International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Analysis of Twitter Media for Public Reaction Identification on COVID-19 Monitoring System using Hybrid Feature Extraction Method

Djoko Adi Widodo*[1], Nur Iksan[1], Budi Sunarko[1]

Submitted: 19/10/2022 **Accepted:** 31/12/2022

Abstract: Several strategies were implemented to prevent COVID-19 spread. However, these steps were not effectively implemented in the community due to low public awareness and lack of discipline in daily life and this indicated a potential threat of continuous exposure to the virus. It was also observed that the pandemic greatly affected other areas besides the health sector ranging from the social, political, religious, and economic aspects to the resilience of the people. These can be observed through direct observation of the community or activities of the people on social media, especially in relation to the socio-economic aspect. Therefore, this research was conducted using social media, specifically Twitter, via the Twitter API to obtain data related to COVID-19 pandemic in Indonesia. In this research, a sentiment analysis method was developed in this research to identify public opinion related to the spread of the COVID-19 virus and its social impact on society. This was achieved using the TF-IDF and Lexical methods for feature extraction and Naive Bayes for classification. This research used a dataset obtained through Twitter using the keyword "COVID-19" in Indonesian and manually labelled using 5 categories of reactions i.e., fear, angry, love, sad, and happy. The prediction accuracy values showed that the proposed TFBS method had a higher accuracy value of 0.85 compared to other methods. The performance was also evaluated by calculating the precision, recall, and F-score values for each extraction method used, and the proposed TFBS method was observed to have the highest values while ME had the lowest.

Keywords: *Sentiment Analysis, COVID-19, TF-IDF, Lexicon Based*

1. Introduction

World Health Organization (WHO) declared the COVID-19 virus a pandemic due to its spread to all countries in the world, thereby, becoming a serious concern for the public, government, and world health institutions [1]. The virus started emerging in Indonesia in early January 2020 and the data from the handling team showed an increasing trend with nearly 600,000 people reported to be exposed and approximately 18,000 died. Several strategies were implemented to prevent further spread such as washing hands, wearing masks, quarantine, social distancing, working from home, and others [2]. However, these steps were not effectively implemented in the community due to low public awareness and lack of discipline in daily life and this indicated a potential threat of continuous exposure to the virus. This led the government to introduce another strategy which involved launching a vaccination program for all elements of society to prevent infection.

This vaccination program was one of the effective efforts made to suppress the spread with the main goal of providing immunity against COVID-19 virus infection and ensuring the formation of herd or group immunity. Furthermore, the spread was also prevented by providing information to increase public awareness on the best way to conduct their activities during the period.

In addition to the health sector, the pandemic also significantly affected the economic, religious, security, resilience, political, and social sectors [3]. These were not identified directly as the case with the health sector but through the observation of community activities using different media such as Instagram, Facebook, Twitter, and others. These media serve as social sensors for communities through the uploading of ideas, thoughts, and opinions in the forms of text, video, photos, and audio. The contents are focused on activities and conditions of the things the people feel and see in their surrounding environment which are subsequently sha red on social media platforms. Meanwhile, sentiment analysis can be used to identify and predict future conditions using the information [4–6]. A sentiment analysis method was developed in this research to identify public opinion related to the spread of the COVID-19 virus and its social impact on society [7–9]. This was achieved using the TF-IDF and Lexical methods for feature extraction and Naive Bayes for classification.

[1] *Department of Electrical Engineering, Faculty of Engineering, Semarang State University – Indonesia*
[2] *Department of Electrical Engineering, Faculty of Engineering, Semarang State University – Indonesia*
[3] *Department of Electrical Engineering, Faculty of Engineering, Semarang State University – Indonesia*
* *Corresponding Author Email: djokoadiwidodo@mail.unnes.ac.id*

