



## Sentiment Analysis of GoPay Application is Improved Using Natural Language Processing Method to Optimize Services

Ibnu Sina Salman Syam<sup>1\*</sup>, Martanto<sup>2</sup>, Arif Dikananda<sup>3</sup>, Dede Rohman<sup>4</sup>

<sup>1,4</sup> Faculty of Informatics Engineering, STMIK IKMI, Cirebon, Indonesia

<sup>2</sup> Faculty of Informatics Management, STMIK IKMI, Cirebon, Indonesia

<sup>3</sup> Faculty of Software Engineering, STMIK IKMI, Cirebon, Indonesia

[syamsalman371@gmail.com](mailto:syamsalman371@gmail.com)<sup>1\*</sup>

---

### Abstract

This study aims to analyze the sentiment of Gopay application user reviews using the Natural Language Processing (NLP) approach based on the Naïve Bayes Multinomial algorithm. User reviews are taken from a public dataset available on the Kaggle platform, which contains review text and rating scores in Indonesian. The stages of this research include data collection, data preprocessing, data transformation, sentiment modeling, and model evaluation. In the data selection stage, only two relevant columns, namely the review text and the assessment score, are retained to ensure a focused analysis. The preprocessing stage involves text cleaning, punctuation removal, tokenization, stop words removal, and stemming to produce clean and structured data. The processed data is then converted into numerical representations using the TF-IDF (Term Frequency-Inverse Document Frequency) technique. The classification process is carried out using the Naïve Bayes Multinomial algorithm, where reviews are categorized into three sentiments: positive, neutral, and negative, based on user rating scores. Model evaluation was carried out using accuracy, precision, recall, and F1-score metrics. The results show that Naïve Bayes' Multinomial algorithm is able to classify sentiment with 90% accuracy, with the best performance on positive sentiment. However, the model shows weaknesses on the neutral sentiment classification, with a very low F1-score, which is caused by an imbalance in the amount of data between categories. This research contributes by generating insights into user perceptions of the Gopay application. The results of the analysis show that most of the reviews are positive, reflecting user satisfaction with the ease and security of the service. Meanwhile, negative reviews indicate technical and customer service issues that require further attention. The study recommends improving the quality of service based on complaints in negative reviews and maintaining the advantages of features that users appreciate. Thus, the results of this research can be used to support strategic decisionmaking in the development of GoPay services.

**Keywords:** *Natural language Processing, Sentiment Analysis, User Reviews*

---

### 1. Introduction

The rapid developments in the field of informatics have had a significant impact on various aspects of life, including financial technology or fintech. Digital wallet applications such as Gopay have become a crucial tool for people to make financial transactions quickly and easily. The success of this application is highly dependent on its ability to understand user needs and sentiments in order to provide optimal service. Sentiment analysis from user reviews is one of the important ways to explore user perceptions and identify areas for service improvement [1]. In the field of education, the use of informatics-based technology has changed conventional teaching methods to be more flexible and accessible, thereby encouraging the improvement of learning quality [2].

The changes that have occurred in this digital era emphasize the importance of further research on the use of information technology to support various sectors of life. Although the development of information technology provides many benefits, there are a number of challenges that need to be overcome, especially in large-scale data management and analysis. One of the prominent issues is the volume of data generated by digital applications, such as the Gopay payment application, which has not been balanced with adequate analytical capabilities. User reviews that are rich in important information related to service quality are often not processed properly, so business decisions are not optimal without the right analytical tools, such as Natural Language Processing (NLP) techniques [3].

Algorithms commonly used in sentiment analysis, such as Naïve Bayes and K-Nearest Neighbor (KNN), often face limitations when it comes to accuracy and interpretation of complex data [1]. The gap between the development of data analytics technology and its application in the business sector, especially in understanding the nuances of sentiment in user reviews, is a relevant issue. Business decisions based on inaccurate sentiment analysis can be detrimental to the company, especially in improving service quality and user experience. Therefore, the development of more effective NLP methods for sentiment analysis is essential to address this problem and answer the literature gap regarding data-driven service optimization from financial app user reviews.

