Sentiment Analysis of Urban Opinions on COVID-19 Handling in Brunei Darussalam Using Lexicon Weighting and Machine Learning Algorithms

[1]Usman Ependi, [2]Wahyu Caesarendra
[1]Faculty of Science Technology, Bina Darma University, Palembang, Indonesia
[2]Information Systems Department, Faculty of Computer Science, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam
[1]u.ependi@binadarma.ac.id, [2]wahyu.caesarendra@ubd.edu.bn

Abstract— Crisis management of Covid-19 is closely related to how government provides policy measures and monitors the health conditions of residents and others. Residents will provide feedback (opinions) for any services provided by the government. The main issue in this area is understanding residents' opinions to become a source of information for sentiment in public policy. This study aims to analyze sentiment on crisis management of covid-19 in Brunei Darussalam. Lexicon weight and machine learning classifiers (random forest, k-nearest neighbors, naive Bayes, and decision trees) are used for handling this issue. The data used in this study comes from resident opinions on the BruHealth application, which is part of Brunei Darussalam Government Services. Based on the experimental results, the sentiment of crisis management of Covid-19 is positive. Lexicon weight is used as a basis for data labelling in machine learning classification. Classification results using random forest, k-nearest neighbors, naive Bayes, and decision trees get a significant accuracy of 83,8%, 73,7%, 55%, and 84,2%, respectively.

Keywords— sentiment analysis, covid-19, lexicon, machine learning

I. INTRODUCTION

The coronavirus (COVID-19) is an infectious disease caused by SARS-CoV-2 that attacks the respiratory system, such as interstitial pneumonia and acute respiratory distress syndrome (ARDS) [1]. Citing Google statistics, from the emergence of Covid-19 in 2019 until December 2022, Covid-19 sufferers reached 435 million people, and 5.95 million died from this virus worldwide. In ASEAN, such as in Indonesia, there are 5.54 million people as sufferers, and 148,000 have died. Malaysia has 3.42 million people as sufferers, and 32,674 are worldwide. There are 2.89 million people suffering and 22,933 in the world.

Meanwhile, Brunei Darussalam has 267 thousand people as sufferers, and 225 have died. These statistics show that casualties in ASEAN countries, especially Indonesia, Malaysia, Thailand, and Brunei Darussalam, have a thigh gap, especially comparing sufferers and casualties for each country. Brunei Darussalam has the least number of sufferers and fatalities compared to other countries. This condition is inseparable from the Brunei Darussalam government's efforts to manage the spread of Covid-19.

Control and management of handling are fundamental problems in overcoming Covid-19. In Brunei Darussalam, control and management of handling Covid-19 have been carried out through the BruHealth application as a form of monitoring community activities. This condition also occurs in other ASEAN countries, such as Indonesia using the PeduliLindungi application, Malaysia using the MySejahtera application, and Thailand using the MorChana application. This condition shows the importance of handling Covid-19, as well as in Brunei Darussalam. However, the government's seriousness is certainly not only limited to how efforts are made but also the public's assessment and perception of the handling management to overcome the spread of Covid-19 itself.

Therefore, it is necessary to assist community assessment in handling Covid-19 to see the linearity between government efforts and citizen perceptions (opinions).

To see how citizen perceptions (opinions) can be used in sentiment analysis. This analysis motto is a process of studying public opinion towards an entity. Mining opinions can also be used in sentiment analysis terminology [2], [3]. Sentiment analysis allows extracting information from social media platforms used to interact. The role of sentiment analysis is also to find innovative information on social media. Therefore, social media has become a treasure of information [3]. This condition follows the media used by the Brunei Darussalam government, such as BruHealth, as a management media for handling Covid-19. Thus, there is the potential for more in-depth information mining related to public sentiment toward government policies.