2. Related Works

Sentiment analysis has become an interesting topic and its methods have been widely developed, especially on Twitter concerning the COVID-19 spread. It was used by Machuca [10] to determine negative and positive sentiments through Machine Learning (ML) and Natural Language Processing (NLP) methods. The Logistic Regression Algorithm was used as an ML algorithm to classify feelings or emotions regarding the COVID-19 topic on Twitter in 2020. Meanwhile, the TF-IDF method was used to extract NLP features and the classification accuracy value was found to be 78.5%. Alabid & Katheeth [11] also developed a model framework to classify public sentiment and opinion regarding the implementation of the COVID-19 vaccination program. The research further compared 2 ML methods including Support Vector Machine (SVM) and Naive Bayes (NB) to classify sentiment polarity into positive, negative, and neutral. The experimental results showed that the NB method had a higher accuracy performance than SVM with 0.81 and 0.75 respectively. Srikanth et al., [12] also conducted a sentiment analysis using a combination of preprocessing methods and word embedding to perform feature extraction. The Deep Belief Neural Network (DBN) algorithm was used to classify tweets through a pseudo-labeling process and the experimental results showed that the DBN algorithm achieved an accuracy value of 90.3% which is better than the other methods. Moreover, Efrilianda [13] developed a sentiment analysis method to determine positive, negative, and neutral polarity using the Backpropagation Multi- Layer Perceptron (MLP) algorithm and the results showed an accuracy value of 70%. It was observed that these studies did not consider the development of a feature extraction method as an important factor in increasing the accuracy of sentiment classification. They also only classified sentiments into negative, positive, and neutral polarities instead of emotions and feelings. Therefore, this research developed a feature extraction method by combining TF-IDF with the lexical method. Sentiment analysis was also developed concerning the community's reactions to the COVID-19 pandemic with a focus on happy, sad, angry, fear, and love categories.

3. Method

Sentiment analysis is also known as opinion mining and it is normally used to process natural discussion or text commonly referred to as NLP to determine the sentiments embedded in an opinion. It usually categorizes the opinion sentences in the form of polarity which include Negative, Positive, and Neutral [9, 14]. It can also be used for classification in other forms such as happy, sad, angry, fear, and love. The steps in the sentiment analysis process are shown in the following Figure 1:

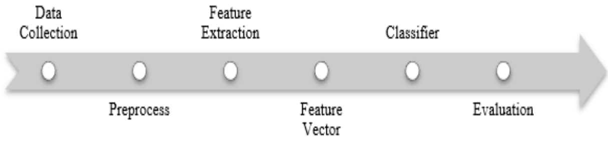

Fig. 1. The steps in the sentiment analysis process

The sentiment analysis consists of 6 stages of data collection, data pre-processing, feature extraction, vectorization, classification, and evaluation. The data collection stage was used to retrieve text data from Twitter through the API [15]. The data pre-processing was used to prepare the collected data for extraction into features [16] and this is important to determine the performance of the sentiment analysis process later. Overall, this stage is divided into two which include general pre-processing associated with the text contained in the tweet and special pre-processing related to the characteristics of the tweet text [17]. Figure 2 shows the stages of data pre-processing conducted in this research.

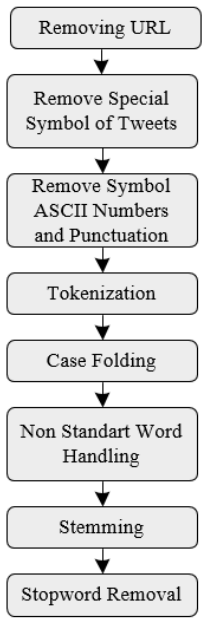

Fig. 2. Pre-processing stages

1. Removal URL by filtering the tweet text to remove URLs such as "HTTP://".
2. Removal of the special symbol in the tweets such as hashtags (#), retweets (RT), and usernames (@).
3. Tokenization such that the tweet text in the form of a sentence was be divided into several constituent words or phrases called tokens using the tokenize() function. The function was obtained from the RegexpTokenizer() package in the NLTK library.
4. Case folding which involved changing some words from the tokenization process to lowercase letters

using the lower() function.

