Analysis of User Reviews for the PeduliLindungi Application on Google Play Using the Support Vector Machine and Naive Bayes Algorithm Based on Particle Swarm Optimization

1st Ali Mustopa
Computer Science
STMIK Nusa Mandiri
Jakarta, Indonesia
14002370@nusamandiri.ac.id

2nd Hermanto
Computer Technology
Bina Sarana Informatika University
Jakarta, Indonesia
hermanto.hmt@bsi.ac.id

3rd Anna
Accounting Information System Of
Campus Pontianak City
Bina Sarana Informatika University
Pontianak, Indonesia
anna.nnz@bsi.ac.id

4th Eri Bayu Pratama
Information System Of Campus
Pontianak City
Bina Sarana Informatika University
Pontianak, Indonesia
eri.ebp@bsi.ac.id

5th Ade Hendini
Information System Of Campus
Pontianak City
Bina Sarana Informatika University
Pontianak, Indonesia
ade.aee@bsi.ac.id

6th Deni Risdiansyah
Information System Of Campus
Pontianak City
Bina Sarana Informatika University
Pontianak, Indonesia
deni.drx@bsi.ac.id

Abstract— Corona Virus 19 (COVID-19) is a contagious viral infection that has now spread to various countries, one of which is Indonesia. Monitoring of the spread of COVID-19 in Indonesia is handled directly by the Government of Indonesia, especially by the Ministry of Communication and Information (KOMINFO) with the creation of the Protected application found on Google Play. Users provide reviews or comments about the application, of course, users will choose applications that have good reviews. However, monitoring reviews from the general public is not easy, because there are so many of them to process. So that the researcher wants to know the extent of the analysis of user reviews of the PeduliLindungi application based on reviews of user comments by using classification techniques, namely the Support Vector Machine (SVM) Algorithm and Naive Bayes Based on Particle Swarm Optimization (PSO). The test results with the accuracy value and AUC value of each, namely for the PSO-based Naive Bayes algorithm the accuracy value = 69.00%, and AUC value = 0.659, while for the PSO-based SVM algorithm the accuracy value = 93.0% and the AUC value = 0.977. For this reason, the application of Particle Swarm Optimization (PSO) -based Support Vector Machine in this study has higher accuracy so that it can be used to provide solutions to sentiment analysis problems in review comments of users of the PeduliLindungi application.

Keywords— Covid-19, Google Play, PeduliLindungi, User reviews, Support Vector Machine (SVM), Naive Bayes, Particle Swarm Optimization (PSO)

I. INTRODUCTION

Coronavirus 19 (COVID-19) is a highly contagious viral infection caused by acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) [1]. It started to become widespread in Wuhan, Hubei Province, China. Most of the patients are epidemiologically linked to the Huanan Seafood Wholesale Market [2]. The spread of COVID-19 has become a major global public health event, linking people's physical and mental health and even mental safety [3]. Indonesia is the fourth most populous country in the world and is expected to be significantly affected by COVID-19 over a longer period [4].

Monitoring the spread of COVID-19 in Indonesia is handled by the Government of Indonesia in various ways, one of which is through the Android application installed by Google Play. This application made by the government is called Pedulilindung, which is a forum for handling and community participation to share location data while traveling so that contact history tracing with COVID-19 sufferers in Indonesia can be carried out immediately. This is certainly a discussion that is often discussed by the public with their respective opinions and opinions through the PeduliLindungi application on Google Play.

Google Play is a digital content service owned by Google which consists of online product stores such as music, songs, books, applications, games, or media players. Google Play can be accessed via the web, android application (Play Store), and Google TV. In Google Play, in addition to online product stores, there is also an assessment feature for customers to provide reviews about the advantages and disadvantages of using the PeduliLindungi application. Various kinds of reviews have been given by users, ranging from smooth and harsh sentences. Google Play users can also provide an assessment in the form of a star score (between 1-5) depending on each user rating. And the last thing, users can also give a score "like" to someone's comment that represents someone's interest or dislike of the application used.

