SENTIMENT ANALYSIS ON SOCIAL MEDIA INSTAGRAM OF DEPRESSION ISSUES USING NB METHOD

Voni Anggraeni Suwito Putri[1], Anik Vega Vitianingsih[2], Rusdi Hamidan[3], Anastasia Lidya Maukar[4], Niken Titi Pratitis[5]
[1,2,3]Department of Informatics, Dr.Soetomo University, Surabaya, Indonesia
[4]Department of Industrial Engineering, President University, Bekasi, Indonesia
[5]Department of Psychology, Universitas 17 Agustus 1945, Surabaya, Indonesia
voniputri2204@gmail.com[1] , vega@unitomo.ac.id[2], rusdi@unitomo.ac.id[3] ,
almaukar@president.ac.id[4] , nikenpratitis@untag-sby.ac[5]

Abstract- In today's digital era, Instagram is one of the main social media platforms used by Generation Z as a means to express themselves. In particular, this platform is often used by adolescents who experience depression to pour out their feelings through uploads accompanied by captions that reflect mental conditions or depression, including feelings of depression and hopelessness. The purpose of this research is to produce a sentiment analysis related to public opinion on the issue of depression and mental health on the Instagram Apps Platform with neutral, positive and negative classifications. This analysis focuses on public opinion to understand sentiment patterns with an approach that combines lexicon-based text labelling to group sentiments at specific weights and the Naive Bayes (NB) method for classification after labelling. The NB method was chosen because it has an effective ability to classify text by producing an accuracy of 82.55% which is higher than previous research with an accuracy of 79%. The data obtained after conducting sentiment analysis is 1300 comment data and has been analyzed, as much as 33% positive, 35.8% neutral and 31.2% negative. The superiority of NB in this study shows its effectiveness in conducting sentiment analysis on Instagram social media. This level of accuracy is a strong basis for further sentiment analysis related to public perception. This study is expected to provide insights for mental health professionals through which deeper understanding and more targeted mental health education can be conducted to reduce stigma, so that people experiencing depression are encouraged to seek help without fear or shame and potentially support mental wellbeing more broadly.

Keywords- Sentiment Analysis, Depression, Mental Health, Naïve Bayes Method

I. INTRODUCTION

The increasing use of social media such as Instagram has made social media platforms not only serve as a means of sharing moments, but also as a place to express feelings and emotional conditions related to mental health and depression [1]. Instagrammers share stories about their struggles with depression through captions, photo or video uploads, as well as hastags such as #depression, #MentalHealth which are increasingly popular [1]. These posts can reflect the mood, outpouring of feelings, and thoughts of users who make Instagram a place to voice what they do not easily express in the surrounding environment. Depression itself is a medical condition that causes feelings of sadness and negatively affects thoughts, behavior, emotions, and overall mental health [2]. The Instagram platform allows individuals

experiencing depression to feel heard and validated through interactions with other users, such as comments and support from the community.

Instagram's features that allow users to make connections without geographical boundaries can help individuals feel connected to others who have had similar experiences [2]. This support has the potential to reduce feelings of isolation often experienced by people with mental illness. Data generated from Instagram posts can be used to understand trends in public sentiment towards depression. Sentiment analysis on text shared by Instagram users allows researchers and health professionals to identify evolving emotional patterns, detect warning signs, and understand changes in people's perceptions of mental health [3]. It is also useful for increasing public awareness, shaping mental health policies, and developing appropriate strategies for providing education and interventions for individuals in need [4]. According to WHO in 2012, about 350 million people suffer from depression, and it is stated that nearly one million people with depression end their lives each year [5].

Based on the literature study, previous research that discusses public opinion related to mental health issues [6]. This research conducted sentiment analysis with hashtags #depression and #MentalHealth using the KNN method resulting in the highest accuracy of 58.39% in classifying positive and negative sentiments on Twitter data, using a data division of 70:30 and a value of k = 5 [6]. While other research [7], with the hashtag #MentalHealth using the Naïve Bayes method shows a higher accuracy of around 79%. This shows that Naïve Bayes is more effective in analyzing sentiment than KNN. [7]The KNN method requires choosing the right k value to affect the accuracy results and requires further evaluation for optimization. In contrast, naïve bayes is simpler, efficient and has more stable accuracy because it is based on statistical probabilities. naive bayes also tends to be faster and requires minimal optimization process compared to KNN. The gap between the two methods lies in accuracy and processing efficiency. Naïve Bayes is superior in terms of accuracy and efficiency on similar data sets, whereas KNN shows limitations in accuracy and requires more complexity. It can be concluded, there has been no analytical research on the issue of depression using lexicon-based methods in identifying sentiments for labelling with certain weights and Naïve Bayes to classify sentiments. This research uses Instagram dataset to classify public opinion on mental health issues.

