

Data-Driven Analysis of Borobudur Ticket Sentiment Using Naïve Bayes

Dedi Kundana¹, Chairani^{2*}

Graduate Program on Informatics Engineering, Computer Science Faculty, IIB Darmajaya, Lampung, Indonesia^{1,2}

e-mail: dedi.2021210016p@mail.darmajaya.ac.id¹, chairani@darmajaya.ac.id²



Kundana, D., & Chairani. (2023). Data Driven Analysis of Borobudur Ticket Sentiment Using Naïve Bayes. *Aptisi Transactions on Technopreneurship (ATT)*, 5(2sp), 221–233.

DOI: <https://doi.org/10.34306/att.v5i2sp.353>

Author Notification
21 August 2023
Final Revised
29 August 2023
Published
04 September 2023

Abstract

The recent growth of social media is hugely influential and plays a significant role in various aspects of people's lives in the digital era. Twitter is a social media network that is widely used in Indonesia. Twitter users can engage in multiple activities, such as communicating with individuals and groups, writing daily activities, promoting businesses, arguing, and expressing ideas about a topic of discussion. At the beginning of June 2022, raising the entrance charge for Borobudur Temple became one of the concerns that caused a lot of conversation in the real world and on other social media platforms, including Twitter. The plan to increase the price of entrance tickets to Borobudur Temple has drawn various pro and con reactions in the community. **This study analyzes** public sentiment toward the planned increase in ticket prices for Borobudur Temple. Sentiment analysis of Twitter data can be implemented using a classification algorithm. The classification algorithms widely used in sentiment analysis research are Nave Bayes (NB) and Decision Tree (DT). The reason for choosing Nave Bayes and Decision Tree is because this algorithm is the most popular algorithm used to process text data classification; the process is simple, efficient, and performs well. **This study's dataset source** was taken from social media sites like Twitter. In comparison to the Decision Tree, which generates a test percentage of 100%, the accuracy of the Naive Bayes approach, based on the evaluation of the test results, produces the highest accuracy number. At the same time, the Decision Tree method's accuracy test yields a test accuracy value of 35.97%.

Keywords: Naive Bayes, Decision Tree, Borobudur Temple Ticket, Social Media Twitter

1. Introduction

The development of social media has recently been very influential and has played a significant role in various aspects of people's lives in the digital era [1] [2] [3]. The development of social media has recently been very influential and has played a significant role in various aspects of people's lives in the digital era [4] [5]. Statista Report shows 18.45 million users of the application founded by Jack Dorsey in Indonesia as of January 2022. This achievement places Indonesia as the 5th most Twitter user in the world. Through Twitter's social media, users can carry out various activities, including communication between individuals and groups, writing daily activities, promoting business, arguing, and expressing opinions related to a topic of discussion. Uploads are status messages (commonly called tweets) limited to 280



Copyright © Dedi Kundana, Chairani

This work is licensed under a [Creative Commons Attribution 4.0](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0)

characters. Tweets written by users are sources that can be used to analyze public sentiment toward the issues discussed because the writing contains sentiments that can be used as a source of evaluation and consideration in making a decision [6] [7].

Various studies have been conducted to analyze public opinion using data from social media Twitter. Research on Twitter sentiment analysis of figures public [8], film review, Performance of the House of Representatives, National BMKG Twitter data review [9], go food sentiment analysis [10], review on Word Cloud-Based Shopee E-Commerce [11], and Indonesian-language Twitter sentiment analysis of peduli lindungi app [12]. Responses written on Twitter can be said to be a sentiment; these responses can be categorized into three: responses that support an event are called positive sentiment, responses that oppose or reject an event are called negative sentiments, and responses that do not support or reject are called neutral. In this study, the emergence of reactions among the Indonesian people, especially on social media and Twitter is interesting to analyze. Sentiment analysis on Twitter data can be implemented using classification algorithms. Classification algorithms widely used in sentiment analysis research are Naïve Bayes (NB) and Decision Tree (DT).

