

IMPROVEMENT OF NAIVE BAYES ALGORITHM IN SENTIMENT ANALYSIS OF SHOPEE APPLICATION REVIEWS ON GOOGLE PLAY STORE

Khoerotul Melina Elistiana^{*1}, Bagus Adhi Kusuma², Pungkas Subarkah³, Hasri Akbar Awal Rozaq⁴

^{1,2,3}Informatics, Faculty of Computer Science, Universitas Amikom Purwokerto, Indonesia

⁴Graduate school of Informatics, Departement of computer science, Gazi University Ankara, Turkiye

Email: ¹khoerotul6@gmail.com, ²bagus@amikompurwokerto.ac.id, ³subarkah@amikompurwokerto.ac.id,
⁴hakbar.rozaq@gazi.edu.tr

(Article received: October 23, 2023; Revision: November 09, 2023; published: Desember 23, 2023)

Abstract

Reviews of the shopee application on the google play store are included in useful information if processed properly. Old or new users can analyze app reviews to get information that can be used to evaluate services. The activity of analyzing application reviews is not enough just to see the number of stars, it is necessary to see the entire contents of the review comments to be able to know the intent of the review. A sentiment analysis system is a system used to automatically analyze a review to obtain information including sentiment information that is part of an online review. The data is classified using Naive Bayes. A total of 1,000 shopee app user reviews were collected to form the sample dataset. The purpose of this study is to determine the sentiment analysis of shopee application reviews in the Google Play Store using the Naive Bayes algorithm. The stages of this research include, data collection, labeling, pre-processing, sentiment classification, and evaluation. In the pre-processing stage there are 6 stages, namely Cleaning, Case folding, Word Normalizer, Tokenizing, Stopword Removal and Stemming. TF-IDF (Term Frequency - Inverse Document Frequency) method is used for word weighting. The data will be grouped into two categories, namely negative and positive. The data will then be evaluated using accuracy parameter testing. The test results show an accuracy value of 81%, this result shows that shopee application reviews tend to be negative.

Keywords: marketplace, Naive Baiyes, Sentiment analysis, TF-IDF.

1. INTRODUCTION

The rapid development of E-commerce in Indonesia has given rise to many marketplaces that offer various products and services to the public. This phenomenon not only reflects the rapid growth of technology, but also creates various opportunities for consumers or businesses. With the increasing number of marketplace options, consumers can easily search and compare products and prices [1]. It also creates healthy competition among marketplaces, encouraging innovation in terms of user experience, services, and promotions.

Currently, the Google Play Store offers a plethora of marketplace applications. Google play store is one of the digital content services owned by Google which contains digital products such as applications, music or songs, books, games and cloud-based media players [2]. In the play store there are many features such as rating and review features where users of products from the play store can provide their opinions on the products they have used.

One of the e-commerce applications found on the play store is the shopee application launched in 2015, based on data from (SimilarWeb, May 2023) shopee users in Indonesia are around 161 million people. Because everyone has different opinions and

thoughts, there are many opinions that are pro and contra to this application. User reviews are very diverse ranging from good and bad talk, user criticism and there are also those who provide suggestions regarding features. Examining reviews can be simplified by observing the star ratings provided by users. However, these star ratings alone do not provide a comprehensive understanding of the complete review content [3]. Review analysis can be done manually by looking at reviews one by one, but if there are too many reviews then we can't do it manually there is a faster and more efficient way to use automated systems such as sentiment analysis [4].

Sentiment analysis is the process of identifying, extracting, and evaluating sentiments or opinions expressed in text or data, particularly in the context of reviews or opinions. The goal of sentiment analysis is to understand people's viewpoints, feelings, or attitudes towards a specific topic or entity. Sentiment analysis can be done manually by humans or automatically using computational algorithms and techniques, such as natural language processing and machine learning. In practical applications, sentiment analysis is often used in customer surveys, social media monitoring, product review analysis, and public feedback analysis to gain insights into perceptions and opinions.

This research was conducted because there are so many reviews given by shopee application users but these reviews are not utilized as well as possible. Even the review is only limited to a review without any action, this causes the shopee application rating to drop on PlayStore.