A number of previous studies have discussed sentiment analysis using the Natural Language Processing (NLP) approach and various algorithms in the context of digital applications. One of the studies conducted by [4] examined user reviews on the Bibit and Bareksa applications using the K-Nearest Neighbor (KNN) algorithm. The study shows that KNN has quite good abilities in sentiment classification.

However, its limitations lie in its degraded performance when dealing with very large and diverse data, and the lack of further exploration of combining NLP with other techniques to improve the accuracy of sentiment analysis results. Another study by [1] conducted a comparison between several algorithms, namely Support Vector Machine (SVM), Naïve Bayes, and Perceptron, in the sentiment analysis of product reviews on Tokopedia. The results of this study show that Naïve Bayes is superior in terms of accuracy compared to other algorithms.

However, this study has limitations in terms of using more in-depth data pre-processing techniques, such as word embedding or more advanced transformer-based models, which can further optimize model performance. Another study by [3] focused on analyzing the sentiment of MyPertamina application user reviews using the KNN and Naïve Bayes algorithms. The results show that Naïve Bayes is superior in data processing speed to KNN, although Naïve Bayes' accuracy still needs to be improved. The study has not yet explored the application of modern NLP techniques, such as Bidirectional Encoder Representations from Transformers (BERT), which have the potential to capture more profound and contextual nuances of sentiment. While these three studies have made a significant contribution to the development of sentiment analysis in digital applications, there is an opportunity for further research. The approach used in previous studies still has limitations, especially in the application of more sophisticated modern NLP techniques and the lack of exploration of more complex data pre-processing methods. Further research may focus on the application of more cutting-edge NLP technology as well as the use of hybrid techniques that combine multiple algorithms to overcome existing limitations and improve the accuracy of sentiment analysis results.

The main purpose of this study is to develop and implement the Natural Language Processing (NLP) method to improve the accuracy of sentiment analysis on Gopay application user reviews. This research aims to overcome the limitations of traditional algorithms used in sentiment analysis, which are often less able to capture the context and nuances of sentiment in depth. Through this research, it is hoped that a more accurate approach can be produced in extracting user sentiment, so that the information generated can support the optimization of Gopay application services. The significance of this research lies in its contribution to filling the knowledge gap related to the application of more effective NLP methods for financial services applications. Practically, this research is also expected to provide benefits for service providers, especially in understanding user feedback more accurately to improve the quality of services provided.

## 2. Research Methods

This study uses the Natural Language Processing (NLP) approach to analyze the sentiment of Gopay application user reviews. The method implemented includes several stages, starting with collecting user review data from the application platform, and then going through a data pre-processing process to clean and prepare the text for analysis. The pre-processing process includes punctuation removal, text conversion to lowercase letters, tokenization, stemming, and deletion of common words (stop words), which aims to improve the quality of text data. After this process, the text will be converted into a numerical representation using the TF-IDF (Term Frequency-Inverse Document Frequency) technique. For sentiment classification, the study used Naïve Bayes' Multinomial algorithm, which is known to be effective for text-based analysis.

The model's performance will be evaluated using metrics such as accuracy, precision, recall, and F1-score to assess the extent to which the model is able to accurately classify sentiment. With this approach, the research is expected to be able to capture user sentiment patterns efficiently, provide reliable results, and be relevant to support decision-making.

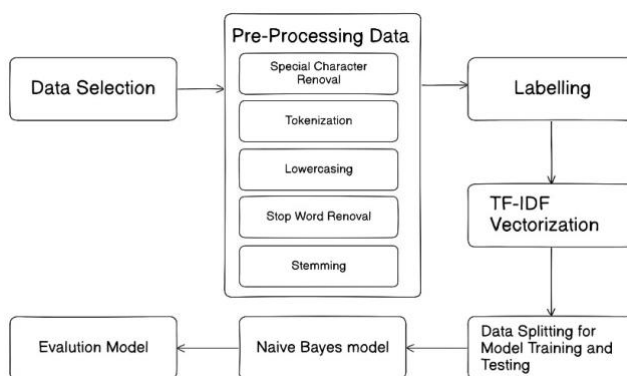


Figure 1: Research method

### 3.1. Results and Discussion

This study uses a method to analyze the review sentiment of Gopay application users using Natural Language Processing (NLP) techniques.

