Collabits Journal
Vol. 2 No. 3 September 2025 : 151-153
E-ISSN : 3046-6709, P-ISSN : 3062-8601
https://publikasi.mercubuana.ac.id/index.php/collabits

Sentiment Analysis Of Reviews Grab Application On Google Playstore Based On Methods Naïve Bayes

Alvian Nugroho[1*], Nisa Febriyani[2], Heri Kurniawan[3]

[1, 2, 3] Informatics Management Study Program, University Raharja, Jakarta, Indonesia

*Coressponden Author: alvian.nugroho@raharja.info

Abstract - This research aims to conduct sentiment analysis of user reviews for the Grab application in the Google Play Store using the Naïve Bayes method. The research uses data in Indonesian language and analyzes sentiment in three classes: positive, neutral, and negative. The Naïve Bayes method is used to classify user reviews into the appropriate sentiment categories. The research utilizes the Google Play Store API and the Google_play_scrapper library to collect user review data. A total of 1195 reviews were successfully collected. The results of the sentiment analysis are expected to provide valuable insights for Grab in improving user experience and the quality of their application services.

Keywords :
Naïve Bayes;
Sentiment Analysis;
Python;

Article History:
Received: 02-07-2025
Revised: 19-08-2025
Accepted: 28-09-2025

Article DOI : 10.22441/collabits.v2i3.30263

1. INTRODUCTION

The very rapid development of technology has a significant impact on humans. One of the most significant impacts is the use of transportation via the internet. The increase in the number of Indonesian internet transportation markets has increased significantly, in 2015, increasing to US$ 900 million in 2018, with a value of US$ 12.7 billion and is predicted to reach US$ 40 billion in 2025. In addition, this research shows that Indonesia is a country that is developing rapidly with the largest internet transportation sector in Southeast Asia, currently, based on a survey by the Association of Electronic Goods Providers. [1]

In this context, this research will examine the implementation of the Naive Bayes Method in sentiment analysis on Grab application reviews on the Google Play Store. Metode Naïve Bayes Classifier ini merupakan salah satu metode pembelajaran mesin yang merupakan teknik sederhana berbasis probabilitas untuk analisis sentimen [2] The main objective of this research is to measure general user sentiment towards the Grab application based on reviews given by users. The results of this sentiment analysis can provide valuable insights for Grab in order to improve user experience and maintain the quality of their application services.

2. CONCEPTUAL STAGE

2.1 Main Theory

1) Naïve Bayes

According to Primajaya & Sari [3] Naïve Bayes is one of the standards for Bayes' theorem. Based on the simplifying assumption that if given an output value, the attribute values are conditionally independent of each other. In other words, the joint observing probability is the product of the individual probabilities given the output value.

2) Python

Python is a programming language that is widely used by large companies and developers to develop various kinds of desktop, web and mobile-based applications. Python was created by Guido van Rossum in the Netherlands in 1990 and its name was taken from Guido's favorite television show Monty Python's Flying Circus. Van Rossum developed Python as a hobby, then Python became a programming language that is widely used in industry and education because it is simple, concise, has intuitive syntax and has an extensive library[4].

3) Sub Vector Machine (SVM)

Support Vector Machine (SVM) is a relatively new technique for making predictions, both in the case of classification and regression. Support Vector Machines are included in the supervised learning class, where in their implementation there is a training stage using sequential SVM training followed by a testing stage [5].

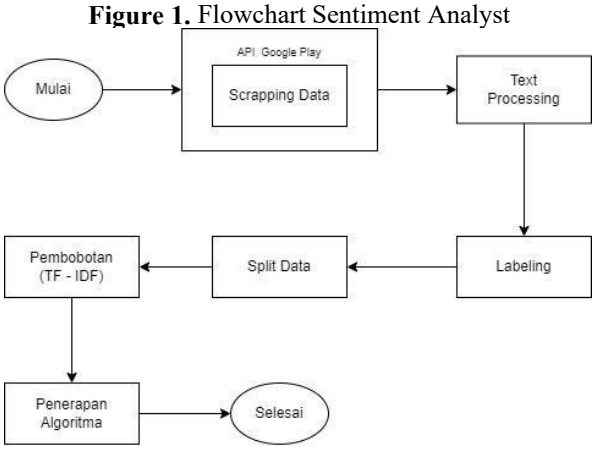
Support Vector Machine (SVM) uses the concept of optimal hyperplane search which is useful for separating two classes of data well. SVM is effective in handling datasets that have many dimensions by utilizing kernel techniques. In the process, SVM only utilizes a few key data points called support vectors to form a classification model, which allows efficient decision making.

