Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes

Hilman Wisnu*, **Muhammad Afif and Yova Ruldevyani**

Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia

*Corresponding author's email: hilman.wisnu@ui.ac.id

Abstract. Indonesia payment behaviour has turned from traditional to digital as the impact of the technology growth. Digital payment usage in Indonesia have increased rapidly in recent years. Many companies offer this service with different terms and tariff options. Social media is one of the places where people can express their feeling and opinion, including Twitter. In this research, sentiment analysis and opinion mining is conducted to see public satisfaction towards the digital payment service in Indonesia (OVO, GO-PAY and LinkAja). This research is using Twitter data and has several stages, which are data crawling from Twitter, data cleaning, feature selection and classification using two machine learning approach (Naïve bayes classifier and K-Nearest Neighbour or KNN). The raw data is processed to get the clean data, and to get the appropriate feature for classification algorithm and then perform classification and validation to the model. As for the classification algorithm, this research finds out that KNN has better accuracy than Naive Bayes. The result of this research also shows that LinkAja and G0-PAY has more neutral sentiment or customers nearly satisfied of the services provided, and OVO has more negative sentiment than neutral sentiment

1. Introduction

Mass adoption of digital payment globally seems rare, however it does exist. Driven by rapid advances in digital payments, the global payment landscape is undergoing a profound transformation [2]. In digital economy era, Central Bank of Indonesia implicated that better service and changes in business model will result in increase efficiency [3]. There are at least three popular private digital payment service providers in Indonesia, there are Go-Pay (launched in 2016), Ovo (launched in 2017), and LinkAja (launched in 2007) [4].

Another digital aspect in today's internet era is social media, where people able to express their opinion freely from anywhere and anytime. Per January 2019, Twitter has at least 326 monthly active users [5]. As one of communication channels to customers, the three digital payment service providers (Go-Pay, Ovo, and Link Aja) also have their official Twitter account (@gopayindonesia since August 2018, @ovo_id since September 2016, @linkaja since May 2015). They used Twitter to engage customers andThe SAC value produced by the S-boxes construction that have been carried out by previous researchers are various. Girija and Singh [7] developed S-boxes with a double random phase encoding (DRPE) system. Farwa et al. [8] built an S-box with a specific nonlinear and iterative map approach. Çavuşoğlu et al. [9] developed an S-box design by placing an 8-bit value taken from a random number generator (RNG). Hussain et al. [10] built S-box construction based on quantum

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

magnets and share the marketing campaign of their digital payment service, which visible from the context of posting.

Number of tweets from entire digital payment service providers Twitter account may be gathered with minimum 4.000 tweets and maximum 40.000 tweets. The growth of textual data sources has been broadly defined as text mining, because we can extract useful information and explore interesting patterns. Nevertheless, social media text mining requires more hard work since the data often contains slang, abbreviation or simply non standard Indonesian language vocabulary.

Since Twitter became a gold mine for organisations to monitor their reputation and brands, researchers have been using text mining for social media sentiment analysis through Twitter. Fitri et al. execute sentiment analysis through Twitter tweets to find out customer satisfaction level from data cellular services, and adopt Naive Bayes classifier algorithm [9]. Other researcher contributes to the sentiment analysis for classifying customers review and applied several machine learning based classification algorithms like Naive Bayes, Maximum entropy and SVM [10]. The research found out Naive Bayes technique gives better result than the other two classification algorithm. Similar to the first research about social media sentiment analysis, Turban et al. analysed about how to measure brand reputation based on people response on their services quality through customer's sentiment analysis from Twitter data [19].

Based on previous Twitter sentiment analysis, researchers use this text mining technique which is sentiment analysis either to know customer's opinion through their tweets or other textual data source [9], [19], or to analyze which classification algorithm has the best performance amongst some algorithms [10], [16], [18]. The limitation was found that previous researchers tend to use Naive Bayes classifier algorithm in sentiment analysis through Twitter. While for research about digital payment in Indonesia [13, 14, 15] more concerned to the customer's intention to use the system, public services performance evaluation, and relationship between characteristics and intention in e-payment adoption. The existing and trend of digital payment services in Indonesia have led to the need of executing this research to discover their reputation through public customer satisfaction, which data gathered from Twitter and analysed with sentiment analysis.