This study aims to classify sentiment values and determine the class of tweet data from Twitter users related to the Borobudur Temple ticket price increase using the Naïve Bayes Algorithm and Decision Tree. The result will be obtained from the percentage of public sentiment towards the discourse on increasing Borobudur Temple ticket prices. The benefits of this research are expected to benefit both the author and readers regarding the sentiment picture of Twitter users towards the plan to increase Borobudur Temple tourist ticket prices. The data used in this study is the tweet data of the increase of Borobudur temples in Indonesian. This data is obtained through the process of crawling twitter data. The number of tweets used is 591, with data taken in June 2022. The contribution of this study analyzes public sentiment towards the discourse of increasing ticket prices of Borobudur Temple using the Naïve Bayes and Decision Tree methods.

2. Research Method

The stages of the research method are needed as a framework and guide for the research process so that the series of research processes can be carried out in a directed, orderly, and systematic manner. The following method proposed in this study is shown in Figure 1.

2.1 Data Collection

The first stage is the process of collecting data on Twitter or called crawling data. Crawling data on Twitter is a process of downloading data in the form of users or tweets from Twitter servers with the help of Application Programming Integration (API) (Eka Sembodo et al., 2016). In this process, the keyword used was "Borobudur ticket increase" with a total data of 591 tweets.

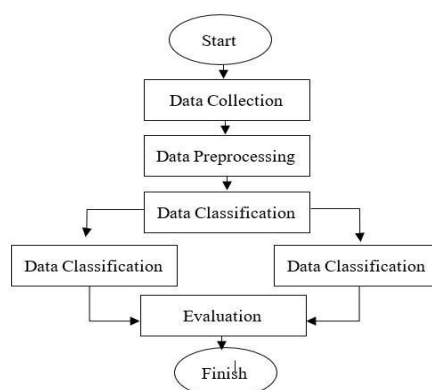


Figure 1. Stages of Research Methods

2.2 Preprocessing Data

Data preprocessing is one of the essential stages in text mining (Alasadi & Bhaya, 2017), which converts raw data into ready-to-use and more structured data with data cleaning and uniformity because the data collected is usually still dirty data. The preprocessing stages applied in this study are normalization, case folding, cleansing, tokenization, stopword removal, transform cases, and labeling. The purpose of some processes at this preprocessing stage is to produce clean tweet data so that the Naïve Bayes & Decision Tree algorithm can be tested more optimally. An illustration of the data preprocessing process is shown in Figure 2. Data Preprocessing has seven stages, including:

1. Normalization
Normalization is a process that aims to correct words that have writing or spelling errors and words that are written abbreviated.
2. Case Folding
Case folding is a technique for replacing capital letters found in data with completely lowercase.
3. Cleansing
Cleansing is a process that aims to eliminate various information that is not needed in the sentiment analysis process in the form of links (http, https, pic.twitter), hashtags, usernames (written @username), and other special characters to obtain better analysis results.
4. Tokenization
Tokenization is the process of breaking each sentence contained in data into pieces of words. The trick is to make spaces as a reference to separate each word.
5. Stopword Removal
Stopword Removal is meaningless words that matter, such as the words "in", "and", "with", "by", and etc. So, stopword removal is a process to eliminate words that do not have valuable meaning to the data to be used.
6. Transform Cases
Transform cases is the process of converting all letters into lowercase or all capital letters.
7. Labelling
Labeling is the processing of stopword removal data results, where the results are given polarity calculations from the comments taken to get a classification, namely positive, negative, and neutral labels.

2.3 Data Sharing

At this stage, the distribution of datasets with labels will be carried out to get train and test data. The dataset division uses a ratio of 80% data train and 20% test data. This is used because, based on previous research, the comparison results of 80% of train data and 20% of test data get good results. The total tweet data is 367 data. This data resulted in 294 data trains and 73 test data.

