Green Intelligent Systems and Applications

Research Article
Volume 5(1), 2025, 51–60
https://doi.org/10.53623/gisa.v5i1.588

Twitter Sentiment Analysis of Mental Health Issues Post COVID-19

Panca Dewi Pamungkasari[1], Sari Ningsih[1], Achmad Pratama Rifai[2]*, Alisyafira Sayyidina Nandila[1], Huu Tho Nguyen[3], Sathish Kumar Penchala[4]

[1]Faculty of Communication and Information Technology Universitas Nasional, Jakarta, Indonesia
[2]Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3]Faculty of Engineering and Technology, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
[4]Indore Institute of Science and Technology, Indore, India

*Correspondence: achmad.p.rifai@ugm.ac.id

SUBMITTED: 16 January 2025; REVISED: 11 March 2025; ACCEPTED: 24 March 2025

ABSTRACT: The Coronavirus Disease 2019 (COVID-19) impacted many aspects of daily life, including mental health, as some individuals struggled to adjust to the rapid changes brought on by the pandemic. This paper investigated sentiment analysis of Twitter data following the COVID-19 pandemic. Specifically, we analyzed a large corpus of tweets to understand public sentiment and its implications for mental health in the post-pandemic context. The Naïve Bayes and Support Vector Machine (SVM) classifiers were used to categorize tweets into positive, negative, and neutral sentiments. The collected tweet data samples showed that 38.35% were neutral, 32.56% were positive, and 29.09% were negative. Results using the SVM method showed an accuracy of 84%, while Naïve Bayes achieved 80% accuracy.

KEYWORDS: Mental health; COVID-19; sentiment analysis; Twitter; Na¨ıve Bayes; support vector machine

1. Introduction

The Coronavirus Disease 2019 (COVID-19), which was declared a pandemic by the World Health Organization (WHO) in March 2020, had a major impact on daily life. The effects of COVID-19 not only harmed physical health but also affected mental well-being [1, 2]. Good mental health was defined as the ability to cope with life's normal pressures while remaining productive [3]. Conversely, mental health problems such as depression, stress, and anxiety increased as a result of the COVID-19 pandemic [4]. The WHO reported that an estimated 5% of adults suffered from depression [1]. Unlike physical health disorders that could be easily detected, mental disorders were more difficult to identify because they were invisible. This often resulted in delayed treatment for individuals with mental health conditions. In addition, the lack of public awareness regarding mental health and the stigma surrounding mental disorders discouraged affected individuals from seeking help at existing health facilities [5]. To cope with these psychological effects, some people shared their concerns, offered advice, and exchanged information about mental health on various social media platforms, including Twitter [6, 7].

As of April 2024, the number of Twitter users in Indonesia reached approximately 24.85 million [2]. Through Twitter, individuals shared their personal experiences on various topics, including mental health. Many users expressed their emotions, described their actions, and shared opinions about their experiences with depression, making it one of the most widely discussed mental health issues. In this paper, a sentiment analysis was conducted on Twitter data regarding mental health issues in Indonesia after the COVID-19 pandemic. In the classification task, we evaluated accuracy values using the Naïve Bayes (NB) and Support Vector Machine (SVM) algorithms. The NB method was widely used in research due to its simplicity and effectiveness across various domains. Meanwhile, the SVM method was employed as a margin-based classification technique for distinguishing between two different classes.

To the best of our knowledge, research on predicting mental health conditions after COVID-19 remained limited. Most existing studies focused on sentiment analysis before and during the pandemic [8, 9]. Therefore, this research contributed in two ways. First, it focused on predicting people's mental health after COVID-19 using Twitter data. Second, to achieve this goal, a Twitter dataset from Indonesia—the largest country in Southeast Asia—was constructed. The remainder of this paper is structured as follows: Section II discusses the methodology, Section III presents the results and discussion, and Section IV concludes the paper.

2. Research Method

Details of the research methodology are illustrated in Fig. 1. The first step involved collecting data from Twitter, followed by data processing and classification using the Naïve Bayes and SVM algorithms. Finally, the classified data were evaluated to assess the performance of the machine learning models used in this study.

