Analyzing the Level of Depression of Twitter Users Using Machine Learning

Putri Armaini[1,a)], Achmad Maududie[1,b)], Priza Pandunata[1,c)]

[1] *Faculty of Computer Science, Jember University Indonesia*

[a)] *putriarmaini@gmail.com*
[b)] *maududie@mail.unej.ac.id*
[c)] *priza@unej.ac.id*

Abstract. Depression is a psychological disorder characterized by changes in an individual's feelings, thoughts, and behaviors. According to the American Psychological Association (APA), those who regularly check their smartphones tend to experience higher levels of stress compared to individuals who spend less time with their phones. In the evaluation of depression symptoms, the Patient Health Questionnaire-9 (PHQ-9) can be used. This study describes a method for collecting depression data based on keywords extracted from the PHQ-9 questionnaire, which can indicate the level of depression. Keywords associated with different levels of depression were identified based on the characteristics linked to PHQ-9. These keywords were then utilized to collect data from Twitter, resulting in a dataset of 79,144 entries covering the period from May 28, 2023, to July 1, 2023. The data was subsequently analyzed using a machine learning approach based on Multinomial Naïve Bayes. The analysis revealed that 45,411 Twitter users did not show signs of depression, 2,385 users indicated mild depression, 5,069 users indicated moderate depression, and 3,636 users indicated severe depression. Interestingly, more than 65% of users who indicated experiencing depression, whether mild, moderate, or severe, tended to be more active in participating in social media conversations.

Keywords: Depression, smartphone use, PHQ-9

INTRODUCTION

Depression is a psychological disorder characterized by changes in an individual's feelings, thoughts, and behaviors. According to the American Psychological Association (APA), those who regularly check their smartphones tend to experience higher levels of stress compared to individuals who spend less time with their phones. The results of the I-NAMHS (Indonesia-National Adolescent Mental Health Survey) show that one in three teenagers in Indonesia has mental health problems while one in twenty teenagers in Indonesia has had a mental disorder in the last 12 months.

Twitter is indeed a platform where people feel comfortable venting, storytelling, discussing, and expressing opinions. It provides a space for individuals to share personal experiences, thoughts, and emotions in a concise and real-time manner. The problem of the research is high prevalence of mental disorders among adolescents in the past 12 months and the growth of messages conveyed through Twitter is very fast and extensive, making it difficult to manually capture the level of depression for each user.

On the other hand, there is a positive correlation between the intensity of social media usage and the level of depression [1]. Using social media more than three times a day can serve as a predictor of mental health issues and poor well-being among adolescents [2]. It is undeniable that the growth of the Internet in Indonesia is often utilized as a platform for users to channel their desires, one of which is expressing opinions on specific topics through social media. Self-expression for an individual can take the form of heartfelt confessions, frustrations, sad emotions, and thoughts in the present moment. Twitter users often feel comfortable unleashing their emotions, thoughts, and feelings in the form of short messages known as tweets. A tweet is a concise message with a character limit of 140 characters [3]. Twitter provides a comfortable space for venting, storytelling, engaging in discussions, and voicing opinions [4].

According to We Are Social, the number of Twitter users in Indonesia as of July 2023 has reached 25.25 million. Twitter is the platform is highly likely to be utilized as a resource for capturing the mental health of social media users, particularly on Twitter. Each user is classified based on their level of depression through the tweets they post. The classification of depression levels can refer to the Patient Health Questionnaire-9 (PHQ-9) questionnaire, which is used to evaluate depressive symptoms, while the classification algorithm refers to a well-established algorithm that has demonstrated high levels of accuracy.

I. H. Agustin (ed.), *Proceedings of the 2nd International Conference on Neural Networks and Machine Learning 2023 (ICNNML 2023)*, Advances in Intelligent Systems Research 183,
https://doi.org/10.2991/978-94-6463-445-7_10

which is used to evaluate depressive symptoms, while the classification algorithm refers to a well-established algorithm that has demonstrated high levels of accuracy.

