Comparison Of Support Vector Machine Radial Base And Linear Kernel Functions For Mobile Banking Customer Satisfaction Analysis

Yogyakarta, Indonesia

Putri Taqwa Prasetyaningrum[1]
Faculty of Information Technology
Universitas Mercu Buana Yogyakarta
Yogyakarta, Indonesia
putri@mercubuana-yogya.ac.id

Nurul Tiara Kadir[2]
Faculty of Information Technology
Universitas Mercu Buana Yogyakarta
Yogyakarta, Indonesia
n.tiara@gmail.com

Albert Yakobus Chandra[3]
Faculty of Information Technology
Universitas Mercu Buana Yogyakarta
Yogyakarta, Indonesia
albertch@mercubuana-yogya.ac.id

Irfan Pratama[4]
Faculty of Information Technology
Universitas Mercu Buana Yogyakarta
irfanp@mercubuana-yogya.ac.id

***Abstract*— Banking services using mobile banking applications, including Indonesian state bank (called BRI). A study on feedback regarding BRI services based on mobile applications was done. In order to compete with other banks, that is used to enhance and modernize the quality of BRI services provided to clients. Based on phenomena that occur in these situations. This study aims to classify comments from users of the BRI Mobile Banking Application on Google Play services into positive and negative comment sentiments. In this study, the Support Vector Machine (SVM) technique is utilized to determine between positive or negative reviews. The sentiment analysis of BRI google play data was carried out by comparing the Radial Basis Function (RBF) kernel function and the Linear kernel. As well as the experiment of adding feature selection, parameters, and n-grams for a period of two years, from January 1st, 2017 to December 31st, 2018. The results of the study using the k-fold cross-validation test, the precision value of the SVM kernel linear is 90.80 percent and the SVM kernel RBF is 90.15 percent. In the RBF kernel, there are 1,816 positive classes and 1,455 negative classes. While the Linear kernel obtained a positive class of 1,734 and a negative class of 1,637.**

Keywords— Sentiment Analysis, Support Vector Machine, Kernel RBF and Kernel Linear

I. INTRODUCTION

Developers of mobile applications are altering their monetization methods for tactical reasons[1]. The example of digital technology that has been promoted as a way to increase financial inclusion [2]. Mobile banking removes the use traditional system such as cash. It has an impact on how individuals business and social aspect relationship[3]. Despite research on the impact of mobile banking applications on promoting financial inclusion, society structure, age, particularly in terms of gender, etc[4]. Business analytic purpose to identify links and make comparisons between the variables that affect of banks day-to-day operations [5].

Recognizing the importance of customer evaluations as a collection of "consumer voices," we present a sentiment analysis and statistical process control method that is integrated. Data to track client complaints to improve service quality. Using data from customer reviews, conduct a control study of customer-centered service quality management. While sentiment analysis provides the systematic identification of customer satisfaction scores from customer review data, statistical process control chart analysis enables the early detection of significant customer complaints and the prevention of service failures.[6].

Machine learning is one way for sentiment analysis. It is a data extraction procedure that integrates computer science and artificial intelligence. Many types of data cannot be effectively analyzed using traditional methods to extract their information or patterns; however, machine learning, an approach to data extraction that combines computer science and artificial intelligence or computer science and computer science, is quite effective at performing data analysis[7]. Using various data processing systems and prediction algorithms, a process prediction strategy has been proposed [8]. The Data Mining process in particular can be used to process data conceptually. One of the best classifiers now in use is SVM. For SVM to determine the classification rules, only a little amount of training data is needed. Classifying new observations frequently requires less computational work and more effective memory[9]. Furthermore, based on three sentiment classifications: positive, negative, and neutral, sentiment analysis employing the support vector machine. This study expected the accuracy classification of the sentiment of mobile banking case study BRI with machine learning algorithm.