For this research, the main focus is to review comments from users of the PeduliLindungi application on Google Play with data retrieval for three months, from April 2020 to June

2020. User reviews of the PeduliLindungi application can be influenced by several things that have not yet come to the attention of the Ministry of Communication and Information (KOMINFO) as the application maker. Starting from the many negative comments that come from users of the PeduliLindungi application in the form of complaints or problems they face when using the application. Some of them sometimes the application provides incorrect information about the security of the user's location from people who have Covid. According to many users, the PeduliLindungi application has not been accurate in providing information on whereabouts whether it is in the red, yellow, or green zone. So that users feel the application only consumes their cellphone quota and battery because they have to keep Bluetooth activated and notifications keep ringing. Another obstacle is that user sessions often expire so they log out automatically and have to input OTP repeatedly to log into the application again. And some also admit that when installing the PeduliLindungi application, their device's internet access speed decreases. These things happen because several features must be improved for the development of the PeduliLindungi application to be more optimal. Many negative comments have been given by the community regarding the PeduliLindungi application services so that users give low ratings & make them do not want to use this application anymore. Based on these constraints, the government especially KOMINFO knows public satisfaction with the PeduliLindungi application so that it can improve and improve service quality. Seeing this, it is necessary to analyze the sentiment comments of the PeduliLindungi application users. So it can be an evaluation for application developers to improve the performance of the application.

Sentiment analysis or opinion mining is a computational study of a person's opinion, behavior, and emotions towards an entity. These entities can describe individuals, events, or topics. Sentiment analysis is a technique in extracting information in the form of a person's view of an issue. It is also said to be a process to understand and process textual data automatically in order to generate sentiment information contained in a sentence. Sentiment analysis brings great influence and benefits, causing research and applications based on sentiment analysis to grow rapidly [5]. Measuring sentiment on open source information is currently an active area of research [6].

Several studies have been conducted in classifying sentiments on reviews available online. Among them, Comparison of Naïve Bayes Algorithm, C4.5, and Random Forest for Service Classification Ojek Online [7]. Sentiment Analysis of Facebook Social Media Comment Data with K-Nearest Neighbor (Case Study on the J&T Ekspress Indonesia Freight Forwarding Service Account) [8]. Sentiment Analysis of User Satisfaction Levels of Indonesian Cellular Telecommunication Service Providers on Twitter with the Support Vector Machine and Lexicon Based Features Method [20]. Analysis of Support Vector Machine Text Mining Model on Positive and Negative Comments for a Comparison Review of Whatsapp vs Bbm [21]. Sentiment Analysis About Gojek and Grab User Reviews Using Svm Algorithm Based on Particle Swarm Optimization [22]. Of

these techniques, the most frequently used for data classification is the Support Vector Machine (SVM) algorithm so that this study uses the SVM algorithm for user reviews of the PeduliLindungi application with user review data sources on Google Play.

Based on previous studies, the Naive Bayes (NB) algorithm and the Support Vector Machine (SVM) are classification algorithms that are widely used by researchers in the field of text mining. These two algorithms are also used in this study in the classification of user reviews of the PeduliLindungi application. With the aim that the selected algorithm is the most accurate so that it can analyze the user reviews of the PeduliLindungi application.

In this research, the data testing process also used RapidMiner, starting with the formation of a model with the data in the first part, namely the sharing of training and testing data. Using 10 fold cross-validation as standard validation, this process splits data randomly into 10 parts. The model that is formed will be tested on the remaining 9 parts of the data. Meanwhile, accuracy can be measured using the confusion matrix, and performance is measured using accuracy and AUC. And will be presented in the form of a ROC curve to classify text in user reviews of the PeduliLindungi application. Of course, it can be seen that the highest level of accuracy in the PeduliLindungi application is in accordance with the results of research using the SVM and NB algorithms based on the Particle Swarm Optimization (PSO) algorithm model.

