Journal of Computer Science

Original Research Paper

Sentiment Analysis on User Reviews of Mutual Fund Applications

[1]Evaristus Didik Madyatmadja, [1]Shinta, [1]Devi Susanti, [1]Florencia Anggreani and [2]David Jumpa Malem Sembiring

[1]*Department of Information Systems, Bina Nusantara University, Jakarta, Indonesia*
[2]*Teknik Informatika, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia*

Article history
Received: 09-03-2022
Revised: 23-06-2022
Accepted: 23-08-2022

Corresponding Author:
Evaristus Didik Madyatmadja
Department of Information Systems, Bina Nusantara University, Jakarta, Indonesia
Email: emadyatmadja@binus.edu

Abstract: The primary goal of this study is to compare the accuracy of the results of sentiment analysis using the Naive Bayes, Support Vector Machine (SVM), and Random Forest methods on one of the mutual fund application's user reviews. The second goal is to identify user reviews of the mutual fund app to gain insight into the topics covered by each sentiment. The user reviews have been collected through a web scraping method on the google play store, then cleaned through several processes of data pre-processing. Feature extraction was performed using TF-IDF along with vectorization using n-grams. The model performance was measured using a confusion matrix. Using a ratio of 80:20 on training and testing data, resulting in an accuracy of 92.7, 93.7 and 94.2% for Naive Bayes, SVM, and Random Forest methods, respectively. Identify the topics covered by each sentiment in user reviews using visualizations. In the positive sentiment of users, the majority discusses the application which is easy and good, especially for novice investors. In negative sentiment, the majority discussed the slow sales process to disbursement of funds and long loading times when opening the application.

Keywords: Sentiment Analysis, User Reviews, Naive Bayes, SVM, Random Forest

Introduction

Investments have been made by humans for a long time consciously to expect results in the future or unconsciously. There are various types of investments in goods or things that investors think is valuable and will be an advantage in the future. The investment allows a person to analyze the benefits and risks of a particular object and put his capital with the belief that what he spends at that time will increase.

The KSEI data shown in Fig. 1, obtained from the publication file on the official website of the PT Kustodian Sentral Efek Indonesia (KSEI, 2021), shows the growth in the number of Indonesian capital market investors from 2018 to August 2021. From the data, it can be seen that the number of mutual fund investors skyrocketed 71% to 5.44 million people. Compared to the end of 2018, which was only 995,510 investors, the number of mutual fund investors in August 2021 has increased more than 5 times. The surge in the number of mutual fund investors shows that the COVID-19 pandemic has accelerated the number of mutual fund investors in line with the optimization of digital channels by industry players. The increasing growth of mutual funds is also due to the convenience offered because they are not managed alone, but are managed by more experienced Investment Managers, so this attracts a lot of interest for beginners to invest in mutual fund instruments.

The Google Play Store allows users to share their experiences with apps in the form of app reviews, which can motivate or discourage other users from downloading the app. Such reviews can also provide valuable information to app developers. Therefore, there is a need to understand user reviews to survive in today's highly competitive mobile app industry (Khalid *et al*., 2015).

According to Liu (2012), sentiment analysis or opinion mining is a field of study that analyzes opinions, sentiments, evaluations, judgments, attitudes, and emotions towards an entity, such as products, services, organizations, individuals, problems, or topics. According to Thakkar and Patel (2015), in the field of sentiment analysis, texts can be classified as positive or negative. There may be multi-valued or binary classes such as positive, negative, and neutral (or irrelevant). This class classification can be used to determine the opinion of most users of the mutual fund application.

According to Aqlan *et al*. (2019), with the huge explosion in information in recent years in the sites of communication, air traffic, and alternative markets, all this huge amount of information cannot be controlled and analyzed using the traditional way, so the scientists and researchers developed a high-efficiency technique to deal with this data. This requires sentiment analysis to process data and know its polarity to determine the right decision.

Research by Himawan and Eliyani (2021) aims to determine the sentiment of public tweet data on the official Twitter account of the DKI Jakarta Provincial Government during the COVID-19 pandemic. This study uses the TF-IDF Vectorizer for word weighting and classification using several methods, namely, random forest classifier with 75.81% accuracy, Naive Bayes algorithm with 75.22% accuracy, and support vector machine algorithm with the highest accuracy score of 77.58%.

Research by Khan and Malik (2018) revealed the sentiment classification of user reviews of various car sales sites in Roman Urdu using Multinomial Naïve Bayes, Bagging, Deep Neural Network, Decision Tree, Random Forest, AdaBoost, k-NN, and SVM Classifiers. The results of this study indicate that the Multinomial Naïve Bayes method has the best performance among all other classifiers, with a precision level of 93%, recall of 96%, f-measure of 95%, and accuracy of 90%.

Another study by Fitri (2020) conducted sentiment analysis on the Ruangguru application by testing three classification models such as Naive Bayes, Random Forest, and Support Vectors Machine. The results showed that the model with the best performance was Random Forest compared to other models with an accuracy rate of 97%.

