ANALISIS SENTIMEN GOOGLE MAPS *REVIEW* WULAN RENT CAR MENGGUNAKAN *SUPPORT VECTOR MACHINE* (SVM)

Septian Wulandari[1*], Dian Novita[2], Agus Wilson[3]
[1,2,3]Teknik Informatika, Universitas Indraprasta PGRI
septian.pmb09@rocketmail.com[1], Dyan.novita21@gmail.com, wilsonaw2580@gmail.com[3]

Submitted May 4, 2026; Revised July 21, 2026; Accepted July 24, 2026

Abstrak

Ulasan pada Google Maps menjadi sumber data penting bagi perusahaan, termasuk bisnis rental mobil, dalam memahami persepsi konsumen. Mengevaluasi data yang cukup besar secara manual menjadi tidak efisien, sehingga diperlukan pemanfaatan teknologi untuk mengevaluasi emosi atau opini yang ada dalam teks (analisis sentimen). Tujuan penelitian ini adalah mengklasifikasikan sentimen ulasan Google Maps pada Wulan Rent Car menggunakan metode *Support Vector Machine* (SVM). Terdapat 443 ulasan, dengan 301 ulasan yang memiliki teks. Tahapan pelabelan menggunakan Bing Liu Opinion Lexicon, *pre-processing*, serta pemberian nilai menggunakan TF-IDF. Hasil pelabelan menunjukkan terdapat 269 positif, 18 netral, dan 14 negatif. Hasil evaluasi model menunjukkan nilai akurasi sebesar 84,75%. Berdasarkan nilai *precision*, *recall*, dan *F1-score*, model menunjukkan performa yang baik pada kelas sentimen positif, namun belum mampu mengklasifikasikan sentimen negatif secara optimal akibat distribusi data yang tidak seimbang. Hasil penelitian menunjukkan bahwa metode SVM dapat diterapkan untuk analisis sentimen ulasan Google Maps pada Wulan Rent Car, namun performa model masih dipengaruhi oleh karakteristik distribusi data, khususnya pada kelas minoritas. Oleh karena itu, penelitian selanjutnya disarankan menerapkan teknik penyeimbangan data, seperti *Synthetic Minority Over Sampling Technique* (SMOTE) atau metode serupa, untuk meningkatkan kemampuan model dalam mengklasifikasikan seluruh kelas sentimen.

Kata Kunci: google maps *review*, sentimen, *Support Vector Machine* (SVM)

Abstract

Google Maps reviews serve as a vital data source for companies including car rental businesses to understand consumer perceptions. Manually evaluating such vast amounts of data is inefficient; therefore, technology-based approaches are required to analyze the emotions or opinions embedded in the text (sentiment analysis). This study aims to classify the sentiment of Google Maps reviews for Wulan Rent Car using the Support Vector Machine (SVM) method. Out of 443 total reviews, 301 contained text and underwent further processing, including labeling via the Bing Liu Opinion Lexicon, data pre-processing, and TF-IDF-based scoring. The labeling results identified 269 positive, 18 neutral, and 14 negative reviews. Model evaluation yielded an accuracy of 84.75%. Based on precision, recall, and F1-score metrics, the model demonstrated strong performance for the positive sentiment class but struggled to optimally classify negative sentiment due to imbalanced data distribution. The findings indicate that while the SVM method applies to sentiment analysis of Wulan Rent Car's Google Maps reviews, model performance remains influenced by data distribution characteristics, particularly regarding minority classes. Consequently, future research should incorporate data balancing techniques—such as the Synthetic Minority Over-Sampling Technique (SMOTE) or similar methods—to enhance the model's ability to classify all sentiment categories effectively.

Keywords : google maps *review*, sentiment, *Support Vector Machine* (SVM)

1. PENDAHULUAN

Perkembangan teknologi dan informasi yang cukup pesat mendorong peningkatan penggunaan platform digital sebagai media penyampaian pendapat dan pengalaman pelanggan. Salah satu platform yang banyak digunakan dalam penyampaian pendapat adalah Google Maps. Pada Google Maps terdapat fitur *review* (ulasan) yang berisikan

ulasan atau opini seseorang ketika telah mengunjungi lokasi tersebut. Ulasan tersebut berisikan berbagai opini mulai dari opini positif hingga negatif yang mencerminkan tingkat kepuasan konsumen terhadap layanan yang telah diberikan perusahaan tersebut. Ulasan opini tersebut tentu membantu pengguna untuk mencegah pengalaman buruk saat menentukan lokasi untuk mewujudkan keinginan mereka [1]. Google Maps *review* juga memudahkan pengguna untuk memberikan penilaian baik secara numerik (*rating*) maupun secara tertulis, serta dapat membagikan foto sehingga memudahkan penggunanya untuk mengakses secara publik [2]. Platform ini tidak hanya memberikan info secara kuantitatif, melainkan juga secara kualitatif yang menggambarkan tanggapan penggunanya terhadap penyajian fasilitas di lokasi tertentu. Sehingga, Google Maps *review* dapat menjadi media atau wadah informasi untuk bertukar pengalaman ketika mengunjungi tempat usaha. Google Maps *review* sering digunakan oleh penggunanya untuk menganalisis kelayakan suatu tempat yang ingin dikunjungi. Sering kali penggunanya menilai suatu tempat dari *rating* yang bagus atau ulasan yang memberikan banyak ulasan positif. Sebaliknya, jika *rating*nya rendah dan banyak ulasan negatif, maka mungkin saja penggunanya tidak tertarik mengunjungi lokasi tersebut. Sehingga, Google Maps *review* dapat menjadi bahan evaluasi suatu tempat atau pelaku usaha untuk menganalisis ulasan yang terdapat di Google Maps *review* secara cepat dan efisien untuk mengembangkan atau mengambil keputusan yang tepat terhadap strategi pemasaran mereka [3].

