Jurnal Komputer, Informasi dan Teknologi Vol: 4, No 2, 2024, Page: 1-10

The Use of Naïve Bayes Algorithm in Sentiment Analysis of Grab Application Reviews

Febri Aditiya[1*], Donny Maulana[2], Edora[3]

[123] Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

DOI:
https://doi.org/10.53697/jkomitek.v4i2.1999
*Correspondence: Febri Aditiya
Email:
febri_aditiya@mhs.pelitabangsa.ac.id

Received: 23-10-2024
Accepted: 22-11-2024
Published: 23-12-2024

Abstract: This study aims to analyze the sentiment of Grab application user reviews using the Naïve Bayes algorithm. In conducting the analysis, how well the Naïve Bayes algorithm can analyze Grab application user review data. Thus, this process refers to the elements that affect the accuracy of the Naïve Bayes analysis. to analyze the sentiment of Grab application user reviews using the Naïve Bayes algorithm. Thus, it not only evaluates the accuracy of the Naïve Bayes method, but also considers the problem of imbalanced class, because in reality not all datasets can be accessed perfectly. To consider the problem of data imbalance, this study requires the Adaptive Synthetic Sampling or ADASYN technique which is used in machine learning to overcome the problem of class imbalance in the dataset. The tools used in processing the algorithm in the method and conducting the analysis use Google Colab. This study focuses on the classification of positive and negative sentiments from user reviews taken by the scraping process from the Google Play Store platform. The analysis process involves data preprocessing, including tokenization, stemming, and word weighting, to improve the accuracy of sentiment classification. Based on the results of this study, the Naïve Bayes model sought an accuracy result of 85.33%, the precision result obtained was 80.55% and the recall result in this research test was 79.09%, from these results by implementing calculations using a confusion matrix and dividing the data into testing data and training data.

Keywords: Sentiment Analysis, Naïve Bayes, ADASYN, Review, Scrapping

Introduction

The Grab application has become one of the most popular transportation and delivery services in various countries, including Indonesia, with 13.8 million users providing diverse reviews about their experiences. This review analysis highlights the ability of algorithms to identify user satisfaction levels by classifying sentiments in Grab app reviews. Specifically, it examines how well the Naïve Bayes algorithm can analyze user review data to classify sentiments as positive or negative. Evaluating the accuracy of this method is crucial, given the need for precise sentiment classification results to support strategic corporate decisions, whether in service improvement or reputation management. Therefore, this study focuses on evaluating the performance of the Naïve Bayes algorithm on Grab review data(Muhamad Anton Permana1), 2023). As a solution for understanding user perceptions and experiences, this research focuses on elements that influence the accuracy of Naïve

Bayes analysis, such as selecting appropriate text features, data preprocessing, and handling words with ambiguous meanings or complex emotional connotations in the Indonesian language (Aji, 2023). The preprocessing and feature selection processes, including techniques such as stemming, tokenization, and word weighting, are expected to provide solutions for enhancing performance.

Thus, this study not only evaluates the accuracy of the Naïve Bayes method but also addresses the issue of imbalanced classes. In reality, not all datasets can be perfectly accessed, especially when dealing with raw data obtained directly through observation or primary data collection. This issue is evident in previous research that faced challenges related to datasets with disproportionate class distributions, known as imbalanced classes. The imbalance in the number of data points available for training the model between one class and another can lead to skewed classification results, impacting overall model performance(Pratama et al., 2022). To address the issue of data imbalance, this study incorporates the use of Adaptive Synthetic Sampling (ADASYN), a machine learning technique designed to handle class imbalance problems in datasets (Kustanto, 2021).

