

SENTIMENT ANALYSIS OF WEST JAVA INTERNATIONAL AIRPORT (BIJB) KERTAJATI ON TWITTER

Ni Putu Oka Intan Yustya Putri*¹, Sri Widiyanesti²

Universitas Telkom, Indonesia

okaintan@gmail.com*¹, widiyanesti.sri@gmail.com²

Abstract: The rapid development of technology allows everything to be accessed by the internet that causes many users of social media and one of the social media is Twitter. An interesting topic to discuss on Twitter is about new and fresh things that attract many users to get involved. One of the things that attract Twitter users is the construction of a new airport, namely Kertajati Airport, which has some problems with airport activities, such as the small number of visitors, lonely conditions of the airport, and decreased number of routes. This study aims to find out Twitter user sentiments towards Kertajati Airport in West Java to know the quality of Kertajati Airport. The method used in this study is sentiment analysis by looking at the calculation of how many positive and negative sentiment have been obtained with the most result so it can reflect the quality of Kertajati Airport and then there is a word cloud to see the spread of word related to sentiment. The results of this study indicate that the quality of the Kertajati Airport cannot be said to be good because the results of the sentiment analysis found that negative sentiments have more percentages than positive sentiments.

Keywords: Sentiment Analysis, Quality Management

INTRODUCTION

Technological sophistication and the number of internet users in Indonesia will create a variety of data and information that arise from the ease of using the internet, namely User Generated Content. Various content created by internet users following what they want will then be shared via social media that can gather a lot of content from other users. One of the most popular social media in Indonesia is Twitter. Indonesia is also claimed to be one of the countries with the largest growth of active Twitter users daily, where the growth of Indonesia is far stronger than the global growth of daily users on Twitter (Clinten, 2019). Twitter is widely used to discuss a particular topic between other users and different from other social media, such as Instagram, Facebook, and others. This is what causes eWOM (electronic word of mouth) which is a method used by Twitter users in reviewing a product or service, both with positive opinions or negative opinions. Various kinds of data and information about opinions, comments, or things that discuss a topic can be processed and sentiment can be

found. Sentiments that are formed can be positive, negative, or neutral. Sentiment analysis is used to find out sentiments about a particular topic.

The discussion of new topics will be interesting things to discuss. One of them is the construction of a new airport in Indonesia and will be operating in 2018 is Kertajati Airport, located in West Java. After operating, Kertajati Airport experienced several problems such as a small number of visitors, lonely conditions of the airport, and target passengers that had not been achieved than previously planned. This will create opinions from many people, especially from Twitter users themselves.

Opinions on Twitter about Kertajati Airport can form electronic word of mouth, whether positive or negative. Analyzing various Twitter user opinions in the form of data and information discussing Kertajati Airport can be done with sentiment analysis that will show positive or negative results. The results of sentiment analysis will show the quality of Kertajati Airport and this may help companies, especially Kertajati Airport to improving the quality of service from the airport and become new

insights for the company. If it gets good analysis results, the company can improve its quality and if it gets bad results, companies especially Kertajati Airport can evaluate and plan innovations to improve services. This study aims to find out Twitter user sentiments towards Kertajati Airport in West Java to know the quality of Kertajati Airport.

Quality has become the hopes and desires of all people especially customer and therefore the pattern of management quality that is considered the most effective must be the most reliable competition strategy (Yamit, 2001). As stated in Goetsch & Davis (2006) quality is the dynamics associated with products, services, people, processes, and environments that meet or exceed expectations. Service quality requires more attention than the quality of a product or item because the distinguishing factor of something intangible is more difficult to identify (Han & Leong, 1996). The quality of an airport as one of the infrastructures used by many people can be included in the service quality, where the airport does not produce a product but provides services that can be felt. The service quality of Kertajati Airport can be seen using sentiment analysis which will be explained later. From this result, we can reflect the service quality of Kertajati Airport included in the category good or bad.

Sentiment analysis is a current research area in text mining, it is the stem of natural language processing or machine learning methods and the important sources of decision making and can be extracted, identified, evaluated from the online sentiment reviews (Vadivukarassi et al, 2017). The use of sentiment analysis at Kertajati Airport is expected to be able to see

Twitter users' opinions about this airport, so the results can reflect the quality of the airport and can be a new insight for the company.

Akram & Kumar (2017) defined social media is an online platform that people use to build social networks or social relations with other people who share similar personal or career interests, activities, backgrounds, or real-life connections. The availability of social media, where social media as a platform to share something, is very useful to find something about Kertajati Airport that can be seen from the opinions shared by social media users about it.

Twitter is a social networking and microblogging online service that allows users to send and receive text-based messages or posts of up to 140 characters called tweets (Jeesmitha & Com, 2019). Twitter has an API which stands for Application Programming Interface. API is the way computer programs "talk" to each other so that they can request and deliver information (Twitter, 2006). Through API, data about Kertajati Airport on Twitter can be found. Tweets related to Kertajati Airport can be taken to see the quality and sentiments of Kertajati Airport.