2. Literature Study

In general, digital or electronic payment (e-payment) refers to transactions that conducted via Internet, although there are many other forms of electronic payment [15]. In service context, overall satisfaction is similar to service quality evaluations and historically, satisfaction has been used to describe loyalty as behavioral intentions (i.e., likelihood of purchasing and recommending [11].

Beside digital payment and customer satisfaction, text mining in a manner analogous to data mining seeks to extract useful information by identifying and exploring interested patterns [8]. Text mining is one of the current developed text mining technique [6]. A line of research that allows to determine people's attitude and opinions in relation to different topics, products, services, events, and their attributes known as sentiment analysis [17]. Often, sentiment analysis was done for studying people's attitudes, emotions toward entities, and opinions toward events or topics [1].

In order to categorize data sets to relevant classes, data mining classification algorithm has emerged, which of two are Naive Bayes and K-Nearest Neighbors.

2.1 Naive Bayes Classifier

Naive Bayes is the most classification methods that can be used to decide sentiment analysis. This Method calculates the probabilities for every factor, then it selects the outcome with the highest probability. There are many applications of Naive Bayes Classifier, such as for text classification, spam filtering, sentiment analysis and many others. Naïve Bayes can be used to make prediction in realtime. Following is Naïve Bayes Formula,

$$P\,(M \mid \mathrm{A}) = \frac{P(A|M)\,P(M)}{P\,(A)}$$

P (M) : The Probablity of the hypothesis M being true, false or neutral. this is known as the prior probability
P (A) : The probability of the evidence
P (MlA) : The probability of the evidence given that hypothesis is true
P (AlM) : The probability of the hypothesis given that the evidence is here

2.2 K Nearest Neighbors (KNN) Classifier

K-Nearest Neighbor (KNN) is one method machine learning who classify objects based on learning data closest to the object. This method is very simpled, easied to represent, had toughness to train data who has a lot noise, and effective for the grouping process. The purpose of this algorithm classifies new objects, attributes and sample training The best k value for this algorithm depends on the data. Primarily, a high k value will reduce the effect of noise on classification, but make the boundaries between each classification increasingly blurred. A good k value can be chosen by parameter optimization, for example by using cross-validation [11].

Special cases where classification is predicted based on training dataclosest (in other words, k = 1) is called the nearest neighbor algorithm. The accuracy of the KNN algorithm is influenced by the presence or absence of irrelevant features or if the value of the feature is not equivalent to its relevance to classification. Most of the research on this algorithm discusses how to select and weight features so that *classification performance is better.*

3. Research Model

This study is about sentiment analysis with experimental study referred to existing research which measure customer’s sentiment analysis from Twitter data, since the good responses from people could awake the desire for a product until it affects intention for buying the products [20]. The research resulted that Naïve Bayes has the best precision rate between other two algorithms [20]. As for the other classifier algorithm, KNN implemented to find out if the proposed KNN will reduce complexity when being used in big data [7]. It is concluded that KNN and Naïve Bayes algorithm have higher precision. Also, sentiment analysis as part of lexicon-based methods require few effort in human-labelled document and is not sensitive to the quantity and quality of training data set [17, 21].

The subject of this research is about digital payment service in Indonesia and there is developed measurement approaches for customer satisfaction known as expectancy-disconfirmation [22]. The adaptation-level theory defines when customers compare actual product and service performance with their prior expectations. Also, there are two approaches to investigate confirmation and disconfirmation of expectations, there are inferred approach and direct method [22]. The direct approach used for summary judgment scales to measure confirmation and disconfirmation. Finally, the proposed research model can be seen in Figure 1.

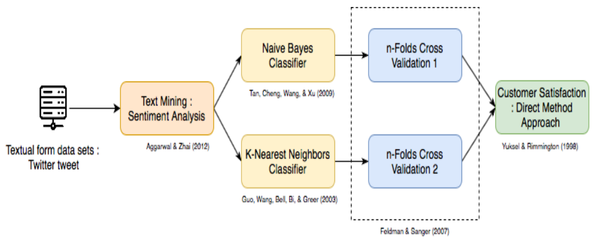

Figure 1. Proposed research model

The model described this research started from the raw textual data source of Twitter, proceed with sentiment analysis and classified with Naïve Bayes and KNN classifier algorithms, validated with n-folds cross validation, and summarize the customer satisfaction. However, KNN is not processed feature selection stage for POS tagging the training and testing data as Naïve Bayes classifier did. In the end, the research figures customer satisfaction of digital payment services in Indonesia, either positive, neutral, or negative.