METHOD

Mental Health

Mental health according to the World Health Organization (WHO). It states that mental health involves the well-being of physical, mental, and social aspects, rather than just being free from illness, disability, and weakness. According to Ihrom in 2008, mental health is the manifestation of personal integrity, harmony with one's true self, and the growth towards self-realization and positive relationships with others. Mental health is demonstrated by one's ability to control oneself, show intelligence, empathy, and sympathy towards others, as well as having a happy outlook on life. According to the Basic Health Research (Riskesdas) 2018, more than 19 million Indonesians aged 15 years and above have emotional mental disorders. Additionally, more than 12 million individuals within the same age range are known to experience depression. This indicates a significant issue related to mental health in Indonesia. It is crucial to understand various mental health disorders. Some common terms are depression, anxiety, stress, and other mental disorders.

Depression

Depression is a psychological disorder characterized by changes in a person's feelings, thoughts, and behaviors. Someone experiencing depression may feel sadness, loneliness, low self-esteem, and withdraw from interactions with their environment [5]. Depression is an emotional and behavioral disorder that occurs together in a specific pattern [6]. The severity of depression can vary depending on symptoms, including changes in behavior and emotions. According to the American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, depression is a psychological disorder characterized by the presence of sadness, feelings of emptiness, sensitivity, along with somatic and cognitive symptoms [7].

The difference between mild, moderate, and severe depression lies in the severity of symptoms and their impact on daily life (American Psychiatric Association, 2013). Here are the differences:

1. Mild depression: Symptoms may include being easily irritated, feeling hopeless, self-loathing, loss of interest in activities, lack of motivation, sleep problems, and changes in appetite.
2. Moderate depression: Symptoms may include feelings of low self-esteem, decreased productivity, restlessness, and excessive anxiety.

Severe depression: Individuals may have trouble carrying out daily activities and have negative thoughts, including suicidal ideation.

The Patient Health Questionnaire-9

The purpose of the nine items on the Patient Health Questionnaire-9 (PHQ-9) is to assess whether a person has had depressed symptoms in the two weeks prior. The elements for major depressive disorder are based on the nine diagnostic criteria. In addition to being able to determine the degree of symptoms, this questionnaire may be used as a dual-purpose tool to screen for depression disorders.

TABLE 1. The patient health questionnaire-9

No	The Patient Health Questionnaire-9 (PHQ-9)
1	little interest or pleasure in doing things?
2	feeling down, depressed, or hopeless?
3	trouble falling or staying asleep, or sleeping too much?
4	feeling tired or having little energy?
5	poor appetite or overeating?
6	feeling bad about yourself or that you're a failure or have let yourself or your family down?

No	The Patient Health Questionnaire-9 (PHQ-9)
7	trouble concentrating on things, such as reading the newspaper or watching television?
8	moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?
9	thoughts that you would be better off dead or of hurting yourself in some way?

Twitter

Social media is an online platform that enables people to generate and distribute content (knowledge, viewpoints, and hobbies) with a larger audience in a variety of contexts (informative, instructional, humorous, critical, etc.) [8]. Through the social networking site and microblogging app Twitter, users can post messages in real time, or "tweets." A tweet can contain up to 140 characters and is a brief message [3]. As per the most recent report by Databoks, as of early 2023, there were approximately 24 million Twitter users in Indonesia. This number positions Indonesia as the nation with the most number of Twitter users worldwide. In addition, as of July 2023, there were 564,1 million Twitter users in Indonesia, placing the country at the top of the global Twitter user base. This data indicates a significant decline in Twitter usage in Indonesia. With an ever-increasing user base, Twitter has become one of the most popular social media platforms in Indonesia for communication and information sharing.

Machine Learning

One technique for gathering data and identifying trends in each text material is text mining. Though it concentrates on unstructured text data, text mining is fundamentally comparable to data mining [9]. More relevant and plentiful information can be obtained by text mining. Numerous establishments depend on this data for their operational procedures. Tokenizing, filtering, stemming, and analysis are the first phases in the text mining process. To get normalized values for the data, term weighting and normalization procedures are then carried out.

In building the model, the Multinomial Naive Bayes algorithm was utilized. Multinomial naive Bayes is one of the text mining methods in the process of text classification using the probability values of a class. The process begins by inputting training data, which can be used for testing data later, to calculate the probability of a class occurring in the training data.

$$P(c) = \frac{N_c}{N_{doc}} \tag{1}$$

c : Category or class
doc : Document
Nc : Number of category c in the training documents
$Ndoc$: Total number of training documents used

Furthermore, the probability that a word belongs to a specific category or class can be calculated using the following equation.

$$P\left(w_i, c\right) = \frac{count\left(w_i, c\right)+1}{\sum w\ count\left(w_i, c\right)+|V|} \tag{2}$$

w : The i-th word in all documents categorized as c
$count\left(w_i, c\right)$: The number of occurrences of a specific word in a category or class
$\sum w\ count\left(w_i, c\right)$: The total count of all words in the class
$|V|$: The total number of unique words in the class The normalized result of a value x

For the next step is every tweet in the Twitter dataset can then be predicted thanks to the utilization of the previously constructed model Every account receives attention in the final step. An account is devoid of depression if it only has one tweet. If an account has several tweets, the analysis relies on the prediction made by the majority of users. As a result, the account is classified as having depression, with the degree of severity based on the majority prediction.