This research aims to classify user comment review sentiments on the PeduliLindungi application with better results for future application development. And can generate the level of accuracy and Area Under Curve to find a better solution. The research results are expected to find ideal criteria to determine the use of appropriate algorithms for classification algorithms that can be used for classifying reviews of user comments for the PeduliLindungi application. In addition, this study limits the data to be classified as only reviews of users who care about being protected on Google Play. The data was taken in only three months, from April 2020 to June 2020. And the review data on the application to be classified is Indonesian text which includes both formal and non-formal languages.

II. LITERATURE REVIEW

A. Data Mining

Data mining is a process that uses artificial intelligence, mathematics, statistics, and machine learning techniques to identify and extract useful information from various large-scale databases [9]. Data mining is also defined as an automatic process of very large data and aims to obtain beneficial relationships or patterns. Data mining can be said to be a decision-making support process that looks for patterns of information in data. This search can be performed by the user [10].

B. Sentiment Analysis

Sentiment analysis is a process to determine the sentiment or opinion of a person which is manifested in text

and can be categorized as positive or negative sentiment [11]. Sentiment analysis refers to the broad fields of natural language processing, computational linguistics, and text mining which aim to analyze the sentiments, opinions, attitudes, evaluations, judgments, and emotions of a person whether the speaker or writer is related to a product, topic, service, organization, individual, or the other activities [12]. Sentiment analysis can also be interpreted as textual information in general divided into factual information (objective) and opinions (subjective) [13]. Sentiment analysis is also computational research of opinions, sentiments, and emotions that are expressed textually [14].

C. Text Mining

The important purpose of text mining is to get high-quality information from text. This is usually done by extracting knowledge and information from patterns in text documents [15]. It is also done through finding patterns and trends in ways such as statistical pattern learning, topic, and statistical language modeling. Text mining generally requires structuring the input text (eg parsing, along with adding some other derived linguistic features and deletions, and subsequent insertion into the database). This is followed by deriving patterns in the structured data, and evaluation and interpretation of the outputs. The "high quality" in text mining usually refers to a combination of relevance, interest, and novelty [16].

D. Support Vector Machine (SVM)

Support Vector Machine (SVM) was first introduced by Vapnik in 1992 as a harmonious series of leading concepts in the field of pattern recognition. As a pattern recognition method, the SVM algorithm is relatively young. However, the evaluation of capabilities in various applications places it as a state of the art in pattern recognition. Currently, SVM is a topic that is growing rapidly. SVM is included in the machine learning method that works on the principle of Structural Risk Minimization (SRM). The goal is to find the best hyperplane that separates two classes in the input space [17].

E. Naive Bayes

This method uses the Bayes theorem, which was discovered by Thomas Bayes 18th century [18]. The NB Classification (NB) is a statistical classification that can be used to predict the probability of membership in a class. According to Wu and Kumar, the NB algorithm is a popular classification method and is included in the ten best algorithms in data mining. NB uses a branch of mathematics known as probability theory to find the greatest chance of possible classifications by looking at the frequency of each classification in the training data [19].

F. Particle Swarm Optimization (PSO) Technique

Particle Swarm Optimization (PSO) is a very simple optimization technique for applying and modifying several parameters. In Particle Swarm Optimization (PSO), there are several techniques for optimization including increasing the attribute weight of all attributes or variables used, selecting attributes (attribute selection), and feature selection [20].

III. RESEARCH METHODS

A. Bussines Understanding

At this stage, it is carried out with an understanding of the object research. The data that will be used in this research is a review of the comments of users of the PeduliLindungi application on Google Play who provide comments, to determine the status of the reviews of users' comments on Google Play. At this stage, data for reviewing comments from PeduliLindungi users are presented in the text on Google Play which will be grouped based on the content of the discussion of each comment review category.