II. RELATED WORK

Banking is a prominent area for empirical and methodological research using AI methodologies[10]. Prior empirical banking research has primarily focused on the factors that influence decision-making[11]. The ability of the bank to

establish and retain client happiness is a critical indicator of its business performance. In this scenario, customer relationship management (CRM) can assist banks in achieving this critical goal. While traditional CRMs are normally run through software systems and databases and are utilized by large banks, there is indication that mobile banking, as a new emergence, is being used[12][13]. Furthermore, the widespread usage of mobile banking by internet users persuades electronic banking to adopt m-banking using a customer relationship management system[14]. The introduction of mobile banking has changed the way individuals contact with one another, prompting businesses and banks to launch their own mobile banking websites in order to communicate and interact with their clients directly[15]. Numerous academics currently feel that customer engagement can be defined as "repeated interactions between consumers and the organization that strengthens the costumer's emotional, mental, or physical investment" as a result of developments in mobile banking and the evolution of its customers[16]. In identifying user concerns and sentiments, the research conducted by Tao(2020) aims to assist users in realizing and understanding mobile application security issues[17]. On the mobile application market has provided applications for innovation and business development[18]. Many OS provide mobile bank application such as App Store and Google Play, are mobile application platforms[19] Sentiment data does not always correspond to stock market prices, and it can be only used to forecast stock prices when the stock is popular among investors [20]. News tagged as positive and negative in this system[21]Increased sales and profitability can improve bank performance. For instance, of the 500 marketing plans that included the usage of mobile banking, those that did not included the platform saw a 24% boost in sales[22].

According to earlier studies, use the selection method to select the kind of kernel function that is most matched to the supplied data. The PSO algorithm is used to determine the relevant parameters, which are then used in an SVM composed of chosen kernel function types for the classification of unknown data. Currently the development of digital technology is increasingly advanced, all forms of information have shifted from conventional (physical) forms to digital forms. This allows us to process the data with certain mechanisms to then generate knowledge that can be used for strategic purposes. In recent years there has been a trending field of science in the world of information technology related to other fields including business, namely Data Science. Data science is a field of science that specifically studies data, especially quantitative or numerical data, both structured and unstructured so that the data can provide an understanding of the problems or facts that exist[23]. Using SVM classification, Silvia Garcia-Mendez et al [2] in 2020. Their work focuses on describing novel systems using machine learning and natural language processing approaches to describe banking transactions for personal financial management. The findings demonstrated that utilizing SVM to combine these detectors generates a high accuracy value as opposed to taking complexity and computation time into account[24]. In comparison, KNN often needs more memory and processing power because it uses all input variables and training data to categorize new observations. Future applications will use convolutional neural networks, decision trees, and maximum entropy machine learning methods[21]. In Oman's research (2019), sentiment analysis uses the Support Vector Machine method with the ratio of training data and test data 70%: 30% performs better than 50%: 50% or 30%: 70% for classifying data on all five mobile payment services provider. This result is based on a model with a ratio of 70%:30% data which has a higher accuracy rate compared to the other two ratios[25].

Based on previous research, we believe machine learning algoritme used Support vector machine radial base functions and linear kernel functions could be refenrence for this case.

III. Material and Method

A. Data

In this study using a dataset obtained from the results of scrapping the Google Play site for the BRI Mobile Banking application, namely BRIMO. The data set contains a collection of reviews from users of the BRIMO application, with a year taken. The data picked is 3,398 data reviews which are then preprocessed data missing value so that the number of data becomes 3,371. We selected data with some criteria based complain or comment such as positive, negative and netral.
The initial data needed is used as an object which is then carried out by the scrapping process through Google Collaboratory to be converted into a .csv format. Here you can see tables and pictures of csv data that will be processed in Google Colab.
The variables that will be used for processing the analysis data are UserName, Content, Created At, and Score which are then labeled as Sentiment column.

B. Research Design

The following stages is the research design of this study.