Results and Discussion

Research Flow

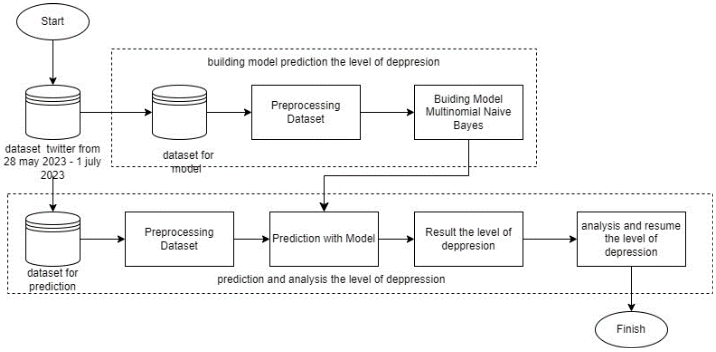

FIGURE 1. Research flow

This research begins by collecting a dataset from Twitter within the timeframe of May 28, 2023, to July 1, 2023. The dataset is then divided into two parts: one for model development and the other for prediction. The data used to build the prediction model for the level of depression undergoes a preprocessing stage. This process aims to clean and organize the data to be used in model construction. Subsequently, a model is built using the Multinomial Naive Bayes method. This method is employed to predict the level of depression based on the tweets in the dataset. The prediction dataset goes through a similar preprocessing stage. Afterward, predictions are made using the previously constructed model. The results of these predictions determine the classification or category to which each tweet belongs. The level of depression for each tweet provides the resulting data, which is then analyzed. By analyzing the results of the predicted level of depression, this research will yield valuable conclusions. These conclusions can provide further understanding of the levels of depression within the context of the Twitter dataset used in this study.

Prepare the dataset for detecting the level of depression

This collection is done by extracting data, which refers to the process of retrieving data from a data source and converting it from one format to another for further processing. In this data collection process, the data will be stored in CSV format. The data collection is obtained through crawling based on keywords extracted from the PHQ-9 questionnaire.

TABLE 2. Keywords extracted from the PHQ-9 questionnaire

No	PHQ-9 Questionnaire	Keywords
1	little interest or pleasure in doing things?	malas, mager, malas gerak
2	feeling down, depressed, or hopeless?	sedih, nangis, murung, putus asa, kecewa
3	trouble falling or staying asleep, or sleeping too much?	insomnia, sulit tidur,
4	feeling tired or having little energy?	lelah, lemah, lesu, cape, lunglai
5	poor appetite or overeating?	tidak nafsu

No	PHQ-9 Questionnaire	Keywords
6	feeling bad about yourself or that you're a failure or have let yourself or your family down?	Tidak percaya diri, malu, takut
7	trouble concentrating on things, such as reading the newspaper or watching television?	Ragu
8	moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?	resah, gelisah, cemas, khawatir, ragu
9	thoughts that you would be better off dead or of hurting yourself in some way?	mati, selfharm

The data collection process in this research uses the tweet-harvest library, which utilizes the author's Twitter auto-token. Subsequently, data crawling is conducted using pre-formed keywords and based on time. This study limits the collection to only the last 100 tweets containing the specified keyword each day. The data for this research was collected from May 28th to July 1st, 2023.

TABLE 3. Dataset crawling

Created-At	**From-User**	**From-User-Id**	**Text**
04-06-23 1:46	biru	1.52E+18	mau selfharm gaada silet mau pake cutter y mnimal kaga berkarat lah memek SOALNY GMW MATII cuma mau ngelampiasin doang
07-06-23 1:30	Who mme who?	9.55E+17	Otak overheat tapi kalo dibawa tidur bawaannya gelisah dan ovt, pantes ya bnyk case org depresi gara2 skripshit ·
25-06-23 22:45	LviaKz	8.97E+17	Alhamdulillah, sedikit cape aja tapi gpp https://t.co/iy3DKbBMzd

The data obtained a total of 79,144 data for this study. The data will be divided into two datasets. 2,261 data will be used for model creation, and the remaining data will be used for the prediction dataset.