At this stage, an understanding will also be made to find the best algorithm method so that it can help the data processing that will be carried out by comparing the results of the NB and SVM algorithms in the review comments from the PeduliLindungi application users.

B. Data understanding

This stage is the process of understanding the data that will be used as the material under study to be carried out to the next stage, namely processing. Protected user review data retrieve 1,364 data from Google Play. The user reviews are then grouped based on the category of comments provided by PeduliLindungi users.

C. Data Preparation

The next stage is to prepare data before the modeling stage or what is called Data Preparation. This second stage is text preprocessing, where the steps will be carried out using two preprocessing applications. First, use the Gata Framework which is accessed via the link http://gataframework.com/textmining. This application can be used for free and easy because you do not have to create an account to use the service and continue preprocessing from RapidMiner. Here are the stages:
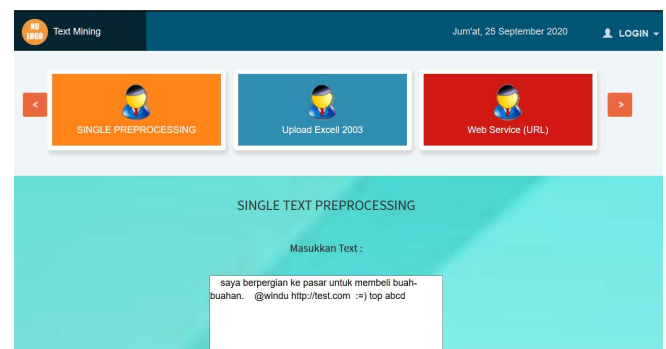

Fig. 1. The View of Gata Framework Text Mining Tools

The Gata Framework is an alternative in Indonesian text pre-processing combined with the RapidMiner application to process Indonesian words. This is because the Rapid Miner application already has a dictionary facility to change acronyms and stopwords, although it is still limited to English, Chinese, and Arabic. Meanwhile, Indonesian is still not available. From the results of pre-processing using the Gata Framework, the dataset will be pre-processed again using the Rapid Miner tools to clean the data so that the results are even better.

D. Remove Duplicates

Remove Duplicates is the next data preparation stage used in the RapidMiner software. Remove duplicates that are used to remove the same or duplicates text. This is done so that the data is not filled with the same text so that it slows down the process of running software to analyze the model.

E. Nominal to Text

Nominal to text is an operator in rapidminer which functions to convert all numbers in text into text. So that the existing numbers will be considered the type of text data, not numeric or nominal. Figure 2 shows how the use of this operator is used in the existing process in RapidMiner.
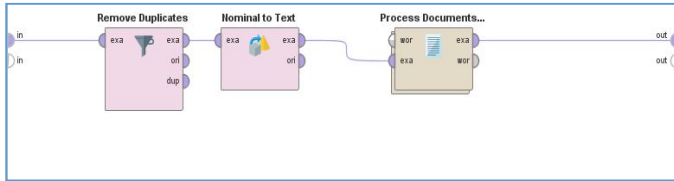

Fig. 2.   Nominal to Text Process View

F. Transform Case

The operators used at this stage are to change the capital letters that are still in the text to be changed to all lowercase letters. This is done so that the typing process that is carried out into the classification model has uniformity of letters and there are no errors in the tokenizing process.

G. Filter Token (By Length)

The token filter is a process that is in data preparation to remove several words (after tokenizing) with a certain character length. In this study, the minimum length of characters used is 4 characters and the maximum length is 25 characters. This means that words that are less than 4 characters long and more than 25 characters long will be removed. To get this result, the operator's parameters are set.
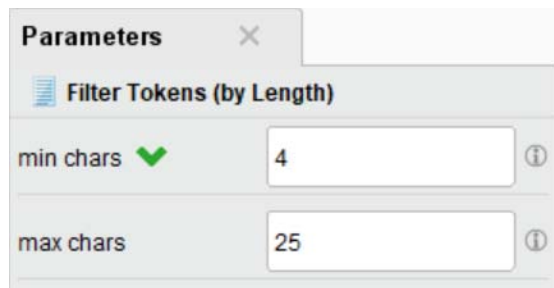

Fig. 3.   The View of Token Filter Process (by Length)

H. Filter Stopword (Dictionary)

Next is the use of the Stopword Removal (by Directory) operator which functions to remove words that are not related to the content of the text. In the previous stage using the Gata framework text mining service, there were some words that the previous service had not been able to remove because they had not been entered as words that had to be deleted. So with the Stopword Removal (by Directory) operator, researchers can register words that should be deleted from the text.

