DOI:10.14421/ijid.2021.2400 IJID (International Journal on Informatics for Development), *e-ISSN :2549-7448*
Vol. 10, No. 1, 2021, Pp. 23-30

Sentiment Analysis and Topic Modeling of Indonesian Public Conversation about COVID-19 Epidemics on Twitter

Muhammad Habibi[1*], Adri Priadana[2] , Muhammad Rifqi Ma'arif[3]
[1]Department of Informatics, Universitas Jenderal Achmad Yani Yogyakarta, Yogyakarta, Indonesia
[2,3]Center of Data Analytic Research and Services, Universitas Jenderal Achmad Yani Yogyakarta, Yogyakarta, Indonesia
[1]muhammadhabibi17@gmail.com, [2]adripriadana3202@gmail.com, [2]muhammad.rifqi@gmail.com

Article History
Received Mar 28th, 2021
Revised Jun 14th, 2021
Accepted Jun 15th, 2021
Published Jun, 2021

***Abstract*— The World Health Organization (WHO) declared the COVID-19 outbreak has resulted in more than six million confirmed cases and more than 371,000 deaths globally on June 1, 2020. The incident sparked a flood of scientific research to help society deal with the virus, both inside and outside the medical domain. Research related to public health analysis and public conversations about the spread of COVID-19 on social media is one of the highlights of researchers in the world. People can analyze information from social media as supporting data about public health. Analyzing public conversations will help the relevant authorities understand public opinion and information gaps between them and the public, helping them develop appropriate emergency response strategies to address existing problems in the community during the pandemic and provide information on the population's emotions in different contexts. However, research related to the analysis of public health and public conversations was so far conducted only through supervised analysis of textual data. In this study, we aim to analyze specifically the sentiment and topic modeling of Indonesian public conversations about the COVID-19 on Twitter using the NLP technique. We applied some methods to analyze the sentiment to obtain the best classification method. In this study, the topic modeling was carried out unsupervised using Latent Dirichlet Allocation (LDA). The results of this study reveal that the most frequently discussed topic related to the COVID-19 pandemic is economic issues.**

Keywords*— *COVID-19; Sentiment Analysis; Topic Modeling; Machine Learning; Natural Language Processing.

1 INTRODUCTION

The coronavirus disease (COVID-19) has spread rapidly throughout the world since it was first discovered in China. The World Health Organization (WHO) declared the COVID-19 outbreak a global health emergency [1]. Based on the results of the COVID-19 situation report originating from the official WHO website on June 1, 2020, the COVID-19 outbreak has resulted in more than six million confirmed cases and more than 371,000 deaths globally [2]. Research related to public health analysis and public conversations on the spread of COVID-19 on social media is also one of the highlights of research globally. Social media can spread disinformation about the virus. It was fueling panic and creating the so-called infodemics [3]. Furthermore, social media has long been recognized as a major spreader of health misinformation [4]. Social media use as a source of information is not regulated. It can lead to health risks through the spread of conspiracy theory, which cause concerns regarding the spread of the COVID-19 conspiracy theory on social media [5]. Analyzing public conversation will help the relevant authorities understand public opinion and information gaps between them and the public [6], helping them develop appropriate emergency response strategies to address existing problems in the community during the pandemic [7]. Moreover, analyzing sentiment analysis can provide information on the population's emotion in different contexts [8].

This study aims to understand public health by analyzing the sentiment and topic modeling of Indonesian public conversations on Twitter about the COVID-19 using the NLP technique. We applied some methods to analyze the sentiment, such as Gaussian Naive Bayes, Multinomial Naïve Bayes, Support Vector Machine, and Random Forest, to obtain the best classification method. In this study, the topic modeling was carried out unsupervised. We applied Latent Dirichlet Allocation (LDA) to conducted topic modeling.

Previous studies of public health and community conversations were carried out using supervised textual data analysis. In 2020 [9], Sear et al. analyzed the emergence and evolution of topics around COVID-19 on Facebook Pages using Latent Dirichlet Allocation (LDA). In that study, LDA was able to identify topics that make sense in a collection of posts from online communities around the vaccine and COVID-19 debates. It was also able to handle big data, and the prompt results were obtained using statistical clustering techniques, instead of having to rely on potentially biased, slow, and expensive human labeling.

Li et al. in 2020 [6] identified and characterized situational information on social media during the COVID-19 pandemic using supervised learning Natural Language Processing (NLP) techniques such as the Support vector machines (SVM) method, Naive Bayes (NB), and Random forest (RF) to learn the types of unlabeled data based on labeled data. The research produced several situational information related to COVID-19 that could be used to build an effective social media-based emergency response program and crisis information system. In 2020 [10], Part et al. investigated the information transmission network and news-sharing behavior related to COVID-19 on Twitter in Korea by classifying the top news channels shared via tweet. Based on this research, social media network analytics were used to monitor public conversations and news dissemination that take place in real time.

