Sentiment Analysis of Ojek Online User Satisfaction Based on the Naïve Bayes and Net Brand Reputation Method

1st Alam Rahmatulloh
Department of Informatics
Siliwangi University
Tasikmalaya, Indonesia
alam@unsil.ac.id

2nd Rahmi Nur Shofa
Department of Informatics
Siliwangi University
Tasikmalaya, Indonesia
rahminurshofa@unsil.ac.id

3rd Irfan Darmawan
Department of Information System
Telkom University
Bandung, Indonesia
irfandarmawan@telkomuniversity.ac.id

4th Ardiansah
Department of Informatics
Siliwangi University
Tasikmalaya, Indonesia
ardicks4@gmail.com

2021 9th International Conference on Information and Communication Technology (ICoICT) | 978-1-6654-0447-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICoICT52021.2021.9527466

***Abstract*—Gojek and Grab are the most popular online motorcycle taxis and are often used today in Indonesia, based on Hootsuite's survey. However, it is not yet known how the response from online motorcycle taxi users. So it is necessary to have a sentiment analysis of online motorcycle taxi users whether they are satisfied or dissatisfied with the drivers and Gojek and Grab companies' services. Twitter with 52% active users of all internet users in Indonesia allows users to write various topics so that to find out the level of user satisfaction with Gojek and Grab. Sentiment analysis can be used as a reference for the development of Gojek and Grab services in the future. They measure the level of satisfaction with the Net Brand Reputation (NBR) method from the Naïve Bayes classification results using the rapid miner tool. The rating with accuracy has an accuracy value of 99.80% for Gojek and 99.90% for Grab. This study shows that more tweets have negative opinions compared to positive opinions for Gojek and Grab. Namely 616 positive opinions and 2317 negative opinions for Gojek drivers, 3560 positive opinions and 6419 negative opinions for Gojek Company. 594 positive opinions, and 1866 negative opinions for Grab drivers. As well as 3516 positive opinions and 4407 negative opinions for Grab Companies. So the results of the sentiment analysis of online motorcycle taxi users are dissatisfaction with either the driver or the company.**

Keywords*—Gojek, Grab, Naïve Bayes, Net Brand Reputation, Sentiment Analysis

I. Introduction

The last few years information technology in Indonesia have transformed, including in the field of transportation [1]. Transportation is an essential means of supporting human activities or mobility every day. It must be well prepared and safe because it is very influential in the economy, delivery of goods or services, passenger transportation, and other things [2]. Various types of traffic in Indonesia, land transportation, are given special attention by the government and users of transportation services. Besides being cheap, land transportation is still the favourite choice of most Indonesian people [3].

Land transportation modes that are easily found and often used in daily life, one of them which is familiarly called Ojek. Ojek is a mode of land transportation using a motorcycle that usually has a base (at the intersection places or in strategic areas such as in front of the station or terminal) to wait for passengers [4]. Currently, there are two types of Ojek services, namely "conventional ojek" and "online ojek" [5]. Technological sophistication and the development of communication tools make it easier for people to affect the shift in the mechanism of conventional transportation service users to online transportation [4].

Some online Ojek in Indonesia includes Gojek, Grab, OK Jack, Indo-Jek, Bang Jek, and others. Gojek and Grab are the most popular online Ojek and are often used today. This is based on a survey conducted by Hootsuite - We Are Social for January 2019 on the Ranking of Mobile Apps By Monthly Active Users in Indonesia, which shows that Gojek and Grab ranks are eighth and ten [6]. Apart from the survey results, we can see on the Google Play Store that the Gojek and Grab applications have been widely downloaded and used. More than 50 million times, the Gojek application was downloaded, while Grab was more than 100 million times.

Each Ojek Online service provider has its advantages that can make users feel satisfied with their service. In addition to excellence, service providers also need to pay attention to the aspects of shortcomings to stay afloat and not move to other service providers. Twitter is one type of social media with 52% active users of all internet users in Indonesia based on a survey conducted by Hootsuite - We Are Social for January 2019 [6]. Twitter allows users to write various topics and discuss problems that occur; unlike some other social media that requires both parties' approval to connect, Twitter enables users to track submissions (called tweets) from other users without getting permission. This is one reason why Twitter is a place for information flow and is used to determine user satisfaction with its services.

How to determine service users' satisfaction for services provided by service providers, both from companies or drivers, it is necessary to do sentiment analysis in which the data to be analyzed is taken from user tweets for each service provider. This Tweet will be categorized based on negative tweets and positive tweets to be used as a reference to determine the level of user satisfaction with service providers.

