COMPILER
Vol. 11, No. 2, November 2022, pp. 83~90
ISSN: 2252-2403 (Print), ISSN: 2549-2403 (Online)
Accredited Fourth Grade by Kemenristekdikti, Decree No:200/M/KPT/2020
DOI: 10.28989/compiler.v11i2.1286

Classification of Twitter User Sentiment Against Government Policy in Overcoming Covid-19 in Indonesia

Hermawan Arief Putranto*, Taufiq Rizaldi, Wahyu Kurnia Dewanto, Rokhimatus Zahro
Information Technology Department, State Polytechnic of Jember, Jember, Indonesia

Article Info

Article history:

Received June 30, 2022
Accepted December 30, 2022
Published December 31, 2022

Keywords:

Sentiment Classification
Covid-19
Twitter
Naïve Bayes
Pre-Processing

ABSTRACT

Sentiment classification is a field of study that analyzes a person's opinions, sentiments, judgments, evaluations, attitudes, and emotions regarding a particular topic, service, product, individual, organization, or activity. The topic that is currently being discussed is Covid-19. Covid19 is a disease caused by a coronavirus, first identified in Wuhan City, China. This disease has spread throughout the world, one of which is Indonesia. Related to this, the Government of Indonesia issued a policy in an effort to break the chain of the spread of the coronavirus. However, this prompted the emergence of various kinds of community responses. One of them is Twitter users, there are pros and cons responses from the community in addressing government policies and causing problems, namely the difficulty of knowing positive, neutral, or negative responses given by the public. Based on the explanation above, sentiment analysis is carried out. This analysis was carried out by utilizing data from Twitter with the keywords *dirumahaja*, *vaksinuntukrakyatindonesia*, psbb, covid, covid19, covidindonesia, *vaksinjakarta*, vaksin, *vaksinPulihkanRI*, and *vaksinDemiLindungiNKRI*. Where the data will be processed through several stages, namely preprocessing, word weighting, and sentiment analysis. The results of the sentiment classification of the majority of Twitter users' responses are neutral, which are 69.2% of the data is classified as neutral sentiment, 30.1% of the data is classified as positive sentiment, and 7% of the data has negative sentiment.

Corresponding Author:

Hermawan Arief Putranto,
Information Technology Department,
State Polytechnic of Jember,
Jl. Mastrip Po Box 164, Jember, East Java, Indonesia
Email: hermawan_arief@polije.ac.id

1. PENDAHULUAN

Twitter is a social media that allows its users to read and send text messages of up to 140 characters, commonly called tweets. Many Twitter users are spread all over the world, including Indonesia. Frequently, people use Twitter to express their opinions or opinions on a phenomenon or event that is happening. Even though at this time the world can be said to be at peace with Coronavirus Disease 2019 or commonly called Covid-19, two years ago, Covid-19 often became a trending topic in the world, including Indonesia. Because of this, Twitter has become a source of raw data, for researchers in the fields of data mining and machine learning, on topics related to Covid-19.

Covid-19 is a disease caused by the coronavirus or Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov-2), first identified in Wuhan City, China [1]. This disease has spread widely throughout the world, including in Indonesia. Related to this, the Government of Indonesia issued a policy as a step to tackle the coronavirus. Among them is an appeal to wear a mask while in public places, keep your distance when in public places, always wash your hands and use hand sanitizers as well as advice to all residents to stay at home, as an effort to break the chain of the spread of the corona. This policy certainly received various reactions from the public, including pros, cons, or neutral. One very easy way to see people's reactions is via Twitter.

However, as a data source, Twitter has its limitations [2]. The raw data obtained is in the form of text data whose structure cannot be ascertained. In addition, the language used is also not a standard language, many uses borrowed words, regional languages, and slang. This certainly affects the results of the analysis, if the data used is too diverse.

In this study, an approach is proposed to deal with problems experienced by raw data, especially text data, namely by carrying out text pre-processing. Furthermore, the data that has been handled will be grouped into sentiment classes, so that the public's response to certain issues related to government policies in handling the Covid-19 case can be known.

