

Mining Customers Opinion on Services and Applications of Mobile Payment Companies in Indonesia Using Sentiment Analysis Approach

Nadhila Idzni Prabaningtyas
Department of Industrial Engineering
Faculty of Engineering, Universitas
Indonesia
Depok, Indonesia
nadhila.idzni@ui.ac.id

Isti Surjandari
Department of Industrial Engineering
Faculty of Engineering, Universitas
Indonesia
Depok, Indonesia
isti@ie.ui.ac.id

Enrico Laoh
Department of Industrial Engineering
Faculty of Engineering, Universitas
Indonesia
Depok, Indonesia
enricolaoh@ui.ac.id

Abstract—The development of technology and digital has also increased the ease of accessing the internet. One aspect of daily life that are affected by the adoption of technology and the internet is the field of payment transactions. Payment transactions are inseparable from everyday life. At this time with the development of technology, payment transactions can be done with the more practical, easy, safe and convenient. The technology is called Financial Technology. Mobile payment is a service that is part of financial technology. The aspects contained in the mobile payment are top up, transfers, cash withdrawals, online payment, and offline payments. Classifications of reviews from Twitter are classified using Support Vector Machine. The results of this study are Go-Pay and OVO must pay attention to every aspect and improve every aspect, of course, to increase customer satisfaction. The accuracy level of the classification model produced for bigram is 92% (Go-Pay) and 93% (OVO). It also shows that sentiment analysis using bigram can improve accuracy level.

Keywords—Text Mining, Sentiment Analysis, N-Grams, Mobile Payment, Support Vector Machine (SVM)

I. INTRODUCTION

Digital and technology are currently developing rapidly and influencing aspects of life today. The development of technology and digital has also increased the ease of accessing the internet. Now the internet seems to be the primary need of every person in carrying out their activities, as well as in Indonesia. The number of internet users in Indonesia in 2017 is 54.68% of the total population of Indonesia. As many as 65.98% of the population whose internet usage duration is every day [1]. This shows that the population in Indonesia is an internet user who uses it for daily activities.

One aspect of daily life that are affected by the adoption of technology and the internet is the field of payment transactions. Payment transactions are inseparable from everyday life. At this time with the development of technology, payment transactions can be done with the more practical, easy, safe and convenient. The technology is called Financial Technology or also commonly called FinTech. Financial Technology is an innovation that combines financial services and technology that aims to attract customers with more user-friendly and efficient products and services. Financial technology companies in Indonesia have also developed. There are several categories of Fintech according to Bank Indonesia Regulation No. 18/40 / PBI / 2016. Companies with the “payment” category

are the most in Indonesia at 39%. Therefore, this research is focused on companies with payment categories in Indonesia.

Mobile payment is one part of the type of financial technology (fintech) that included in the payment category. The mobile payment service is a transaction conducted online using cellular devices such as mobile phones or smartphones that can be safely carried out for various payment sizes, transaction locations and transaction times [2]. Mobile payment services can be used in daily life as well as business activities such as bill payments, data, and credit package purchases, virtual transfers, restaurant payments, online and public transportation payments and others. In addition, mobile payment provides a lot of convenience in transactions, can reduce costs for transactions and increase the security of electronic payments [3]. Of the several mobile payment companies located in Indonesia, Go-Pay and OVO are the latest mobile payment companies which appeared in 2016 and 2017 [4]. Both are server-based mobile payment service companies, which the medium is hardware or mobile device connected to the internet and the service offered are top-up, transfer (remittances), withdrawals, online and offline payments.

There is competition between these two companies because the services they offer are similar and usually in the same location. Therefore, each company needs to maintain the quality of services and applications offered to maintain customer satisfaction. Customer satisfaction can be seen or assessed after using the service.

The development of technology and digital increases the explosive growth of social media, each individual and organization can issue opinions that can influence public sentiments and emotions which have an impact on the social and political system. Therefore, seeing customer reviews from social media is one of technique to see and assess customer satisfaction. From social media can get a good understanding of opinions, complaints, and question from customers [5]. Twitter is one of the Top 5 social media that is a forum for expressing opinions easily [6]. Twitter users in Indonesia in 2016 amounted to 24.3 million which made Indonesia the third highest Twitter user country. This services and applications of mobile payment review are analyzed to knowing whether mobile payment companies in Indonesia have managed to meet the expectations of mobile payment users. Reviews from Twitter are certainly diverse and unstructured. The right method for extracting useful information from non-measured data is Data Mining.