People nowadays also increasingly use emoticons in texts to express feelings or recapitulate their words [25]. Lwin et al., in 2020 [11], examined the worldwide trend of four emotions (fear, anger, sadness, and joy) and the underlying narratives of those emotions from 20 million data tweets during the COVID-19 pandemic. From this research, it can be seen that public feelings shifted sharply from fear to anger during the pandemic, while sadness and joy also emerged. Jelodar et al., in 2020 [12], utilized the NLP technique to classify public sentiment regarding COVID-19 on the Reddit platform. The study also quantified the polarity of COVID-19 comments regarding sentiment and opinion analysis. The results of these studies indicate that the method applied produces an accuracy of 81.15%. Samuel et al., in 2020 [13], identified public sentiment related to the COVID-19 pandemic on Twitter by utilizing the machine learning (ML) classification method. One of the ML methods used in this study is the Naïve Bayes method, which provides a reliable classification accuracy of 91% for short Tweets. These studies prove that the ML-based NLP technique is proven to be used to analyze public perception.

2 METHOD

2.1. Tweet Data Acquisition and Preprocessing

The data in this study were obtained from the Twitter streaming API between March 2 and May 8, 2020. We chose this date following President Joko Widodo's announcement of the first two confirmed cases of COVID-19 infection on March 2, 2020 [14]. During that time, we managed to collect 477,011 tweets related to COVID-19 using the keywords " wabah COVID-19" and " virus corona". Twitter streaming API provides an interface to acquire a complete set of tweet attributes. However, in this experiment, only some attributes were used. Table 1 outlines the tweet attributes used in this research.

Table 1. attributes of tweet data from twitter API.

Tweet Attributes	Description
Tweet ID	unique Twitter user identity
Screen Name	username of the Twitter account.
Tweet Text	a post on the social media site Twitter
Timestamp	a sequence of characters or encoded information identifying when a certain event occurred
Retweet	a re-posting of a Tweet
Likes	used to show appreciation for a Tweet

After data acquisition, the next step was data pre-processing. Generally, raw tweets contain a lot of noise, misspelled words, meaningless words, including various abbreviations and slang words. These words often disturbed the data on the sentiment and obstructed the classification model's performance. Therefore, tweets should be pre-processed before the features were actually extracted. The followings are the steps of the pre-processing:

- Tokenization: the process of dividing a text into specific parts.
- Normalization: taking text to its standard form. The general normalization techniques used are as follows:
  - Case folding: change capital letters to lowercase
  - Elimination of periods in terms. For example, S.K.M. to SKM
  - Remove hyphens in a term. For example, health workers-doctors to health workers doctors.
- Cleaning: The steps in cleaning tweet data are as follows:
  - Remove URL contained in the tweet
  - Remove the @ sign in the username
  - Remove hashtags (#) contained in the tweet
  - Remove the numbers contained in the tweet
  - Remove punctuation marks, such as question marks, exclamation points, periods, and others.
  - Remove Unicode and symbols
- Stopword Removal: remove words that are deemed meaningless using the stopword list.

2.2. *Sentiment Analysis*

Sentiment analysis or opinion mining is a field of study that analyzes people's opinions, sentiments, evaluations, judgments, attitudes, and emotions towards product entities, organizational services, individuals, issues, events, topics, and attributes [15]. Sentiment analysis is often referred to as subjectivity analysis, opinion mining, and assessment extraction with several relationships related to affective computing, namely computer recognition and emotional expression [16].

Different machine learning methods can be used to classify tweets based on sentiment polarity. The techniques used in this study are as follows:

2.2.1 *Gaussian Naïve Bayes:* Gaussian Naïve Bayes algorithm makes use of the Bayes Theorem of conditional probability [17]. Gaussian Naïve Bayes is easier to use because you only need to calculate the mean and standard deviation of the training data. Classify each tweet and calculate each feature's probability in the tweet for each positive, negative, and neutral class label. Class labels are assigned to tweets based on the sentiment label with the most significant sentiment [18]. The normal distribution calculation is shown in Equation 1.

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}) \qquad (1)$$

2.2.2 *Multinomial Naïve Bayes:* the Naive Bayes Theorem is the essence of the Naive Bayes classifier, which uses possibilities that can then be used to classify data sets. Events X and Y, calculation of the Multinomial Naive Bayes is shown in the Equation 2.

$$P(X|Y) = \frac{P(Y|X).P(X)}{P(Y)} \qquad (2)$$

Naïve Bayes' Multinomial Classification is a modified version of a text-related algorithm mainly used for classification, taking into account different features such as word frequency.