The problem in sentiment analysis is understanding the words (phrases) in public opinion regarding handling covid-19. This understanding is important to explore so that sentiment determination can be done correctly. Various analytical studies of sentiments related to covid-19 have been carried out a lot, such as public sentiment towards vaccination [3], public opinion about vaccines [4], community pros and cons related to vaccines [5], and public sentiment related to covid-19 [6]. These various studies show that in determining the polarity of positive, neutral, and negative sentiments, using manual methods validated by linguists. One's knowledge strongly influences the determination of polarity in that way; ultimately, subjective judgments arise.

Based on the problems raised, this study aims to analyze the sentiment of handling COVID-19 in Brunei Darussalam using a lexicon weighting approach. A lexicon is a store of words in long-term memory related to grammar in composing phrases and sentences containing information such as parts of speech [6]. In terms of sentiment analysis, the lexicon acts as a comparison in determining sentiment outcomes through predetermined word weights. For this reason, this research uses two types of Lexicon SentiStrength [7] to determine sentiment. As for polarity classification, it uses machine learning algorithms random forest, k-nearest neighbors, naïve Bayes, and decision trees. The data used to see public sentiment regarding managing covid-19 is sourced from reviews (opinions) on the BruHealth application.

II. METHOD

This research stage consists of three stages: data preparation, sentiment weight, and machine learning classifier. Data preparation consists of data collection and preprocessing, as shown in Figure 1. Sentiment weight is the weighting process using the SentiStrength lexicon, and the machine learning classifier uses random forest, k-nearest neighbors, naïve Bayes, and decision trees.

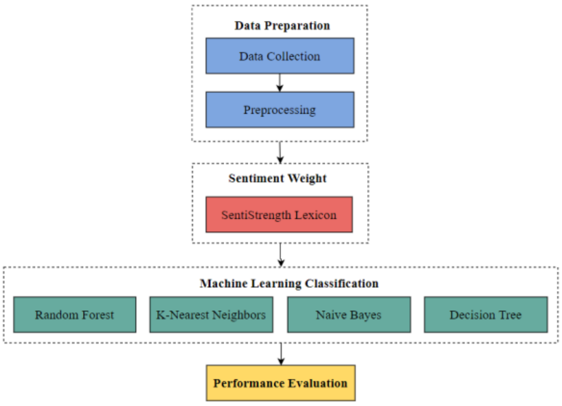


Figure 1. Proposed Methods

A. DATA PREPARATION

The data preparation stage is the stage for preparing data. This stage consists of two stages: data collection and preprocessing. Data collection in this study was through scraping the application for handling the Covid-19 crisis in Brunei Darussalam in the form of an opinion (review) on the BruHealth application with the ID: egnc.moh.bruhealth. The data collected with the newest filter was 1,457 raw data. The column used in this study is a content column that contains public opinions (reviews) related to covid-19 handling in Brunei Darussalam.

Data preprocessing is a process to make data easier to process so that it is easy for computers to understand [12]. The data preprocessing techniques [8] used in this study consisted of removing taq and duplicates, case folding, steaming, stopword, and normalization. The results of the data preprocessing obtain 1,177 rows of data from 1,457 raw data that were successfully crawled, as shown in Figure 2.

Figure 1. Data Preparation

B. SENTIMENT WEIGHT

Sentiment weight follows the rules of lexicon weight based on lexicon SentiStrength, as shown in Table 1 [9]. Weighting is done by looking at the words in the lexicon and comparing them with those in the data row. Furthermore, the accumulated weights become the final value for each row of data. Determination of data row to be positive, negative, or neutral based on Equation 1. Where if the weighting result is greater than zero (> 0), then it is positive; if the weighting result is less than zero (< 0), then it is negative, and if the weighting result is equal to zero (= 0), then it is neutral.

