Public Perception for the use of Digital Wallet in Indonesia using Social Network Analysis

Desanty Ridzky [1,*] Mohammad Isa Irawan [2]

[1] *Institut Teknologi Sepuluh Nopember Surabaya, Indonesia*
[2] *Institut Teknologi Sepuluh Nopember Surabaya, Indonesia*
Corresponding author : desanty.19092@mhs.its.ac.id, mii@its.ac.id

ABSTRACT

The development of financial technology is currently growing by payment transactions can be done easily using digital wallets, so customers do not need to use cash to make some payments. The competition of fintech service providers in Indonesia is getting more competitive, based on the varied services offered as the company's strategy to maintain customer satisfaction. The problem arises in how the public perceives the benefits of mobile payment service providers GoPay and OVO as the top digital wallet in the second quarter of 2019. Companies can use social media to see how the company works by public perception because the public uses Twitter to convey messages, comments, and complaints about the services provided by the company. This study aims to analyze the public views of the two companies by combining sentiment analysis using Naïve Bayes to identify positive and negative perceptions and social network analysis to know the interaction between the public. Sentiment analysis using TF-IDF feature selection and Naïve Bayes classification generate GoPay accuracy is 80.42%, and OVO is 82.44%. Positive sentiment towards the use of GoPay and OVO has a higher value. The network on SNA shows that nodes that are not connected to the gojekindonesia and ovo_id node have more positive perceptions, namely perceptions related to promos, cashback and discounts on the use of GoPay and OVO.

Keywords: perception, GoPay, OVO, Naïve Bayes, social network analysis

1. INTRODUCTION

E-Commerce Statistics Data in 2019 verify that online business development over the last five years has driven digital changes in the business world [1]. The online business impacts the payment system digitalization process using financial technology, especially in digital wallets. Indonesia has implemented payments using digital wallets to make the transaction more efficiently for customers. The competition between digital wallet companies is getting more challenging because the number of digital wallet development companies in Indonesia increases. Based on survey results by Snapcart in 2019, the most favourite wallets in Indonesia is OVO and GoPay In the second quarter of 2019, a survey by iPrice also proves that monthly active users using digital wallet are GoPay and OVO [2]. This survey shows that OVO and GoPay are the top two digital wallets interested in the public. The use of digital wallets is support by the development of the internet, which give new changes for society. Nowadays, internet development also encourages people to use social media to get information about services or products offered by online business companies. Social media use to communicate between companies and the public in conveying their perceptions about the service or products provided by the company [3]. This perception relates to the sentiments shared by the people on social media.

This study analyses sentiment analysis based on public perception using GoPay and OVO digital wallet use TF-IDF feature selection and Naïve Bayes. The TF-IDF method is a feature extraction option with the best performance in various classifications, one of which is Naïve Bayes[4]. The category used in conducting sentiment analysis in this study is Naïve Bayes with the Gaussian model. The Gaussian model is an easy classification model by utilizing numerical data whose calculations only count on mean and standard deviation [5], [6]. Positive sentiment expresses good perception, and negative sentiment contains a lousy perception of service or products. The classification data will be analyzed using the social network analysis method to understand the interaction and relation between objects in graph visualization. It will be studying the thing of the graph produces a knowledge network [7]. Knowledge that can be analyzed using the SNA method is to determine company position or rank of the company based on SNA network properties [8]. Companies need to know their position to consider developing and perfecting the strategy to target customer focus, purposed to maintain or improve their rankings influenced by public perception based on SNA properties [9]. Using SNA can also identify actors who have an essential role in a network that relate to other actors of GoPay and OVO [10]. The novelty of this study is to understand the public

perception through a combination of sentiment analysis and social network analysis, their role in influencing the use GoPay and OVO digital wallet on Twitter, and determine the company's position based on public interaction related to their perception about GoPay and OVO.

2. LITERATURE REVIEW

2.1 Digital Wallet

Payment services with digital wallets are gaining increasing popularity in the transition phase towards a future with innovations in technology. In this era, the digital wallet is helping payments to be more efficient with digital coupons, transaction receipts, where these receipts can be stored on the user's smartphone using a digital wallet [11]. A previous study was conducted to determine customers' factors and behavioural intentions using digital wallets namely. The most significant influence is the habit factor [12]. In this study, case studies are on GoPay and OVO. PT Dompet Karya Anak Bangsa is a GoPay digital wallet in Gojek balance electronic money that can be used for all transaction services on the Gojek application and outside the application. OVO is a digital wallet from PT Visionet International managed by Lippo Group in collaboration with online transportation, provider application namely Grab to make payments for service on the Grab application or outside the application can using OVO.