2.1. Data collection.

The data used in this study were collected from Twitter using Tweet-Harvest, a tool that crawls Twitter data through the Application Programming Interface (API). One key advantage of Tweet-Harvest is its ability to retrieve a large volume of data while requiring only an authentication token.

Algorithm 1 Tweet Harvesting Pseudocode

1: BEGIN
2: Define filename
3: Define search_keyword '"*mental health*" OR "*depre- sion*" lang:id until:2023-06-26'
4: Define limit as 500
5: Define command as "npx –yes tweet-harvest@2.2.8 -o
{filename} -s {search_keyword} -l {limit} –token
{twitter_auth_token}"
6: Execute command
7: END = 0

Data collection was conducted from June to October 2023 by filtering Indonesian-language keywords, specifically "mental health" (*kesehatan mental*) and "depression" (*depresi*). To optimize efficiency, the data were collected in multiple batches over five months rather than retrieving them all at once. Using the program code outlined in Algorithm 1, the collected tweets were stored in a CSV file format.

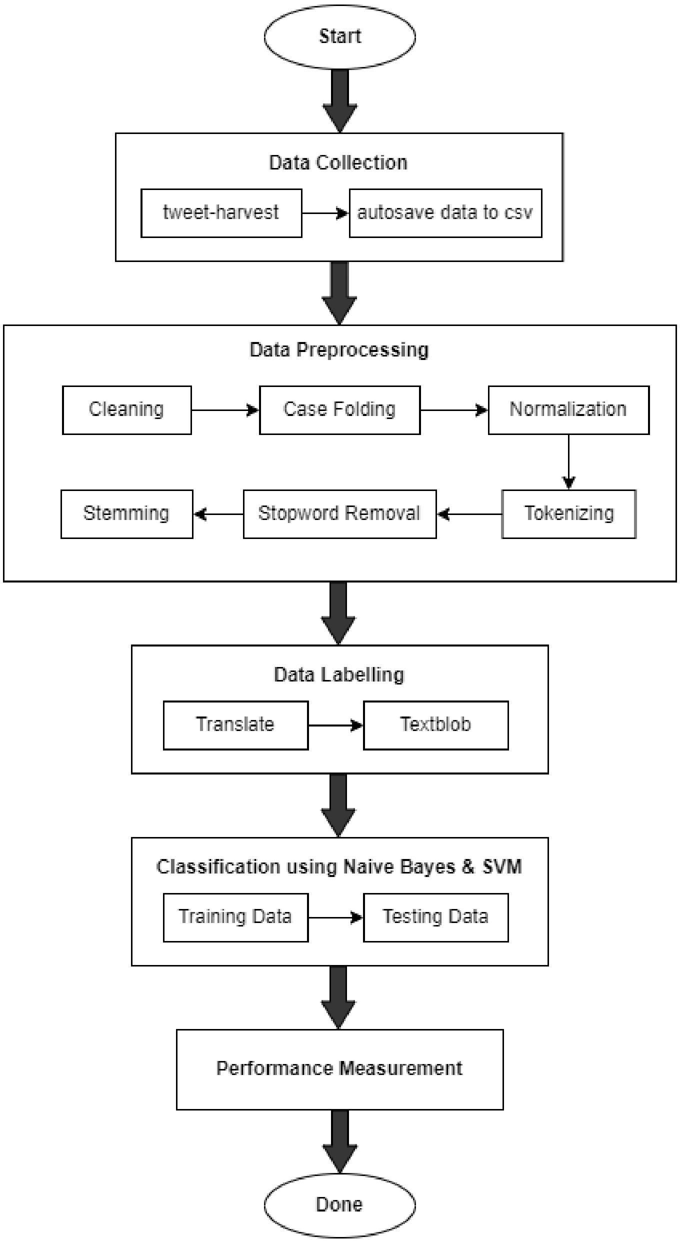

Figure 1. Research Method.

2.2. Data preprocessing.

Several steps were used to preprocess the data collected from Twitter. Data preprocessing was performed to eliminate irrelevant information or transform data into a format that is easier for the system to process. The preprocessing stages included filtering, case folding, normalization, tokenization, stopword removal, and stemming. Filtering is the process of cleaning the characteristics of user names, hashtags, URLs, punctuation, and symbols. Case Folding is performed to change all capital letters in the text to lowercase. Normalization changes irregular words into regular words. Tokenization is performed to divide each word in a sentence into its own word units. Stopword removal is the process of selecting or filtering words in text that have no meaning, such as conjunctions and adverbs, which are not needed in data modeling. Stemming is performed to change suffixed words into basic words.