Based on several previous studies, it was found that there was a gap in the results of the research on the sentiment analysis method that produced the highest accuracy.

The existence of this research gap and the need to understand user reviews makes the authors interested in conducting a comparative study of sentiment analysis methods on the opinions of mutual fund application users using the Naive Bayes, Support Vector Machine (SVM), and Random Forest methods, and also identified user reviews of the mutual fund app to gain insight into the topics covered by each sentiment.

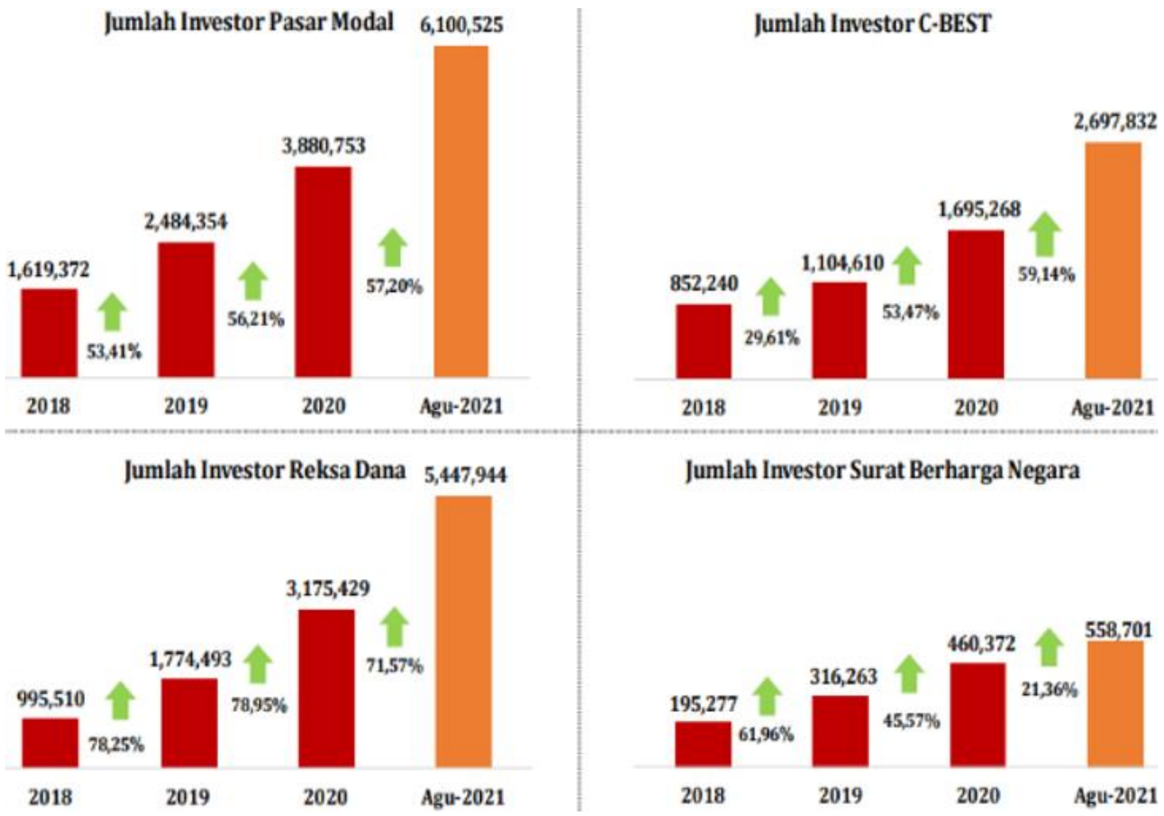

Fig. 1: Indonesian investor growth chart (KSEI, 2021)

A. *Sentiment Analysis*

Sentiment Analysis or Opinion Mining is used to learn user behavior by identifying, analyzing, and extracting subjective text that consists of opinion, preferences, and user sentiment (Nanli *et al*., 2012). There are 2 types of sentiment analysis approach, Lexicon Based Approach, and the Machine Learning Approach.

– Lexicon Based Approach

Lexicon-based approaches use lexicon sentiment to describe polarity (positive, negative, and neutral) from the textual content. This approach is easier to understand and implement (Sadia *et al*., 2018)

– Machine Learning Approach

The machine learning approach is a collection of methods that enable computers to automate data-driven model building and programming through the systematic discovery of statistically significant patterns in the available data (Bhavsar *et al*., 2017)

B. *Text Preprocessing*

Text preprocessing is an important part of the NLP system because the identified characters, words, and sentences in this stage are fundamental units passed for further processing stages. Sometimes, text data contains some special formats, such as number formats, date formats, and general words that are unlikely to help text mining. By passing the text preprocessing stage, those words can be removed (Gurusamy and Kannan, 2014). Generally, text preprocessing stages are case folding, stop words removing, tokenizing, and stemming (Jumeilah, 2017).