5. Nonstandard word handling which involved returning the detected unstandardized words such as errors in the use of spelling, abbreviations, slang words, and elongated words to standard form.
6. The stemming process plays an important role in finding the root word of a derivative word. This process will remove affixes to a word. The method used in this process will be different in each language. In Indonesian usage, the stemming process uses the Literature library, while in English it uses the NLTK library.
7. Stopword Removal stage which involved removing common words that do not have a significant effect on the tweet sentence. This was achieved through the Literature and NLTK libraries.

The next stage is extraction which is very fundamental in the sentiment analysis process [18]. It involves converting the pre- processed text into vector form and was conducted in this research using the TF-IDF approach which normally extracts sentence features based on the frequency of word occurrence. Moreover, the frequency term was calculated by determining a feature based on the consistent appearance of a word. The features represent a document and have a high weight. In contrast to TF-IDF, the frequency of a word occurrence in a document is usually compared with an entire database. In a situation the word appears too often in several documents, it cannot be used as a feature because it is considered a general word that does not represent a specific document. The TF-IDF weight was calculated using the following Equations (1) and (2):

$$tf.idf_{t.d} = tf_{td} \times idf_t \tag{1}$$

$$idf_t = \log \frac{N}{df_t} \tag{2}$$

The lexical method was also applied in this research and it involves using the word meaning to determine the sentence features. The accuracy of the sentiment analysis is higher when the word meaning is more complete [19, 20]. The use of this method usually focuses on applying word lists such as “clean”, “healthy”, “good”, and “safe” to determine positive polarity and others such as “sloppy”, “bad”, “and slow” to determine the negative polarity. This feature extraction for the English lexicon was conducted using the library from the Google Translate API and the lexicon was translated first after which the sentiment scores were calculated from sentiwordnet using Equations (3) and (4).

$$Score = \sum \frac{(PosScore - NegScore)}{tot_{index}} \tag{3}$$

$$Sentence_{Score} = \sum Score \tag{4}$$

Meanwhile, the feature extraction of the Indonesian language lexicon was conducted by calculating the appearance of the word opinion in each sentence after which the positive and negative sentiments were determined using Equations (5) and (6).

$$Score = \sum frek_{opini_{word}} \tag{5}$$

$$Sentence_{Score} = \sum_{i=0}^{n} \sum Score_i \tag{6}$$

This means TF-IDF and Lexical methods which are statistical and semantic approaches respectively were combined in this research to determine the weighting of the feature extraction process. The proposed sentiment analysis model is presented in the following Figure 3.

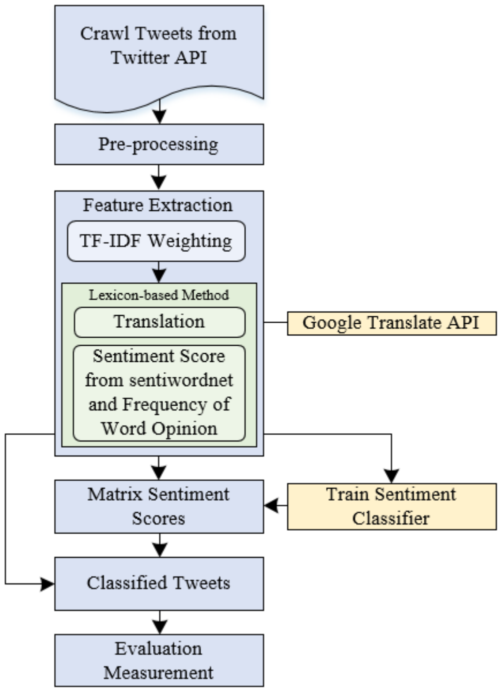

Fig. 3. Proposed Sentiment Analysis Model

At the classification stage, the Naive Bayes algorithm was used to determine the tweets consisting of 5 classes which include fear, angry, love, sad, and happy. Furthermore, the performance evaluation was conducted by comparing the Feature-Based Sentiment algorithm using TF-IDF and Lexicon (TFBS) with several other methods such as Maximum Entropy (ME), Feature-Based Sentiment (FBS), and Augmented Feature-Based Sentiment (AFBS).