The rapidminer tools uses Naïve Bayes classification to produce accuracy, precision, recall, and error rate values. This value will be used as a reference to find out the satisfaction level of online Ojek users based on service providers, namely Gojek and Grab.

This study will then measure the level of satisfaction of online Ojek users between Gojek and Grab based on sentiment analysis from Twitter using the Net Brand Reputation method and the Naïve Bayes classification.

II. Related Work

Sentiment analysis or opinion mining is a computational study of people's opinions, sentiments, and emotions through entities and attributes expressed in text form [7]. Sentiment analysis is part of data mining where the data can be classified based on positive or negative opinions [8].

Research on sentiment analysis has been done before. Kristiyanti et al.'s research stated that the Naïve Bayes algorithm is superior to the Support Vector Machine algorithm in classifying public opinion with Indonesian text on Twitter for the prospective governor of the West Java Period 2018-2023 [9]. In their research, Parveen & Pandey stated that the Naïve Bayes algorithm's performance would increase when changing emoticons into meaningful words [10].

Vidya et al. measured the satisfaction of mobile provider users in Indonesia based on reputation and classification. The method used to determine reputation is the Net Brand Reputation (NBR), while the process for classification is Support Vector Machine, Naïve Bayes, and Decision Tree [11].

Based on related research [9], [10] and [11], this study will perform the classification of naïve Bayes and the calculation of net brand reputation (NBR) on opinion data from Twitter to determine the level of customer satisfaction. The addition of emoticon conversions in the preprocessing process was carried out to improve Naïve Bayes' performance.

III. Methodology

The tool used in this study uses Twitterscraper, Microsoft Excel, and Rapidminer. Twitterscraper is a python library for retrieving data from Twitter. Microsoft Excel is used to preprocess and calculate Net Brand Reputation (NBR), while Rapidminer is used in the Naïve Bayes classification. Figure 1 shows the stages of the study.

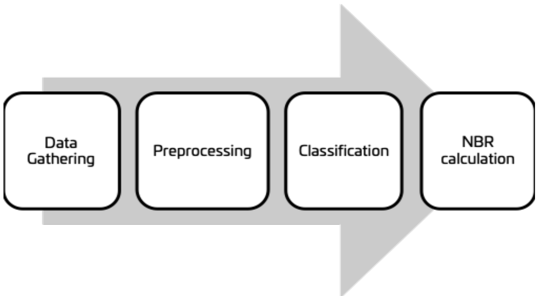

Fig. 1. Methodology

A. Data Gathering

Data retrieval is done using Twitter scraper [12] with the keywords "@gojekindonesia or gojekindonesia" for Gojek and "@GrabID or GrabID" for Grab. They are starting from the beginning of September to the end of December 2019. A Twitter scraper's advantage is that there is no need to use Twitter API for its use, making it easier to retrieve data. Still, the disadvantage of Twitter scrapper is that the number of withdrawals depends on internet speed. Table I is the data taken from Twitter.

TABLE I. Data Gathering from Twitter

Data	Keywords	Values
Gojek	@gojekindonesia or gojekindonesia	125178 *tweet*
Grab	@GrabID or GrabID	130802 *tweet*

A. Preprocessing

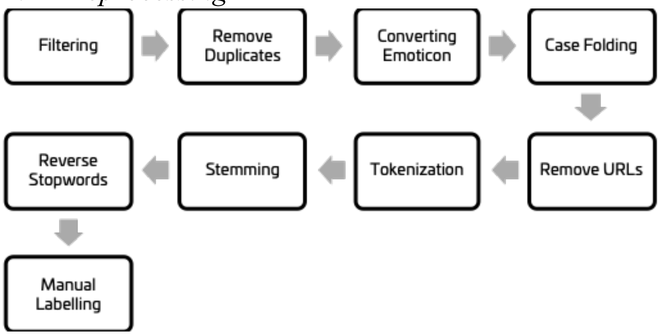

Fig. 2. Preprocessing Step

Preprocessing is done to avoid incomplete data, interference with data, and inconsistent data [13]. Figure 2 shows the preprocessing stages.

1) Filtering is a stage for selecting data according to the data needed.

2) Remove Duplicates, the stage of deleting data considered to have the same tweet and retrieving one from the same tweet.

3) Converting Emoticons, change emoticons into meaningful words [10]. A list of converting emoticons is presented in table II.