The use of the Naïve Bayes algorithm for sentiment analysis in this study offers a more accurate and objective approach to classifying user reviews into positive and negative categories. This algorithm is capable of handling large volumes of review data automatically and has demonstrated high accuracy in previous studies (Lestari & Saepudin, 2021). Manual analysis alone is not feasible for processing such large volumes of data. Therefore, this study aims to provide a more accurate evaluation of the Grab application based on real user experiences (Hyuningtyas, 2023). This approach offers a comprehensive overview of the application's quality through genuine reviews. The researchers suggest that the Naïve Bayes algorithm holds significant potential for sentiment analysis of user reviews for the Grab application, as it demonstrates reliable accuracy in sentiment classification. To further enhance its performance, the method requires effective preprocessing steps and feature selection techniques, such as tokenization, stemming, and appropriate word weighting. The variables influencing the performance of the Naïve Bayes algorithm in sentiment analysis are critical in determining whether they can improve sentiment classification outcomes. With proper implementation, this algorithm can provide an efficient solution for companies to understand user perceptions of their services (Rexie, 2023). This study is expected to serve as a foundation for further development in utilizing the Naïve Bayes algorithm, as well as other methods, to enhance sentiment analysis accuracy within the online transportation industry (Hermanto, 2020).

This study aims to analyze user perceptions of the Grab application based on reviews provided on the Google Play Store by categorizing these reviews into positive and negative sentiment groups. Additionally, it seeks to evaluate the performance of the Naïve Bayes algorithm in conducting sentiment analysis, including measuring the accuracy of the algorithm in classifying user reviews by sentiment. Furthermore, the research aims to develop additional indicators for assessing the quality of online transportation applications by utilizing sentiment analysis as an alternative or complement to traditional metrics, such as star ratings and download numbers (Ernawati, 2020). Lastly, the study seeks to uncover insights from user reviews, such as specific criticisms, suggestions, and complaints, which are often overlooked in star-based assessments, thereby providing a deeper understanding of user experiences.

Methodology

The data used in this study consists of textual data in the form of user reviews of the Grab application obtained from the Google Play Store platform. This data is qualitative in nature, as it comprises written comments that reflect users' opinions, evaluations, and experiences with the Grab app. Each review contains text expressing user sentiments, such as complaints, praise, or suggestions, often accompanied by a numerical rating (stars). The data will undergo sentiment analysis, where reviews are classified into positive and negative sentiment categories (Putri, 2020). The process will also involve data extraction and preprocessing, including the removal of irrelevant words and text normalization, to facilitate the analysis.

The data used in this study consists of a collection of user reviews of the Grab application from the Google Play Store. These reviews include text containing users' opinions about their experiences with the Grab service, which were obtained through a data scraping process, along with star ratings given as evaluations of the app. The textual data covers a range of sentiments, including positive comments appreciating the app's features and services, as well as negative comments containing complaints or issues experienced by users. The data was collected over a specific time period to ensure adequate variation and representation. Additionally, for sentiment analysis using the Naïve Bayes method, the review data will undergo preprocessing steps such as punctuation removal, stopword elimination, and text normalization to ensure the model can process and classify sentiments more accurately (Mussalimun, 2021).

This study will go through several stages, organized systematically and progressively. The sequence of research processes includes every step necessary to achieve the research objectives, from data collection to result analysis. Each stage will be described in detail in the methodology section, providing a clear understanding of the workflow used, followed by testing with the Naïve Bayes algorithm.

The process begins with scraping review data. After the data is obtained, a dataframe will be created from the collected dataset, which consists of user reviews, including review ratings. This data will be gathered from the Play Store, ensuring diversity and representativeness. Next, the labeling process will convert the review rating values into positive and negative sentiments (Habal, 2023). After obtaining sentiment values, the data preprocessing step will be carried out, including tokenization, stopword removal, and stemming, to prepare the text to be cleaner and more structured.

Subsequently, a vectorization technique, such as Term Frequency-Inverse Document Frequency (TF-IDF), will be used to convert the text into a numerical representation that can be understood by the Naïve Bayes algorithm. Once the data is ready, the Naïve Bayes model will be applied to classify the sentiment of the reviews. The data will be split into training and testing datasets. The training data will then be used to calculate the probabilities of each sentiment class based on the words in the reviews. The following are the stages of research used in this study, as outlined below:

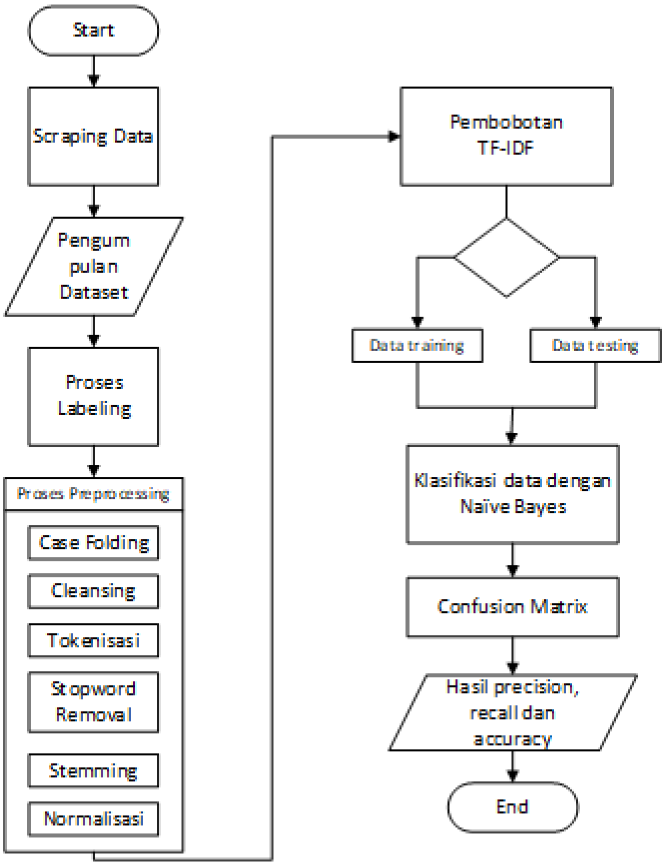

Figure 1. Research Methodology

In this study, data was collected using the scraping technique, utilizing the *google_play_scraper* library, which allows for the automatic extraction of Grab app reviews from the Google Play Store. The data collection process took place on October 14, 2024, with the goal of obtaining a large volume of data quickly and efficiently. This technique was chosen because it minimizes errors that often occur in manual data collection and ensures the data is more accurate and organized. After the data was successfully collected, the next step involved labeling the sentiment of the reviews based on the ratings given by users. Reviews with ratings below 4 were labeled as negative, while those with ratings of 4 or higher were labeled as positive. Out of the 1,500 reviews collected, 558 were identified as positive, and 942 as negative (Pratmanto, 2020).

Next, the collected data underwent a preprocessing process to ensure better data quality and readiness for analysis. This process was carried out using Google Colab to facilitate and speed up the data processing. The preprocessing stage included several steps, such as data cleaning and text normalization. One of the text normalization steps is *case folding*, which converts all letters to lowercase to avoid unnecessary differences between uppercase and lowercase letters (Durachman, 2024). Then, data cleaning was performed to remove irrelevant content, such as URLs, emoticons, and non-alphanumeric characters, ensuring the data is cleaner and ready for further analysis. Finally, the cleaned data underwent *tokenization*, where the text was broken down into smaller tokens, such as words or phrases, that can be analyzed by the algorithms used in this research. All these steps are aimed at improving the accuracy and efficiency of sentiment analysis for Grab app reviews.

Result and Discussion

Dataset

The dataset in this study consists of data collected from the Grab app on the Google Play Store using the web scraping technique. The data was downloaded on October 14, 2024, from the application with the ID 'com.grabtaxi.passenger'. The collection process involved scraping essential elements such as ratings and user reviews. Web scraping was selected because it allows for the automatic and efficient gathering of large amounts of data, resulting in 1,500 reviews along with their corresponding ratings, which served as the foundation for the analysis (Mustopa, 2020).

Following data collection, the gathered data underwent a preprocessing stage to ensure quality and consistency before being used in the analysis model. After data selection, only the review content and sentiment were retained, leading to a final dataset consisting of 1,500 entries, with 558 positive sentiments and 942 negative sentiments. This dataset effectively represents user perceptions of the Grab app on the Google Play Store and can be utilized to build accurate predictive or analytical models.

Table 1. Dataset Collection

Query	Data Scraped	Sentiment
'com.grabtaxi.passenger'	1,500	Positive
		Negative

Testing Process

In the sentiment analysis testing of Grab app reviews, the Naive Bayes method was utilized, with Google Colab serving as the platform for the tests. User review data was used as input for the sentiment analysis model, with the objective of classifying the sentiment of each review as either positive or negative. Google Colab was chosen because it provides a cloud-based programming environment that allows Python code execution with access to reliable computational resources.

Before testing, the review data, which had undergone preprocessing steps such as text cleaning, removal of stopwords, and stemming, was transformed into numerical representations using vectorization techniques like TF-IDF (Term Frequency-Inverse Document Frequency). Following this, the Naive Bayes algorithm was applied as the main classification method (Shanmugapriyaa, 2023). Naive Bayes was selected due to its effectiveness in handling large text datasets with inter-feature relationships, enabling quick and accurate classifications.

In the testing phase, the dataset was split into training and testing data. The training data was used to train the model to recognize sentiment patterns in the text. After training, the testing data was used to assess the model's performance using evaluation metrics such as accuracy, precision, and recall. The results from this Google Colab-based testing provided insights into how well the Naive Bayes model could identify positive and negative sentiments in user reviews of the Grab app, as well as its accuracy in predicting sentiments for new reviews.

Testing Results

In this testing process, the Python programming language was used, and the Naive Bayes algorithm was applied with Google Colab as the tool for processing the tests. The confusion matrix results from using the Naive Bayes model are presented below:

Table 2. Confusion Matrix Results

Predicted Data	Negative	Positive
Negative	169	23
Positive	21	87

The values in the confusion matrix represent the classification results: 87 true positive (TP) values, 169 true negative (TN) values, 21 false positive (FP) values, and 23 false negative (FN) values. These values were used to calculate the following evaluation metrics:

Accuracy

The accuracy of the Naive Bayes model in this test was 85.33%, as calculated below:
Accuracy = (TP + TN) / (TP + TN + FP + FN) = (87 + 169) / (87 + 169 + 21 + 23) = 256 / 300 = 0.8533 = 85.33%

Precision

The precision of the model was calculated as 80.55%, as shown below:
Precision = TP / (TP + FP) = 87 / (87 + 21) = 87 / 108 = 0.8055 = 80.55%

Recall

The recall for the model was 79.09%, calculated as:
Recall = TP / (TP + FN) = 87 / (87 + 23) = 87 / 110 = 0.7909 = 79.09%
The overall evaluation results from the confusion matrix and the calculation of accuracy, precision, and recall using the Naive Bayes model, along with the ADASYN (Adaptive Synthetic Sampling) technique for the confusion matrix calculation, are summarized in the table below:

Table 3. Testing Data Evaluation Results

Accuracy	Precision	Recall
85.33%	80.55%	79.09%

Sentiment Classification Results

The sentiment analysis classification results are based on the testing model used in this analysis. The following represents the distribution of reviews according to their sentiment classification from the Grab app reviews.

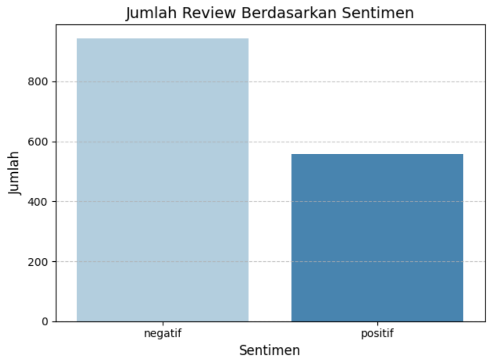

Figure 2. Sentiment Review Distribution

From the graph shown, it can be concluded that, out of the 1500 Grab app reviews tested, there were 942 negative sentiments and 558 positive sentiments. This indicates that the number of negative reviews significantly outweighs the number of positive reviews. This suggests that the majority of users tend to provide less favorable feedback.