2.2 Sentiment Analysis

Sentimental is a feeling, attitude, or emotion-based on perception or opinion. Sentiment Analysis is mainly to know thoughts expressing or implying positive or negative sentiment on a particular word or sentence. Determine an individual's perception of something based on sentences shared in social media, usually using sentiment analysis. A previous study was conducted to determine customer perception of well-known clothing brands and then conclude that online users compare with other brands while choosing a particular brand of interest [13].

2.3 Gaussian Naïve Bayes

Naïve Bayes classifier is a classification with probability and statistical methods on machine learning that is most effective for predicting a positive or negative weighted sentiment. Classification using Naïve Bayes is often used for text categorization to combine word probabilities and categorize documents [14], [15]. This study using the Gaussian Naïve Bayes model that supports continuous-valued features, compute mean and standard deviation from training data. Classification is done in every tweet and calculates the probabilities in every feature present in the tweet for each class positive and negative label. The labelling is assigned based on the tweet that has the most significant probability score. The

equation for Gaussian Naïve Bayes used in sentiment analysis programming [6].

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right) \qquad (1)$$

2.4 Cross-validation

Cross-validation is a method commonly used to evaluate the predictive model performance. In evaluating k-fold cross-validation, the data divide into two parts, namely some parts for training data and predictive data tested in other regions. This study using the k-10 fold to get the best validation because k with the k-10 fold can give a low bias result, best variance and computational cost are not significant. Using k-10 fold is done based on the results of previous experiments that testing with a fold number of 10 is the best choice for accurate estimates [16].

2.5 Social Network Analysis

Social network analysis or SNA is knowledge to understand graphs representing the relationship between a node in a network and process the relationship to become the data easily understood in the graph form. The SNA network will form a node as an actor or other object, and also the relationship formed are called edges. In this type of business analysis, social network analysis is used as a communication flow within the organization and can improve the relationship between a company and their customers.

Social network analysis has network properties that can be used to determine and analyze the network model by scores or weights. SNA network properties are size, modularity, diameter, average degree and average path length. The size of network size indicates the number of nodes in a network representing the number of users. The number of an edge represents the relationship that occurs in a social network. Modularity measures how different a group is in a network and how many groups are forming and how different the groups are. The diameter shows the maximum distance between one node and another connected node on a network. The average degree is an essential network property in the social network because the average degree shows the number of edges connecting a node to other nodes. The average path length is the average distance from all the shortest paths in the network whose score or weight can determine and analyze the network.

A network has a centrality by network connection pattern calculation that can be measured to illustrate the importance node and influence of nodes on the network. Centrality in SNA is as follows [17]:

a. Betweenness centrality, calculated by analyzing the shortest path between nodes and summing the number of these paths that pass through each particular node

b. Closeness centrality is a measure of the average distance from a node to another node on the network

c. Degree centrality is metric to measure the number of interconnected nodes, or degree centrality is the score of the bond to a particular node.

d. Eigenvector centrality measures the importance of a node in a graph from its adjacent or neighbouring significant node. If a node connects to a crucial node, then the eigenvector centrality score will be higher than a node connected to an unimportant node. Based on the number of connections and formed, metric eigenvector centrality measures a node's influence in the network [18].

The previous study has used SNA to visualize information on actors or users who have an influential position in a community by the centrality score and determine how connected the network structure of these actors is in the network graph [10].

3. METHODOLOGY

This stage is crawling data collected data on Twitter related to the object of this study. The crawled datasets are replies, retweets, and mentions from users contain the keyword "GoPay" and OVO. The data collection period starts from February 1, 2019 to February 30, 2021. Data retrieved through Twitter using Twitter API with python 3.9.2. Data is filtered again by deleting tweets not relevant to this study, such as spam tweets. After deleting useless tweets, the number of datasets for further analysis GoPay is 1,892 tweets, and OVO is 1,883 tweets.