Bagi perusahaan rental mobil, Google Maps *review* menjadi sumber data yang sangat berharga karena memuat persepsi pelanggan secara langsung dan *real-time*. Namun, jumlah ulasan yang terus bertambah membuat proses pengkajian yang dilakukan dengan cara manual menjadi tidak efektif dan sensitif terhadap subjektivitas. Sehingga, diperlukan suatu metode yang tepat untuk memproses dan menganalisis data ulasan tersebut, yaitu melalui analisis sentimen. Dengan melakukan analisis sentimen, perusahaan atau pelaku bisnis dapat mengawasi secara langsung tanggapan pelanggan terhadap komoditas atau jasa mereka, seperti perusahaan rental mobil. Wulan Rent Car yang merupakan perusahaan di bidang penyedia layanan sewa mobil. Wulan Rent Car telah beroperasi sejak tahun 2010 dan berlokasi di Depok, Jawa Barat. Wulan Rent Car telah berhasil mengumpulkan banyak *review* di Google Maps, dengan *rating* yang telah dimiliki sebesar 4,9, yang tentunya merupakan *rating* yang cukup tinggi, yang diperkuat oleh ratusan testimoni yang terus meningkat seiring dengan berjalannya waktu. Hal ini tentunya menunjukkan bahwa Google Maps *review* berperan sangat penting dalam membangun serta mempertahankan reputasi di platform digital. Oleh karena itu, Google Maps *review* dapat menjadi acuan dalam mengevaluasi setiap layanan yang telah diberikan.

Analisis sentimen diekspresikan melalui teks dan bagaimana sentimen tersebut dianalisis berdasarkan sentimen positif dan negatif [4]. Bukan hanya itu, analisis sentimen dapat digunakan untuk mendeteksi sifat dan susunan kata [5]. Dengan menerapkan analisis sentimen, perusahaan dapat mengetahui kecenderungan persepsi pelanggan terhadap layanan mereka secara lebih sistematis dan terukur. Banyak sekali metode *machine learning* diterapkan dalam menganalisis sentimen seperti *Support Vector Machine* (SVM), *Naive Bayes*, *Deep learning*, dan *Convolution Neural Networks* (CNN) [6]. SVM adalah metode yang banyak diterapkan dalam mengkategorikan teks pada data yang berukuran besar. Metode SVM juga merupakan metode yang efisien [7]. Penelitian terdahulu yang

berkaitan dengan analisis sentimen telah dilakukan, seperti yang dilakukan oleh [4] yaitu mengoptimalkan analisis sentimen pada aplikasi Glints menggunakan metode SVM yang memperoleh hasil yang akurat, yaitu sebesar 92% dengan presisi 87%, recall 86%, dan F1-Score 86%. Hal ini mencerminkan evaluasi bahwa metode SVM dapat beroperasi dengan optimal pada setiap kategori sentimen. Kemudian, penelitian lainnya juga dilakukan oleh [4] yaitu mengkaji sentimen pada tanggapan produk Daviena yang ada di Shopee dengan menggunakan metode SVM dan memperoleh hasil 94% akurasi dengan perbandingan data *training* dan data *testing* yaitu sebesar 90:10. Sedangkan, penelitian lain yang dilakukan [8] menerapkan mentode *Naive Bayes* untuk analisis sentimen pada ulasan di aplikasi Ruangguru mendapatkan hasil terdapat akurasi 80% yaitu 8 kasus sentimen yang benar dari 10 kasus uji yang dilakukan. Di samping itu, penelitian lainnya yang dilakukan oleh [9] yaitu menerapkan metode *Naive Bayes* dan SVM pada ulasan sentimen pada aplikasi Pluang dan menghasilkan metode SVM menunjukkan keakuratan sebesar 99,5%, presisi 99,67%, *recall* 99,33%, dan *F1-score* 99,5%, sedangkan *Naive Bayes* memiliki akurasi sebesar 99,25%, presisi 99,44%, *recall* 99,06%, dan *F1-score* 99,25%. Sehingga dapat disimpulkan bahwa SVM memiliki kemampuan membedakan teks positif dan negatif yang lebih baik dibandingkan dengan Naive Bayes.

Berdasarkan penelitian terdahulu yang telah dijelaskan di atas, data yang digunakan pada analisis sentimen banyak berasal dari aplikasi ataupun *marketplace* seperti Glints, Ruangguru, Shopee, dan Pluang. Sedangkan, penelitian yang menggunakan sumber data seperti Google Maps Review, khususnya pada sektor jasa rental mobil, masih sangat terbatas. Kenyataannya, Google Maps Review berisikan pengalaman atau tanggapan pelanggan secara *real time* serta sangat berpengaruh pada reputasi dan keputusan para calon pelanggan untuk memilih penyedia jasa.
Di samping itu, penggunaan algoritma SVM menjadi relevan digunakan pada penelitian ini karena secara empiris dan teoritis dapat bekerja dengan baik pada jumlah dataset yang tidak terlalu banyak, relatif memiliki ketahanan terhadap *overfitting*, dan dapat menghasilkan kinerja yang lebih baik dibandingkan dengan algoritma Naive Bayes. Metode SVM dipilih karena dapat mereduksi dimensi pada teks, serta dapat menghasilkan tingkat akurasi yang cukup tinggi jika dibandingkan dengan metode klasifikasi lainnya. Selain itu, SVM efektif dalam menemukan *hyperplane* terbaik yang memisahkan data ke dalam kelas-kelas yang berbeda.