### 3.2. Data Selection

The dataset used in this study is taken from the Kaggle platform, a platform that provides various public datasets for research purposes. The dataset titled "GoPay App Reviews Indonesian" includes GoPay app user reviews in Indonesian.

### 3.3. Pre-processing Data Using NLP Techniques

After the data is collected, pre-processing is carried out on the text of user reviews. This step is critical in the NLP-based sentiment analysis process to ensure that the data is in a form that is ready to be processed by the machine learning model.

- Removal of special characters
- Tokenization to break down text into individual words
- Convert lowercase letters to maintain data consistency.
- Removal of stop words
- Stemming or Lemmatization to change a word to its basic form.

	content	clean_content
0	Mantap	mantap
1	Bagus dan mudah digunakan	bagus mudah
2	Good	good
3	Sangat membantu orang menengah mantap semoga s...	bantu orang tengah mantap moga sukses gopay
4	Ga ribet .. Sukaaaa	ga ribet sukaaaa
5	Sangat membantu	bantu
6	Mantap pakay gopay	mantap pakay gopay
7	mantap mantap mantap	mantap mantap mantap
8	Layanan sang Mudah Dan Sangat Memuaskan	layan sang mudah puas
9	Berbisnis	bisnis
10	Mohon di buka kembali goppy saya 🙏🙏🙏	mohon buka goppy
11	Cepat	cepat
12	keren, dan sangat terbantu	keren bantu
13	Sangat mudah	mudah
14	Sip dah	sip dah
15	Gopay yang sangat mantap respon cepat	gopay mantap respon cepat
16	Memudahkan driver..mantap	mudah drivermantap
17	Keren	keren
18	Mudah dan sangat ringkas	mudah ringkas

Figure 2: Pre-processing Data

### 3.4. Labeling

The labeling process is an essential stage in sentiment analysis that aims to classify data into certain categories according to research needs. In the context of the GoPay application review dataset, labeling is carried out by using the score column as a reference to determine the sentiment category. A score of 4-5 is defined as a positive sentiment because it reflects satisfaction or a good assessment of the application. A score of 3 is categorized as neutral sentiment, which represents the user's views that are not inclined towards satisfaction or dissatisfaction. Meanwhile, a score of 1-2 is classified as a negative sentiment because it indicates user dissatisfaction with the app. The categories generated from this process are used as variable targets in the development of machine learning models.

	clean_content	label
0	mantap	positif
1	bagus mudah	positif
2	good	positif
3	bantu orang tengah mantap moga sukses gopay	positif
4	ga ribet sukaaaa	positif
5	bantu	positif
6	mantap pakay gopay	positif
7	mantap mantap mantap	positif
8	layan sang mudah puas	positif
9	bisnis	positif
10	mohon buka goppy	positif

Figure 3: Labeling

#### 3.4.2. Feature Extraction with TF-IDF Vectorization

After the pre-processing stage, the next step is to extract features from the review text using the TF-IDF (Term Frequency-Inverse Document Frequency) technique. TF-IDF converts the words in the review text into numerical representations that can be incorporated into the classification model.

```
[33] # Step 4: TF-IDF Vectorization
vectorizer = TfidfVectorizer()
x = vectorizer.fit_transform(df['clean_content'])
y = df['label'] # Label sentimen
```

Figure 4: Feature Extraction with TF-IDF Vectorization

### 3.5. Data Sharing for Model Testing training

Before training the model, the data is divided into training data and test data. Training data is used to train the model, while test data is used to evaluate its performance.

```
x_train, x_test, y_train, y_test = train_test_split(x,
y, test_size=0.2, random_state=42)
```

