Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification

Marchenda Fayza Madjid[1)*], Dian Eka Ratnawati[2)], Bayu Rahayudi[3)]
[1,2,3)]Universitas Brawijaya, Malang, Indonesia
[1)] fayzaaa_@ub.ac.id, [2)] dian_ilkom@ub.ac.id, [3)] ubay1@ub.ac.id

Submitted : Jan 25, 2023 | **Accepted** : Jan 31, 2023 | **Published** : Feb 2, 2023

Abstract: A review is an assessment given by someone based on certain aspects, such as the delivery of stories, pictures, effects, or visuals. Users can provide reviews which help the company know the quality of the application. However, reviews cannot be used as a reference for rating, because there are still users who provide reviews that are irrelevant to the rating given. This study aims to carry out sentiment analysis in order to classify the application user review data. The sentiment classification process begins with collecting and labeling 700 data. The data then goes through a text preprocessing, word weighting with TF-IDF, and classification using the Support Vector Machine and Naïve Bayes Classifier. The results produce the highest accuracy in the comparison of training and test data of 90%:10%. Support Vector Machine algorithm is capable of providing high accuracy with RBF kernel, γ=1, and C=10. The results obtained using 10-fold cross validation give an accuracy value of 92.86%, a precision value of 92.88%, a recall value of 92.88%, a specificity value of 94.73%, and f-measure of 92.76%. Naïve Bayes Classifier method is able to provide high accuracy by using Multinomial Naïve Bayes Classifier. The results obtained using 10-fold cross validation give an accuracy value of 92.54%, a precision value of 92.55%, a recall value of 92.51%, a specificity value of 93.9%, and f-measure of 92.44%. Based from the result, it can be concluded that the classification using the Support Vector Machine is superior to the Naive Bayes Classification.

Keywords: Naïve Bayes Classifier; Reviews; Sentiment Analysis; Support Vector Machine; Text Preprocessing; TF-IDF

INTRODUCTION

The Finder Institute conducted a survey on the number of digital banking users in 2021 and showed that by 2026, as many as 39% of adults in Indonesia or 74,785,062 people are predicted to use digital banking applications (Laycock, 2021). A survey conducted by (PwC Indonesia, 2019) concluded that the increasing number of digital banking application users has encouraged banking companies to start developing digital banking applications. This research will implement sentiment analysis on the Allo Bank application.

The results of the investigation carried out by the author provide the result that there are still many users who complain and have difficulties with registration, gifts that are not sent, and QRIS scanning. The users also consider that the application feels too fast for release, even though there are still features that are not ready. Besides that, reviews on the Google Play Store cannot really be used as a reference for companies to assess the service of the applications they make. This happens because there are still users who give a 5 (highest) star rating, but with negative or irrelevant reviews for the rating given.

Based on the problems that have been described, sentiment analysis is needed to classify Allo Bank application user review data based on reviews on the Google Play Store. Sentiment analysis has the goal of obtaining classification results from a review, both negative and positive reviews (Brahimi et al., 2021). The classification algorithm methods used are Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC). Previous research was chosen with the topic of sentiment analysis which compared the accuracy values of the two classification algorithms, namely the Support Vector Machine and the Naïve Bayes Classifier. Previous research conducted by Mustakim & Priyanta (2022), regarding sentiment analysis reviews on the KAI Access application made a comparison of the two classification methods, namely NBC and SVM. The conclusion from research conducted by Mustakim & Priyanta (2022), is that classification using the SVM algorithm is capable of

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

producing superior accuracy. The resulting accuracy is 91.63%, compared to the Naïve Bayes Classifier method which gives an accuracy value of 83.47%. Another research was done by Kristiyanti, et al (2019) concluded that the Naïve Bayes Classifier method is able to provide superior accuracy values. The resulting accuracy is 94%, compared to the Support Vector Machine classification algorithm which produces an accuracy of 75.50%. Research by Kristiyanti, et al (2019) has an object that is related to sentiment analysis on the topic of public elections for the election of candidates for Governor of West Java for the 2018-2023 period on Twitter.

Based on the background that has been described, the author intends to conduct a research on the topic of sentiment analysis of Allo Bank application users obtained from Google Play Store review data. The classification method that the author uses is the Support Vector Machine classification algorithm and the Naïve Bayes Classifier. The use of the two methods aims to compare the level of accuracy between the two classification algorithms to be used. A comparison of the two algorithms was also carried out because in previous studies, there were cases where the SVM algorithm provided superior accuracy compared to NBC, and vice versa. The collected data will then be processed by TF-IDF weighting. The use of the TF-IDF weighting method is intended to give each word a weight that can determine the amount of calculation of how far the relationship between words is. This research is expected to be used by PT. Allo Bank Indonesia Tbk to improve quality and service and serve as the basis and reference for the company PT. Allo Bank Indonesia Tbk to innovate and improve in the future.