Performance evaluation was conducted by calculating the accuracy, precision, recall, and F-Score. The accuracy value was calculated by comparing the number of correct data and the total amount of data. A higher accuracy value indicates a better method. Moreover, precision, recall, and F-score were measured to determine any deviations in the data. Precision value was determined by comparing the number of relevant classification data with the total amount of data from a particular class. The recall value was obtained from the comparison of the amount of relevant data with the total relevant data while the F-Score value was determined by calculating the average value of precision and recall. The equations used to calculate precision, recall, and f-score values are stated as Equations (7), (8) and (9).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (7)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (8)$$

$$F - Score = 2 * \frac{Precision * Recall}{Precission + Recall} \quad (9)$$

4. Result and Discussion

This research used a dataset obtained through Twitter using the keyword "COVID-19" in Indonesian. The dataset was retrieved through a crawling process using the Tweepy library in Python for as many as 3148 lines. It was subsequently passed through the normalization stage to eliminate duplicated data such as the tweets reposted by other users or retweets (RT). The dataset was manually labelled using 5 categories of reactions as shown in Table 1.

Table 1. Labelled Dataset using 5 categories of reactions

Text Tweet (Original)	Text Tweet (after translated in English)	Labels
Positif covid-19 bukan kutukan dan bukan aib. Terima kasih sudah berbagi dan tetap semangat..😭	Positive covid-19 is not a curse and not a disgrace. Thank you for sharing and keep the spirit. 😭	Happy
sedih banget kalo denger cerita perawat yang jadi satgas covid-19. banyak pasien covid yang bandel dan susah dibilangin ternyata	It's really sad to hear the story of the nurse who became the Covid-19 task force. many covid patients are stubborn and hard to say it turns out	Sad
gak respect banget.. ini ada sodara meninggal di RS padahal karena riwayat stroke, udah dimandiin dan diangkat pas beres malah keluar surat positif COVID-19?! Sumpah baru pertama kali ngeliat langsung kasus orang meninggal dicovidkan dan bener adanya 😂😂	really no respect.. this is a friend who died in the hospital even though because of a history of stroke, he was washed and picked up when it was done, instead a positive letter for COVID-19 came out?! I swear it's the first time I've seen a case of someone who died being COVID-19 and it's true	Angry
Besok mau di vaksin covid :(Hrusnya seneng ga sih Kan nntinya dpt perlindungan, jujur malah takut :(	Tomorrow I'm going to be vaccinated against covid :(Should I be happy or not, I'm going to get protection, to be honest I'm scared :(	Fear
Rindu suasana belajar mengajar sebelum covid mewabah di negri ini	Missing the teaching and learning atmosphere before the covid outbreak in this country	Love

Fig. 4. Example of a text tweet on Twitter

The tweets contain some useless information or noise such as hashtags (#), Html (http://www), mentions (@), numbers, local language usage, and abbreviations. The initial stage in the sentiment analysis process is the pre-processing which involves removing the URL, special symbol, punctuation, tokenization, case folding, and stemming as shown in Figure 2. The useless information was removed by omitting URLs, special symbols, ASCII numbers, and punctuations. Figure 4 shows an example of a tweet containing noise and this means there is a need for pre- processing. The tweet text in Figure 4 was pre-processed and the results are presented in the following Table 2.