Table 1. SentiStrength Lexicon Word and Weight

No	Word	Weight
1	abandon	-2
2	abate	-2
3	abdicate	-2
4	abhor	-4
...	...	...
2843	yum	4
2844	yummy	4
2845	zest	2
2846	zestril	1

$$\text{sentiment}\begin{cases}\text{positive} = \text{weight} > 0\\ \text{negative} = \text{weight} < 0\\ \text{neutral} = \text{weight} = 0\end{cases} \quad (1)$$

C. MACHINE LEARNING

Machine learning is an algorithm designed to emulate human intelligence through environmental learning. Machine learning plays an important role in the era of big data and has been successfully applied to classification in many fields [10]. For this reason, the study used three machine learning algorithms for the sentiment classification process consisting of random forest (RF), k-nearest neighbors (KNN), Naive Bayes (NB), and Decision Trees (DS). The three algorithms are commonly used in machine learning classifications, including random forest for heart disease classification [11], k-nearest neighbors for Balinese script classification [12], naïve bayes for the selection of majors for students [13], and decision trees for anti-LGBT campaign classification [14]. Furthermore, the training scenario in the classification uses 90% of the data for training and 10% for testing. At the same time, the extraction feature uses the term frequency-inverse document frequency (TFIDF).

D. PERFORMANCE EVALUATION

Performance evaluation is the process of reviewing the results of the classification that has been carried out. In this study, we used confusion matric to assess whether the classification results performed well or vice versa. The matric provisions are TF (true negative), FP (false positive), TN (true negative), and FN (false negative) [15]. Based on these values, we evaluate performance by looking at accuracy, precision, recall, and f1-score. Calculations for each performance evaluation as shown in Equations 2 to 5 [16].

$$\text{Accuracy} = \frac{(TP+TN)}{(TP + FP + FN + TN)} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{F-Measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

ISSN 2442-7888 (online) DOI 10.24167/Sisforma

III. RESULTS AND DISCUSSION

Based on the stages of research that have been explained, starting from data preparation to performance evaluation, the research results can be explained based on two main components: sentiment distribution based on lexicon weighting and classification performance based on each type of machine learning algorithm.

A. SENTIMENT IN CRISIS MANGEMENT OF COVID-19

The determination of sentiment toward the crisis management of covid-19 handling in Brunei Darussalam is based on positive and negative words in each row of data. The positive and negative words are weighted based on the Lexicon SentiStrength as in Table 1. Sentiment determination can also be influenced by words appearing on each data row; every word in the data row has a weight. Thus, it will affect the sentiment result and sentiment weight. Figure 3 shows which frequently occur words most affect Sentiment weight.

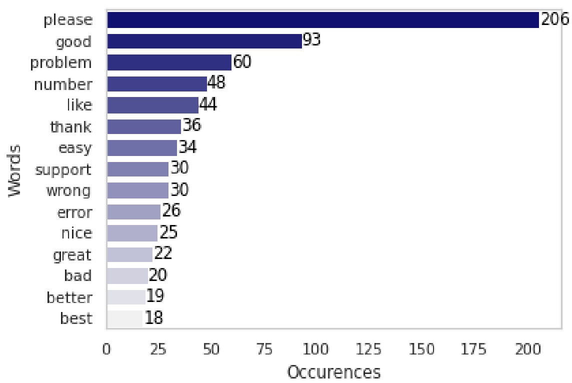
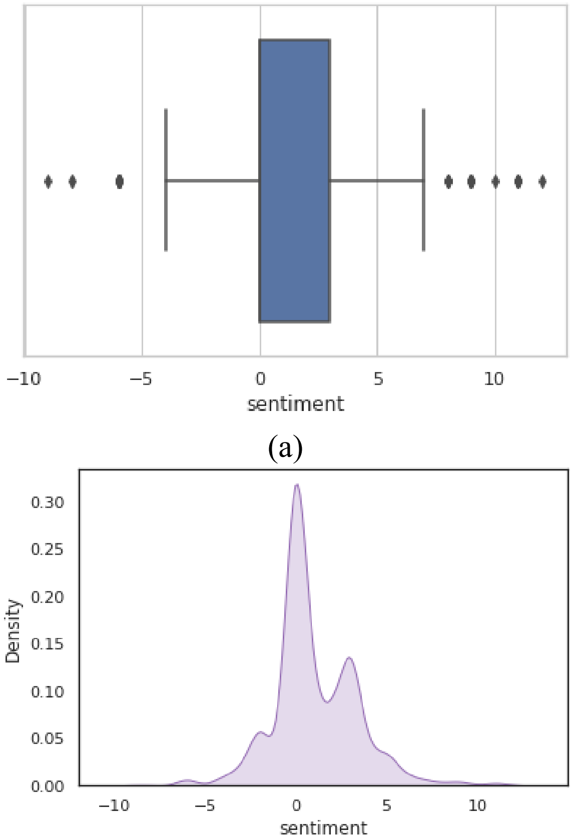