Text Mining is part of Data Mining. Text Mining is a method for getting meaningful information from a large number of review text. This method is also called intelligent text analytics which involves several techniques ranging from data retrieval or extracting data, natural language processing (NLP), classification methods that are machine learning, and others [7].

Sentiment Analysis or Opinion Mining is a field related to text mining which has the same goal to process a large volume of user reviews from the platform or social media using computational techniques [8]. Sentiment analysis can analyze people's opinions, evaluations, emotions, and sentiments towards things that are objects such as products and services and their attributes. Therefore, there are many industries that utilize content such as reviews, forum discussions, comments, and posts contained on social media to conduct sentiment analysis and know the public's opinion to support decision-making in industrial activities. Sentiment analysis has been used in every field from a product, health services, financial services, political elections and social events [9].

Previous research has been carried out with a sentiment analysis approach to reviews in the field of mobile phones [10], reviews related to the food price crisis in India [11], and reviews on online sports forums [12]. Most in these researches using the method of sentiment analysis for unigram or one word. There was research that proposed efficient algorithms show that approaches to extraction and use of bigrams can significantly improve analytical performance [13]. According to [14], models that use word pairs show higher accuracy results from models using only one word. Then research related to sentiment analysis with initial word pairs or bigrams is done in various fields and shows more accurate results than analysis using one word.

In the research comparing four methods which are Naïve Bayes, Maximum Entropy, Stochastic Gradient Descent (SGD), and Support Vector Machine (SVM) in the review sentiment classification showed that for unigram (one word) and bigram (two-word pair) produced better accuracy than for the classification of four grams and five grams [15]. In one research showed that the use of n-grams can increase the level of accuracy of classification, classification using Support Vector Machine with bigrams (two-words pair) resulted in an accuracy rate of 87% whereas with unigrams only produced 81.1% [16].

The main contribution of the paper can be stated as doing aspect-based sentiment analysis with word pairs or bigrams of services and applications of mobile payment reviews. Then group it based on aspects related to mobile payment such as top-up, transfer, cash withdrawals, online & offline payments [4]. The evaluation that will be processed comes from social media platforms, namely Twitter written in English.

II. RESEARCH METHODOLOGY

The structure of this research is data collection, data processing that divided into two steps which are pre-processing data and generate n-gram model. And the last is the model evaluation. The brief method is shown in Fig. 1.

A. Data Collection

During this stage, collected all the information needed to perform an analysis according to the research objectives. Data on Twitter extracted using the Twitter API named as "tweepy". There are several steps that must be taken in retrieving this data, including through dev.twitter.com which will create a new app, at this stage the researcher must have a twitter account first. Then it will get a security id, secret number, access token id, and secret token number that must be input into the auth file which is a pycharm file. Then put the filter function into the streaming file according to what data needed to process in this research. The auth file and streaming file will be run using python. The results of extracting data from Twitter will be in the new file.

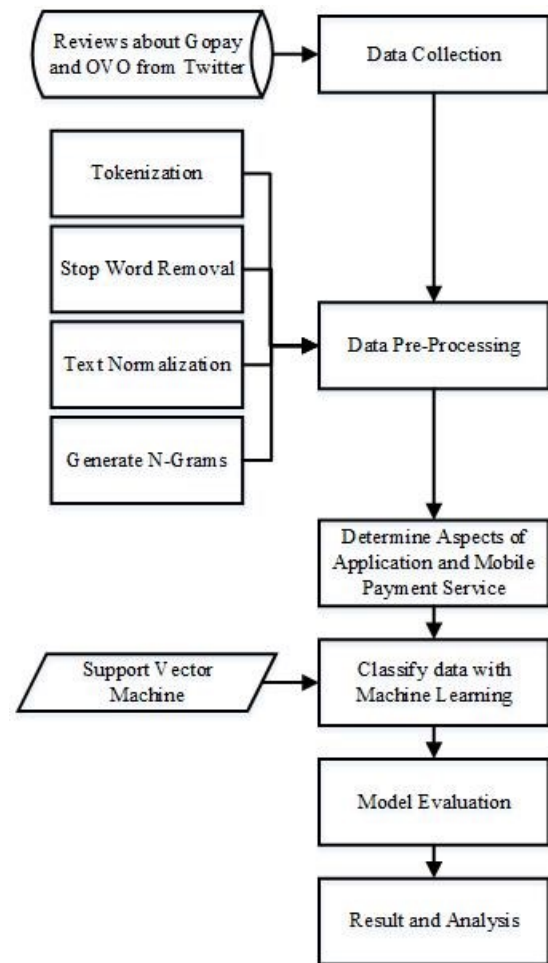


Fig. 1. Structure of Methodology

The processed data is data in 2019, searched based on the keywords "Go-Pay" and "OVO" which are two mobile payment companies in Indonesia that are the object of this research.