LITERATURE REVIEW

Term Frequency – Inverse Document Frequency

According to Qaiser & Ali (2018), the weighting of the word is a stage that aims to provide indicators to each word that is adjusted to the level of categories of each word. The words in the data will be converted into numeric so that each word can be weighted from one document to another(Amrizal, 2018). One of the most widely used methods in research is the Terquency-Inverse Document Frequency (TF-IDF) method. TF-IDF is a combination method of two word weighting methods, namely Terquency (TF) and Inverse Document Frequenzy (IDF). Formula Calculating Terquency-Inverse Document Frequency (TF-IdF) can be seen in equation 2.1.

$$tf - idf_{(t,d)} = \frac{f_{k,d}}{N_d} \times \log \frac{N}{df_{k,N}} \quad (2.1)$$

Support Vector Machine

Support Vector Machine is a classification algorithm that searches for the maximum maximum marginal hyperplane (MMH) or the best separating hyperplane by dividing the two classes in the input space. The level of accuracy in the SVM method is influenced by the use of kernel parameters and functions.

The training dataset for this algorithm is described in the form (xi, yi), where xi is a tuple and yi is illustrated as a label having class i = 1, ..., N. This method has a formula that can be seen in equation 2.2 and equation 2.3(Mahendrajaya et al., 2019)..

$$f(x) = wx + b \quad (2.2)$$

$$f(\mathrm{x}) = \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a_i y_i K(x,x_i)} + \mathrm{b} \quad (2.3)$$

Naïve Bayes Classifier

Naïve Bayes Classifier is a statistical classifier algorithm that predicts the probability of membership in a particular class (Achmad & Girsang, 2020). There are two stages in carrying out classification processing on text, including the training and testing stages. In the training phase, word analysis processing occurs which will always appear in documents that can represent a document. After that, the probability determination for each category will be carried out. The testing phase is carried out by determining the category value of a document for the word that appears the most in the document(Jurafsky & Martin, 2014). The formula for performing classification calculations can be seen in equation 2.6.

$$P(j|t) = P(j) \times P(t_1|j) \times P(t_2|j) \times \ldots \times P(t_n|j) \quad (2.4)$$

METHOD

In conducting this research, the stages of the research were divided into 8 stages which can be seen in Figure 3.1.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

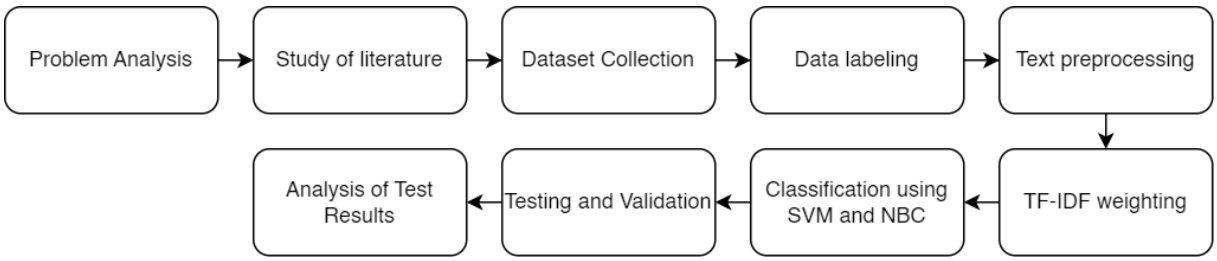

Fig. 1 General Research Method

The research stage begins with analyzing the problem and then looking for a solution to the problem. The results of problem identification will be illustrated in an Ishikawa Diagram (Fishbone). The next stage is followed by a literature study, then dataset collection from user reviews data for the Allo Bank application from the Google Play Store. The data collection (scraping) stage will be carried out using the google-play-scraper library. The data that has been successfully collected will go through a data labeling process to manually identify each comment. This data labeling process will be carried out by 3 actors with Indonesian language expert backgrounds. The data will then go through a text preprocessing process, TF-IDF weighting, classification by implementing the Support Vector Machine and Naïve Bayes Classifier algorithms. Finally, the data will go through a testing process which contains a validation process with 10-fold cross validation.

