 1. Desain Penelitian

Penelitian ini dilaksanakan melalui 6 tahapan yang merujuk pada Gambar 1. Tahapan diawali dengan Business Understanding untuk memahami kebutuhan bisnis dan mendefinisikan masalah analisis sentimen layanan GoRide, dilanjutkan dengan Data Understanding untuk mengeksplorasi karakteristik data ulasan

pengguna dari Google Play Store menggunakan teknik web scraping (Adhi Putra, 2021). Tahap Data Preparation mencakup preprocessing data (casefolding, tokenizing, stopwords removal, stemming), data labeling, dan feature extraction menggunakan TF-IDF untuk mengubah teks menjadi representasi numerik (Joergensen E Munthe et al., 2022). Tahap Modelling menerapkan algoritma klasifikasi Support Vector Machine (SVM) dengan hyperparameter tuning untuk menentukan kombinasi parameter terbaik seperti kernel, C, dan gamma, dimana SVM bekerja dengan menemukan hyperplane optimal yang memisahkan dua kelas sentimen menggunakan kernel trick untuk menangani data non-linier (Iskandar & Nataliani, 2021). Tahap Evaluation mengukur performa model menggunakan metrik confusion matrix meliputi accuracy, precision, recall, dan F1-score. Tahap terakhir adalah Deployment dengan mengimplementasikan model dalam bentuk aplikasi visualisasi berbasis Streamlit untuk memudahkan interpretasi hasil klasifikasi sentimen. Gambar 2. menyajikan detail dari alur proses penerapan klasifikasi Support Vector Machine (SVM).



Gambar 2. Klasifikasi Support Vector Machine

HASIL DAN PEMBAHASAN

A. Hasil Implementasi Alur Kerja CRISP-DM

1. Hasil Business Understanding

Peningkatan jumlah pengguna layanan GoRide diikuti pertumbuhan signifikan ulasan yang menjadi sumber informasi penting untuk mengevaluasi layanan dan memahami kepuasan pelanggan. Namun, volume dan kompleksitas data membuat analisis manual tidak efisien. Pendekatan analisis sentimen menggunakan algoritma Support Vector Machine (SVM) diterapkan karena kemampuannya dalam klasifikasi data secara presisi dan telah terbukti efektif dalam berbagai domain aplikasi digital (Alamsyah et al., 2024; Utami et al., 2023).

Pengelompokan ulasan menjadi sentimen positif dan negatif memberikan wawasan relevan untuk pengambilan keputusan strategis. Model diimplementasikan ke dalam sistem berbasis Streamlit untuk mempermudah visualisasi dan pemahaman opini publik mengenai layanan GoRide.

2. Hasil Data Understanding

Proses pemahaman data diawali dengan pengambilan data melalui teknik crawling menggunakan library google-play-scraper pada Python yang difokuskan pada ulasan aplikasi Gojek dengan penyaringan kata kunci "goride" atau "GORIDE". Teknik crawling data digunakan untuk mengekstraksi data ulasan aplikasi mobile merupakan pendekatan standar dalam penelitian analisis sentimen berbasis Google Play Store (Adhi Putra, 2021; Idris et al., 2023). Berikut kode program crawling data yang dikembangkan:

```

1 import pandas as pd
2 from google_play_scraper import Sort, reviews
3 from datetime import datetime
4 import time
5
6 # Bagian 1: Konfigurasi dan Fungsi Inti untuk Scraping
7 CONFIG = {
8     "TARGET_YEAR": 2025,
9     "TARGET_MONTH": 4,
10    "BATCH_SIZE": 200,
11    "LANG": 'id',
12    "COUNTRY": 'id',
13    "APP_ID": 'com.gojek.app',
14    "KEYWORDS": ["Goride", "goride", "go-ride", "Go-ride", "GO RIDE", "go ride", "ride", "Ride"]
15 }
16
17 def scrape_and_filter_reviews(config):
18     """
19     Mengambil dan memfilter ulasan berdasarkan kata kunci, bulan, dan tahun target.
20     Fungsi ini menjalankan loop untuk mengambil data secara batch dan
21     menghentikannya saat ulasan sudah berada di luar periode target.
22     """
23     year, month, keywords = config["TARGET_YEAR"], config["TARGET_MONTH"], config["KEYWORDS"]
24     start_date = datetime(year, month, 1)
25     end_date = datetime(year + 1, 1, 1) if month == 12 else datetime(year, month + 1, 1)
26
27     all_reviews = []
28     continuation_token = None
29     stop_early = False
30
31     while not stop_early:
32         try:
33             result, continuation_token = reviews(
34                 config["APP_ID"],
35                 lang=config["LANG"],
36                 country=config["COUNTRY"],
37                 sort=Sort.NEWEST,
38                 count=config["BATCH_SIZE"],
39                 continuation_token=continuation_token
40             )
41             if not result or not continuation_token:
42                 break
43
44             for review in result:
45                 review_date = review['at']
46                 review_text = review.get('content', '')
47                 if review_date < start_date:
48                     stop_early = True
49                     break
50                 if (review_text and start_date <= review_date < end_date and
51                     any(keyword in review_text for keyword in keywords)):
52                     all_reviews.append({
53                         'review_id': review['reviewId'],
54                         'user_name': review['userName'],
55                         'rating': review['score'],
56                         'review_text': review_text,
57                         'date': review_date.strftime('%Y-%m-%d %H:%M:%S'),
58                     })
59         except Exception as e:
60             print(f"Error encountered: {e}")
61             break
62
63     return pd.DataFrame(all_reviews)