B. Data Pre-Processing

After collecting the data, the reviews must go through the stages of processing reviews, which include tokenization, stop words removal, text normalization. And extract features so that reviews can be given to the classification algorithm [17].

- **Tokenization**

Tokenization is a process that divides a given text into smaller and more meaningful elements. In this research, we tokenize the sentence into words (tokens).

- **Stop Words and Punctuation Marks Removal**

The process of eliminating the words in the text which has no useful meaning or often called a stop word and signs such as hashtag (#), URL, certain punctuation marks (emojicons), and others.

The words included in the stop words are "a", "or", "am", "I", "as", and others. The word "not" in this research was omitted from the stop word list because it can affect and have a significant impact on classification.

- **Text Normalization**

The process of arranging all the same characters into a form that is small or big letters. Then, also equate the word form by correcting abbreviated words or misspellings so that new tokens are not formed.

C. N-Gram Model

N-Grams are a contiguous sequence of n items from an extract text or document and in this occurrence, the n-grams refer to words. The n-gram model is used for sentiment analysis to help analyze the sentiment of a text or document. N indicates the number of words, unigram or $n = 1$, the number of words is one, then bigram or $n = 2$, the number of words is two.

Example: "top up Go-Pay via Mandiri online was successful, but the balance did not increase"

TABLE I. EXAMPLE OF UNIGRAM AND BIGRAM

Unigram	top, up, Go-Pay, via, mandiri, online, was, successful, but, the, balance, did, not, increase
Bigram	top up, up Go-Pay, Go-Pay via, via mandiri, mandiri online, online was, was successful, successful but, but the, the balance, balance did, did not, not increase

D. Support Vector Machine (SVM)

Support Vector Machine or SVM is one technique for classifying a document / text. SVM has been proven to generally outperform Naive Bayes in the effectiveness of text categorization [18]. This technique attempts to find the best classifier or hyperplane to separate two types of objects. The best hyperplane is in the middle between two sets of objects from two classes which are in this research the two classes are positive and negative classes [19].

To assess the performance of classification, it can be seen from the evaluation criteria which consist of accuracy, precision, recall and f-measure. The resulting accuracy level shows how many words are correctly classified. The higher the level of accuracy can indicate that the model is more accurate in classifying positive and negative reviews. Then the precision is the accuracy between the data requested and

the data provided by the system. Then recall is the success rate of the system in rediscovering data or information.

In this research, the classification was carried out for two classes, positive and negative. The resulting precision value indicates the number of true positive labeled sentences compared to the number of positively suspected sentences. Then the resulting recall value shows the number of positively labeled sentences that are correctly predicted compared to the number of positive samples. And produced a f-measure value that can represent the overall accuracy of the calculation of recall and precision.

III. RESULT AND ANALYSIS

Test data needed in this process. This process is to test the accuracy of the classification model that classify positive and negative reviews of mobile payment service and application. Test data contains reviews with positive or negative labels independently determined.

A. Classification Result of Go-Pay Dataset

Fig. 2. Shows the sentiment graph from the review on Twitter. It can be seen in each aspect that more tweets are negative than positive. This shows that Go-Pay users are not satisfied with the quality of services provided by the Go-Pay company, there are still obstacles when using the service. In addition, there may also be a tendency for someone to express their opinion on the company's Twitter and mention to be notified because of a complaint and looking for a solution.

All aspects contained in the graph must take a step in improving the quality of its services to increase customer satisfaction as well. But the aspect that has the most negative tweets is the "Top Up" aspect and then the "Cash Withdrawals" aspect.