User-Generated Content (UGC) is an online product review content about a product or service uploaded on social media in a certain form by non-professionals or general consumers (Onny & Kusumawati, 2019). User-Generated Content (UGC) helps in searching data about Kertajati Airport, where the existence of UGC can create content related to Kertajati Airport which contains various expressions of Twitter users.

Based on the explanation above, the framework is available in Figure 1.

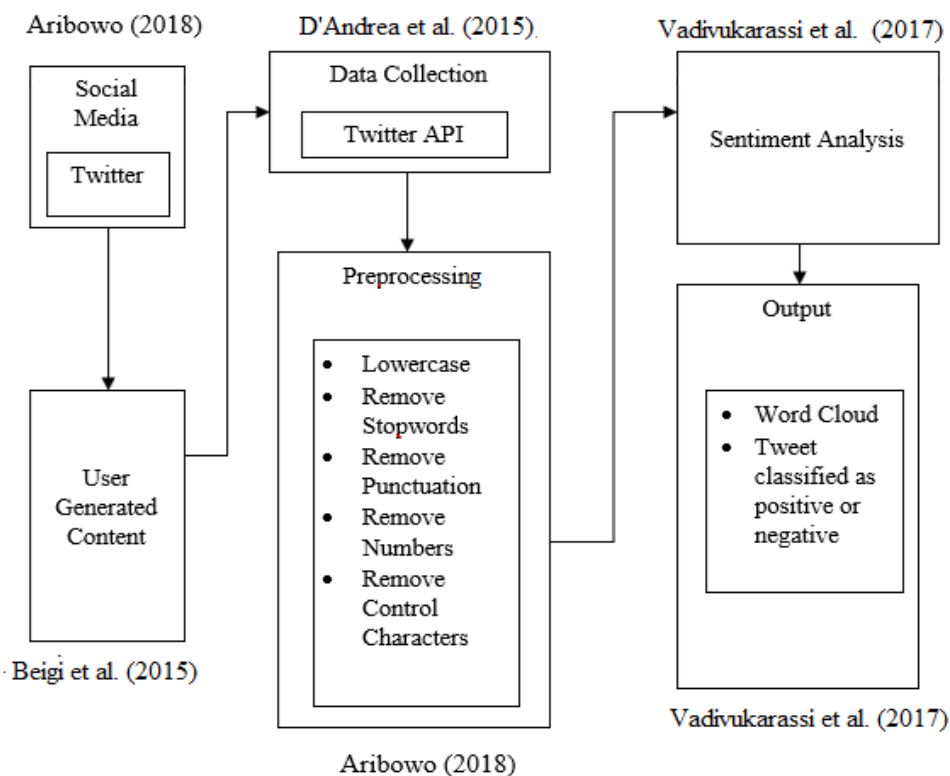


Figure 1. Framework
 Source: Aribowo (2018)

METHODS

The type of study is descriptive. A descriptive study is undertaken in order to ascertain and be able to describe the characteristics of the variables of interest in a situation (Sekaran, 2003). This study will describe the sentiment analysis of West Java International Airport.

The research methodology of this study is qualitative and in accordance with the statement from Sekaran (2003) that qualitative research involving analysis of data or information that is descriptive in nature and not readily quantifiable.

The timing of this study was included in the cross-sectional study. A cross-sectional study is a study that can be done in which data are gathered just once, perhaps over a period of days or weeks or months, in order to answer a research question (Sekaran, 2003). As in this study, the data are gathered just once over periods from September to December 2019.

The population of this study is the contents of a tweet that discusses West Java International Airport, namely Kertajati Airport. The sample of this study is the contents of a tweet about Kertajati Airport in September to December 2019 and will be processed to get cleaner data.

This data is obtained by crawling on Twitter using RStudio. After the data is obtained, the next step is to clean up the data by removing some irrelevant content in the form of retweets and URLs in data about Kertajati Airport. After that, the thing to do is to clean the data again, which is preprocessing data that can make the data cleaner. The first step of preprocessing data is lowercase. Lowercase is the process of changing tweets that have been found from capital letters to lowercase letters. After lowercase is stopwords, the process of deleting or eliminating words that have no meaning from a tweet to make it easier in the analysis process, such as conjunctions. The next steps of

preprocessing data are removing punctuation, removing numbers, and removing control characters.

After preprocessing data, the next step is the sentiments classifications. In this step, subjective sentences are classified in positive, negative, good, bad; like-dislike, but classification can be made by using multiple points (D'Andrea et al, 2015). The clean data about Kertajati Airport then classified, whether included in the positive or negative.

The sentiment analysis process requires applications that are used to process data. The application or software that will be used in processing data to find sentiment analysis is RapidMiner which can produce results from sentiment analysis, whether included in the positive or negative sentiment of tweets related to Kertajati Airport. In addition to analyzing sentiments, the output in this study is the word cloud, which in analyzing the word cloud also requires software to process data so that the word cloud can be

displayed properly. The software used to process tweets into a word cloud is to use Orange.