IV. RESULTS AND DISCUSSION

A. Modeling Stages

It is the stage of selecting a mining technique by determining the algorithm to be used. The tool used is the Rapid Miner 9.1 version. The results of the model testing carried out were analyzing reviews of PeduliLindungi user comments on Google Playstore with positive and negative comment data labels using the NB and SVM algorithms based on PSO to get the best accuracy value. The following is the RapidMiner model design used, namely:

B. Testing Stages

The settings and use of operators and parameters in the Rapid Miner framework greatly affect the accuracy and the model formed, more clearly the testing of the NB algorithm and the SVM Based on PSO is as follows:

The process of this model is the NB algorithm and SVM Based on PSO, a process of 10 fold validation is carried out, can be seen in Figure 4:
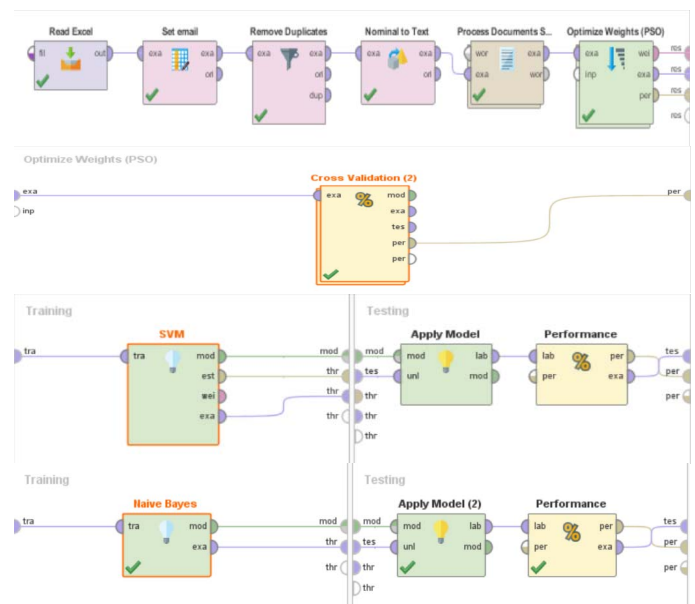

Fig. 4. SVM Algorithm Testing Model Based on PSO Using RapidMiner

Figure 4 above is a PSO-based SVM algorithm testing model using rapidminer, starting from entering data then determining the role that will determine the label there and nominal text then processing the document. In this tester, the data used is clean data that has gone through preprocessing. The data is retrieved using the Read Excel operator which is saved in Excel format (.xlsx). Then enter into the algorithm design model, enter the SVM operator cross-validation, in which there is an algorithm calculation, then the model is applied, after that, it enters the performance results, then the results of the accuracy of the value appear.

C. Evaluation Model

The evaluation stage aims to determine the use-value of the model that was successfully created in the previous step. For evaluation, 10-fold cross-validation is used. From the results of model testing, the algorithm used is to produce an Accuracy (Confusion Matrix) and AUC (Area Under Curve) value. Then get the results of the ROC graph with the AUC value (Area Under Curve).

