International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Sentiment Analysis of Lalamove Application Users Against Google Playstore Reviews Using Support Vector Machine Algorithm with Phyton Programming Language

Putri Aprilyani[1], Cut Maisyarah Karyati [2]
[1, 2]Business Information System Technology and Engineering Faculty, Computer Science and Information Technology Faculty
University of Gunadarma, Jakarta, Indonesia Margonda Raya Street Number 100, Depok, East Java, 16424

Abstract**—***Lalamove app is an online goods delivery service platform available in Indonesia and can be downloaded on Google Play Store. Every new user who uses the Lalamove app is likely to want to know reviews from previous users on the Google Play Store. Sentiment analysis is needed to find out the assessment of the Lalamove application, whether the reviews tend to be positive or negative. The stages of the sentiment analysis process are divided into several stages, namely data retrieval, data labeling, preprocessing, word weighting, classification process and data visualization. The preprocessing stage is divided into several stages, namely case folding, data cleaning, tokenizing, stopword removal and normalization. The next process is to divide the data into 80% training data and 20% testing data and then calculate TF-IDF for further classification using the Support Vector Machine method. The accuracy obtained is 91.6% with a recall of 97.7%, precision of 91.7%, f1-score of 94.6%, and there is an error rate of 0.9 or 9%.*

Keywords**—** *Confusion Matrix, Google PlayStore, Machine Learning, Sentiment Analysis, Support Vector Machine, Web Scapping,*

I. INTRODUCTION

Indonesia's technology that continues to develop rapidly, the use of smartphones is increasing and popular. Make people do the activity of buying goods online by downloading on the PlayStore. In Indonesia there are many slicing services that cannot be separated from the various problems faced by shipping service companies [1]. The main factor affecting the shortage of delivery service services on smartphone users is the opinion of customers on the PlayStore reviews. Reviews can be accessed openly, making it easier for people to compare delivery services to be used [2].

Using text analysis to collect data from the web and social media sites is known as sentiment analysis. Sentiment analysis is used to analyze public opinions about the rise of freight forwarding services in Indonesia, there are many reviews of freight forwarding service applications on the Google PlayStore. This review aims to provide a brief review of the application we want to use and see the shortcomings or advantages of the application. Currently, there are more and more one million reviews on the Lalamove application (28/04/2024 14:44 WIB). Of the many reviews given, if you look at the reviews one by one manually, it takes a long time and is also ineffective. Therefore, sentiment analysis is needed to process the opinion data from the review into positive and negative sentiments. By using sentiment analysis, it is expected that companies and developers can improve and improve the quality of the application [3].

In the increasingly competitive delivery business, Lalamove's delivery service continues to innovate and grow. Application is a 24/7 on-demand delivery application that provides an effective, affordable, and reliable delivery solution [4]. Providing services such as moving goods transport services, pickup rental, blindvan rental, truck rental, box car rental, and instant delivery of goods. Lalamove's choice of payment methods is complete with credit/debit card, Bank (ATM), Ovo,and LinkAja payment types. To improve the performance of the company, customer satisfaction is a very important aspect. Companies need to focus on what customers think is important to achieve customer satisfaction. Therefore, the company should conduct an assessment to determine what factors affect customer assessment and how to meet customer expectations.