4. Research Methodology

This study concentrated on textual information of the message social media and used step by step methodology for filtering text information to find customer sentiment analysis and customer satisfaction level digital payment services in Indonesia. This study also adopted hybrid techniques that combine both main techniques in order to improve classification performance and accuracy [17, 21]. Supervised machine learning method implemented, with KKN and Naïve Bayes algorithm.

Unsupervised learning implemented for creating training data. We used Twitter dataset of each digital payment service providers Twitter official account. However, for the data mining process implemented, we referred to Cross-Industry Standard Process for Data Mining (CRISP-DM) that served as a nonproprietary standard methodology for data mining [23]. As shown in Figure 2, this framework consists of six stages, from business understanding to deployment.

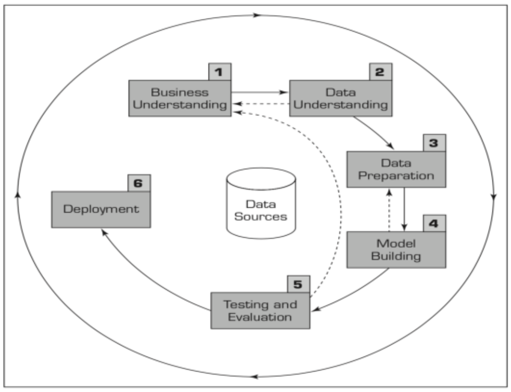

Figure 2. CRISP-DM data mining process

The main process of data mining technique adopted in this study are as described below

4.1 Business Understanding

The purpose of this study is to understand how satisfied customers of digital payment service providers (Go-Pay, Ovo, and Link Aja) to the digital payment service, which captured from their expression in Twitter tweets. This study demanded answer of “How is customer satisfaction of each digital payment services in Indonesia : a case of Go-Pay, Ovo, Link Aja don social media twitter?”. Another question to be solved and understood through this study is “How to measure customer satisfaction level using sentiment analysis with Naive Bayes and KNN classifier algorithm?”.

4.2 Data Understanding

Data understanding is done by crawling the data sets, with purpose to collect or download data from a database server Twitter, data in database contained opinion sentences on tweet digital payment service in official Twitter account (Gopay @gopayindonesia, Ovo @ovo_id, and Link Aja @linkaja) in Indonesian language. The comment selection stage performed and the brief description as below to find negative, positive or neutral score. The date range of collected tweets are January to April 2019.

Social media comments are taken using the application programming language in Python using PyQuery and the URLlib library, because there are limitations to TwitterAPI. The number of tweets from @gopayindonesia resulted in 2095, 4.001 from @ovo_id, and 31541 from @linkaja. Collected Tweets are stored in the new excel file and separated from the previous raw data.

4.3 Data Preparation
After collecting data as main feature of business understanding, next step is to prepare the data for analysis. This process include data cleaning, data transformation, and data reduction.

4.3.1 Data Cleaning
The first process manages basic cleaning operations, which consist in invalid tweet, disturbing elements or symbols. In order to provide only significant information, a clean tweet should not contain URLs, hashtags (i.e. #gojek) or mentions (i.e. @gojek). This process made valid to training data;

4.3.2 Data Transformation
Second step is to fold the case and filter. Case folding means lowering all the words include first words in sentence. While filtering uses stopword list Indonesian Language library (KBBI) to rearrange database comment sentence;

4.3.3 Data Reduction
The next operation is stemming and lemmatization, to remove the word repeated in sequence. This process will results in normalized words, two words written in a different way (i.e. sayyyyyaaaa and saya) will become equals. After the words are normalized, tokenization or bag-of-words creation divides a sentence into several words according to Indonesia Dictionary (KBBI). They are using commas to separate and identify individual words are separate commas (,).