D. The Accuracy Value of the NB Algorithm and SVM Based on PSO

Based on the results of experiments conducted using user comment review data of the PeduliLindungi application on Google Play. In this experiment, using the NB algorithm and SVM based on PSO. Using 1364 user review data that produces the accuracy value (Confusion Matrix) as follows:

accuracy: 69.00% +/- 7.09% (micro average: 69.00%)			
	true Positif	true Negatif	class precision
pred. Positif	125	49	71.84%
pred. Negatif	75	151	66.81%
class recall	62.50%	75.50%	

Fig. 5. The Accuracy Value of PeduliLindungi with the NB + PSO Algorithm

The number of True Positives (TP) is 125 records classified as positive and False Positives (FP) are 75 records classified as Negative. Next, 151 True Negatives (TN) records were classified as Negative and 49 False Negative (FN) records were classified as Negative. Based on table 1 above shows that the level of accuracy of user reviews of the PeduliLindungi application using the NB algorithm based on PSO is 69,00%.

accuracy: 93.00% +/- 3.50% (micro average: 93.00%)			
	true Positif	true Negatif	class precision
pred. Positif	181	9	95.26%
pred. Negatif	19	191	90.95%
class recall	90.50%	95.50%	

Fig. 6. The Accuracy Value of PeduliLindungi with the SVM+ PSO Algorithm

The number of True Positives (TP) is 181 records classified as positive and False Positives (FP) are 19 records classified as Negative. Furthermore, 191 records of True Negatives (TN) were classified as Negative and 9 records of False Negatives (FN) were classified as Negative. Based on table 2 above shows that the level of accuracy of user reviews of the PeduliLindungi application using the SVM algorithm based on PSO is 93,00%.

E. The AUC value of PeduliLindungi Application with NB and SVM Algorithm Based on PSO.

1. The following are the results of testing data on user reviews of PeduliLindungi with 1364 data using the NB Algorithm Based on PSO resulting in AUC = 0.659:
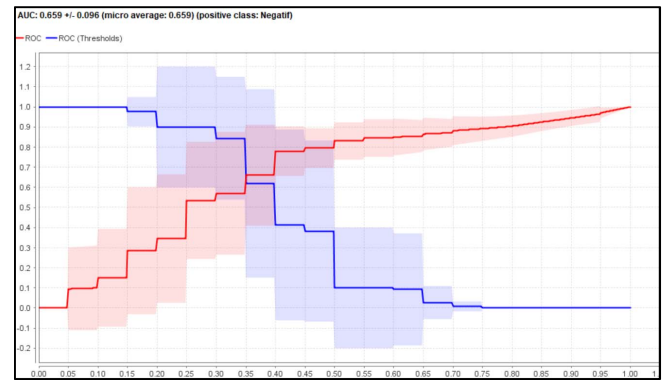

Fig. 7. RCO Naive Bayes Curve + PSO

2. The following are the results of testing data for Protected users with 1364 data using the SVM algorithm based on PSO resulting in AUC = 0.977:
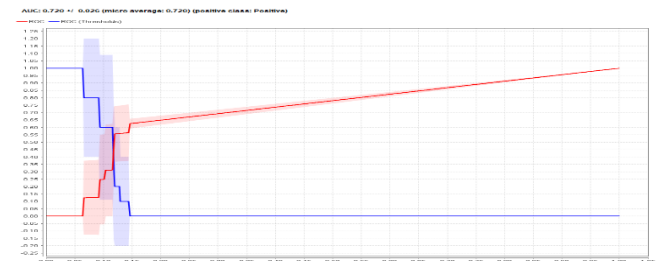

Fig. 8. ROC SVM Curve + PSO

F. Recall and Precision Values of PeduliLindungi with NB and SVM Algorithms Based on PSO.

1. The recall value with the results of testing data from users of PeduliLindungi using SVM + PSO amounted to 95,50%. Meanwhile, for NB + PSO resulted in a recall of 75,50%. The following are the results of the data processing.

recall: 95.50% +/- 3.50% (micro average: 95.50%) (positive class: Negatif)			
	true Positif	true Negatif	class precision
pred. Positif	181	9	95.26%
pred. Negatif	19	191	90.95%
class recall	90.50%	95.50%	