In a review on social media platforms, in Wasim and Hassan Mourad's research (2022) with sentiment analysis of the iPhone release using a support vector machine to determine user responses on twitter related to the release of the latest iPhone and based on the results of accuracy classification which can be 89.21%, precision 92.43%, Recall 95.53%, and F1 Score 93.95% [5] Then research by Mahmud Isana, Gregorius Nathanael, and Bens Pardamean (2023) with TikTok application user sentiment analysis using the vader and SVM methods based on classification evaluation results with the SVM algorithm resulted in positive precision values of 65.6%, neutral 24.7%, and negative 9.7% [6]. Furthermore, e-Commerce Product Review using Support Vector Machine (SVM) Based TF-IDF by Siti Fidyanti Nurfadila and Riyanto Jayadi the results of the analysis resulted in an accuracy of 98.9% [7]. Followed by research by Irma Surya and Irvan Abraham (2023) with sentiment analysis of the shoppe application using the support vector machine algorithm, shoppe application review data was classified into positive and negative comments about the application and resulted in an accuracy of 98% [8]. Finally, sentiment analysis on ovo payments using the Support Vector Machine (2022) concluded that the accuracy value obtained from the application review was 76.50% accuracy SVM and an increase of 6.25% by using SVM with PSO (Particle Swarm Optimization) of 82.75% [9].

Putri Aprilyani and Cut Maisyarah Karyati, "Sentiment analysis of Lalamove application users against Google Playstore reviews using Support Vector Machine algorithm with Python programming language" *International Research Journal of Advanced Engineering and Science*, Volume 10, Issue 3, pp. 35-39, 2025.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Based on the description above, the user sentiment analysis of freight forwarding application (Lalamove) was conducted on PlayStore reviews using Support Vector Machine algorithm with Python Programming Language. This research aims to benefit Lalamove company and make Lalamove the best and most used delivery service application in Indonesia.

II. METHOD

A. Collect Data

This study uses web scraping techniques for data retrieval by collecting Lalamove application user review data on Google Playstore using Jupyter Notebook with Python programming language. Here is a Lalamove app review link https://play.google.com/store/search?q=lalamove&c=apps at the initial stage, data retrieval is done by scraping using the help of libraries in the Python programming language, namely the google_play_scraper library with the app module

B. Data Labeling

Data labeling stage against the dataset already retrieved using google_play_scraper. Labeling of positive and negative data is done automatically on 1194 data based on stars given on Lalamove app reviews on Google PlayStore. On the labeling is divided into two namely, positive data labeling is based on the number of more than 3 stars and negative labeling is based on the number of smaller than 3 stars.

C. Preprocessing

The preprocessing stage is carried out to select data that has value by selecting previous data that is processed into a machine learning model so as not to affect the analysis results and get better accuracy results. Preprocessing stages are case folding, data cleansing, tokenization, stopword generation, and normalization [10].

D. Data Training dan Data TestingPreprocessing

Data training is data used to learn an algorithm, while data testing is data used to determine the performance of the algorithm that has been done when finding new data that has never been seen before. In this analysis data training and data testing is divided into 80% data training and 20% data testing [11].

E. Weighting of Word (TF-IDF)

Weighting of word technique in this analysis is TF-IDF (Term Frequency – inverse Document Frequency). Before performing the classification using SVM each token on each data to be classified must be converted into vector form. The conversion of tokens into numerical vector form is called word weighting. The TF-IDF equation is calculated from the TF Value multiplied by the IDF value, then the TF-IDF equation as follows:

$$tfidftd = tftd \times idftd \quad (1)$$

$tfidftd$: Term Weights
$tftd$: Term Frequency word (t) on document (d)
$idftd$: Invers Document Frequency word (t)

Term Frequency is the weight of a word (t) as the number of occurrences in a document(d). Term Frequency ignores the order in which words appear, and most importantly the number of occurrences of each word(t) in the document(d). The inverse document frequency is designed to mitigate the impact of words that are too frequent in a document. The result of word weighting using tfidf technique is a term document matrix with dimensions [12].