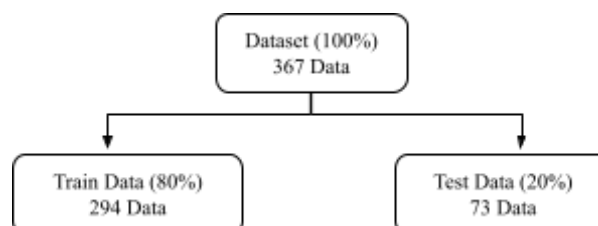


Figure 2. Data Sharing

2. 4 Naïve Bayes classification

Naive Bayes is a classification or classification of data that calculates the probability of an available dataset [13]. According to [14], Naive Bayes is a classification to predict future odds by probability and statistical methods according to previous experience. The advantage of the Naive Bayes algorithm is that the data needed to determine approximate parameters in the classification process using this method only requires a small amount of training data [15]. While according to [16]. The advantage of Naive Bayes is that it is easy to implement and in many cases, gives good results, then the disadvantage is that it is not related between features or is independent, while in reality, the relationship must exist and cannot be modeled by the Naive Bayesian Classifier. This classification is based on Bayes' theorem. Naive Bayes assumed that the effect of attribute values on a particular class was independent of the importance of other attributes. The process begins by entering the training data. The formula of the Naive Bayes algorithm is shown in the following equation (1):

(1)

Information:

c : Category or class
 doc : Document
 Nc : Number of category c's in the training document
 Ndoc : The total number of training documents used

The subsequent calculation of the probability that the word i belongs to a particular category or class can be done using the following equation (2):

(2) $P(H|X) =$

Where :

X : Sample data with unknown class (label)
 H : The hypothesis that X is data with probability class (label) C
 P(H|X) : The probability that the hypothesis is correct (valid) for the observed sample data X
 P(X|H): The probability of sample data X, when it is assumed that the hypothesis is valid
 P(H) : The probability of hypothesis H
 P(X): Observed sample data odds

2.5 Decision Tree Classification

Decision Tree is the most widely used algorithm for classification problems [17]. The Decision Tree algorithm is powerful, popular, logic-based, and easy to understand [18]. The interesting thing about Decision Tree is the use of a tree structure that serves to represent the rules formed from the results of classification [19]. Decision Tree uses a supervised machine learning method, a learning process where new data is classified based on existing training samples. The gain value is the gain information used to find variables/attributes in the dataset (S) used as root/node and branch nodes and is the attribute with the highest gain value. You can use the concepts of entropy, Gini coefficient, and misclassification to find information gain. The maximum profit value obtained from the attributes of the dataset (training data) is first used to find attributes that are worthy of the root (decision tree) of the decision tree. Then the process of searching for attributes that will be branches is repeated until we find the leaf that is the label of the class. The entropy for which the information is obtained with the entropy value is an expression for calculating the uniformity of attributes (A) from the sample data (S).

$$(3) Entropy(S) = \sum_{i=1}^n p_i \log_2 p_i$$

S = Case set in a dataset

A = Features (attributes)

n = number of partition attributes

S Pi = proportion of Si to S

So:

$$Entropy(S) = \sum_{i=1}^n p(i|s) \log_2 p(i|s)$$

(4)

Gain (S, A) is the information of attribute A in the example collection

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Entropy(S_i)$$

(5)

Entropy (S_i)

Information Gain with Gini Index Value

$$Gini(t) = 1 - \sum_{i=1}^n [p(i|s)]^2$$

(6)

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Gini(S_i)$$

Information Gain dengan classification error, value C. Error obtained from the smallest attribute value of the label class

$$C. Error(S) = 1 - \max_i [p(i|s)]$$

(7)

$$Gain(S, A) = Error(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Error(S_i)$$

(8)

A decision tree has roots/nodes, branches, and leaves like a tree.

2.6 Evaluation

The data testing results are evaluated using the Confusion Matrix table, namely accuracy. The Confusion Matrix represents the predictions and actual conditions of the algorithm-generated data. Accuracy is used for evaluation to determine the correct prediction ratio to overall data.