```

Gambar 3. Proses Crawling dan Filtering Data Ulasan Goride

```

2 def main():
3     month_name = datetime(CONFIG["TARGET_YEAR"], CONFIG["TARGET_MONTH"], 1).strftime("%B")
4     print(f"Memulai scraping ulasan GoRide untuk {month_name} {CONFIG["TARGET_YEAR"]}...")
5
6     start_time = time.time()
7     reviews_df = scrape_and_filter_reviews(CONFIG)
8     elapsed_time = time.time() - start_time
9
10    if reviews_df.empty:
11        print("Tidak ada ulasan GoRide yang ditemukan untuk periode ini.")
12        return
13
14    csv_filename = f"%ulasan_goride_{CONFIG["TARGET_YEAR"]}_%{CONFIG["TARGET_MONTH"]}%2d.csv"
15    reviews_df.to_csv(csv_filename, index=False, encoding="utf-8")
16
17    print("\n--- Proses Selesai ---")
18    print(f"Total waktu eksekusi: {elapsed_time:.2f} detik")
19    print(f"Total ulasan ditemukan: {len(reviews_df)}")
20    print(f>Data berhasil disimpan ke file: {csv_filename}")
21    print("\nPreview 5 data terakhir:")
22    print(reviews_df.head())
23
24 if __name__ == "__main__":
25     main()

```

Gambar 4. Proses Penarikan Data Ulasan Goride

Implementasi program scraping berhasil mengekstraksi ulasan relevan periode Januari hingga Desember 2024 yang disimpan dalam format CSV. Dataset memiliki struktur data dengan atribut meliputi username, review, dan rating. Hasil crawling data

Tabel 1. Dataset Ulasan Pengguna Layanan Goride

No	Username	Ulasan	Rating
1	Ayun Hawa	Aneh, voucher goride gocar gamuncul semua padahal saya punya banyak voucher yang muncul cuma voucher gosend diperbaiki deh tolong	1
2	Supri Yanto	Kok buka aplikasi goride hilang2 mulu...	5
3	Ruwy Masyari	Gojek keren bisa pesen sekalian 2 goride satu waktu makasih gojek 😊👍👍	5
4	Achmad Sayid	Sekarang udah mahal goride	3
5	Resa Agustin	Kesalahan dari sistem gojek saldo Gopay kepotong otomatis ketika melakukan transaksi order GoRide GoRide nya BATAL tp saldo Gopay kepotong dan gak balik lagi. Udh bikin laporan dan disuruh tunggu 2x24 jam tp saldo ga masuk lagi ke akun gue. Hadeuuu rumit	1
....			
....			
659	Wardah Muthia	Saya beli paket goplus tapi voucher go ride tidak bisa di pakai	1

dataset dapat dilihat pada Tabel 1.

3. Hasil Data Preparation

Preprocessing data merupakan tahapan krusial untuk membersihkan dan menormalisasi teks agar dapat diproses secara optimal oleh model machine learning (Haviana & Poetro, 2022). Tahapan ini meliputi lima langkah berurutan yang dijelaskan sebagai berikut.

3.1. Preprocessing Data

3.1.1. Case Folding

Pada tahap preprocessing data, case folding dilakukan untuk mengubah seluruh huruf dalam teks menjadi huruf kecil. Tujuan dari langkah ini adalah untuk menyeragamkan penulisan kata agar tidak terjadi perbedaan makna dari variasi huruf dan menjaga konsistensi data (Dwianto & Sadikin, 2021). Berikut proses dan hasil case folding dalam preprocessing data:

Tabel 2. Preprocessing Data - Case Folding

No	Ulasan	Setelah Case Folding
1	Aneh, voucher goride gocar gamuncul semua padahal saya punya banyak voucher yang muncul cuma voucher gosend diperbaiki deh tolong	aneh, voucher goride gocar gamuncul semua padahal saya punya banyak voucher yang muncul cuma voucher gosend diperbaiki deh tolong
2	Kok buka aplikasi goride hilang2 mulu...	kok buka aplikasi goride hilang2 mulu...
3	Gojek keren bisa pesen sekalian 2 goride satu waktu makasih gojek 😊👍👍	gojek keren bisa pesen sekalian 2 goride satu waktu makasih gojek 😊👍👍
4	Sekarang udah mahal goride	sekarang udah mahal goride
5	Kesalahan dari sistem gojek saldo Gopay kepotong otomatis ketika melakukan transaksi order GoRide GoRide nya BATAL tp saldo Gopay kepotong dan gak balik lagi. Udh bikin laporan dan disuruh tunggu 2x24 jam tp saldo ga masuk lagi ke akun gue. Hadeuuu rumit	kesalahan dari sistem gojek saldo gopay kepotong otomatis ketika melakukan transaksi order goride goride nya batal tp saldo gopay kepotong dan gak balik lagi. udh bikin laporan dan disuruh tunggu 2x24 jam tp saldo ga masuk lagi ke akun gue. hadeuuu rumit

3.1.2. Cleansing

Tahap tokenizing memisahkan teks menjadi token berupa kata atau frasa untuk mengubah data teks menjadi format terstruktur yang dapat diproses secara komputasional (Pranata et al., 2022). Hasil cleansing data dapat dilihat pada Tabel

3.

Tabel 3. Preprocessing Data - Cleansing

No	Sebelum Cleansing	Setelah Cleansing
1	Aneh, voucher goride gocar gamuncul semua padahal saya punya banyak voucher yang muncul cuma voucher gosend diperbaiki deh tolong	aneh voucher goride gocar tidak muncul semua padahal saya punya banyak voucher yang muncul hanya voucher gosend diperbaiki deh tolong
2	Kok buka aplikasi goride hilang2 mulu...	kenapa buka aplikasi goride hilang melulu
3	Gojek keren bisa pesen sekalian 2 goride satu waktu makasih gojek 😊👍👍	gojek keren bisa pesan sekalian goride satu waktu terima kasih gojek
4	Sekarang udah mahal goride	sekarang sudah mahal goride
5	Kesalahan dari sistem gojek saldo Gopay kepotong otomatis ketika melakukan	kesalahan dari sistem gojek saldo gopay kepotong otomatis ketika melakukan

transaksi order GoRide GoRide nya BATAL tp saldo Gopay kepotong dan gak balik lagi. Udh bikin laporan dan disuruh tuunggu 2x24 jam tp saldo ga masuk lagi ke akun gue. Hadeuuu rumit	transaksi order goride goride nya batal tapi saldo gopay kepotong dan tidak balik lagi sudah bikin laporan dan disuruh tuunggu x jam tapi saldo tidak masuk lagi ke akun saya aduh rumit
--	--

3.1.3. Tokenisasi

Tahap tokenizing memisahkan teks menjadi token berupa kata atau frasa untuk mengubah data teks menjadi format terstruktur yang dapat diproses secara komputasional (Pranata et al., 2022). Hasil tokenisasi dapat dilihat pada Tabel 4.