Figure 5: Model Testing Training

### 3.6. Model Naïve Bayes

Naïve Bayes is a classification model that is often used for sentiment analysis due to its simplicity and effectiveness in the classification of texts. effective machine learning for text analysis. This model calculates the probability of each sentiment category (positive, negative, neutral) based on the given features.

```
model = MultinomialNB()
model.fit(x_train, y_train)
```

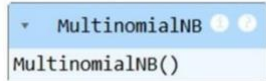


Figure 6: Model Naive Bayes

### 3.7. Model Evaluation

Model evaluation is a crucial step in the development of machine learning algorithms to assess the model's ability to make predictions on new data that has never been used during the training process. This evaluation aims to ensure that the model does not experience overfitting, which is a condition where the model is only able to work well on the training data but fails on the test data. In this context, various evaluation metrics are used, such as accuracy, precision, recall, and F1-score. As a complement, the confusion matrix is used to provide a detailed view of the distribution of correct and false predictions for each class.

	precision	recall	f1-score	support
negatif	0.80	0.62	0.70	3223
netral	0.20	0.00	0.00	839
positif	0.91	0.98	0.95	20589
accuracy			0.90	24651
macro avg	0.64	0.53	0.55	24651
weighted avg	0.87	0.90	0.88	24651

Confusion Matrix:

```
[[ 1996   2 1225]
 [  162   1  676]
 [   337   2 20250]]
```

Figure 7: Model Evaluation

## 4. Conclusion

This research focuses on the application of the Natural Language Processing (NLP) method to improve the sentiment analysis of GoPay application user reviews. In this study, NLP was used to automatically identify user sentiment, which was divided into three main categories: positive, neutral, and negative. Using the Naïve Bayes model, the results show that NLP can provide a more accurate analysis of user perception of the application. The NLP technology applied in this study shows that the model used, the Naïve Bayes Multinomial Model, shows a high overall accuracy, which is 90%, with excellent performance in the positive category.

However, the model's performance in the neutral category is very weak, which is the model's main weakness in sentiment classification. This study shows that the NLP method is effective in fulfilling the research objective, which is to increase sentiment analysis on Gopay user reviews to optimize the service. With the results achieved, this study succeeded in providing recommendations that can be used as a basis for the Gopay development team to improve and improve the service according to user needs and perceptions.

The results of the analysis show that the majority of user reviews have positive sentiment, which reflects satisfaction with Gopay's services. Dominant words such as "easy," "steady," "transactional," and "secure" in the word cloud indicate that the aspects of convenience, speed, and security are the features that users appreciate the most. The distribution of sentiment shows the dominance of positive reviews, while negative and neutral reviews are much less. Word cloud on positive sentiment provides insight into the aspects of services that are valued, while negative reviews indicate areas of service that need improvement. The results of this analysis provide a basis for strategic decision-

making, such as improving user complaints identified in negative reviews and improving features that are the strengths of the GoPay service.