C. *Feature Extraction*

Feature extraction is the process of selecting a set of features by using some effective methods to reduce the dimension of feature space. During feature extraction, there will be a document selection stage to choose the part of the document that reflects information on the content, and weight calculation. As a data preprocessing method of learning algorithms, feature extraction can improve the accuracy of learning algorithms and shorten the processing time (Liang *et al*., 2017).

D. *Text Classification*

Text Classification is the task of dividing a set of documents into several classes where each document can belong to one or multiple classes (Sebastiani 2002). Text classification assigns a collection of text into categories that have been predefined before. To perform the process of text classification, inputted data are divided into two parts: Training data and testing data.

There are two approaches to data splitting that researchers commonly use. First Holdout Method, this method is simply dividing a set of data into two subsets usually with a 2:1 ratio, 2 for training data and 1 for testing data. And the other method is K-Fold Cross Validation, data randomly assigned into K groups of the same size. Each K group is chosen to be testing data and others are being training data. This method takes some time to finish because the process is repeated.

Text classification techniques used in this study:

1. Naive Bayes

Naive Bayes is a classification technique that uses probability and statistics, introduced by Thomas Bayes. This technique predicts future chances based on experience before. Naive Bayes classifies documents into certain classes using a probabilistic model with an independent assumption of the term. The advantage of using naïve Bayes is that it only requires a small amount of training data to determine the estimated parameters needed in the classification process (Madyatmadja *et al*., 2019).

Multinomial Naive Bayes is one of the classification models from the Naive Bayes method. This model captures terms' frequency by representing documents using Bag of Words. Documents are treated as a set of words with the frequency of each word:

2. Support Vector Machine (SVM)

The main objective of the SVM technique is to decide the hyperplane that will divide classes. Classes are usually symbolized as -1 and +1. The figure below illustrates the hyperplane, the red square is class -1 and the yellow circle is class +1.

Figure 2(a) shows an alternative dividing line between the two classes (discriminant boundaries). The best dividing line is the one with the maximum hyperplane margin. Margin is the distance between the hyperplane and the closest pattern in each class. The closest pattern is called a support vector. In Fig 2(b), the circled pattern is the support vector for each class. Meanwhile, the thick line in Fig. 2(b) is the best hyperplane because it is in the middle of the two

DOI: 10.3844/jcssp.2022.885.895

classes. The process of finding the location of this hyperplane is the core of the SVM method (Fachruddin, 2015).

Data used in this study can be linearly separable, this study uses the Linear Kernel model to transform instances into hyperplanes to maximize the distance between Training Samples:

3. Random Forest

This technique is an ensemble. The base classifier of Random Forest is a Decision Tree, which is a combination of several decision trees (Kulkarni and Sinha, 2014). As its name, RF creates a forest with trees. The accuracy of the RF technique depends on each tree's strength and dependence's measurement of trees. It's efficient on large databases.

The tree was built using information gain and gini index to determine the root node and rule. The more trees created, the stronger a forest in other words accuracy of the model is more accurate (Han *et al*., 2012).

This technique is shaped like a tree structure where there is a roots, branches, and leaves. Each node showed testing for an attribute, each branch represents the result from testing.

E. Confusion Matrix

Confusion Matrix is used to measure the performance of the classification technique. It contains information that compares classification results that've been done by systems with the actual classification result as shown in Table 1. There are four matrixes of Confusion Matrix:

- True Positive (TP), a condition where actual data is in positive class and classification models predict also positive
- True Negative (TN), a condition where actual data is in negative class and classification models predict also negative
- False Positive (FP), is a condition where actual data is in a negative class but classification models predict is positive
- False Negative (FN), a condition where actual data is in positive class but classification models predict is negative