Tabel 4. Preprocessing Data - Tokenisasi

No	Sebelum Tokenisasi	Setelah Tokenisasi
1	aneh voucher goride gocar tidak muncul semua padahal saya punya banyak voucher yang muncul hanya voucher gosend diperbaiki deh tolong	['aneh', 'voucher', 'goride', 'gocar', 'tidak', 'muncul', 'semua', 'padahal', 'saya', 'punya', 'banyak', 'voucher', 'yang', 'muncul', 'hanya', 'voucher', 'gosend', 'diperbaiki', 'deh', 'tolong']
2	kenapa buka aplikasi goride hilang melulu	['kenapa', 'buka', 'aplikasi', 'goride', 'hilang', 'melulu']
3	gojek keren bisa pesan sekalian goride satu waktu terima kasih gojek	['gojek', 'keren', 'bisa', 'pesan', 'sekalian', 'goride', 'satu', 'waktu', 'terima', 'kasih', 'gojek']
4	sekarang sudah mahal goride	['sekarang', 'sudah', 'mahal', 'goride']
5	kesalahan dari sistem gojek saldo gopay kepotong otomatis ketika melakukan transaksi order goride goride nya batal tapi saldo gopay kepotong dan tidak balik lagi sudah bikin laporan dan disuruh tuunggu x jam tapi saldo tidak masuk lagi ke akun saya aduh rumit	['kesalahan', 'dari', 'sistem', 'gojek', 'saldo', 'gopay', 'kepotong', 'otomatis', 'ketika', 'melakukan', 'transaksi', 'order', 'goride', 'goride', 'nya', 'batal', 'tapi', 'saldo', 'gopay', 'kepotong', 'dan', 'tidak', 'balik', 'lagi', 'sudah', 'bikin', 'laporan', 'dan', 'disuruh', 'tuunggu', 'x', 'jam', 'tapi', 'saldo', 'tidak', 'masuk', 'lagi', 'ke', 'akun', 'saya', 'aduh', 'rumit']

3.1.4. Stopword Removal

Stopword removal dalam preprocessing data yaitu penghapusan kata-kata umum yang memiliki frekuensi tinggi namun tidak memberikan makna penting dalam analisis seperti kata hubung atau kata depan (Rokhman et al., 2021). Dapat dilihat pada Tabel 5. hasil stopwords removal.

Tabel 5. Preprocessing Data - Stopword Removal

No	Sebelum Stopword Removal	Setelah Stopword Removal
1	['aneh', 'voucher', 'goride', 'gocar', 'tidak', 'muncul', 'semua', 'padahal', 'saya', 'punya', 'banyak', 'voucher', 'yang', 'muncul', 'hanya', 'voucher', 'gosend', 'diperbaiki', 'deh', 'tolong']	['aneh', 'voucher', 'goride', 'gocar', 'tidak', 'muncul', 'voucher', 'voucher', 'muncul', 'voucher', 'gosend', 'diperbaiki', 'tolong']

2	['kenapa', 'buka', 'aplikasi', 'goride', 'hilang', 'melulu']	['buka', 'aplikasi', 'goride', 'hilang', 'melulu']
3	['gojek', 'keren', 'bisa', 'pesan', 'sekalian', 'goride', 'satu', 'waktu', 'terima', 'kasih', 'gojek']	['gojek', 'keren', 'bisa', 'pesan', 'goride', 'terima', 'kasih', 'gojek']
4	['sekarang', 'sudah', 'mahal', 'goride']	['mahal', 'goride']
5	['kesalahan', 'dari', 'sistem', 'gojek', 'saldo', 'gopay', 'kepotong', 'otomatis', 'ketika', 'melakukan', 'transaksi', 'order', 'goride', 'goride', 'nya', 'batal', 'tapi', 'saldo', 'gopay', 'kepotong', 'dan', 'tidak', 'balik', 'lagi', 'sudah', 'bikin', 'laporan', 'dan', 'disuruh', 'tuunggu', 'x', 'jam', 'tapi', 'saldo', 'tidak', 'masuk', 'lagi', 'ke', 'akun', 'saya', 'aduh', 'rumit']	['kesalahan', 'sistem', 'gojek', 'saldo', 'gopay', 'kepotong', 'otomatis', 'transaksi', 'order', 'goride', 'goride', 'nya', 'batal', 'saldo', 'gopay', 'kepotong', 'tidak', 'bikin', 'laporan', 'disuruh', 'tuunggu', 'x', 'jam', 'saldo', 'tidak', 'masuk', 'akun', 'aduh', 'rumit']

3.1.5. Stemming

Tahap dari preprocessing data yaitu stemming mengubah kata menjadi bentuk dasarnya dengan menghilangkan imbuhan sehingga sistem dapat mengenali kata yang memiliki makna dasar sama untuk meningkatkan efisiensi pemodelan (Iskandar & Nataliani, 2021). Hasil stemming dapat dilihat pada Tabel 6.