4.4 Model Building
In modeling and classifying data, Naive Bayes and KNN used. Both algorithm known to have best performance and reliability amongst data mining classification algorithm. The used of two algorithm named as cross-validation technique, it is done after manual classification assessment for the training data set. The assessment is carried out by two respondents.

4.5 Testing and Evaluation
Model to test data has been developed in previous step, then testing the data executed with n-fold cross-validation technique. The data folded several times until it gives the best accuracy. This study uses K-5 to K-20 where data is sought to find perfect multiples. After best accuracy found from between the n-fold, the sentiment from results chosen and defined.

4.6 Deployment
Final process of data mining technique proposed in this research is to create the summary of entire mining process and sentiment analysis result. The report will conclude about customer satisfaction toward digital payment services in Indonesia. Further recommendations also provided to complete what lacks from this study.

5. Analysis and Result
5.1 Business Understanding
First step is crawling the data from Twitter official account of each digital payment service provider analysed in this research (Go-Pay, Ovo, and Link Aja). The process include from accessing PyQuery library and URLLib to gather all the tweets for the past four months to convert the text result to CSV format. URLLib will response the HTML tag served by running PyQuery library in JSON form, In the end, the columns in final document are username, date, retweets, favorites, text, geo, mentions, hashtags, id, permalink.

5.2 Data Preparation
The visualization of preparation steps for tweets gathered from three Twitter official accounts (@gopayindonesia, @ovo_id, @linkaja) can be seen in Table I.

5.3 Model Building
During the time data prepared, model building also conducted concurrently. This model will be used for training the data and will be tested afterwards. The flow of model building can be seen on Figure 3.

Table 1. Data Preparation Procedure

Procedure	**Data**
Real Data	Thx min @ gojekindonesia @ gopayindonesia saldo go-pay sy udah nambah.. pic.twitter.com/LBhPlFwrkp
After Cleansing Data	Thx min saldo go-pay sy udah nambah..
After Case Folding Data	thx min saldo go-pay sy udah nambah..
After Filtering Data	thx min saldo go-pay sy udah nambah..
After Stemming and Lemmatization Data	Terimakasih admin saldo go-pay saya sudah tambah
After Tokenization Data	“terimakasih”, “admin”, “saldo”,“go-pay”, “saya”,” sudah”,“tambah”

5.4 Testing and Evaluation
5.4.1 Data Training and Testing
The training and testing of data executed with cross validation method, Naive Bayes and KNN. We also used K-fold to evaluate model performance with multiple 5 folds. Guo et al. described n-fold cross-validation is a commonly used method to smooth out variations in the corpus [11]. We divided the entire document into multiple 5 parts, and the training and testing process run 5, 10, 15, and 20 times, where different part of collection used each time as test set. This process were done until 20-folds has reached average accuracy level from the whole iteration. Table II showed the number of training data provides and number of testing data (training data or number of folds) to know level of customer satisfaction from each digital payment service providers.

5.4.2 Algorithms Analysis and Evaluation
This research uses Naive Bayes and KNN classification algorithm, the result shows Naive Bayes has the best data accuracy when doing 20-folds cross validation. In Table 3, there is classification of each algorithms accuracy after classifying digital payment service providers Twitter tweets. From the table we can see GO-PAY’s best accuracy (83.50%) reached after running 20-folds with KNN algorithm. For OVO and LinkAja, the accuracy reached its best (84.00% and 91.00%) after 20-folds cross validation with KNN algorithm.