In this research, we will conduct sentiment analysis on reviews of the Shopee application accessible on the Google Play Store. The goal is to evaluate Shopee app users by utilizing the Naive Bayes technique with TF-IDF weighting to classify them into negative and positive categories. Furthermore, this study seeks to assess the precision of the Naive Bayes algorithm in performing sentiment analysis of Shopee application reviews.

Furthermore, as for previous research related to this research, it will be used as the author's research material and used as a reference. In the following, research that discusses sentiment analysis in reviews.

First research was carried out by arif rahman, Ema Utami, Sudarmawan come Amikom University Yogyakarta in 2021 entitled "Analysis of Sentiment on Google Play Store Apps utilizing Naive Bayes Algorithm and Genetic Algorithm" aims for determining level accuracy between the two algorithms then for the dataset used comes from several applications such as shopee, ruang guru, pedia Shopee and gojek. In this study the pedia store dataset have the highest accuracy rate of 96.87 percent, then for Shopee obtained an overall accuracy 96.53 percent, Ruangguru obtained the 95.54%, and Gojek obtained an accuracy of 96.54%. The research collected datasets using the web scrapper application from Google Chrome then the labeling was done manually in the Ms Excel application [5].

Second research was conducted by dedi darwis, nery siskawati, zaenal abidin from teknokrat "In 2021, the University of Indonesia conducted research titled 'Utilization of the Naive Bayes Algorithm for Analyzing Sentiments in National BMKG Twitter Data'. Python 3.74 was employed as the primary programming tool, and a dataset comprising 1179 tweets was used for testing. The data was divided into two categories: training data and testing data, with a ratio of 70:825. It's worth noting that the dataset was imbalanced, with more positive reviews than negative ones. Notably, the author opted not to perform feature normalization, resulting in relatively low accuracy test results when applying the Naive Bayes method for classification namely 69.97% [6].

Third research was conducted by styawati, nirwana hendrastuty, auliya Rahman isnain, ari yanti rahmadhani from teknokrat university of Indonesia in 2021 entitled "Analysis of Public Sentiment Towards the Pre-Employment Card Program on Twitter with the Support Vector Machine Method". Aims to analyze public opinion with data obtained on twitter social media using Support Vector Machine (SVM). Then to measure the performance of SVM classification, the Confusion Matrix method is used.

In this study, two kernels were compared, namely linear and RBF kernels. The evaluation results show that the precision of the linear kernel is 98.67%, the precision is 98%, the recall is 99%, and the F1 score reaches 98%. Meanwhile, the precision of the RBF kernel is 97%, the recall is 98.67%, and the F1 score is also 98%. The accuracy of the RBF kernel reached 98.34%. From these results, it can be concluded that the public opinion of Twitter users towards the pre-employment card program during the pandemic tends to be neutral, with a neutrality rate of 98.34%. When viewed from the accuracy aspect, the linear kernel is more accurate than the RBF kernel with an accuracy of 98.67% and 98.34% respectively [7].

Fourth studies was carried out with the aid of Billy, Helen, Enda Esyudha retrieved tanjungpura university in 2018, discussing Indonesian online product opinions to achieve records such as sentiment records that's a part of the web evaluation. The information is run through a classification process the use of the Naive Bayes approach. The sentiment analysis gadget is going via 5 levels, namely records series, pre-processing, phrase weighting, model building, and sentiments class. within the phrase weighting degree, the TF-IDF (term Frequency-Inverse document Frequency) approach usage. Facts is assessed into five lessons: strongly bad, terrible, neutral, tremendous, and strongly advantageous. moreover, the statistics might be evaluated the usage of a confusion matrix by using measuring accuracy, bear in mind, and precision. The take a look at outcomes show that in the 3-magnificence take a look at (bad, neutral, and superb) the first-rate performance is completed with ninety% education facts and 10% test data, ensuing in an accuracy of 78%, bear in mind of 93.33%, and precision of 77%. inside the 5-class take a look at, the high-quality performance turned into also finished with 90% schooling statistics and 10% take a look at information, with accuracy reaching 59.33%, remember 58.33%, and precision 59.33% [8].