2.3. Data labeling.

After data collection and preprocessing, the next step was data labeling. Labeling was performed to categorize the data into three sentiment classes: positive, negative, and neutral. The labels were assigned based on polarity scores calculated using TextBlob. TextBlob utilizes an English-language lexicon model, such as WordNet, which is significantly more advanced for sentiment analysis compared to models in other languages. Therefore, the dataset needed to be translated into English before classification. Once translated, each tweet was classified based on its sentiment polarity [10].

2.4. Na¨ıve Bayes (NB).

Naïve Bayes (NB) is a machine learning algorithm that applies probability calculations using the Bayesian approach. Bayes' theorem in the NB algorithm combines prior probability and conditional probability into a formula that calculates the likelihood of each possible classification [11]. In other words, it is a classification technique that integrates probability and statistical methods with a "naïve" assumption that attributes are independent. The NB algorithm demonstrates high accuracy and **is** easy to train, making it efficient for large datasets with low error rates [12]. Key advantages of this classification method include its ability to quickly compute probabilities and its reliance on probabilistic hypotheses [13].

Let (x_i, y_i) be the data sample pair, where x_i is the i-th feature vector which is conditionally independent of each other given the class $y_i \in \{C_1, \cdots, C_j, \cdots, C_{|C|}\}$. The probability of predicting class C_j is given by [14].

$$P\ (C_j \mid \mathrm{x}) = \frac{P(C_j)_{i=1}^{IT_n} P(\mathrm{x}_i \mid C_j)}{P(\mathrm{x})} \tag{1}$$

where n is the number of data samples. The Bayes optimal classifier is given by:

$$y^i = argmax_{C_j} P\big(C_j\big)_i^{IT} P\ (\mathbf{x}_i \mid C_j) \tag{2}$$

2.5. Support vector machine (SVM).

SVM method is a type of supervised learning machine learn- ing that was introduced by Cortes and Vapnik in 1995, after a series of development [15]. The SVM creates a boundary that separates two classes. The linear classifier that separates class +1 and class −1 is given by f (x) = $\mathrm{w}^T \mathrm{x} + b$, where x is the weight vector, b is the bias, and $(\cdot)^T$ is the transpose operator.

Further development of the SVM considers the trade off between the margin and the number of mistakes on the training data. Therefore, slack variables ξ are introduced. The optimization problem now becomes:

$$min_{\mathrm{w},\xi\mathrm{i}} = ||\mathbf{x}||^2 + c_i^n \xi_i, \tag{3}$$

$$\mathrm{s.t.}\, y_i(\mathrm{w}^T \mathrm{x}_i + b) \geq 1 - \xi_i, \tag{4}$$

where c is the regularization parameter. Small c value makes large margin. Note that for multi-class classification task, either One-vs-One (OVO) or One-vs-All (OVA) configuration was used. Kernel strategy can be used to maximize non-linear sepa- ration between classes. Some kernel functions that were commonly used in the SVM include linear, polynomial, radial basis function, and sigmoid. In this paper, OVO configuration and linear kernel were used.

2.5. Performance evaluation.

Model evaluation for $|C|$-class classification task was per- formed using various macro average metrics, including accu- racy, precision, recall, and F1 score, which are mathematically defined, respectively, by:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{5}$$

$$\mathrm{Pr} = \frac{1^{|C|}}{|C|_{i=1}} \frac{TP_i}{TP_i + FP_i}, \tag{6}$$

$$Re = -\frac{1^{|C|}}{|C|_{i=1}} \frac{TP_i}{TP_i + FN_i}, \tag{7}$$

$$F1 = \frac{2\, x\, \mathrm{Pr}\, x\, Re}{Pr + Re}, \tag{8}$$

where TP , TN , FP , and FN are true positive, true negative, false positive, and false negative samples, respectively.