Examples of word weighting using TFIDF technique will be done using training data that can be seen in the table 3.11

TABLE 1. Data training examples for TFIDF organizations

Document	Review
d1	The existence of Lalamove is very helpful with Lalamove's delivery, which remains successful and successful
d2	It's really hard to find a driver, but when you do find one, they ask for more money

F. Support Vector Machine Classification

At this stage using Document Term Matrix from training data and train_Y labeling with train_Y labeling for 0 equals negative and 1 equals positive. SVM requires completion of optimization to obtain a linear hyperland. Optimization in this study using modules from the library sklearn.svm the svm module [12]. The equation used to obtain a linear hyperland is as follows:

$$\min_{W,b,\varepsilon} \frac{1}{2} W^T W + C \sum_{i=1}^{Ntrain} \varepsilon_i \quad (1)$$

$$s.t.\, Y_{train(i)}\left(W^T \theta\left(dtm_{train(i)}\right) + b\right) \geq 1 - \varepsilon_i \quad (2)$$

$$\varepsilon_i \geq 0$$

Classification process using SVM optimization is done using SVM then get the value of accuracy and tested using Confusion matrix algorithm Confusion Matrix is a table with 4 different combinations of predictive value and actual value [13]. There are four terms that represent the results of the classification process in the confusion matrix, namely True Positive, True negative, False positive, and False negative. Here is a figure of the confusion matrix 4 combination of prediction and actual that can be seen in the figure 1.

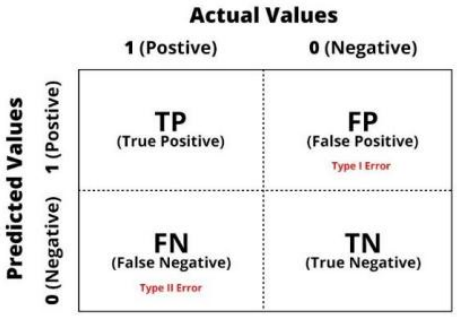

Figure 1. Combination of Confusion Matrix 4

The terms in the confusion matrix:
- True Positive: Correctly predicted positive data
- False Positive: Incorrectly predicted positive data
- False Negative: Negative Data predicted as positive data
- True Negative: positive data predicted as negative data

III. RESULT AND DISCUSSION

A. Data Collection Result

The Data taken using the google_play_scraper library is 796 review data. Then the scraper data is inserted into

Putri Aprilyani and Cut Maisyarah Karyati, "Sentiment analysis of Lalamove application users against Google Playstore reviews using Support Vector Machine algorithm with Python programming language" *International Research Journal of Advanced Engineering and Science*, Volume 10, Issue 3, pp. 35-39, 2025.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Microsoft Excel automatically with the format file Comma Separed Value (.csv).

Table 2. Result of Collect Data Scrapper

Content	Score
All the features are great, and the payment is cheaper than other apps. The only thing I don't like is that I can't top up my wallet with the amount I want; I can only set it in the app. I also can't transfer funds from my wallet. This is very frustrating! PLEASE fix this in the wallet menu.	3
What is this, the wallet must contain a minimum of 1 million, people only move once, so they don't use the extra money, but if we press the order we can't cancel it and it's already in a hurry to enter the driver's cellphone. It's not nice to cancel it.	1

B. Data Labeling Result

The labeling of positive data is based on the number of more than 3 stars and the labeling of negative data is based on the number of less than 3 stars. Data labeling is given a number 1 for positive data and a number 0 for negative data. From 1194 data there were 337 positive data and 857 negative data.

Table 3. Result of Data Labeling

Content	Score	Sentiment
All the features are great, and the payment is cheaper than other apps. The only thing I don't like is that I can't top up my wallet with the amount I want; I can only set it in the app. I also can't transfer funds from my wallet. This is very frustrating! PLEASE fix this in the wallet menu.	3	0
What is this, the wallet must contain a minimum of 1 million, people only move once, so they don't use the extra money, but if we press the order we can't cancel it and it's already rushing into the driver's cellphone. It's not nice to cancel it.	1	0

C. Preprocessing Result

Preprocessing is a stage of data processing before the data is processed into a machine learning model. The preprocessing stage is divided into several stages, namely case folding, cleaning data, tokenizing, making stopwords, and normalization to produce data that is ready to be classified using a Support Vector Machine.

