KOMTEKINFO JOURNAL
LPPM Universitas Putra Indonesia YPTK Padang
Jl. Raya Lubuk Begalung, Padang, West Sumatera, Indonesia, Zip Code: 25221
Volume: 12, Issue: 4, Page: 239 - 245, Desember 30th, 2025, e-ISSN: 2502-8758
Available online at website: https://jkomtekinfo.org/ojs/index.php/komtekinfo

Comparison of Random Forest and Support Vector Machine Learning Algorithms in Sentiment Analysis of Gojek User Reviews

Tesa Vausia Sandiva[1]✉, Arip Kristiyanto[2]
[1] Department of Information System, Pamulang University, Serang, 42183, Indonesia
[2] Department of Information System, Pamulang University, Serang, 42183, Indonesia
dosen03220@unpam.ac.id

Abstract

The development of digital technology has brought significant changes across various sectors of life, including transportation. One of the most popular modes of transportation among the public today is online motorcycle taxis, such as Gojek. Gojek continues to innovate to meet customer needs more effectively and to expand its range of services. This study aims to identify the number of positive, neutral, and negative sentiments in a user review dataset, as well as to evaluate the performance of the algorithms used—namely, SVM and Random Forest. The analysis was conducted on 10,000 customer reviews from the Play Store application, resulting in 2,057 positive sentiments, 1,135 neutral sentiments, and 6,295 negative sentiments. The classification model compared the SVM algorithm with the Random Forest algorithm, and the results show that Random Forest achieved better performance, with 91% accuracy compared to SVM's 89%. These findings demonstrate that Random Forest performs better in handling word distribution within review texts than the SVM method.

Keywords: Sentiment analysis, Go-jek, Random-forest, SVM, Machine Learning

1. Introduction

The development of digital technology has brought significant changes across various sectors of life, including transportation. App-based services have become the primary solution for meeting people's mobility needs, offering greater convenience, efficiency, and speed. Gojek, one of the pioneers of online transportation services in Indonesia, has evolved into an integrated digital ecosystem that offers a wide range of services, from transportation and food delivery to digital payments [1][2].

With over 100 million downloads and a 4.6-star rating on the Google Play Store, Gojek has become the preferred choice for users in Indonesia and across the Southeast Asian region [3]. User reviews on digital platforms such as the Google Play Store play a crucial role in reflecting user perceptions, satisfaction levels, and experiences regarding the quality of the services provided. This information can be utilized by developers to make strategic decisions aimed at improving service quality [4].

As the number of users continues to grow, the review feature has become a valuable source of information for evaluating customer satisfaction levels. The large volume of user comments can provide insights into customer perceptions of a particular service and help identify areas for improvement. This can be achieved through sentiment analysis of user comments posted on digital platforms such as the Google Play Store. However, user reviews are often subjective and exhibit complex variations in sentiment, requiring analytical methods capable of classifying opinions into positive, negative, or neutral categories. Sentiment analysis, a subset of natural language processing (NLP), aims to detect and categorize opinions or emotions expressed in text into corresponding sentiment groups [5].

The development of machine learning, as a branch of artificial intelligence (AI), has also advanced rapidly and has been widely adopted across various industrial sectors [6]. Support Vector Machine (SVM) and Random Forest are among the most widely used algorithms in sentiment analysis due to their proven effectiveness in text classification [7][8]. SVM has the capability to separate data into two distinct classes by identifying the optimal hyperplane that maximizes the margin between them [9]. Meanwhile, Random Forest is a decision tree–based algorithm derived from the CART (Classification and Regression Trees) method, in which multiple decision trees are combined to enhance classification accuracy. NLP technology enables large-scale automated analysis of review texts to identify sentiment patterns such as positive, negative, or neutral [10].