Figure 3. Top 15 Most Often Occured Words

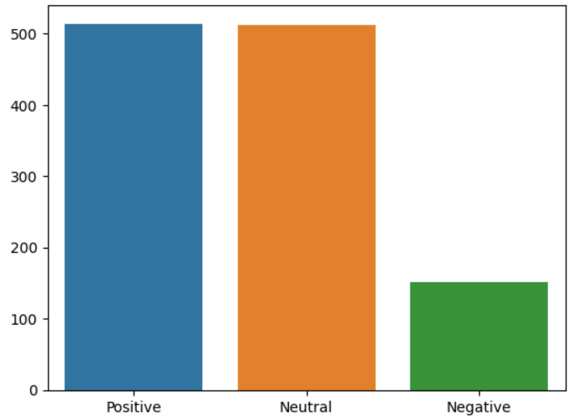
Based on the weight of the word lexicon SentiStrength, as in Table 1, the weighting process for each row of data is carried out. The weighting results show that the positive values for each data row start from 1 through 18, while the negative values start from -1 to -8, as shown in Figure 4(a). Meanwhile, the sentiment distribution based on lexicon weights shows neutral and positive, as shown in Figure 4(b).



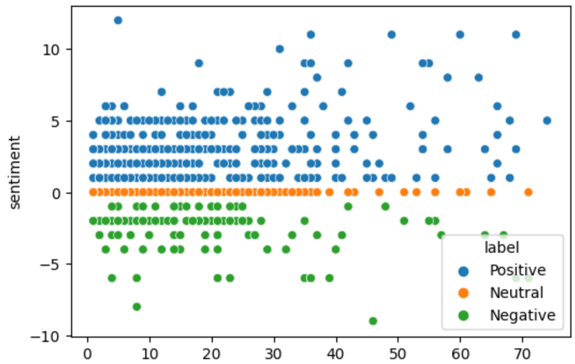
(b)

Figure 4. (a) Sentiment Weight and (b) Sentiment Distribution

The results of the sentiment distribution, as shown in Figure 4, show that public sentiment towards handling the covid-19 crisis management in Brunei Darussalam is positive. This condition is in line with statistical data on lower fatalities compared to other ASEAN countries such as Indonesia, Malaysia, and Thailand. To determine the sentiment's polarity based on the sentiment distribution results shown in Figure 4. Determination of sentiment polarity based on equation 1. Sentiment polarity results will be used for machine learning classification to see polarity accuracy performance. Figure 5 shows the polarity of public opinion sentiment (review) related to handling covid-19 crisis management in Brunei Darussalam through the BruHealth application. The results of this polarity also show a similarity between the value of sentiment and the polarity of sentiment itself.



(a)



(b)

Figure 5. Sentiment Polarity: (a) number of sentiment, (b) distribution of sentiment

B. CLASSIFICATION OF MACHINE LEARNING

In this study, four machine learning algorithms were utilized for classification: Random Forest, K-Nearest Neighbors (KNN), Naïve Bayes, and Decision Trees.

The classification was performed on data labeled based on sentiment polarity: positive, negative, and neutral as illustrated in Figure 5. The classification results demonstrated varying levels of accuracy and F1-scores across the different algorithms. The Random Forest algorithm achieved the highest accuracy, with 84% for training and 83.8% for testing. In contrast, the KNN algorithm recorded the lowest accuracy, with 55% for both training and testing.