TABLE II. Converting Emoticons

Emoticons				Converting
:d	:-d	:)	:-)	Smile
:(	:-(	:[	:-[	Sad
x(	x-(	X(	X-(	Angry
:x	:-x	:X	:-X	Love
:-	:/	:-/		Confused
: (	:" (	:’(		Crying
:D	:-D	=D	=-D	Laughing
B|	B-|	8|	8-|	Cool

4) Case Folding, homogenize letter shapes from large to small [14].

5) Remove URLs, stage of removing a URL to reduce data that is not needed.

6) Tokenization, cutting the input string based on the words that make it up [13].

7) Stemming is the stage of looking for root words by removing affixes to a word [15]. Stemming steps are carried out using Sastrawi libraries [16].

8) Remove Stopwords; stopwords are words that do not have a specific meaning [14]. The remove stopwords stage is performed using the Sastrawi library [16].

9) Manual Labeling, tweets will be labeled as positive and negative.

Data from the preprocessing stage is divided into two parts, namely for drivers and companies, so we get 2933 tweets for Gojek drivers, 9979 tweets for Gojek companies, 2460 tweets for Grab drivers, and 7923 tweets for Grab companies.

B. Classification

The naïve Bayes classifier is based on the Bayes theorem discovered by Thomas Bayes in the 18th century [17]. The Naïve Bayes theorem can be seen in equation (1) [18].

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

C. Classification Evaluation

Classification stages for drivers or companies Gojek and Grab will produce accuracy, precision, recall, and error rate values. These values will be compared with each other to determine which classification is best based on each Gojek and Grab.

D. Calculation of Net Brand Reputation (NBR)

Net Brand Reputation (NBR) is the net reputation value of a brand digitally [11]. The purpose of the NBR is to simplify the measurement of consumer loyalty. The NBR formula can be seen in equation (2).

$$\left(\frac{Positive\ mentions - negative\ mentions}{Positive\ mentions + negative\ mentions}\right) x100\% \quad (2)$$

Net Brand Reputation (NBR) calculations are only based on the number of positive and negative tweets labeled by Gojek and Grab datasets. The NBR value can be between -100 and 100; the higher the positive number, the NBR indicates that more tweets are considered positive.

IV. Result and Analysis

A. Preprocessing

The first stage of processing the results of withdrawal is filtering. The early stage of filtering is to look for tweets with the keyword @gojekindonesia for Gojek and @GrabID for Grab because not all tweets resulting from data retrieval have the word. The results of filtering obtained 56635 tweets for Gojek and 45648 tweets for Grab. The filtering stage is followed by searching and deleting tweets with the words @jokowi and @prabowo because tweets that have the word have meant not to the user's response about the service provided. The results of this advanced filtering eliminated 451 tweets for Gojek and 115 tweets for Grab, so that 56184 tweets were obtained for Gojek and 45533 tweets for Grab.

The second stage is to remove duplicates. This stage eliminates 4555 tweets for Gojek and 5131 for Grab so that it gets 51629 tweets for Gojek and 40402 tweets for Grab.

The third stage is converting emoticons, turning tweets with emoticon symbols into words that have meaning, as described in Table II.

The fourth stage is case folding, changing all tweets that have uppercase letters to lowercase letters.

The fifth stage is to remove the URL; at this stage, tweets with URLs like HTTP, HTTPS, or pic.twitter.com are deleted. Before the next step, the removal duplicates stage is done again; when the URL is removed, several tweets have the same data. The results of remove duplicates delete 575 tweets for Gojek and 387 tweets for Grab to get 51052 tweets for Gojek and 40015 tweets for Grab.

The sixth stage is tokenization, the characters in tweets like (?), (.), (/), (,), (!), (:), (-), (<), (>), ()) , ((), ([), (]), ('), ("), (=), (*), (%), (_), (;), (^), (#), (&), (\), (+), ($) and (|) are removed.

The seventh stage is stemming; at this stage, removes the affixes to a word, such as the word "thwart" to the word "fail," and so forth.

The eighth stage is removed stopwords; at this stage, it eliminates words that do not have specific meanings, such as what, why, how, and so forth.

The ninth stage is manual Labeling; at this stage, each tweet is given a label, whether included in the positive label or negative label, with the number of Labeling for each positive label and negative label of 55 names. The positive and negative label results are then compared if the positive label is more than the negative label. The tweet is included in a positive tweet or class, and otherwise. The number of data changes from the first stage to the last step is presented in Table III.