In addition, Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC) can also be used to quantify the performance of machine learning models. The ROC plots the True Positive Rate (TPR), also known as recall, as a where n is the number of data samples. The Bayes optimal function of False Positive Rate (FPR) which is given by:

$$FPR = -\frac{1^{|C|}}{|C|_{i=1}} \frac{FP_i}{FP_i + TN_i} \tag{9}$$

Note that AUC defines the ability of a model to distinguish between classes.

3. Results and Discussion

3.1. Dataset labeling.

Data from Twitter was collected using the keywords "mental health" and "depression" (in the Indonesian language), resulting in 2,506 tweets. As explained earlier, the data was preprocessed, and the comment labeling process was carried out using TextBlob. Sentiment was assessed based on the words contained in a tweet. A tweet was labeled as positive if the sentiment score column showsed a value greater than 0 (value > 0), A tweet was labeled as negative if the value was less than 0 (value < 0), A tweet was labeled as neutral if the value was equal to 0 (value = 0). Table 1 presents examples of tweets that were labeled as positive, negative, and neutral. Figure 2 shows the label distribution of the obtained tweet data. It can be seen that 961 tweets were classified as neutral, 816 tweets were classified as positive, and 729 tweets were classified as negative. The final labeled dataset was processed using machine learning, specifically the Naïve Bayes and SVM algorithms.

Table 1. Sentiment analysis of tweets.

Tweet	Sentiment score	Sentiment
I hope my depression won’t last long and my life will end happily.	0.15	Positive
Praying to get through the depression phase when I reach the lowest point in life.	0.0	Neutral
Life is messed up because of people, I’m depressed and sad.	-0.35	Negative

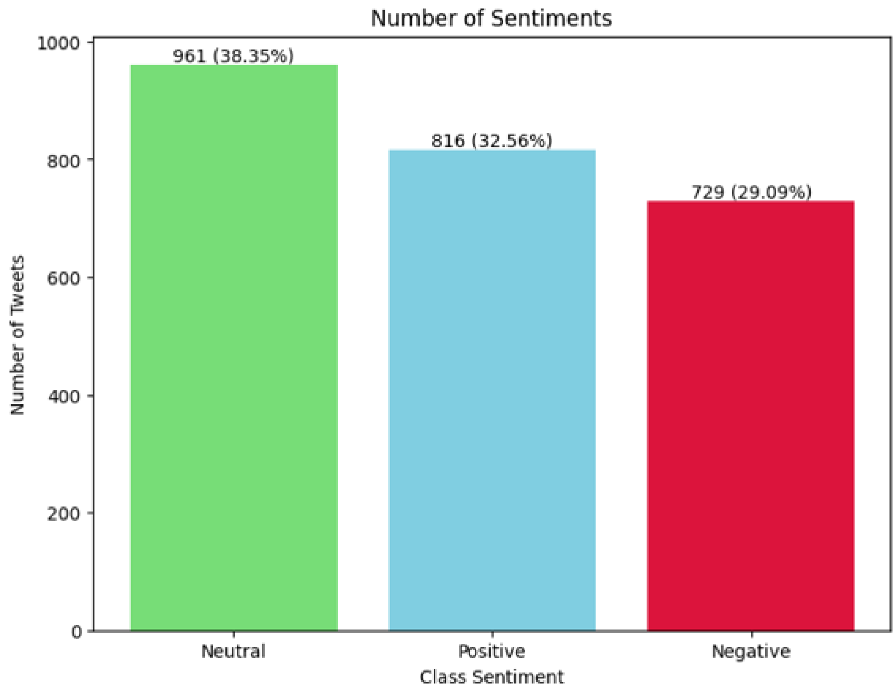

Figure 2. Label distribution.

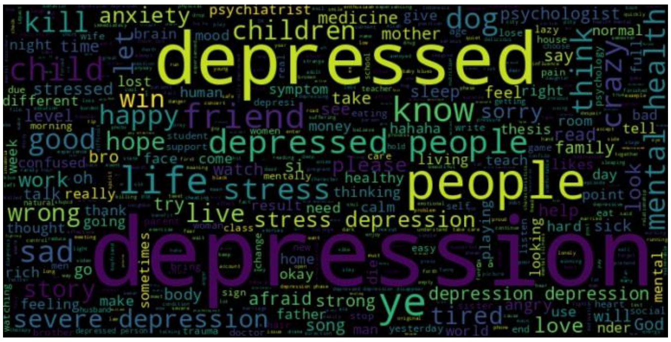

Figure 3. Wordcloud sentiment of mental health.