Therefore, this study aims to produce a sentiment analysis of public opinion on the issue of depression on Instagram with neutral, positive, and negative sentiment classification, in order to provide more accurate insight and understanding of public views on mental health and depression. The dataset is taken from the hashtags #depression, #MentalHealth, covering various related cases. This research differs from previous studies that focus more on Twitter as the main platform for sentiment analysis related to mental health. Twitter is known for its short message characteristics, whereas Instagram gives users more freedom to express themselves. This study chooses Instagram which allows for a deeper understanding of acceptance and disclosure through complex visuals and text. By integrating visual and textual analysis, this study is expected to overcome the limitations of previous studies in conducting textual analysis on Twitter. And, this study is also expected to contribute to expanding the understanding of mental health and depression issues on social media and make Instagram a significant new data source in studies related to mental health sentiment.

II. SIGNIFICANCE OF THE STUDY

A. Literature Study

Sentiment analysis is a type of classification that analyzes text based on well-defined sentiment opinion orientations and also involves the process of detecting the structural polarity of text to determine whether a particular text has a positive, negative, or neutral value [8]. The most frequently used method in sentiment analysis research is NB, which is applied as a social media context platform [1]. New research published in the journal "Sentiment Analysis Of Mental Health Using K-Nearest Neighbors On Social Media Twitter," shows how effective the KNN algorithm is in classifying data according to a high degree of accuracy [6]. In contrast, research [7] used the NB method for classification in sentiment analysis. This study has an accuracy rate of about 79% in classifying sentiments related to mental health on social media twitter. This research focuses on public opinion and under-researched areas such as the Instagram platform using the naive bayes method. Although NB is a popular method, there are still not many studies that use this method to analyze the Instagram platform, especially on the topic of depression issues.

1) Text Pre-processing. Pre-Processing is an important step in preparing data before using it in a model [9]. This stage has a significant impact on the data collection process because often the data used is not in good condition for data collection. This procedure is also necessary to handle data variations that can affect the quality of research data [9]. Stages in pre-processing include:

a) Data Cleaning. In this process, irrelevant or biased data can be detected and processed to improve the quality and accuracy of the data set used.

b) Case Folding. In this stage, all text is converted into a standardized form (lowercase). This data filtering process results in data that has been cleaned and prepared for processing to the next level.

c) Tokenization. The process of tokenization is to divide the text into tokens or small units such as words or phrases. The first step in this process is to input the review data, then check the combined data, and then separate each combined word so that the review data can be separated correctly.

d) Stopword. The purpose of this step is to remove words from the text analysis process that are considered meaningless or contribute very little. To properly customize the comments, the algorithm first imports the comment dataset, verifies each word, and then removes frequently occurring words.

e) Stemming. The stage of removing affixes such as prefixes and suffixes. In this process, the entered data is reviewed and then checked whether the words in the data are rough. If not, the next step is to remove prefixes. However, if the word is rough, the data will not be processed further.

2) Term Frequency-Inverse Document Frequency (TF-IDF): Term Frequency (TF) is a measure of the frequency of occurrence of a word across a collection of documents calculated as part of this term weighting technique [10]. All documents in the dataset are used to calculate the inverse document frequency (IDF), which includes the total number of documents containing the phrase. By measuring how often feature t appears in document d, TF is determined. According to reference [10], The IDF value is obtained by dividing the total number of documents (n) by the number of documents (DFt) that contain a feature (t). The following are the formulas of this method equation expressed in (1), (2), and (3) [10].

Term Frequency (TF)

$$TF_{t,d} = F(t,d) \quad (1)$$

Inverse Document Frequency (IDF)

$$IDF_t = \log\left(\frac{N}{DF_{(t)}}\right) \quad (2)$$

Term Frequency-Inverse Document Frequency (TF - IDF)

$$W_{t,d} = W_{tf_{t,d}} \times IDF_t \quad (3)$$

3) NB: NB is a classification technique based on the principle of probability, invented by Thomas Bayes, a British statistician [11]. Bayes' theorem, also known as the NB algorithm, is able to predict the likelihood of future events based on past data [11]. The NB method excels in efficiency because it is able to perform the classification process accurately despite using a relatively small amount of data for the necessary parameter estimation. In Bayes' terms, X is considered as "evidence" that describes the measurement of a particular attribute, such as attribute n. While Y is a hypothesis, such as that the data tuple X belongs to class F. In the classification problem, the main objective is to determine P(Y|X), which is the probability that the hypothesis Y is true, given the "evidence" or observed data tuple X. P(Y|X) is the posterior probability of Y conditioned on X [12]. The formula can be seen in Equation (4).