Table 1. Confusion Matrix

Class	Actual Positive	Actual Neutral	Actual Negative
Positive Prediction	True Positive (TP)	False Positive1 (FP1)	False Positive2 (FP2)
Negative Predictions	False Neutral1 (FNt1)	True Neutral (TNt)	False Neutral2 (FNt2)
Negative Predictions	False Negative1 (FNg1)	False Negative2 (FNg2)	True Negative (TNg)

Information:

1. True Negative / Actual Negative = actual data that is in the negative class and the model has predicted negative
2. True Positive = actual data that is in the positive class and the model has predicted positive
3. True Neutral / Actual Neutral = actual data that is in the neutral class, and the model has predicted neutral
4. False Negative = actual data that is in the positive or neutral class, but the model has predicted negative
5. False Positive / Positive Prediction = actual data that is in the negative or neutral class, but the model has predicted positive

- 6. False Neutral = actual data that is in the negative or positive class, but the model has predicted neutral

3. Findings

3.1. Data Collection

The first stage is the process of collecting data on Twitter or called crawling data. Crawling data on Twitter is a process of downloading data in the form of users or tweets from Twitter servers with the help of Application Programming Integration (API) (Eka Sembodo et al., 2016). In this process, the keyword used was "Borobudur ticket increase" with a total data of 591 tweets. The following is the process of crawling data at the data collection stage:

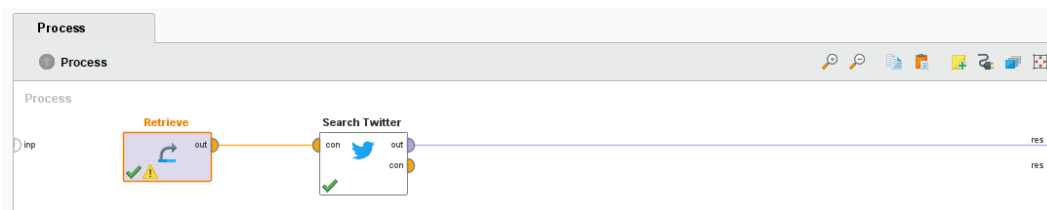


Figure 3. Twitter Data Crawling Process

3.2. Preprocessing Data

Preprocessing is one of the most important stages in text mining (Alasadi & Bhaya, 2017) which is used to convert raw data into ready-to-use and more structured data with data cleaning and uniformity because the data collected is usually still dirty data. The preprocessing stages applied in this study are cleansing, tokenization, transform cases, stopword removal, filter tokens (by length), and labeling.

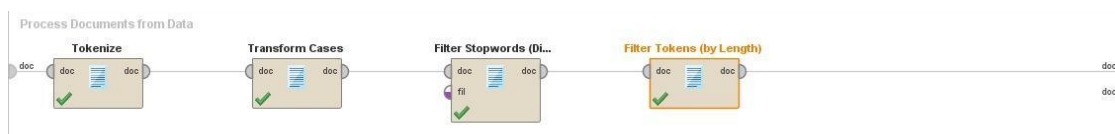


Figure 4. Data Preprocessing

3.2.1. Cleansing

Cleansing is a process that aims to eliminate various information that is not needed in the sentiment analysis process in the form of links (http, https, pic.twitter), hashtags, usernames (written @username), numbers, and other special characters to obtain better analysis results.

Table 2. Text Comparison Before and After Cleansing Process

Processing	Text
Text before cleansing	RT @Bambang_DP: After seeing this video, it seems that the entrance ticket to Borobudur temple so 750k makes sense. https://t.co/atjgGnUdCT
Text after cleansing	After seeing this video, it seems that the entrance ticket to Borobudur temple so k makes sense

Source: Research Results (2023)

3.2.2 Tokenization

Tokenization is the process of breaking each sentence contained in data into pieces of words. The trick is to make spaces as a reference to separate each word.