This analysis was carried out by utilizing data from Twitter with the keywords vaksinuntukrakyatindonesia, psbb, covid, covid19, *covidindonesia, vaksinjakarta,* vaksin, *vaksinPulihkanRI*, and v*aksinDemiLindungiNKRI*. Where the data obtained will be processed through several stages of text preprocessing consisting of data cleaning, tokenizing, filtering, normalization, and stemming [3]. Next, word weighting is carried out, to map each word in the tweet into a word-vector form. The classification method used in the analysis process is the Naïve Bayes Classifier method. For classification into multiple classes, the Naïve Bayes Classifier has a better level of accuracy than other classifier models [4]. Finally, the system accuracy testing process is carried out using the n-fold cross-validation method.

The use of Twitter microblogging as a source of research data has been widely carried out, including research by Winda Estu Nurjanah, Rizal Setya Perdana, and Mochammad Ali Fauzi with the title " *Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet*" [5]. Television show viewers often give opinions on television shows through social media, one of which is Twitter. Public opinion on Twitter towards television shows has an important role because it can be used to carry out sentiment analysis in predicting people's evaluation of a television program, whether it is positive or negative. This study uses the K-Nearest Neighbor Algorithm. From the results of accuracy testing using textual weighting, it was obtained 82.50%, using non-textual weighting 60%, and using a combination of the two 83.33% with the value k = 3 and the exact multiplication constant $\alpha = 0.8$ and $\beta = 0$

Next is Rizka Ardiansyah's research with the title *"Analisis Sentimen Calon Presiden dan Wakil Presiden Periode 2019-2024 Pasca Debat Pilpres di Twitter*" [6]. Twitter and Facebook have become common elements in political campaigns and elections, especially for the 2019 Indonesian presidential election. For 6 periods 2019-2024, there are two presidential and vice-presidential candidates namely Ir. H. Joko Widodo - Prof. Dr. K.H. Ma'ruf Amin and Lieutenant General (Ret.) H. Prabowo Subianto – H. Sandiaga Uno. B.B.A., M.B.A. the two candidates running for election sparked a lot of public opinion regarding which candidate is the most suitable to become the next presidential term. The purpose of this study is to summarize public opinion voiced through social media regarding the election of candidates for President and Vice President of Indonesia for the 2019-2024 period after the debate on the presidential election. This study uses text mining and machine learning methods to retrieve and classify the polarity of opinions from data sources. The algorithm used is the Naïve Bayes Classifier. The results of this study are the presidential candidate Joko Widodo - Ma'ruf Amin obtained positive sentiment of 25%, a negative sentiment of 4.5%, and neutral sentiment of 70.5%. Meanwhile, the Prabowo Subianto Sandiaga Uno pair received 5.1% positive sentiment, 2.5% negative sentiment, and 92.4% neutral sentiment

.

2. METHODOLOGY

As shown in Figure 1, the approach used in this research is divided into four main stages, namely data collection, text pre-processing, word weighting and sentiment classification.

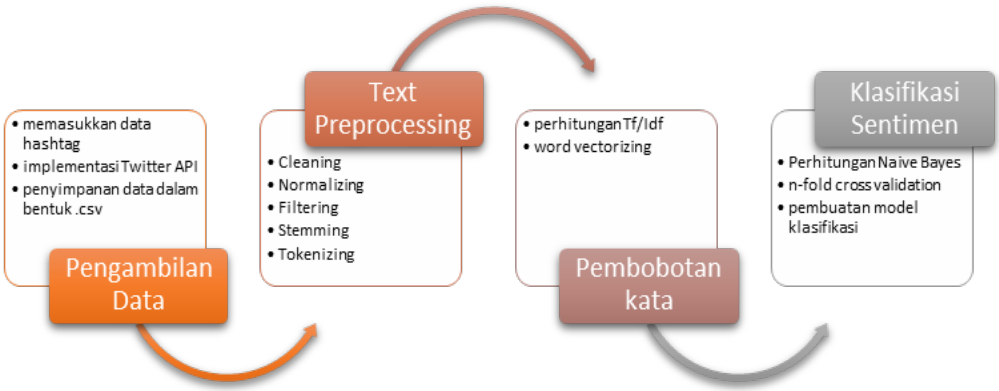

Figure 1. Research Methodology

2.1 Data Extraction

The process of collecting Twitter data is done using a crawling technique. Crawling is a technique of collecting data on a website by entering a Uniform Resource Locator (URL). The implementation is done using the python language by utilizing the twepy library and the Twitter API. To be able to retrieve Twitter data, we must first have a Customer Key and Access Token. This Customer Key and Access Token is obtained by creating a new application at http://developer.twitter.com/.