## References

- [1] M. Amien, "History and Development of Natural Language Processing (NLP) Techniques in Indonesian: A Review of the History, Development of Technology, and Applications of NLP in Indonesian," no. 2007, p. 1–7, 2023, [Online]. Available on: <http://arxiv.org/abs/2304.02746>
- [2] M. Diki Hendriyanto, A. A. Ridha, and U. Enri, "Sentiment Analysis of Mola Application Reviews on Google Play Store Using Support Vector Machine Algorithm," *J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, hal. 1–7, 2022.
- [3] M. Aulia and A. Hermawan, "Comparative Analysis of SVM, Naïve Bayes, and Perceptron Algorithms for Sentiment Analysis of Tokopedia Product Reviews," *J. Media Inform. Budidarma*, vol. 7, no. 4, hal. 1850–1859, 2023, doi:10.30865/mib.v7i4.6839.
- [4] D. Maulana and E. Budiarto, "Security Analysis of ChatGPT Information System Using Natural Language Processing (NLP) Algorithms," *J. Technol. Inform. and Komput.*, vol. 9 no. 1 plow. 592–604, 2023, doi:10.37012/jatik.v9i1.1678.
- [5] W. Kurnia, "Sentiment Analysis of E-Commerce Applications Based on User Reviews Using the Stochastic Gradient Descent Algorithm," *J. Technol. and Sist. Inf.*, vol. 4, no. 1, hal. 138–143, 2023.
- [6] A. D. Adhi Putra, "Sentiment Analysis on User Reviews of Bibit and Bareksa Applications with KNN Algorithm," *JATISI (Jurnal Tek. Inform. and Sist. Information)*, Vol. 8, No. 2, Hal. 636–646, 2021, Yogurt: 10.35957/National.V8E2.962.
- [7] N. P. Eki, Indrianti, and W. C. Wihandika, "Sentiment Analysis on User Reviews of Online Mandiri Applications Using the Modified Term Frequency Scheme and Naïve Bayes Method," *J. Developer. Technol. Inf. and Komput Science.*, vol. 5, no. 3, hal. 1039–1047, 2021.
- [8] I. Huda, "Implementation of Natural Language Processing (NLP) for Location Search Applications," *J. Nas. Technol. Apply.*, Vol. 3 No. 2 Hull. 15 2021, doi: 10.22146/JNTT.35036.
- [9] E. Wita, "BEES: Bulletin of Electrical and Electronics Engineering Application of Natural Language Processing to Identify Ambiguous Sentences in Newspapers Applying the Shift Reduce Parsing Method ARTICLE INFORMATION," vol. 2, no. 2, p. 63–66, 2021.
- [10] N. Nurwanda, N. Suarna, and W. Prihartono, "The Application of NLP (Natural Language Processing) in the Analysis of Telegram User Sentiment on Playstore," *JATI (Journal of Mhs. Tek. Inform.*, Vol. 8, No. 2, Hal. 1841–1846, 2024, Yogurt: 10.36040/nation.V8E2.8469.
- [11] D. T. Ksatria, Y. Yunefri, and L. L. Van FC, "Sentiment Analysis of Mypertamina Application User Reviews on Google Playstore Using K-Nearest Neighbor and Naïve Bayes," *Technol. Inf.*, vol. 2, no. 1, p. 213–227, 2023, [Online]. Available on: <https://journal.unilak.ac.id/index.php/Semaster/article/view/18526>
- [12] D. O. Sihombing, "Implementation of Natural Language Processing (NLP) and Cosine Similarity Algorithm in Automated Essay Exam Assessment," *J. Sist. Comput. and Inform.*, vol. 4, no. 2, hal. 396, 2022, doi:10.30865/json.v4i2.5374.
- [13] H. Irsyad and A. Taqwiyim, "Analysis of User Reviews of Plant Diagnosis Applications on the Play Store Using Naïve Bayes," *Bul. Ilm. Inform. Technol.*, vol. 1, no. 2, hal. 43–51, 2022.
- [14] V. R. Prasetyo, N. Benarkah, and V. J. Chrisintha, "Implementation of Natural Language Processing in Making Chatbots in the Information Technology Program of the University of Surabaya," *Teknika*, vol. 10, no. 2, hal. 114–121, 2021, to date: 10.34148/10i2.370.
- [15] N. A. Mutiara, A. Fitri, K. S. Rizwine, S. S. Vetryc, and K. Fihri, "Implementation of NLP in the Creation of Chatbot Customer Service Publisher Case Study Journal LARISMA," vol. 1, no. 1, p. 13–17, 2024.
- [16] Supriady, A. Eldita, and S. Dewinta, "APPLICATION OF NLP ALGORITHM ON SAWALA INTERACTIVE DISCUSSION PLATFORM WITH CHATBOT FOR ULBI STUDENTS D3 Informatics Engineering Study Program, University of Logistics and International Business Abstract This research focuses on applying Natural Language Processing AI," *J. Tek. Inform.*, vol. 16, no. 2, hal. 41–45, 2024.