Various studies have been conducted on the application of machine learning algorithms in

Submitted: 20-10-2025 | Revised: 25-11-2025 | Accepted: 26-12-2025 | Published: 30-12-2025

sentiment analysis. One study applied the SVM algorithm to analyze application reviews and achieved a high level of accuracy[11][12][13]. This demonstrates the effectiveness of SVM in sentiment classification using online transportation service user data. The Random Forest algorithm was also applied to analyze Gojek review sentiments and produced competitive results [14]. A comparison of several algorithms, including Naive Bayes, SVM, Logistic Regression, and Random Forest, showed that the performance of each algorithm is highly influenced by the characteristics of the dataset [15][16]. This indicates a significant difference in accuracy and efficiency between SVM and Random Forest. However, there is still limited research that directly compares the two algorithms comprehensively, taking into account factors such as preprocessing, data volume, and sentiment distribution in Gojek user reviews. [17].

Based on the above description, this study aims to compare the performance of the Support Vector Machine (SVM) and Random Forest algorithms in the sentiment analysis of Gojek user reviews. It also seeks to examine the influence of various factors, such as text preprocessing, dataset size, and class distribution, on classification results. The findings of this research are expected to make a meaningful contribution to the development of machine learning–based sentiment analysis systems and serve as a valuable reference for the digital services industry in gaining a deeper understanding of user opinions.

2. Methods

This study analyzed user reviews of the Gojek application available on the Google Play Store. The stages of the research are presented in Figure 1.[18][19][20].

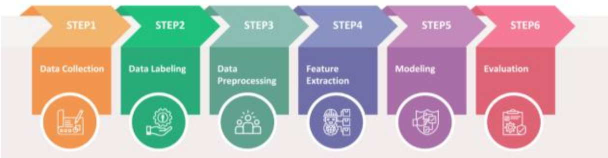

Figure 1. Method

1. This study utilized secondary data in the form of user reviews of the Gojek application obtained from the Google Play Store, specifically collected by the researcher to address the research problem. The data collection process involved scraping 10,000 user comments or reviews from the Play Store using the google-play-scraper library available in the Python programming language.
2. Sentiment analysis employs supervised learning methods, which require a labeled or annotated dataset. The labeling process is essential for enabling supervised learning. In this study, a sentiment column was added with the following criteria: scores of 1–2 represent negative sentiments and are coded as -1, a score of 3 represents a neutral sentiment and is coded as 0, and scores of 4–5 represent positive sentiments and are coded as 1.
3. The next stage is data preprocessing, in which the data is prepared for analysis. Preprocessing is necessary to ensure that the data meets optimal quality standards before being used for analysis. This process consists of several steps, including data cleaning, tokenization, stopword removal, and stemming..
4. At this stage, the Term Frequency–Inverse Document Frequency (TF-IDF) method is used to convert the collection of raw text documents into a numerical representation.
5. At this stage, the prepared text data is classified and modeled. Before the modeling process, the preprocessed data is weighted using the Term Frequency–Inverse Document Frequency (TF-IDF) method. TF-IDF is a weighting technique used in text analysis to measure the importance of a word or phrase within a document. After the weighting process is completed, the modeling process continues with the application of the SVM and Random Forest algorithms.

3. Results and Discussions

3.1 Data collection

In this study, data were obtained from Gojek user reviews on the Play Store using web scraping techniques. A total of 10,000 data entries were collected.

Figure 2. Web Scraping Result Data

3.2 Data Labeling

This labeling process is essential for supervised learning methods. A sentiment column was added with the following criteria: scores of 1–2 represent negative sentiment and are coded as -1, a score of 3 represents neutral sentiment and is coded as 0, and scores of 4–5 represent positive sentiment and are coded as 1.

	content	score	Year	Month	Day	sentiment
3466	Sampai saat ini tidak ada kendala apapun. Mula...	5	2018	9	14	1
7393	Baru coba Go-Food sih, sejauh ini cukup baik m...	4	2018	9	14	1
3760	Untuk masalah promo go pay ketentuannya beruba...	1	2018	9	15	-1
6001	Saya senang menggunakan aplikasi gojek khususn...	2	2018	9	15	-1
2704	Sekadar memberi saran ya,,tarifnya untuk para ...	5	2018	9	15	1

Figure 3. Data labeling results

3.3 Preprocessing

The next stage is data preprocessing, in which the data are prepared for analysis. Preprocessing is essential to ensure that the data meet optimal quality standards before being used for analysis. This process involves several steps, including data cleaning, tokenization, stopword removal, and stemming.