3.1 Processing Data

Data preprocessing was eliminating irrelevant data in this study. The steps involved in data processing are as follows :

- Cleansing, the process to remove punctuation, hashtags and others unnecessary in the analysis process

- Data labelling, data class labelling will be carried out into two classes, namely positive and negative

- Case folding, a process to change all uppercase letters to lowercase

- The tokenizing process to separate or cut sentences into several parts

- Stopword Removal, removing the words that have no meaning and are irrelevant in the study

- Stemming, mapping and parsing the form of a word into the basic word form

3.2 Feature Selection TF-IDF

Term frequency-inverse document frequency I a selection feature used in this study. TF-IDF weighting is done by assigning a weight to the relationship of a word or term to the document. TF-IDF calculation to find out how important a word is in the document or the group of words. The feature extraction obtained in this study uses the top 100 terms to optimize the results. In calculating the document frequency, the term frequency-inverse document frequency will be calculating the top 100 terms. The top 100 term is the term that affects the context of the sentence in a document. The following formula 2 is the calculation of the weight in TF-IDF.

$$TFIDF = TF \times IDF \qquad (2)$$

TF or Term frequency is the frequency of occurrence of term and inverse document frequency or IDF.

3.3 Classification

The dataset used for classification is training data as input to the testing process and testing data for classification. This study's testing data were divided into k-10 fold based on the cross-validation method for evaluating the classification performance. Cross-validation with the number of folds is ten will calculate the testing data for the amount of data with a specific classification divided by the number of the fold at the beginning. In this study, 851 training data with a positive label, and 852 negative labels, so the testing data will be processed is 95 for positive class. The negative class is 94 on GoPay digital wallet. In the OVO digital wallet, the positive class training data is 847, the testing data with the positive label is 94, and the negative class training data is 948. Based on the training data, the number of testing data is 94. Next, calculate the probability score using Gaussian Naïve Bayes and the selection features obtained. Based on the probability score, if the positive probability score is higher than negative, then the test data is classified as positive and vice versa.

3.4 Classification Evaluation

The next stage is accuracy testing. Accuracy testing is using the cross-validation method with the k-10 fold. Iterations in this testing process were done ten times using a combination of 10 parts of the variation of training data and GoPay and OVO test data. There is something false statement in the classification process. For example, a tweet gets a negative label in classification even though the tweet has a positive sentence is called a false negative. If the class of tweets in classification is positive and the label is positive, it is

called a true positive. The number of false statements is summarized in the confusion matrix. The following Table 1 is an example of a confusion matrix on GoPay

Table 1 Example confusion matrix on Gopay

		Predicted: Positive	Predicted: Negative
Actual Positive	:	80	14
Actual Negative	:	23	72

Based on the number of summarized false statements in Table 1, this study calculates the classification accuracy using formula 3

$$\text{Accuracy: } \frac{TP+TN}{TP+TN+FN+FP} \quad (3)$$

TP defined true positive that condition is predictive positive and actual is positive. TN is a true negative, with the result predicted is negative and actual negative. False positive or FP is a condition when indicated positive and actual is negative. FN defined false negative, the situation when predicted is negative and actual is positive. Using the formula to calculate accuracy perception using GoPay and OVO, the results obtained are as Table 2.

Table 2 Accuracy score GoPay and OVO

K fold	Digital Wallet	
	Gopay	OVO
1	72.10 %	78.30%
2	75.78%	76.19%
3	73.54%	77.24%
4	75.13%	71.27%
5	**80.42%**	78.19%
6	78.30%	68.61%
7	75.66%	75.53%
8	75.66%	80.31%
9	78.30%	75%
10	76.19%	**82.44%**

Table 2 shows that GoPay has the highest accuracy in fold 5, 80.42%, and OVO in fold 10 is 82.44%. The testing data from the highest accuracy score in GoPay and OVO is used for the following analysis. Based on GoPay and OVO's highest accuracy results, the percentage with positive and negative classifications for the use of GoPay and OVO can be seen in the following **Figure 1** compares

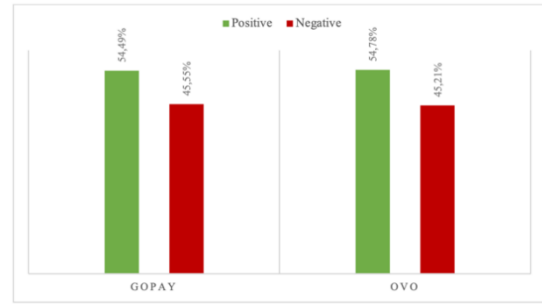the rate of positive and negative classifications using GoPay and OVO

Figure 1 Percentage of positive and negative classifications of GoPay and OVO

Based on Figure 1, the results of the sentiment analysis of the classification of public tweets on the use of the two digital wallets got a higher value in the positive classification than the negative classification value.