The dataset is divided into three classes

From the dataset, it will be manually labeled based on categories that correspond to mild PHQ-9 depression. The data will be divided into three classes: 708 data in class 1 (depression), 991 data in class 2 (moderate depression), and 864 data in class 3 (severe depression). The labeled dataset will be used to create a model using the multinomial naive Bayes algorithm, which includes both training and testing data.

TABLE 4. Labeling guidelines

No	**PHQ-9 Questionnaire**	**Keywords**	**Class**
1	little interest or pleasure in doing things?	malas, mager, malas gerak	mild depression
2	feeling down, depressed, or hopeless?	sedih, nangis, murung, putus asa, kecewa	mild depression
3	trouble falling or staying asleep, or sleeping too much?	insomnia, sulit tidur,	mild depression
4	feeling tired or having little energy?	lelah, lemah, lesu, cape, lunglai	moderate depression

No	PHQ-9 Questionnaire	Keywords	Class
5	poor appetite or overeating?	tidak nafsu	mild depression
6	feeling bad about yourself or that you're a failure or have let yourself or your family down?	Tidak percaya diri, malu, takut	moderate depression
7	trouble concentrating on things, such as reading the newspaper or watching television?	Ragu	moderate depression
8	moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?	resah, gelisah, cemas, khawatir	moderate depression
9	thoughts that you would be better off dead or of hurting yourself in some way?	mati, selfharm	severe depression

TABLE 5. Dataset for model

Created-At	From-User	From-User-Id	Text	Label
04-06-23 1:46	biru	1.52E+18	mau selfharm gaada silet mau pake cutter y mnimal kaga berkarat lah memek SOALNY GMW MATII cuma mau ngelampiasin doang	3
07-06-23 1:30	Who mme who?	9.55E+17	Otak overheat tapi kalo dibawa tidur bawaannya gelisah dan ovt, pantes ya bnyk case org depresi gara2 skripshit ·	2
25-06-23 22:45	LviaKz	8.97E+17	Alhamdulillah, sedikit cape aja tapi gpp https://t.co/iy3DKbBMzd	2

Creating and validating a depression classification model

Depression classification models are made via multinomial naive Bayes. There are several steps to the technique, the first being preparation, which includes preprocessing, weighing, normalization, and data modeling. Multinomial Naïve Bayes is used to classify the data into three groups throughout the model construction process: mild depression, moderate depression, and severe depression. In this study, the data is split using the train_test_split function from the sklearn library, with 90% going to training and 10% to testing. The model is then generated using the sklearn.naive_bayes module and multinomialNB(). The modeling results are contained in the variable "predicted". As a result, 0.876 is the accuracy of the model.

Making predictions on the entire data obtained from crawling

A dataset of 76,483 is used in the prediction step to provide predictions based on the model that was previously developed. Preprocessing for the depression prediction stage begins in a manner akin to that of the model construction preprocessing step before. This phase may require a lot of time. After that, TF-IDF is used for weighting, and MinMaxScaler is used for normalization. The pre-existing model is used to make predictions after weighting yields the vector values.

Predictions are thus made using the twitter data that is now available. The tweet data will be categorized into three classes based on the prediction results: mild depression, moderate depression, and severe depression.

```
                                                tweet  prediksi
0          and he still pulled the capek out three times         2
1         capek capek war nyantol calo nya luar nalar ceee         2
2        live near capek verdean diaspora their descend...         2
3                                    capek malam ngakakin         3
4                             capek banget tonton nappeun eomma         3
...                                                   ...       ...
76479                                relax yuk biar selfharm         2
76480          mati lampu hujan tau mentioning of selfharm         2
76481                     menang tidur kadang suka selfharm sih         1
76482  nyokap tinggal sbmptn kayak kenal selfharm dll...         3
76483  mintak balas orang kecewa kecewa perhati jaga ...         3

[76484 rows x 2 columns]
```

FIGURE 2. Predict depression

Based on figure 2. The model's predictions aim to determine the depression level indicated by sentences. In the first data, sentences show mild depression, in the second data, sentences show moderate depression, and in the third data, sentences show severe depression. This pattern persists for a total of 76,483 data points.