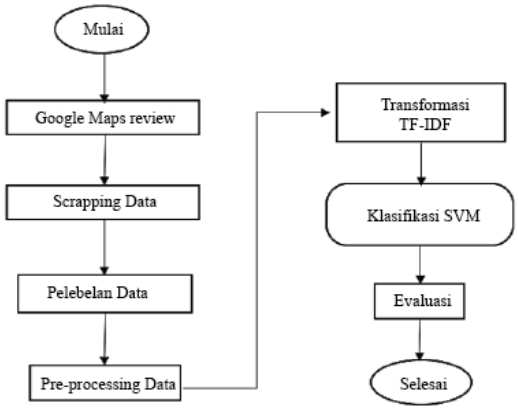
Berdasarkan aspek-aspek itulah, urgensi penelitian ini memperjelas kontribusi ilmiah yang membedakannya dari penelitian terdahulu. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut dengan menerapkan metode *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen pada ulasan Google Maps Review Wulan Rent Car sebagai dasar evaluasi kualitas layanan dan pendukung pengambilan keputusan bisnis.
Penggunaan metode SVM dalam analisis sentimen terhadap Google Maps *review* pada Wulan Rent Car diharapkan mampu memberikan hasil pengkategorian yang cukup akurat dan mengoptimalkan perusahaan dalam memahami kebutuhan serta harapan konsumen. Dengan demikian, perusahaan dapat meningkatkan kualitas layanan, memperbaiki kekurangan, serta menjaga daya saing strategis dalam kompetisi pasar yang semakin kompetitif.

2. METODE PENELITIAN

Data yang dihimpun pada penelitian ini dihasilkan dari google maps *review* pada perusahaan Wulan Rent Car dengan

menggunakan web pada laman www.nocodeserpapi.com diperoleh sebanyak 443 dataset yang berisikan review dari pelanggan sampai dengan bulan Agustus 2025. Kemudian dataset disimpan dalam bentuk Excel untuk memudahkan proses selanjutnya.

Selanjutnya, dataset dipartisi menjadi dua subset yaitu data *testing* dan data *training*. Data *training* berfungsi sebagai dasar pada pemodelan klasifikasi SVM, sedangkan data *testing* berfungsi untuk menilai performa model pada data yang tidak terlihat. Pembagian data menjadi dua bagian memudahkan peneliti dalam mengukur seberapa baik model SVM dalam mengklasifikasikan ulasan sentimen pada Google Maps R*eview*.



Gambar 1. Alur Penelitian

Pelabelan Data

Pada proses pelabelan data pada ulasan dilakukan dengan mengategorikan menjadi tiga kategori, yaitu positif, netral, dan negatif. Namun, untuk memudahkan pelabelan, dataset yang berisi ulasan dalam bahasa Indonesia diubah ke bahasa Inggris. Proses pelabelan menggunakan *Bing Liu Opinion Lexicon* yaitu melalui fungsi *get_sentiments* ("*bing*") pada *package tidytext* yang terdapat pada *software* Rstudio. *Bing Liu Opinion Lexicon* merupakan kamus yang memuat kata-kata opini beserta kategori sentimen atau ukuran numerik yang mempresentasikan polaritas dan kekuatan kata yang terkait dengannya [10].

Proses pelabelan dilakukan dengan melihat setiap kata pada ulasan. Diberikan label positif untuk ulasan yang terdapat kata "*good, great, excellent, love,* dll", dan diberikan label negatif untuk ulasan yang terdapat kata "*bad, terrible, hate,* dll", sedangkan diberikan label netral untuk kata seperti *"car", "time", "staff*", dll.

Pre-Processing Data

Langkah *pre-processing* diterapkan untuk mengoptimalkan proses analisis data dengan tahapan pada *pre-processing* yaitu [11]:

1. *Cleaning* yaitu menghilangkan tanda baca yang tidak berguna seperti titik (.), titik dua (:), koma (,), dan lain sebagainya.
2. *Case Folding* yaitu mengkonversi menjadi huruf kecil untuk menyeragamkan karakter pada dataset
3. *Normalization* yaitu melakukan standarisasi kata non baku menjadi kata baku.
4. *Stopword Removal* yaitu menghapus kata tugas seperti kata hubung "dan", "atau", "ini", dan lain sebagainya.
5. *Tokenization* yaitu memisahkan kalimat menjadi kata-kata yang terpisah.

Proses tersebut dilakukan untuk memastikan dataset yang digunakan telah bersih, seragam, serta siap untuk dianalisis lebih lanjut dengan metode *Support Vector Machine* (SVM).

Term Frequency-Inverse Document Frequency (TF-IDF)

Setelah tahap *pre-processing* data, data diubah ke bentuk numerik melalui pembobotan kata. Pada proses ini, bobot diberikan pada setiap kata yang siap untuk dianalisis dengan dua konsep utama, yaitu TF (Ter*m Frequency*) dan IDF (*Inverse Document Frequency*). Metode ini mwngintegrasikan representasi TF dalam dokumen dengan frekuensi dokumen terbaik IDF sebagai basis dalam

pembobotan kata [12]. TF merupakan menghitung seberapa sering kata di seluruh kumpulan data, sedangkan IDF menilai signifikasi sebuah kata di seluruh kumpulan data. Proses normalisasi ini memastikan bahwa kata dominan menerima nilai TF yang lebih tinggi. Sederhananya, frekuensi kemunculan yang tinggi dalam dokumen akan memperoleh bobot yang lebih besar. Sedangkan nilai IDF membantu membedakan kata mana yang lebih spesifik dan unik dalam sebuah dokumen.

Penggabungan TF-IDF memberikan bobot yang mencerminkan pentingnya sebuah kata dalam dokumen tertentu dan mempertimbangkan seluruh kumpulan data. Berikut merupakan rumus dalam menghitung bobot (W) terhadap kata kunci adalah [13]:

$$W_{t,d} = TF_{t,d} * IDF_t \quad (1)$$

Keterangan:
d =dokumen ke- d
t = kata ke- t dari kata kunci
W = bobot dokumen ke- d terhadap kata ke- t

Metode *Support Vector Machine* (SVM)

Metode *Support Vector Machine* (SVM) adalah salah satu metode atau model untuk klasifikasi. SVM menggunakan fungsi *hyperplane* pada dataset sehingga membentuk daerah-daerah pada kelas. Prinsipnya, metode ini membangun *hyperplane* yang mempunyai margin yang sama dan tidak cenderung mendekati daerah dari salah satu kelas [14]. Proses klasifikasi memiliki dua langkah, yaitu proses pengujian dan proses pelatihan. Proses pelatihan digunakan untuk membuat model pada set pengujian [15].