4. RESULT AND DISCUSSION

4.1 Social Network Analysis on GoPay

An analysis of the interaction graph between society on Twitter will use social network analysis at this stage. The graph analysis of interactions between society on Twitter takes from the testing data of 189 tweets classified in the previous step. The graph forms nodes, which are Twitter accounts. The connection between nodes edges that form of interaction containing classified positive and negative perceptions from people using GoPay and OVO. The following Figure 2 is a visualization of the graph of social interaction using GoPay
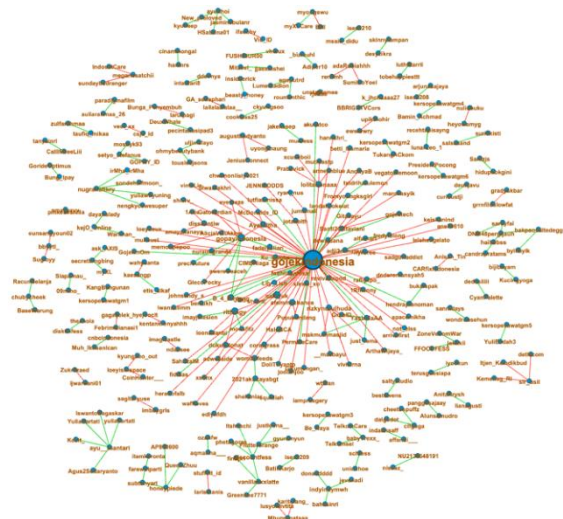

Figure 2 Social network analysis using GoPay

The graph is formed by nodes that are Twitter users and edges that connect each node. Green edges are people who talk about using GoPay with positive classification and red edges with negative classification. The number

of edges is 245, and nodes are 325. Positive classification GoPay is widely discussed by one node with another individual and has no relationship with the gojekindonesia node. Positive classification using Gopay through direct interaction between the society is related to the discount promotion offered, cashback, topping up GoPay balances, and recommending each other to make purchases and payments with GoPay. GoPay graph shows that the gojekindonesia node has a larger size than the other nodes because the gojekindonesia node has high connectivity with other nodes. The gojekindonesia node will be analyzed further regarding the perceptions discussed by the public about using GoPay with the gojekindonesia node. The following Figure 3 is a node that is connected to the gojekindonesia node.
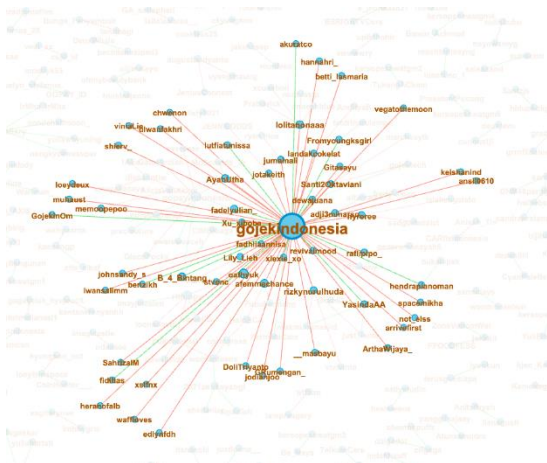

Figure 3 Graph of nodes connected to node GoPay

The graph shows that the use of GoPay has many connections with the use of the Gojek application. It is seen from the interactions carried out on the gojekindonesia node. Discussions about the GoPay graph, the gojekindonesia node has more negative classification than positive classification. The negative classification tweet using GoPay related to the node, which is dominantly associated with bad perceptions, contains complaints about GoPay balances being cut off; however, the drivers do not pick up passengers, GoPay transactions are often unsuccessful, but the balance is also deducted, filling GoPay balance not successful and the balance does not enter the gojek account, GoPay balance that has been removed but not returned, fraud in the name of gojek, and GoPay transactions out of user control. In addition to knowing the perception of the gojekindonesia node, social network analysis can also be used for companies to identify nodes that have an important role based on the highest centrality score on the GoPay graph. The following Table 3 is the node with the highest centrality calculation result from the GoPay graph.