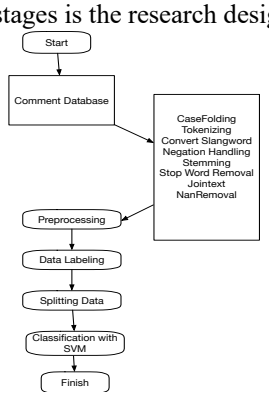


Figure 1 Research Design

Figure 1 discribe our study research, firstly we collect data in database than preprocessing, data labelling, splitting data, and classification.

C. Preprocessing

Data preprocessing is a process to convert unstructured data into structured data so that the analysis process is easier and the classification results are more accurate. Various methods are applied in the process:

1. CaseFolding Used to uniform the data so that it takes the form of lowercase, it aims to be standart using the commands in Python. In this process discribed the selecting top 10 comments.
2. Tokenizing Used to remove special text, spaces, mentions, tags, links, hashtags, delete incomplete urls, remove numbers, punctuation marks, remove spacing or spaces in each word, remove emoji, delete duplicate characters or letters.
3. Convert Slangword(Normalization) Used to convert words into standard form. In the non-standard language normalization stage, the writer uses the slangword.xls dictionary.
4. Negation Handling (Negation)Used to overcome the problem of negation of a word. In this process to overcome the problem of negation of a word, negation is very influential on the polarity of other words and the accuracy of a sentiment analysis. There is an example of the sentence "This application is not good, often errors" the meaning of the sentence means dislike. However, in the Natural Language Processing stopwords process, the word is not included in the words contained in the stopwords dictionary. So that if the stopwords process is run, the word no will disappear and the sentence will change to the sentence "This application is good, often errors". Of course this will result in an error that should not be ignored. If it is ignored, then the analysis process will contain reviews from users who are categorized into a negative class and will be assessed as a positive class, thus making the classification prediction less accurate. After the negation process is carried out, the sentence "This application is not good, often errors" will become "This application is not good, often errors". Words that are at the beginning or end of words that are included in the list of negation words that have been determined by the author will be combined into one word. A list of negation words that are made such as the words less, no, not yet, don't, and not.
5. Stemming Used to change words into root words. Stemming is the process of changing additional words into basis. Before doing the stemming process, then need to install the literary library. Literature is a nlp library specifically for indonesian language
6. Stopwords Removal Used to remove words that are not important. This process is useful for removing unimportant words. In this process also adding some words in the stopwords list like the word : yg, dg, dgn, rt, kalo, biar, biki, klo, dan sebagainya.
7. JointText Used to return words that are split into sentences. This process is to restore words that were previously in the form of word fragments to their original form, namely to become a complete sentence.
8. NanRemoval Used to remove empty or null data. Nan data is empty or null data. By writing the df.dropna() syntax, the empty data in the scrapping results is deleted so that there is a difference in the amount of data used

D. Dataset Labeling

After the preprocessing process is complete, then the next process is review data labeling. This process is done by applying a dictionary lexicon obtained from https:github.com/masdevid/ID-OpinionWords. Dictionary The lexicon used has more than 3,371 common words in Bahasa Indonesia which is accompanied by negative sentiments and positive sentiments. Define labels based on values total words. If a word has a value = 0, then the word is not included in the lexicon dictionary, while if it has a value = 1 then is included in the positive class on the contrary if the value is obtained -1 then it is included in the negative class. The results of the review data labeling process can be seen in the table. 1

Table 1. Review Labeling Result

Class	Amount of Data
Positive	1.734
Negative	1.637

E. Splitting data

A good classification process is supported by several things in order to obtain machine learning model capable of predicting new data classes, namely data needed for the training and testing process. The dataset is divided into training data and test data using K-Fold Validation where the value of k = 1 0, and using a ratio of 0.1 . This shows that 90% of the data used for the training process and 10% for the testing process. K-Fold Validation This is used so that the training process is more accurate so that the prediction results will be better good.

