



# INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : [www.joiv.org/index.php/joiv](http://www.joiv.org/index.php/joiv)



## Sentiment Analysis of Neobank Digital Banking Using Support Vector Machine Algorithm in Indonesia

Kusnawi <sup>a,\*</sup>, Majid Rahardi <sup>a</sup>, Van Daarten Pandiangan <sup>a</sup>

<sup>a</sup> Faculty of Computer Science, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

Corresponding author: \*[khusnawi@amikom.ac.id](mailto:khusnawi@amikom.ac.id)

**Abstract**—Currently, in the industrial era 4.0, information and communication technology is very developed, whereas, in this era, there is an increase in complex activities, one of which is in the banking sector. With the ease and efficiency of online finance, people want to switch to using digital banks. Neobank is an online savings and deposit application from Bank Neo Commerce (BCN) that the public can use by using the Internet. One of the online services is mobile banking which can be used by both Android and iOS versions of customers. Users can review Neobank's performance and services through the Google Play Store to improve and evaluate Neobank's performance. Neobank application reviews on the Google Play Store are increasing. Therefore, a review analysis is needed by conducting a sentiment analysis on Neobank's review. The data amounted to 3159 user reviews collected from reviews of the Neobank application on the Google Play Store. This study aims to classify Neobank user review data, including positive or negative sentiments. The method used in this study is an experimental method using the Support Vector Machine algorithm. The accuracy results obtained using the Support Vector Machine algorithm are 82.33%, which is owned by the scenario of 90% training data and 10% test data. The precision results are 82%, and recall is 81%. Future studies can add datasets from various sources so that there are even more datasets so as to increase the accuracy of model classification.

**Keywords**—Sentiment; Neobank; support vector machine; play store.

Manuscript received 15 Feb. 2023; revised 1 Apr. 2023; accepted 2 Jun. 2023. Date of publication 30 Jun. 2023. International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



### I. INTRODUCTION

The development of digital technology is increasingly rapid. Even now, almost all activities carried out by the community are increasingly easy because of the very advanced steps of digital technology. Therefore, Indonesia may be aggressive in carrying out several programs and accelerating changes in digital technology, which is almost echoed as a national program for almost all industries. This change in digital technology has also led to the banking industry, which is carried out independently by the government and supported by the owners of the banking business themselves[1]–[4].

The rapid development of the banking industry and the large number of internet users in Indonesia (196,714,070 people) show that this can be a big market opportunity for the banking industry[5]. Therefore, this is the trigger for the presence of digital banks in Indonesia. The presence of digital banks is currently very influential on some people in Indonesia, and this is because digital banks aim to increase operational activities' efficiency and customer service quality. Using electronic means as the foundation allows customers or prospective customers to obtain information, communicate, register, open

accounts, transactions, and close accounts independently. Digital banking services want to make it easier for customers to manage their finances, not waste time and energy, and even cost money to take care of banking services.

Every banking industry that has a digital bank certainly has its own strategy to continue to survive during very tight competition and compete to provide the best service for the people who are its customers. Banking industry companies need to know what features or facilities are advantages and disadvantages according to users to be used as benchmarks in evaluations to improve the quality of the company[6]–[9]. One of the public companies engaged in banking is Bank Neo-Commerce (BNC). In providing services to their customers, Neo Commerce Bank opens much input for their services, ranging from social media accounts to applications provided by Neo Commerce Bank, namely NeoBank as a place for customer activities in conducting transactions. Currently, NeoBank on the Google Play Store has been downloaded more than 10 million with a rating of 3.9 and 150 thousand reviews. The fairly low rating, accompanied by various negative and positive reviews, indicates that the services provided by Neo

Commerce Bank have not fully met the expectations of NeoBank users. Therefore, an analysis is needed, namely sentiment analysis using the Super Vector Machine (SVM) classification method through more than 3000 data reviews. It is hoped that the results obtained can be used as material for evaluating the NeoBank application to improve and meet the needs of its users.

## II. MATERIALS AND METHOD

### A. Neobank

Neobank is a development as well as innovation in the field of technology, namely finance, which provides digital banking services without having branches or physical forms like most other banks. Neobank comes online and is used online too. This is done so that technology-savvy customers do not hesitate to manage their finances through a mobile application. Neobank is a digital bank product founded by PT. Bank Neo Commerce, Tbk, on January 09, 1990. Neobank began to be known and quickly touched its market after the Financial Services Authority had confirmed its upgraded status as a book bank II in 2020. More than 10 million downloads are available on the Google Play Store, and there are 2 million more Neobank daily active users until the end of 2021. The Neobank application was also in the Top Charts position on the Google Play Store in free applications in all categories in 2021. Even today, public interest in downloading the Neobank application is increasing. The advantages that Neobank has compared to other digital banks are that opening an account can be done without any fees, the interest on savings offered is quite high, customers are also facilitated by the transaction features that are quite easy to understand, and can take advantage of the Neobank application as a means of compiling expenses and income and withdrawals money can be made through outlets such as Indomaret and Alfamart[10].