3.3.1 Data Cleaning

This stage involves the removal of unnecessary components to reduce noise in the data. The remaining data consist only of alphabetical characters. The source code for the data cleaning process is shown below.

Menghapus URL dari kolom konten

```
df['content'] = df['content'].str.replace('https\S+', ' ', case=False)
```

Merubah Teks Jadi Lower Case

```
df['content'] = df['content'].str.lower()
```

Menghapus mention

```
df['content'] = df['content'].str.replace('@\S+', ' ', case=False)
```

Menghapus Hashtag

```
df['content'] = df['content'].str.replace('#\S+', ' ', case=False)
```

Figure 4. Source Code data cleaning

Figure 4 shows the Python code used for data cleaning, which includes removing URLs, converting text to lowercase, deleting mentions and hashtags, and removing special characters.

3.3.2 Tokenizing

This stage involves breaking sentences into smaller units or individual words, a process known as tokenization. We use the NLTK (Natural Language Toolkit) library to perform this task. The content column contains the original text (review sentences). The .apply(regexp.tokenize) function applies the tokenization process to each row in the content column. The results of this process a list of individual words are stored in a new column named content_token. Figure 5 displays the tokenization results.

```
from nltk.tokenize import RegexpTokenizer #import package RegexTokenizer
regexp = RegexpTokenizer('\w+')
df['content_token']=df['content'].apply(regexp.tokenize) #menambahkan kolom content_token pada dataframe df
df.head(3)
```

	content	score	Year	Month	Day	sentiment	content_token
3466	sampai saat ini tidak ada kendala apapun. mula...	5	2018	9	14	1	[sampai, saat, ini, tidak, ada, kendala, apapu...
7393	baru coba go-food sih, sejauh ini cukup baik m...	4	2018	9	14	1	[baru, coba, go, food, sih, sejauh, ini, cukup...
3760	untuk masalah promo go pay ketentuannya beruba...	1	2018	9	15	-1	[untuk, masalah, promo, go, pay, ketentuannya,...

Figure 5. Tokenizing Results

3.3.3 Remove stopwords

This stage removes words that do not carry significant meaning and are unnecessary for analysis. The lambda function filters each word (item) in the list, and the condition if item not in stopwords ensures that only words that are not stopwords are retained. Figure 6 shows the results of the stopword removal process.

```
# Remove stopwords
df['content_token'] = df['content_token'].apply(lambda x: [item for item in x if item not in stopwords])
df.head(3)
```

	content	score	Year	Month	Day	sentiment	content_token
3466	sampai saat ini tidak ada kendala apapun. mula...	5	2018	9	14	1	[kendala, apapun, topup, order, gocar, gofood,...
7393	baru coba go-food sih, sejauh ini cukup baik m...	4	2018	9	14	1	[coba, go, food, sih, pesan, estimasi, hargany...
3760	untuk masalah promo go pay ketentuannya beruba...	1	2018	9	15	-1	[promo, go, pay, ketentuannya, berubah, ubah, ...