F. Classification on Support Vector Machine (SVM)

This process aims to determine the negative class and positive class. In the text classification process, SVM has kernel functions and parameters whose value can be determined and the difference in the value of the parameter will be effect on the performance of the model. Therefore, the testing process is carried outin the form of tuning parameters for each kernel function so that it can produce the best machine learning models. The tests carried out are assigning a value to each the parameters used are the value of C, Gamma, Percentile, and N-gram

SVM function in this research by using RBF kernel SVM. In analyzing the RBF kernel function, the cost (C) and Gamma (γ) parameters are optimized. The analysis was carried out

using a sample of 3,398 data from BRIMO which was divided into two parts, 90% as a training dataset and another 10% as a test dataset. In determining the best parameters for the RBF kernel, trial and error was also carried out.

IV. Result And Discussion

The preprocessing stage in this study consists of several processes, namely case folding, data cleansing, tokenizing, converting slangwords (normalization), negation handling, stemming, and stopwords removal. The following are the results of case folding preprocessing. Case Folding is used to uniform the data so that it is in the form of a lowercase. Displayed for the top 10 records. Can be seen in Figure 2.

```
Case Folding Result :

0    bri saya kenapa lagi sih... terus terkendala d...
1    sangat menyebalkan. sering terblokir tiba tiba...
2    saya buka rekening online via brimo,sudah dpt ...
3    masa udah ganti pasword, sukses. untuk log in ...
4    saya mau daftar ke bank terdekat, malah ribet,...
5    setelah update, user terblokir 🤦. saya buka re...
6    maaf saya kasih bintang 2 karna seringkali tib...
7    ribet, gak kayak aplikasi bank yg lain, kalo m...
8    gk jelas lah,, kalau memang di upgrade upgrade...
9    bri sangat tidak transparan terhadap kekuranga...
Name: content, dtype: object
```

Figure 2. Case Folding Results

Before the tokenizing process is carried out, you must install several libraries and modules such as punkt, regex, string, itertools, nltk. In this tokenization process there is a process to remove special text, spaces, mentions, tags, links, hashtags, delete incomplete urls, remove numbers, punctuation marks, remove spacing or spaces in each word, remove emoji, remove duplicate characters or letters. The results displayed are the top 10 data. Can be seen in Figure 3.

```
Tokenizing Result :

0    [bri, saya, kenapa, lagi, sih, terus, terkenda...
1    [sangat, menyebalkan, sering, terblokir, tiba,...
2    [saya, buka, rekening, online, via, brimosudah...
3    [masa, udah, ganti, pasword, sukses, untuk, lo...
4    [saya, mau, daftar, ke, bank, terdekat, malah,...
5    [setelah, update, user, terblokir, saya, buka,...
6    [maf, saya, kasih, bintang, karna, seringkali,...
7    [ribet, gak, kayak, aplikasi, bank, yg, lain, ...
8    [gk, jelas, lah, kalau, memang, di, upgrade, u...
9    [bri, sangat, tidak, transparan, terhadap, kek...
Name: content, dtype: object
```

Figure 3. Tokenizing Results

After the data preprocessing process is carried out, the data labeling process is carried out. Labeling uses preprocessed data using a lexicon dictionary which is calculated based on the sum of the weights of each word. If the sum of the weights obtained results in (-) minus or (<1) less than one, it will be included in the negative label. Meanwhile, if on the contrary, if the sum of the weights of each word produces a result of more than 1 or (> 1) then it will be included in the positive label. Can be seen in Figure 4.