2.2.3 *Support Vector Machine:* SVM is a Support Vector Machine. It is a non-probabilistic binary linear classifier. For a training set of points (x_i, y_i), x is a feature vector, and y is the class. To determine the maximum margin hyperplane that divides the points with $x_i = 1$ and $x_i = 1$. The equation of the hyperplane is: $w.x + b = 0$. For a data set consisting of features set and labels set, an SVM classifier builds a model to predict the classes for the new examples. It assigns a new case or data points to one of the categories [19].
Algorithm:

a) Define an optimal hyperplane
b) Extend step I for nonlinearly separable problems
c) Map data to high dimensional space where it is easy to classify with linear decision surfaces.

2.2.4 *Random Forest:* Random forest classifier is an ensemble learning classification algorithm. It is very similar to the decision tree but contains a multitude of decision trees, and the class label is the mode value of the classes predicted by individual decision trees. This algorithm is efficient in handling large datasets and thousands of input variables without their deletion. Random forest uses majority vote and returns the class label with maximum votes by the individual decision trees. Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads [18]. The decision tree method can be related to a rule-based system. Some rules shows up in the decision tree algorithm when the training data file with targets and features is provided. The similar set rules can be utilized to exhibit the prediction on the test dataset [19]

2.3. *Topic Modeling*

Topic modeling is statistical-based processing to find a set of topics within a certain collection of text documents. Those topics give an intuitive syntactic representation of ideas being discussed along with the text document [20]. Hence, a collection of documents can be broken down into several categories according to the similarity of their topic model. In the textual form, the topic model is represented by a set of keywords that best represent certain parts of text documents. In recent years, topic modeling has gained wide attention in natural language and machine learning communities [21]. The main reason for using topic modeling in this experiment is its flexibility and reliability to reveal the hidden topic within a huge number of text documents. Topic modeling is a robust approach to find the statistical correlation between words within the text corpus to form a syntactically sound topic represented by a set of words. The main objective of topic modeling is to discover the use of words along with the documents and how to relate them with other words in different segments of the document [22].

The basic idea of topic modeling is to find a set of words that frequently co-appear together within a certain range in a different segment of document corpus according to statistical measurements. From the perspective of the topic model, a collection of document corpus consists of a set of distinct topics, and every topic consists of a set of distinct words or keywords [23]. Furthermore, for each word in each topic, there is a probability distribution on how certain significant words contribute to constructing the semantic meaning of certain topic [22]. Fig. 1 is a well-known example to illustrate the topic model topology of text documents created by Blei [20]. From Fig. 1, we can see that each topic contains a set of words that have a certain probability that shows the significance of those words within the topic. Then, each document contains a set of terminologies that represent a set of topics.

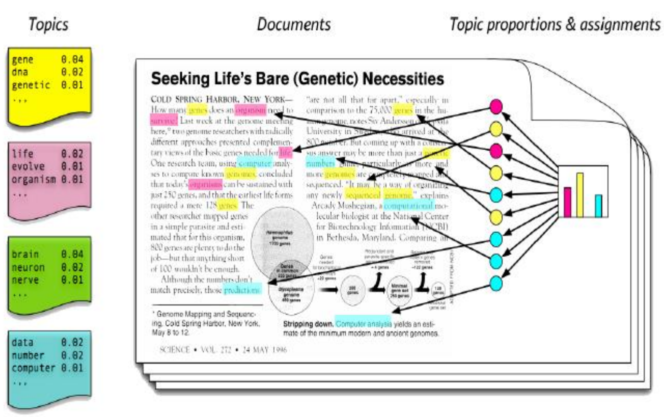

Figure 1. A general illustration of the topic model [20]

There are several algorithms to implement the topic modeling approach. One of the most widely used algorithms for topic modeling and being employed in this experiment is Latent Dirichlet Allocation (LDA). LDA is a generative probabilistic model that operates over the discrete data. Thus, algorithm is well suited to the text data [20]. From the machine learning perspective, LDA is categorized as an unsupervised learning approach since that approach automatically categorizes the unlabeled document into several groups according to the similarity of their terms probability distributions [24].

Topic modeling is a robust statistical-based approach to infer a set of topics that occurred within a text document collection. Nevertheless, several criteria need to be considered during the topic of model generation. The number of topics (*k*) should be determined at the beginning [25]. That could be a serious problem if there is no prior knowledge of the related documents. Hence, in this experiment, we iteratively generate various *k*-number of topics and evaluate each experiment to determine the best number of *k*. Then, for evaluating the best number of k-topic, we employed a metric named topic coherence. The coherence metric calculates how one term and another within a topic are statistically correlated with each other. The higher the correlation score, the higher the coherence. It indicates a better quality of the generated topic.

3 RESULT AND DISCUSSION

3.1. Sentiment Analysis

We have performed a term weighting scheme using several different machine learning classifications. The evaluation methods used to look at the performance of the model classifier are precision, recall, accuracy, and F1-Score [26]. Classifier performance evaluation is useful to see which classification method has a better performance value.