The fifth study conducted by Apriani & Gustian in 2019 aims to evaluate the performance of accuracy, class recall, and AUC of sentiment analysis using the Tokopedia application by applying the Naive Bayes algorithm. The results of this study show that the precision value for class 1 is 1. On the other hand, in class recall, the result is 95.49% (positive class: negative class) and the AUC value is 0.980. From the analyzed reviews, it can be seen that comments with negative sentiments have a percentage of 63.53%, while comments with positive sentiments are only about 36.37% [9].

The sixth research in 2022 entitled "Two-Way CNN-RNN Architecture with Enhanced Group-Wise and Attention Mechanism for Cryptocurrency Sentiment Analysis". This research uses Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and attention mechanism methods to perform sentiment analysis related to

2.4. Data Mining

Data mining is the step-by-step process of extracting relevant, hidden, and potentially valuable information from large and complex data sets [12]. The main purpose of data mining is to uncover new knowledge and information that can provide added value or competitive advantage in a domain or industry. This research uses one of the existing algorithms in data mining, namely naive baiyes.

Naive baiyes is a working method for the most popular classifiers with a good average accurization rate. Naive Baiyes classification technique is based on simple probabilities and is designed to be used with the assumption of independence between explanatory variables. This algorithm emphasizes probability learning, here is the formula of naive baiyes[13].

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{2}$$

This formula explains that X is evidence, H the hypothesis, then P(H | X) is part of the probability that hypothesis H is true if accompanied by evidence X, or vice versa, P(H | X) enters the probability of hypothesis H in the presence of condition X. P(X | H) can be interpreted as the probability of evidence X in the attendance of hypothesis H, then P(H) is the initial ability of hypothesis H, while P(X) is the initial probability of evidence X..

3. RESULTS AND DISCUSSION

3.1. Data Selection

1. Data Scraping refers to the process of extracting information or data from a source or website. This is done by using algorithms or specialized software that automatically extracts data from web pages and stores it in a format that can be further used and analyzed. Data scraping methods can be done manually, where a person manually copies and pastes information from a website into a file or database, or automatically using specialized software or scripts designed to automatically retrieve data from web pages[14]. In this research, scraping shopee application review data is done automatically using the API retrieved from Google Play Store. The results of this data collection are 1000 data, then the data that has been taken is stored in a csv format file.

```
#scrape ulasan
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.shopee.id',
    lang='id',
    country='id',
    sort=Sort.MOST_RELEVANT,
    count=1000,
    filter_score_with=None
)
```

Figure 2. Script for scraping data

	userName	score	at	content
0	Alda EnFebriant	4	2023-09-13 13:47:37	Sangat baik. Sebagai saran. Untuk list produk ...
1	Aizzan RaQilia	2	2023-09-04 15:04:57	Aplikasi lemot, terlalu banyak iklan, sekarang...
2	Arif Fathul Andalus	1	2023-08-29 03:13:23	Aplikasi pengganggu, neipon berkali-kali menaw...
3	Francisca	1	2023-09-12 06:20:07	Fitur beri tip untuk kurir shopee express regu...
4	Feby Wulandari	1	2023-09-09 08:16:30	Shopee makin kesini makin bad pelayanannya, ap...

Figure 3. Scraping data results

2. Data labeling, also known as data tagging, involves assigning a specific category or label to each data point or sample in a data set. The purpose of data labeling is to give context or meaning to the data so that it can be used to train or test machine learning models[15]. In this research, we divided them into negative and positive labels and the subsequent labelin process is done automatically using the python programming language.

2	Promonya bagus. Tapi tolong jasa kirimnya kebi...	5	Positif
3	Kalau bisa toko online harus bisa bersaing seh...	1	Negatif
4	Akun belanja kesayangan, karena banyak promo d...	5	Positif
5	shopee memang gk jelas,padahal cuman liat aj,t...	1	Negatif
6	Maaf aku kasih bintang 1 soalnya makin lama ko...	1	Negatif
7	Ada beberapa toko yg lama mengantar barang pad...	1	Negatif
8	Sekarang cek barang dan toko lebih mudah. Pili...	5	Positif
9	tadinya males banget belanja via shopee, selal...	5	Positif
10	Aplikasi Shopee mmbuat belanjaku semakin mudah...	5	Positif
11	Makin kesini makin kesana lemot nya gak ketulu...	2	Negatif

Figure 4. Results of labeling negative and positive reviews.