Table 4. Result of Case Folding

Ulasan	Case Folding
All the features are great, and the payment is cheaper than other apps. The only thing I don't like is that I can't top up my wallet with the amount I want; I can only set it in the app. I also can't transfer funds from my wallet. This is very frustrating! PLEASE fix this in the wallet menu.	All the features are good and the payment is also cheaper than other apps. The only thing I don't like is that I can't top up my wallet with the nominal amount I want, it can only be determined in the app. I also can't transfer the balance in the app's wallet. It's very detrimental!!! Please fix this in the wallet menu.

Table 5. Result of Cleaning Data

Hasil Case Folding	Cleaning Data
All the features are good and the payment is also cheaper than other apps. The only thing I don't like is that I can't top up my wallet with the nominal amount I want, it can only be determined in the app. I also can't transfer the balance in the app's wallet. It's very detrimental!!! Please fix this in the wallet menu.	All the features are good and the payment is also cheaper than other apps, the only thing I don't like is that when I top up my wallet, the nominal amount cannot be according to our needs, it can only be determined in the app and I can't transfer the balance in my wallet. The app is very detrimental. Please fix it in the wallet menu

Table 6. Result of Tokenizing

Hasil Cleaning Data	Tokenizing
All the features are good and the payment is also cheaper than other apps, the only thing I don't like is that when I top up my wallet, the nominal amount cannot be according to our needs, it can only be determined in the app and I can't transfer the balance in my wallet. The app is very detrimental. Please fix it in the wallet menu	['all', 'features', 'ok', 'and', 'payment', 'also', 'more', 'cheap', 'from', 'apps', 'other', 'only', 'which', 'don't', 'like', 'that', 'my', 'top', 'up', 'wallet', 'nominal', 'don't', 'can', 'according to', 'needs', 'which', 'we', 'want', 'only', 'can', 'determined', 'in', 'app', 'and', 'no', 'can', 'transfer', 'balance', 'which', 'there', 'in', 'wallet', 'app', 'very', 'disadvantageous', 'please', 'so that', 'can', 'fixed', 'on', 'menu', 'wallet']

Table 7. Result of Stopword

Tokenizing Results	Removal of Stopwords
['all', 'features', 'ok', 'and', 'payment', 'also', 'more', 'cheap', 'from', 'apps', 'other', 'only', 'which', 'don't', 'like', 'that', 'my', 'top', 'up', 'wallet', 'nominal', 'don't', 'can', 'according to', 'needs', 'which', 'we', 'want', 'only', 'can', 'determined', 'in', 'app', 'and', 'no', 'can', 'transfer', 'balance', 'which', 'there', 'in', 'wallet', 'app', 'very', 'disadvantageous', 'please', 'so that', 'can', 'fixed', 'on', 'menu', 'wallet']	['features', 'ok', 'payment', 'also', 'cheap', 'apps', 'likes', 'sy', 'top', 'up', 'wallet', 'nominal', 'according to', 'needs', 'determined', 'app', 'tf', 'balance', 'wallet', 'app', 'loss', 'please', 'fix', 'menu', 'wallet']

Table 8. Result of Normalisasi

Stopword Removal Results	Normalization
['features', 'ok', 'payment', 'also', 'cheap', 'apps', 'likes', 'sy', 'top', 'up', 'wallet', 'nominal', 'according to', 'needs', 'determined', 'app', 'tf', 'balance', 'wallet', 'app', 'loss', 'please', 'fix', 'menu', 'wallet']	['features', 'ok', 'payment', 'also', 'cheap', 'apps', 'like', 'me', 'top', 'up', 'wallet', 'nominal', 'according to', 'needs', 'determined', 'app', 'tf', 'balance', 'wallet', 'app', 'loss', 'please', 'fix', 'menu', 'wallet']

D. Training and Testing Data Result

The results of training data and testing data are divided into 80% training data and 20% testing data. Training Data were used as many as 955 data and testing data as many as 239 data.