Figure 3 shows a visualization in the form of a word cloud. The word cloud displays words that were frequently used by Twitter users when commenting on mental health after COVID-19, making it easier to identify commonly appearing terms. Based on the word cloud results, several words were frequently used, such as "depression," "people," and "life." This demonstrates that mental health has become a topic of discussion within the Indonesian community following COVID-19.

3.2. Result.

In the classification task, the dataset was first divided into two parts: training data and testing data, with a ratio of 80% for training and 20% for testing. Table 2 shows the details of the training and testing datasets.

Table 2. Total of training and testing data.

Data	Ratio (%)	Total
Training Data	80	2004
Testing Data	20	502
Total	100	2506

Fig. 4 shows the confusion matrices for the Naïve Bayes and SVM algorithms. As seen in Fig. 4a, the Naïve Bayes classifier model correctly predicted 115 positive labels (true positives), 18 false positive labels (false positives), and 15 neutral positive labels (neutral positives), with a total of 148 positive label predictions. Additionally, there were 22 correct negative label predictions (true negatives), 116 incorrect negative label predictions (false negatives), and 14 neutral negative label predictions, with a total of 152 negative label predictions. Meanwhile, there were 19 correct neutral label predictions (true neutrals), 14 incorrect neutral label predictions (false neutrals), and 169 neutral label predictions (neutral neutrals), with a total of 202 neutral label predictions.

Based on Fig. 4b, the SVM classifier model correctly predicted 122 positive labels (true positives), 18 false positive labels (false positives), and 9 neutral positive labels (neutral positives), with a total of 149 positive label predictions. There were 15 correct negative label predictions (true negatives), 120 incorrect negative label predictions (false negatives), and 11 neutral negative label predictions, with a total of 146 negative label predictions. Meanwhile, there were 19 correct neutral label predictions (true neutrals), 10 incorrect neutral label predictions (false neutrals), and 178 neutral label predictions (neutral neutrals), with a total of 207 neutral label predictions.

Table 3. Performance evaluation Naive Bayes.

Predicted / Actual	Precision	Recall	F1-Score	AUC
Positive Label	78%	74%	76%	88%
Neutral Label	84%	85%	84%	91%
Negative Label	76%	78%	77%	87%
Macro Average	79%	79%	79%	
Weighted Average	80%	80%	80%	
Accuracy	80%			

Tables 3 and 4 summarize the performance of the Naïve Bayes and SVM models, respectively. For the Naïve Bayes classifier model, it can be seen that for positive labels, the precision value was 78%, the recall value was 74%, the F1-score was 76%, and the AUC value was 88%. For negative labels, the precision value was 76%, the recall value was 78%, the F1-score was 77%, and the AUC value was 87%. Moreover, for neutral labels, the precision value was 84%, the recall value was 85%, the F1-score was 84%, and the AUC value was 91%. Overall, the accuracy value was 80%.