Berdasarkan hasil perhitungan TF-IDF, maka pencarian *hyperplane* pada SVM dapat dilakukan dengan menggunakan rumus [13]:

$$f(\boldsymbol{x}): \boldsymbol{w}^T \boldsymbol{x} + b = 0 \quad (2)$$

Keterangan:
$\boldsymbol{x}$ = vektor fitur TF-IDF
$\boldsymbol{w}$ = vektor bobot hasil penelitian
b = bias

Evaluasi

Evaluasi diperlukan pada penelitian ini untuk mengevaluasi performa model yang diusulkan. *Confusion matrix* merupakan teknik penilaian yang digunakan pada penelitian ini. *Confusion matrix* adalah suatu perhitungan dari algoritma *supervised learning* yang menggunakan *machine learning* dalam labeling dataset. Pelabelan yang digunakan pada metode ini adalah *actual value* dan *predicted value*. *Actual value* adalah nilai yang sesuai dengan data yang sudah ada. *Predicted value* merupakan prediksi dari hasil pemodelan. Hasil *confusion matrix* pada dilihat dengan menggunakan gambar berikut ini:

		Actual value	
		Positif	Negatif
Predicted value	Positif	TP	FP
	Negatif	FN	TN

Sumber: [13]

Gambar 2. Gambar *Confusion Matrix*

Gambar 2 menunjukkan *confusion matrix* dengan keterangan TP = *True* Positif, TN = *True* Negatif , FN = *False* Negatif, dan FP = *False* Positif. Berdasarkan *confusion matrix* diatas, maka nilai *acuracy, precision, recall*, dan *F1-score* dapat dihitung dengan menggunakan rumus [14]:

$$acuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = \frac{2*Precision*Recall}{Precision*Recall} \quad (6)$$

3. HASIL DAN PEMBAHASAN

Data yang dikumpulkan melalui proses *scraping* data terdapat 443 dataset ulasan. Ulasan dataset yang diperoleh masih dalam bentuk Bahasa Indonesia, sehingga perlu dilakukan perubahasan menjadi Bahasa Inggris agar memudahkan dalam pengolah data. Berikut merupakan dataset yang telah dikumpulkan melalui proses *scraping*:

	rating	user	review.en
1	5	{ "name": "Ahmad Fajri Shauti", "link": "https://www...	Amazing experience!! Best unit, best service, and bes
2	5	{ "name": "Sakti Nugroho", "link": "https://www.goo...	Wulan Rent Car never disappoints. They have the best
3	5	{ "name": "Elsa Risfadona", "link": "https://www.goo...	Rental subscription. Great cars and affordable prices,
4	5	{ "name": "Nindha yonna nf", "link": "https://www.g...	It's been 5 years of renting and it never fails, always c
5	5	{ "name": "Amelia Sherly Ayuningtyas", "link": "https...	best rental car in depok city..
6	5	{ "name": "Tria Hadi", "link": "https://www.google.c...	Nice place, car, and sweet septi
7	5	{ "name": "Refka McCallister", "link": "https://www....	Joss
8	5	{ "name": "Desy putri rejeki", "link": "https://www.g...	Good
9	5	{ "name": "Gordon Sinaga", "link": "https://www.go...	Nice car
10	5	{ "name": "Cheap Used Goods", "link": "https://www...	Best rental car 🤟🙏
11	5	{ "name": "maria veronika maylinda sari", "link": "htt...	Best wedding car ❤❤❤
12	5	{ "name": "Fiqih Ramadhan", "link": "https://www.g...	Good quality, good price!
13	5	{ "name": "Reza Aminudin", "link": "https://www.go...	The best👌
14	5	{ "name": "Zaenal Abidin", "link": "https://www.goo...	Okay👍
15	5	{ "name": "Riztha Ronanda", "link": "https://www.go...	Good!
16	5	{ "name": "Amrul Hanifah", "link": "https://www.goo...	👍🔥
17	5	{ "name": "Hikmah Ratna Pratiwi", "link": "https://w...	Recommended!

Gambar 3. Potongan Hasil *Scrapping* Data Google Maps *Review* Wulan Rent Car

Dari 443 dataset yang telah dikumpulkan, terdapat 301 dataset yang memiliki ulasan teks, sedangkan pada 142 dataset tidak terdapat ulasan teks. Sehingga, pada penelitian ini 301 dataset dilakukan pada proses pelabelan selanjutnya. Proses pelabelan dilakukan dengan melihat setiap kata pada ulasan. Diberikan label positif untuk ulasan yang terdapat kata "*good, great, excellent, love,* dll", dan diberikan label negatif untuk ulasan yang terdapat kata "*bad, terrible, hate,* dll", sedangkan diberikan label netral untuk kata seperti *"car", "time", "staff*", dll. Pada proses pelabelan ini menggunakan *script*:

```
#kamus sentimen
library(tidytext)

sentiment_lexicon <- get_sentiments("bing")

#pelabelan kata
df_sentiment <- df_token %>%
  mutate(sentiment = case_when(
    word %in% positive_words ~ "positive",
    word %in% negative_words ~ "negative",
    TRUE ~ "neutral"
  ))

#hitung sentimen per review bing
df_sentiment <- df_token %>%
  inner_join(sentiment_lexicon, by = "word")

#label akhir
df_label <- df_label %>%
  mutate(label = case_when(
    score > 0 ~ "positif",
    score < 0 ~ "negatif",
    TRUE ~ "netral"
  ))
```

