Jurnal TAM (Technology Acceptance Model) Volume 12, Nomor 1, Juli 2021 Hal. 75-79
p-ISSN : 2339-1103
e-ISSN : 2579-4221

SENTIMENT ANALYSIS OF PERFORMANCE EFFECTIVENESS OF MALIOBORO PEDESTRIAN USING SENTISTRENGTH METHOD ON TWITTER

Uning Lestari[1], Debby Anugrahni[2]
[1,2]Department of Informatics, Institut Sains & Teknologi AKPRIND Yogyakarta
[1,2]Jl. Balapan No.28, Klitren, Gondokusuman, Yogyakarta City, Daerah Istimewa Yogyakarta, Indonesia
Email: uning@akprind.ac.id[1], debbyanugrahni98@gmail.com[2]

Article history:
Received: June 11, 2021
Revised: June 22, 2021
Accepted: July 1, 2021

Keywords:
Sentiment;
Analysis;
Pedestrian;
Malioboro;
Sentistrength;
Twitter

Abstract
Sentiment analysis is a study to analyze public opinions, sentiments, evaluations, attitudes, and emotions towards services, products, public issues, organizations, general topics, etc. Sentiment analysis is a computational research of various opinions and emotions expressed textually and opinions in the form of text can be obtained through social media such as Twitter. The Malioboro area as one of the famous tourist destinations in Yogyakarta has pedestrian facilities for visitors. In the area, there are many pedestrian facilities including pedestrian paths, sidewalks, zebra crossings and parking. This study aims to measure the effectiveness of the use of the pedestrian area in Malioboro based on opinions on Twitter. This study uses the Sentistrength method. The results shows that from 3,572 Tweet data from 2016 to 2020, the results of Positive sentiment are 55.81%, the results of Neutral sentiment are 36.18% and the results of negative sentiment are 8.01%.

I. INTRODUCTION

Nowadays, social media has become a necessity for some people to support their daily activities. One of the social media continuing to grow today is Twitter. In 2019, Twitter was ranked 4th for social media users after Facebook (50.7%), Instagram (17.8%) Youtube (15.1%) and Twitter (7.1%)[1]. Along with the development of social media, Twitter is not only used to communicate but also is a medium for conveying feelings, opinions, thoughts and criticisms freely conveyed from the public. Exploring public reactions on social media is a strategic effort to obtain feedback, but it is not easy to do. Users need a long time to read thousands of tweets while sorting their sentiments, so an automatic extractive sentiment summary is needed [2].

Sentiment analysis is closely related to human life because it has an influence on habits. Opinions become the basis for various matters of personal life, public policy, and assessment of service quality in an organization. Sentiment analysis or opinion mining is a field of study that analyzes public opinions, sentiments, evaluations, attitudes, and emotions towards services, products, public issues, organizations, general topics, and so on. Sentiment analysis is a computational research of various opinions and emotions expressed textually and opinions in the form of text can be obtained from social media [3].

The process of sentiment analysis is to group the text into sentences or documents and then determine the opinions expressed in the sentences or documents analyzed whether they are positive, negative, or neutral. Sentiment analysis or opinion mining refers to a broad field of natural language processing, computational linguistics and text mining which has the aim of analyzing opinions, sentiments, evaluations, attitudes, judgments and emotions of a person whether they are concerned with a topic, product, service, organization, individual, or certain activities. [4]. Sentiment analysis is also often used in tourism research, for example to measure the level of reputation and performance of a tourism place. Tourists often upload their personal experiences or opinions about certain tourism locations on social media. This is a source of big data that can be used to measure the performance of tourism places [5].

One of the famous tourist sites in Yogyakarta is Malioboro. The Malioboro area is a center for selling Yogyakarta souvenirs and shops. The number of visitors in the Malioboro area continues to increase. So it needs to be supported by adequate public facilities. Because of this, it is related to the level of service for visitors in the Malioboro area. One of these public facilities is pedestrian facilities. Pedestrians can mean people walking in road traffic spaces [6]. This facility must meet the standards so that it can be used effectively and optimally. these facilities has not worked yet effectively and optimally, because of the factors that hinder the performance of these facilities. So it is necessary to determine the effectiveness of performance on pedestrian facilities along the area in Malioboro.

This study is conducted to measure the effectiveness of the use of the Malioboro area's facilities using data from Twitter social media

because on Twitter there are many opinions about the Malioboro area from visitors or the public having visited Maliboro. This sentiment analysis research uses the sentistrength method.