Table 3 Centrality calculation on GoPay graph

	Node	Total
Degree Centrality	gojekindonesia	57
Betweenness Centrality	gojekindonesia	5434.0
Closeness Centrality	Iamprasgery, wtftian, pinkeuvanilla, amaliyayy, salsaaand	1.0
Eigenvector centrality	gojekindonesia	1.0

Based on the score of degree, betweenness, and eigenvector centrality, the gojekindonesia node is the ideal node to influence the public towards its statement because gojekindonesia is the central node graph with the highest centrality. Companies can use this gojekindonesia account to promote their products because gojekindonesia is a node that influences the graph of GoPay usage. Many nodes own the closeness centrality score but only mention five nodes with the highest centrality in the table. These nodes form their group and communicate directly and get a high score of closeness centrality. Node with the highest closeness centrality has fast access to other nodes because it has the shortest route and closes to access other nodes. A company can use the node account to disseminate information or promote through this node also spread more quickly.

4.2 Social Network Analysis on OVO

The next stage is the analysis OVO graph. The centrality score obtained from the OVO graph is based on social interaction on Twitter from 188 testing data that has been classified. The following Figure 4 is a graph of the interaction between users on the use of OVO.
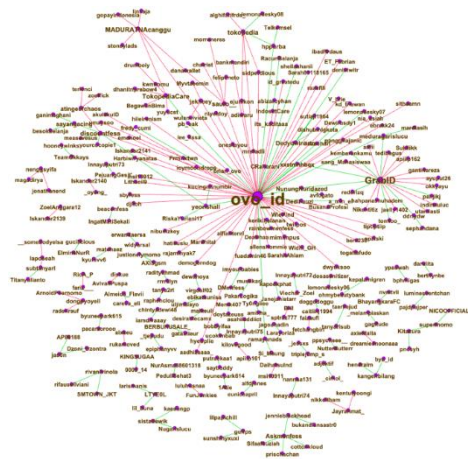

Figure 4 Social network analysis using OVO

A node is a Twitter user and edges that connect each node based on their perception in the graph. As with GoPay, the green edges on the OVO graph are people who talk about using OVO with positive classification

and red edges with negative classification. The number of edges formed on the OVO graph is 227, and the node is 287. The graph shows that the ovo_id node has many relationships with other nodes. Meanwhile, many nodes are not connected to the ovo_id node having positive classification perceptions because each node is directly related to discuss their perceptions regarding the use of OVO. The positive classification of discussions on using OVO between nodes is about promotion, discount, cashback, payments using OVO points and balance transfer using OVO.

GrabID and Tokopedia nodes also have connectivity with other nodes, where the edges of several nodes connect these edges to the ovo_id node. This connection is because people use OVO on their Grab and Tokopedia accounts, and the discussion regarding the use of OVO has a relationship with the GrabID and Tokopedia nodes. Based on the graph formed, this study analyzed the perceptions discussed by the public about using OVO with the ovo_id node. The following Figure 5 is a node connected to the ovo_id node.
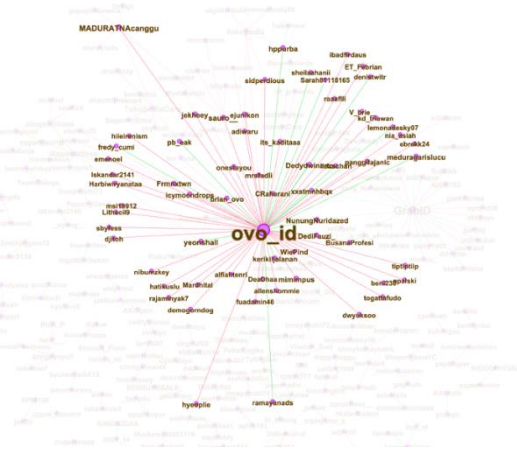

Figure 5 Graph of nodes connected to node ovo_id

The graph shows that conversations about OVO use related to ovo_id have more weighted edges with negative classifications than positive classification. The negative classification of a tweet using OVO related to the ovo_id node is dominantly related to complaints about interference with the error system in the OVO application. The other is topping up OVO balance not entered, OVO premiere account upgrades which are very difficult and not available in online service, OVO service promotion cannot use, and slow service in responding to customer problems. The next stage is to identify the highest centrality in the OVO graph to determine the effect of a node. The following Table 4 is the node with the highest centrality score calculation result from the OVO graph.