G. Weighting of Word (TF-IDF)

The final result of the TF-IDF calculation will produce a document term matrix (dtm) on the calculation for training data and testing data. Formation of document term matrix using scikit-lern library function is tfidfvectorizer function. The resulting Output from the calculation of TF-IDF can be seen in figure 2 and 3.

```
(0, 4983)     0.17187117298588597
(0, 4876)     0.12607494428617488
(0, 4875)     0.05420267254440694
(0, 4683)     0.05065865825446848
(0, 4663)     0.10642842319986415
(0, 4482)     0.14808072095802957
(0, 4481)     0.2280740450349787
(0, 4273)     0.08709734991071
(0, 4262)     0.11393372825958709
(0, 4233)     0.12031252245258595
(0, 4231)     0.07228143050142165
(0, 4134)     0.13995904353889668
(0, 4044)     0.14808072095802957
(0, 3972)     0.10822115606422154
(0, 3708)     0.08063203065384209
(0, 3655)     0.14808072095802957
(0, 3499)     0.08910207374301106
(0, 3283)     0.09254432394714233
```

Figure 2. Document Term Matrix Data Training

Putri Aprilyani and Cut Maisyarah Karyati, "Sentiment analysis of Lalamove application users against Google Playstore reviews using Support Vector Machine algorithm with Python programming language" *International Research Journal of Advanced Engineering and Science*, Volume 10, Issue 3, pp. 35-39, 2025.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

```
(0, 4875)     0.11635366945324949
(0, 4576)     0.1869667266583975
(0, 4530)     0.22107432287925352
(0, 4415)     0.20440105093562763
(0, 3997)     0.10648235085799682
(0, 3981)     0.31787612030815526
(0, 3779)     0.19733629340490977
(0, 3616)     0.17946982142653078
(0, 3185)     0.10228405930030335
(0, 2946)     0.20440105093562763
(0, 2343)     0.2408334476129098
(0, 2304)     0.10019296650508852
(0, 2108)     0.16818859506694234
(0, 1997)     0.15097534240435084
(0, 1899)     0.1960526657146847
(0, 1878)     0.23739174450775458
(0, 1820)     0.13772794498689442
(0, 1321)     0.07771433122829895
```

Figure 3. Document Term Matrix Data Testing

E. *Support Vector Machine Clasiffication Result*

After the data is trained and get the classification model using the Support Vector Machine, the next step is to test the testing data with the model that has been made previously with the Support Vector Machine. In this study the testing data used for 239 data. The test results show accuracy with Support Vector Machine of 91.6%. Data accuracy testing is done by calculating the truth of the sentiment of the actual results with the sentiment of the predicted results. At 239 data will be tested with the algorithm confusion matrix.

Table 9. Confusion Matrix

Confussion Matrix		Hasil Aktual	
		0 (Negatif)	1 (Positif)
Hasil Prediksi	1 (Positif)	177 (TP)	4 (FN)
	0 (Negatif)	16 (FP)	42 (TN)

After getting the results of true Positive, true negative, false positive, and true negative it will be calculated accuracy, precision, recall, f1-score and error rate. The test result with confusion matrix displays the output as figure 4.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.92      0.98      0.95       181
           1       0.91      0.72      0.81        58