### B. Google Play Store

Google Play Store is an official application store created by Google that can be used on Android devices to run it. This Android application store was published on October 22, 2008, with the previous name Android Market, and then in March 2012, it was changed to Google Play Store. This application store offers several free and paid programs, such as music, books, movies, and television, which can be directly accessed via Android devices, websites, and Google TV[11]–[14].

### C. Sentiment Analysis

Sentiment analysis is a technique of extracting text data that is usually opinionated and aims to analyze, understand, and process text data with entity relationships, such as products, services, organizations, people, and a particular topic[15]–[18].

### D. Online Review

One form of electronic word of mouth (E-WOM) or word of mouth communication is Online Review. Online reviews are one of the real media that can be used for consumers to see input from other consumers in the form of pictures or sentences on a product, tool, or service that the company has implemented. Online Review is a tool that consumers can use to find information that will later affect purchasing decisions

or product use [19]. Online review is also a communication channel that can be a fairly effective promotion method used by companies because this method does not require high costs. It is enough through satisfied customers with the company's products, tools, and services[20]. With reviews from satisfied users, the resulting product that has been produced will become a product recommendation to the public, who are consumers or users of the product. The company also gets feedback from customers who have used the company's products which can then be used as a design material to improve company productivity[21].

### E. TF-IDF

The Term Frequency-Inverse Document Frequency (TF-IDF) method is a method of assigning weight to the relationship of a word (term) to a document[22]–[24]. This technique combines two weight calculation techniques, namely Term Frequency (TF), which is the frequency of the number of occurrences of the word (t) in the sentence (d). Document Frequency (DF) is the number of sentences in which a word (t) appears. Here is the TF-IDF formula:

$$TF-IDF(q, d) = \sum_{t \in q} tf(t, d) \times idf(t, d) \quad (1)$$

### F. Cosine Similarity

Cosine Similarity is a method used to calculate the level of similarity between two objects[25]. Mathematical calculations are used to calculate the most dominating similarity between two fields. The standard calculation used to calculate the similarity of two documents between  $x_1$  and  $x_2$  is represented by the cosine size between vectors with the following formula:

$$\text{Cosine Similarity } (x_1 + x_2) = \frac{V(x_1) \cdot V(x_2)}{|V(x_1)||V(x_2)|} \quad (2)$$

$|V(x_1)|$  = vector length  $x_1$

$|V(x_2)|$  = vector length  $x_2$

$V(x_1) \cdot V(x_2)$  = dot product of  $x_1$  and  $x_2$

### G. Support Vector Machine

Support Vector Machine Classification is a machine learning system using a hypothetical space in the form of a linear function in a large-dimensional feature space, trained with a learning algorithm based on optimization theory by implementing the learning deviation generated from the outgoing statistical learning theory[26]. Support Vector Machine is included in the type of classifier, which is binary, linear, and nonlinear probabilistic. The first underlying understanding of the classification with SVM is finding the optimal hyperplane line. Support Vector Machine is a classification method that uses Machine Learning (Supervised Learning) to predict classes through models or patterns based on the results of the training process. In predicting or separating one class from another, grading is carried out to find a hyperplane or decision boundary with a predetermined margin. This plays a role in separating positive sentiment reviews (marked +1) and negative reviews (marked -1). The largest margin will be determined by calculating the maximum distance from the hyperplane, the dividing line between the two classifications. Then the hyperplane can see

and separate the resulting pattern against the values of the two different input spaces. After that, it is necessary to calculate the entire input space to get the closest value to the hyperplane. After getting the closest value for each class hyperplane, the Support Vector can be done[27]–[29].

#### H. Confusion Matrix

A confusion Matrix is a performance measurement for a machine learning classification problem where the output of this confusion matrix can be in the form of two or more classes[16]. A confusion matrix is a table consisting of the number of rows of data being tested that are predicted to be correct and one of the classification models, and this aims to determine the performance of a classification model used.