Table 4 Centrality calculation on GoPay graph

	Node	Total
Degree Centrality	ovo_id	63
Betweenness Centrality	ovo_id	8530.667
Closeness Centrality	Luluhusnaa, FunJunkies, mustikari, PakarLogika, erwandaersa	1.0
Eigenvector centrality	ovo_id	1.0

After getting the node with the highest centrality score, the ovo_id account is the OVO digital wallet service provider company's account. Based on the highest centrality score obtained, the ovo_id account can also be useful for companies to spread company service information or product promotions because the ovo_id is a node that affects the OVO graph. The highest closeness centrality score in the OVO graph is the same as GoPay because many nodes own it. These nodes form their close group and communicate directly. These nodes have fast access to other connected nodes. Based on the proximity and the shortest path, if the company want to spread the information, it will be spread more quickly.

4.3 Company Position

At this stage, the company's position or ranking will be identified based on society conversation about using GoPay and OVO by comparing the network property score in social network analysis. The following Table 5 is the score of GoPay and OVO network properties

Table 5 Network properties score of GoPay and OVO

Network Property		Network Property Score	
		GoPay	**OVO**
Size	Nodes	325	287
	Edges	245	227
Modularity		0.874	0.816
Diameter		10	8
Average Degree		1.508	1.582
Average Path Length		3.318	3.489

Table 5 shows that the first network property is size, and size explains that a good and active network has a larger size. GoPay has a higher score of size. Next, the calculation of the network properties of the modularity is based on the number of groups formed. If many groups are formed, they will become different communities, and

this makes the groups need more specifications on the GoPay and OVO networks in each community. Therefore, a smaller value is a better modularity value, namely OVO. The following network property is a diameter. The smaller the diameter formed, the easier the nodes to communicate with each other due to the shorter distance. The OVO digital wallet has a smaller diameter compared GoPay digital wallet. The fourth network property is the average degree. If more edges connect between nodes, it will be faster and easier the disseminate information. The digital wallet that has the highest average degree is OVO, with an average degree of 1,582. The following network property is the average path length. The fewer accounts that are crossed, the network has a strong connection. Therefore GoPay has the lowest average score of 3,318.

5. CONCLUSION

Based on the discussion related to the use of social network analysis methods in this study, the conclusion of this study is :

1. Sentiment analysis using TF-IDF feature extraction and the Naïve Bayes method of public perception of the use of GoPay and OVO through validation of results ten times, the highest accuracy gain of 80.42% on GoPay and OVO of 82.44%. Based on the classification that has been done, the percentage value of positive responses on the two digital wallets is higher than the classification of negative perceptions, namely the positive percentage of 54.49% on GoPay and OVO of 54.78%. In contrast, the negative perception of GoPay is 45.55% and 45.21% on OVO.

2. Mapping of interactions on the network using the social network analysis method produces nodes that are not connected to the nodes of gojekindonesia and ovo_id dominant containing positive perceptions that can be maintained by GoPay and OVO service providers regarding promos and cashback offered by GoPay and OVO digital wallets. The nodes connected to the gojekindonesia node and ovo_id, which are the official accounts of the digital wallet provider, show a perception with dominant negative classification, namely complaints against the use of the GoPay and OVO applications.

3. Based on the closeness centrality score obtained, GoPay companies can use Iamprasgery, wtftian, pinkeuvanilla, amaliyayy, salsaaand nodes. OVO companies can use Luluhusnaa, FunJunkies, mustikari, PakarLogika, erwandaersa nodes to promote products and spread information related

to company services because the nodes have fast access that is connected to other nodes.

4. Based on the comparison score of community interaction network properties on Twitter social media using SNA consisting of size, diameter, modularity, average degree, and average path length, it is concluded that OVO digital wallet has a superior position. It is because the OVO digital wallet is supported by banking financial institutions so that it is more secure for the public in conducting digital transactions.

AUTHORS' CONTRIBUTIONS

This study is arranged and designed by All authors. DR conducted the experiments, analyzed the data and wrote the paper. The final version of this manuscript is approved by All authors.

ACKNOWLEDGMENTS

REFERENCES

[1] Badan Pusat Statistik, *Statistik E-Commerce 2019*. 2019. [Online]. Available: https://www.bps.go.id/publication/2019/12/18/fd1e96b05342e479a83917c6/statistik-e-commerce-2019.html

[2] Vivin Dian Devita, "Siapa Aplikasi E-wallet dengan Pengguna Terbanyak di Indonesia?" https://iprice.co.id/trend/insights/e-wallet-terbaik-di-indonesia/

[3] M. Yadav and Z. Rahman, "Measuring consumer perception of social media marketing activities in e-commerce industry: Scale development & validation," *Telemat. Inform.*, vol. 34, no. 7, pp. 1294–1307, Nov. 2017, DOI: 10.1016/j.tele.2017.06.001.