3.2. Pre-Processing

Derived from the completed labeling process, the subsequent section presents the outcomes of the pre-processing phase.

	content	score	Label	text_clean	text_Stopword	text_tokens
0	Selama ini pakai shopee cod sampe nya cepat bg...	2	Negatif	selama ini pakai shopee cod sampe nya cepat bg...	pakai shopee cod sampe nya cepat bgtter km ...	[pakal, shopee, cod, sampe, nya, cepat, bgtter...
2	Promonya bagus. Tapi tolong jasa kirimnya kebi...	5	Positif	promonya bagus tapi tolong jasa kirimnya kebih...	promonya bagus tolong jasa kirimnya kebih perh...	[promonya, bagus, tolong, jasa, kirimnya, kebi...
3	Kalau bisa toko online harus bisa bersaing seh...	1	Negatif	kalau bisa toko online harus bisa bersaing seh...	toko online bersaing sehat harga standart harg...	[toko, online, bersaing, sehat, harga, standar...
4	Akun belanja kesayangan, karena banyak promo d...	5	Positif	akun belanja kesayangan karena banyak promo da...	akun belanja kesayangan promo diskon kadang ap...	[akun, belanja, kesayangan, promo, diskon, kad...
5	shopee memang gk jelas,padahal cuman liat aj,t...	1	Negatif	shopee memang gk jelaspadahal cuman liat ajtau...	shopee gk jelaspadahal cuman liat ajtau nya ud...	[shopee, gk, jelaspadahal, cuman, liat, ajtau,...

Figure 5. Results from pre-processing

3.3. Text Transformation

The text transformation or attribute generation stage is a step in the process of obtaining the required document representation[16]. In this stage, the author performs feature extraction using the TF-IDF method, which is a technique for generating attributes from

documents. This process aims to transform the document text into a numerical representation that can be used for further analysis.

```
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf_vectorizer = TfidfVectorizer()
tfidf_train = tfidf_vectorizer.fit_transform(X_train)
tfidf_test = tfidf_vectorizer.transform(X_test)
print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)

(721,)
(721,)
(181,)
(181,)
```

Figure 6. Weighting TF-IDF

3.4. Data Mining

After all steps completed, calculations will be made to see accuracy rate using naive baiyes algorithmic approach. The accuracy rate when performing sentiment analysis using this method is quite good with an accuracy value of 81%, then there are also results of precision 80%, recall 97%, f1-score 88% of negative comment data and precision 86%, recall 45%, f1-score 60% of positive comment data.

```
confusion_matrix:
[[121  4]
 [ 30 25]]
```

	precision	recall	f1-score	support
Negatif	0.80	0.97	0.88	125
Positif	0.86	0.45	0.60	55
accuracy			0.81	180
macro avg	0.83	0.71	0.74	180
weighted avg	0.82	0.81	0.79	180

Figure 7. Calculation results with the naive baiyes method.

3.5. Frequency of Word Appearance

Frequently occurring words in user reviews can provide information about the app. This time the author will provide an overview of user review based on positive and negative reviews. Visible in Figure 8 and Figure 9.



Figure 8. Result frequency of word appearance reviews positif



Figure 9. Result frequency of word appearance reviews negatif

4. DISCUSSION

Of the findings obtained from the research that has been carried out, namely sentiment analysis using the naive baiyes algorithm, several things can be used as evaluation material, among others, In this study, the Naive Baiyes algorithm proved to be an accurate algorithm because it produced an accuracy value of 81%. The advantage of this research is that it has good accuracy, precision and recall values. Then the calculation of precision, recall and f1-score is also carried out based on positive and negative reviews. Then there is previous research conducting sentiment analysis with several algorithms where the result is the naive baiyes algorithm which has the highest algorithm until the accuracy rate reaches 65,26 % [17]. Therefore, it can be concluded that naive baiyes is accurate enough to be used. The drawback in this study is that it uses a fairly small dataset so that the results are less satisfying, for further researchers can add the number of dataset.