Google Colab Output Results

The following are the results of the calculations performed using the Naive Bayes algorithm in Google Colab, which yielded the accuracy, precision, and recall values shown in Figure **3.** The results from the manual calculations using formulas and the algorithmic computation are consistent, confirming that the Naive Bayes model's outputs align with the manually derived metrics.

```
Data Train setelah ADASYN: 1503
Data Test: 300
Confusion Matrix :
                 Predicted Negative  Predicted Positive
True Negative                   169                  21
True Positive                    23                  87
Accuracy   : 0.8533333333333334
Precision  : 0.8055555555555556
Recall     : 0.7909090909090909
```

Figure 3. Naive Bayes Algorithm Output

Based on the results from the confusion matrix, the model is able to correctly classify the majority of data, both for positive and negative classes. However, there are still some cases of false positives and false negatives. The relatively high accuracy value (85.33%) indicates that the model performs well overall in classifying the data. The high precision value (80.55%) shows that when the model predicts a review as positive, it is likely to be correct. The slightly lower recall value (79.09%) suggests that the model may miss some positive reviews, which could be a concern if the main goal is to identify all positive reviews. The graph showing the sentiment distribution of reviews reveals a significant imbalance between negative and positive classes, with negative reviews dominating. This indicates that most users tend to leave negative feedback (Al-Hagree, 2022).

The Naive Bayes model developed demonstrates good performance in classifying sentiment in Grab app reviews. The use of the ADASYN technique helped address the data imbalance issue, allowing the model to perform better for the minority class (positive). The high precision value indicates that the model is reliable in identifying actual positive reviews. However, the slightly lower recall suggests that the model needs further improvement to capture all positive reviews. The developed model can serve as a tool to monitor customer sentiment periodically, enabling the company to identify emerging trends and issues and take corrective actions. Although the model has provided good results, there is still room for improvement. For example, experimenting with other algorithms or fine-tuning parameters could potentially enhance the model's performance (Kristiyanti, 2020).

Conclusion

The Naïve Bayes algorithm has proven to be effective in classifying the sentiment of Grab app user reviews, achieving an accuracy of 85.33%. This indicates the strong potential of the model in performing sentiment analysis, particularly in automatically identifying both positive and negative reviews. The high precision value (80.55%) suggests that predictions for positive reviews tend to be accurate. However, the slightly lower recall value (79.09%) indicates that some positive reviews may be missed, which could impact the model's overall ability to identify all positive reviews. The preprocessing steps, such as tokenization, stemming, and word weighting, play a critical role in enhancing the model's performance. These steps help the algorithm understand the meaning and context of each word, especially in Indonesian, where words can have varied meanings and emotional connotations (Maheswari, 2019).

Additionally, the use of the ADASYN technique has been effective in addressing the data imbalance, making the model more responsive in detecting the minority class (positive reviews). Although the model demonstrates good performance, limitations such as false positives, false negatives, and the lower recall for positive reviews still exist. Therefore, further research is recommended to explore alternative methods or optimize parameter tuning to enhance the model's performance. Other algorithms could also be explored to compare results with Naïve Bayes (Abdulameer, 2024).

References

Abdulameer, A. G. (2024). Naïve Bayes algorithm for timely fault diagnosis in helical gear transmissions using vibration signal analysis. *International Journal on Interactive Design and Manufacturing*. https://doi.org/10.1007/s12008-024-02037-0

Aji, S. (2023). The algorithm comparison of support vector machine and Naive Bayes in sentiment analyzing the Tiktok application. *AIP Conference Proceedings*, *2714*. https://doi.org/10.1063/5.0129009

Al-Hagree, S. (2022). Arabic Sentiment Analysis on Mobile Applications Using Levenshtein Distance Algorithm and Naive Bayes. *2022 2nd International Conference on Emerging Smart Technologies and Applications, ESmarTA 2022*. https://doi.org/10.1109/eSmarTA56775.2022.9935492