    accuracy                           0.92       239
   macro avg       0.92      0.85      0.88       239
weighted avg       0.92      0.92      0.91       239
```

Figure 4. Output Confusion Matrix Testing

1. Accuracy

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \times 100\ \% \quad (3)$$

$$\text{Accuracy} = \frac{177+42}{177+16+4+42} \times 100\ \%$$

$$\text{Accuracy} = 0.913 \times 100\% = 91{,}6\ \%$$

2. Precision

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\ \% \quad (4)$$

$$\text{Precision} = \frac{177}{177+16} \times 100\ \%$$

$$\text{Precision} = 0.917 \times 100\% = 91{,}7\%$$

3. Recall

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\ \% \quad (5)$$

$$\text{Recall} = \frac{177}{177+4} \times 100\ \%$$

$$\text{Recall} = 0.977 \times 100\% = 97.7\%$$

4. F1-Score

$$\text{F1- score} = \frac{2 \times precision \times recall}{precision + recall} \times 100\ \% \quad (6)$$

$$\text{F1- score} = \frac{2 \times 0.917 \times 0.977}{0.917 + 0.977} \times 100\ \%$$

$$\text{F1- score} = \frac{1.793}{0.917 + 0.977} \times 100\ \%$$

$$\text{F1- score} = 0.946 \times 100\ \% = 94.6\%$$

5. Error Rate

$$\text{Error Rate} = 1 - \text{Akurasi} \quad (7)$$

$$\text{Error Rate} = 1 - 0.913 = 0.087$$

$$\text{Error Rate} = 0.087 \times 100\% = 9\%$$

F. *Data Visualization Result*

At this stage will be done data visualization in the form of a pie chart that aims to determine the positive and negative sentiment data in the pie chart before and after using the classification Support Vector Machine.

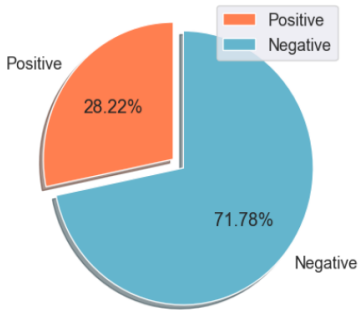

Figure 5. Diagram Pie Before Using Classification SVM.

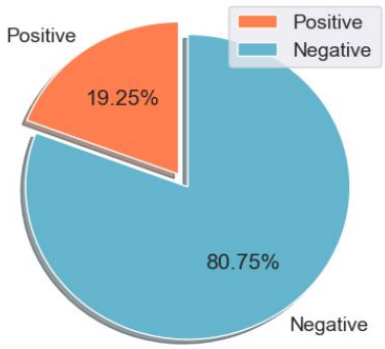

Figure 6. Diagram Pie After Using Classification SVM.

IV. CONCLUSION

Based on the sentiment analysis of lalamove app delivery service user reviews from Google PlayStore reviews using the support vector machine algorithm for this study, it was found that lalamove apps tend to have negative views. with a total positive sentiment of 337 reviews and a total negative sentiment of 857 reviews. The preprocessing stage carried out in the data processing of 1194 Lalamove application reviews was successfully carried out with the stages of case folding, cleaning data, tokenizing, stopwords, and normalization with the results of the dataset divided into two parts 80% and 20%,

Putri Aprilyani and Cut Maisyarah Karyati, "Sentiment analysis of Lalamove application users against Google Playstore reviews using Support Vector Machine algorithm with Python programming language" *International Research Journal of Advanced Engineering and Science*, Volume 10, Issue 3, pp. 35-39, 2025.

namely 955 training data and 239 testing data used to classify using the support vector machine method. After conducting research using the Support Vector Machine (SVM) algorithm model, it can be concluded that SVM proved to be a good enough classification algorithm to analyze sentiment in lalamove application reviews with an accuracy of 91.6%, recall of 97.7%, precision of 91.7%, f1-score of 94.6%, and an error rate of 0.9 or 9%.

REFERENCES

[1] Roji, F. F., Ginasta, N. G., Cahyan, Y., Rahayu, D., & Ramdani, D. (2023). Review analysis of satusehat application using support vector machine and latent dirichlet allocation modeling. *RISTEC: Research in Information Systems and Technology*, *4*(1), 50-62.Hamad, M. M., Salih, M. A., & Jaleel, R. A. (2021). Sentiment analysis of restaurant reviews in social media using naïve bayes. *Applications of Modelling and Simulation*, *5*, 166-172.