When examining the F1-scores, the Decision Tree algorithm emerged as the top performer, with an average F1-score of 75.3%. The KNN algorithm, once again, lagged, with the lowest average F1-score of 49%. Notably, the Decision Tree algorithm also showed the smallest gap between training, F1-score, and testing performance, with only a 7% difference. This is in stark contrast to the Random Forest algorithm, which, despite its high accuracy, exhibited a 13% gap.

C. DISCUSSION

In this analysis of public sentiment toward COVID-19 crisis management in Brunei Darussalam, key insights were uncovered by examining positive and negative words weighted through the Lexicon SentiStrength method. The frequency and presence of specific words in each data entry significantly influenced sentiment scores, as shown in Figure 3, which highlights the top 15 impactful words. This demonstrates the critical role of word weighting in sentiment analysis and provides a comprehensive framework for understanding public opinion during crises.

The results from the weighting process, depicted in Figure 4(a), reveal a wide range of sentiment scores, with positive values between 1 and 18, and negative values from -1 to -8. This distribution indicates that while positive sentiments are prevalent, negative sentiments are present but less frequent. Figure 4(b) further illustrates this sentiment distribution, showing a dominance of neutral and positive sentiments, suggesting that the public generally holds a favorable view of the COVID-19 crisis management efforts in Brunei Darussalam.

The sentiment distribution results from Figure 4 align with statistical data indicating lower fatalities in Brunei compared to other ASEAN countries like Indonesia, Malaysia, and Thailand. This positive public sentiment likely reflects the effectiveness of Brunei's crisis management strategies. The combination of low fatality rates and effective pandemic handling has contributed to this overall positive sentiment, highlighting the importance of effective communication and robust health measures in shaping public opinion during health crises.

Further, the sentiment polarity results, derived from the sentiment distribution and presented in Figure 5, offer a deeper understanding of public opinion. By applying equation 1, sentiments were classified into positive, negative, and neutral categories. These polarity results not only align with overall sentiment values but also provide a nuanced perspective of public sentiment, validating the accuracy of the sentiment analysis approach used in this study.

The study also evaluated four machine learning algorithms—Random Forest, K-Nearest Neighbors (KNN), Naïve Bayes, and Decision Trees—for sentiment classification. Using data labeled by sentiment polarity, the performance of these algorithms was assessed, revealing significant differences in accuracy and F1-scores.

The Random Forest algorithm achieved the highest accuracy, with 84% for training and 83.8% for testing, demonstrating its ability to handle complex datasets effectively. However, the algorithm also showed a 13% gap between training, F1-score, and testing performance, indicating potential overfitting issues.

Conversely, the Decision Tree algorithm showed the best overall performance, with an average F1-score of 75.3% and a minimal performance gap of 7%. This consistency across training, F1-score, and testing metrics suggests a well-balanced model that effectively captures the nuances of sentiment data. The small performance gap indicates that the Decision Tree algorithm generalizes well to new data, making it a reliable choice for sentiment classification.

In contrast, the KNN algorithm had the lowest performance, with 55% accuracy for both training and testing, and an average F1-score of 49%. This poor performance suggests that KNN may not be suitable for sentiment classification in this context, likely due to its sensitivity to the choice of neighbors and the nature of the sentiment data.

Finally, the Decision Tree algorithm emerged as the most effective for sentiment classification in this study. Its consistent performance and minimal gap between training and testing metrics highlight its reliability and robustness. These findings suggest that the Decision Tree algorithm is a suitable tool for sentiment analysis in crisis management scenarios, providing valuable insights into public opinion and aiding in the development of effective communication strategies during health crises.

IV. CONCLUSION

This study presents the use of a lexicon-based approach for sentiment analysis to reduce human subjectivity in determining sentiment polarity. By applying the SentiStrength lexicon, the analysis of COVID-19 crisis management in Brunei Darussalam through the BruHealth application revealed predominantly positive sentiments. This positive sentiment indicates public approval of the handling measures.