Figure 6. Stopword Removal

3.3.4 Stemming

Stemming is performed by removing affixes or word endings. The results of the stemming process are shown below. Figure 7 shows that words containing affixes or morphological variations have been reduced to their base or root form. For example, different word forms are standardized into a single base term.

	content	score	Year	Month	Day	sentiment	content_token	stemmed
3466	sampai saat ini tidak ada kendala apapun. mula...	5	2018	9	14	1	[kendala, apapun, topup, order, gocar, gofood,...	[kendala, apa, topup, order, gocar, gofood, go...
7393	baru coba go-food sih, sejauh ini cukup baik m...	4	2018	9	14	1	[coba, go, food, sih, pesan, estimasi, hargany...	[coba, go, food, sih, pesan, estimasi, harga, ...
3760	untuk masalah promo go pay ketentuannya beruba...	1	2018	9	15	-1	[promo, go, pay, ketentuannya, berubah, ubah, ...	[promo, go, pay, tentu, ubah, ubah, cash, back...
6001	saya senang menggunakan aplikasi gojek khususn...	2	2018	9	15	-1	[senang, aplikasi, gojek, gofood, tolonglah, v...	[senang, aplikasi, gojek, gofood, tolong, vouc...
2704	sekadar memberi saran ya,,tarifnya untuk para ...	5	2018	9	15	1	[saran, ya, tarifnya, pengemudi, tolong, diper...	[saran, ya, tarif, kemudi, tolong, perhati, ti...
4799	saya suka dengan promo cashback alfamart, tp k...	2	2018	9	16	-1	[suka, promo, cashback, alfamart, tp, bbrp, al...	[suka, promo, cashback, alfamart, tp, bbrp, al...
9360	aplikasi sih ok2 saja tapi tolong untuk go sen...	5	2018	9	17	1	[aplikasi, sih, ok2, tolong, go, send, drivern...	[aplikasi, sih, ok2, tolong, go, send, drivern...
9112	dear hasil karya anak bangsa, udah sebulan leb...	1	2018	9	18	-1	[dear, hasil, karya, anak, bangsa, udah, sebul...	[dear, hasil, karya, anak, bangsa, udah, bulan...
8190	selamat siang gojek, saya mau memberitahuan ke...	3	2018	9	19	0	[selamat, siang, gojek, memberitahuan, keluhan...	[selamat, siang, gojek, memberitahuan, keluh, ...
8036	seharusnya aplikasi ini jadi sebuah solusi unt...	3	2018	9	20	0	[aplikasi, solusi, transportasi, yg, kekuranga...	[aplikasi, solusi, transportasi, yg, kurang, o...

Figure 7. Stemming Result

3.4 Feature Extraction

The Term Frequency–Inverse Document Frequency (TF-IDF) method is used to convert a collection of raw text documents into numerical representations. The results of the TF-IDF process are shown below.

```
    appversion   at  content  repliedat  replycontent  reviewcreatedversion  \
0          0.0  0.0      0.0        0.0           0.0                   0.0
1          0.0  0.0      0.0        0.0           0.0                   0.0
2          0.0  0.0      0.0        0.0           0.0                   0.0
3          0.0  0.0      1.0        0.0           0.0                   0.0
4          0.0  0.0      0.0        0.0           0.0                   0.0
5          0.0  0.0      0.0        0.0           0.0                   0.0
6          0.0  0.0      0.0        0.0           0.0                   1.0
7          0.0  1.0      0.0        0.0           0.0                   0.0
8          0.0  0.0      0.0        0.0           1.0                   0.0
9          0.0  0.0      0.0        1.0           0.0                   0.0
10         1.0  0.0      0.0        0.0           0.0                   0.0

    reviewid  score  thumbsupcount  userimage  username
0        1.0    0.0            0.0        0.0       0.0
1        0.0    0.0            0.0        0.0       1.0
2        0.0    0.0            0.0        1.0       0.0
3        0.0    0.0            0.0        0.0       0.0
4        0.0    1.0            0.0        0.0       0.0
5        0.0    0.0            1.0        0.0       0.0
6        0.0    0.0            0.0        0.0       0.0
7        0.0    0.0            0.0        0.0       0.0
8        0.0    0.0            0.0        0.0       0.0
9        0.0    0.0            0.0        0.0       0.0
10       0.0    0.0            0.0        0.0       0.0
```

Figure 8. TF-IDF Result

3.5 Modeling

After splitting the dataset into training and testing data, with 7,500 of the 10,000 entries used for training and

2,500 for testing, the modeling phase was carried out using the SVM and Random Forest algorithms.