Table 3. Text Comparison Before and After Tokenization Process

Processing	Text
Text before <i>tokenization</i>	After seeing this video, it seems that the entrance ticket to Borobudur temple so k makes sense
Text after <i>tokenization</i>	After seeing this video, it seems that the entrance ticket to Borobudur temple so k makes sense

Source: Research Results (2023)

3.2.3 Transform Cases

Transform Cases is the process of converting all uppercase letters into lowercase letters so that they can relate to sentiment.

Table 4. Text Comparison Before and After Transforming Cases Process

Processing	Text
Text before performed <i>transform cases</i>	After seeing this video, it seems that the entrance ticket to Borobudur temple so k makes sense
Text after <i>transforming cases</i>	After seeing this video, it seems that the entrance ticket to Borobudur Temple so K makes sense

Source: Research Results (2023)

3.2.4 Stopword Removal

At this stage, the operator used is a stopwords filter (dictionary) because the dataset used is in Indonesian. In this process, data is entered into a list of words including stopwords then the file is uploaded in the stopwords filter operator (dictionary). At this stage, irrelevant words will be removed such as the word after, this, it seems, so, are words that have no meaning of their own if separated from other words and are not related to sentiment-related adjectives.

Table 5. Text Comparison Before and After Stopword Filter Process (Dictionary)

Processing	Text
Text before <i>stopword filter (dictionary)</i>	After seeing this video, it seems that the entrance ticket to Borobudur Temple so K makes sense
Text after <i>stopword filter (dictionary)</i>	Seeing the video of Borobudur Temple entrance ticket makes sense

Source: Research Results (2023)

3.2.5 Filter Token (By Length)

In this process, words that have a length of less than 4 or more than 25 will be deleted, such as the words "in", "and", "with", "by", and so on which are words that have no meaning if separated from other words and are not related to adjectives related to sentiment.

Table 6. Text Comparison Before and After Process Filter Token (By Length)

Processing	Text
Text before doing <i>filter token (by length)</i>	After seeing this video, it seems that the entrance ticket to Borobudur temple so k makes sense
Text after done <i>filter token (by length)</i>	After seeing the video, it seems that the entrance ticket to Borobudur Temple makes sense

Source: Research Results (2023)

3.2.6 Labelling

Labeling is the processing of data results after cleansing, where the results are given labeling calculations from the comments taken, so as to get a classification, namely positive, negative, and neutral labels. From 591 tweets, after preprocessing the data can be obtained 367 twitter data. The following are the labeling results of 367 twitter data.

Table 7. Labeled Tweet data

Data	Sentiment	Text
1	Positive	The announcement of Borobudur Temple ticket rates which were planned to reach thousands is now not going up, guys, even though it was canceled, the government has prepared a number of rules for visitors
2	Neutral	PUPR Minister Basuki Hadimuljono stated that the government had decided to cancel the increase in Borobudur Temple tickets
...
366	Positive	Nuri explained that the plan to increase the price of entrance tickets to Borobudur Temple was a shock to the community This ticket pricing plan is also not based on research
367	Negative	It's really arrogant statements try what competent six set the price of borobudur entrance tickets etc

Source: Research Results (2023)

3.3. Data Sharing

At this stage, the distribution of datasets that already have labels will be carried out to get train data and test data. Dataset division uses a ratio of 80% data train and 20% test data. This is used because based on previous research the comparison results of 80% data train and 20% data test get good results [20] [21] [22]. The total tweet data is 367 data. This data resulted in 294 data trains and 73 test data.