TABLE III. PREPROCESSING

Preprocessing	**Amount of tweet**	
	Gojek	***Grab***
Filtering	56184	45533
Remove Duplicates	51629	40402
Converting Emoticon	51629	40402
Case Folding	51629	40402
Remove Url	51052	40015
Stemming	51052	40015
Remove Stopword	51052	40015
Manual Labelling	51052	40015

B. Classification

TABLE IV. DATASET DRIVER

Driver	**The amount of data**		**Total Dataset**
	Data Training	***Data Testing***	
Gojek	2433	500	2933
Grab	1960	500	2460

TABLE V. DATASET COMPANY

Company	**The amount of data**		**Total Dataset**
	Data Training	***Data Testing***	
Gojek	8979	1000	9979
Grab	6923	1000	7923

The preprocessing stage result dataset will be divided into testing data and training data. The data sharing depends on the number and number of dataset drivers and companies. The data distribution for drivers is determined as 500 data testing, and the rest is for training, while the data sharing for companies is determined as 1000 data for testing data, and the rest is for training data. The division of data for drivers and companies is presented in table IV and table V.

Classification tests using Naïve Bayes were performed once. Table VI shows that the object's driver found three negative class data that were predicted as positive, while for Grab driver found 2 data in the positive class that were predicted as negative.

Table VII shows that for Gojek, companies found one negative class data that was predicted as positive and one positive class data predicted as negative; for Grab companies found 1 data in the positive class that was predicted as negative.

TABLE VI. DRIVER PREDICTIONS

Predictions	Class			
	Gojek		***Grab***	
	Positive	***Negative***	***Positive***	***Negative***
positive	118	3	114	2
negative	0	379	0	384

TABLE VII. COMPANY PREDICTIONS

Predictions	Class			
	Gojek		***Grab***	
	Positive	***Negative***	***Positive***	***Negative***
positive	399	1	397	0
negative	1	599	1	602

TABLE VIII. CLASSIFICATION OF DRIVERS

Driver	Accuracy	Precision	Recall	Error Rate
Gojek	99.40%	97.52%	100%	0.60%
Grab	99.60%	100%	98.28%	0.40%

Table VIII classification results for Gojek drivers show an accuracy value of 99.40%, because of a total of 500 pieces of classification data, 3 negative class data that are wrongly predicted as positive classes, so the level of prediction error (classification error) is 0.60%. The result of precision is 97.52% because when doing classification, 118 positive class data are predicted to be correct as positive classes from the total data are predicted to be positive (3 negative classes are wrongly predicted as positive classes and 118 positive classes are predicted to be correct as positive classes), while the result of the recall is 100% because 118 positive classes that are predicted to be correct as positive classes out of the total positive class (118 positive classes that are predicted to be correct as positive classes and 0 positive classes that are predicted incorrectly as negative classes).

The classification results for Grab drivers show an accuracy value of 99.60% because of a total of 500 pieces of classification data, 2 positive class data are wrongly predicted as negative classes, so the level of prediction error (classification error) is 0.40%. The result of precision is 100% because when classifying, 114 positive class data are predicted to be correct as positive classes from the total data that are predicted to be positive (0 negative classes are predicted incorrectly as positive classes and 114 positive classes are predicted correctly as positive classes), while the result of recall is 98.28% because 114 positive classes are predicted to be correct as positive classes out of the total positive class (114 positive classes that are predicted to be correct as positive classes and 2 positive classes that are wrongly predicted as negative classes).

The classification results for drivers can be concluded that the Grab driver is better than the Gojek driver based on accuracy, which indicates that Grab's accuracy value is higher than the accuracy value for Gojek.

TABLE IX. CLASSIFICATION OF COMPANY

Company	Accuracy	Precision	Recall	Error Rate
Gojek	99.80%	99.75%	99.75%	0.20%
Grab	99.90%	100%	99.75%	0.10%

The classification results in table IX for the Gojek company show an accuracy of 99.80% because of a total of 1000 classification data, 1 negative class data is wrongly predicted as a positive class and 1 positive class data that is wrongly predicted as a negative class, so the level of prediction error (classification) error) of 0.20%. The result of precision is 97.75% because when classifying, 399 positive class data are predicted to be correct as a positive class from the total data that are predicted to be positive (1 negative class is wrongly predicted as a positive class and 399 positive classes are predicted to be correct as a positive class), while the result of recall was 99.75% because 399 positive classes were predicted to be correct as positive classes out of the total positive class (399 positive classes that were predicted to be correct as positive classes and 1 positive class which was predicted incorrectly as negative classes).