Two methods of measurement commonly used for classification performance analysis are "recall" and "precision". Both measures are useful for feature extraction, opinion phrase extraction, and information retrieval [27]. Precision is the proportion of the relevant prediction result. Recall is the proportion of relevant examples correctly classified. The F1-score is a harmonic average of precision and recall and is always closer to the smaller of the two—the higher the F1-score, the better the classifier's predictive power [28].

The results of the precision and recall calculations can be seen in Fig. 2. Fig. 2 summarizes all the classifiers' precision and recall using different term weighting schemes of all the datasets. Fig. 2 shows that the support vector machine method has the highest precision and recall values, namely 81%. It indicates that SVM is a better classifier than other classifiers.

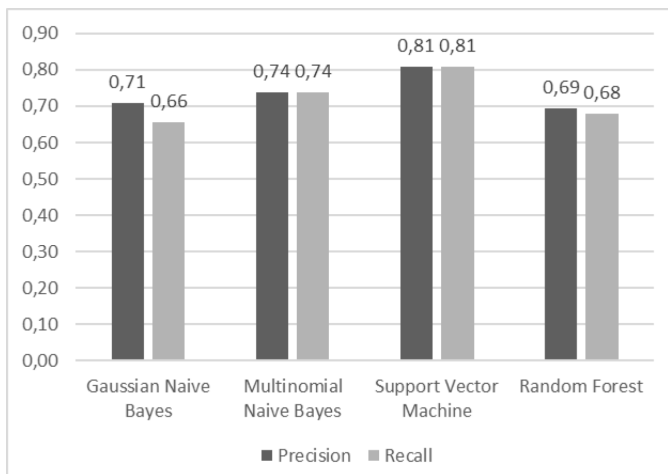

Figure 2. Precision and recall comparison of each algorithm

This section presents the comparative results of the accuracy and f1-scores of the four classification methods used. The results of using the four methods for COVID-19 tweet Classification are presented in Fig. 3. Interestingly, we found that tweets were correctly classified (81% accuracy) for the Support Vector

Machine method. Also, the SVM F1-score has the highest value as well. A higher F1-Score dan accuracy indicates a better classifier. The Naive Bayes multinomial classification method ranks second with an accuracy value of 74% and F1-score. In the third position is the Random Forest method with an accuracy value of 68% and an f1-score of 67%. The fourth position is the Gaussian Naive Bayes classification method with an accuracy value of 65% and an f1-score of 63%. Based on the results of the accuracy value and F1-score, the method that will then be used to classify the polarity of sentiments related to COVID-19 tweets is the Support Vector Machine. Similar to our analysis results, several studies have shown that the Support Vector Machine provides more accurate results than other methods [29], [30],[31].

After obtaining the best classification method, the next step was to classify sentiments. Sentiment classification objectives had been set, namely, positive and negative sentiments. The classification model that had been made during the training stage was used to classify it. All tweets about the COVID-19 were labeled as sentiments automatically. Sentiment classification results can be seen in Fig. 4.

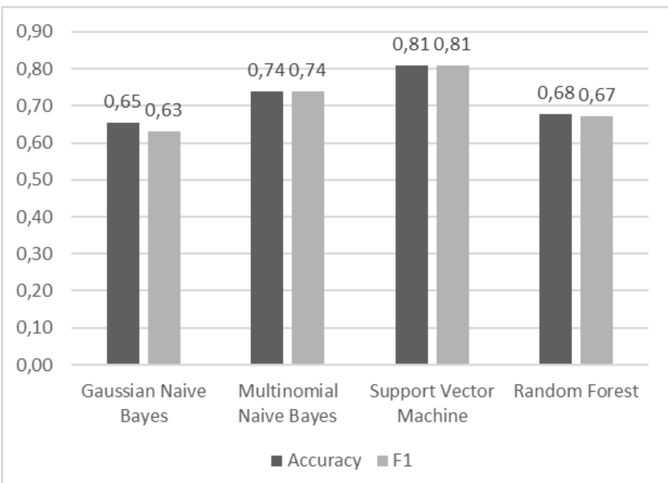

Figure 3. Accuracy and F1-score comparison of each algorithm

Analysis of Twitter data in this study displays positive and negative sentiments of tweets related to the COVID-19, as shown in Fig. 4. The tweets used in this study were taken when the new COVID-19 began to spread in Indonesia. The public comments mostly on the pros and cons of government policies related to large-scale social restrictions. The sentiment analysis results show that constant public sentiment for three months tends to be negative regarding the COVID-19. Negative sentiment has an average percentage of above 60% in three months, while positive sentiment has a rate of 36% to 40%. It can be seen that even though negative tweets have a higher percentage of positive tweets, but the tendency of negative tweets to have a decrease in number every month. At the same time, positive tweets tend to increase every month. This indicates a slight change in people's perceptions regarding COVID-19 from being negative to positive.