Gambar 4. *Script* Pelabelan Dataset

Dataset yang telah diperoleh masih berbentuk kalimat yang mengandung tanda baca, singkatan kata, kata hubung, dan kata-kata yang terkadang sulit untuk dimengerti. Sehingga diperlukan tahapan selanjutnya, yaitu tahapan *pre-processing*. Tahapan *pre-processing* adalah *cleaning, case folding, normalization, stopword removal*, dan *tokenization*. Tahapan *pre-processing* data dilakukan dengan menggunakan *software* Rstudio. Tabel 1 menunjukkan contoh hasil pelabelan setiap kata. Pada dataset, terdapat 122 kata dengan label negatif dan 890 kata dengan label positif. Tabel 2 merupakan contoh hasil pelabelan dan *pre-processing* ulasan Wulan rent Car dan diperoleh hasil 14 sentimen negatif, 18 sentimen netral, dan 269 sentimen positif.

Tabel 1. Contoh Potongan Hasil Pelabelan Kata

Kata	Sentimen	Kata	Sentimen
amazing	positif	*love*	positif
disappoints	negatif	*nice*	positif
affordable	positif	*recommended*	positif
fails	negatif	*scratches*	negatif
clean	positif	*dents*	negatif

Tabel 2. Contoh Hasil Pelabelan dan *Pre-Processing* Ulasan Wulan Rent Car

Review.en	*Preprocessing*	positif	negatif	Score	Label
The car is in good condition, the engine is in good condition,,	*"car", "good", "condition", "engine", "driven", "uphill"*	2	0	2	positif

it can still be driven uphill					
The price can be said to be cheap compared to other rental prices	*"price", "cheap", "compare", "rental"*	0	1	-1	negatif
Easy conditions, no. hassle	*"easy", "condition", "hassle"*	1	1	0	netral
The interior condition is still quite good, even though some of the seats are torn	*"interior", "condition", "quite", "good", "seat", "torn"*	1	0	1	positif

Metode *Term Frequency-Inverse Document Frequency* (TF-IDF) adalah tahapan dalam vektorisasi numerik dengan memberikan nilai pada dataset. Kemudian, penggabungan nilai TF-IDF memberikan bobot yang mencerminkan pentingnya sebuah kata dalam dokumen tertentu dan mempertimbangkan seluruh kumpulan data. Hasil nilai TF, IDF, dan TF-IDF terlihat pada Tabel 3.

Tabel 3. Hasil nilai TF, IDF, dan TF-IDF

Kata	TF	IDF	TF-IDF
good	1,000000	1,4673356	1,46733557
rent	0,04651163	1,5206816	0,07072937
wulan	0,04651163	1,4804077	0,6885617
best	0,33333333	1,9824996	0,66083320
rental	0,04545455	1,2364300	0,05620136
service	0,04545455	1,0489671	0,04768032

Setelah melakukan proses pelabelan data, *pre-processing*, dan pembobotan, tahapan selanjutnya adalah melakukan Tahapan kategorisasi berbasis metode SVM ini memisahkan sentiment ke dalam kelas positif, netral, dan negatif. Tahapan ini dilakukan untuk menilai efektivitas model menggunakan *confusion matrix*. Penggunaan *confusion matrix* ini dinilai sangat penting untuk mengevaluasi kinerja model, yaitu dengan mengevaluasi kinerja prediksi dengan hasil klasifikasi . *Confusion* m*atrix* memberikan nilai untuk TP = *True* Positif, TN = *True* Negatif, FN = *False* Negatif, dan FP = *False* Positif. Penilaian inilah yang diimplementasikan untuk mengukur *matrix accuracy, precision, recall*, dan *F1-score*. Pengujian data dipartisi menjadi dua subset yaitu data *testing* dan data *training*. Kombinasi yang dilakukan pada penelitian ini adalah 20% data *testing* dan 80% data *training*. Sehingga diperoleh hasil pengujian SVM pada pada Gambar 4. Hasil evaluasi menggunakan metode SVM diperoleh *accuracy* sebesar 0,8475 atau 84,75%, *precision* 0,9411765 atau 94,11765%, *recall* 0,88888889 atau 88,888889%, dan *F1-score* 0,9142857 atau 91,42857%. Matriks *precision* mengevaluasi seberapa akurat model dalam memprediksi kelas positif. Untuk kelas negatif, *precision*-nya 0, kelas netral 33,33333%, dan kelas positif 94,11765%. *Recall* digunakan untuk mengevaluasi kemampuan model dalam mendeteksi keseluruhan contoh posistif yang ada. Untuk kelas negatif bernilai 0%, kelas netral 66,66667%, dan kelas positif 88,888889%. *F1-score* memadukan nilai *precision* dan *recall* menggunakan perhitungan rata-rata harmonis untuk menghindari bias pada penilaian. Untuk kelas negatif NaN artinya posisi atau observasi pada dataset tidak memiliki nilai atau kosong, kelas netral 44,44444%, dan kelas positif 91,42857%.

Hasil tersebut menunjukkan bahwa model SVM belum mampu mengidentifikasi ulasan dengan sentiment negatif secara optimal. Hal ini disebabkan oleh distribusi data yang tidak seimbang, yaitu jumlah ulasan negatif sebesar 14 data jauh lebih sedikit dibandingkan ulasan positif sebesar 269 data, dan netral sebesar 18 data. Ketidakseimbangan tersebut menyebabkan model lebih banyak mempelajari pola dari

kelas mayoritas sehingga cenderung mengklasifikasikan data ke kelas positif.