After that is the program coding process using Jupyter Notebook tools. The data taken is only Indonesian-language tweets, using the hashtags at home, vaccines for the people of Indonesia, psbb, covid, covid19, covidindonesia, vaccinesjakarta, vaccines, vaccines Recover RI, and vaccines for the sake of protecting the Republic of Indonesia. Of the 1000 tweets that were successfully obtained, only 271 data were used for making mathematical models. This number is data that has undergone cleaning and normalization, and is considered to best represent the class to be tested, namely neutral, positive and negative. While the source of the data is taken randomly, both from ordinary users and online media. The crawled tweet data will later be saved in a file with the .csv format and named data_twitter.csv.

2.2 Text Preprocessing

The data that has been obtained must go through a text pre-processing first so that the data is prepared and ready to be analyzed. This process uses several Python libraries, so first, we must import the required libraries. The library consists of the pandas library which generally functions to create tables, change data dimensions, check data, and so on. Regular expression library (regex), which is a string of characters used to search strings or text using patterns. The Library Natural Language Toolkit (NLTK) is a python library for working with text modeling. And the Sastrawi library is a library that is used specifically for the Indonesian language stemming process. After that, the tweet data resulting from the crawling process in the form of a .csv file is uploaded to the Dataframe using pandas.

2.2.1 Cleaning

Cleaning data is the process of removing noise in the data so that it does not interfere with data processing. Included in this cleaning process are deleting emoticons, and punctuation marks, changing capital letters to lowercase letters, and removing empty characters, mentions, URLs, hashtags, and numbers. As well as removing tabs, new lines, and back slices. In this cleaning process, in general, we use the re. sub-function to replace one or more matches with the selected string.

2.2.2 Normalizing

Normalizing is the process of uniform terms that have the same meaning but are written differently due to writing errors, abbreviations, or "slang". Normalization data is created manually and saved in the normalization.xlsx file.

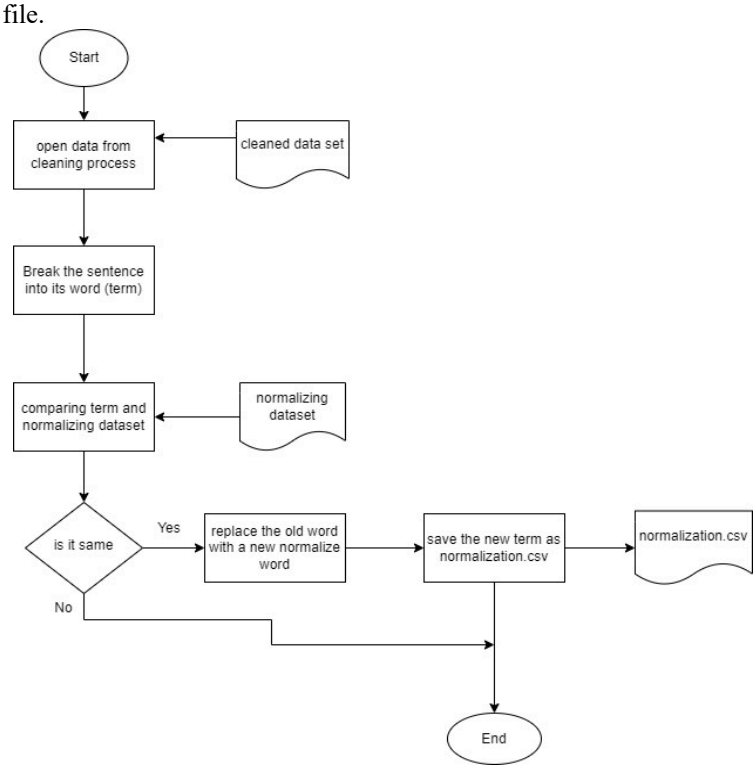

Figure 2. Term normalizing flowchart

In general, the flow of the normalization process is to check the terms or words in the tweet data, whether the word is the same as the word in the normalization data dictionary or not. If yes, then the word will be replaced according to the word contained in the normalized data dictionary and if not, then the word remains unchanged.