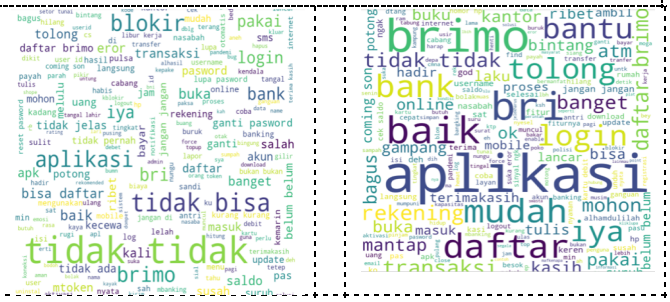
	content	userName	at	sentiment	Weight
0	bri kendala saldo transfer uang sat cek saldo ...	Dahmelia Napitupulu	2021-08-18 17:15:34	negative	0 + -1 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + ...
1	sebal biokir pagi pagi lancar siang blokir dib...	Fitriyani Purwati	2021-08-20 14:27:58	negative	0 + -1 + 0 + 0 + 1 + 0 + -1 + 0 + 0 + -1 + 0 ...
2	buka rekening online via brimosudah nomor reke...	Ady Suprianto	2021-08-18 11:32:24	positive	0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0...
3	ganti pasword sukses log in pasword tidak tida...	yudhie sasongko	2021-08-20 08:23:35	negative	0 + 0 + (1 * -1) + 0 + 0 + 0 + -1 + -1 + 1 ...
4	daftar bank ribet ktp buku rekening atm nomor ...	Mahkluk Nocturnal	2021-08-22 14:03:58	negative	0 + 0 + -1 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + ...

Figure 4. Tokenizing Results

Here are the results from experiments on both kernels. Experiments by adding chi square selection features, N-Gram extraction features, and using C and Gamma parameters as well as experiments without adding features. Then the following results are produced.

The following is a visualization of the word predictions that often appear in both kernels. The results of this visualization can later be used as an evaluation chart for the developer. Visualization of Word Prediction in Linear Kernel Negative dan positive can seen in Figure 5.

Figure 5. Visualization of Word Prediction in Linear Kernel



After doing an analysis using the RBF kernel function and the Linear kernel from the review data of the BRIMO mobile banking application. The most appropriate kernel function to utilize to assess the precision of sentiment analysis class categorization will then be determined. The following is a summary table of the accuracy values of the two kernel functions after the experiment with the addition of feature selection, n-gram, C and gamma parameters and experiments without additional features. Can be seen in Table 2 and Table 3.

Table 2. Experiment Result

Classification Model	Accuracy(%)	
	Training	Testing
SVM linear kernel with N-grams (1,2), Chi-Square	97.40	92.80

feature selection (75% percentile), parameters C = 1 and gamma = 0.01		
SVM linear kernel without N-grams, feature selection, C and gamma parameter parameters	95.81	90.26
SVM kernel RBF with N-grams (1,2), Chi-Square feature selection (percentile 100%), parameters C = 100 and gamma = 1	98.90	93.00
SVM kernel RBF with N-grams (1,2), Chi-Square feature selection (percentile 100%), parameters C = 100 and gamma = 1	98.00	90.43

From table 2. explained training dataset Taking training dataset and testing the dataset is done randomly with the help of python software. Support Vector Machine (SVM), a classification method used in data mining, will be employed, as described in the study methodology in the previous chapter, to discover the best accuracy value. The results based on table 2, RBF kernel has the best accuracy value. Utilization of the RBF kernel with parameters gamma=1 and C=100, N-Gram (1,2), Chi Square feature selection (Percentile 100%) is able to produce an accuracy value of 96.70% on training data and 92.43% on testing data. For the comparison results obtained the results shown in table 3.

Table 3. Accuracy Comparison.

Classification Model	Class	Accuracy (%)		
		Precision	Recall	F1-Score
SVM kernel Linear	0	93.90	90.59	92.22
	1	90.80	94.05	92.40
SVM kernel RBF	0	93.33	90.59	91.94
	1	90.15	93.45	92.08

Based on table 3 explained the accuracy value of the two kernel functions after experimenting with the addition of feature selection, n-gram, C and gamma parameters and experiments without additional features. Nilai 0 menunjukan positif , dan 1 negative, dalam komparisi ini menunjukan SVN kernel liner memiliki akurasi sintimen tertinggi,baik yang sentimen negatif yaitu 92.40, maupun yang sentimen positif yaitu 92.22.
Based on table 3 explained the accuracy value of the two kernel functions after experimenting with the addition of feature selection, n-gram, C and gamma parameters and experiments without additional features. A value of 0 indicates positive and 1 is negative, in this comparison it shows the SVN kernel linier has the highest sentiment accuracy, both negative sentiment is 92.40 and positive sentiment is 92.22.