Durachman, Y. (2024). Analysis Sentiment of Public Opinion on Social Media Using Naïve Bayes and TF-IDF Algorithms. *2024 3rd International Conference on Creative Communication and Innovative Technology, ICCIT 2024*. https://doi.org/10.1109/ICCIT62134.2024.10701191

Ernawati, S. (2020). Comparison of Naïve Bayes Algorithm with Genetic Algorithm and Particle Swarm Optimization as Feature Selection for Sentiment Analysis Review of Digital Learning Application. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012040

Habal, B. G. M. (2023). Classifying User Experience (UX) of the M-Commerce Application Using Multinomial Naive Bayes Algorithm. *ACM International Conference Proceeding Series*, 135–142. https://doi.org/10.1145/3639233.3639333

Hermanto. (2020). Gojek and Grab User Sentiment Analysis on Google Play Using Naive Bayes Algorithm and Support Vector Machine Based Smote Technique. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012102

Hyuningtyas, R. Y. (2023). Particle swarm optimization for feature selection in sentiment analysis on the application of digital payments OVO using the algorithm of Naive Bayes. *AIP Conference Proceedings*, *2714*. https://doi.org/10.1063/5.0129011

Kristiyanti, D. A. (2020). E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012079

Kustanto, N. S. (2021). Sentiment Analysis of Indonesia's National Health Insurance Mobile Application using Naïve Bayes Algorithm. *2021 International Conference on Artificial Intelligence and Big Data Analytics, ICAIBDA 2021*, 38–42. https://doi.org/10.1109/ICAIBDA53487.2021.9689696

Lestari, S., & Saepudin, S. (2021). Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes. *Seminar Nasional Sistem Informasi Dan Manajemen Informatika*, 163–170.

Maheswari, S. (2019). Heart disease prediction system using decision tree and naive bayes algorithm. *Current Medical Imaging*, *15*(8), 712–717. https://doi.org/10.2174/1573405614666180322141259

Muhamad Anton Permana1), S. W. S. S. 3). (2023). Analisis Sentimen Penggunaan Aplikasi VideoConference Pada Ulasan Google Play StoreMenggunakan Metode Nbc (Naive Bayes Classifier). *Jurnal Riset Sistem Informasi Dan Teknologi Informasi (JURSISTEKNI)*, *5*(1), 178–191.

Mussalimun. (2021). Comparison of K-N earest Neighbor (K-NN) and Naïve Bayes Algorithm for Sentiment Analysis on Google Play Store Textual Reviews. *2021 8th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2021*, 180–184. https://doi.org/10.1109/ICITACEE53184.2021.9617217

Mustopa, A. (2020). Analysis of user reviews for the pedulilindungi application on google play using the support vector machine and naive bayes algorithm based on particle swarm optimization. *2020 5th International Conference on Informatics and Computing, ICIC 2020*. https://doi.org/10.1109/ICIC50835.2020.9288655

Pratama, I., Chandra, A. Y., & Presetyaningrum, P. T. (2022). Seleksi Fitur dan Penanganan Imbalanced Data menggunakan RFECV dan ADASYN. *Jurnal Eksplora Informatika*, *11*(1), 38–49. https://doi.org/10.30864/eksplora.v11i1.578

Pratmanto, D. (2020). App Review Sentiment Analysis Shopee Application in Google Play Store Using Naive Bayes Algorithm. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012043

Putri, D. A. (2020). Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012085

Rexie, J. A. M. (2023). k-mer Based Prediction of Gene Family by applying Multinomial Naïve Bayes Algorithm in DNA sequence. *AIP Conference Proceedings*, *2914*(1). https://doi.org/10.1063/5.0175878

Shanmugapriyaa, D. (2023). A Comparative Study of Sentiment Analysis on Flipkart Dataset using Naïve Bayes Classifier Algorithm. *14th International Conference on Advances in Computing, Control, and Telecommunication Technologies, ACT 2023*, *2023*, 1110–1118. https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85174312884&origin=inward