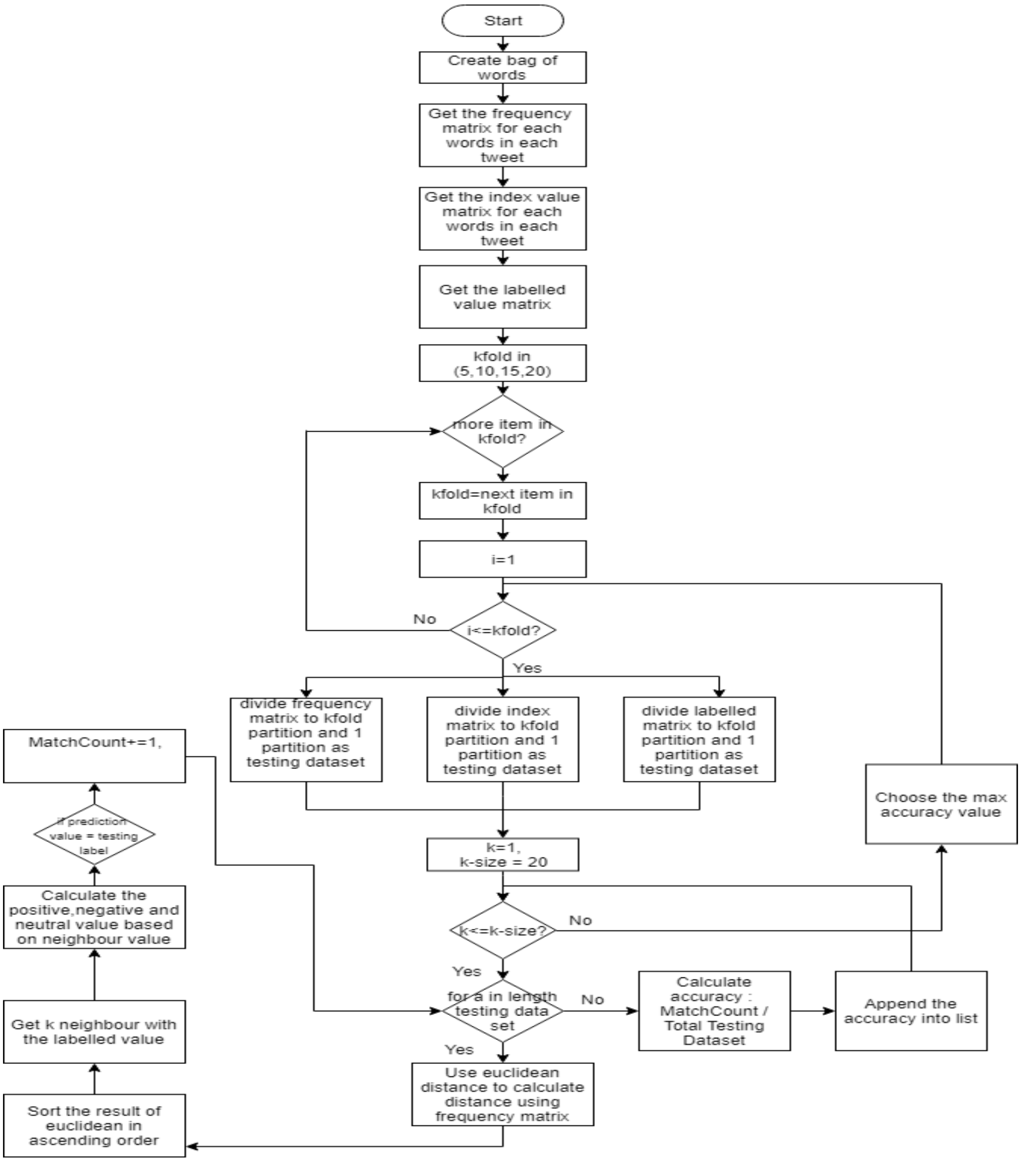

Figure. 3. Model building

5.5 Deployment: Analysis Data and Measuring Customer Satisfaction

After running some stages of crawling data, data preparation, model building, and testing and evaluation, we analyze the data result and measure customer satisfaction toward digital payment service providers in Indonesia through customer's tweets. Table III showed KNN gives the best accuracy for all digital payment service providers, specifically on 20-folds. The result showed that GO-PAY is the only one service providers that has positive sentiment from their customer Twitter tweets (3 positive comments). For neutral sentiment, LinkAja has the most for 91 neutral comments. Meanwhile, the negative sentiment took the most (55 negative comments) by OVO.