TABLE I  
CONFUSION MATRIX

	Actual True	Actual False
True Predicted	True Positive (TP)	False Positive (FP)
False Predicted	False Negative (FN)	True Negative (TN)

There are four terms that represent the results of the classification process in the confusion matrix, which can be

seen in Table 1. The performance evaluation measurement uses precision, recall, f1-score, and accuracy techniques to see the advantages of the developed model.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1\ Score = \frac{2(Precision)(Recall)}{(Precision+Recall)} \quad (5)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

#### I. Dataset

This type of study uses a quantitative approach. Quantitative studies are a type of research that is systematic and numerical and must also be structured from the start. This quantitative study was conducted to examine the reviews contained in the Neobank application, which was the sample in the study, which would then be identified to find a decision. Quantitative research uses hypotheses as research instruments. The research flow in this study can be seen in Figure 1.

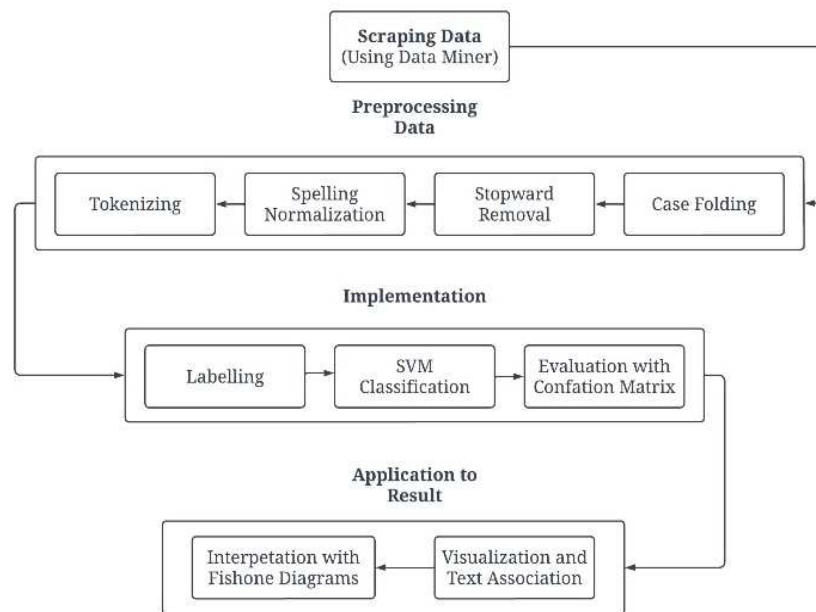


Fig. 1 System Overview of Sentiment Analysis

This study begins by preparing the dataset requirements taken from online review comments on the google play store on the Neobank application. The dataset used in this study is 3159 data consisting of 1207 positive labels, 917 negative labels, and 1035 neutral labels. Data retrieval using scraping method with Google Chrome Extensions, namely Data Scraper version 5.6.2.6. The dataset source link is as follows <https://play.google.com/store/apps/details?id=com.bnc.finance&hl=in&gl=US&showAllReviews=true>.

In this study, data collection was carried out online using the Data Scraper application. This technique, called data scraping, takes the URL or application link contained in the Google Play Store, which will then be used as raw data. An example of a scraping dataset can be seen in Table 2.

TABLE II  
RESULTS OF DATASET SCRAPING

Username	Rating	Time	Content
Suyatno Yuti	5	2022-04-20 12:03:40	Jujur loh ini penilaian saya pribadi neo
Indra Diyet	5	2022-04-20 12:02:41	Bagus banget
Iyan Official	4	2022-04-20 11:55:33	Isi pulsa kok lama sekali, aduh saldo kepotong tapi pulsa proses terus... gimana ni

### J. Pre-processing Data

The results of the data obtained at the data scraping stage cannot be used directly because it still has the form of a word or sentence that has not been structured [15], [24], [27]. At this stage, the sentence is changed into a structured one using tools such as Python. The following is the preprocessing stage:

1) *Case Folding*: At this level, it works to adjust letters or change to lowercase letters and delete punctuation in sentences. This is to say those which have upper and lower case letters are not processed because they have different meanings, which can be seen in Table 3.

TABLE III  
RESULTS OF CASE FOLDING PROCESS

Data before Case Folding	Data after Case Folding
APLIKASI yang sangat BAGUS	aplikasi yang sangat bagus
Bank digital paling NOMOR 1	bank digital paling nomor 1

2) *Stopword Removal*: Stopword removal functions to remove words that do not affect the sentence without reducing the meaning of the sentence. Stopword removal reduces the dataset size because there will be a reduction in data Friday. Here are the changes to the review sentence:

TABLE IV  
RESULTS OF STOPWORD REMOVAL

Data before Stopword	Data after Stopword
Sangat menyenangkan menggunakan aplikasi ini	menyenangkan menggunakan aplikasi
Aplikasi nya bagus sangat membantu	aplikasi bagus membantu

3) *Spelling Normalization*: Correct the abbreviations or slang words in the sentences at this stage. For example, the word "tdk" is changed to "no", "lg" is changed to "again". Here is an example of correcting a review sentence.