Figure 9 presents the code used to create and train a Random Forest Classifier model on the training data (X_train, y_train). The output of the RandomForestClassifier() confirms that the model has been successfully created.

```
random_forest_classifier = RandomForestClassifier()
random_forest_classifier.fit(X_train,y_train)

RandomForestClassifier
RandomForestClassifier()
```

Figure 9. Modeling Random Forest

Figure 10 shows the code used to create and train an SVM classification model using the training data (X_train, y_train). The SVC() model uses the RBF kernel by default.

```
svm = SVC()
svm.fit(X_train,y_train)

SVC
SVC()
```

Figure 10. Modeling SVM

3.6 Evaluation

The following are the results of the evaluation of the Random Forest and SVM models using the confusion matrix and the metrics of accuracy, precision, recall, and F1-score.

Figure 11 presents the evaluation results of the Random Forest Classifier algorithm, demonstrating excellent performance with an accuracy of 91%. This indicates that the training process and model parameters are performing optimally on the dataset.

	precision	recall	f1-score	support
-1	0.85	0.92	0.88	1593
0	0.97	0.93	0.95	1525
1	0.91	0.87	0.89	1604
accuracy			0.91	4722
macro avg	0.91	0.91	0.91	4722
weighted avg	0.91	0.91	0.91	4722

Figure 11. Random Forest model evaluation

Figure 12 presents the evaluation results of the SVM algorithm, showing an accuracy of 89%.

	precision	recall	f1-score	support
-1	0.78	0.96	0.86	1593
0	0.99	0.88	0.93	1525
1	0.96	0.84	0.89	1604
accuracy			0.89	4722
macro avg	0.91	0.89	0.90	4722
weighted avg	0.91	0.89	0.90	4722

Figure 12. Evaluation of SVM model

From the results above, the Random Forest model achieved an accuracy of 0.91 and an F1-score of 0.91, indicating that it performs better in handling the given data distribution. The SVM model achieved an accuracy of 0.89 and an F1-score of 0.90, which is relatively close to the performance of Random Forest.

Figure 13. presents the confusion matrix of the Random Forest model, indicating robust classification performance. The high values along the diagonal demonstrate that most instances were correctly classified. The model effectively distinguished the data into three classes, with classification counts of 1,465, 1,416, and 1,403, respectively.

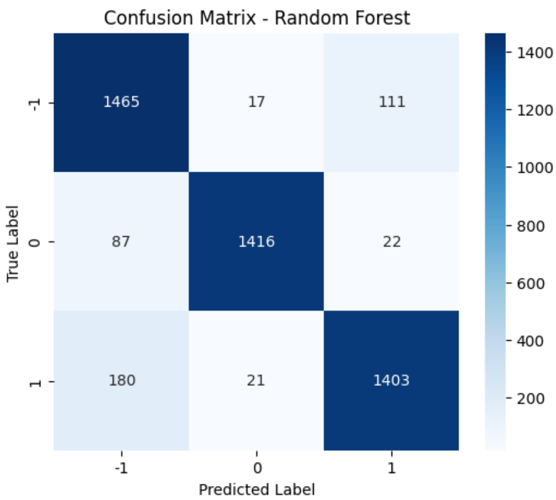

Figure 13. Confusion matrix random forest

Figure 14 presents the confusion matrix of the SVM model, where the main diagonal values (1,536; 1,342; 1,345) represent the number of correctly classified instances. The results indicate that the SVM model performs effectively, particularly in accurately distinguishing class 0.