Based on the results of data analysis that has been found, conclusions can be drawn regarding the research conducted and this can lead to new insights for the company. It also can see how the quality of Kertajati Airport, especially regarding service quality based on the results of sentiment analysis produced.

RESULTS AND DISCUSSION

The total amount of data obtained from the process of retrieving or crawling as many as 6444 tweets using the Twitter API. The entire data then enters the preprocessing stage. The next step is processing data using RapidMiner and getting the performance value of sentiment analysis using Naive Bayes. The results of sentiment analysis performance can be seen in the table below.

Table 1. Performance Value of Naïve Bayes Classification in Kertajati Airport Data

Recall	71,73%
Precision	64,52%
Accuracy	70,70%
F-Measure	0,679
Kappa	0,318

Source: processed data (2019)

The value obtained by each component above can be said quite well with a large enough number. This shows that sentiment analysis is appropriate

and can be used with the naïve Bayes method. Then sentiment analysis and results are obtained in accordance with the table below.

Table 2. Sentiment Analysis of Kertajati Airport

Positive Sentiment	230 (22,91%)
Negative Sentiment	774 (77,09%)
Data	1004

Source: processed data (2019)

The table above shows the results from the amount of data regarding Kertajati Airport processed using RapidMiner totaling 1004, which can be divided into two sentiments. The result

of positive sentiment is 230 or 22.91% while the result of negative sentiment is 774 or 77.09%. Negative sentiment gave the biggest gains based on the table above. These results indicate that

cannot be said to be good. This result reflects that the quality that is felt by many customers or the public regarding Kertajati Airport, especially on social media Twitter can still be said to be not good enough.

Word cloud about Kertajati Airport produces words that are widely used by Twitter users, namely kertajati, bandung, bijb, lonely, far, international, flight, access, and others. There are words lonely, far, flight, and access that attract attention and are related to service quality. These results are consistent with sentiment analysis which shows more negative results, so it can be said that by looking at the word cloud, the quality of service from Kertajati Airport is still not good enough. Word cloud also reflects the shortcomings and the core of the problems felt by Kertajati Airport users, namely regarding airport activity and airport distances which are strengthened by the size of the two-word sizes seen in the word cloud compared to other words.

Suggestions are given based on the research carried out using a variety of sentiment analysis methods, such as KNN (K-nearest neighbor), SVM (support vector machine), lexicon-based and other methods that can make research more extensive in the use of available methods and search for data with more time span time, where if more data is obtained, the research conducted can be more accurate and more representative.

REFERENCES

- Akram, W., & Kumar, R. (2017). A Study on Positive and Negative Effects of Social Media on Society. *International Journal of Computer Sciences and Engineering*, Vol 5, Issues 10, 347-354.
- Aribowo, A.S. (2018). Analisis Sentimen Publik pada Program Kesehatan Masyarakat menggunakan Twitter Opinion Mining. Retrieved from <https://journal.uin.ac.id/snimed/article/view/11877>.
- Clinton, B. (2019). Pengguna Aktif Harian Twitter Indonesia Diklaim Terbanyak. Retrieved from <https://tekno.kompas.com>.
- D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, Tools, and Application for Sentiment Analysis Implementation. *International Journal of Computer Application*, Vol. 125, No. 3, 26-33.
- Goetsch, D.L., & Davis, S.B. (2006). *Quality Management: Introduction to Total Quality Management for Production, Processing, and Services, edisi ke lima*. New Jersey, United State of America: Pearson Education, Inc.
- Han, F., & Leong, D. (1996). *Productivity Service & Quality*. Singapore: Prentice Hall Simon & Schuster (Asia) Pte Ltd.
- Jeesmitha, P.S., & Com, M. (2019). The Impact of Social Media. *International Journal of Scientific Research and Engineering Development*, Vol. 2, Issue. 1, 229-235.
- Landis, J.R., & Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, Vol. 33, No. 1, 159-174.
- Onny, I.Y., & Kusumawati, A. (2019). Pengaruh User Generated Content (UGC) dan Brand Equity pada Green Purchase (Studi Pada Konsumen The Body Shop di Brazil dan Indonesia). *Jurnal Administrasi Bisnis*, Vol. 73, No. 1, 187-195.
- Sekaran, U. (2003). *Research Methods for Business: A Skill Building Approach, edisi ke empat*. New York, United State of America: John Willey & Sons, Inc.
- Twitter. (2006). Tentang API Twitter. Retrieved from <https://help.twitter.com>.
- Vadivukarassi, M., Puviarasan, N., & Aruna, P. (2017). Sentimental Analysis of Tweets Using Naïve

Bayes Algorithm. *World Applied Science Journal*, Vol. 35, No. 1, 54-59.

Yamit, Z. (2001). *Manajemen Kualitas Produk & Jasa*, edisi pertama. Yogyakarta, Indonesia: Ekonosia.