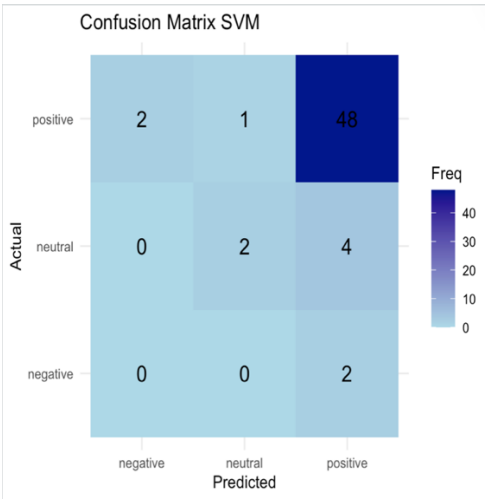
```
> evaluation
                                Class Precision    Recall  F1_Score
Class: negative Class: negative 0.0000000 0.0000000       NaN
Class: neutral   Class: neutral 0.3333333 0.6666667 0.4444444
Class: positive Class: positive 0.9411765 0.8888889 0.9142857

Statistics by Class:

                     Class: negative Class: neutral Class: positive
Sensitivity                  0.0000        0.66667          0.8889
Specificity                  0.9649        0.92857          0.4000
Pos Pred Value               0.0000        0.33333          0.9412
Neg Pred Value               0.9649        0.98113          0.2500
Prevalence                   0.0339        0.05085          0.9153
Detection Rate               0.0000        0.03390          0.8136
Detection Prevalence         0.0339        0.10169          0.8644
Balanced Accuracy            0.4825        0.79762          0.6444

      Overall Statistics

                  Accuracy : 0.8475
                    95% CI : (0.7301, 0.9278)
       No Information Rate : 0.9153
       P-Value [Acc > NIR] : 0.9743

                     Kappa : 0.2468
```

Gambar 5. Hasil Klasifikasi SVM

Sensitivity merupakan kemampuan model mendeteksi data positif, artinya seberapa banyak data benar-benar positif berhasil dikenali sebagai positif. Hasil yang diperoleh adalah kelas negatif 0%, kelas netral 66,667%, dan kelas positif 88,89%. *Specificity* merupakan kemampuan model mendeteksi data negatif, artinya seberapa banyak data negatif yang dikenali dengan benar sebagai negatif. Untuk kelas negatif 96,49%, kelas netral 92,857%, dan kelas positif 40%. *Positive predictive value* adalah seluruh hasil prediksi positif yang selaras dengan label aktualnya. Untuk kelas negatif 0%, untuk kelas netral 33,33%, dan kelas positif 94,12%. *Negative predictive value* merupakan prediksi negatif, artinya dari semua yang diprediksi negatif, berapa yang benar-benar negatif. Untuk kelas negatif 96,49%, kelas netral 98,113%, dan kelas positif 25%. *Prevalence* adalah proporsi data positif dalam dataset, artinya seberapa banyak data positif dibandingkan dengan seluruh data. Untuk kelas negatif 3,39%, kelas netral 98,113%, dan kelas positif 91,53%. *Detection rate* merupakan proporsi data yang benar-benar positif dan berhasil dideteksi artinya seberapa banyak dari selurih datang yang merupakan TP. Untuk kelas negatif 0%, kelas netral 33,90%, dan kelas positif 81,36%. *Detection prevalence* merupakan proporsi data yang diprediksi sebagai positif, artinya berapa banyak data yang diprediksi oleh model (benar atau salah). Untuk kelas negatif 3,39%, kelas netral 10,169%, dan kelas positif 86,44%. *Balanced accuracy* adalah rata-rata kemampuan model pada kelas positif dan negatif, artinya menggabungkan sensitivity dan *specificity* agar adil jika data tidak seimbang. Untuk kelas negatif 48,25%, kelas netral 79,762% dan kelas positif 64,44%.

Penelitian ini tidak seimbang, yaitu terdiri dari 269 sentimen positif, 18 sentimen netral, dan 14 sentimen negatif. Kondisi ini tentu menyebabkan nilai NIR mencapai 91,53%, sehingga model selalu memprediksi kelas mayoritas positif dan dapat memperoleh akurasi yang cukup tinggi. Sementara itu, model SVM pada penelitian ini menghasilkan nilai akurasi sebesar 84,75% yang berada di bawah nilai NIR. Oleh karena itu, ketidakseimbangan dataset memengaruhi kemampuan model dalam mengenali kelas minoritas. Kemudian, evaluasi model tidak hanya didasarkan pada akurasi, melainkan juga mempertimbangkan *precision, recall, F-1 score,* dan *confusion matrix* untuk memberikan gambaran performa model pada setiap kelas.

p-ISSN: 2527 - 9661
e-ISSN: 2549 - 2837



Gambar 6. Visualisasi Hasil *Confusion Matrix*

Pada Gambar 6, model berhasil mengkategorikan 0 data pada kelas negatif dengan akurat, namun memiliki kekeliruan pada 2 data kelas positif. Model berhasil mengkategorikan 2 data pada kelas netral dengan akurat, namun memiliki kekeliruan pada 1 data kelas positif. Model berhasil mengkategorikan 48 data kelas positif, namun memiliki kekeliruan pada 4 data kelas netral dan 2 data kelas negatif.

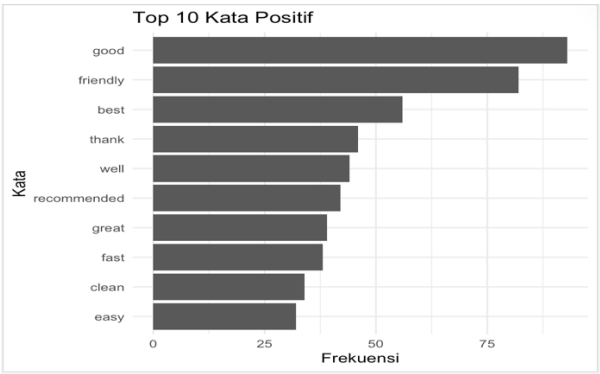
Word cloud adalah visualisasi dari kumpulan kata yang menggambarkan frekuensi kemunculan kata atau dataset. Dalam visualisasi *word cloud*, untuk kata yang sering muncul akan menampilkan gambar yang lebih besar. Berdasarkan hasil klasifikasi sentimen diperoleh frekuensi kata untuk sentimen kelas positif dan negatif. Kemudian, frekuensi tersebut diurutkan berdasarkan kata-kata yang memiliki frekuensi yang sering timbul pada setiap sentimen kelas positif dan negatif divisualisasikan dengan menggunakan *word cloud* yang terlihat pada Gambar 7 dan 8.