The classification results for the Grab company shows an accuracy of 99.90%, because of a total of 1000 classification data, 1 positive class data is wrongly predicted as a negative class, so the level of prediction error (classification error) is 0.10%. The result of precision is 100% because when classifying, 397 positive class data are predicted to be correct as positive classes from the total data that are predicted to be positive (0 negative classes are predicted incorrectly as positive classes and 397 positive classes are predicted correctly as positive classes), while the result of the recall was 99.75% because 397 positive classes were predicted to be correct as positive classes out of the total positive class (397 positive classes that were predicted to be correct as positive classes and 1 positive class which was predicted incorrectly as negative classes).

The company's classification results can be concluded that the Grab driver is better than the Gojek driver based on the accuracy value, which indicates that Grab's accuracy value is higher than the accuracy value for Gojek.

The accuracy value is used as a reference as the final result of classification because accuracy predicts the closeness between the algorithm used and the original value. Table VIII and Table IX are the results of the classification of drivers and companies.

C. Calculation of Net Brand Reputation (NBR)

Welding to determine whether the tweet is included in a positive or negative tweet is done in the manual Labeling preprocessing stage. The total tweets for Gojek and Grab are presented in table X.

TABLE X. NET BRAND REPUTATION

Tweet	Driver		Company	
	Positive	*Negative*	*Positive*	*Negative*
Gojek	616	2317	3560	6419
Grab	594	1866	3516	4407

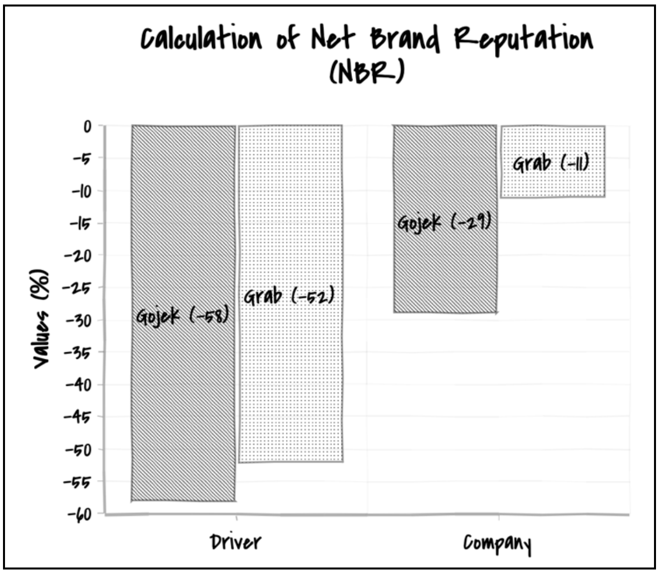

Fig. 3. NBR calculation for Gojek and Grab

Net Brand Reputation (NBR) calculation results for drivers with a value of -58% for Gojek and -52% for Grab, while the results of the calculation of NBR for companies with a value of -29% for Gojek and -11% for Grab. Calculation results for drivers and companies are presented in Figure 3.

Based on Figure 3 that can be seen, both produce negative values, which means that Gojek and Grab are still lacking in service. Both the driver and the two companies.

V. CONCLUSIONS

The results showed that the online motorcycle taxi companies that this study discussed were Gojek and Grab, which had many users using their services. However, they do not yet know the public's perception of the quality of service they provide to their customers. Usually, opinions submitted by users are ignored by companies. From the keywords @gojekindonesia, gojekindonesia, @GrabID, and GrabID, we extract positive and negative opinions then process these sentiments with a classification algorithm. The classification results show that Naïve Bayes produces an accuracy value of 99.80% for Gojek and 99.90% for Grab. Meanwhile, the NBR calculation results show that the Grab company has a better user satisfaction score than Gojek. Although the conclusion from the data that has been processed in this study results in customer dissatisfaction with the two companies.

Further studies must create a dashboard system that can monitor real-time NBR scores and data retrieval using scraping techniques. Because the amount of data retrieval dramatically affects the results of the classification and calculation of NBR. So that user feedback data can be known more precisely as input for online motorcycle taxi companies.

REFERENCES

[1] A. A. Anwar, “Online vs Konvensional: Keunggulan dan Konflik Antar Moda Transportasi di Kota Makassar,” *Jurnal Etnografi Indonesia*, vol. 2, pp. 220–246, 2017.