[4] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The Impact of Features Extraction on the Sentiment Analysis," *Procedia Comput. Sci.*, vol. 152, pp. 341–348, Jan. 2019, DOI: 10.1016/j.procs.2019.05.008.

[5] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, Jan. 2019, DOI: 10.1016/j.procs.2019.11.181.

[6] A. Rane and A. Kumar, "Sentiment Classification System of Twitter Data for US Airline Service Analysis," in *2018 IEEE 42nd Annual Computer Software and Applications Conference*

(COMPSAC), Jul. 2018, vol. 01, pp. 769–773. DOI: 10.1109/COMPSAC.2018.00114.

[7] B. Ufuk, U. Bakan, and T. Kan, "Understanding the Usage Characteristics of Twitter in the UK Universities: A Social Network Analysis (SNA) Approach," *11th LUMEN Int. Sci. Conf. Commun. Action Transdiscipl. Ethical Soc.*, 2018, Accessed: Feb. 22, 2021. [Online]. Available: https://proceedings.lumenpublishing.com/ojs/index.php/lumenproceedings/article/view/108

[8] M. F. Firdaus, Z. K. A. Baizal, M. K. Bratawisnu, and H. A. Gusman, "Brand Awareness Using Network Modeling Method," in *2020 8th International Conference on Information and Communication Technology (ICoICT)*, Jun. 2020, pp. 1–5. DOI: 10.1109/ICoICT49345.2020.9166245.

[9] L. Steinhoff, D. Arli, S. Weaven, and I. Kozlenkova, "Online relationship marketing," *J. Acad. Mark. Sci.*, vol. 47, pp. 369–393, May 2019, DOI: 10.1007/s11747-018-0621-6.

[10] M. Grandjean, "A social network analysis of Twitter: Mapping the digital humanities community," *Cogent Arts Humanit.*, vol. 3, p. 1171458, Apr. 2016, DOI: 10.1080/23311983.2016.1171458.

[11] Z. Bezhovski, "The Future of the Mobile Payment as Electronic Payment System," *Eur. J. Bus. Manag.*, vol. 8, no. 8, pp. 127–132, 2016.

[12] M. Widodo, M. I. Irawan, and R. A. Sukmono, "Extending UTAUT2 to Explore Digital Wallet Adoption in Indonesia," in *2019 International Conference on Information and Communications Technology (ICOIACT)*, Jul. 2019, pp. 878–883. DOI: 10.1109/ICOIACT46704.2019.8938415.

[13] A. Rasool and A. Kamyab, "(PDF) Twitter Sentiment Analysis: A Case Study for Apparel Brands," *J. Phys. Conf. Ser. Artif. Intell. Netw. Inf. Technol.*, Mar. 2019, DOI: 10.1088/1742-6596/1176/2/022015.

[14] R. Jagdale, V. Shirsat, and S. Deshmukh, "Sentiment Analysis on Product Reviews Using Machine Learning Techniques," 2019, pp. 639–647. DOI: 10.1007/978-981-13-0617-4_61.

[15] R. Ireland and A. Liu, "Application of data analytics for product design: Sentiment analysis of online product reviews," *CIRP J. Manuf. Sci. Technol.*, vol. 23, pp. 128–144, Nov. 2018, DOI: 10.1016/j.cirpj.2018.06.003.

[16] Max Kuhn and Kjell Johnson, *Applied Predictive Modeling*. 2013. Accessed: Jun. 17, 2021. [Online]. Available: https://www.academia.edu/41760919/Applied_predictive_modeling_max_kuhn_kjell_johnson

[17] L. Ronoh, K. Rabah, and M. D, "Using SNA Centrality Metrics to Detect Suspicious Social Media Users to Aid Law Enforcement Agencies in Kenya," *Min. Forensic Evid. Law Enforc. Soc. Media Platf.*, vol. 2, pp. 1–19, Jan. 2019.

[18] A. Bihari and M. Pandia, "Eigenvector centrality and its application in research professionals' relationship network," presented at the 2015 1st International Conference on Futuristic Trends in Computational Analysis and Knowledge Management, ABLAZE 2015, Feb. 2015. DOI: 10.1109/ABLAZE.2015.7154915.