V. CONCLUSSION

The experimental results show that the use of complete preprocessing, parameter settings, and n-gram feature selection can improve the performance of the model used compared to without the addition of features. From the analysis results obtained visualization of words that often appear between the two kernel models, from the sentiment class it can be seen which words often appear and can be used as evaluation material for BRI. Based on the results testing obtained the best accuracy on the RBF kernel with parameters gamma =1 and C=100, N-Gram (1 ,2), while feature selection Chi Square (Percentile 1 00%) produces an accuracy value on the training data of 96.70% and testing data 92.43%. In the RBF kernel, there are 1,455 negative classes and 1,816 positive classes. While the Linear kernel obtained a negative class of 1,637 and a positive class of 1,734. This automation system is still running on Google Colab, so it cannot retrieve review data in real time.

VI. FUTURE RESEARCH

Future work of this research is to be carried out with larger datasets to ensure the accuracy of the analysis. It is necessary to do a comparison using other optimization methods as a comparison in determining the best optimization method

REFERENCES

[1] Y. J. Lee, H. Ghasemkhani, K. Xie, and Y. Tan, "Switching decision, timing, and app performance: An empirical analysis of mobile app developers' switching behavior between monetization strategies," *J. Bus. Res.*, vol. 127, no. February, pp. 332–345, 2021, doi: 10.1016/j.jbusres.2021.01.027.

[2] M. Chibba, "Financial inclusion, poverty reduction and the millennium development goals," *Eur. J. Dev.*

Res., vol. 21, no. 2, pp. 213–230, 2009, doi: 10.1057/ejdr.2008.17.

[3] T. Correa and I. Pavez, "Digital Inclusion in Rural Areas: A Qualitative Exploration of Challenges Faced by People From Isolated Communities," *J. Comput. Commun.*, vol. 21, no. 3, pp. 247–263, 2016, doi: 10.1111/jcc4.12154.

[4] A. van Klyton, J. F. Tavera-Mesías, and W. Castaño-Muñoz, "Innovation resistance and mobile banking in rural Colombia," *J. Rural Stud.*, vol. 81, no. October 2020, pp. 269–280, 2021, doi: 10.1016/j.jrurstud.2020.10.035.

[5] I. González-Carrasco, J. L. Jiménez-Márquez, J. L. López-Cuadrado, and B. Ruiz-Mezcua, "Automatic detection of relationships between banking operations using machine learning," *Inf. Sci. (Ny).*, vol. 485, pp. 319–346, 2019, doi: 10.1016/j.ins.2019.02.030.

[6] J. Kim and C. Lim, "Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses," *Adv. Eng. Informatics*, vol. 49, no. April, p. 101304, 2021, doi: 10.1016/j.aei.2021.101304.

[7] R. K. Tripathi, A. S. Jalal, and S. C. Agrawal, "Suspicious human activity recognition: a review," *Artif. Intell. Rev.*, vol. 50, no. 2, pp. 283–339, 2018, doi: 10.1007/s10462-017-9545-7.

[8] D. A. Neu, J. Lahann, and P. Fettke, "A systematic literature review on state-of-the-art deep learning methods for process prediction," *Artif. Intell. Rev.*, no. 0123456789, 2021, doi: 10.1007/s10462-021-09960-8.

[9] T. Sontayasara *et al.*, "Twitter sentiment analysis of bangkok tourism during covid-19 pandemic using support vector machine algorithm," *J. Disaster Res.*, vol. 16, no. 1, pp. 24–30, 2021, doi: 10.20965/jdr.2021.p0024.

[10] M. Doumposa, Zopounidisab, C. Gounopoulosc, E. Platanakisc, and W. Zhangc, "Operational research and artificial intelligence methods in banking," *Eur. J. Oper. Res.*, pp. 1–16, 2022.