TABLE V  
RESULTS OF SPELLING NORMALIZATION

Data before Spelling	Data after Spelling
aplikasi yg sgt bagus baik kernn	aplikasi yang sangat bagus baik keren

4) *Tokenizing*: At this stage, the text is separated into words using a comma (,). Here is an example of changing sentences review:

TABLE VI  
RESULTS OF TOKENIZING

Data before Tokenizing	Data after Tokenizing
Saya sangat suka	'saya', 'sangat', 'suka'
Aplikasinya mantap bagus	'aplikasinya', 'mantap',
Proses transfer cepat	'bagus', 'proses', 'transfer', 'cepat'

### K. Data Labeling

Labeling is a way of grouping that can be done manually & using Lexicon-based. Manual labeling is done by knowing the domain of individual and language understanding. The lexicon-based method is done using a dictionary called the Lexicon dictionary. The Lexicon Dictionary is a dictionary that is used as a reference to group these words containing opinions or not that have been equipped with word weights. A lexicon is a collection of well-known sentiment words and

collected. Labeling requires a dictionary or lexicon containing the word that contains sentiments which is called sentiment dictionaries.

TABLE VII  
RESULTS OF LABELLING PROCESS

Username	Content	Polarity_score	Polarity
Iyan Official	isi pulsa kok lama sekali adauh saldo kepotong...	-15	negative

### III. RESULTS AND DISCUSSION

If the training and test data have been divided accordingly, the next step is to classify using the Support Vector Machine method (SVM), at this stage, using a linear kernel classification. Classification non51 linear using the linear kernel, polynomial, sigmoid, and radial kernel functions basis functions (RBF). The kernel is a function used in the SVM method used to solve a problem. The parameter selection stage optimization aims to obtain maximum results, hence the need for testing using the SVM function. Through the four scenarios that have been carried out in the linear kernel SVM algorithm, the accuracy results obtained from the four scenarios are:

TABLE VIII  
SUPPORT VECTOR MACHINE SCENARIO RESULTS

Scenario	Accuracy
60:40	77,37%
70:30	79,02%
80:20	78,71%
90:10	82,33%

Table 8 is the result of the accuracy value in the SVM algorithm. Based on the results obtained values using linear and non-SVM algorithms linear with the comparison of training data and test data 90:10 resulting in accuracy the largest is 82.33%. In the evaluation stage, predictions are made using test data and evaluation with a confusion matrix.

TABLE IX  
CONFUSION MATRIX RESULTS

	Precision	Recall	F1-Score
Positive	89%	82%	86%
Neutral	79%	73%	76%
Negative	77%	88%	82%
Average	82%	81%	81%
Accuracy		82%	

#### A. Visualization and Association

The visualization and association process uses positive and negative sentiments in this stage. This stage aims to obtain information that often appears based on the perception of NeoBank application users. Stage visualization with bar charts and word cloud to make the process easier analysis [30]–[33]. While the association stage aims to find out words that are often used appear on positive, neutral, or negative sentiments.

1) *Positive Sentiment*: The results obtained are the distribution of the word on the positive sentiment.

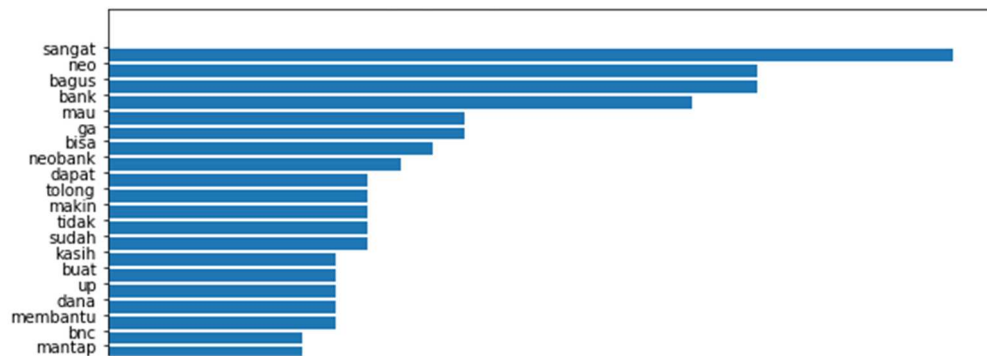


Fig. 2 Frequently Appearing Words (Positive)