Fig. 9. Recall SVM + PSO

recall: 75.50% +/- 9.34% (micro average: 75.50%) (positive class: Negatif)			
	true Positif	true Negatif	class precision
pred. Positif	125	49	71.84%
pred. Negatif	75	151	66.81%
class recall	62.50%	75.50%	

Fig. 10. Recall NB + PSO

2. The following is the value of precision with the results of testing data reviews of users who care to be protected using SVM + PSO is 91,13%. Meanwhile, NB + PSO produces a precision of 67,32%.

precision: 91.13% +/- 4.60% (micro average: 90.95%) (positive class: Negatif)			
	true Positif	true Negatif	class precision
pred. Positif	181	9	95.26%
pred. Negatif	19	191	90.95%
class recall	90.50%	95.50%	

Fig. 11. Precision SVM + PSO

precision: 67.32% +/- 7.68% (micro average: 66.81%) (positive class: Negatif)			
	true Positif	true Negatif	class precision
pred. Positif	125	49	71.84%
pred. Negatif	75	151	66.81%
class recall	62.50%	75.50%	

Fig. 12. Precision NB + PSO

G. Comparison of Accuracy

Based on the results of experiments conducted using user review data for the PeduliLindungi application on Google Play, using the NB and SVM based on PSO using 1364 filtered comment review data, the resulting accuracy, and AUC values are as follows:

TABLE I.   COMPARISON OF PERFORMANCE ALGORITHMS

	Accuracy	AUC	Recall	Precision
NB+PSO	69.00%	0.659	75,50%.	67,32%
SVM+PSO	**93.00%**	**0.977**	**95,50%**	**91,13%**

From the results of the comparison of the performance of the two algorithms above, the SVM+PSO-based test results produce a higher accuracy value than the Naive Bayes algorithm PSO based. So that the application of the PSO-based SVM algorithm in this research is said to be better and can be used to provide solutions to the problems of sentiment analysis in reviews of users of the PeduliLindungi application on Google Play.

V. CONCLUSION

In this study, after preprocessing and model testing using data mining methods, namely Naive Bayes and Support Vector Machine based on Particle Swarm Optimization, it can be seen that the accuracy value to determine that the analysis of review comments from users of the PeduliLindungi. It can be proven by looking at the accuracy value and the AUC value from each other. For the PSO-based NB algorithm, with an accuracy value = 69.00% and AUC value = 0.659, while for the PSO-based SVM algorithm, with an accuracy value = 93.0% and AUC value = 0.977. And from this research, it can be seen that the level of accuracy produced by the NB algorithm and PSO-based Support Vector Machine, the superior algorithm that has the highest accuracy value is the PSO-based SVM algorithm compared to the PSO-based NB algorithm. In other related research [22] which uses the SVM algorithm using user review data from Gojek and Grab on Google Play as much as 1,360 data. The results of these studies indicate a high accuracy value using a Support Vector Machine based on Particle Swarm Optimization up to 73.09% for user reviews of the grab application. For this reason, in this study, the application of Support Vector Machine based on Particle Swarm Optimization has also resulted in higher accuracy so that it can be used to provide solutions to sentiment analysis problems in review comments from users of the PeduliLindungi application on Google Play. For future research, it can take a larger amount of data in more than three months on Google Play, so it will get even better accuracy values, and also necessary to analyze application user reviews based on 1-5 star rates and the number of 'likes' obtained. In addition, it is hoped that further research can add or replace other classification methods to compare with other Supervised Learning methods from the proposed model, such as the use of the C-45 Algorithm, Random Forest, or other algorithms.

REFERENCES

[1] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses," *J. Adv. Res.*, vol. 24, pp. 91–98, 2020.

[2] L. Yang *et al.*, "Epidemiological and clinical features of 200 hospitalized patients with coronavirus disease 2019 outside Wuhan, China: A descriptive study.," *J. Clin. Virol.*, vol. 129, no. March, p. 104475, 2020.