Tabel 6. Preprocessing Data - Stemming

No	Sebelum Stemming	Setelah Stemming
1	['aneh', 'voucher', 'goride', 'gocar', 'tidak', 'muncul', 'voucher', 'muncul', 'voucher', 'gosend', 'diperbaiki', 'tolong']	['aneh', 'voucher', 'goride', 'gocar', 'tidak', 'muncul', 'voucher', 'muncul', 'voucher', 'gosend', 'baik', 'tolong']
2	['buka', 'aplikasi', 'goride', 'hilang', 'melulu']	['buka', 'aplikasi', 'goride', 'hilang', 'melulu']
3	['gojek', 'keren', 'bisa', 'pesan', 'goride', 'terima', 'kasih', 'gojek']	['gojek', 'keren', 'bisa', 'pesan', 'goride', 'terima', 'kasih', 'gojek']
4	['mahal', 'goride']	['mahal', 'goride']
5	['kesalahan', 'sistem', 'gojek', 'saldo', 'gopay', 'kepotong', 'otomatis', 'transaksi', 'order', 'goride', 'goride', 'nya', 'batal', 'saldo', 'gopay', 'kepotong', 'tidak', 'bikin', 'laporan', 'disuruh', 'tuunggu', 'x', 'jam', 'saldo', 'tidak', 'masuk', 'akun', 'aduh', 'rumit']	['salah', 'sistem', 'gojek', 'saldo', 'gopay', 'potong', 'otomatis', 'transaksi', 'order', 'goride', 'goride', 'nya', 'batal', 'saldo', 'gopay', 'potong', 'tidak', 'bikin', 'lapor', 'suruh', 'tuunggu', 'x', 'jam', 'saldo', 'tidak', 'masuk', 'akun', 'aduh', 'rumit']

Berdasarkan hasil preprocessing yang telah dilakukan pada Data Preparation diperoleh dataset yang telah melalui berbagai tahap preprocessing data sehingga dapat digunakan untuk analisis lebih lanjut. Tampilan data sebelum dan sesudah melalui proses preprocessing dapat dilihat pada Tabel 7.

Tabel 7. Hasil Dataset setelah Preprocessing Data

No	Sebelum Preprocessing Data	Setelah Preprocessing Data
1	Aneh, voucher goride gocar gamuncul semua padahal saya punya banyak voucher yang muncul cuma voucher gosend diperbaiki deh tolong	aneh voucher goride tidak muncul voucher muncul voucher baik tolong
2	Kok buka aplikasi goride hilang2 mulu...	buka aplikasi goride hilang melulu
3	Gojek keren bisa pesen sekalian 2 goride satu waktu makasih gojek 🤖👍👍	gojek keren bisa pesan goride terima kasih gojek
4	Sekarang udah mahal goride dan gofood	mahal goride
5	Kesalahan dari sistem gojek saldo Gopay kepotong otomatis ketika melakukan transaksi order GoRide GoRide nya BATAL tp saldo Gopay kepotong dan gak balik lagi. Udh bikin laporan dan disuruh tuunggu 2x24 jam tp saldo ga masuk lagi ke akun gue. Hadeuuu rumit	salah sistem gojek saldo gopay potong otomatis transaksi order goride goride nya batal saldo gopay potong tidak bikin lapor suruh tuunggu x jam saldo tidak masuk akun aduh rumit

3.1.6. Labeling Data

Pemberian label pada dataset dilakukan dengan pendekatan semi-otomatis menggunakan model pre-trained IndoBERT (mdhugol/indonesia-bert-sentiment-classification) untuk mengidentifikasi kategori sentimen pada setiap ulasan sebagai langkah awal sebelum pelatihan model SVM. Pendekatan transfer learning menggunakan model pre-trained telah terbukti efektif untuk tugas klasifikasi sentimen dengan akurasi tinggi (Haviana & Poetro, 2022). Proses labeling dilakukan secara batch dengan ukuran 100 ulasan per iterasi untuk menghindari error pada proses komputasi. Gambar 5 menunjukkan implementasi kode program untuk pseudo-labeling otomatis menggunakan pipeline sentiment analysis.

```

1 from transformers import pipeline, AutoTokenizer, AutoModelForSequenceClassification
2
3 # STEP 2: Load IndoBERT model dan tokenizer
4 model_name = "mdhugol/indonesia-bert-sentiment-classification"
5 tokenizer = AutoTokenizer.from_pretrained(model_name)
6 model = AutoModelForSequenceClassification.from_pretrained(model_name)
7
8 # STEP 3: Load dataset CSV
9 df = pd.read_csv("goride_review.csv")
10 print(f"Jumlah ulasan: {len(df)}")
11 df.head()
12
13 # STEP 4: Buat pipeline analisis sentimen
14 sentiment_pipeline = pipeline("sentiment-analysis", model=model, tokenizer=tokenizer)
15
16 # STEP 5: Lakukan pseudo-labeling (berlabel otomatis)
17 # Proses batch untuk menghindari error terlalu banyak permintaan
18 batch_size = 100
19 labels = []
20
21 for i in range(0, len(df), batch_size):
22     batch = df["review_text"].iloc[i:i+batch_size].tolist()
23     preds = sentiment_pipeline(batch)
24     labels.extend([pred["label"] for pred in preds])
25
26 df["pseudo_label"] = labels

```

Gambar 5. Pelabelan Dataset Dengan Model Pre-trained BERT

Untuk memastikan relevansi hasil pelabelan, verifikasi manual dilakukan untuk memastikan akurasi pelabelan berdasarkan konteks ulasan. Hasil pelabelan menghasilkan dua kategori sentimen (positif dan negatif) seperti ditunjukkan pada Tabel 8.