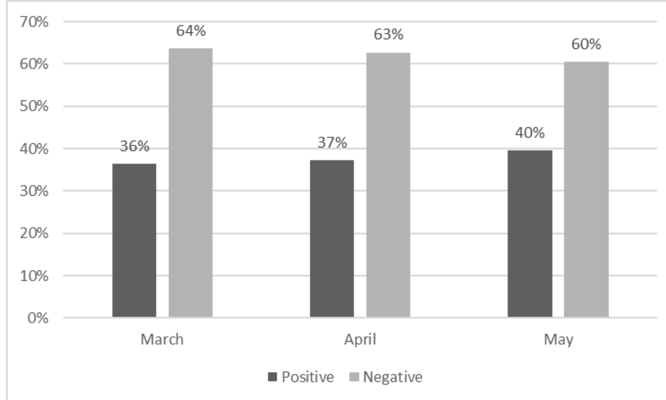

Figure 4. Positive and negative sentiment comparison by month

3.2. *Topic Analysis*

As mentioned in the previous part, topic modeling in our experiment was divided into two parts. The first part was constructing the LDA model, consisting of a higher number of topics that sound statistically. Then by using lexical similarity, we further grouped the topics with similar words/terms. For the first stage of topic modeling, we determined the best number of k-topic for our LDA model. To this objective, we calculated a topic coherence of the various number of k-topic iteratively. Fig. 5 depicts the coherence scores in various initial number of topic. In this experiment, we tried to generate 1 to 60 topics from our corpus. From Fig. 5, it can be seen that the coherence score significantly increased after reaching 15 topics. Then, it goes up and down until 60 topics. The coherence score reaches the peak in the graphic at 59 topics. Thus, we can determine that the best *k* is 59 topics. We then focused on this number for further analysis.

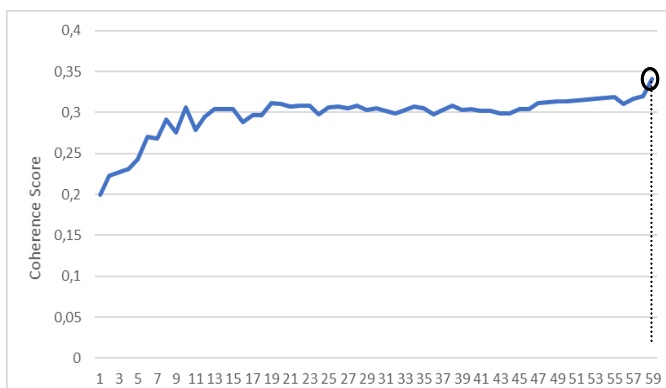

Figure 5. Coherence score of each number of topics

To better understand the structure of the words constructing every topic on the topic model, we then visualized the distance map of those 59 topics. This visualization gave an insight into how topics related to each other based on their shared terms/words. Fig. 6 shows the distance map of the 59 topics generated by the LDA algorithm over the document corpus used in this experiment. On Fig. 6, we can see that several topics have some distance from each other, and some other topics are overlapped with another. The overlapped topic shows that there are similar words that construct different topics. This 59 k-topics with the visualization in Fig. 6 are the result of the first two stages of topic modeling employed in this research.

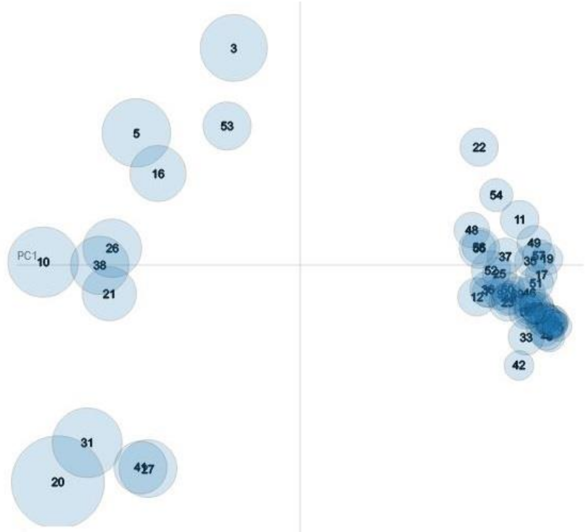

Figure 6. Term clusters visualization resulted by LDA

The first stage of topic modeling has been described in the previous paragraph, which is to determine the best number of k-topics for the LDA model. The second stage of topic modeling is to combine overlapping topics into one larger topic. The basic idea of topic representation in LDA is a topic consisting of a set of words. Therefore, overlapping topics will generally share a few words in their representation. Instead of manually selecting overlapping topics based on visualizations, we used binary cosine similarity to define aggregate topics. Thus, topics that share the most keywords in their topic representation will have a high similarity score and are considered similar topics. Based on Fig. 6, we can see that topics that have a long distance between topics mean that these topics have no relationship. However, topics with close or even overlapping distances between topics mean that these topics have relationship or may have similar topic themes. Table 2 summarizes the topic categories of the aggregation process as a result of the second phase of topic modeling in this study.