Sentiment analysis research using Twitter social media data has been carried out by several previous researchers. The written opinion is about the condition of the tourist attractions in the Sangiran Museum, Sragen used to measure the level of public opinion on the effectiveness and facilities of the Sangiran Museum. This is aimed at the Sangiran Museum's brand reputation [7]. Sentiment analysis, in addition to using Twitter opinion data, also uses a website facility built to write online opinions about a place, for example for a hotel. Sentiment analysis is one solution to overcome the problem of grouping opinions or reviews into positive or negative opinions automatically. One of the methods used is the Support Vector Machine (SVM) method [8].

In addition to the SVM method, the sent-strength method is widely used for text classification. The Sentistrength method is an algorithm as well as an opinion mining program using a dictionary/lexicon-based approach. Sentistrength dictionary/lexicon contains terms and the weight of the sentiment strength [9][10]. The Centristrength method produces a sentiment strength score and can classify sentiments into positive, neutral, and negative according to the values contained in the tweets delivered [2].

II. RESEARCH METHODS

Below is the sentiment analysis process for the effectiveness of the use of Malioboro pedestrians from Twitter opinion data which can be seen in Figure 1. This stage includes the start of the data collection and labeling process, then the data preprocessing stage, tweet classification, and evaluation of the classification results.

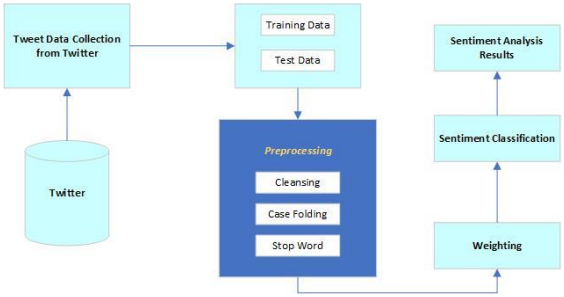

Figure 1. Research Stages

2.1 Twitter Data Collection Process

The Twitter data collection process uses several hashtags, especially complaints, praise or just opinions on the pedestrians in the Malioboro area. The keywords chosen are Pedestrian Malioboro and Sidewalk Malioboro. This data collection is done using the Command Prompt (CMD) application modified by adding Python. In this process the data will be collected in Ms.Excel format. The data collected in Ms.Excel format consists of twitter user username, date and time the tweet was uploaded, the text of the tweet, the number of retweets, favorites (the number of twitter users who liked the tweet), mentions, hashtags and URLs.

Example:

```
agungnurmajid humas_jogja;2017-12-28 14:36;0;0;"
Does that mean the west door will be raised to the same height as
the Malioboro pedestrian?
";;;;"946223363378130944";https://twitter.com/agungnurmajid/st
atus/946223363378130944

andri000me_15;2017-12-22 22:03;0;0;"09 okezonenews [FOTO]
Sellers who sell on the Pedestrian Line make the Malioboro area
look chaotic http://okz.me/CYQeN
pic.twitter.com/M7V4Oa3DsC";;;;"944161403505729536";https://t
witter.com/andri000me_15/status/944161403505729536
```

2.2 Preprocessing Data

The data cleaning stage is very important before entering the stage of providing sentiment analysis on the SentiStrength program by removing the same data, checking for inconsistent data and checking for incorrect or inappropriate data.

a. Data Cleaning
   At this stage, noise removal consists of HTML characters, emoji icons, hashtags, username (@username), mentions, URLs, email names, and so on.
b. Case Folding
   Case folding is a process of changing the contents of a tweet that originally contained uppercase letters and then changed to lowercase letters. This process uses NetBeans software.
c. Stopword Removal
   Stopword removal is the process of removing words that are not needed in the classification. Punctuation marks or hyphens that have absolutely no meaning will be removed because they have no impact on sentiment analysis.

2.3 Classification Process with SentrStrenght Method

At the classification stage, sentiment analysis is carried out using the SentiStrength method identifying a person's Tweet whether the twitter contains the value of a positive, negative, or neutral sentiment analysis. In this method, it will be weighted with words ranging from -5 to +5 depending on the type of words that are in the sentence of a data. Determining the point of each word in SentiStrength has a reference or dictionary called a lexicon or sentiment dictionary. The result of determining the weight by SentiStrength is in the form of a Ms. Excel

file with sentiment descriptions for each tweet data. This classification is carried out using the JetBrains PyCharm 2020 software by using a Query that has been set for sentiment analysis specially made to perform an Indonesian sentiment analysis.

The available dictionaries are Idiom Dictionary, Emoticon Dictionary, Sentiment Dictionary, Booster Word Dictionary, Negation Dictionary, and Question Word Dictionary. Each word in the dictionary has been assigned a weight from -5 to +5. Many dictionaries are available to determine the value weight of a word. This study uses a dictionary compiled by Devid Haryalesmana Wahid and Azhari SN [2].