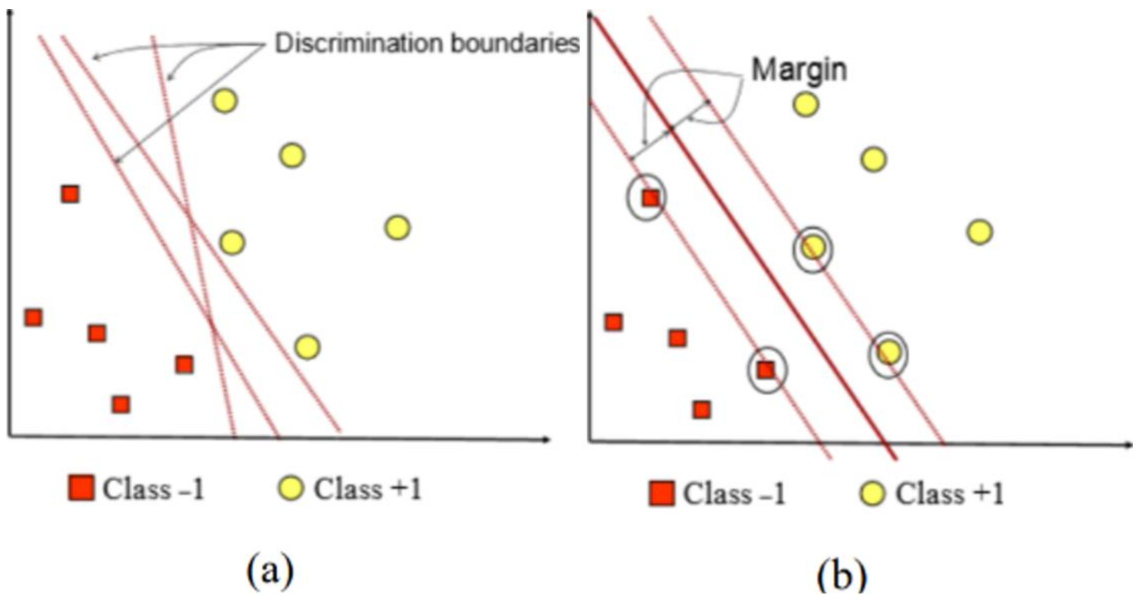

Fig. 2: Hyperplane Illustration (Fachruddin, 2015)

Table 1: Confusion matrix

		Predicted class		
		Yes	No	Total
Actual Class	Yes	TP	FN	P
	No	FP	TN	N
	Total	P'	N'	P + N

F. Matplotlib

A whole reference to make static visualization, animation, and python interactive. Its features are suitable for plotting. The plot is the graphic representation of a collection of data, it shows relationships between variables.

G. Related Works

There are many studies related to sentiment analysis that has been carried out using a machine learning approach. According to Nguyen *et al.* (2018), the Machine Learning approach in the Supervised Learning category is superior to the Lexicon Based approach.

Research by Saepulrohman *et al*. (2021) states that the results of the study will be more accurate if the author avoids words that have the same or double meaning and uses more datasets to increase the accuracy of the data to be studied. The addition of slang dictionary data in the Stop word process, because in the reviews of application users on the Google Play Store there are still many using less standardized language.

Research by de Godoi Brandão and Calixto (2019) showed that the use of N-Gram and TF-IDF for feature extraction for training SVM classifiers had a positive impact on accuracy results when compared to studies using other characteristic extraction methods. No other preprocessing tasks like stemming, stop word removal, noise removal, normalization, etc., are applied. Therefore, the authors suggest analyzing the impact of these tasks in future research.

Materials and Methods

The methodology is used by the author to analyze, work on and solve the problems at hand. The theoretical framework or scientific framework is the scientific method that will be applied in conducting research. The research stage is used, namely studying the literature, collecting data, processing data, analyzing data, and making reports.

Collecting Dataset

Figure 3 shows the data collection methods used. The data collected and used in this research is the review and rating data from the mutual fund application users that is retrieved from the Google Play Store website. This stage consists of a web scraping stage and export data. The web scraping stage is performed by using Python and Google Play Scraper Library. Then, the result of web scraping was saved in xlsx format. Data attributes and the type collected from this method are shown in Table 2.

Text Preprocessing

Figure 4 shows the steps or methods used in text preprocessing. In the review selection stage, web scraping results will be selected manually to reduce the content that is not relevant to sentiment analysis, such as emoticons, referral codes, etc. The data produced in this stage will only contain users' reviews and ratings. The data was retrieved within one year period, starting from 29 September 2020 until 29 September 2021. The data retrieved was only reviews with 1, 2, 4, and 5 ratings, because 3 is considered neutral.

The data selected will be labeled manually, where review rated 1 and 2 will be labeled "Negative" and review rated 4 and 5 will be labeled "Positive".

The review data retrieved has different languages, which are Indonesian and English. So, the data will be converted into Indonesian. Translation to the Indonesian stage was done so the data used only one language, to avoid duplication of review.

In the normalization stage, the review will be normalized manually. The review data consists of abbreviations, slang words, or typo words, this can affect the result of sentiment analysis. The review data will be normalized according to Kamus Besar Bahasa Indonesia (KBBI).

The Case Folding stage will convert words into lowercase and other characters. The case folding stage was performed by the programming language, Python, with the NLTK library.

In the Stopwords Removal stage, common words that do not have any meaning or do not affect the classification will be deleted. The stop words removal stage was performed using Python with the NLTK library.

The review sentences will be broken into smaller units, called words. Each smaller unit is called a token. Tokenization was performed to make the text meaning interpretation easier and as the first stage of the stemming process.

The stemming stage is to convert the words from tokenization into the basic words, by removing the affix on the words. Every word that has already been converted into basic words will be combined back into the sentence.