RESULT

In this section, the researcher will explain the results of the research by excecuting the testing scenario of classification using both Support Vector Machine and Naïve Bayes Classifier. Tests are carried out to obtain a large level of accuracy of a system using different types and parameters. In this study, the design of the test scenarios carried out was divided into two parts, namely the test scenarios which were carried out using the Support Vector Machine algorithm and the test scenarios which were carried out using the Naïve Bayes Classifier algorithm. Prior to classification, the dataset will go through a text preprocessing process to remove the noise in each review sentences. The cleaned sentences then will go through a word weighting process using TF-IDF that can be automated calculate with the help of a library called Sklearn using class TFIDFVectorizer. This library is already doing an automated mathematic calculation using a formula that can be seen on Equation 2.1.

Both classification will be calculated automatically with the help of the Sklearn library that calls svm function for SVM and naive_bayes function for Naive Bayes Classifier. In SVM calculations using the Sklearn library, it already includes the calculations in Equation 2.2, and customizable kernel types $K(x, x_i)$ will be calculated and included in Equation 2.3. Same implementation with SVM, the calculation of Naive Bayes Classifier using the Sklearn library will also automatically calculated the prior value ($P(j)$), likelihood value ($P(t_1 \mid j) \times P(t_2 \mid j) \times \ldots \times P(t_n \mid j)$), just like what was written on Equation 2.4. Testing using the Support Vector Machine will be divided into 4 test scenarios. The test is a comparison test of training data and test data, SVM kernel comparison test, parameter γ and parameter C testing. Furthermore, testing using the Naïve Bayes Classifier algorithm will be divided into 2 test scenarios, namely testing comparison of training data and test data, as well as testing using different types of Naïve Bayes Classifier.

Table 1 Comparison of Accuracy Between Types of SVM Kernels

Kernel Types	Means
Linear	93.6%
Polynomial	93.9%
RBF	94.3%

Table 2 Accuracy Testing Result Using γ=1 and C=10 Parameters

γ=1 and C=10	Trial Accuracy Values										
	1	2	3	4	5	6	7	8	9	10	Means
	95.7%	91.4%	92.9%	94.3%	100%	95.7%	94.3%	94.3%	98.6%	94.3%	95.3%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 3 10-Fold Cross Validation Test Results on Support Vector Machine

Fold Values	Accuracy	Precision	Recall	Specificity	F-Measure
1	98.41%	98.39%	98.49%	100%	98.41%
2	90.48%	90.46%	90.46%	90%	90.46%
3	90.48%	90.15%	91.32%	96.15%	90.36%
4	90.48%	89.9%	89.9%	92.31%	89.9%
5	93.65%	94.6%	93.33%	100%	93.57%
6	92.06%	92.09%	92.09%	90.63%	92.06%
7	95.24%	95%	95.37%	96.3%	95.16%
8	95.24%	95.3%	95.15%	96.97%	95.22%
9	93.65%	93.6%	93.86%	91.18%	93.64%
10	88.89%	89.29%	88.81%	93.75%	88.84%
Means	92.86%	92.88%	92.88%	94.73%	92.76%

Table 4 Comparison of Accuracy Between Types of Naïve Bayes Classifier

Types	Means
Gaussian	86%
Multinomial	94.9%

Table 5 10-Fold Cross Validation Test Results on Naïve Bayes Classifier

Fold Values	Accuracy	Precision	Recall	Specificity	F-Measure
1	95.24%	95.21%	95.3%	96.67%	95.23%
2	90.48%	90.71%	90.3%	86.67%	90.42%
3	87.3%	86.79%	87.47%	88.46%	87.04%
4	92.06%	91.37%	91.99%	92.31%	91.65%
5	93.65%	94.6%	93.33%	100%	93.57%
6	90.48%	90.67%	90.42%	93.75%	90.46%
7	95.24%	95%	95.37%	96.3%	95.16%
8	95.24%	95.33%	95.15%	96.97%	95.22%
9	95.24%	95.15%	95.33%	94.12%	95.22%
10	90.48%	90.67%	90.42%	93.75%	90.46%
Means	92.54%	92.55%	92.51%	93.9%	92.44%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