Gambar 7. *Word Cloud* Sentimen Positif

Sumber: Peneliti, 2026

Gambar 8. *Word Cloud* Sentimen Negatif

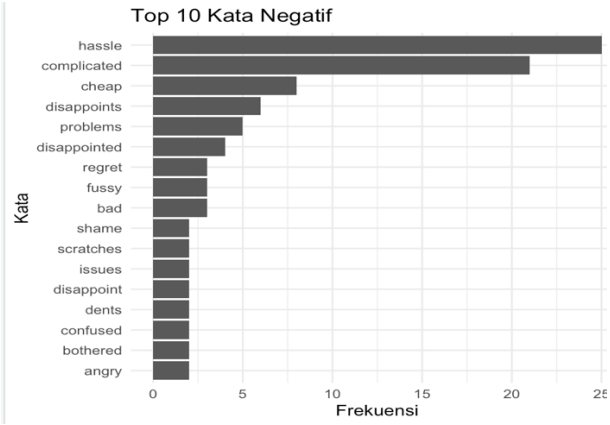
Pada *Word Cloud* sentimen positif diperoleh frekuensi untuk setiap kata terlihat pada Gambar 9. Sedangkan, *word cloud* sentimen kelas negatif diperoleh frekuensi setiap kata dapat dilihat pada Gambar 10.



Gambar 9. Frekuensi Kata Sentimen Positif

Pada Gambar 9 diperoleh 10 kata pada sentimen positif yang sering muncul sebagai berikut: "*good*" 93 kali, "*friendly*" sebanyak 82 kali, "*best*" 56 kali, "*thank*" 46 kali, "*well*" 44 kali, "*recommended*" 42 kali, "*great*" 39 kali, "*fast*" 38 kali, "*clean*" 34

kali, "*easy*" 32 kali. Kemudian pada *word cloud* sentiment negatif diperoleh frekuensi pada Gambar 10.



Gambar 10. Frekuensi Kata Sentimen Negatif

Pada Gambar 10 diperoleh 10 kata sentiment negatif yang sering muncul sebagai berikut: "*hassle*" sebanyak 25 kali, "*complicated*" 21 kali, "*cheap*" 8 kali, "*disappoints*" 6 kali,"*problems*" 5 kali, "*disappointed*" 4 kali, "*bad*" 3 kali, "*fussy*" 3 kali, dan "*regret*" 3 kali. Sehingga dapat dianalisa bahwa:

1. Kata-kata yang muncul di sentimen positif dan negatif tidak ada kemunculan kata yang sama.
2. Kata-kata yang terdapat pada sentimen positif yang paling sering muncul adalah "*good*", artinya konsumen merasa puas terhadap berbagai aspek layanan rental mobil, mulai dari kondisi kendaraan, kualitas pelayanan, sampai dengan pengalaman secara keseluruhan.
3. Kata-kata yang terdapat pada sentimen negatif yang paling sering muncul adalah "*hassle*" biasanya sering digunakan untuk menggambarkan pengalaman yang tidak nyaman, misalnya proses ribet, banyak aturan atau prosedur yang menyulitkan.
4. Pada kata sentimen positif, kata yang terendah di antara 10 kata positif adalah kata "*easy*". Hal ini menggambarkan kemudahan dalam proses penyewaan mobil, mulai dari pemesanan hingga pelayanan.
5. Pada kata sentimen negatif, kata yang terendah muncul pada 10 kata negatif, yaitu "*regret*". Hal ini menggambarkan adanya penyesalan dari pelanggan setelah menggunakan layanan rental mobil artinya pengalaman yang diperoleh tidak sesuai dengan harapan.
6. Pada kamus *Bing Liu Opinion Lexicon*, kata "*cheap*" dikategorikan sebagai sentimen negatif karena dalam bahasa Inggris memiliki makna "murahan" atau "berkualitas rendah". Padahal dalam konteks bisnis rental mobil, kata "*cheap*" harusnya bermakna positif karena menggambarkan harga yang terjangkau. Munculnya kata "*cheap*" pada sentimen kelas negatif merupakan konsekuensi dari penggunaan kamus leksikon yang bersifat umum dan tidak mempertimbangkan konteks domain. Hal ini tentu menjadikan salah satu keterbatasan karena kamus ini belum mampu memahami makna kata berdasarkan konteks domain maupun hubungan antarkata dalam satu kalimat. Oleh karena itu, untuk penelitian selanjutnya penggunakan kamus sentimen perlu disesuaikan dengan domain polaritas kata sehingga diharapkan dapat meningkatkan kualitas pelabelan sentimen.

4. SIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen ulasan Google Maps review Wulan Rent Car menggunakan algoritma SVM maka diperoleh beberapa kesimpulan yaitu: Hasil pelabelan data menunjukkan bahwa sebagian besar ulasan pelanggan Wulan Rent Car memiliki

sentimen positif. Dari total data ulasan yang digunakan, diperoleh 269 ulasan positif, 18 ulasan netral, dan 14 ulasan negatif. Dominasi sentimen positif menunjukkan bahwa sebagian besar pelanggan memberikan pengalaman yang baik terhadap layanan Wulan Rent Car, baik dari aspek pelayanan, kualitas kendaraan, maupun pengalaman penggunaan jasa.