Fig. 3 Wordcloud (Positive)

TABLE X  
POSITIVE SENTIMENT WORD ASSOCIATION RESULTS

Antecedents	Consequents	Confidence	Lift
(luar)	(semoga, neobank, terima, aplikasi, transaksi)	1.000000	58.911017
(makin, luar)	(semoga, neobank, terima, aplikasi, transaksi)	1.000000	58.911017
(terbaik)	(yang)	0.757426	5.732439

2) *Neutral Sentiment*: The results obtained are the distribution of the word on the neutral sentiment.

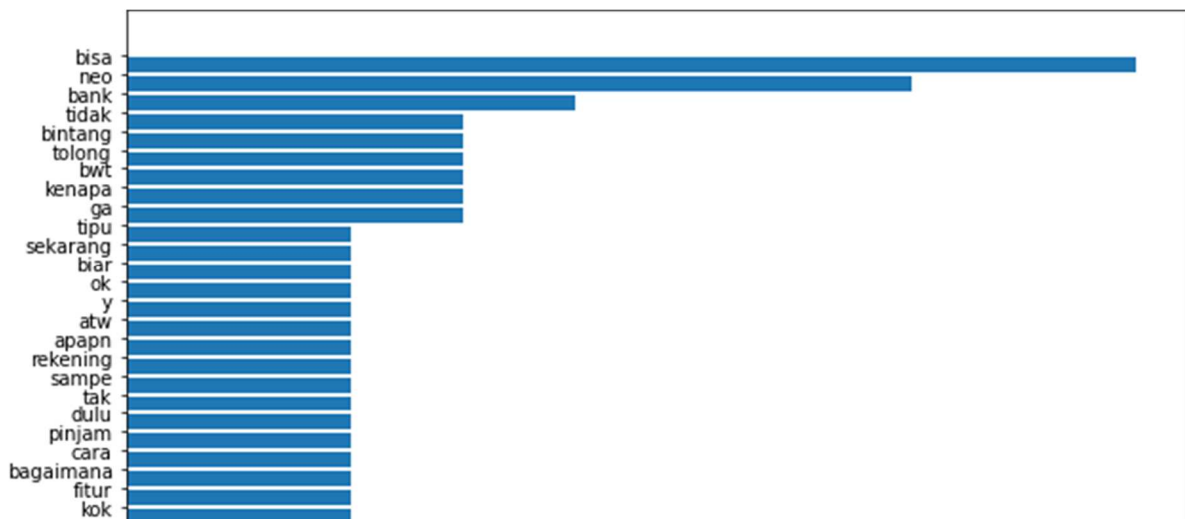


Fig. 4 Frequently Appearing Words (Positive)



Fig. 5 Wordcloud (Neutral)

TABLE XI  
NEUTRAL SENTIMENT WORD ASSOCIATION RESULTS

Antecedents	Consequents	Confidence	Lift
(neo, uang)	(di, bank, tidak)	1.000000	62.626667
(cara)	(bagaimana)	1.000000	59.455696
(bagaimana)	(cara)	1.000000	59.455696
(uang)	(di, bank, neo, tidak)	0.821168	51.427007

3) *Negative Sentiment*: The results obtained are the distribution of the word on the neutral sentiment.



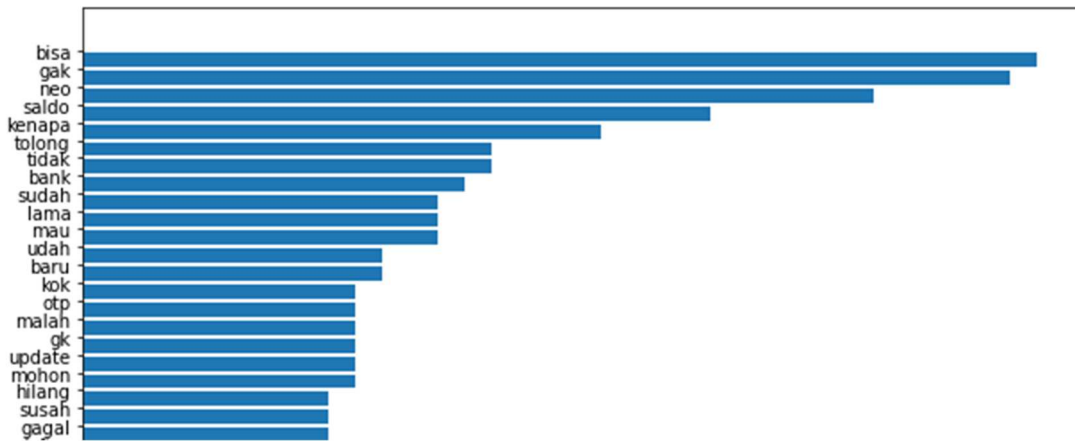


Fig. 6 Frequently Appearing Words (Negative)