[3] W. Cai, B. Lian, X. Song, T. Hou, G. Deng, and H. Li, "A cross-sectional study on mental health among health care workers during the outbreak of Corona Virus Disease 2019," *Asian J. Psychiatr.*, vol. 51, no. March, p. 102111, 2020.

[4] R. Djalante *et al.*, "Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020," *Prog. Disaster Sci.*, vol. 6, p. 100091, 2020.

[5] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," *Integer J. Maret*, 2017.

[6] L. L. Dhande and G. Patnaik, "Review of Sentiment Analysis using Naive Bayes and Neural Network Classifier," *Int. J. Sci. Eng. Technol. Res.*, vol. 03, no. 07 Mei 2014, pp. 1110–1113, 2014.

[7] Hermanto, S. J. Kuryanti, and S. N. Khasanah, "Comparison of Naïve Bayes Algorithm, C4 . 5 and Random Forest for Service Classification Ojek Online," *J. Publ. Informatics Eng. Res.*, vol. 3, no. April 2019, pp. 266–274, 2019.

[8] A. Salam, J. Zeniarja, and R. S. U. Khasanah, "Analisis Sentimen Data Komentar Sosial Media Facebook Dengan k-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspress Indonesia)," *Pros. SINTAK 2018*, no. ISBN: 978-602-8557-20-7, 2018.

[9] J. Ipmawati, Kusrini, and E. Taufiq Luthfi, "Komparasi Teknik Klasifikasi Teks Mining Pada Analisis Sentimen," *Indones. J. Netw. Secur.*, 2017.

[10] E. N. Sari, "Analisa Algoritma Apriori untuk Menentukan Merek Pakaian yang Paling Diminati pada Mode Fashion Group Medan," *Pelita Inform. Budi Darma*, 2013.

[11] N. Muchammad Shiddieqy Hadna, P. Insap Santosa, and W. Wahyu Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, 2016.

[12] U. M. Boy, "Analisis Sentimen Pada Twitter Menggunakan Text Mining Skripsi," 2014.

[13] C. B. Safari, *Skripsi Analisis Sentimen Menggunakan Metode Klasifikasi SVM*. Jakarta: Universitas Pendidikan Indonesia, 2012.

[14] J. Ipmawati, Kusrin, and E. T. Luthfi, "Komparasi Teknik Klasifikasi Teks Mining Pada Analisis Sentimen," *IJSE –Indonesian J. Softw. Eng.*, vol. Vol. 2 No, 2016.

[15] H. Hermanto, A. Mustopa, and A. Y. Kuntoro, "Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, 2020.

[16] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. 2016.

[17] A. S. Nugroho, A. B. Witarto, and D. Handoko, "Support Vector Machine – Teori dan Aplikasinya dalam Bioinformatika," *Kuliah Umum IlmuKomputer.Com*, 2003.

[18] Suyanto, *Data Mining untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung, 2017.

[19] A. Mukminin and D. Riana, "Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah," *J. Inform.*, 2017.

[20] I. Rahmayuni, "Perbandingan performansi algoritma c4.5 dan cart dalam klasifiksi data nilai mahasiswa prodi teknik komputer politeknik negeri padang," *Teknoif*, 2014.

[21] Agus darmawan, Syamsiah (2016). Analisis Model Support Vector Machine Textmining Pada Komentar Positif Dan Negatif Untuk Review Perbandingan Whatsapp Vs BBM, Jurnal String Vol. 1 No. 1 Tahun 2016, ISSN: 2527 – 9661.

Hermanto, H., Kuntoro, A., Asra, T., Nurajijah, N., Effendi, L., & Ocanitra, R. (2020). Sentiment Analysis On Gojek And Grab User Reviews Using Svm Algorithm Based On Particle Swarm Optimization. Pilar Nusa Mandiri: Journal Of Computing And Information System, 16(1), 117-122.