Table 2. Cleansing at the Pre-processing Stage

Preprocessing stage	Original Tweet	Tweet in English
Tweet	" Orang-orang begitu mana paham covid-19, dijelasin berulangpun tetep ngeyel. What should i do? . . . Yaudah biarin aja,kalo positif apalagi sampe bergejala ga akan diurus. Simpan tenagamu untuk yang benar-benar membutuhkan ya nakes ☹. Kita kuat kita bisa survive"	" How do people understand covid-19, even when they explain it again and again, it's still annoying. What should I do? . . . Well, just let it be, if it's positive, especially if you have symptoms, it won't be taken care of. Save your energy for those who really need it, guys. We are strong we can survive"
Preprocessing Results	"orang paham covid jelas ulang udah biar saja positif gejala urus simpan tenaga butuh ya nakes kuat survive"	"people who understand that covid is being explained again, they keep yelling, what should i do, let's just leave it if it's positive until the symptoms don't take care of saving energy, we need strong health workers to survive"
Tokenization Result	[orang] [paham] [covid] [jelas] [ulang] [udah] [biar] [saja] [positif] [gejala] [urus] [simpan] [tenaga] [butuh] [nakes] [kuat] [survive]	[people] [understand] [covid] [clear] [repeat] [already] [let it be] [just] [positive] [symptoms] [manage] [save] [energy] [need] [nakes] [strong] [survival]

Several processes were implemented at the pre-processing stage and these include the deletion of emoticons, URLs, special symbols, ASCII symbols, and punctuation marks. Furthermore, the text characteristics were handled using different methods such as tokenization, case folding, non-standard word handling, stemming, and stop-word removal. The results were later forwarded to the feature extraction stage to produce a scoring matrix by combining the TF-IDF and lexicon methods.

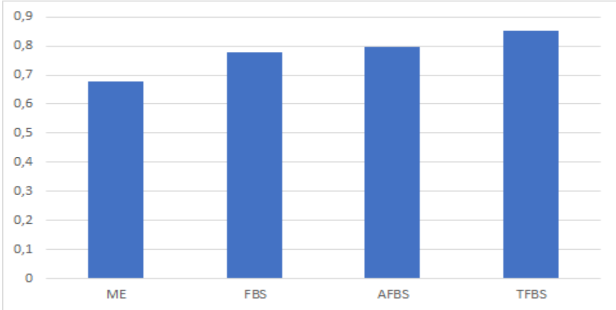

Fig. 5. The accuracy values obtained from using different feature extraction methods

Figure 5 shows the accuracy values obtained from using different feature extraction methods. It is important to note that the focus was placed on the tweet classification using five classes of fear, angry, love, sad, and happy. The results showed that the proposed TFBS method had a higher accuracy value of 0.85 compared to other methods.

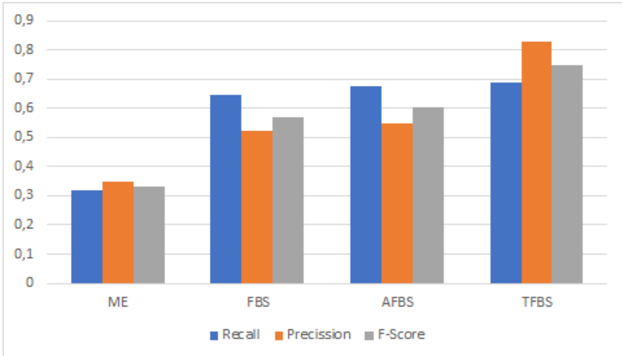

Fig. 6. Result of the Recall, Precision, and F-Score measurements for each method

As shown in Figure 6, the performance was also evaluated by calculating the precision, recall, and F-score values for each extraction method used, and the proposed TFBS method was observed to have the highest values while ME had the lowest.