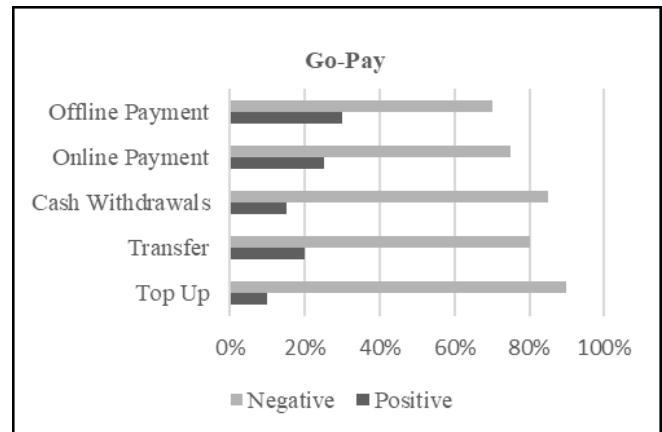


Fig. 2. Sentiment Graph of Go-Pay Reviews

Table 2 shows the evaluation parameters of the classification of Go-Pay reviews from Twitter. The accuracy level of unigram is 89%, while for bigram is 92%. It shows sentiment analysis with Bigram could increase the level of accuracy of classification, according to previous research. Precision value for unigram is 78% and for bigram is 83%. The classification model is quite good in determining the number of positively labeled sentences that are correct with the value of precision.

Recall value generated from this classification model is 74% for unigram and 76% for bigram. The success rate of the system in rediscovering data or information in this study, which is recall value is directly proportional to the value of precision produced. Then, the F-Measure, which is an average of precision and recall, that obtained of Go-Pay reviews for unigram is 76% and for bigram is 79%.

The result of classification Go-Pay dataset is the accuracy level for bigram which is 89%, is higher than the accuracy level of unigram. And review that comes from Twitter show every aspect of the service of Go-Pay require quality improvements to enhance customer satisfaction as well.

TABLE II. EVALUATION PARAMETERS OF GO-PAY DATASET

	Evaluation Parameter		Accuracy
Unigram	Precision	78%	89%
	Recall	74%	
	F-Measure	76%	
Bigram	Precision	83%	92%
	Recall	76%	
	F-Measure	79%	

B. Classification Result of OVO Dataset

Fig. 3. Shows the sentiment graph from the OVO review on Twitter. From that graphic, every aspect has more negative tweets than positive tweets. This shows that OVO users are not satisfied with the quality of services provided by them. When using OVO services, there are still problems and other difficulties. OVOs must think about the improvements they need to make.

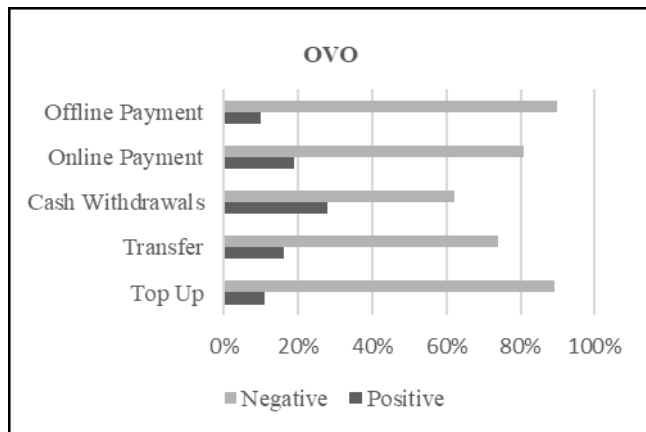


Fig. 3. Sentiment Graph of OVO Reviews

The evaluation parameters of classification OVO reviews from Twitter is shown in Table 3. The accuracy level of unigram is 89%, while for bigram is 93%. It is the same as the results of the Go-Pay reviews classification and previous research, where this study resulted in a higher level of accuracy when sentiment analysis uses bigram (word pairs). Precision value for unigram is 79% and for bigram is 86% which indicates that the classification model is good enough to determine the number of the correct positive

sentence. Recall value generated from this classification model is 76% for unigram and 74% for bigram. And the F-Measure, which is an average of precision and recall for unigram is 77% and for bigram is 80%. Overall the classification model produces good results and in accordance with previous research.

TABLE III. EVALUATION PARAMETERS OF OVO DATASET

	Evaluation Parameter		Accuracy
Unigram	Precision	79%	89%
	Recall	76%	
	F-Measure	77%	
Bigram	Precision	86%	93%
	Recall	74%	
	F-Measure	80%	

IV. CONCLUSION

This research focused on doing aspect-based sentiment analysis with word pairs or bigrams of services and applications of mobile payment reviews. Also group it based on aspects related to mobile payment such as top-up, transfer, cash withdrawals, online & offline payment. The data used are Go-Pay and OVO reviews that are collected from Twitter. Go-Pay and OVO are the latest mobile payment companies which appeared in 2016 and 2017. Both are server-based mobile payment service companies, which the medium is a hardware or mobile device connected to the internet. All data processing starts from pre-processing to the classification is done using python software.