Fig. 7 Wordcloud (Negative)

TABLE XII  
NEGATIVE SENTIMENT WORD ASSOCIATION RESULTS

Antecedents	Consequents	Confidence	Lift
(kepada)	(apk, gagal, untuk,	1.000000	61.462295
(bug)	saya, kalo, bug,	0.600394	36.901575
(kalo)	neo,	0.381727	23.461827
(gagal)	ini)	0.291309	17.904489

### B. Result Interpretation

This stage analysis with Fishbone Diagram based on the results of the negative review sentiment obtained information that is a constraint or problem submitted by NeoBank application users.

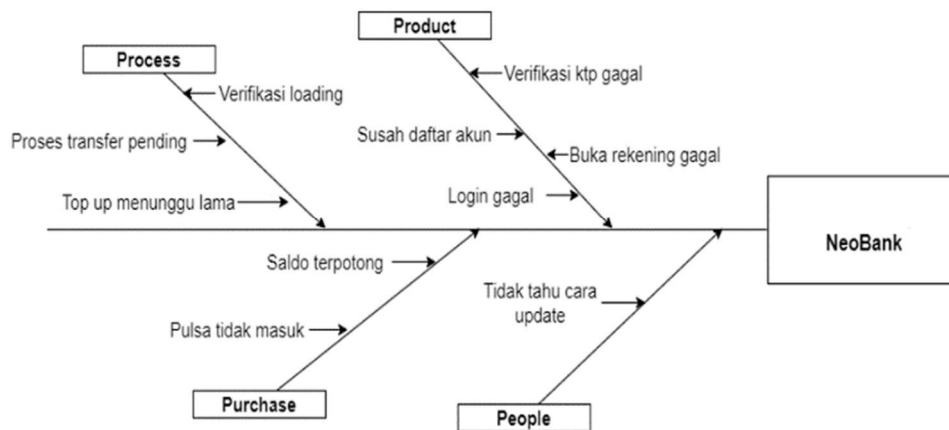


Fig. 8 Fishbone Diagram

## IV. CONCLUSIONS

The classification process uses the Support Vector Machine algorithm to get the highest accuracy value which is 82.33%, owned by a scenario of 90% training data and 10% test data as large as 82.33%. In the positive association rule, some words have the highest frequency, namely the word 'terbaik', which is associated with the words 'semoga' and 'sejak'. This shows that the Neobank application is one of the best online banking they have used, and they hope that the application will provide the best service. The association rule

of negative sentiment shows that the words 'baru' and 'aplikasi' are associated with the words 'bisa' and 'tidak'. Then the fishbone diagram also shows the people factor containing the review sentence 'Tidak tahu cara update'. The results of the association rules and fishbone diagrams further strengthen the statement that it is true that there is a lack of education on usage in the Neobank application as an important facility that needs to be provided to users.

## ACKNOWLEDGMENT

We are grateful to the research department of AMIKOM Yogyakarta University for supporting this research.