Model SVM menghasilkan akurasi sebesar 84,75% dalam melakukan klasifikasi sentimen. Berdasarkan evaluasi menggunakan *precision*, *recall*, dan *F1-score*, model menunjukkan performa yang baik dalam mengenali kelas positif dengan nilai *F1-score* sebesar 91,43%. Namun, performa model pada kelas negatif masih rendah dengan nilai *precision* dan *recall* sebesar 0%. Hal ini menunjukkan bahwa model mengalami kesulitan dalam mengenali kelas minoritas akibat distribusi data yang tidak seimbang.

Ketidakseimbangan jumlah data antarkelas menjadi salah satu faktor yang memengaruhi performa model. Jumlah ulasan positif yang jauh lebih banyak dibandingkan dengan ulasan negatif dan netral menyebabkan model lebih cenderung mempelajari pola dari kelas mayoritas. Oleh karena itu, nilai akurasi secara keseluruhan perlu diinterpretasikan secara hati-hati dan tidak dapat menjadi satu-satunya indikator keberhasilan model.

Hasil analisis sentimen dapat memberikan informasi tambahan bagi Wulan Rent Car dalam memahami persepsi pelanggan terhadap layanan yang diberikan. Informasi mengenai kecenderungan sentimen positif, netral, maupun negatif dapat dimanfaatkan sebagai bahan evaluasi untuk mempertahankan kualitas layanan, memperbaiki kekurangan, serta meningkatkan strategi pelayanan kepada pelanggan.

Untuk penelitian selanjutnya, diperlukan pendekatan tambahan untuk menangani permasalahan ketidakseimbangan data, seperti penggunaan metode resampling (SMOTE atau *oversampling*), pembobotan kelas (*class weighting*), maupun penggunaan metode klasifikasi lain sebagai pembanding. Pendekatan tersebut diharapkan dapat meningkatkan kemampuan model dalam mengenali kelas minoritas, khususnya sentimen negatif.

DAFTAR PUSTAKA

[1] Saiful Nur Budiman, Sri Lesanti, and Erwan, “Analisis Sentimen Berdasarkan Hasil Review Lokasi Google Map Menggunakan Natural Language Toolkit TextBlob dan Naïve Bayes,” *JAMI: Jurnal Ahli Muda Indonesia*, vol. 5, no. 2, pp. 114–126, Dec. 2024, doi: 10.46510/jami.v5i2.311.

[2] M. Alfiza, . D., N. O. Sipayung, and R. Alhamdi, “Pengaruh Rating dan Review Melalui Google Reviews Terhadap Keputusan Pembelian Konsumen di Arch Alley Batam,” *Jurnal Ekonomi Manajemen dan Bisnis (JEMB)*, vol. 4, no. 2, pp. 949–961, Dec. 2025, doi: 10.47233/jemb.v4i2.3889.

[3] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, “Visualisasi Word Cloud Hasil Analisis Sentimen Berbasis Fitur Layanan Aplikasi Gojek Dengan Support Vector Machine,” *Jurnal Serina Sains, Teknik dan Kedokteran*, vol. 2, no. 1, pp. 61–70, Mar. 2024, doi: 10.24912/jsstk.v2i1.32058.

[4] F. Reza Pradhana, A. Musthafa, and I. Fitria, “Analisis Sentimen Ulasan Produk Daviena Di Shopee Menggunakan Algoritma Support Vector Machine,” Sukoharjo: SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2024, Nov. 2024, pp. 1–13.

[5] P. M. Aryanto and R. Mardhiyyah, “Analisis Sentimen Terhadap Review Google Maps Jogja City Mall Menggunakan Algoritma

Support Vector Machine," *Journal of Computer System and Informatics (JoSYC)*, vol. 6, no. 1, p. 2535, Nov. 2024, doi: 10.47065/josyc.v6i1.6049.

[6] F. Rahmasari, N. Rahaningsih, R. D. Dana, and C. L. Rohmat, "Optimasi Analisis Sentimen Aplikasi Glints Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5681.

[7] I. S. Aisah, B. Irawan, and T. Suprapti, "Algoritma Support Vector Machine (SVM) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, Dec. 2023.

[8] Y. Aliya Rohim, V. Atina, and Sundari, "Analisis Sentimen terhadap Ulasan Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes," *JIIP (Jurnal Ilmiah Ilmu Pendidikan)*, vol. 7, no. 6, Jun. 2024, [Online]. Available: http://Jiip.stkipyapisdompu.ac.id

[9] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 375–384, Feb. 2024, doi: 10.57152/malcom.v4i2.1206.

[10] F. Bravo-Marquez, E. Frank, and B. Pfahringer, "Building a Twitter Opinion Lexicon from Automatically-annotated Tweets," 2016. [Online]. Available: http://sentic.net/

[11] A. Hamzah and R. Yanwastika Ariyana, "Klasifikasi Emosi Berbasis Emolex dari Komentar Evaluasi Akademik Mahasiswa Emolex-Based Classification of Emotions from Academic Evaluation Comments," *Techno.com*, vol. 23, no. 2, pp. 455–466, May 2024.

[12] M. Nurjannah, Hamdani, and I. Fitri Astuti, "Penerapan Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) Untuk Text Mining," *Jurnal Informatika Mulawarman*, vol. 8, no. 3, p. 110, Sep. 2013.

[13] Farras Naufal Majid and Sulastri, "Analisa Sentimen Aplikasi Pedulilindungi dengan Metode NBC dan SVM," *JURNAL ELEKTRONIKA DAN KOMPUTER*, vol. 16, no. 1, pp. 100–108, Jul. 2023, [Online]. Available: https://journal.stekom.ac.id/index.php/elkom■page100

[14] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *SMATIKA Jurnal*, vol. 10, Dec. 2020.

[15] D. Septhya, K. Rahayu, S. Rabbani, V. Fitria, Y. Irawan, and R. Hayami, "Implementation of Decision Tree Algorithm and Support Vector Machine for Lung Cancer Classification," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 1, pp. 15–19, Apr. 2023.