Table 2 shows that the topic aggregation process in stage two of the topic modeling work in this study resulted in six categories of topics. Each topic generated consists of several sets of words that have different occurrence weights. Topics that have the same occurrence of words are considered to have the same topic category, we can see which words are the most significant in building a topic category. This method helps to distinguish topics that can be interpreted semantically and topics that are the result of human interpretation [32]. Based on the similarity of the occurrence of words in the 59 topics, six categories of topics were obtained as shown in Table 2 of the topic description column. The topic distance map illustrates how many topics can be grouped by proximity. Of the six topic categories, two of them have a dominant share with more than one hundred thousand related tweets. The other two accounted for a fraction with some related tweets under thirty thousand. Fig. 7 shows the overall proportion of all topic categories.

Based on the proportion illustrated in Fig. 7, topic category 5 is the highest rate issue discussed by an Indonesian citizen. Table 2 shows that topic category 5 contains approximately 114 thousand tweets that mainly talk about the economic impact of COVID-19 pandemics. The second biggest issue Indonesian citizens talk about on Twitter is the policy on lockdown (topic category 1) with more than a hundred thousand tweets or 23% of total tweets acquired. These top 2 topic categories are related to each other, in which lockdown policy will immediately affect Indonesian citizen's economic situation. In terms of any policy taken by either central and local government, this topic also gain a lot of attention from Indonesian Twitter users with more than 86 thousand tweets. That issue (topic category 2) is placed in the fourth rank with a 19% related tweet from all acquitted tweets within our datasets.

Table 2. List of Topic Identified from LDA Term Clusters.

Topic Category	Associated LDA Term Cluster	Topic Description	Number of Tweet
Category 1	20,27,31,41	Citizen response to lockdown and access restriction option	101,268
Category 2	5,16	Citizen opinion to policy and action taken by local government in general	86,104
Category 3	10,21,26,36	Citizen conversation about health protocol to prevent the spread of corona virus	95,946
Category 4	53	Public appeal on the importance of educating people on how to respond to COVID-19 pandemics	24,765
Category 5	1,2,4,6,7,8, 9,11,12,13, 14,15,17,18 ,19,22,23,2 4,25,28,29, 30,32,33,34 ,35,37,38,3 9,40,42,43, 44,45,46,47 ,48,49,50,5 1,52,53,54, 55,56,57,58 ,59	Citizen story about the economic impact which already occurred	114,571
Category 6	3	Public appeal to society in order to have better awareness to surrounding environment	27,788

The third highest issue discussed by Indonesian citizens on Twitter related to COVID-19 pandemics is the health protocol to prevent the spread of COVID-19 (topic category 3) with more than 95 thousand tweets (21%). The next issues in topic category six and topic category four consecutively construct the bottom two outlined in Table II in terms of the number of tweets. Topic category six talks about how citizens post a campaign to encourage people to be more aware of their surrounding environment to tackle the spread of the virus. This topic is related to topic 4 in which people on Twitter try to educate society about COVID-19.

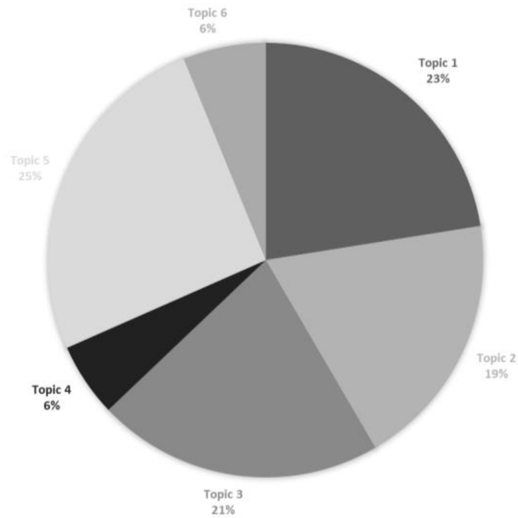

Figure 7. Proportion of tweets associated with each topic

3.3. Sentiment Analysis over Topic Category

In this section, we outlined the sentiment in each topic, which is already mentioned in the previous part of this paper, to complete the perspective of analysis. Fig. 8 outlines the sentiment proportion of each topic. From Fig. 8, it can be easily pointed out that almost all topics have a bigger portion of tweets with a negative tone rather than non-negative (positive sentiment). From six topics, only topic category six has more tweets with non-negative sentiment, while the rest has a negative tweet as the majority. If we look at Table III, topic category six discusses the public appeal to society to have a better awareness of the surrounding environment. Most tweets in this topic category tend to educate society with positive words to raise a better engagement in the environmental awareness campaign to prevent the spread of the COVID-19.