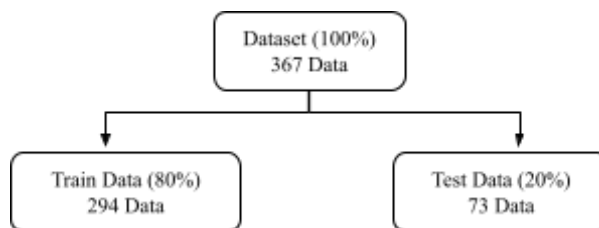


Figure 5. Data Sharing

3.4. Naive Bayes Test Results

The following stages are the design of the Naive Bayes method model testing process used, namely: :

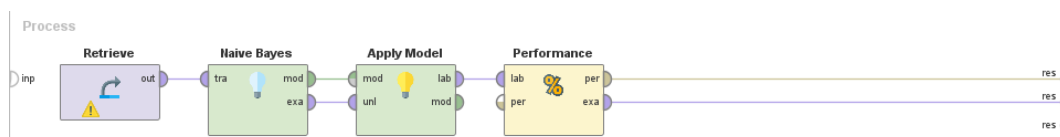


Figure 6. Naive Bayes Model Testing

Based on the test results above, the results can be seen in the following table:

Table 8. Naive Bayes Model Test Results Data

No	Sentiment	Prediction	Text
1	Positive	Positive	The announcement of the announcement of Borobudur Temple ticket fares which were planned to reach Rp750 thousand is now not going up, guys, even though it was canceled, the government has prepared a number of rules for visitors
2	Neutral	Neutral	PUPR Minister Basuki Hadimuljono stated that the government had decided to cancel the increase in Borobudur Temple tickets
....
366	Positive	Positive	Nuri explained that the plan to increase the price of entrance tickets to Borobudur Temple was a shock to the community This ticket pricing plan is also not based on research
367	Negative	Negative	What a proud statementx try sya there dsna of course sya will tnya kompetensix set the price of borobudur entrance tickets etc

3.5 Decision Tree Test Results

Stages of classification The following is the design of the Decision Tree method model testing process used, namely:

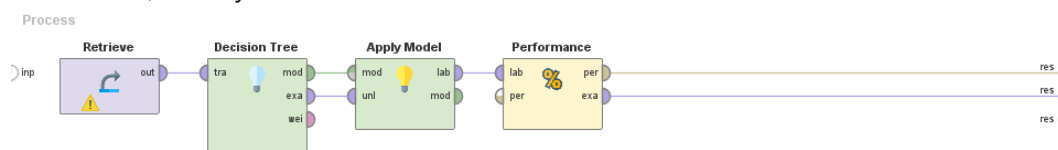


Figure 7. Decision Tree Model Testing

Based on the test results above, the results can be seen in the following table:

Table 9. Decision Tree Model Test Results Data

No	Sentiment	Prediction	Text
1	Positive	Positive	The announcement of Borobudur Temple ticket fares which were planned to reach Rp750 thousand, is now not going up, guys; even though it was canceled, the government has prepared a number of rules for visitors.
2	Neutral	Positive	PUPR Minister Basuki Hadimuljono stated that the government had decided to cancel the increase in Borobudur Temple tickets.
....
366	Positive	Positive	Nuri explained that the plan to increase the price of entrance tickets to Borobudur Temple was a shock to the community. This ticket pricing plan is also not based on research.
367	Negative	Positive	It's really arrogant to say, try me to have funds, of course I will ask what is the competence to set the price of Borobudur entrance tickets, etc.

Source: Research Results (2023)

3.6. Evaluation

Evaluation of the data testing results using the Confusion Matrix table, namely accuracy. The Confusion Matrix represents the predictions and actual conditions of the data generated by the algorithm used. Accuracy is used for evaluation to know the ratio of correct predictions to overall data.

Table 10. Confusion Matrix Naive Bayes

Accuracy: 100.00%				
	True Positive	True Neutral	True Negative	Class Precision
Pred. Positif	132	0	0	100.00%
Pred. Neutral	0	123	0	100.00%
Pred. Negative	0	0	112	100.00%
Class Recall	100.00%	100.00%	100.00%	

Source: Research Results (2023)

Based on the table with confusion matrix above, using the naïve Bayes method to check tweet data, the results obtained are an accuracy rate of 100% or 367 of the correct words and 0% error, which is as many as 0 words or commonly called there is also no error in detection. The positive class data that corresponds to the optimistic prediction is 132 data. The neutral class data that matches the neutral prediction is 123 data. The negative class data that corresponds to the negative prediction is 112 data.