Features Extraction

Figure 5 shows the method used in feature extraction, which begins with vectorization. In the vectorization stage, we used the N-Gram method to perform feature extraction using TF-IDF. The N-Gram method used is Unigram and Bigram. Unigram is n-gram with n value. One n unit is equal to one n-gram. While in Bigram, one n unit is equal to two n-grams. The combination of Unigram and Bigram used is to increase the accuracy level of the sentiment analysis results.

DOI: 10.3844/jcssp.2022.885.895

The feature extraction used in this research is Term Frequency-Inverse Document Frequency (TF-IDF), which is to calculate the frequency score of each word. TF-IDF describes the importance of a term to a document in a collection or corpus. The implementation of TF-IDF and N-Gram before modeling is intended to give more accurate analysis sentiment results.

Sentiment Analysis Classification

Figure 6 shows the stages in conducting sentiment analysis. In this stage, the result from the previous stage will be split into two different categories by using the Holdout method, where data will be split into Training data and Testing data. Data split by 80:20 ratio.

The classification methods used in this research are Naive Bayes, Support Vector Machine (SVM), and Random Forest. Naive Bayes method classification by calculating each word frequency and the highest probability of each category in training data to define the attribute that will be used for testing data. Support Machine Vector method classification by mapping the non-linear to convert training data into a higher dimension. Random Forest method classification by creating decision trees as an ensemble, where each decision tree spits out a class prediction, and the most voted class, becomes the model's prediction.

Evaluation and Comparison

The evaluation and comparison method used in this research is the Confusion Matrix method. By using a confusion matrix, each classification method's performance can be evaluated by accuracy in identifying every tuple of different classes. The classes can be represented as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). After the classification result is gained, the next stage is to calculate scores using the Precision, Recall, F1-Score, and accuracy of each classification method.

Visualization

Sentiment analysis comparison result was visualized using matplotlib to plot data into Clustered Bar Chart, Bar Chart, and Word Cloud. Clustered Bar Chart enables the comparison of series in a category, which makes it easier to see the comparison of each classification method's performance. Keyword visualization in each sentiment used Bar Chart and Word Cloud. This visualization enables us to see the keywords that appear most often based on the word frequency.

Results and Discussion

Dataset used in this study consists of 2.391 user reviews on Google Play Store from September 2020 to September 2021. This number is cleaned data that has been processed in several text processing stages. Then, feature extraction is performed on the cleaned data. To continue sentiment analysis classification techniques, data are split into Data Training and Data Testing with an 80:20 ratio, resulting in the distribution of 1912 training data and 479 testing data as shown in Table 3.

Fig. 3: Data collection methods

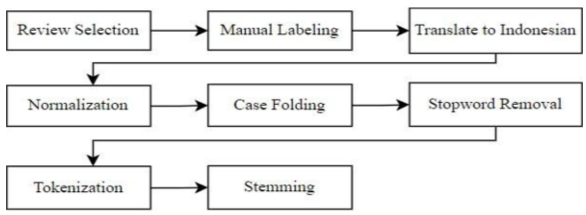

Fig. 4: Text preprocessing methods

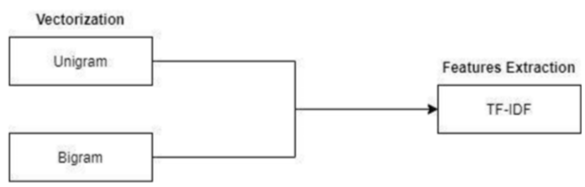

Fig. 5: Feature extraction methods

Table 2: Data attribute and data type

Attribute	Description	Type
Reviewed	Random character to identify the review	String
User name	Reviewer name used in google play store	String
User image	Reviewer profile image	String
Content	User review	String
Score	Rating	Integer
Thumbs up count	Number of reviews likes	Integer
Review created version	App version code	Integer
	When a user posts the review	Datetime
Reply content	App owner or its team response to review	String
Replied at	When app team response	Datetime

To create classification models, each technique performs on the Training dataset, and to have the accuracy score of the model its test on the Testing dataset. Python programming language and Scikit-Learn package library are used in this study to perform all the processes including the classification. The classification model's performance results are shown below:

A. Naive Bayes

Using the Multnomial NB model and function from scikit-learn, data are classified based on keywords. The evaluation result of the Naive Bayes Technique using the confusion matrix produce 431 TP, 31 FP, 13 TN, and 4 FN as shown in Table 4.

This technique shows the average precision value is 92%, an average recall value is 93%, f1-score is 91%, and the accuracy average of the Naive Bayes technique is 93%, where the calculations are shown in Table 5.

B. Support Vector Machine

SVM creates a classification model using Support Vector Classifier (SVC) with linear kernel from python scikit-learn module. After the classifier is created, we perform prediction on Testing data. The evaluation result of the SVM Technique using the confusion matrix produce 428 TP, 23 FP, 21 TN, and 7 FN as shown in Table 6. The SVM model used in this study shows a better performance than the Naïve Bayes model, with an accuracy of 94%. The calculation of the evaluation result can be seen in Table 7.