The rest of the topic other than that of category six is mostly interpreted in negative forms. The two biggest issues related to COVID-19 pandemics (topic category 1 and 5) are related to the immediate impact of COVID-19 pandemics on the economic aspect of Indonesian citizens. This economic issue is considered as the most crucial factor in Indonesian lives. Hence, since COVID-19 has a negative impact on their financial and economic situation, most of the economic-related conversation mostly delivered in negative interpretations. Topic three, which discusses the government policy in response to the COVID-19 pandemic, also has a bigger portion of negative sentiment. In topic three, Indonesian Twitter users tend to relate the government policy with the negative impact on the economy. A negative sentiment is not only dominant in economic-related conversation on Twitter but also in conversation about health. In categories three and four, the negative tweets point to a group of citizens who did not obey the health protocols defined by the government.

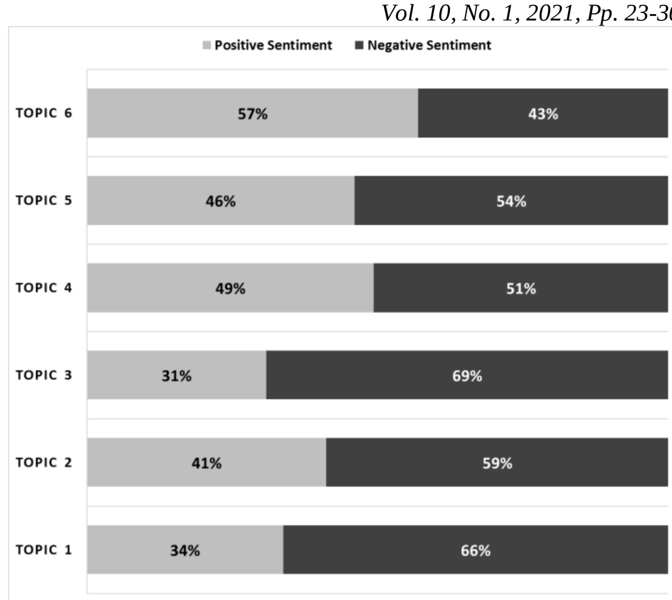

Figure 8. Sentiment proportion of each topic

Referring to the research objectives that have been described in the introduction section, this study has been successful in analyzing public health, topic modeling of Indonesian public conversations about the COVID-19 pandemic on Twitter, and sentiment over the topic category. This study revealed the dominant topic related to the COVID-19 and discovered the sentiment of each topic.

4 CONCLUSION

In this paper we addressed a number of issues discussed by Indonesian citizens on Twitter and revealed the corresponding sentiment related to COVID-19 pandemics. We collected a large number of tweets posted by Indonesian Twitter users and analyze those tweets with text mining and machine learning approach. We used four machine learning algorithms to identify sentiment. We then figured out the topics contained within our corpus datasets with statistical approach named Latent Dirichlet Allocation. The results from those machine learning and statistical approach are interpretable and understandable. On the third section, we discussed the successfully identified issue or topic category. From the discussion, we can conclude that most of Indonesian citizens, based on the Twitter data, have a bigger concern on economic issue rather than health-related issue. When they talked about government policy, the main concern was also on economic aspect rather than health aspect. Hence, there is an urgency for government or other responsible parties to increase the COVID-19 prevention campaign efforts, specifically to raise the level of awareness of Indonesian citizens on health aspect.

ACKNOWLEDGMENT

This study's data source is from median-analytics.com (a social media analysis platform) owned by the study center and data analysis service at Universitas Jenderal Achmad Yani Yogyakarta. The author would like to thank the Faculty of Engineering and Information Technology, Universitas Jenderal Achmad Yani Yogyakarta, for the support in the form of internal research funding assistance in the 2020 Applied scheme.

REFERENCES

[1] World Health Organization (WHO), "The Coronavirus Disease 2019 (COVID-19) Situation Report.," *WHO*, 2020. .

[2] WHO, "Coronavirus disease (COVID-19) Situation Report – 133," 2020.

[3] A. Jarynowski, M. Wojta-Kempa, and V. Belik, "Trends in Perception of COVID-19 in Polish Internet," *medRxiv*, p. 2020.05.04.20090993, May 2020.

[4] H. J. Larson, "The biggest pandemic risk? Viral misinformation," *Nature*, vol. 562, no. 7726, pp. 309–310, Oct. 2018.

[5] D. Allington, B. Duffy, S. Wessely, N. Dhavan, and J. Rubin, "Health-protective behaviour, social media usage, and conspiracy belief during the COVID-19 public health emergency," *Psychol. Med.*, pp. 1–7, 2020.

[6] L. Li *et al.*, "Characterizing the Propagation of Situational Information in Social Media during COVID-19 Epidemic: A Case Study on Weibo," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 2, pp. 556–562, Apr. 2020.