5. CONCLUSION

The research findings indicate that an accuracy rate of 81% was attained. Additionally, there were precision results of 80% and recall of 97%, resulting in an f1-score of 88% for negative comments. For positive comments, the precision was 86%, recall was 45 percent, and f1-score was 60%. The survey employed The TF-IDF combined with the Naive Bayes algorithm weighting technique to categorize user reviews into either negative or positive sentiments. Furthermore, the sentiment analysis of Shopee app users reveals a prevailing negative sentiment towards Shopee's services. This suggests that there are areas where Shopee could improve, particularly in terms of features and other aspects. Based on the analysis, it can be concluded that the data sourced from the Google Play Store platform for the Shopee application indicates a higher frequency of negative reviews compared to positive ones. The negative reviews often revolve around issues with the application's features, shipping process, and frequent slow updates. Conversely, positive feedback frequently highlights various promotions offered by Shopee, such as free shipping, product discounts, and cashback.

BIBLIOGRAPHY

- [1] D. Pranitasari and A. N. Sidqi, "Analisis Kepuasan Pelanggan Elektronik Shopee menggunakan Metode E-Service Quality dan Kartesius," *Jurnal Akuntansi dan Manajemen*, vol. 18, no. 02, pp. 12–31, 2021, doi: 10.36406/jam.v18i02.438.
- [2] N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 1, pp. 47–54, 2022.
- [3] P. Subarkah *et al.*, "SENTIMENT ANALYSIS ON REVIEWS OF WOMEN 'S TOPS ON SHOPEE," vol. 9, no. 1, pp. 126–133, 2023, doi: 10.33480/jitk.v9i1.4179.INTRODUCTION.
- [4] E. H. Muktafin, K. Kusriani, and E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *Jurnal Eksplora Informatika*, vol. 10, no. 1, pp. 32–42, 2020, doi: 10.30864/eksplora.v10i1.390.
- [5] A. Rahman, E. Utami, and S. Sudarmawan, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 1, pp. 60–71, 2021, doi: 10.31603/komtika.v5i1.5188.
- [6] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naïve Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [7] N. Hendrastuty *et al.*, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, pp. 150–155, 2021.
- [8] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naïve Bayes," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 4, no. 2, p. 113, 2018, doi: 10.26418/jp.v4i2.27526.
- [9] R. Apriani *et al.*, "Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia," *Jurnal Rekayasa Teknologi Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019.
- [10] G. C. Habek, M. A. Toçoğlu, and A. Onan, "Bi-Directional CNN-RNN Architecture with Group-Wise Enhancement and Attention Mechanisms for Cryptocurrency Sentiment Analysis," *Applied Artificial Intelligence*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2145641.
- [11] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, "Analisis Sentimen Menggunakan Naïve Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps," *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 4, pp. 28–38, 2022, doi: 10.35969/interkom.v16i4.192.
- [12] P. M. S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, no. 1, pp. 9–19, 2022, doi: 10.25008/janitra.v2i1.142.
- [13] A. Ridwan, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, vol. 4, no. 1, pp. 15–21, 2020, doi: 10.47970/siskom-kb.v4i1.169.
- [14] D. D. A. Yani, H. S. Pratiwi, and H. Muhandi, "Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace," *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, vol. 7, no. 4, p. 257, 2019, doi: 10.26418/justin.v7i4.30930.
- [15] H. Syah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
- [16] L. Nguyen-Thi-Huong, H. Nguyen-Viet, A. Nguyen-Phuong, and D. Van Nguyen, "How does digital transformation impact bank performance?," *Cogent Economics and Finance*, vol. 11, no. 1, 2023, doi: 10.1080/23322039.2023.2217582.
- [17] J. Muliawan, E. Dazki, and R. D. Kurniawan, "SENTIMENT ANALYSIS OF INDONESIA'S CAPITAL CITY RELOCATION USING THREE ALGORITHMS: NAÏVE BAYES, KNN, AND RANDOM FOREST," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 5, pp. 1227–1236, 2023, doi: 10.52436/1.jutif.2023.4.5.347.