Tabel 8. Hasil Pelabelan Dataset

No	Ulasan	Label
1	aneh voucher goride tidak muncul voucher muncul voucher baik tolong	Negatif
2	buka aplikasi goride hilang melulu	Negatif
3	gojek keren bisa pesan goride terima kasih gojek	Positif
4	mahal goride	Negatif
5	salah sistem gojek saldo gopay potong otomatis transaksi order goride goride nya batal saldo gopay potong tidak bikin lapor suruh tunggu x jam saldo tidak masuk akun aduh rumit	Negatif

3.2. Feature Extraction

Feature extraction mengkonversi ulasan hasil preprocessing menjadi bentuk numerik menggunakan teknik Term Frequency-Inverse Document Frequency (TF-IDF) untuk menghasilkan representasi fitur yang dapat diproses dalam model klasifikasi. TF-IDF telah terbukti efektif sebagai metode ekstraksi fitur untuk analisis sentimen karena kemampuannya memberikan bobot lebih tinggi pada kata-kata diskriminatif yang membedakan antar kelas sentimen (Joergensen E Munthe et al., 2022; Utami et al., 2023). Gambar 6 menunjukkan alur proses feature extraction yang diterapkan.

```

1 from sklearn.feature_extraction.text import TfidfVectorizer
2
3 # TF-IDF Feature Extraction
4 tfidf = TfidfVectorizer(
5     ngram_range=(1,2),
6     max_features=1000,
7     min_df=2,
8     max_df=0.95,
9     sublinear_tf=True
10 )
11
12 X_tfidf = tfidf.fit_transform(df['preprocessing_text'])

```

Gambar 6. Proses Feature Extraction

Proses ini diimplementasikan menggunakan TfidfVectorizer dengan beberapa parameter khusus untuk mengoptimalkan representasi fitur. Tabel 9. menyajikan parameter yang digunakan dalam feature extraction.

Tabel 9. Parameter Feature Extraction TF-IDF

Parameter	Nilai
ngram_range	(1,2)
max_features	1000
min_df	2
max_df	0.95
Sublinear_tf	True

Hasil proses ini menghasilkan 1000 fitur teks yang siap digunakan pada tahap pemodelan. Tabel 10. menunjukkan contoh hasil feature extraction dalam bentuk matriks TF-IDF.

Tabel 10. Hasil Feature Extraction

Document_id	Term	TF-IDF
0	muncul	0.587920
0	voucher	0.526966
0	aneh	0.347235
0	voucher goride	0.302120
0	goride tidak	0.254053
0	tolong	0.200205
0	baik	0.199515
0	tidak	0.121079
0	goride	0.75609
1	aplikasi goride	0.478957
...	...	...
659	beli paket	0.453337

4. Hasil Modelling

Pemodelan algoritma merupakan tahap utama dalam analisis sentimen untuk membangun model klasifikasi berdasarkan data ulasan hasil preprocessing. Proses diawali dengan pembagian data (data splitting) untuk memisahkan data latih dan uji, kemudian dilanjutkan penerapan algoritma Support Vector Machine (SVM) untuk melatih model dalam mengenali pola sentimen.

4.1. Data Splitting

Tahap awal pemodelan adalah pembagian data (data splitting) untuk memisahkan dataset menjadi data latih (training set) dan data uji (testing set) untuk melakukan memvalidasi model secara objektif dan mencegah overfitting. Gambar 7 menunjukkan proses splitting data yang diterapkan.

```
1 X = df['ulasan_bersih']
2 y = df['label']
3
4 # Bagi data menjadi 90% data latih dan 10% data uji
5 X_train, X_test, y_train, y_test = train_test_split(
6     X, y, test_size=0.1, stratify=y, random_state=42
7 )
8
9 print(f"Jumlah data latih: {len(X_train)}")
10 print(f"Jumlah data uji: {len(X_test)}")
```

Gambar 7. Proses Data Splitting

Proporsi pembagian data menggunakan rasio 90:10 dengan 593 data latih dan 66 data uji dari total 659 data seperti disajikan pada Tabel 11.

Tabel 11. Skema Pembagian Data

Komponen	Jumlah Data	Persentase
Data Latih (Train)	593	90%
Data Uji (Test)	66	10%
Total Data	659	100%

4.2. Klasifikasi Support Vector Machine (SVM)

Klasifikasi SVM diimplementasikan untuk membangun model prediksi sentimen dengan optimasi parameter yang tepat. Model dilatih menggunakan dataset latih hasil preprocessing dan feature extraction untuk mempelajari pola karakteristik pembeda sentimen pada setiap ulasan. Optimasi kinerja dilakukan melalui hyperparameter tuning menggunakan metode GridSearchCV seperti ditunjukkan pada Gambar IV.13.

```

1 # Membuat pipeline untuk menggabungkan langkah-langkah pemodelan
2 pipeline = ImbPipeline([
3     ('tfidf', TfidfVectorizer(ngram_range=(1, 2),
4                             max_features=1000,
5                             min_df=1,
6                             max_df=1.0,
7                             sublinear_tf=True)),
8
9     ('smote', SMOTE(random_state=42)),
10
11     ('svm', SVC(probability=True, random_state=42))
12 ])
13
14 # Menentukan parameter yang akan diuji untuk menemukan kombinasi terbaik
15 param_grid = {
16     'svm__C': [0.1, 1, 10],
17     'svm__kernel': ['linear', 'rbf', 'poly', 'sigmoid'],
18     'svm__gamma': ['scale', 'auto']
19 }
20
21 # Membuat objek GridSearchCV untuk melakukan pencarian hiperparameter dengan cross-validation
22 grid_search = GridSearchCV(
23     pipeline, param_grid=param_grid,
24     scoring='f1', cv=5, verbose=1, n_jobs=-1
25 )
26
27 # Melatih model dengan data latih
28 print("Memulai pencarian hiperparameter terbaik...")
29 grid_search.fit(X_train, y_train)
30
31

```

Gambar 8. Hyperparameter Tuning dengan GridSearchCV

Proses tuning melibatkan tiga hyperparameter utama yaitu C, kernel, dan gamma dengan ruang pencarian nilai berdasarkan eksperimen awal. Parameter C diuji dengan nilai [0.1, 1, 10], kernel dengan ['linear', 'rbf', 'poly', 'sigmoid'], dan gamma dengan ['scale', 'auto']. Hasil pengujian berbagai konfigurasi parameter menghasilkan kombinasi optimal yang memberikan kinerja klasifikasi terbaik sebagaimana disajikan pada Tabel IV.12.