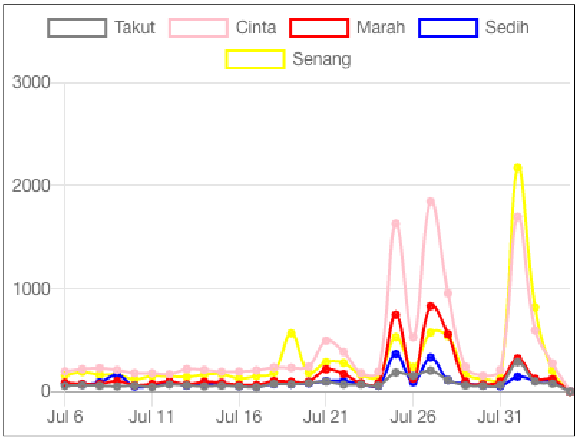

Fig. 7. Trend of public reactions in 06 July 2022 to 04 Aug 2022

The next experiment was to classify public reactions during the COVID-19 pandemic using the TFBS method with a focus on happy, sad, fear, angry, and love categories. It was discovered from the crawling analysis in Figure 7 that public reactions associated with "love" were higher than others. On July 24-31, 2022 the reactions to the pandemic were observed to be normal with the assumption that it could not attack Indonesia because of the ongoing vaccination program and the creation of herd immunity. The people thought those that were not disciplined with the implementation of the health protocols and refused vaccination were most susceptible to the infection. The "angry" class was found in second place on July 25-27 2022 due to the increase in COVID-19 cases in some areas that have returned to normal activities including schools that have been closed to implement online learning. It was also discovered that other viruses such as monkeypox were becoming an issue in these communities.

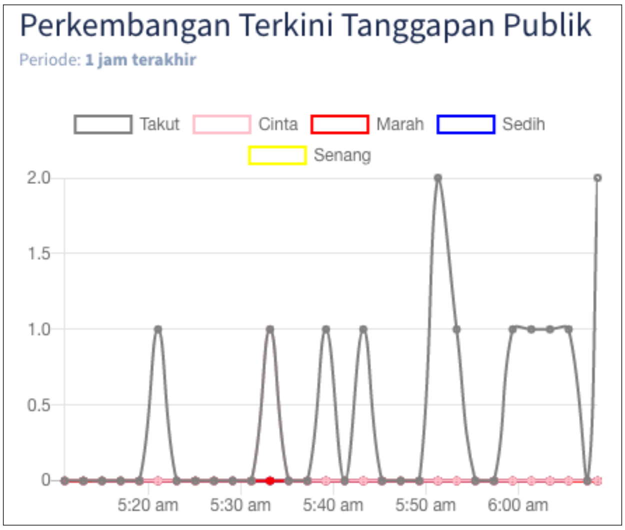

Fig. 8. Real- time on the dashboard of the monitoring system every 2 minutes in the last 1 hour

Moreover, in Figure 8, the trend of public reactions was also visualized in real- time on the dashboard of the monitoring system with a focus on every 2 minutes in the last 1 hour. It was discovered that the trend from 05.20 AM to approximately 06.00 tends more toward fear reactions. This simply indicates the general feeling of the people in the community regarding COVID-19.

The classifications are indicated by visualization in the form of a word cloud as shown in Figures 9 to 13. It was discovered that the words often observed in each class vary. For example, the frequent words in "happy" class include health, protocol, and discipline, those in "sad" are infected, sick, and fever, "love" had health, protocol, and discipline, "fear" recorded results, examination, and ministry of health while "angry" had infected, vaccinated, and positive.

Fig. 9. Popular topic of Word Cloud that associated with happy reaction

Figure 9 shows the trend of public reactions in happy category through the word cloud consisting of several frequent words that are widely discussed in the community and formed the opinion of the people. Some of these include protocol, health, masks, and discipline that were trending because, at the time, the enforcement of health protocol discipline was the main thing and had become an issue of public awareness.

Fig. 10. Popular topic of Word Cloud that associated with sad reaction

Figure 10 indicates the public reactions for the sad category are dominated by frequent words such as infected, sick, fever, positive, and cough. They were appearing often at the time because work and school activities of people were conducted offline. This led to several positive cases and other symptoms of COVID-19.

Fig. 11. Popular topic of Word Cloud that associated with love reaction

Figure 11 shows that the frequent words in the “love” category include health, protocol, discipline, progress, increase, deltacron, variant, and mask. These are almost the same as those observed in the happy category but the love aspect focuses on the hope for the better. During this period, new variants of COVID-19 began to appear such as omicron and deltacron but the public expresses their love while maintaining health protocols and discipline through the usage of masks.