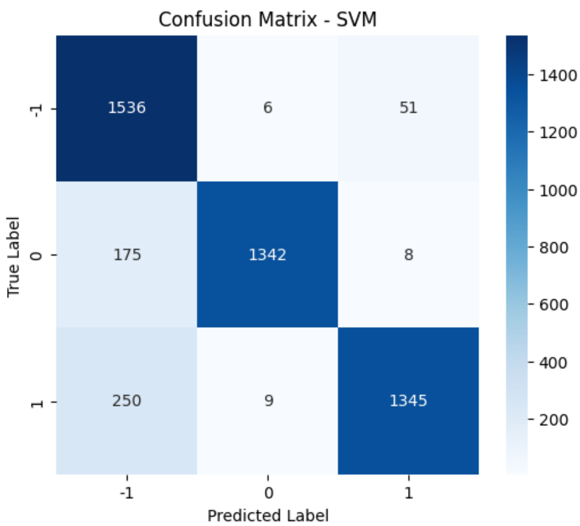

Figure 14. Confusion matrix SVM

Analysis and comparison of the performance of the SVM vs Random Forest models based on the Confusion Matrix above can be seen in Table 1.

Table 1. SVM vs Random Forest based on Confusion Matrix

Class	SVM Model	Random Forest Model	Conclusion
-1	1536 correct, 57 incorrect.	1465 correct, 128 incorrect	SVM is slightly better for class -1
0	1342 correct, 175 incorrect	1416 correct, 87 incorrect	Random Forest is better for class 0
1	1345 correct, 250 incorrect	1403 correct, 180 incorrect	Random Forest is also better for class 1

The confusion matrix results show that the Random Forest model is able to identify classes “0” (neutral) and “1” (positive) more accurately than the SVM model. Although SVM performs slightly better in recognizing class “-1” (negative), it produces a higher number of misclassifications between classes compared to Random Forest.

3.7 Visualisasi

This visualization stage utilizes the Word Cloud library. Word Cloud is a Python library used to visualize text-based data by displaying the most frequently occurring words in a document. In this study, the data visualization is divided into three categories based on sentiment labels: positive, neutral, and negative.

Figure 15 displays the word cloud for positive sentiments. Prominent words such as *“gojek”, “aplikasi”, “driver”, “bantu”, “tolong”, “pesan”, “cepat”, “layanan”* indicate the terms most frequently used by users.

Figure 15. Word cloud sentimen positif

Figure 15 displays the word cloud for netral sentiments as “driver”, “aplikasi”, “pesan”, “gojek”, “harga”, “bayar”, “tolong”, “pakai”, “gak”, “order”.

Figure 16. Word cloud sentimen netral

Figure 17 shows the dominant word is “aplikasi”, “driver”, “pesan”, “gak”, “udah”, “kecewa”, “tolong”, “bayar”, “order”, “masuk”, “akun”, “cancel”. Word “kecewa”, “gak”, “salah”, “susah”, “rugi”, “blokir” illustrates the word cloud for negative sentiments, which indicates user dissatisfaction or issues encountered during the use of the application.

Figure 17. Word cloud sentimen negatif

4. Conclusions

After implementation and testing, several important findings can be concluded from this study. The classification of sentiment reviews on the Gojek application in the Play Store shows that the Random Forest method has better performance than SVM with 91% accuracy compared to SVM with 89% accuracy. These results prove that Random Forest is superior in handling word distribution in review text compared to the SVM method, especially in classifying positive, neutral and negative sentiment. Based on the results of the sentiment analysis, it is clear that Gojek user sentiment tends to be negative. with 6295 negative sentiments, 2570 positive sentiments and 1135 neutral sentiments.

The results of the confusion matrix indicate that the Random Forest algorithm is more accurate in identifying classes “0” (neutral) and “1” (positive) compared to SVM. Although SVM performs slightly better in recognizing class “-1” (negative), it produces a higher number of misclassifications between classes than Random Forest.