Table 11. Confusion Matrix Decision Tree

Accuracy: 35.97%				
	True Positive	True Neutral	True Negative	Class Precision
Pred. Positif	132	123	112	35.97%
Pred. Neutral	0	0	0	0.00%
Pred. Negative	0	0	0	0.00%
Class Recall	100.00%	0.00%	0.00%	

Source: Research Results (2023)

Based on the table with the confusion matrix above, using the decision tree method to check the tweet data, the results obtained are an accuracy rate of 35.97% or 132 of 367 correct words and 64.03% of errors, which is as many as 235 of 367 words in its detection. The positive class data that corresponds to the positive prediction is 132 data. The neutral class data predicted positive is 123 data. The negative class data predicted into positive is 112 data.

4. Conclusion

Based on the analysis conducted in this study, it can be concluded that this study classifies sentiment on tweets with the Text Mining process using the Naïve Bayes and Decision Tree methods. Analysis of the calculation results of the Naive Bayes and Decision Tree methods was carried out based on data sources obtained from the Twitter crawling API, with 591 data then pre-processed into 367 data, 80% (294 data) used as training data, and 20% (73 data) as testing data. The accuracy results of the Naive Bayes method based on the evaluation of test results produce the highest accuracy value compared to the Decision Tree, which produces a test percentage of 100%. At the same time, the accuracy results of the Decision Tree method resulted in a test accuracy value of

35.97%.