Tabel 12. Nilai Parameter Terbaik Model SVM

Hyperparameter	Nilai Uji	Nilai Terbaik
C	[0.1, 1, 10]	0.1
Kernel	['linear', 'rbf', 'poly', 'sigmoid']	Linear
Gamma	['scale', 'auto']	Scale

Konfigurasi parameter optimal diperoleh pada nilai C = 0.1, kernel = linear, dan gamma = scale yang selanjutnya digunakan sebagai parameter final dalam pelatihan model SVM untuk dataset ulasan pengguna layanan GoRide seperti ditampilkan pada Gambar 9.

```

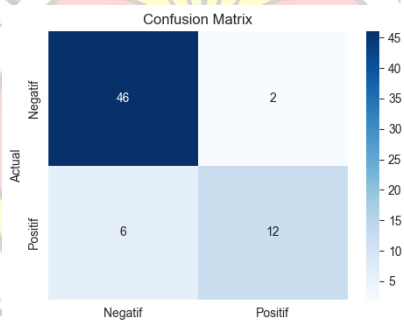
1 # Modelling SVM dengan parameter terbaik
2 best_svm_model = SVC(
3     C=0.1,
4     kernel='linear',
5     gamma='scale',
6     probability=True,
7     random_state=42
8 )
9
10 best_svm_model.fit(X_train_balanced, y_train_balanced)
11 y_pred_final = best_svm_model.predict(X_test_tfidf)
12

```

Gambar 9. Modelling Algoritma SVM

5. Hasil Evaluation

Evaluasi performa model SVM dilakukan menggunakan data uji independen untuk mengukur kemampuan klasifikasi sentimen. Hasil confusion matrix disajikan pada Gambar 8 yang menunjukkan distribusi prediksi model terhadap kelas aktual.



Gambar 10. Confution Matrix Model SVM

Berdasarkan confusion matrix pada Gambar 8, model menghasilkan True Positive (TP) sebanyak 12, True Negative (TN) sebanyak 46, False Positive (FP) sebanyak 6, dan False Negative (FN) sebanyak 2 dari total 66 data uji. Tabel 3 menyajikan perhitungan metrik evaluasi berdasarkan nilai confusion matrix.

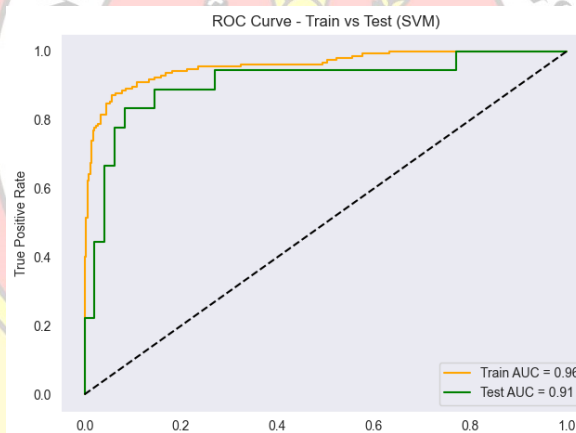
Tabel 13. Confution Matrix Ulasan Goride

	True Positive	True Negative	Class Precision
Pred.Positive	12	6	66.67%
Pred.Negative	2	46	95.83%
Class Recall	85.71%	88.46%	

Perhitungan dari tabel matrix evaluasi:

- $Precision = \frac{TP}{TP+FP} = \frac{12}{12+6} = 0.6667 = 66.67\%$
- $Recall = \frac{TP}{TP+FN} = \frac{12}{12+2} = 0.8571 = 85.71\%$
- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{12+46}{12+46+6+2} = 0.8788 = 87.88\%$
- $F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.6667 \times 0.8571}{0.6667 + 0.8571} = 0.7500 = 75.00\%$

Selain evaluasi berdasarkan confusion matrix, performa model juga diukur menggunakan kurva ROC (Receiver Operating Characteristic) dan nilai AUC (Area Under Curve) untuk menilai kemampuan model dalam membedakan kelas positif dan negatif pada berbagai threshold. Gambar 9 menunjukkan kurva ROC untuk data training dan testing dengan nilai AUC sebesar 0.96 pada data training dan 0.91 pada data testing, mengindikasikan kemampuan diskriminasi model yang sangat baik dan konsisten antara fase training dan testing.

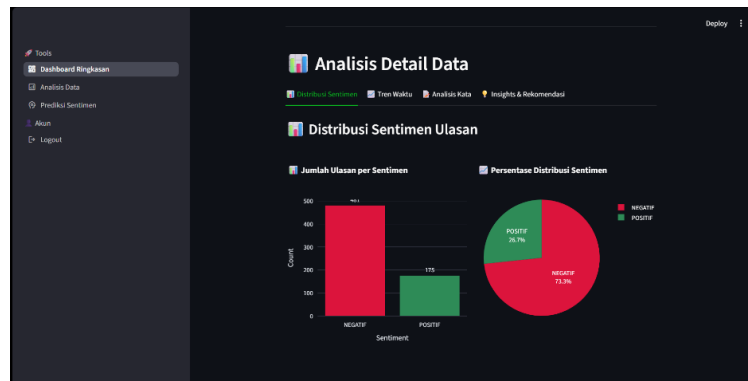


Gambar 11. ROC-AUC Model SVM

6. Hasil Deployment

Tahap deployment merupakan implementasi model SVM ke dalam aplikasi web bernama SentimentGo untuk menganalisis sentimen publik terhadap layanan GoRide. Aplikasi ini menyediakan antarmuka interaktif yang memudahkan interpretasi hasil analisis sesuai rancangan sistem. Dashboard ringkasan menampilkan distribusi dan proporsi ulasan positif maupun negatif melalui

visualisasi yang informatif, memungkinkan pengguna memperoleh gambaran persepsi publik secara cepat dan efisien (Gambar 1).