Future work on this research could involve applying more rigorous data filtering techniques and exploring or comparing additional algorithms.

References

[1] Annatasya, M. H. Hasmira, Difo Hanggoro, AB Sarca Putera, Ayu Adriyani, And Nasywa Salsabila Kamal, "Perkembangan Teknologi Informasi Menciptakan Inovasi Di Bidang Transportasi Online: Ojek Online," *Social Empirical*, Vol. 1, No. 2, Pp. 154–160, Dec. 2024, Doi: 10.24036/Scemp.V1i2.39.

[2] A. Karim, A. Asrianto, M. Ruslan, And M. Said, "Gojek Accelerate Economic Recovery Through The Digitalization Of Msmes In Makassar," *The Winners*, Vol. 24, No. 1, Pp. 23–31, 2023, Doi: 10.21512/Tw.V24i1.9388.

[3] D. Ginna, B. Mogontha, A. Aliabbas, And S. H. Rancamaya, "Analisa Perbandingan Ekspansi Gojek Ke Empat Negara (Filipina, Singapura, Thailand Dan Vietnam)".

[4] S. Sabouri, "Impact Of Ride-Hailing Services On Travel Behavior," *Paper Knowledge . Toward A Media History Of Documents*, Vol. 3, No. 2, 2021.

[5] S. A. Mahira, I. Sukoco, C. S. Barkah, N. Jamil, A. Novel, And J. A. Bisnis, "Teknologi Artificial Intelligence Dalam Analisis Sentimen: Studi Literatur Pada Perusahaan Kata.Ai," 2023.

[6] Tesa Vausia Sandiva, S. Defit, And G. W. Nurcahyo, "Implementasi Algoritma C4.5 Untuk Prediksi Penerima Beasiswa Program Indonesia Pintar," *Jurnal Komtekinfo*, Pp. 354–362, Sep. 2024, Doi: 10.35134/Komtekinfo.V11i4.582.

[7] R. Wahyudi And G. Kusumawardana, "Analisis Sentimen Pada Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine," *Jurnal Informatika*, Vol. 8, No. 2, Pp. 200–207, 2021, Doi: 10.31294/Ji.V8i2.9681.

[8] A. Kristiyanto, "Journal Of Information System Analisis Data Tiktok Untuk Menilai Pengaruh Terhadap Kebiasaan Pengguna Dengan Support Vector Machine," Vol. 1, No. 1, Pp. 14–20, 2025.

[9] I. S. K. Idris, Y. A. Mustofa, And I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Mengunakan Algoritma Support Vector Machine (SVM)," *Jambura Journal Of Electrical And Electronics Engineering*, Vol. 5, No. 1, Pp. 32–35, 2023, Doi: 10.37905/Jjeee.V5i1.16830.

[10] S. N. Almuayqil, M. Humayun, N. Z. Jhanjhi, M. F. Almufareh, And D. Javed, "Framework For Improved Sentiment Analysis Via Random Minority Oversampling For User Tweet Review Classification," *Electronics (Switzerland)*, Vol. 11, No. 19, Pp. 1–17, 2022, Doi: 10.3390/Electronics11193058.

[11] D. Nugraha And D. Gustian, "Analisis Sentimen Penggunaan Aplikasi Transportasi Online Pada Ulasan Google Play Store Menggunakan Algoritma Svm (Support Vector Machine)," *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, Vol. 1, No. 1, Pp. 99–106, 2023.

[12] M. N. Muttaqin And I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine Dan K Nearest Neighbor," *UNNES Journal Of Mathematics*, Vol. 10, No. 2, Pp. 22–27, 2021, Doi: 10.15294/Ujm.V10i2.48474.

[13] E. Suryati, Styawati, And A. A. Aldino, "Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM)," *Jurnal Teknologi Dan Sistem Informasi*, Vol. 4, No. 1, Pp. 96–106, 2023, Doi: Https://Doi.Org/10.33365/Jtsi.V4i1.2445.