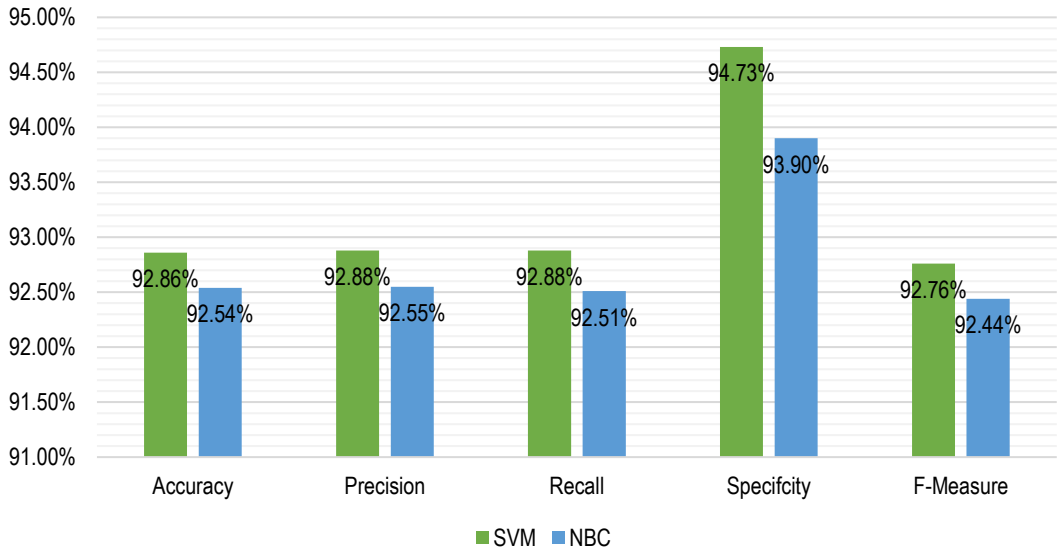

Fig. 2 Comparison of SVM and NBC Classification Results

Fig. 3 Word Cloud Visualization of Positive Label Using WordCloud Library on Google Colaboratory

Fig. 4 Word Cloud Visualization of Negative Label Using WordCloud Library on Google Colaboratory

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.12161
e-ISSN : 2541-2019
p-ISSN : 2541-044X

DISCUSSIONS

Testing and analysis is carried out to calculate the accuracy value of the classification that has been applied. The test will calculate the accuracy of the Support Vector Machine and Naïve Bayes Classifier classification algorithms. The test will be categorized using data labeled with two sentiments, namely data that enters the positive class and data that enters the negative class. The dataset used in the testing process is 700 data with a division of 350 classes with positive labels and 350 classes with negative labels. The testing process is carried out 10 times and then the average value will be taken.

In Table 1, the accuracy of the highest predicted score is generated using the RBF kernel, which is 94.3%, while in the Polynomial kernel and Linear kernel, the accuracy obtained is 93.9% and 93.6% respectively. Based on the accuracy results obtained, it can be concluded that the RBF kernel can provide higher accuracy. The RBF kernel can produce a high level of accuracy because it can move data into a high-dimensional feature space that depends on the magnitude of C and γ. Optimal C and γ parameters that will help this kernel provide high accuracy.

In Table 2, the highest prediction score accuracy is at γ=1 and C = 10, which is 95.3%. Based on the accuracy results obtained, it can be concluded that in the research dataset, the high accuracy value of the prediction process using the SVM algorithm can be influenced by the use of optimal γ and C values, namely 1 for γ and 10 for C.

Based on the test results that can be seen on Table 3 dan Table 5, it can be concluded that the results of the accuracy from both Support Vector Machine and Naive Bayes algorithm model, which is 92.86% and 92.54%, are considered to have very good performance in the classification process of the Allo Bank application review dataset.

In Table 4, the highest prediction score accuracy is generated using Multinomial NBC, which is 94.9%, while for Gaussian NBC, the accuracy obtained is 86%. Based on the accuracy results obtained, it can be concluded that Multinomial NBC can provide much higher accuracy than using Gaussian NBC. Multinomial NBC can provide high accuracy because it is suitable for document classification problems, which are able to sort out which documents belong to a certain class with discrete data types.

Based on the test results, there are several factors that affect the accuracy of the Support Vector Machine and Naïve Bayes Classifier algorithms. In the Support Vector Machine algorithm, the factors that affect the accuracy value are differences in the amount of training data and test data, differences in kernels, and differences in the use of parameters γ and C. Large amounts of training data, use of RBF kernels, use of parameters γ and C are optimal. greatly increase the accuracy of the SVM algorithm. In the Naïve Bayes Classifier algorithm, the factors that affect the accuracy value are the difference in the amount of training data and test data, and the use of the NBC type of algorithm. The large amount of training data and the use of the NBC type, namely the Multinomial Naïve Bayes Classifier, can provide a high degree of accuracy.