III. RESULTS AND DISCUSSION

The researcher collected data from Twitter social media from January 1, 2016 to June 5, 2020 using two keywords. From the combination of these keywords, 3,572 data were obtained.

Next, the researcher carries out the preprocessing process. Data preprocessing is a cleaning stage to remove duplicate data or those that have similarities in the data. At this stage the researcher cleans the data by removing punctuation marks (period, comma, question mark, exclamation mark, and so on), and also removing the hashtag symbol (#), the @ symbol, doing stopwords, and also doing case folding. Preprocessing is done using a function found in Ms.Excel and using NetBeans IDE 8.2. The results of some Tweet data after preprocessing can be seen in Table 1.

Table 1. Tweet data after preprocessing

No	Text
1.	West door will be raised in height so that pedestrians in Malioboro
2.	Merchants selling pedestrian paths in the chaotic Malioboro area
3.	Enjoy the Pedestrian Malioboro lur
4.	It's raining in the morning, Jogja
5.	Malioboro Pedestrian Bicycle Storage Bikes can be accessed for free by visitors
6.	Evening guys, the art attractions of Malioboro WIB Pedestrian Malioboro
7.	Reading the news on the Pedestrian area on the east side of Malioboro is over
8.	Malioboro Pedestrian Line Revitalization is delayed
9.	Pedestrian face and Malioboro arrangement
10.	Blind Street, Pedestrian Street, Malioboro Street
11.	Malioboro Pedestrian, Malioboro Coffee Night

After getting the tweet data and doing preprocessing, the researcher then carries out the sentiment analysis process using the SentriStrength method. SentiStrength is a method for classifying sentiment analysis based on the value of the words contained in a sentence based on the sentiment dictionary or lexicon contained in the SentiStrength program. Based on the data obtained by the author, sentiment is divided into three categories, Positive, Negative and Neutral sentiment. The data classified sentiment can be seen in Table 2.

Public opinion on Malioboro Pedestrians is very diverse where positive, negative, and neutral sentiments become emotions that they express on social media twitter. The results of the percentage analysis of sentiment can be seen in Figure 2. Figure 2 shows that the number of tweet data held is more positive than neutral and negative sentiment. Seen in positive sentiment, there are 1,345 tweet data with a percentage of 55.81%. In neutral sentiment there are 872 tweet data with a percentage of 36.18%. The lowest is negative sentiment which has 193 tweet data with a percentage of 8.01%.

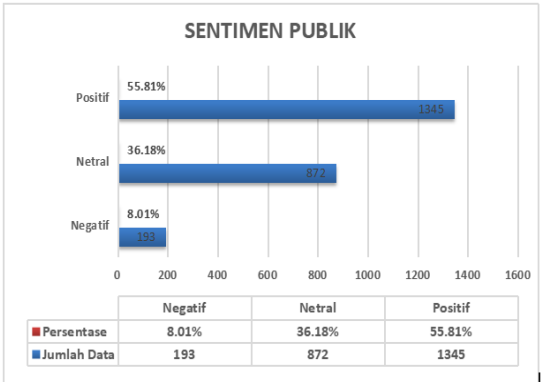

Figure 2. Result of sentiment analysis in Malioboro pedestrian

Table 2. Sentiment Analysis Results Using SentiStrength

No.	Text	Sentiment
1.	Smell [-4] As easy as the Malioboro pedestrian path appears	Negative
2.	Malioboro pedestrians	Neutral
3.	Jogja doesn't hang out on the Malioboro sidewalk, there are already many long benches, I don't feel like a gang	Neutral
4.	Beautiful Malioboro [5] with west side pavement finish	Positive
5.	Happy Face [4] Malioboro Stroller Friendly Sidewalk Expansion	Positive
6.	In the rainy season, it's good to sit on the sidewalk bench in	Positive

	Malioboro and enjoy [4] the warmth of Ronde	
7.	Suroto street, you know, from gramed sudirman kridosono, the sidewalk and garden nets were renovated, which saw chairs that resembled malioboro sidewalks placed on the middle edge	Neutral
8.	That's cool [4] street art action on the sidewalks of malioboro, jogja	Positive
9.	feel like sleeping on the sidewalk bench in malioboro	Neutral
10.	Editing Jogja, I got a chance to see a famous road [4] Jogja I hope [4] I can follow the protocol road in Jakarta to imitate the wide Malioboro sidewalks	Positive
11.	Malioboro night but I prefer [4] the sidewalks of Malioboro during the day	Positive
12.	Yasmine is a beautiful little girl [5] who is sad [-2] on the sidewalks of Malioboro hehehee	Positive
13.	The presence of Mbak Farssky's WAWA doll and the widening Malioboro sidewalk like a chasm [-3] hell [-5] for disobedient children	Negative
14.	Kopiluewak band at the Malioboro Night Festival tomorrow, Sunday, the pedestrian crossing area	Neutral