C. Random Forest

Random forest classification creates decision trees on random data samples. How RF classification works is each tree voted on a class Evaluation result of Random Forest Technique using confusion matrix produce 429 TP, 22 FP, 22 TN, and 6 FN as shown in Table 8. We use 20 trees to create the forest, and it shows 94% accuracy in this model. The calculation of the evaluation result can be seen in Table 9.

D. Sentiment Analysis Technique Comparison

Based on the performance score evaluation of each classification model for each aspect which is precision, recall, f1-score, and accuracy, it shows that Random Forest has the highest performance score, then followed by the SVM model, and the Naive Bayes model come in last.

To find the best classification model, we can assume the accuracy percentage of each model.

Table 10 shows the percentage detail of the performance score of each classification model. From the table, the highest accuracy result has by the Random Forest model. To have a better visualization of the performance comparison of each model, we present the comparison with clustered bar chart using matplotlib. In Fig. 7 it can be seen that Random Forest Model's bar is higher than the others.

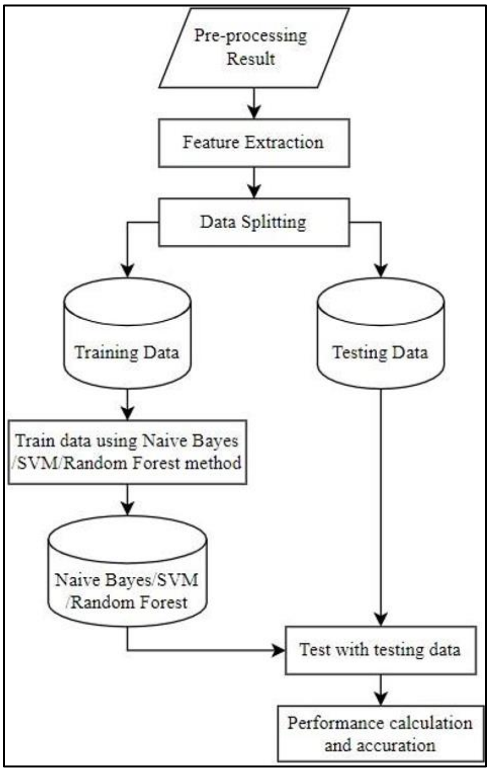

Fig. 6: Sentiment analysis classification methods

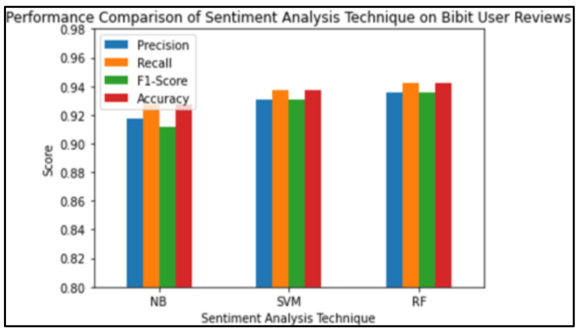

Fig. 7: Performance comparison in a clustered bar chart

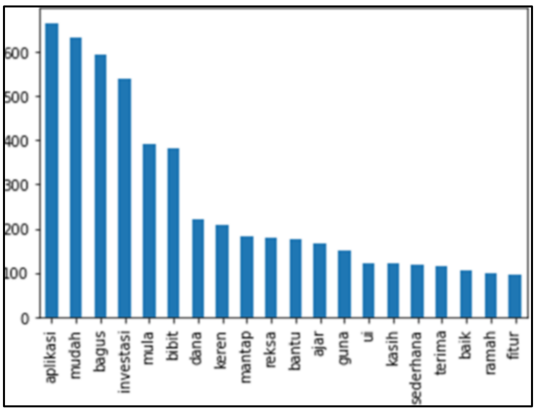

Fig. 8: Positive sentiment reviews in the bar chart

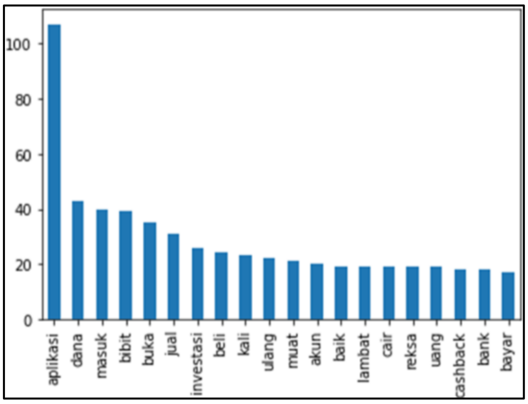

Fig. 9: Negative sentiment reviews in a bar chart

Fig. 10: Positive sentiment review in the word cloud

Fig. 11: Negative sentiment review in the word cloud

Table 3: Numbers of data split for each label

Label	Data numbers
Train	1912
Test	479

Table 4: Confusion Matrix of Naïve Bayes Model

	Confusion matrix	
	Predicted_positive	Predicted_negative
Positive	431	4
Negative	31	13

E. Sentiment Analysis Result

The topics most discussed by users in reviews that have positive or negative sentiments related to the experience of using the application can be identified by visualizing the application user review dataset that has been cleaned so that it only contains words that are meaningful to analyze and visualize. Every word is visualized into a bar chart using matplotlib which displays the 20 words with the most frequently referred to by users.