Making conclusions about the level of depression for each account

Based on the results of the prediction values, conclusions will be drawn. The conclusions will only be derived from tweet data from usernames that have more than one tweet. Therefore, usernames that only have one tweet will be concluded as not experiencing depression. For usernames with multiple tweets, conclusions will be based on the frequency of predicted values.

In this study, the following conclusions were drawn: 45,411 Twitter users or accounts were found to not experience depression. Furthermore, 2,385 Twitter users or accounts were identified as experiencing mild depression. Moderate depression was observed in 5,069 Twitter users or accounts. Lastly, 3,636 Twitter users or accounts were found to experience severe depression.

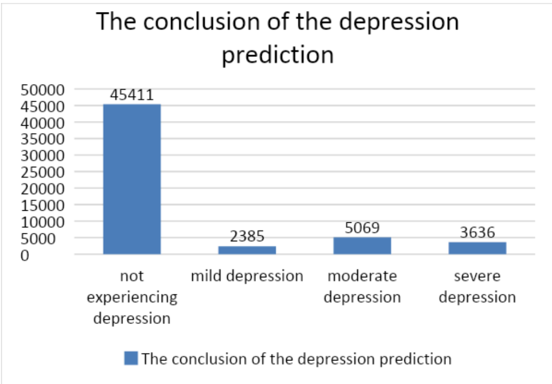

FIGURE 3. The conclusion of the depression prediction

Create a distribution of the number of tweets per account for each level of depression

An evaluation of the scores will be conducted to obtain conclusions from the distribution of tweets across the classes of severe depression, moderate depression, and mild depression.

This stage will provide results in the form of a pie chart. The distribution of mild depression class, based on the number of tweets per account, yields the following results: 22.4% of accounts have 2-3 tweets, 6.5% have 4-5 tweets, 1.9% have 6-7 tweets, and 69.3% have more than 7 tweets.

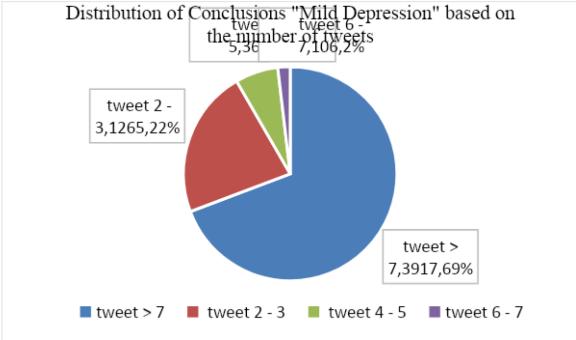

FIGURE 4. Distribution of conclusion “Mild Depression” based on the number of tweet

The distribution process also applies to the moderate depression level. This stage will provide results in the form of a pie chart. The distribution of the moderate depression class, based on the number of tweets per account, yields the following results: 16% of accounts have 2-3 tweets, 6% have 4-5 tweets, 3.4% have 6-7 tweets, and 74.6% have more than 7 tweets.

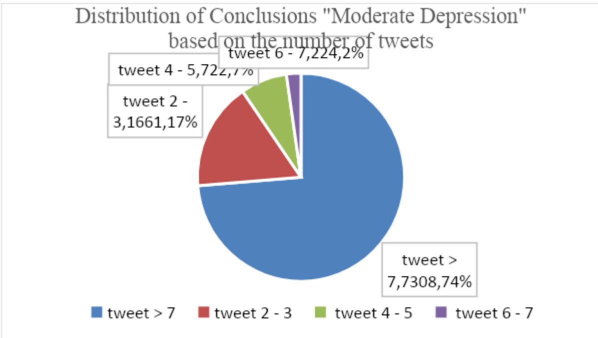

FIGURE 5. Distribution of cunclusions “moderate depression” based on the number of tweets

The distribution process also applies to the severe depression level. This stage will provide results in the form of a pie chart. The distribution of the severe depression class, based on the number of tweets per account, yields the following results: 16.8% of accounts have 2-3 tweets, 7.3% have 4-5 tweets, 2.3% have 6-7 tweets, and 73.7% have more than 7 tweets.