2.2.3 Filtering

Filtering is eliminating words that appear in large numbers, but are considered to have no meaning (stop word) [7]. Here the filtering process is carried out twice, the first using the nltk library and the second using a custom literary library. The custom stop words list is stored in the stopwords.txt file. This stop word list can be customized according to our needs or desires.

Broadly speaking, the flow of the filtering process is carried out by creating a variable to accommodate the list of stop words originating from Python and the list of custom stop words, after which the term or word in the tweet data is checked whether the word is the same as the data contained in the stop word list variable or not. If yes, then the word will be deleted and if not, then the word will remain.

2.2.4 Stemming

Stemming is the process of changing words into basic words or removing affixes. To do Indonesian stemming, we can use the literary library. This library is a special library for the Indonesian language stemming process. This stemming process is carried out by implementing filtered data (custom) in the stem() function contained in the literary library.

2.2.5 Tokenizing

Tokenizing is a process carried out to cut/break sentences into several parts/words called tokens. A sentence or data can be split into words (tokens) with the word_tokenize() class in the nltk library [8] [9]. A sentence or data can be split into words (tokens) with the word_tokenize() class in the nltk library. The tokenizing process is carried out by implementing the stemming result data in the word_tokenize() function.

2.3 Term Weighting

The tweet data that has gone through the preprocessing process is then calculated for the term or word weight value. Here the word weighting is calculated using the TF-IDF (Term Frequency and Inverse Document Frequency) method. However, before weighting the words is done, the tweet data from the preprocessing results must be manually labeled and the polarity determined. Polarity 1 means positive sentiment, polarity 0 means neutral sentiment, and polarity -1 means negative sentiment.

After the process of determining the polarity is complete, the next process is to weigh the words or terms. Word weighting is calculated using the CountVectorizer() function found in the sklearn library. This function is used to convert text features or word data into a vector or numeric representation, so that said data can be analyzed.

Also at this stage, the preprocessing result data is calculated by the frequency of occurrence of words in each document. Furthermore, the number 1 (Laplace smoothing) is added, namely to remove the number 0 so as not to damage the calculation process being carried out. Calculate the total class on the term and calculate the total term as a whole. After that, calculate the weight of the word or term in each class. So that the weight of the term or word is produced [10][11].

2.4 Sentiment Classification

At the sentiment classification stage, a training process is carried out on the classification model to be made, using tweet data that has previously gone through the process of weighing words and determining polarity. The tweet data was taken from February 20 2021 to April 28, 2021, because that's when the research began. The data is split or divided into two parts, namely 80% training data or training data and 20% test data or testing data. Furthermore, the training process for the model uses training data and the Multinomial Naïve Bayes algorithm as the classification method. The implementation of this algorithm uses the MultinomialNB() function contained in the sklearn library.

After the model training process is complete, predictions are made on the testing data to find out whether the performance of the model is good or not. In this testing process, it is hoped that the model can classify the sentiment of the testing data correctly according to the contents of the tweet data.

3. RESULT AND DISCUSSION

From the test results using test data, a confusion matrix was obtained which can be seen in Figure 3.

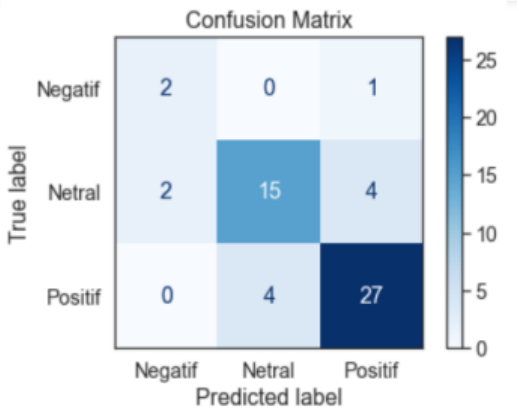

Figure 3. Confusion Matrix

The figure shows that for negative data that are classified as truly negative are 2 tweet data, neutral data that are correctly classified as truly neutral are as many as 15 tweet data, and positive data that are classified as true positive are as many as 27 tweet data. From the confusion matrix, Precision, Recall, F1 Score, and Accuracy values can be obtained, as shown in Figure 4.

```
Report dari hasil prediksi data testing
              precision    recall  f1-score   support

          -1       0.50      0.67      0.57         3
           0       0.79      0.71      0.75        21
           1       0.84      0.87      0.86        31

    accuracy                           0.80        55
   macro avg       0.71      0.75      0.73        55
weighted avg       0.80      0.80      0.80        55

    Accuracy : 80.00%

    F1 Score : 80.06%

    Confusion Matrix:
     [[ 2  0  1]
     [ 2 15  4]
     [ 0  4 27]]
```