3. RESEARCH METHODOLOGY

In the research chosen is comparative analysis. This research aims to analyze sentiment in the review text of the Grab application on Google Playstore using the Naïve Bayes method. This research design involves comparing positive, negative and neutral sentiments contained in Grab application user reviews. The data collected via the Google Playstore API was 1195 pieces and then processed and analyzed using the Naïve Bayes and SVM methods. The results of this sentiment analysis are then compared to see the differences and patterns that emerge in user reviews.

4. RESULT

Figure 1. Flowchart Sentiment Analyst



The flowchart above is the flow of the research process using the Python programming language

Figure. 2. Classification Report Naïve Bayes

Classification Report (Data Latih - Naive Bayes):

	precision	recall	f1-score	support
Positif	0.99	0.62	0.76	301
negatif	0.85	1.00	0.92	654
accuracy			0.88	955
macro avg	0.92	0.81	0.84	955
weighted avg	0.89	0.88	0.87	955

Classification Report (Data Uji - Naive Bayes):

	precision	recall	f1-score	support
Positif	1.00	0.50	0.67	68
negatif	0.83	1.00	0.91	171
accuracy			0.86	239
macro avg	0.92	0.75	0.79	239
weighted avg	0.88	0.86	0.84	239

In the image above, the classification report results from training and test data on the Naïve Bayes algorithm. The results of the training data get an accuracy value of 88%, false negative precision of 89%, true positive precision of 99%, negative recall of 100%, and positive recall of 62%, then the results of the classification report on the test data get an accuracy value of 86%. false negative precision is 83%, true positive precision is 100%, negative recall is 100%, and positive recall is 50%.

Figure. 3. Classification Report SVM

Classification Report (Data Latih - SVM):

	precision	recall	f1-score	support
Positif	1.00	0.87	0.93	301
negatif	0.94	1.00	0.97	654
accuracy			0.96	955
macro avg	0.97	0.93	0.95	955
weighted avg	0.96	0.96	0.96	955

Classification Report (Data Uji - SVM):

	precision	recall	f1-score	support
Positif	0.93	0.59	0.72	68
negatif	0.86	0.98	0.92	171
accuracy			0.87	239
macro avg	0.89	0.79	0.82	239
weighted avg	0.88	0.87	0.86	239

from training data and test data on the SVM algorithm. The results of the training data get an accuracy value of 96%, false negative precision of 94%, true positive precision of 100%, negative recall of 100% and positive recall of 100%, then the results of the classification report on the test data get an accuracy value of 87%, precision false negative was 86%, true positive precision was 93%, negative recall was 98% and positive recall was 59%.

Using 1195 data and divided into 2 types of data, namely training data and test data with a ratio of 80 : 20 (955 : 239). Sentiment classification using the SVM algorithm produces better performance compared to the Naïve Bayes algorithm. The SVM model is able to provide high accuracy on both training data and test data, with an accuracy value on test data of 89.53%. However, both Naïve Bayes and SVM show quite good abilities in classifying sentiment from text.

Results from the classification report show that the SVM model tends to have better precision in identifying negative and positive sentiment, especially on training data. This is shown by the high precision and recall values, especially in positive sentiment classification.

Although the SVM model has high accuracy on training data, there is a slight decrease in performance on test data. This indicates the potential for overfitting in the SVM model, which can be overcome by using regularization techniques or adjusting model parameters.

REFERENCES

[1] APJII. (2018). DAFTAR ISI B. PERILAKU PENGGUNA INTERNET INDONESIA.
[2] Watrianthos, R., Suryadi, S., Irmayani, D., Nasution, M., & Simanjorang,
[3] E. F. S. (2019). Sentiment Analysis Of Traveloka App Using Naïve Bayes Classifier Method. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY
[4] Primajaya, A., & Sari, B. N. (2018). Random Forest Algorithm for Prediction of Precipitation. Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM), 1(1), 27–31.
[5] Romzi, M., & Kurniawan, B. (2020). PEMBELAJARAN PEMROGRAMAN PYTHON DENGAN PENDEKATAN LOGIKA ALGORITMA (Issue 2).
[6] Julianto, Y., & Setiabudi, D. (2022). Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine.