Analysis Of Public Sentiment About Covid-19 In Indonesia On Twitter Using Multinomial Naive Bayes And Support Vector Machine

F J Damanik 1, and D B Setyohadi 2,

1 MagisterInformatika Universitas Atma Jaya Yogyakarta Yogyakarta, Indonesia, 55281
205303284@students.uajy.ac.id
2 MagisterInformatika Universitas Atma Jaya Yogyakarta Yogyakarta, Indonesia, 55281
djoko.budiyanto@uajy.ac.id

Email: doemanik@gmail.com

Abstract. The Covid-19 case which initially occurred in China until now has spread to all continents, including Indonesia. Indonesia itself was first affected by the Covid-19 virus in February 2020 which made a crowd in cyberspace including in particular Twitter. Everyone is free to give their opinion or opinions about Covid-19 so that it raises many opinions, not only positive or neutral opinions but also negative opinions regarding Covid-19. Social media is now not only used as a means of making friends or making friends, but also used for other activities. The purpose of this research is to build a sentiment analysis method with the theme of the Covid-19 Pandemic on Twitter social media with the most optimal and maximum accuracy. Meanwhile, the benefit is to help Indonesians conduct research on public opinion on Twitter which contains positive, neutral, or negative sentiments. In this study, Text Mining has been used using the Support Vector Machine and Multinomial Naive Bayes method, which is one of the methods used to conduct sentiment analysis. The results of the Support Vector Machine and Naive Bayes Multinomial method used for sentiment analysis are useful for obtaining information and knowledge about Covid-19. The results obtained are that the opinion of the Indonesian people regarding Covid-19 is that the Positive and Neutral amount is more than the Negative with an average of 40% Positive and Neutral while the Negative is 20%. For the calculation of the F-1 average value, it can be seen that the Support Vector Machine method is the best model in this study with a value of 93%. Little difference with the Naive Bayes Multinomial method of 92%. It is hoped that this research will continue with more data and use other sentiment analysis methods in order to find out the average results of the opinions of the Indonesian people regarding Covid-19. Keywords - Sentiment Analysis, Multinomial Naive Bayes, Twitter, Data Mining, Covid-19.

1. Introduction

Technology is currently very rapidly developing, which makes an event that occurs can spread quickly. One of the most popular technologies at this time is the internet because, with the availability of the internet, information that is scattered and important data can be found and stored easily and quickly. This information can be obtained from various existing information portal sites. Most of the information is in the form of text [1] [2].

Social media, especially Twitter, is now a very popular communication tool among internet users in Indonesia. Twitter users, based on data from PT Bakrie Telecom, have 19.5 million users in Indonesia out of a total of 500 million global users.
There is no mistaking it now that Twitter social media is an effective and efficient place for sharing. The Coronavirus, known as Covid-19, was first discovered in Wuhan City in December 2019. Since the spread of this virus, many people have talked about it on Twitter, which has brought attention to the international community. The large number of victims who left until its rapid spread made people afraid and worried and the emergence of various opinions about Covid-19, including in Indonesia. It is the opinions of the community that is then analyzed to find out what has positive, neutral, or negative characteristics. Data mining is a process where information is extracted from data collected in the statistical field. Data mining is an analysis process from different views to get conclusions so that it becomes valuable information that can increase profits at minimal costs [3] [4].

Sentiment analysis or opinion mining is the process of understanding, extracting, and processing textual data automatically to get sentiment information contained in an opinion sentence. In this study, sentiment analysis was conducted to see a public opinion about the Covid-19 Pandemic. The magnitude of the influence and benefits of sentiment analysis causes research and applications based on sentiment analysis to develop rapidly. Even in America, there are about 20-30 companies that focus on sentiment analysis services [5].

The method used is the Support Vector Machine (SVM) and Multinomial Naive Bayes method. This method is used to find the best average of public opinion about Covid-19. The results obtained are the number of scores the average value obtained from the precision and recall where the highest result is a good or good method in analyzing the problem. These results are obtained by using Jupiter preprocessing with the help of the Python library. Opinion data is taken from every public Twitter tweet with covid-19 or something related to covid-19 with the help of the Rapidminer Studio application.

Rapidminer Studio is an application that deals directly with data retrieval to provide data preparation, machine learning, deep learning, text mining, and predictive analysis service. The way it works is to use the Twitter Key that you get by registering it on Twitter. After getting the Twitter Key, the Key is entered and linked with the keywords you want to search. After that wait and the keyword results will come out of the Rapidminer Studio program [6] [7].

Sentiment analysis or opinion mining is the process of understanding, extracting and processing textual data automatically to get sentiment information contained in an opinion sentence. This paper aims to find out how the opinion of the public's sentiment towards Covid-19 tweets which will produce a percentage of positive, negative, and neutral tweets. So that the Indonesian people know how the impact of the Corona virus in Indonesia in public opinion through the social Twitter application. The results obtained in this study is that using the Support Vector Machine and Multinomial Naive Bayes method received a response or an opinion that the Indonesian people about Covid-19 are positive with numbers 2004 and Negative 999 and Neutral 1997 from 5000 data. For the percentage of sentiment analysis results on Covid-19, it can be seen that positive and neutral responses are more than negative responses with an average of 40% for positive and neutral responses while for negative responses on average 20%. For the calculation results of the average F-1 score, it is shown that the Support Vector Machine method is the best model in this study with a value of 93%. the slight difference with the Multinomial Naive Bayes method of 92%.

For both models, the prediction results that often miss are between Neutral predictions and negative predictions, this may be due to the unevenness of the total data in each label in the dataset, it can be seen that the positive and neutral labels are almost the same, but we know that the impact of Covid-19 is brought a negative impact. Therefore, for future research, the dataset must be carefully prepared with the amount of data that may be greater than 5000 datasets, provided that the data must be balanced which will provide better results to be classified and the sentiment results of the value of Covid-19 are known.

2. Literature Study

2.1. Related Works

In this research study, research that is already similar or that which already exists is needed. The existence of this research uses research data from questionnaires, original data, and data from tweets on Twitter. This study is entitled "Analysis of Newspaper Framing of the Nigerian Governors"
Election Crisis of 2013 Forum "which aims to determine whether Twitter is a provider of a rich database of the spatial information about users who broadcast their opinions, sentiments, and real-time activities. In this paper, researchers seek to investigate the holistic influence of land use and time period on public sentiment. The IBM Watson Alchemy API is used to measure sentiment scores sent via tweets on a large scale. Then the researchers analyzed the sentiment scores statistically across time and space. The results show a clear spatiotemporal pattern of user sentiment. A higher sentiment score is particularly evident in commercial and public areas, during the day/night and on weekends [8].

This research is entitled "Analysis of Public Sentiment Against Corona Virus Based on Opinion from Twitter Based on Web Scraper" which aims to find out how the impact of the coronavirus in Indonesia is in accordance with public opinion. This study uses the Web Scraper method with data taken from Twitter tweets. The results of this study indicate that the percentage of Indonesian people's opinions is 79% negative, 11% neutral, and 10% positive [4].

This research is entitled "Analysis of Public Sentiment About the Development of the Indonesian E-Sport Industry on Twitter Using Multinomial Naive Bayes" which aims to determine the impact of the development of the world of E-Sport among the Indonesian people. This study uses the Naive Bayes Classifier, Multinomial Naive method with data taken from Twitter tweets. The results of this study indicate that the percentage of Indonesian people's opinion on E-Sport with the Bayes Naive Multinomial method for PUBG is 68% positive and Mobile Legend is 66% positive [5].

This research is entitled "Software Requirements Classification Using Machine Learning Algorithms" which aims to show a comparison between a text feature extraction technique and machine learning algorithms. Farmers aim to compare which system is better by using the Logist Regression (LR) method, Support Vector Machine (SVM), Multinomial Naive Bayes (MNB), and k-Nearest Neighbors (kNN). The results obtained are that the use of the Logist Regression (LR) method is better than other methods with a result of 0.91 [6].

This study is entitled "Comparison of Selection Features in the Support Vector Machine Algorithm for Review Sentiment Analysis" which aims to determine public opinion about tourist destinations by conducting comparative studies with several SVM methods such as the Support Vector Machine (SVM) Model, Particle Swarm Optimization (SVM-PSO), and Generate Algorithm (SVM-GA). By using Twitter as much as 300 data, 150 positives, and 150 negatives. The printed data will be processed by performing the Tokenization, Stemming, Stopword Removal, and Generate NGrams processes. The results of this study indicate that the Support Vector Machine model produces 75.33%, while the Particle Swarm Optimization (SVM-PSO) produces 88.67% and the Generate Algorithm (SVM-GA) produces 84.33%. This study shows that the Particle Swarm Optimization (SVM-PSO) based Support Vector Machine model produces higher values than other models. [9].

This study is entitled “Using Twitter to Better Understand the Spatiotemporal Patterns of Public Sentiment A Case Study in Massachusetts, USA” which aims to determine Twitter provides a rich database of spatial-temporal information about users broadcasting their opinions, feelings, and activities in real-time. In this article, researchers seek to investigate the holistic influence of land use and timing on public sentiment. IBM Watson Alchemy API is used to measure sentiment scores sent by large-scale tweets. Next, the researchers analyzed the sentiment scores statistically over time and space. The results show a clear spatiotemporal pattern of user sentiment. A higher sentiment score was particularly evident in commercial and public spaces, day and night, and on weekends [10].

This research is entitled "It’s Getting Personal: Personalisation of political campaigns in four Prague districts during the 2018 Czech Senate elections" which is where In October 2018 Senate elections were held in the Czech Republic. In the capital city of Prague, 41 candidates - both party members and independents - are contesting votes from voters in four districts. The purpose of this article is to analyze the election campaigns carried out in these four districts on the online realm of the social media site Facebook. Through complementary quantitative and qualitative methods, this text focuses on the communication of the candidates themselves, but also the reactions of voters in the social media environment. Using qualitative content analysis of topics covered by candidates, analysis of user comment sentiment, and quantitative analysis of posting frequency and followers, this article examines whether candidates who lead active personalization campaigns are more successful than candidates who communicate with the public only sporadically and with less personalization [11].
This study entitled "Exploring US Shifts in Anti-Asian Sentiment with the Emergence of COVID-19" aims to find out about an analytical report on the impact of anti-racism in which anecdotal shows that anti-racism and racist attitudes towards Asians are increasing. in response to COVID-19. Racism can have significant social, economic, and health implications, but there has been a little systematic investigation into the emergence of anti-Asian bias. Method: We used Twitter's Live Streaming Application Programming Interface to collect 3,377,295 tweets related to US racism from November 2019 to June 2020. This was done using vector machine support, a supervised machine learning model. The accuracy for identifying negative emotions when comparing machine learning models with manual-labeled tweets is 91%. Result: The percentage of negative tweets referring to Asians increased by 68.4%. In contrast, the proportion of negative tweets mentioning other racial/ethnic minorities remained relatively stable during this period, dropping less than 1% for tweets mentioning black people and increasing by 2% for tweets. refers to Latinx. Common topics that emerged during the content analysis of a random sub-sample of 3,300 tweets included: racism and error, anti-racism, and the impact of everyday life [12].

In today's world, social media is everywhere, and everyone is in contact with it every day. Within the scope of this article, we conclude and analyze the sentiments and manifestations of users of the Twitter social media platform, based on the main trends with Natural Language Processing and by Sentiment Classification using Artificial Neural Networks. Trained models work much more accurately, with a smaller margin of error, in determining emotional polarity in today's 'modern' often with ambiguous tweets. This study uses a Recurrent Neural Network (RNN) to classify emotions in tweets. Researchers developed a model to analyze the emotional nature of tweets, use repetitive neural networks for emotional predictions, look for relationships between words, and tag them with positive or negative emotions. The result is that the RNN model carries a small amount of data in neutral or completely reduced to zero. This proves that the researcher's model is capable of making decisions and categorizing in several directions even based on small details. Our comparisons were made primarily against TextBlob, which also performs very well and provides stable results, but is often a neutral result above 30% compared to the RNN model, which we cannot use as useful for further evaluations as for our RNN. Model [13].

This research entitled "Mathematical modeling for the prediction of cerebral white matter lesions based on clinical examination data" which aims to find out which patients have the highest risk of White Matter disease using the Naive Bayes method, Mathematical Models, and Vector Machines without the help of MRI (Magnetic resonance imaging) machine. By using data from patients affected by White Matter disease as many as 1904 patients consisting of 988 men and 916 women at Shin Takeo Hospital, Tokyo City, Japan. The result of this study is that the methods used above have not confirmed whether the highest score is the patients who are most at risk for White Matter disease. But with the existence of this method model helps doctors to better know which patients are at the highest risk of White Matter before the patient is put into an MRI (Magnetic resonance imaging) machine [14].

3. Methodology
   A. Multinomial Naive Bayes

   Multinomial Naïve Bayes is one of the existing models in Naïve Bayes which is useful for classifying text. The method used by Multinomial Naïve Bayes is supervised learning, where any existing data needs to be labeled before conducting training.

   Here is a basic explanation of the Multinomial Naïve Bayes:

   \[
P(w_i | c_j) = \frac{\text{count}(w_i, c_j) + 1}{(\sum_{w \in V} \text{count}(w, c_j)) + |V|}
   \]

   - \(P(w_i | c_j)\) : The probability of a word \(i\) falling into category \(j\).
   - \(\text{count}(w_i, c_j)\) : The number of query words that appear in a class or category.
   - \(1 : +1\) is given to avoid zero values.
   - \(\sum_{w \in V} \text{count}(w, c_j)\) : The sum of all the words in the class or category \(c_j\).
B. Confusion Matrix

This section will explain the calculation using the Confusion Matrix, which is a method for testing an object to distinguish which one is right and what is wrong. In calculating the Confusion Matrix, four calculations are performed, namely Accuracy, Precision, Recall, and Specificity (f1-score). The following is the calculation formula for Accuracy, Precision, Recall, and Specificity as well as priorities in Machine Learning [15]:

1. **Accuracy**

   On accuracy comparing the ratio of True Positive (TP) and True Negative (TN) on all data with formulas:

   \[
   \text{Accuracy} = \frac{\sum_{i=1}^{n} (TP_i + TN_i)}{\sum_{i=1}^{n} (TP_i + TN_i + FP_i + FN_i)} \times 100\% 
   \]  

2. **Precision**

   On Precision comparing the True Positive (TP) ratio on the data, namely positive predictions with formulas:

   \[
   \text{Precision} = \frac{\sum_{i=1}^{n} (TP_i)}{\sum_{i=1}^{n} (TP_i + FP_i)} \times 100\% 
   \]  

3. **Recall**

   On recall comparing the True Positive (TP) ratio on the whole data, namely True Positive with the formula:

   \[
   \text{Recall} = \frac{\sum_{i=1}^{n} (TP_i)}{\sum_{i=1}^{n} (TP_i + FN_i)} \times 100\% 
   \]  

4. **F-1 Score**

   The F-1 score is the result of the average value obtained from precision and recall using the formula:

   \[
   F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} 
   \]  

The following is the preprocessing stage by retrieving the data from RapidMiner Studio:

1. **Collecting Tweet Data**

   The data taken in tweets on Twitter is 5,000 data from January 2020 to September 2020 using the RapidMiner Studio application. The data taken from Twitter's tweets are data with the keyword “Covid-19”. Tweets or data taken are Indonesian words. The results of 5000 tweet data about public opinion on Covid-19 are shown in Figure 1.

   ![](Figure1.png)  

**Figure 1.** Example of Twitter Tweet Data from RapidMiner Studio with Keywords Covid-19

2. **Preprocessing Data**

   After searching for data and collecting tweet data from RapidMiner Studio, the next step is to conduct data preprocessing. This step is made to eliminate double words or empty words. The following are the steps in conducting data preprocessing, namely [5]:

- |v| : The sum of all unique words in all categories.
a. Normalization Character Repeat
   This stage is used to delete the same characters or there is more than one character.

b. Normalization Spell
   This stage is used to correct the abbreviated words to be correct as they are, I became caveats with the Python ModSpellChecker library.

c. Tokenization Text
   This token text is obtained by creating a twitter account and the Tokenize Text stage can break down tweet data in the form of sentences into words.

d. Constructions Extend
   This stage is used to be able to solve word abbreviations and tell the meaning of these abbreviations such as etc. to and so on by using the Python Constructions library.

e. Stemming
   This stage is used with the help of the Python StemmerFactory library to change word data to be more basic and more understanding.

f. Removing Special Characters
   This stage is used to remove special character words such as @, $, %, and others in Twitter tweet data.

g. Removing Stopwords
   This stage is used with the help of the Python Literature library to remove words that are not standard or meaningless in Twitter tweets.

h. Normalization Corpus
   This stage is used to combine the words that have been carried out by the 7 stages above and the results become the core words of the tweet data. For the results can be seen in Figure 2.

Figure 2. Tweet Dataset Preprocessing Results

3. Weighing and Labeling
   At this stage, weights, and labels using Excel. The tweet data that has been stored in Excel can show the results of the value of public sentiment towards Covid-19, which results are positive, negative, and neutral.

Figure 3. Rated Dataset and Labels

Figure 3.3 shows the results of the data set in this study, amounting to 5000 tweet data. Then the data will be labeled with positive results totaling 2004 data, negative labels totaling 999 data and neutral labels totaling 1997 data and you can see the results in Figure 4.
4. Evaluation Model

In this section to know how to calculate sentiment in Excel:

| Sentiment | Score |
|-----------|-------|
| Negative  | -1    |
| Neutral   | 0     |
| Positive  | +1    |

Figure 5. Sentiment Weighted Scoring

Figure 5 shows the calculation of the sentiment weight where Positive gets +1, for Negative is given a weight of -1 and for Neutral is given a weight of 0. Calculations using Excel to get positive and negative results can be used using Excel formulas as shown in Figure 6 is a positive result and for Figure 7 is the result for negative, like the image below:

\[
\text{Positive Result Calculations} = \frac{\text{SUMPRODUCT}((\text{LEN}(A2)-\text{LEN}((\text{UPPER}(A2)))); \text{UPPER}((\text{Positif!$A$1:$A$103};"")))}{\text{LEN}((\text{Positif!$A$1:$A$103}))}
\]

\[
\text{Negative Result Calculations} = \frac{\text{SUMPRODUCT}((\text{LEN}(A2)-\text{LEN}((\text{UPPER}(A2)))); \text{UPPER}((\text{Negatif!$A$1:$A$2336};"")))}{\text{LEN}((\text{Negatif!$A$1:$A$2336}))}
\]
It can be seen in Figure 8 taken above that the public's response to their opinion about Covid-19 is negative. Meanwhile, if we take data from Twitter and do preprocessing using Rapid Miner, it shows that from the 5000 tweet data taken, it shows that public opinion shows a greater positive result than negative. This shows that the data used is not large enough and needs to be tested again with more data and other methods in order to find out which method is better for finding public sentiment analysis.

4. Result and Analysis

At this stage, the model used for test data classification is the Support Vector Machine algorithm and the Multinomial Naïve Bayes algorithm, which are algorithms that focus on text classification. The model evaluation stage, which discusses the test results to determine the validity of the tweet obtained. The results issued using the Jupyter Preprocessing Application contained 4 results in the form of Accuracy, Precision, Recall, and F-1 Score [5] [16]:

Here are the results of the two methods, namely the Support Vector Machine and Multinomial Naïve Bayes:

1. Support Vector Machine (SVM)

Figure 9 shows the results for the Support Vector Evaluation Model, where the score for Accuracy is 0.93, for Precision is 0.93, for Recall is 0.93 and the F-1 Score is 0.93.

![Figure 7. Negative Result Calculations](image1)

![Figure 8. Trends in Public Sentiment Against Covid-19](image2)

![Figure 9. Results of the Support Vector Machine Evaluation Model](image3)
Figure 10 is an image of the result of a configuration matrix with a Support Vector Machine. In the figure, there are 659 true positives, 282 true negatives, 34 false positives, 11 false negatives. Neutral Negative totals 20.

![Confusion Matrix](image)

**Figure 10.** Confusion Matrix Results with Support Vector Machine

2. Multinomial Naive Bayes (MNB)

Figure 11 shows the results for the Naive Bayes Multinomial Evaluation Model, where the score for Accuracy is 0.92, for Precision is 0.92, for Recall is 0.92 and the F-1 Score is 0.92.

|                      | precision | recall | f1-score | support |
|----------------------|-----------|--------|----------|---------|
| Negatif              | 0.90      | 0.79   | 0.84     | 336     |
| Neutral              | 0.99      | 0.92   | 0.95     | 652     |
| Positif              | 0.87      | 0.98   | 0.92     | 682     |

**Figure 11.** Results of the Multinomial Naive Bayes Evaluation Model

Figure 12 is the resulting image of the confusion matrix with Multinomial Naive Bayes. In the figure, there are 666 True Positives, 265 true Negatives, 64 False Positives, 15 false-negative ones. Negative Neutral amounted to 7.
10

Figure 12. Confusion Matrix Results with Multinomial Naive Bayes

The results obtained in this study is that using the Support Vector Machine and Multinomial Naive Bayes method received a response or an opinion that the Indonesian people about Covid-19 are positive with numbers 2004 and Negative 999 and Neutral 1997 from 5000 data. For the percentage of sentiment analysis results on Covid-19, it can be seen that positive and neutral responses are more than negative responses with an average of 40% for positive and neutral responses while for negative responses on average 20%. For the calculation results of the average F-1 score, it is shown that the Support Vector Machine method is the best model in this study with a value of 93%. the slight difference with the Multinomial Naive Bayes method of 92%.

5. Discussion

This discussion explains the discussion of sentiment analysis of public opinion. In Joshua Restu's [5] paper using the Multinomial Naive Bayes method, the f-1 score for PUBG is 61%, for Mobile Legend is 79%. It can be seen that this number does not reach a good number because the value is only up to 80%. Meanwhile, for Kurniawan and Apriliani's [4] paper using the Web Scraper method, the results were 79% positive, 11% negative, and 10% neutral. The positive results of Kurniawan and Apriliani's paper got good results because the positive results were higher than the negative results.

Meanwhile, the paper I created got an F-1 Score with Multinomial Naive Bayes of 92% and for Support Vector Machine of 93%. With a Positive result of 40%, for Negative it is 11% while Neutral is 39%. This figure shows that it is not optimal because the positive and neutral values are not much different. Maybe if the data is added above 5000 tweets, it will produce more than the above results. But to use Confusion Matrix calculations get perfect results because the results are above 90% using both the Multinomial Naive Bayes (MNB) method and the Support Vector Machine (SVM).

In the paper, the researchers showed that the value of MNB was 92% and SVM was 93%. Why is the result of SVM higher than MNB is because the SVM method according to other paper research shows that the calculation of the SVM method uses Probabilistic or means the calculation for documents whose data is large enough while for MNB calculations it shows geometry which means the calculation used for training, which was if there is a lot of training it will produce less good results, whereas if the training is little it will produce good results.

6. Conclusion

The conclusion that I get is that the machine learning algorithms from the Support Vector Machine and Multinomial Naive Bayes methods are suitable for calculating public sentiment analysis
of the opinions they make on social media. But this method has advantages and disadvantages where it must use large enough data to produce better values. And it is also necessary to test other methods to get better results by what method should be used. For future research, it is hoped that more sentiment analysis can be carried out with quite a lot of data taken because the Covid-19 case is continuing until 2021, to get better and better results so that the results of public opinion sentiment towards Covid-19 can be seen.

7. References
[1] Dickson, “10 Situs Jejaring Sosial Terpopuler di Dunia,” 10 Jejaring Sos. Terpopuler di Dunia, pp. 1–9, 2020.
[2] Kementrian Komunikasi dan Informasi, “Kominfo : Pengguna Internet di Indonesia 63 Juta Orang,” Website Resmi Kementeri. Komun. dan Inform. RI, pp. 1–9, 2013.
[3] R. T. L. I. H. C.-19-9 Juli, “Total : 919,” Isu Hoaks Corona Virus, p. 919, 2020.
[4] R. Kurniawan and A. Apriliani, “Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scraper,” Jurnal Instek (Informatika Sains dan Teknologi), vol. 5, no. 1. p. 67, 2020.
[5] J. R. Valentino, “Analisis Sentimen Masyarakat Tentang Perkembangan Industri E-Sport Indonesia Pada Twitter Menggunakan Multinomial Naive Bayes,” 2020.
[6] E. Dias Canedo and B. Cordeiro Mendes, “Software Requirements Classification Using Machine Learning Algorithms,” Entropy, vol. 22, no. 9. p. 1057, 2020.
[7] M. Mardalius, “Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means,” Jurteksi, vol. 4, no. 2, pp. 123–132, 2018.
[8] D. O. Anyadike, “Analysis of Newspaper Framing of the Nigerian Governors’ Forum Election Crisis of 2013,” Mediterr. J. Soc. Sci., vol. 6, no. 6, pp. 315–327, 2015.
[9] Y. T. Arifin, “Komparasi Fitur Seleksi Pada Algoritma Support Vector Machine Untuk Analisis Sentimen Review ,” vol. 3, no. September, pp. 191–199, 2016.
[10] X. C. Id, P. Macnaughton, Z. Deng, J. Yin, X. Zhang, and J. G. Allen, “Using Twitter to Better Understand the Spatiotemporal Patterns of Public Sentiment : A Case Study in Massachusetts , USA,” 2018.
[11] O. Brunnerová, “It ’ s Getting Personal : Personalisation of political campaigns in four Prague districts during the 2018 Czech Senate elections,” vol. 15, pp. 279–308, 2019.
[12] T. T. Nguyen et al., “Exploring U . S . Shifts in Anti-Asian Sentiment with the Emergence of COVID-19,” pp. 1–13, 2020.
[13] L. Nemes and A. Kiss, “Social media sentiment analysis based on,” J. Inf. Telecommun., vol. 0, no. 0, pp. 1–15, 2020.
[14] Y. Shinkawa, T. Yoshida, Y. Onaka, M. Ichinose, and K. Ishii, “Mathematical modeling for the prediction of cerebral white matter lesions based on clinical examination data,” PLoS One, vol. 14, no. 4, pp. 1–16, 2019.
[15] R. Arthana, “diprioritaskan dalam Machine Learning,” pp. 1–9, 2020.
[16] M. Kaya, G. Fİdan, and I. H. Toroslu, “Sentiment Analysis of Turkish Political News,” no. December, 2012.