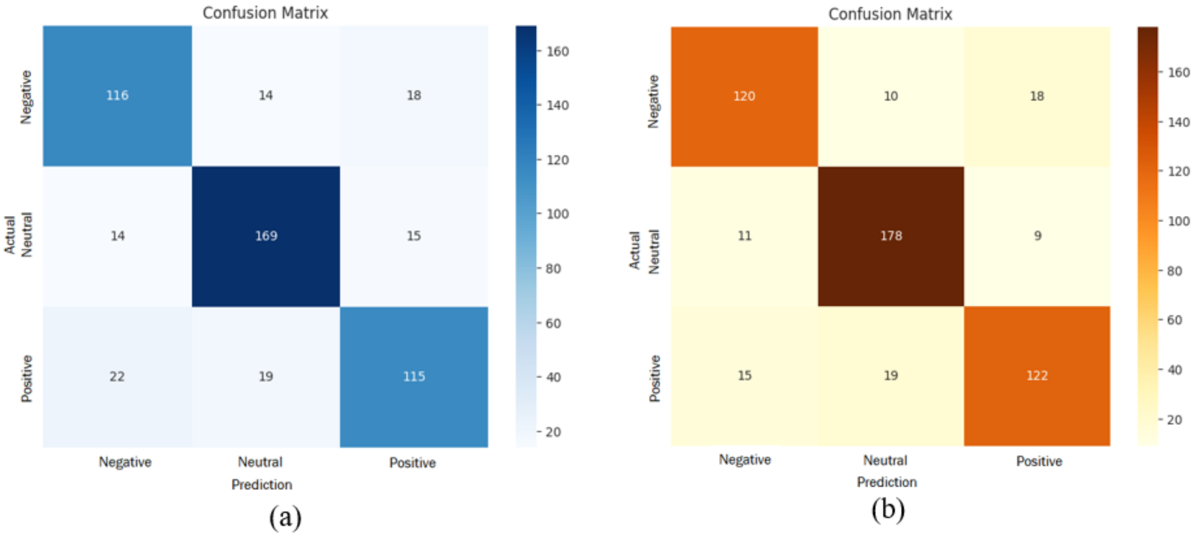

Figure 4. Confusion Matrix Naıve Bayes (a); Confusion Matrix SVM (b).

Table 4. Performance evaluation SVM.

Predicted / Actual	Precision	Recall	F1-Score	AUC
Positive Label	82%	78%	80%	94%
Neutral Label	86%	90%	88%	95%
Negative Label	82%	81%	82%	92%
Macro Average	83%	83%	83%	
Weighted Average	84%	84%	84%	
Accuracy	84%			

Meanwhile, for the SVM classifier model, it can be seen that for positive labels, the precision value was 82%, the recall value was 78%, the F1-score was 80%, and the AUC value was 94%. For negative labels, the precision value was 82%, the recall value was 81%, the F1-score was 82%, and the AUC value was 92%. Furthermore, for neutral labels, the precision value was 86%, the recall value was 90%, the F1-score was 88%, and the AUC value was 95%. The overall accuracy value was 84%.

It can be seen that SVM performed better than Naïve Bayes in sentiment analysis with a dataset of tweets from Twitter, as it provided more accurate and precise predictions. SVM outperformed Naïve Bayes in Twitter sentiment analysis due to its ability to handle highdimensional text data efficiently. SVM was also more robust against noisy and imbalanced datasets, which are common in Twitter sentiment data. It effectively managed slang, emojis, and abbreviations, which often misled Naïve Bayes. With kernel functions, SVM adapted to complex decision boundaries, improving classification accuracy. These advantages made SVM a superior choice for Twitter sentiment analysis, achieving higher precision and recall than Naïve Bayes.

4. Conclusions

In this research, we conducted a sentiment analysis using the Naïve Bayes classifier and SVM to determine Twitter users' responses to the keywords *mental health* and *depression*. The accuracy obtained was 84% for SVM, while the Naïve Bayes algorithm achieved an accuracy of 80%. From these results, the predominant sentiment was neutral, reflecting a balance of positive and negative opinions from the public concerning mental health issues in Indonesia in the post-COVID-19 era. It is hoped that the results of this research will serve as a reference for the public to understand the overall sentiment, the stigma surrounding mental health issues in society, and public attitudes toward maintaining mental health after the COVID-19 pandemic in Indonesia. Additionally, the findings from this sentiment analysis can be used as evaluation and advocacy material for relevant stakeholders, such as the government, health institutions, the media, and the general public, to improve mental health policies and address the challenges faced by society.

Author Contibution

Panca Dewi Pamungkasari: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing; Sari Ningsih: Conceptualization, Formal analysis, Investigation, Methodology, Experiment, Validation, Visualization, Writing - original draft,; Achmad Pratama Rifai: Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing - original draft, Writing - review & editing; Alisyafira Sayyidina Nandila: Experiment, Validation, Writing - original draft, Writing - review & editing, Nguyen Huu Tho: Resources, Writing - review & editing, Sathish Kumar Penchala: Resources, Writing - review & editing,

References

[1] Birmingham, W.C.; Wadsworth, L.L.; Lassetter, J.H.; Graff, T.C.; Lauren, E.; Hung, M. (2023). Covid-19 lockdown: Impact on college students' lives. *Journal of American College Health*, *71*(3), 879–893. https://doi.org/10.1080/07448481.2021.1909041.