[2] A. N. Soleh, C. Harini, and Djamaludin, “The Effect Of Service Quality, Price And Trust To Customer Satisfaction Users Of Transportation Services Online Ojek,” *Journal of Management*, 2018.

[3] I. Farida, A. Tarmizi, and Y. November, “Analisis Pengaruh Bauran Pemasaran 7P Terhadap Go-Jek Online,” *Jurnal Riset Manajemen Dan Bisnis*, vol. 1, no. 1, pp. 31–40, 2016.

[4] R. Fathy, “Modal Sosial Ojek Pangkalan : Adaptasi Terhadap Aplikasi Online Transportasi Publik 1 Social Capital of the Ojek Pangkalan : Adaptation To Online Public Transportation Aplication,” vol. 20, no. 2, pp. 191–210, 2018. DOI: 10.14203/jmb.v20i2.635

[5] F. D. Amajida, “Kreativitas Digital Dalam Masyarakat Risiko Perkotaan: Studi Tentang Ojek Online ‘Go-Jek’ Di Jakarta,” *Informasi*, vol. 46, no. 1, p. 115, 2016. DOI: 10.21831/informasi.v46i1.9657

[6] H. we are social, “Hootsuite We are Social Indonesian Digital Report-2019,” 2019.

[7] B. Liu, “Sentiment Analysis and Opinion Mining,” no. May, pp. 1–168, 2012.

[8] L. Dey, S. Chakraborty, A. Biswas, B. Bose, and S. Tiwari, “Sentiment Analysis of Review Datasets Using Naïve Bayes‘ and K-NN Classifier,” *International Journal of Information Engineering and Electronic Business*, vol. 8, no. 4, pp. 54–62, 2016. DOI: 10.5815/ijieeb.2016.04.07

[9] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin, and L. Marlinda, “Comparison of SVM Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter,” *2018 6th International Conference on Cyber and IT Service Management, CITSM 2018*, no. Citsm, pp. 1–6, 2019. DOI: 10.1109/CITSM.2018.8674352

[10] H. Parveen and S. Pandey, “Sentiment analysis on Twitter Data-set using Naive Bayes algorithm,” *Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016*, pp. 416–419, 2017. DOI: 10.1109/ICATCCT.2016.7912034

[11] N. A. Vidya, M. I. Fanany, and I. Budi, “Twitter Sentiment to Analyze Net Brand Reputation of Mobile Phone Providers,” *Procedia Computer Science*, vol. 72, pp. 519–526, 2015. DOI: 10.1016/j.procs.2015.12.159

[12] R. R. el Akbar, R. N. Shofa, M. I. Paripurna, and Supratman, “The Implementation of Naïve Bayes Algorithm for Classifying Tweets Containing Hate Speech with Political Motive,” in *2019 International Conference on Sustainable Engineering and Creative Computing (ICSECC)*, 2019, pp. 144–148 [Online]. DOI: 10.1109/ICSECC.2019.8907208

[13] A. Fathan Hidayatullah and A. Sn, “ANALISIS SENTIMEN DAN KLASIFIKASI KATEGORI TERHADAP TOKOH PUBLIK PADA TWITTER,” *Seminar Nasional Informatika*, vol. 2014, no. semnasIF, pp. 115–122, 2014.

[14] S. Anastasia and B. Indra, “Twitter sentiment analysis of online transportation service providers,” *2016 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2016*, pp. 359–365, 2017. DOI: 10.1109/ICACSIS.2016.7872807

[15] A. Rahmatulloh, N. I. Kurniati, I. Darmawan, A. Z. Asyikin, and D. Witarsyah J, “Comparison between the Stemmer Porter Effect and Nazief-Adriani on the Performance of Winnowing Algorithms for Measuring Plagiarism,” *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 4, p. 1124, Aug. 2019 [Online]. DOI: 10.18517/ijaseit.9.4.8844

[16] K. K. Purnamasari and I. S. Suwardi, “Rule-based Part of Speech Tagger for Indonesian Language,” *IOP Conference Series: Materials Science and Engineering*, vol. 407, p. 012151, Sep. 2018 [Online]. DOI: 10.1088/1757-899X/407/1/012151

[17] F. Handayani and S. Pribadi, “Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110,” *Jurnal Teknik Elektro*, vol. 7, no. 1, pp. 19–24, 2015.

[18] J. Han, M. Kamber, and J. Pei, *Data mining: Data mining concepts and techniques*. 2013. DOI: 10.1109/ICMIRA.2013.45