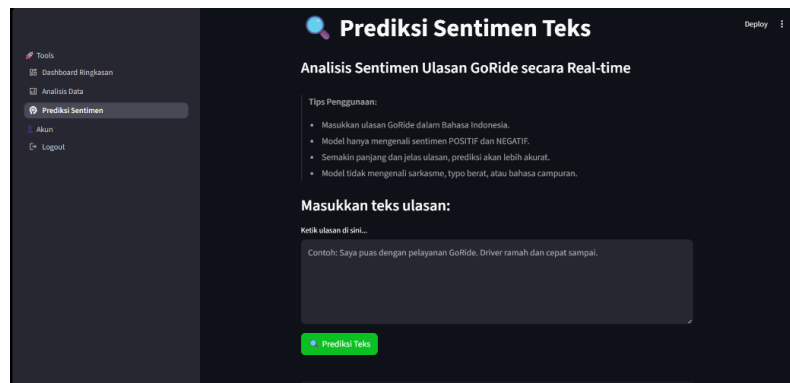
Gambar 12. 1 Fitur Dashboard Ringkasan

Fitur Analisis Data memfasilitasi pengunggahan dataset CSV yang diproses otomatis melalui preprocessing dan klasifikasi model SVM, menghasilkan visualisasi interaktif untuk pemantauan sentimen berkala (Gambar 2).



Gambar 13. Fitur Analisis Data

Halaman Prediksi Sentimen menyediakan klasifikasi real-time dimana pengguna memasukkan teks ulasan yang diproses melalui preprocessing standar dan menghasilkan prediksi beserta tingkat kepercayaan sebagai validasi performa model (Gambar 3).



Gambar 13. Fitur Prediksi Sentimen

Uji Sistem Aplikasi SentimentGo

Pengujian fungsionalitas sistem dilakukan untuk memastikan setiap fitur aplikasi berjalan dengan baik menggunakan metode Black Box Testing yang berfokus pada perspektif pengguna akhir tanpa melihat struktur kode internal. Tabel 4 menyajikan hasil pengujian untuk setiap skenario pengujian aplikasi SentimenGo.

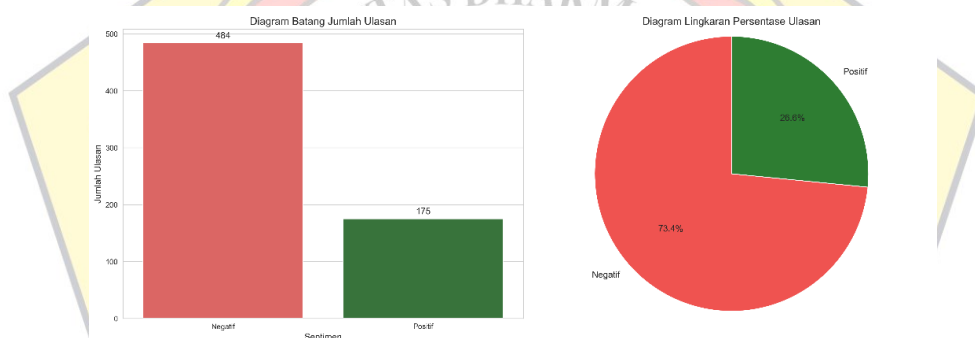
Tabel 14. Black Box Testing Case Aplikasi SentimentGo

No	Skenario Pengujian	Test Case	Hasil yang Diharapkan	Kesimpulan
1	Autentikasi Masuk	Login dengan email dan password valid	Pengguna berhasil login dan diarahkan ke dashboard	Valid
2	Autentikasi Masuk	Login dengan email salah	Pesan error "Email tidak ditemukan" muncul	Valid
3	Autentikasi Masuk	Login dengan password salah	Pesan error "Kata sandi salah" muncul	Valid
4	Autentikasi Masuk	Login dengan email belum verifikasi	Pesan verifikasi email muncul	Valid
5	Daftar Akun	Registrasi dengan data valid	Akun berhasil dibuat dan email verifikasi dikirim	Valid
6	Daftar Akun	Registrasi dengan email sudah terdaftar	Pesan error "Email sudah terdaftar" muncul	Valid
7	Daftar Akun	Registrasi dengan password lemah	Pesan error "Kata sandi terlalu lemah" muncul	Valid
8	Reset Password	Reset password dengan email terdaftar	Email reset password dikirim	Valid
9	Reset Password	Reset password dengan email tidak terdaftar	Pesan error "Email tidak ditemukan" muncul	Valid
10	Analisis Data	Upload file CSV valid dan proses analisis	Hasil analisis dan visualisasi berhasil muncul	Valid
11	Analisis Data	Upload file CSV tanpa kolom teks	Pesan error validasi muncul	Valid

12	Prediksi Sentimen	Input teks ulasan valid	Hasil prediksi sentimen dan confidence muncul	Valid
13	Prediksi Sentimen	Input teks kosong	Pesan error input kosong muncul	Valid

B. Visualisasi Diagram dan Wordcloud

Berdasarkan hasil klasifikasi model SVM terhadap 659 ulasan pengguna layanan GoRide yang telah melalui tahap preprocessing dan feature extraction, analisis distribusi sentimen dilakukan untuk memahami pola persepsi pengguna secara keseluruhan. Diagram batang dan pie chart digunakan untuk menyajikan proporsi sentimen secara visual dan memudahkan interpretasi data. Distribusi sentimen ulasan GoRide secara keseluruhan disajikan pada Gambar 13.



Gambar 14. Visualisasi Distribusi Sentimen

Gambar 13 menunjukkan sentimen negatif mendominasi dengan 484 ulasan (73.4%), sedangkan sentimen positif sebanyak 175 ulasan (26.6%). Ketidakseimbangan ini mengindikasikan adanya permasalahan signifikan yang perlu mendapat perhatian dalam layanan GoRide. Untuk mengidentifikasi aspek dominan dari setiap kategori sentimen, dilakukan analisis mendalam menggunakan wordcloud dan frekuensi kata seperti ditunjukkan pada Gambar 14 dan Gambar 15.