[2] Mastrorosa, I.; Del Duca, G.; Pinnetti, C.; Lorenzini, P.; Vergori, A.; Brita, A.C.; Camici, M.; Mazzotta, V.; Baldini, F.; Chinello, P.; Mencarini, P.; Giancola, M.L.; Abdeddaim, A.; Girardi, E.; Vaia, F.; Antinori, A. (2023). What is the impact of post-covid-19 syndrome on health-related quality of life and associated factors: a cross-sectional analysis. *Health and Quality of Life Outcomes*, *21*(1), 28. https://doi.org/10.1186/s12955-023-02107-z.

[3] Fusar-Poli, P.; Salazar de Pablo, G.; De Micheli, A.; Nieman, D.H.; Correll, C.U.; Kessing, L.V.; Pfennig, A.; Bechdolf, A.; Borgwardt, S.; Arango, C.; van Amelsvoort, T. (2020). What is good mental health? a scoping review. *European Neuropsychopharmacology*, *31*, 33–46. https://doi.org/10.1016/j.euroneuro.2019.12.105.

[4] Cullen, W.; Gulati, G.; Kelly, B.D. (2020). Mental health in the COVID-19 pandemic. *QJM: An International Journal of Medicine*, *113*(5), 311–312. https://doi.org/10.1093/qjmed/hcaa110.

[5] Henderson, C.; Evans-Lacko, S.; Thornicroft, G. (2013). Mental illness stigma, help seeking, and public health programs. *American Journal of Public Health*, *103*, 777–780. https://doi.org/10.2105/ajph.2012.301056.

[6] Jansli, S.M.; Hudson, G.; Negbenose, E.; Erturk, S.; Wykes, T.; Jilka, S. (2022). Investigating mental health service user views of stigma on twitter during covid-19: a mixed-methods study. *Journal of Mental Health*, *31*(4), 576–584. https://doi.org/10.1080/09638237.2022.2091763.

[7] Benrouba, F.; Boudour, R. (2023). Emotional sentiment analysis of social media content for mental health safety. *Social Network Analysis and Mining*, *13*(1), 17. http://doi.org/10.21203/rs.3.rs-2170906/v1.

[8] Colella, S.; Dufourt, F.; Hildebrand, V.A.; Vivès, R. (2023). Mental health effects of covid-19 lockdowns: A twitter-based analysis. *Economics & Human Biology*, *51*, 101307. https://doi.org/10.1016/j.ehb.2023.101307.

[9] Beierle, F.; Pryss, R.; Aizawa, A. (2023). Sentiments about mental health on twitter—before and during the covid-19 pandemic. *Healthcare*, *11*(21), 2893. https://doi.org/10.3390/healthcare11212893.

[10] Hazarika, D.; Konwar, G.; Deb, S.; Bora, D.J. (2020). Sentiment analysis on twitter by using textblob for natural language processing. The International Conference on Research in Management & Technovation, 63–67. http://doi.org/10.15439/2020KM20.

[11] Fitri, F.; Nasrun, M.; Setianingsih, C. (2018). Sentiment analysis on the level of customer satisfaction to data cellular services using the naive bayes classifier algorithm. IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), 201–206. http://doi.org/10.1109/IOTAIS.2018.8600870.

[12] Testas, A. (2023). Naive Bayes Classification with Pandas, Scikit-Learn, and PySpark; Apress: Berkeley, USA, 281–297. https://doi.org/10.1007/978-1-4842-9751-3_11.

[13] Seref, B.; Bostanci, G.E. (2018). Sentiment analysis using naive bayes and complement naive bayes classifier algorithms on hadoop framework. International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). http://doi.org/10.1109/ISMSIT.2018.8567243.

[14] Taheri, S.; Mammadov, M. (2013). Learning the naive bayes classifier with optimization models. *International Journal of Applied Mathematics and Computer Science*, *23*(4), 787–795. http://doi.org/10.2478/amcs-2013-0059.

[15] Montesinos López, O.A.; Montesinos López, A.; Crossa, J. (2022). Support Vector Machines and Support Vector Regression; Springer: Cham, Switzerland. pp. 337–378. http://doi.org/10.1007/978-3-030-89010-0_9.