The classification of sentiment polarity using machine learning algorithms demonstrated significant accuracy for the Random Forest, Naïve Bayes, and Decision Tree algorithms. However, the K-Nearest Neighbors (KNN) algorithm did not perform as well, which may be attributed to the imbalance in polarity data classes. This highlights a key area for future research: addressing class imbalance issues in sentiment polarity to improve the accuracy and reliability of sentiment classification models.

ACKNOWLEDGMENT

The heading of the Acknowledgment section and the References section must not be numbered.

REFERENCES

[1] F. Landi *et al.*, "Post-COVID-19 global health strategies: the need for an interdisciplinary approach," *Aging Clin. Exp. Res.*, vol. 32, no. 8, pp. 1613–1620, 2020, doi: 10.1007/s40520-020-01616-x.

[2] B. Liu, "Social Network Analysis," in *Web data mining: Exploring hyperlinks, contents, and usage data*, vol. 1, Heidelberg: Springer, 2011, pp. 269–309. doi: 10.1007/978-3-642-19460-3.

[3] D. Sharma, M. Sabharwal, V. Goyal, and M. Vij, "Sentiment analysis techniques for social media data: a review," in *First international conference on sustainable technologies for computational intelligence*, 2020, pp. 75–90.

[4] B. Laurensz and Eko Sediyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.

[5] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 2, pp. 1–9, 2021.

[6] A. Prihatini, "Semantic network of the word association in the field of law," *Litera*, vol. 18, no. 3, pp. 430–446, 2019.

[7] A. L. Fairuz, R. D. Ramadhani, and N. A. F. Tanjung, "Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 1, pp. 42–51, 2021, doi: 10.20895/dinda.v1i1.180.

[8] V. Chitraa and D. A. S. Davamani, "A Survey on Preprocessing Methods for Web Usage Data," *Int. J. Comput. Sci. Inf. Secur.*, vol. 7, no. 3, pp. 78–83, 2010, doi: 10.48550/arXiv.1004.1257.

[9] S.-U. Hassan *et al.*, "Predicting literature's early impact with sentiment analysis in Twitter," *Knowledge-Based Syst.*, vol. 192, p. 105383, 2020, doi: https://doi.org/10.1016/j.knosys.2019.105383.

[10] I. El Naqa and M. J. Murphy, "What Is Machine Learning?," in *Machine Learning in Radiation Oncology*, Cham: Springer International Publishing, 2015, pp. 3–11. doi: 10.1007/978-3-319-18305-3_1.

[11] P. R. Togatorop, M. Sianturi, D. Simamora, and D. Silaen, "Optimizing Random Forest using Genetic Algorithm for Heart Disease Classification," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 13, no. 1, p. 60, 2022, doi: 10.24843/lkjiti.2022.v13.i01.p06.

[12] I. W. A. S. Darma, "Implementation of Zoning and K-Nearest Neighbor in Character Recognition of Wrésastra Script," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 10, no. 1, p. 9, 2019, doi: 10.24843/lkjiti.2019.v10.i01.p02.

[13] A. Saleh and F. Nasari, "Implementation Equal-Width Interval Discretization in Naive Bayes Method for Increasing Accuracy of Students' Majors Prediction," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 9, no. 2, p. 104, 2018, doi: 10.24843/lkjiti.2018.v09.i02.p05.

[14] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, 2019, doi: https://doi.org/10.1016/j.procs.2019.11.181.

[15] P. R. Togatorop and A. Fauzi, "Klasifikasi Penggunaan Masker Wajah Menggunakan Squeezenet," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, pp. 397–406, 2022, doi: 10.35957/jatisi.v9i1.642.

[16] S. Devella, Y. Yohannes, and F. N. Rahmawati, "Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 2, pp. 310–320, 2020, doi: 10.35957/jatisi.v7i2.289.