SIMPULAN

Implementasi metode Support Vector Machine (SVM) dalam analisis sentimen ulasan GoRide menunjukkan hasil yang efektif untuk mengklasifikasikan 659 ulasan tidak terstruktur menjadi kategori positif dan negatif melalui kerangka kerja CRISP-DM. Model SVM yang dikembangkan menunjukkan performa klasifikasi yang baik dengan tingkat accuracy 87.88%, precision 66.67%, recall 85.71%, dan F1-score 75%, serta nilai AUC sebesar 0.91 yang mengindikasikan kemampuan diskriminasi model sangat baik. Model ini berhasil diimplementasikan menjadi aplikasi fungsional berbasis web bernama SentimenGo untuk merepresentasikan visualisasi analisis dan prediksi sentimen secara interaktif. Hasil analisis sentimen berhasil mengidentifikasi aspek-aspek kunci yang mempengaruhi persepsi pengguna berdasarkan visualisasi diagram, wordcloud, dan analisis TF-IDF. Sentimen positif sangat dipengaruhi oleh aspek penawaran promosi dan efisiensi layanan yang ditandai dengan kata seperti "promo", "goride", dan "driver". Sebaliknya, sentimen negatif berpusat pada aspek waktu tunggu dan kendala teknis aplikasi yang tercermin dari kata seperti "tidak", "lama", "batal", dan "aplikasi", menunjukkan perlunya perbaikan pada performa platform dan sistem pemesanan untuk meningkatkan kepuasan pengguna layanan GoRide.

DAFTAR PUSTAKA

- Adhi Putra, A. D. (2021). Sentiment Analysis on User Reviews of the Bibit and Bareksa Application with the KNN Algorithm. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(2), 636–646.
- Alamsyah, A. S., Budiman, S. N., & Romadhona, R. D. (2024). Penerapan Algoritma Support Vector Machine Untuk Analisis Sentimen Data Ulasan Aplikasi Binance Pada Google Play Store. *Journal Of Social Science Research*, 4(5), 6145–6158.
- Dwianto, E., & Sadikin, M. (2021). Analisis Sentimen Transportasi Online pada Twitter Menggunakan Metode Klasifikasi Naïve Bayes dan Support Vector Machine. *Format : Jurnal Ilmiah Teknik Informatika*, 10(1), 94. <https://doi.org/10.22441/format.2021.v10.i1.009>
- Hakim, A. R. ... Syarif, A. R. (2023). Analisis Perbandingan Algoritma Machine Learning Terhadap Sentimen Analisis Pemindahan Ibu Kota Negara. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(2), 179–185. <https://doi.org/10.35870/jtik.v7i2.701>
- Harimurti, C., & Sofyan, M. (2022). The Merging of Gojek and Tokopedia Forms Go To. Is This a

Threat to National Banking and Fintech? *Management, Business and Social Science (IJEMBIS) Peer-Reviewed-International Journal*, 2(3).
<https://cvodis.com/ijembis/index.php/ijembis><https://cvodis.com/ijembis/index.php/ijembis/article/view/380>

Haviana, S. F. C., & Poetro, B. S. W. (2022). Deep Learning Model for Sentiment Analysis on Short Informal Texts. *Indonesian Journal of Electrical Engineering and Informatics*, 10(1), 82–89.
<https://doi.org/10.52549/ijeei.v10i1.3181>

Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, 5(1), 32–35. <https://doi.org/10.37905/jjee.v5i1.16830>

Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1120–1126. <https://doi.org/10.29207/resti.v5i6.3588>

Joergensen E Munthe, C., Astuti Hasibuan, N., & Hutabarat, H. (2022). Penerapan Algoritma Text Mining Dan TF-IDF Dalam Menentukan Promo Produk Pada Marketplace. *Resolusi : Rekayasa Teknik Informatika Dan Informasi*, 2(3), 110–115.
<https://doi.org/10.30865/resolusi.v2i3.309>

Katadata.co.id. (2022). *Riset INDEF: Gojek Pimpin Pasar Taksi dan Ojek Online Tahun Ini*. 2022-12-07. <https://katadata.co.id/digital/startup/63903b6aec7e5/riset-indef-gojek-pimpin-pasar-taksi-dan-ojek-online-tahun-ini>

Mehta, P., & Pandya, S. (2020). A Review On Sentiment Analysis Methodologies, Practices And Applications. *International Journal of Scientific and Technology Research*, 9(2), 601–609.

Pranata, A., Budianita, E., Yusra, & Cynthia, E. P. (2022). Klasifikasi Sentimen Terhadap Maxim Menggunakan Algoritma SVM Pada Media Sosial Twitter. *Klasifikasi Sentimen Terhadap Maxim Menggunakan Algoritma SVM Pada Media Sosial Twitter*, 5(3), 332–341.
<https://ojs.serambimekkah.ac.id/jnkti/article/view/4253>

Rizki, M. F., Auliasari, K., & Primaswara Prasetya, R. (2021). Analisis Sentiment Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 5(2), 548–556. <https://doi.org/10.36040/jati.v5i2.3808>

Rokhman, K. A., Berlilana, B., & Arsi, P. (2021). Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online. *Journal of Information System Management (JOISM)*, 3(1), 1–7.
<https://doi.org/10.24076/joism.2021v3i1.341>

Utami, M. A. A. T., Silvianti, P., & Masjukur, M. (2023). Algoritme Support Vector Machine untuk Analisis Sentimen Berbasis Aspek Ulasan Game Online Mobile Legends: Bang-Bang. *Xplore: Journal of Statistics*, 12(1), 63–77. <https://doi.org/10.29244/xplore.v12i1.1064>