Fig. 12. Popular topic of Word Cloud that associated with fear reaction

Figure 12 shows that the words trending in the fear category are dominated by the ministry of health, health, laboratory, examination, discipline, and protocol. The focus was placed more on the issue of an increase in COVID-19 patients, the emergence of new variants and the monkeypox virus during the period, and the need for the Ministry of Health to respond through appropriate policies. The public was always vigilant, afraid, and kept reminding each other to always be disciplined in maintaining health protocols.

Fig. 13. Popular topic of Word Cloud that associated with angry reaction

The trending words related to the fear category were found in Figure 13 to be dominated by words such as sick, affected, positive, death, and vaccine. The cause of these reactions was observed to be almost the same as the sad category.

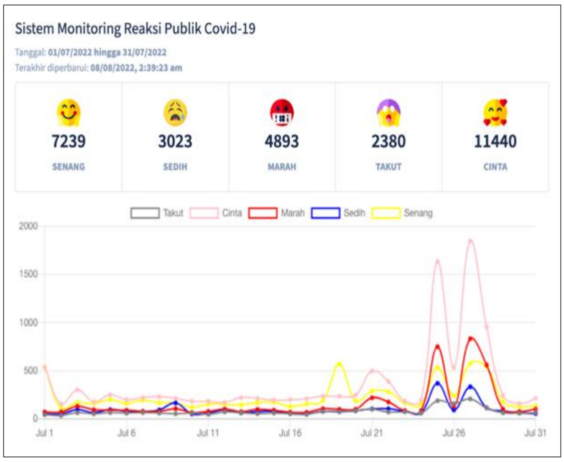

Fig. 14. Visualization of the COVID-19 public reaction monitoring system report

Some words in the word cloud from each category that are believed to have appeared due to their high frequency of occurrence in text tweets written by users. This means they reflect the opinion developed in the community. Figure 14 shows a visualization of the report from the COVID-19 public reaction monitoring system with a focus on the tweet recapitulation of each reaction category. It also indicated the graph of people's reactions in real-time which can also be determined at a certain time.

5. Conclusion

This research developed a sentiment analysis model using the TF-IDF and lexical methods to identify public reactions to COVID- 19. The prediction accuracy measurement showed that the proposed method was more accurate than the others due to its value of 0.854. It was also discovered from the Recall, Precision, and F-Score evaluation that the ME method had poor performance while the proposed TFBS had the best followed by AFBS and FBS. The TFBS also ranked higher compared to the other methods because it correctly identified and classified tweets. Moreover, the experiments conducted to classify the tendency of public reactions in dealing with the COVID-19 pandemic using the method showed the possibility of identifying and visualizing several categories on the dashboard of the community reaction monitoring system.

References

[1] R. P. Pratama and A. Tjahyanto, "The influence of fake accounts on sentiment analysis related to COVID-19 in Indonesia," in Procedia Computer Science, 2021, vol. 197, no. 2021, pp. 143–150, doi: 10.1016/j.procs.2021.12.128.

[2] A. D. Fakhira, A. S. Pawitra, K. C. Diyanah, E. A. Wikurendra, I. Nagy, and N. S. M. Abdeljawad, "Awareness of Doing 3M (Wearing Mask, Physical Distancing, Washing Hands) During Pandemic Era in Rural and Urban Families," J. Kesehat. Lingkung., vol. 13, no. 2, p. 94, 2021, doi: 10.20473/jkl.v13i2.2021.94-101.