[14] N. Z. Rania And R. D. Syah, "Analisis Sentimen Terhadap Aplikasi Gojek Pada Play Store Menggunakan Metode Random Forest Classifier," *Jurnal Ilmiah Informatika Komputer*, Vol. 29, No. 2, Pp. 144–153, 2024, Doi: 10.35760/Ik.2024.V29i2.11877.

[15] O. S. D. Fadhillah, J. H. Jaman, And C. Carudin, "Perbandingan Naive Bayes, Support Vector Machine, Logistic Regression Dan Random Forest Dalam Menganalisis Sentimen Mengenai Tiktokshop," *Jurnal Informatika Dan Teknik Elektro Terapan*, Vol. 13, No. 1, Pp. 840–847, Jan. 2025, Doi: 10.23960/Jitet.V13i1.5746.

[16] A. Pinkan Mareta *Et Al.*, "Perbandingan Metode Naïve Bayes, Decision Tree, Dan Knn Dalam Analisis Sentimen Aplikasi Gojek Di Playstore," *Jurnal Sistem Informasi*, Vol. 7, No. 2, Pp. 725–734, 2025, Doi: 10.31849/Zjf8x279.

[17] T. Ariaji, S. N. Wahyuni, And M. Ikhsan, "Perbandingan Algoritma SVM Dan RF Pada Analisis Sentimen Menggunakan Pendekatan Machine Learning," *Journal Automation Computer Information System*, Vol. 5, No. 1, Pp. 80–92, May 2025, Doi: 10.47134/Jacis.V5i1.107.

[18] I. Aida Sapitri And M. Fikry, "Pengklasifikasian Sentimen Ulasan Aplikasi Whatsapp Pada Google Play Store Menggunakan Support Vector Machine," *Jurnal TEKINKOM*, Vol. 6, No. 1, Pp. 1–7, 2023, Doi: Https://Doi.Org/10.37600/Tekinkom.V6i1.773.

[19] K. D. Indarwati And H. Februariyanti, "Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier," *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, Vol. 10, No. 1, 2023, Doi: 10.35957/Jatisi.V10i1.2643.

[20] N. Sidauruk, N. Riza, And Rd. N. Siti Fatonah, "Penggunaan Metode Svm Dan Random Forest Untuk Analisis Sentimen Ulasan Pengguna Terhadap Kai Access Di Google Playstore," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 7, No. 3, Pp. 1901–1906, Nov. 2023, Doi: 10.36040/Jati.V7i3.6899.

Biographies of Authors (10 PT)

	. **Tesa Vausia Sandiva** He works at Pamulang University, Serang, as a lecturer and researcher. He was born on March 5, 2001, in South Solok. He pursued a degree in Computer Science, completing both his undergraduate and master's studies at Universitas Putra Indonesia YPTK Padang. The university awarded him the titles of Bachelor of Computer Science (S.Kom.) and Master of Computer Science (M.Kom.) in the field of Computer Science. He can be contacted via email at **dosen03220@unpam.ac.id**. His office is located at Jl. Raya Jakarta Km 5 No.6, Kalodran, Walantaka District, Serang City, Banten 42183, Indonesia. His areas of specialization include Machine Learning and Data Processing.
	**Arip** Kristiyanto is a dedicated lecturer in the Computer System Departement in Faculty Universitas Pamulang , Serang, Indonesia. He earned his Bachelor's Degree from Universitas Pamulang in the Informatics Engineering program in Faculty of Technics. Then he pursued a Master's Degree at STMIK Eresha, specializing in Informatics Engineering. unique identifier, Scopus ID, is 59515683600. Her research spans a wide range of areas, including the Internet of Things, machine learning, data mining, and fuzzy logic. Arip Kristiyanto welcomes communication and collaboration, and he can be reached via email at dosen10027@unpam.ac.id.