In Fig. 8 it can be seen that the 20 words most discussed by users who left positive sentiment reviews were 'aplikasi', 'mudah', 'bagus', 'investasi', 'mula', 'bibit', 'dana', 'keren', 'mantap', 'reksa', 'bantu', 'ajar', 'guna', 'ui', 'kasih', 'sederhana', 'terima', 'baik', 'ramah', and 'fitur'.

In Fig. 9 it can be seen that the 20 words most discussed by users who left negative sentiment reviews were 'aplikasi', 'dana', 'masuk', 'bibit', 'buka', 'jual', 'investasi', 'beli', 'kali', 'ulang', 'muat', 'akun', 'baik', 'lambat', 'cair', 'reksa', 'uang', 'cashback', 'bank', and 'bayar'.

The words of positive or negative sentiment reviews from mutual fund app users can also be visualized in the form of a word cloud, where words that have more frequency will be displayed in a larger size than other words. The results of word cloud visualization with positive sentiment can be seen in Fig. 10 and negative sentiment can be seen in Fig. 11.

Conclusion and Future Work

In this study, the primary goal is to compare the accuracy of the results of sentiment analysis using the Naive Bayes, Support Vector Machine (SVM), and Random Forest methods on user reviews of the mutual fund application. The user reviews have been collected through a web scraping method on the google play store, then cleaned through several processes of data pre-processing. Feature extraction was performed using TF-IDF along with vectorization using a combination of n-grams, namely unigram and bigram. The model performance was measured using a confusion matrix. At the training: Testing data ratio of 80:20, the accuracy of the Naive Bayes, SVM, and Random Forest methods is 92.7, 93.7, and 94.2%, respectively.

The second goal is to identify user reviews of the mutual fund app to gain insight into the topics covered by each sentiment using visualizations. In the positive sentiment of users, the majority discusses the application which is easy and good, especially for novice investors. In negative sentiment, the majority discussed the slow sales process to disbursement of funds and long loading times when opening the application.

For future work, this study has suggested some scenarios: (1) use Indonesian corpus in the sentence normalization process, (2) use a balanced dataset or balance the data set using Over sampling or Under sampling technique, and (3) use other NLP libraries such as TextBlob, Spacy, and CoreNLP from Stanford as a comparison, (4) use other training and testing data sharing methods such as the K-Fold method as a comparison, (5) use other evaluation methods such as AUC-ROC Curve as a comparison.

Table 1: Classification report of naïve bayes model

	Classification Report			
	Precision-Recall	F1-Score	Support	
0	0.76	0.30	0.43	44
1	0.93	0.99	0.96	435
Accuracy			0.93	479
Macro avg	0.85	0.64	0.69	479
Weighted avg	0.92	0.93	0.91	479

Table 2: Confusion matrix of SVM model

	Confusion Matrix	
	Predicted_Positive	Predicted_Negative
Positive	428	7
Negative	23	21

Table 3: Classification report of SVM model

	Classification Report			
	Precision-Recall	F1-Score	Support	
0	0.75	0.48	0.58	44
1	0.95	0.98	0.97	435
Accuracy			0.94	479
Macro avg	0.85	0.73	0.77	479
Weighted avg	0.93	0.94	0.93	479

Table 4: Confusion matrix of random forest model

	Confusion Matrix	
	Predicted_Positive	Predicted_Negative
Positive	429	6
Negative	22	22

Table 5: Classification report of random forest model

	Classification Report			
	Precision-Recall	F1-Score	Support	
0	0.79	0.50	0.61	44
1	0.95	0.99	0.97	435
accuracy			0.94	479
macro avg	0.87	0.74	0.79	479
weighted avg	0.94	0.94	0.94	479

Table 6: Comparison of Classification Model's Result

Model	Precision	Recall	F-1 Score	Accuracy
Naive bayes	0.917	0.927	0.912	0.927
Support vector machine	0.931	0.937	0.931	0.937
Random forest	0.936	0.942	0.936	0.942

Acknowledgment

This study is supported by Research and Technology Transfer Office, Bina Nusantara University.

Author's Contribution

Evaristus Didik Madyatmadja: Lead research project, coordinate developer, doing experiment, be an instructor, data analysis and writing the manuscript.

Shinta: Advise research project, design the research methodology, data analysis, writing manuscript and proof reading.

Devi Susanti: Advise research project, design the experiment, data analysis and writing manuscript.

Florencia Anggreani: Advise data analysis and writing manuscript.