[3] A. K. and M.-V. Dorothée Allain-Dupré, Isabelle Chatry and Michalun, "The territorial impact of COVID-19 :

Managing The Crisis Across Levels of Government," 2020. [Online]. Available: https://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-covid-19-managing-the-crisis-across-levels-of-government-d3e314e1/.
[4] L. Nahar, Z. Sultana, N. Iqbal, and A. Chowdhury, "Sentiment Analysis and Emotion Extraction: A Review of Research Paradigm," 1st Int. Conf. Adv. Sci. Eng. Robot. Technol. 2019, ICASERT 2019, no. December, 2019, doi: 10.1109/ICASERT.2019.8934654.
[5] Z. Drus and H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," Procedia Comput. Sci., vol. 161, pp. 707–714, 2019, doi: 10.1016/j.procs.2019.11.174.
[6] A. Pertiwi, A. Triayudi, and E. T. E. Handayani, "Sentiment Analysis of the Impact of Covid-19 on Indonesia's Economy through Social Media Using the ANN Method," J. Mantik, vol. 4, no. 1, pp. 605–612, 2020, [Online]. Available: https://iocscience.org/ejournal/index.php/mantik.
[7] P. Yadav and D. Pandya, "SentiReview: Sentiment Analysis based on Text and Emoticons," in International Conference on Innovative Mechanisms for Industry Applications, 2017, pp. 467–472.
[8] S. Peng, G. Wang, and D. Xie, "Social Influence Analysis in Social Networking Big Data: Opportunities and Challenges," IEEE Netw., vol. 31, no. 1, pp. 11–17, 2017, doi: 10.1109/MNET.2016.1500104NM.
[9] A. Go, L. Huang, and R. Bhayani, "Twitter Sentiment Analysis," 2009.
[10] C. R. Machuca, C. Gallardo, and R. M. Toasa, "Twitter sentiment analysis on coronavirus: Machine learning approach," J. Phys. Conf. Ser., vol. 1828, no. 1, 2021, doi: 10.1088/1742-6596/1828/1/012104.
[11] N. N. Alabid and Z. D. Katheeth, "Sentiment analysis of twitter posts related to the covid-19 vaccines," Indones. J. Electr. Eng. Comput. Sci., vol. 24, no. 3, pp. 1727–1734, 2021, doi: 10.11591/ijeecs.v24.i3.pp1727-1734.
[12] J. Srikanth, A. Damodaram, Y. Teekaraman, R. Kuppusamy, and A. R. Thelkar, "Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks," Comput. Intell. Neurosci., 2022, doi: https://doi.org/10.1155/2022/8898100.
[13] D. A. Efrilianda, E. N. Dianti, and O. G. Khoirunnisa, "Analysis of twitter sentiment in COVID-19 era using fuzzy logic method," J. Soft Comput. Explor., vol. 1, no. 1, pp. 1–5, 2021, doi: 10.52465/joscex.v2i1.12.
[14] T. N. Simanjuntak and S. Pramana, "Sentiment Analysis on Overseas Tweets on the Impact of COVID-19 in Indonesia," Indones. J. Stat. Its Appl., vol. 5, no. 2, pp. 304–313, 2021, doi: 10.29244/ijsa.v5i2p304-313.
[15] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," Artif. Intell. Rev., 2022, doi: https://doi.org/10.1007/s10462-022-10144-1.
[16] T. N. Prakash and A. Aloysius, "Data Preprocessing in Sentiment Analysis using Twitter Data," Int. Educ. Appl. Res. J., vol. 3, no. 7, 2019.
[17] A. F. Hidayatullah and M. R. Ma'arif, "Pre-processing Tasks in Indonesian Twitter Messages," J. Phys. Conf. Ser. Vol. 801, Int. Conf. Comput. Appl. Informatics - IOP Publ. Ltd, 2016, doi: 10.1088/1742-6596/801/1/012072.
[18] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The impact of features extraction on the sentiment analysis," Procedia Comput. Sci., vol. 152, pp. 341–348, 2019, doi: 10.1016/j.procs.2019.05.008.
[19] Kusrini and M. Mashuri, "Sentiment Analysis In Twitter Using Lexicon Based and Polarity Multiplication," 2019, doi: 10.1109/ICAIIT.2019.8834477.
[20] [N. Medagoda, S. Shanmuganathan, and J. L. Whalley, "Sentiment Lexicon Construction," 2015, [Online]. Available: 10.1109/ICNC.2015.7378094.