Figure 4. Predictive results of test data

After that, the accuracy of the system was tested using the k-fold method. K-fold is one of the popular Cross Validation methods by folding k data and repeating (iteration) the experiment as much as k as well [12]. The dataset will later be divided into 10 parts or k = 10, so that 271 data is divided into 10 folds where the ratio of training data and testing data is 80% : 20%. So, from the 271 data used, the division is obtained as follows, 216 training data and 55 testing data. Testing using partitioned data will be repeated 10 times with a different testing data position for each iteration.

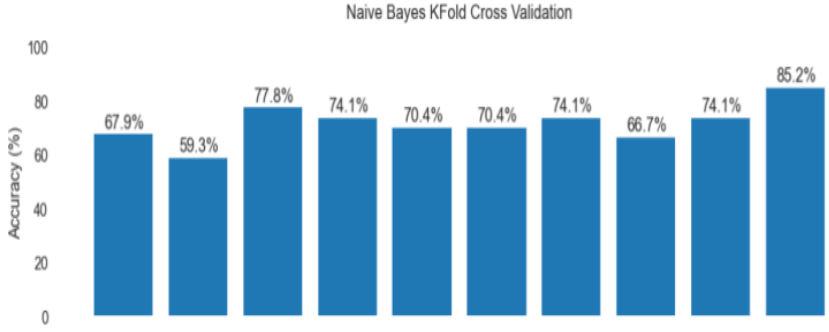

Figure 5. k-fold Cross validation result.

In Figure 5, it can be seen that the highest accuracy obtained was the 10th fold, which was 85.2%, followed by the 3rd fold, which was 77.8%, the 4th, 7th, and 9th folds had the same value, which was 74.1%. the 5th and 6th folds have the same value, which is 70.4%, the 1st fold is 67.9%, the 8th fold is 66.7%, and the lowest is the 2nd fold, which is 59.3%. As well as obtained an average accuracy of all folds that is equal to 71.97%.

As for the F1 score, it can be seen in Figure 6, the highest F1 score was obtained at the 9th fold of 85.19%, followed by the 3rd fold of 78.22%, the 8th fold of 77.22%, the 2nd fold of 76.26%, the 4th fold is 75.89%, the 6th fold is 73.43%, the 5th fold is 68.58%, the 7th fold is 66.75%, the 65.37% fold, F1 The lowest score is obtained in the 2nd fold of 56.24%. And the average F1 score for all folds is 72.31%.

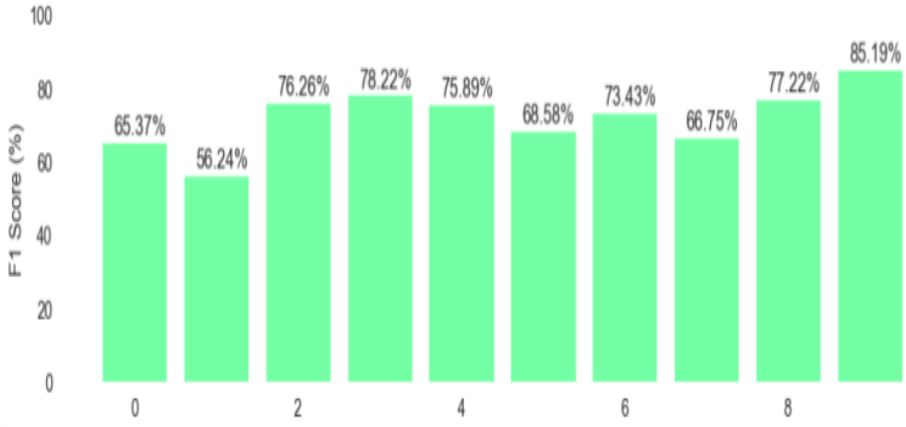

Figure 6. F1 Score calculation result

The expected result in the implementation of this model is that the model can classify the testing data sentiment correctly according to the contents of the tweet data.