David Jumpa Malem Sembiring: Advise research project, design the application, data analysis, writing manuscript, and proof reading.

Ethics

The authors confirm that this manuscript has not been published elsewhere and that no ethical issues are involved.

References

Aqlan, A. A. Q., Manjula, B., & Lakshman Naik, R. (2019). A study of sentiment analysis: Concepts, techniques, and challenges. In *Proceedings of International Conference on Computational Intelligence and Data Engineering* (pp. 147-162). Springer, Singapore. http://doi.org/10.1007/978-981-13-6459-4_16

Bhavsar, P., Safro, I., Bouaynaya, N., Polikar, R., & Dera, D. (2017). Machine learning in transportation data analytics. In *Data Analytics for Intelligent Transportation Systems* (pp. 283-307). Elsevier. http://doi.org/10.1016/B978-0-12-809715-1.00012-2

de Godoi Brandão, J., & Calixto, W. P. (2019, September). N-Gram and TF-IDF for Feature Extraction on Opinion Mining of Tweets with SVM Classifier. In *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-5). IEEE. http://doi.org/10.1109/idap.2019.8875900

Fachruddin, M. I. (2015). *Perbandingan metode random forest classification dan support vector machine untuk deteksi epilepsi menggunakan data rekaman electroen cephalograph (EGG)* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).

Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, *18*(1), 71-80. http://doi.org/10.26623/transformatika.v18i1.2317

Gurusamy, V., & Kannan, S. (2014). Preprocessing Techniques for Text Mining.

Han, J., Kamber, M., & Pei, J. (2012). Outlier detection. *Data Mining: Concepts and Techniques*, 543-584.

Himawan, R. D., & Eliyani, E. (2021). Perbandingan akurasi analisis sentimen tweet terhadap pemerintah provinsi dki jakarta di masa pandemi. *jepin (Jurnal Edukasi dan Penelitian Informatika)*, *7*(1), 58-63. http://doi.org/10.26418/jp.v7i1.41728

Jumeilah, F. S. (2017). Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *1*(1), 19-25: http://doi.org/10.29207/resti.v1i1.11

Khalid, M., Asif, M., & Shehzaib, U. (2015). Towards improving the quality of mobile app reviews. *International Journal of Information Technology and Computer Science (IJITCS)*, *7*(10), 35. http://doi.org/10.5815/ijitcs.2015.10.05

Khan, M., & Malik, K. (2018, April). Sentiment classification of customer's reviews about automobiles in roman urdu. In *Future of Information and Communication Conference* (pp. 630-640). Springer, Cham. http://doi.org/10.1007/978-3-030-03405-4_44

KSEI. (2021). Statistik_Publik_Agustus_2021.pdf. KSEI: https://www.ksei.co.id/files/Statistik_Publik_Agustus_2021.pdf

Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: A review. *EURASIP Journal on Wireless Communications and Networking*, *2017*(1), 1-12. http://doi.org/10.1186/s13638-017-0993-1

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1-167. http://doi.org/10.2200/s00416ed1v01y201204hlt016

Madyatmadja, E. D., Olivia, J., & Sunaryo, R. F. (2019). Priority analysis of community complaints through e-government based on social media. *Int. J. Recent Technol. Eng.*, *8*(3), 1-5. http://doi.org/10.35940/ijrte.c5011.098319

Nanli, Z., Ping, Z., Weiguo, L. I., & Meng, C. (2012, November). Sentiment analysis: A literature review. In *2012 International Symposium on Management of Technology (ISMOT)* (pp. 572-576). IEEE. http://doi.org/10.1109/ISMOT.2012.6679538

Nguyen, H., Veluchamy, A., Diop, M., & Iqbal, R. (2018). Comparative study of sentiment analysis with product reviews using machine learning and lexicon-based approaches. *SMU Data Science Review*, *1*(4), 7.

Sadia, A., Khan, F., & Bashir, F. (2018, February). An overview of the lexicon-based approach for sentiment analysis. In *2018 3rd International Electrical Engineering Conference (IEEC 2018)* (pp. 1-6).

Saepulrohman, A., Saepudin, S., & Gustian, D. (2021). Analisis Sentimen Kepuasan Pengguna Aplikasi Whatsapp Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine. *is Best Account. Inf. Syst. Inf. Technol. Bus. Enterp*, *6*(2), 91-105. http://doi.org/10.52005/rekayasa.v8i1.107

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, *34*(1), 1-47. DOI.org/10.1145/505282.505283

Thakkar, H., & Patel, D. (2015). Approaches for sentiment analysis on Twitter: A state-of-art study. *arXiv preprint arXiv:1512.01043*. https://www.researchgate.net/publication/269762971_Sentiment_Analysis_and_Opinion_Mining

Kulkarni, V. Y. & Sinha. P. K. (2014). Effective Learning and Classification Using Random Forest Algorithm. *International Journal of Engineering and Innovative Technology (IJEIT).* 3(11): 267-273