[2] Hamad, M. M., Salih, M. A., & Jaleel, R. A. (2021). Sentiment analysis of restaurant reviews in social media using naïve bayes. Applications of Modelling and Simulation, 5, 166-172.

[3] Wulandari, S., & Hasan, F. N. (2024). Analisis Sentimen Masyarakat Indonesia Terhadap Pengalaman Belanja Thrifting Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes. *Jurnal Media Informatika Budidarma*, *8*(2), 768-776.

[4] Suryati, E., Styawati, S., & Aldino, A. A. (2023). Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi dan Sistem Informasi*, *4*(1), 96-106 .

[5] Bourequat, W., & Mourad, H. (2021). Sentiment analysis approach for analyzing iPhone release using support vector machine. *International Journal of Advances in Data and Information Systems*, *2*(1), 36-44

[6] Isnan, M., Elwirehardja, G. N., & Pardamean, B. (2023). Sentiment analysis for TikTok review using VADER sentiment and SVM model. *Procedia Computer Science*, *227*, 168-175

[7] Hadju, S. F. N., & Jayadi, R. I. Y. A. N. T. O. (2021). Sentiment analysis of indonesian e-commerce product reviews using support vector machine based term frequency inverse document frequency. *Journal of Theoretical and Applied Information Technology*, *99*(17), 4316-4325.

[8] Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Mengunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, *5*(1), 32-35.

[9] Hyuningtyas, R. Y., Sari, R., & Yusnaeni, W. (2023, May). Particle swarm optimization for feature selection in sentiment analysis on the application of digital payments OVO using the algorithm of Naive Bayes. In *AIP Conference Proceedings* (Vol. 2714, No. 1). AIP Publishing.

[10] A. Ghafoor, A. S. Imran, S. M. Daudpota, Z. Kastrati, R. Batra, M. A. Wani *et al.*, "The impact of translating resource-rich datasets to low-resource languages through multi-lingual text processing," *IEEE Access*, vol. 9, pp. 124 478–124 490, 2021.

[11] Aisah, I. S., Irawan, B., & Suprapti, T. (2023). Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(6), 3759-3765.

[12] F.Pedregosa,G.Varoquaux,A.Gramfort,V.Michel,B.Thirion,O.Grisel,M. Blondel,P.Prettenhofer,R.Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[13] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, "Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix," *J. Inform. Upgris*, vol. 6, no. 2, pp. 66–75, 2021, doi: 10.26877/jiu.v6i2.6552.

[14] Alzahrani, M. E., Aldhyani, T. H., Alsubari, S. N., Althobaiti, M. M., & Fahad, A. (2022). Developing an Intelligent System with Deep Learning Algorithms for Sentiment Analysis of E-Commerce Product Reviews. *Computational Intelligence and Neuroscience*, *2022*(1), 3840071.

[15] Artanto, F. A. (2024). Support Vector Machine Berbasis Particle Swarm Optimization Pada Analisis Sentimen Anggota KPPS. *JURNAL FASILKOM*, *14*(1), 75-79.

[16] Hidayat, M. N., & Pramudita, R. (2024). Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT. *Information Management For Educators And Professionals: Journal of Information Management*, *8*(2), 161-170.

[17] Kaur, P., & Edalati, M. (2022). Sentiment analysis on electricity Twitter posts. *arXiv preprint arXiv:2206.05042*. \

[18] S. Shaikh, S. M. Daudpota, A. S. Imran, and Z. Kastrati, "Towards improved classification accuracy on highly imbalanced text dataset using deep neural language models," *Applied Sciences*, vol. 11, no. 2, p. 869, 2021.

[19] U. Khaira, R. Johanda, P. E. P. Utomo, and T. Suratno, "Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength," Indonesian Journal of Artificial Intelligence and Data Mining, vol. 3, no. 1, p. 21, 2020.