References

- [1] A. S. Harahap M A, "Tren Penggunaan Media Sosial Selama Pandemi Di Indonesia," 2020.
- [2] J. Hendra and N. Laugu, "Eksistensi Media Sosial, Youtube, Instagram dan Whatsapp Ditengah Pandemi Covid-19 Dikalangan Masyarakat Virtual Indonesia," *Baitul Ulum J. Ilmu Perpust. dan Inf.*, vol. 4, no. 1, 2020, [Online]. Available: <https://databooks.com>.
- [3] D. A. Rifandi and I. Irwansyah, "Perkembangan Media Sosial pada Humas Digital dalam Industri 4.0," *J. SIMBOLIKA Res. Learn. Commun. Study*, vol. 7, no. 2, pp. 141–151, Oct. 2021, doi: 10.31289/simbollika.v7i2.5376.
- [4] R. S. Rohman, R. A. Saputra, and D. A. Firmansaha, "Komparasi Algoritma C4.5 Berbasis PSO Dan GA Untuk Diagnosa Penyakit Stroke," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 1, p. 155, 2020, doi: 10.24114/cess.v5i1.15225.
- [5] F. Solihin, S. Awaliyah, A. Muid, and A. Shofa, "Pemanfaatan Twitter Sebagai Media Penyebaran Informasi Oleh Dinas Komunikasi dan Informatika," *J. Pendidik. Ilmu Pengetah. Sos.*, vol. 1, no. 13, pp. 52–58, 2021, [Online]. Available: <http://e-journal.upr.ac.id/index.php/JP-IPS>.
- [6] I. W. Sudiartana, W. G. S. Parwita, and ..., "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Sosial Media Twitter," *SmartAI Bul. Artif. ...*, 2022, [Online]. Available: <https://ejournal.abivasi.id/index.php/SmartAI/article/view/2>.
- [7] M. Alfi, R. Reynaldhi, and Y. Sibaroni, "Analisis Sentimen Review Film pada Twitter menggunakan Metode Klasifikasi Hybrid SVM, Naive Bayes, dan Decision Tree," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 10127–10137, 2021.
- [8] T. T. Widowati et al., "Analisis Sentimen Twitter Terhadap Tokoh Publik Dengan Algoritma Naive Bayes Dan Support Vector," *J. SIMETRIS*, vol. 11, no. 2, 2020.
- [9] D. D. Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (Dpr) Pada Twitter Menggunakan Metode Naive Bayes Classifier," vol. 10, no. 1, pp. 34–40, 2022.
- [10] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," vol. 15, no. 1, pp. 131–145.
- [11] M. I. Petiwi, A. Triayudi, and ..., "Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naive Bayes dan Support Vector Machine," *J. Media ...*, 2022, [Online]. Available: <http://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/3530>.
- [12] J. J. A. Limbong, I. Sembiring, K. D. Hartomo, U. Kristen, S. Wacana, and P. Korespondensi, "Analisis Klasifikasi Sentimen Ulasan Pada E-Commerce Shopee Berbasis Word Cloud Dengan Metode Naive Bayes Dan K-Nearest Analysis of Review Sentiment Classification on E-Commerce Shopee Word Cloud Based With Naive Bayes and K-Nearest Neighbor Methods," *J. Teknol. Inform. dan Ilmu Komput.*, vol. 9, no. 2, pp. 347–356, 2022, doi: 10.25126/jtiik.202294960.
- [13] R. Akmalia, I. Slamet, and H. Pratiwi, "Analisis Sentimen Twitter Berbahasa Indonesia Terhadap Aplikasi PeduliLindungi dengan Algoritma SVM, KNN, dan Regresi Logistik."
- [14] Krismayani, I., Jumino, J., & Wasisto, J. (2021). Selected Webliographies on Indonesian Cultural Heritages. In *E3S Web of Conferences* (Vol. 317, p. 05008). EDP Sciences.
- [15] Aung, S. M. T., & Lwin, N. (2022). 2022/74 "What do the Official Chinese Media's Mixed Messages on the Myanmar Coup Mean?" by Su Mon Thazin Aung and Nan Lwin..
- [16] Nugroho, Y., & Adrianto, D. W. (2022). 2022/69 "The Nusantara Project: Prospects and Challenges" by Yanuar Nugroho and Dimas Wisnu Adrianto.
- [17] Prasetyo, L., Artawa, K., Mulyawan, I. W., Purnawati, K. W., & Suardiana, I. W. (2023). COVID-19 PREVENTION SIGNS IN THE INDONESIA'S SUPER PRIORITY TOURIST DESTINATION. *Sabbhata Yatra: Jurnal Pariwisata dan Budaya*, 4(1), 45-55.
- [18] Triyono, J., & Nugroho, D. Y. (2022). Pengaruh Wacana Kenaikan Tarif Candi Borobudur Terhadap Minat Wisatawan dan Jumlah Pemesanan Jasa Travel. *Jurnal Manajemen Perhotelan dan Pariwisata*, 5(3), 248-253.

-
- [19] Rifki, M., & Imelda, I. (2022). ANALISIS SENTIMEN WACANA KENAIKAN HARGA TIKET CANDI BOROBUDUR MENGGUNAKAN MULTINOMIAL NAÏVE BAYES. *JIKO (Jurnal Informatika dan Komputer)*, 5(2), 156-163.
- [20] Heslinga, J. H., Yusuf, M., Damanik, J., & Stokman, M. (2023). Future strategies for tourism destination management: post COVID-19 lessons observed from Borobudur, Indonesia. *Journal of Tourism Futures*.
- [21] D. Juhaeni and A. Wibowo, "Penerapan Metode Naïve Bayes Untuk Wacana Kenaikan Harga Tiket Candi Borobudur Pada Twitter," 2022.
- [22] M. Rifki and I. Imelda, "Analisis Sentimen Wacana Kenaikan Harga Tiket Candi Borobudur Menggunakan Multinomial Naïve Bayes," *J. Inform. dan Komputer (Akreditasi KEMENRISTEKDIKTI, vol. 5, no. 2, 2022, doi: 10.33387/jiko*.