Table 2. Number of training and testing data sets

Testing	Training Data	Testing Data
GO-PAY	1900 comments	100 comments
OVO	1900 comments	100 comments
Link Aja	1900 comments	100 comments

The conclusion of customer satisfaction measurement with sentiment analysis through customer Twitter tweets is no positive sentiment from all digital payment service providers, LinkAja has the most neutral sentiment and the second most neutral sentiment is GO-PAY. For negative sentiment, OVO has the most for 55, GO-PAY has 22, and LinkAja has 0. GO-PAY and LinkAja customers tend to be more neutral toward services provided by GO-PAY and LinkAja through their Twitter official account. While for OVO, customers tend to be less satisfied due to more negative sentiment than neutral or positive. The overall sentiments number can be seen on table 4.

Table 3. Classification algorithm accuracy

	GO-PAY		**OVO**		**Link Aja**	
K	Naïve Bayes	KNN	Naïve Bayes	KNN	Naïve Bayes	KNN
5	62.53%	74,74%	75.60%	73.25%	69.5%	85.25%
10	67.21%	77.31%	75.70%	78.50%	70.00%	87.50%
15	70.15%	80.62%	76.05%	79.10%	70.55%	88.72%
20	70.71%	83.50%	75.75%	84.00%	70.60%	91.00%

The conclusion of customer satisfaction measurement with sentiment analysis through customer Twitter tweets is no positive sentiment from all digital payment service providers, LinkAja has the most neutral sentiment and the second most neutral sentiment is GO-PAY. For negative sentiment, OVO has the most for 55, GO-PAY has 22, and LinkAja has 0. GO-PAY and LinkAja customers tend to be more neutral toward services provided by GO-PAY and LinkAja through their Twitter official account. While for OVO, customers tend to be less satisfied due to more negative sentiment than neutral or positive. The overall sentiments number can be seen on Table 4.

6. Discussion and Implication

This study provides theoretical and practical implications. For theoretical implications, this study applies sentiment analysis to find customer satisfaction from digital payment services in Indonesia with percentage accuracy and conclusion opinion mining (ect. positif, negative and neutral), by used method to datasets with a large scale will produce a higher accuracy effect and real of conclusion opinion. The conclusion from previous studies [17], the number of K-N folds is greater then the accuracy is greater, in this research not applied in @OVO_ID result, it depends on the clean of data. Naive Bayes method accuracy is lower than the KNN method with same as amount dataset, prove that the statement from previous research [9] wasn't applied for the case of comparison between Naive Bayes and KNN algorithm.

Table 4. Sentiment analysis digital payment services

	K	Naive Bayes Classifier			KNN Classifier		
		Positive	Neutral	Negative	Positive	Neutral	Negative
GO-PAY	5	0	333	0	12	218	60
	10	0	186	0	0	126	24
	15	6	128	0	0	88	16
	20	2	87	0	3	56	22
OVO	5	1	191	136	0	96	197
	10	0	105	60	0	55	102

	15	0	8	115	0	34	72
	20	0	33	56	0	29	55
Link Aja	5	0	284	18	0	341	0
	10	0	149	12	0	174	1
	15	1	102	4	0	115	3
	20	1	81	4	0	91	0

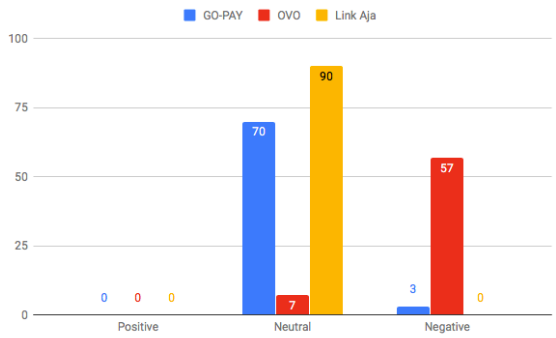

Figure 4. Digital payment providers testing result

For practical implications, this study suggests digital payment service providers to increase service response complaint in application digital payment used twitter official accounts, so that the number of positive comments may increase. Related to previous practical implications, digital payment service providers have high chance to realize positive sentiment goals, they are implemented by automatic text mining techniques and their customer service tools for complaint. These will increase positive response and number of satisfied customers, due to fast respond by providers with the fact customers tend to need the fast response.