[11] K. Rajaratnam, P. Beling, and G. Overstreet, "Regulatory capital decisions in the Context of consumer loan portfolios ga," *J. Oper. Res. Soc.*, vol. 68, no. 7, pp. 847–858, 2017, doi: 10.1057/jors.2016.7.

[12] J. H. Wu and S. C. Wang, "What drives mobile commerce? An empirical evaluation of the revised technology acceptance model," *Inf. Manag.*, vol. 42, no. 5, pp. 719–729, 2005, doi: 10.1016/j.im.2004.07.001.

[13] Z. He, "Rivalry, Market Structure and Innovation: The Case of Mobile Banking," *Rev. Ind. Organ.*, vol. 47, no. 2, pp. 219–242, 2015, doi: 10.1007/s11151-015-9466-z.

[14] H. Hamidi and M. Safareeyeh, "A model to analyze the effect of mobile banking adoption on customer interaction and satisfaction: A case study of m-banking in Iran," *Telemat. Informatics*, vol. 38, no. June 2018, pp. 166–181, 2019, doi: 10.1016/j.tele.2018.09.008.

[15] M. N. Hajli, J. Sims, M. Featherman, and P. E. D. Love, "Credibility of information in online communities," *J. Strateg. Mark.*, vol. 23, no. 3, pp. 238–253, 2015, doi: 10.1080/0965254X.2014.920904.

[16] R. J. Brodie, L. D. Hollebeek, B. Jurić, and A. Ilić, "Customer engagement: Conceptual domain, fundamental propositions, and implications for research," *J. Serv. Res.*, vol. 14, no. 3, pp. 252–271, 2011, doi: 10.1177/1094670511411703.

[17] C. Tao, H. Guo, and Z. Huang, "Identifying security issues for mobile applications based on user review summarization," *Inf. Softw. Technol.*, vol. 122, no. October 2019, p. 106290, 2020, doi: 10.1016/j.infsof.2020.106290.

[18] H. Liao and C. Yuan, "the Effect of Operating Systems and Types of Apps on Purchase Intention," *Issues Inf. Syst.*, vol. 16, no. Iii, pp. 156–163, 2015, doi: 10.48009/3_iis_2015_156-163.

[19] A. Ghose and S. P. Han, "Estimating demand for mobile applications in the new economy," *Manage. Sci.*, vol. 60, no. 6, pp. 1470–1488, 2014, doi: 10.1287/mnsc.2014.1945.

[20] K. Guo, Y. Sun, and X. Qian, "Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market," *Phys. A Stat. Mech. its Appl.*, vol. 469, pp. 390–396, 2017, doi: 10.1016/j.physa.2016.11.114.

[21] T. Yu and K. T. Nwet, "Comparing SVM and KNN algorithms for myanmar news sentiment analysis system," *PervasiveHealth Pervasive Comput. Technol. Healthc.*, pp. 65–69, 2020, doi: 10.1145/3379247.3379293.

[22] J. Boehmer and S. Lacy, "Sport News on Facebook: The Relationship Between Interactivity and Readers' Browsing Behavior," *Int. J. Sport Commun.*, vol. 7, no. 1, pp. 1–15, 2014, doi: 10.1123/ijsc.2013-0112.

[23] V. Dhar, "Data Science and Prediction," *Commun. ACM*, vol. 56, no. 12, pp. 64–73, 2013, doi: 10.1145/2500499.

[24] S. Garcia-Mendez, M. Fernandez-Gavilanes, J. Juncal-Martinez, F. J. Gonzalez-Castano, and O. B. Seara, "Identifying Banking Transaction Descriptions via Support Vector Machine Short-Text Classification Based on a Specialized Labelled Corpus," *IEEE Access*, vol. 8, pp. 61642–61655, 2020, doi: 10.1109/ACCESS.2020.2983584.

[25] O. F. Prihono and P. K. Sari, "Comparison Analysis Of Social Influence Marketing For Mobile Payment Using Support Vector Machine," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, pp. 367–374, 2019, doi:

10.22219/kinetik.v4i4.921.