The results of classification Go-Pay and OVO reviews are similar. The accuracy level of bigram in both datasets are higher than the unigram. This research indicates that classification uses Support Vector Machine (SVM) with a n-grams approach consisting of two words (bigram) resulting in a higher level of accuracy than one word (unigram).

The review taken from Twitter shows that every aspect of the services of the Go-Pay and OVO companies still has not succeeded in making users satisfied. Users still say a lot of negative tweets rather than their positive tweets. Therefore, Go-Pay and OVO companies need to pay attention to and improve those aspects, especially those with the most negative tweets. For Go-Pay, the most are the top up aspect, while for OVO the most is the offline payment. Every company needs to improve the quality of its services to increase customer satisfaction as well.

ACKNOWLEDGMENT

Authors would like to express gratitude and appreciation to Universitas Indonesia for funding this research through PIT-9 Research Grants Universitas Indonesia No: NKB-0061/UN2.R3.1/HKP.05.00/2019

REFERENCES

- [1] "Infografis Penetrasi & Perilaku Pengguna Internet 2017," Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) & Teknopreneur Indonesia, rep., 2018.

- [2] J. Ondrus and Y. Pigneur, "A Disruption Analysis in the Mobile Payment Market," *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 2005.
- [3] C. J. Hoofnagle, J. M. Urban, and S. Li, "Mobile Payments: Consumer Benefits & New Privacy Concerns," *SSRN Electronic Journal*, 2012.
- [4] "Mobile Payments in Indonesia: Race to Big Data Domination," MDI Ventures & Mandiri Sekuritas, Jakarta, rep., 2017.
- [5] T. P. Singh and Dr. RatnaSinha, "The Impact of Social Media on Business Growth and Performance in India," *International Journal of Research in Management & Business Studies*, vol. 4, no. 1, Mar. 2017.
- [6] S. Wakade, C. Shekar, K. J. Liskza, and C.-C. Chan, "Text Mining for Sentiment Analysis of Twitter Data," *International Conference on Information and Knowledge Engineering*, 2012.
- [7] L. Feng, Y. K. Chiam, and S. K. Lo, "Text-Mining Techniques and Tools for Systematic Literature Reviews: A Systematic Literature Review," *2017 24th Asia-Pacific Software Engineering Conference (APSEC)*, 2017.
- [8] I. Surjandari, M. S. Naffisah, and M. I. Prawiradinata, "Text Mining of Twitter Data for Public Sentiment Analysis of Staple Foods Price Changes," *Journal of Industrial and Intelligent Information*, vol. 3, no. 3, 2014.
- [9] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool, 2012.
- [10] R. Aggarwal and L. Gupta, "A Hybrid Approach for Sentiment Analysis using Classification Algorithm," *International Journal of Computer Science and Mobile Computing*, vol. 6, no. 6, pp. 149–157, 2017.
- [11] S. Goyal, "Sentimental Analysis of Twitter Data using Text Mining and Hybrid Classification Approach," *International Journal of Advance Research, Ideas, and Innovations in Technology*, vol. 2, no. 5, 2016.
- [12] N. Li and D. D. Wu, "Using text mining and sentiment analysis for online forums hotspot detection and forecast," *Decision Support Systems*, vol. 48, no. 2, pp. 354–368, 2010.
- [13] C.-M. Tan, Y.-F. Wang, and C.-D. Lee, "The use of bigrams to enhance text categorization," *Information Processing & Management*, vol. 38, no. 4, pp. 529–546, 2002.
- [14] J. Qiu, C. Liu, Y. Li, and Z. Lin, "Leveraging sentiment analysis at the aspects level to predict ratings of reviews," *Information Sciences*, vol. 451–452, pp. 295–309, 2018.
- [15] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," *Expert Systems with Applications*, vol. 57, pp. 117–126, 2016.
- [16] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery," *Proceedings of the twelfth international conference on World Wide Web - WWW 03*, 2003.
- [17] M. Hofmann and A. Chisholm, *Text mining and visualization case studies using open-source tools*. Boca Raton: CRC Press, Taylor & Francis Group, 2016.
- [18] T. Joachims, "Text categorization with Support Vector Machines: Learning with many relevant features," *Machine Learning: ECML-98 Lecture Notes in Computer Science*, pp. 137–142, 1998.
- [19] B. Santosa and A. Umam, *Data Mining & Big Data Analytics*, vol. 1. Yogyakarta: Penebar Media Pustaka, 2018.