In Figure 2, the highest accuracy value is 92.86% in the classification method with SVM. Based on the tests carried out, it can be concluded that the level of performance accuracy for classifying using SVM has a higher magnitude when compared to using the NBC algorithm. SVM is superior because the concept of the SVM algorithm itself can visualize linear data into high dimensions where data distribution and predictions will certainly be good. The use of the RBF kernel which uses optimal C and γ values in the testing process can also help in increasing the accuracy of this algorithm.

In Figure 3, it can be seen that for sentiments that are categorized as positive the words that often appear are “mudah” 117 times, “bagus” 85 times, “bantu” 54 times, “mantap” 51 times, and “diskon” 52 time. The positive label means that there are users of the Allo Bank application who provide support and like the services provided by Allo Bank. A positive label also means a user's reaction or attitude towards a product or service that tends to be enthusiastic, happy and excited.

In Figure 4, it can be seen that in sentiment with a negative label category, the words that often appear are “tidak” 307 times, “masuk” 130 times, “salah” 62 times, “susah” 61 times, and “daftar” 58 times. The negative label means that there are users of the Allo Bank application who react to a product or service that tend to be angry, annoyed, and frustrated.

CONCLUSION

The sentiment analysis process is carried out using the Support Vector Machine and Naïve Bayes Classifier algorithms. The analysis process begins by scraping data on reviews of the Allo Bank application on the Google Play Store, then the data will be processed manually. Second, the data will go through a text preprocessing process and word weighting using TF-IDF, then the weighting results will be entered into two algorithm models, namely Support Vector Machine or Naïve Bayes Classifier. Sentiment classification will also include several factors or parameters for comparison. In the Support Vector Machine method, the accuracy value is 94.29%, the precision value is 94.48%, the recall value is 94.19%, the specificity is 95.27%, and the f-measure is 94.18%. In

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

the Naïve Bayes Classifier method, the accuracy value is 93.97%, the precision value is 93.92%, the recall value is 94.04%, the specificity is 93.52%, and the f-measure is 93.94%. Based on the tests that have been carried out, the two algorithms can provide very good performance in classifying Allo Bank application reviews. In addition, obtained the results of a comparison between the two classification methods that have been done. Classification using the Support Vector Machine algorithm produces superior accuracy values than using the Naïve Bayes Classifier algorithm.

REFERENCES

Achmad, R., & Girsang, A. S. (2020). Implementation of naive bayes classifier algorithm in classification of civil servants. *Journal of Physics: Conference Series*, *1485*(1). https://doi.org/10.1088/1742-6596/1485/1/012018

Amrizal, V. (2018). PENERAPAN METODE TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DAN COSINE SIMILARITY PADA SISTEM TEMU KEMBALI INFORMASI UNTUK MENGETAHUI SYARAH HADITS BERBASIS WEB (STUDI KASUS: HADITS SHAHIH BUKHARI-MUSLIM). *JURNAL TEKNIK INFORMATIKA*, *11*(2). https://doi.org/10.15408/jti.v11i2.8623

Brahimi, B., Touahria, M., & Tari, A. (2021). Improving sentiment analysis in Arabic: A combined approach. *Journal of King Saud University - Computer and Information Sciences*, *33*(10). https://doi.org/10.1016/j.jksuci.2019.07.011

Jurafsky, D., & Martin, J. (2014). Speech and Language Processing. In *Speech and Language Processing.* (Vol. 3).

Kristiyanti, D. A., Umam, A. H., Wahyudi, M., Amin, R., & Marlinda, L. (2019). Comparison of SVM Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter. *2018 6th International Conference on Cyber and IT Service Management, CITSM 2018*. https://doi.org/10.1109/CITSM.2018.8674352

Mahendrajaya, R., Buntoro, G. A., & Setyawan, M. B. (2019). ANALISIS SENTIMEN PENGGUNA GOPAY MENGGUNAKAN METODE LEXICON BASED DAN SUPPORT VECTOR MACHINE. *KOMPUTEK*, *3*(2). https://doi.org/10.24269/jkt.v3i2.270

Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, *181*(1). https://doi.org/10.5120/ijca2018917395

Laycock, R. (2021, October 7). *Penggunaan Perbankan Digital 2021*. Retrieved August 17, 2022, from Finder: https://www.finder.com/id/neo-bank

Mustakim, H., & Priyanta, S. (2022, April 30). Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 16 (2)*, 113-124. doi:https://doi.org/10.22146/ijccs.68903

PwC Indonesia. (2019). *PwC Survey: Digital Banking in Indonesia 2018.* Retrieved January 7, 2023, from pwc: https://www.pwc.com/id/en/pwc-publications/industries-publications/financial-services-publications/digital-banking-survey-2018.html

*name of corresponding author