7. Conclusions And Future Works

This research purpose is to define the customer satisfaction of digital payment services in Indonesia. Therefore, the data samples were gathered from three digital payment service providers Twitter official account. Refer to the result, it is known that KNN classifier algorithm has better accuracy than Naive Bayes. As for the customer satisfaction, LinkAja has the most neutral sentiment on their customer's tweets and 0 negative sentiment. Despite of OVO with the most negative sentiment, 0 positive and neutral sentiment on their customer's tweets. While GO-PAY customers tend to be more neutral than not satisfied. Based on the result, there are several suggestions we propose for the future research, 1) To apply other classification algorithm that best known in Twitter sentiment analysis method or compare results of Naive Bayes and a well known Support Vector Machine (SVM), since both algorithms often compared in previous research to know people's sentiment; 2) To execute more data pre-processing and pre-classification stage such as part of speech (POS) tagging; 3) To research about customer satisfaction toward digital payment services in Indonesia with other methodology, not limited with text mining methodology.

References

[1] Aggarwal, C. C., & Zhai, C. (Eds.). (2012). Mining text data. Springer Science & Business Media

[2] Anonymous. (2013). Digital Payments Transformation: From transactions to consumer interactions. Accenture

[3] Anonymous. (2017). Digital Payment Transformation. Central Bank of Indonesia

[4] Anonymous. (2018). Indonesia Social Media Trend 2018: 2nd Semester. Jajak Pendapat

[5] Anonymous. (2019). Digital 2019 : Global Digital Overview. Smart Insight

[6] Budi, S. (2017). Text Mining for Sentiment Analysis of Film Review using K-Means Algorithm (in Bahasa). Techno. Com,16(1), 1-8

[7] Deng, Z., Zhu, X., Cheng, D., Zong, M., & Zhang, S. (2016). Efficient kNN classification algorithm for big data. Neurocomputing, 195, 143-148

[8] Feldman, R., & Sanger, J. (2007).The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge university press

[9] Fitri, F. S., Si, M. N. S., & Setianingsih, C. (2018, November). Sentiment Analysis on the Level of Customer Satisfaction to Data Cellular Services Using the Naive Bayes Classifier Algorithm. In 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)(pp. 201-206). IEEE

[10] Gautam, G., & Yadav, D. (2014, August). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In 2014 Seventh International Conference on Contemporary Computing (IC3)(pp. 437-442). IEEE

[11] Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003, November). KNN model-based approach in classification. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"(pp. 986-996). Springer, Berlin, Heidelberg

[12] Gustafsson, A., Johnson, M. D., & Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. Journal of marketing, 69(4), 210-218

[13] Hermana, B., & Silfianti, W. (2011). Evaluating e-government implementation by local government: digital divide in internet based public services in Indonesia. International Journal of Business and Social Science, 2(3)

[14] G. Liu, W. Yang, W. Liu, and Y. Dai, "Designing S-boxes based on 3-D four-wing autonomous chaotic system," Nonlinear Dyn., vol. 82, no. 4, pp. 1867–1877, 2015

[15] Junadi[a], S. (2015). A model of factors influencing consumer's intention to use e-payment system in Indonesia. Procedia Computer Science, 59, 214-220

[16] Kharde, V., & Sonawane, P. (2016). Sentiment analysis of twitter data: a survey of techniques. arXiv preprint arXiv:1601.06971

[17] Kolchyna, O., Souza, T. T., Treleaven12, P. C., & Aste12, T. (2015). Methodology for Twitter Sentiment Analysis. arXiv preprint arXiv:1507.00955

[18] Saif, H., He, Y., & Alani, H. (2012, November). Semantic sentiment analysis of twitter. In International semantic web conference(pp. 508-524). Springer, Berlin, Heidelberg

[19] Turban, E., Sharda, R., & Delen, D. (2014). Business intelligence and analytics: systems for decision support. Pearson Higher Ed

[20] Vidya, N. A., Fanany, M. I., & Budi, I. (2015). Twitter sentiment to analyze net brand reputation of mobile phone providers. Procedia Computer Science, 72, 519-526

[21] Vohra, S. M., & Teraiya, J. B. (2013). A comparative study of sentiment analysis techniques. Journal JIKRCE, 2(2), 313-317

[22] Yüksel, A., & Rimmington, M. (1998). Customer-satisfaction measurement: Performance counts. Cornell Hotel and Restaurant Administration Quarterly, 39(6), 60-70