## REFERENCES

- [1] H. Luo, L. Lin, K. Chen, M. F. Antwi-Afari, and L. Chen, "Digital technology for quality management in construction: A review and future research directions," *Dev. Built Environ.*, vol. 12, p. 100087, Dec. 2022, doi: 10.1016/J.DIBE.2022.100087.
- [2] J. Wang, X. Ma, J. Zhang, and X. Zhao, "Impacts of digital technology on energy sustainability: China case study," *Appl. Energy*, vol. 323, p. 119329, Oct. 2022, doi: 10.1016/J.APENERGY.2022.119329.
- [3] J. C. Kouladom, M. A. K. Wirajing, and T. N. Nchofoung, "Digital technologies and financial inclusion in Sub-Saharan Africa," *Telecomm. Policy*, p. 102387, Jun. 2022, doi: 10.1016/J.TELPOL.2022.102387.
- [4] M. Iranmanesh, M. Ghobakhloo, M. Nilashi, M. L. Tseng, E. Yadegaridehkordi, and N. Leung, "Applications of disruptive digital technologies in hotel industry: A systematic review," *Int. J. Hosp. Manag.*, vol. 107, p. 103304, Oct. 2022, doi: 10.1016/J.IJHM.2022.103304.
- [5] Asosiasi Penyelenggara Jasa Internet Indonesia, "Laporan Survei Internet APJII 2019 – 2020," *Asos. Penyelenggara Jasa Internet Indonesia*, 2020. <https://apjii.or.id/survei>.
- [6] W. A. Alkhowaiter, "Digital payment and banking adoption research in Gulf countries: A systematic literature review," *Int. J. Inf. Manage.*, vol. 53, p. 102102, Aug. 2020, doi: 10.1016/J.IJINFOMGT.2020.102102.
- [7] H. A. Alnemer, "Determinants of digital banking adoption in the Kingdom of Saudi Arabia: A technology acceptance model approach," *Digit. Bus.*, vol. 2, no. 2, p. 100037, Jan. 2022, doi: 10.1016/J.DIGBUS.2022.100037.
- [8] N. A. Windasari, N. Kusumawati, N. Larasati, and R. P. Amelia, "Digital-only banking experience: Insights from gen Y and gen Z," *J. Innov. Knowl.*, vol. 7, no. 2, p. 100170, Apr. 2022, doi: 10.1016/J.JIK.2022.100170.
- [9] F. Ernawan, A. Aminuddin, D. Nincarean, M. F. A. Razak, and A. Firdaus, "Three Layer Authentications with a Spiral Block Mapping to Prove Authenticity in Medical Images," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 4, pp. 211–223, 2022, doi: 10.14569/IJACSA.2022.0130425.
- [10] L. V. Gorodianska, T. Nosenko, and V. Vember, "Neobanks operations and security features," *2019 IEEE Int. Sci. Conf. Probl. Infocommunications Sci. Technol. PIC S T 2019 - Proc.*, pp. 839–842, Oct. 2019, doi: 10.1109/PICST47496.2019.9061268.
- [11] T. Sharma and D. Rattan, "Malicious application detection in android — A systematic literature review," *Comput. Sci. Rev.*, vol. 40, p. 100373, May 2021, doi: 10.1016/J.COSREV.2021.100373.
- [12] J. Pang and J. Bian, "Android malware detection based on naive bayes," *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 2019–October, pp. 483–486, 2019, doi: 10.1109/ICSESS47205.2019.9040796.
- [13] P. K. Tiwari, S. R. Basireddy, and T. Velayutham, "Identification of Possibly Intemperate Permission Demands in Android Apps," *Proc. 2nd Int. Conf. Innov. Pract. Technol. Manag. ICIPTM 2022*, pp. 101–106, 2022, doi: 10.1109/ICIPTM54933.2022.9753830.
- [14] A. Aminuddin, "Android Assets Protection Using RSA and AES Cryptography to Prevent App Piracy," *2020 3rd Int. Conf. Inf. Commun. Technol. ICOIACT 2020*, pp. 461–465, 2020, doi: 10.1109/ICOIACT50329.2020.9331988.
- [15] M. Rahardi, A. Aminuddin, F. F. Abdulloh, and R. A. Nugroho, "Sentiment Analysis of Covid-19 Vaccination using Support Vector Machine in Indonesia," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 6, p. 2022, 2022, doi: 10.14569/IJACSA.2022.0130665.
- [16] S. Zahoor and R. Rohilla, "Twitter Sentiment Analysis Using Machine Learning Algorithms: A Case Study," *Proc. - 2020 Int. Conf. Adv. Comput. Commun. Mater. ICACCM 2020*, pp. 194–199, 2020, doi: 10.1109/ICACCM50413.2020.9213011.
- [17] E. Y. Sari, A. D. Wierfi, and A. Setyanto, "Sentiment Analysis of Customer Satisfaction on Transportation Network Company Using Naive Bayes Classifier," *2019 Int. Conf. Comput. Eng. Network, Intell. Multimedia, CENIM 2019 - Proceeding*, vol. 2019–November, Nov. 2019, doi: 10.1109/CENIM48368.2019.8973262.