[7] L. (Lucy) Yan and A. J. Pedraza-Martinez, "Social Media for Disaster Management: Operational Value of the Social Conversation," *Prod. Oper. Manag.*, vol. 28, no. 10, pp. 2514–2532, Oct. 2019.

[8] M. Minicucci Ibañez, R. R. Rosa, and L. N. F. Guimarães, "Sentiment Analysis Applied to Analyze Society's Emotion in Two Different Context of Social Media Data," *Intel. Artif.*, vol. 23, no. 66, pp. 66–84, Dec. 2020.

[9] R. F. Sear *et al.*, "Quantifying COVID-19 Content in the Online Health Opinion War Using Machine Learning," *IEEE Access*, vol. 8, pp. 91886–91893, 2020.

[10] H. W. Park, S. Park, and M. Chong, "Conversations and medical news frames on twitter: Infodemiological study on COVID-19 in South Korea," *J. Med. Internet Res.*, vol. 22, no. 5, p. e18897, May 2020.

[11] M. O. Lwin *et al.*, "Global Sentiments Surrounding the COVID-19 Pandemic on Twitter: Analysis of Twitter Trends.," *JMIR public Heal. Surveill.*, vol. 6, no. 2, p. e19447, May 2020.

[12] H. Jelodar, Y. Wang, R. Orji, and H. Huang, "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach," *IEEE J. Biomed. Heal. Informatics*, pp. 1–1, Jun. 2020.

[13] J. Samuel, G. G. M. N. Ali, M. M. Rahman, E. Esawi, and Y. Samuel, "COVID-19 public sentiment insights and machine learning for tweets classification," *Inf.*, vol. 11, no. 6, pp. 1–23, 2020.

[14] "Kasus Covid-19 Pertama, Masyarakat Jangan Panik | Indonesia.go.id." .

[15] B. Liu, *Sentiment Analysis and Opinion Mining*, no. May. Morgan & Claypool Publishers, 2012.

[16] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1, 2008.

[17] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." 2010.

[18] A. Rane and A. Kumar, "Sentiment Classification System of Twitter Data for US Airline Service Analysis," in *Proceedings - International Computer Software and Applications Conference*, 2018, vol. 1, pp. 769–773.

[19] S. G. Kanakaraddi, A. K. Chikaraddi, K. C. Gull, and P. S. Hiremath, "Comparison Study of Sentiment Analysis of Tweets using Various Machine Learning Algorithms," in *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, 2020, pp. 287–292.

[20] D. M. Blei, "Probabilistic topic models," *Commun. ACM*, vol. 55, no. 4, pp. 77–84, Apr. 2012.

[21] I. Vayansky and S. A. P. Kumar, "A review of topic modeling methods," *Inf. Syst.*, vol. 94, p. 101582, Dec. 2020.

[22] Y. Hu, J. Boyd-Graber, B. Satinoff, and A. Smith, "Interactive topic modeling," *Mach. Learn.*, vol. 95, no. 3, pp. 423–469, Oct. 2014.

[23] Y. Zuo *et al.*, "Topic modeling of short texts: A pseudo-document view," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, vol. 13-17-August-2016, pp. 2105–2114.

[24] H. Jelodar *et al.*, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimed. Tools Appl.*, vol. 78, no. 11, pp. 15169–15211, Jun. 2019.

[25] D. Newman, H. Jey, Lau, K. Grieser, and T. Baldwin, "Automatic Evaluation of Topic Coherence," 2010.

[26] A. A. Anees, H. Prakash Gupta, A. P. Dalvi, S. Gopinath, and B. R. Mohan, "Performance analysis of multiple classifiers using different term weighting schemes for sentiment analysis," in *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, 2019, pp. 637–641.

[27] J. U. Kundale and N. J. Kulkarni, "Language independent multi-class sentiment analysis," in *Proceedings - 2019 5th International Conference on Computing, Communication Control and Automation, ICCUBEA 2019*, 2019.

[28] A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," in *Procedia Computer Science*, 2020, vol. 167, pp. 589–598.

[29] A. M. Rahat, A. Kahir, and A. K. M. Masum, "Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset," 2020, pp. 266–270.

[30] P. Pankaj, P. Pandey, M. Muskan, and N. Soni, "Sentiment Analysis on Customer Feedback Data: Amazon Product Reviews," in *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019*, 2019, pp. 320–322.

[31] A. Rahman and M. S. Hossen, "Sentiment Analysis on Movie Review Data Using Machine Learning Approach," in *2019 International Conference on Bangla Speech and Language Processing, ICBSLP 2019*, 2019.

[32] K. Stevens, P. Kegelmeyer, D. Andrzejewski, and D. Buttler, "Exploring topic coherence over many models and many topics," in *Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP)*, 2012, no. July, pp. 952–961.