Figure 3. Visualization of positive tweet worldcloud

Figure 4. Visualization of negative tweet worldcloud

Figure 3 and Figure 4 show the results of worldcloud visualization which is a comparison between positive and negative sentiment tweet data. The image shows the words appearing frequently in the entire tweet.

Figure 3 shows the positive words most often appearing. There are the words “pedestrian”, “comfortable”, “enjoying”, “clean”, “change” and others. This shows that the public or people who have visited Malioboro are happy and grateful for the revitalization of Malioboro, so that every year it brings a good impact to Malioboro visitors with changes made by the government related to Malioboro pedestrian facilities. Visitors feel comfortable and can enjoy Malioboro which makes tourists out of town miss and want to come back with a very charming Malioboro atmosphere.

Figure 4 shows the most frequent negative words. There are the words “malioboro”, “trash”, “arrangement”, “sad”, “waste”, “damage” and others. This shows that before the Malioboro area underwent changes as it is today, visitors who have visited Malioboro are disturbed by the smell of garbage and the ineffective arrangement of street vendors that should be carried out by the relevant agencies. Visitors also feel sad about the pedestrian path and the revitalization that has been carried out for a very long time, so this disturbs the comfort of visitors.

IV. CONCLUSION

Community sentiment towards Malioboro Pedestrian has various responses. There are those who have positive, negative responses and even just joking which will be categorized as neutral sentiments. In this research, the responses that have the highest accuracy are Positive sentiments, 1,345 (55.81%), Neutral sentiments, 872 (36.18%) and the lowest is Negative sentiments, 193 (8.01%). Neutral sentiment in the SentiStrength program is considered not to contain sentiment or emotion in the text so that neutral sentiment cannot be used as a reference for readers to make a decision or policy. This shows that

the effectiveness of Malioboro pedestrian performance is quite effective, but further and in-depth studies are still needed, both through social media and discussions, to improve pedestrian public facilities in Malioboro. This is an input for evaluation materials to the relevant agencies in the city of Yogyakarta.

The development of this research needs to be done by analyzing more valid data. Vocabulary in the SentiStrength program dictionary must also be updated by following the development of languages that are relevant to their era because nowadays many languages are difficult to categorize due to a lack of vocabulary that supports and is relevant in their time to produce more accurate data.

REFERENCES

[1] APJII, "Penetrasi & Profil Perilaku Pengguna Internet Indonesia Tahun 2018", *Apjii*, bl 51, 2019.

[2] D. H. Wahid en A. SN, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity", *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol 10, no 2, bl 207, 2016.

[3] L. Yue, W. Chen, X. Li, W. Zuo, en M. Yin, "A survey of sentiment analysis in social media", *Knowl. Inf. Syst.*, 2019.

[4] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. 2015.

[5] A. R. Alaei, S. Becken, en B. Stantic, "Sentiment Analysis in Tourism: Capitalizing on Big Data", *Journal of Travel Research*. 2019.

[6] Peraturan Menteri Pekerjaan Umum Nomor : 03/PRT/M/2014, "Pedoman Perencanaan, Penyediaan, dan Pemanfaatan Prasarana dan Sarana Jaringan Pejalan Kaki di Kawasan Perkotaan", *Menteri Pekerj. Umum Republik Indones.*, vol 2013, bl 8, 2014.

[7] F. P. Nugroho en Y. D. Pambudi, "Analisa Brand Reputation Wisata Daerah Menggunakan Sentimen Data Twitter (Studi Kasus: Museum Sangiran Kabupaten Sragen)", *J. Inf.*, vol 6, no 1, bll 40–45, 2020.

[8] E. Indrayuni, "Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization", *J. Evolusi Vol. 4 Nomor 2 - 2016*, 2016.

[9] M. R. Islam en M. F. Zibran, "SentiStrength-SE: Exploiting domain specificity for improved sentiment analysis in software engineering text", *J. Syst. Softw.*, 2018.

[10] O. A. M. Ghaleb en A. S. Vijendran, "An enhancement of the public sentiment analysis on social networking by improving sentiment analysis tools", *Int. J. Intell. Eng. Syst.*, 2018.