- [18] S. Ranjan and S. Mishra, "Comparative Sentiment Analysis of App Reviews," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, Jul. 2020, doi: 10.1109/ICCCNT49239.2020.9225348.
- [19] H. Mehyar, M. Saeed, H. B. A. Al-Ja'afreh, and R. Al-Adaileh, "The impact of electronic word of mouth on consumers purchasing intention," *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 2, pp. 183–193, 2020.
- [20] N. Colmekcioglu, R. Marvi, P. Foroudi, and F. Okumus, "Generation, susceptibility, and response regarding negativity: An in-depth analysis on negative online reviews," *J. Bus. Res.*, vol. 153, pp. 235–250, Dec. 2022, doi: 10.1016/J.JBUSRES.2022.08.033.
- [21] J. Qin and M. Zeng, "An integrated method for product ranking through online reviews based on evidential reasoning theory and stochastic dominance," *Inf. Sci. (Nij.)*, vol. 612, pp. 37–61, Oct. 2022, doi: 10.1016/J.INS.2022.08.070.
- [22] A. Yaqin, M. Rahardi, and F. F. Abdulloh, "Accuracy Enhancement of Prediction Method using SMOTE for Early Prediction Student's Graduation in XYZ University," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 6, pp. 418–424, 2022.
- [23] M. Rahardi, F. F. Abdulloh, and W. S. Putra, "A Blind Robust Image Watermarking on Selected DCT Coefficients for Copyright Protection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 719–726, 2022.
- [24] H. Hairani, M. Innuddin, and M. Rahardi, "Accuracy Enhancement of Correlated Naive Bayes Method by Using Correlation Feature Selection (CFS) for Health Data Classification," *2020 3rd Int. Conf. Inf. Commun. Technol. ICOIACT 2020*, pp. 51–55, 2020, doi: 10.1109/ICOIACT50329.2020.9332021.
- [25] P. P. Gokul, B. K. Akhil, and K. K. M. Shiva, "Sentence similarity detection in Malayalam language using cosine similarity," *RTEICT 2017 - 2nd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. Proc.*, vol. 2018–January, pp. 221–225, Jul. 2017, doi: 10.1109/RTEICT.2017.8256590.
- [26] F. F. Abdulloh, M. Rahardi, A. Aminuddin, S. D. Anggita, and A. Y. A. Nugraha, "Observation of Imbalance Tracer Study Data for Graduates Employability Prediction in Indonesia," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 8, pp. 169–174, 2022.
- [27] A. C. Emcha, Widyawan, and T. B. Adji, "Quotation extraction from Indonesian online news," *2019 Int. Conf. Inf. Commun. Technol. ICOIACT 2019*, pp. 408–412, 2019, doi: 10.1109/ICOIACT46704.2019.8938558.
- [28] Z. Mustaffa, M. H. Sulaiman, F. Ernawan, Y. Yusof, and M. F. M. Mohsin, "Dengue outbreak prediction: Hybrid meta-heuristic model," *Proc. - 2018 IEEE/ACIS 19th Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distributed Comput. SNPD 2018*, pp. 271–274, Aug. 2018, doi: 10.1109/SNPD.2018.8441095.
- [29] T. M. Ma, K. Yamamori, and A. Thida, "A Comparative Approach to Naïve Bayes Classifier and Support Vector Machine for Email Spam Classification," *2020 IEEE 9th Glob. Conf. Consum. Electron. GCCE 2020*, pp. 324–326, 2020, doi: 10.1109/GCCE50665.2020.9291921.
- [30] A. Aminuddin and F. Ernawan, "AuSR1: Authentication and self-recovery using a new image inpainting technique with LSB shifting in fragile image watermarking," *J. King Saud Univ. - Comput. Inf. Sci.*, Feb. 2022, doi: 10.1016/J.JKSUCI.2022.02.009.
- [31] A. Aminuddin and F. Ernawan, "AuSR2: Image watermarking technique for authentication and self-recovery with image texture preservation," *Comput. Electr. Eng.*, vol. 102, no. April, p. 108207, 2022, doi: 10.1016/J.COMPELECENG.2022.108207.
- [32] R. M. Amir Latif, M. Talha Abdullah, S. U. Aslam Shah, M. Farhan, F. Ijaz, and A. Karim, "Data scraping from google play store and visualization of its content for analytics," *2019 2nd Int. Conf. Comput. Math. Eng. Technol. iCoMET 2019*, pp. 1–8, 2019, doi: 10.1109/ICOMET.2019.8673523.
- [33] K. Kusnawi and A. Hendra Wijaya, "Sentiment Analysis of Pancasila Values in Social Media Life Using the Naive Bayes Algorithm," *Proc. - 2021 Int. Semin. Appl. Technol. Inf. Commun. IT Oppor. Creat. Digit. Innov. Commun. within Glob. Pandemic, iSemantic 2021*, pp. 96–101, Sep. 2021, doi: 10.1109/ISEMANTIC52711.2021.9573194.