4. CONCLUSION

Based on the analysis and testing that has been done, the conclusion that can be drawn is that sentiment analysis is carried out in several stages, namely crawling or data collection, preprocessing, word weighting, sentiment analysis, and data visualization. The implementation of the Naive Bayes Classifier method in this sentiment analysis is to use the MultinomialNB() function contained in the sklearn library. In this study, 271 data were used to construct the classification model, then the model was tested again with 1000 data. From the application of this model, it was found that the response of the majority of Twitter users was neutral, namely 692 data classified as neutral sentiment, 301 data classified as positive sentiment, and 7 data classified as negative sentiment.

This sentiment analysis application can run well, according to its function. This is indicated by the results of system testing which states that all features have been tested and display results that are as expected. The classification process is more accurate if a large number of training data or training data are used. This is shown from the precision and recall results of positive sentiments which have high values, namely 0.84 and 0.87, considering that positive sentiments have the most training data compared to neutral and negative sentiments. The implementation of the Naive Bayes algorithm in this study resulted in an accuracy value of 80%, as well as a precision of 84% positive sentiment, 50% precision of negative sentiment and 79% precision of neutral sentiment.

REFERENCES

[1] F. Alamsyah, "Covid-19: penyebab, penyebaran dan pencegahannya," *Indones. Sch. Netw.*, pp. 5–9, 2020, [Online]. Available: https://eprints.uai.ac.id/1711/.

[2] D. Rustiana and N. Rahayu, "Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 1, pp. 113–120, 2017, doi: 10.24176/simet.v8i1.841.

[3] H. A. Putranto, O. Setyawati, and W. Wijono, "Pengaruh Phrase Detection dengan POS-Tagger terhadap Akurasi Klasifikasi Sentimen menggunakan SVM," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 5, no. 4, 2016, doi: 10.22146/jnteti.v5i4.271.

[4] D. Xhemali, C. J. Hinde, and R. G. Stone, "Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages," *Int. J. Comput. Sci.*, vol. 4, no. 1, pp. 16–23, 2009, [Online]. Available: http://cogprints.org/6708/.

[5] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah

Retweet," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 12, pp. 1750–1757, 2017.

[6] R. Ardiansyah, "Analisis Sentimen Calon Presiden Dan Wakil Presiden Periode 2019-2024 Pasca Debat Pilpres Di Twitter," *Sci. Comput. Sci. Informatics J.*, vol. 2, no. 1, p. 21, 2019, doi: 10.22487/j26204118.2019.v2.i1.13068.

[7] A. P. Natasuwarna, "Tantangan Menghadapi Era Revolusi 4.0 - Big Data dan Data Mining," in *Seminar Nasional Hasil Pengabdian Kepada Masyarakat*, 2019, pp. 23–27.

[8] S. A. Qutsiah, M. K. Sophan, and Y. F. Hendrawan, "Aplikasi Pembelajaran Matematika Dasar Bangun Datar Menggunakan Python Pada Perangkat Bergerak," *SCAN - J. Teknol. Inf. dan Komun.*, vol. 11, no. 3, pp. 13–22, 2016, [Online]. Available: http://www.ejournal.upnjatim.ac.id/index.php/scan/article/view/868.

[9] R. Mudhar, "Pembangunan Sistem Informasi Helpdesk Ticketing System Menggunakan Django Framework (Studi Kasus: SMK Saradan)," UIN Syarif Hidayatullah, 2016.

[10] A. N. Rahimah, D. S. Rusdianto, and M. T. Ananta, "Pengembangan Sistem Pengelolaan Ruang Baca Berbasis Web Dengan Menggunakan Django Framework (Studi Kasus : Ruang Baca Fakultas Ilmu Komputer Universitas Brawijaya)," *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 5, pp. 4439–4446, 2019.

[11] A. A. S. Daironi and W. Yustanti, "Perbandingan Penggunaan NoSQL Mongodb Dan Mysql Pada Basis Data Forum Komunikasi," *Manaj. Inform.*, vol. 6, no. 1, pp. 134–142, 2016.

[12] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.