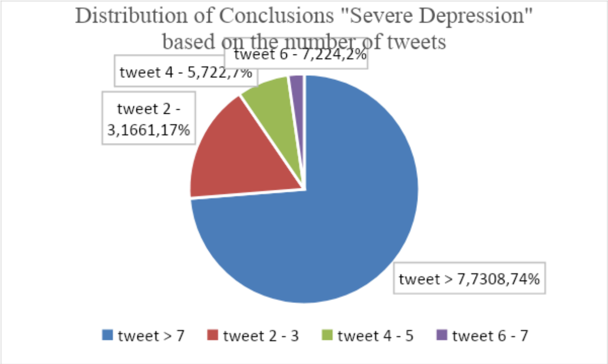

FIGURE 6. Distribution of conslusions “Severe Depression” based on the number of tweets

In the distribution results based on the number of tweets per account, it is observed that the majority of accounts have a distribution of more than 7 tweets. Therefore, the conclusion drawn is that approximately 69% of users who are indicated to have depression, whether it is mild, moderate, or severe, tend to tweet more than 7 times.

Conclusion

The application of this model resulted in an accuracy score of 0.876, with three classes: mild depression, moderate depression, and severe depression. The created model can predict tweet data by classifying them into three categories: mild depression, moderate depression, and severe depression.

The analysis of the depression levels among Twitter users shows that more than 69% of users who are indicated to have mild, moderate, and severe depression tend to engage in commenting on Twitter, with a frequency of more than seven comments.The application of this model resulted in an accuracy.

The conclusion drawn from the prediction results is as follows: 45,411 Twitter users or accounts do not experience depression, 2,385 Twitter users or accounts experience mild depression, 5,069 Twitter users or accounts experience moderate depression, and 3,636 Twitter users or accounts experience severe depression. This indicates that a total of 11,090 individuals, which accounts for 20% of Twitter users, are indicated to have depression.

This finding is in line with the research conducted by I NAMHS, which states that one in three Indonesian teenagers has a mental health problem. It is crucial to address this issue in Indonesia as it has the potential to become a significant concern by the time Indonesia reaches its "Golden Indonesia" goal in 2045. Taking proactive measures to address mental health issues is essential to ensure the well-being and productivity of the population.

Acknowledgment

This research was supported by University of Jember. I thank all those who have helped my research.

References

1. Abidah and A. Aziz, “Hubungan Antara Intensitas Penggunaan Media Sosial dan Tingkat Depresi pada Mahasiswa,” 2020. [Online]. Available: http://journal.uny.ac.id/index.php/acta-psychologia
2. K. E. Riehm et al., (2019), “Associations Between Time Spent Using Social Media and Internalizing and Externalizing Problems Among US Youth,” *JAMA Psychiatry*, vol. 76, no. 12, p. 1266, Dec. 2019, doi: 10.1001/jamapsychiatry.2019.2325.
3. *Data Analytics for Social Microblogging Platforms*. Elsevier, (2023). doi: 10.1016/C2021-0-01247-5.
4. D. Inayah and F. Law Purba, “IMPLEMENTASI SOCIAL NETWORK ANALYSIS DALAM PENYEBARAN INFORMASI VIRUS CORONA (COVID-19) DI TWITTER,” 2020.

5. A. Kupferberg and G. Hasler, (2023), "The social cost of depression: Investigating the impact of impaired social emotion regulation, social cognition, and interper-sonal behavior on social functioning," *J Affect Disord Rep*, vol. 14, p. 100631, Dec. 2023, doi: 10.1016/j.jadr.2023.100631.
6. A. Kupferberg, L. Bicks, and G. Hasler, (2016), "Social functioning in major depres- sive disorder," Neurosci Biobehav Rev, vol. 69, pp. 313–332, Oct. 2016, doi: 10.1016/j.neubiorev.2016.07.002.
7. J. C. Tolentino and S. L. Schmidt, (2018), "DSM-5 Criteria and Depression Severity: Implications for Clinical Practice," Front Psychiatry, vol. 9, Oct. 2018, doi:10.3389/fpsyt.2018.00450.
8. G. F. Khan, *Social Media for Government*. Singapore: Springer Singapore, 2017. doi: 10.1007/978-981-10-2942-4.
9. D. Antons, E. Grünwald, P. Cichy, and T. O. Salge, (2020), "The application of text mining methods in innovation research: current state, evolution patterns, and development priorities," *R&D Management*, vol. 50, no. 3, pp. 329–351, Jun. 2020, doi: 10.1111/radm.12408.