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = P(X|Y)P(Y) \quad (4)$$

4) Confusion Matrix: Confusion Matrix is a process that is commonly used to calculate accuracy levels in deep learning algorithms. The matrix is used in testing to estimate the correct and incorrect objects, thus allowing the calculation of accuracy, precision and recall values [13]. Confusion Matrix also represents the performance of the classification model by comparing the predicted results with the actual values in the form of a table [14]. The confusion matrix calculation can be seen in table 1.

TABLE 1. CONFUSION MATRIX

	Prediction Class		
	Positive	Negative	Neutral

	Positive	TP	FN	FN
Actual Data	Negative	FP	TP	TN
	Neutral	FP	TN	TP

The above statement can be explained that TP (True Positive) is a positive classification result and the actual data is true as positive, FP (False Positive) is a negative classification result but the actual data is positive, TN (True Negative) is a negative classification result but the actual data is negative, and FN (False Negative) is a positive classification result but the actual data is negative.

The following is the formula for calculating the confusion matrix to measure the precision of the system's accuracy in determining relevant documents, Recall is the value of the system's accuracy in obtaining actual relevant documents and accuracy to get the accuracy of decision making by the system in determining document relevance. as in Equations (5), (6), and (7).

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (7)$$

B. Research Methods

In this research, the sentiment analysis approach is carried out through a series of structured and systematic stages to ensure the accuracy and validity of the results obtained. The process starts from data collection through web crawling techniques to the final evaluation stage. Each stage has an important role in preparing the data and ensuring that the analysis results truly reflect the public opinion recorded on Instagram social media. The stages in the research.

The process starts with the user entering the Instagram post link as input. The system then retrieves comments through web crawling using the python programming language. The comments obtained are raw data ready for further processing. After the data is obtained through crawling, the next step is to pre-process the data to prepare the data before analysis. This process includes case folding, data cleaning, stopwords, stemming, and tokenization. After the data is ready, the next process is labelling with a lexicon-based method, to determine whether the comment has a negative, positive, or neutral sentiment. Then, the labeled data will be classified using naïve bayes to classify the data based on sentiment probability. The last stage is evaluation, this process includes the calculation of TF - IDF to assess the weight of each word, the creation of a confusion matrix to measure classification accuracy, and the

visualization of the results in graphical form. This process provides a more accurate insight into public sentiment towards mental health and depression issues on the Instagram platform.

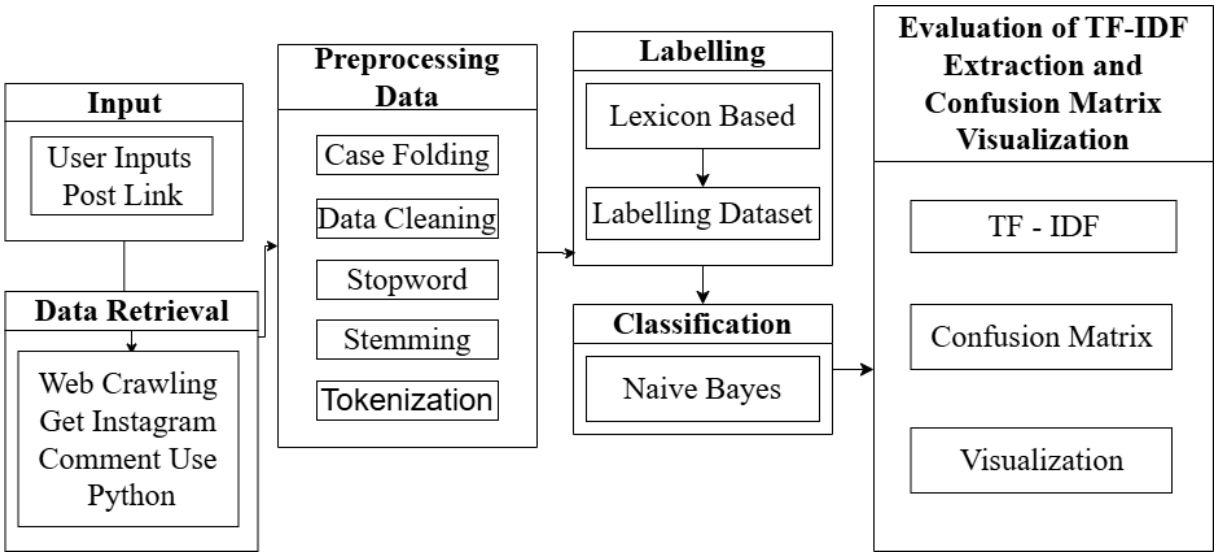

Figure 1. Research Methodology

C. Data Collection

Data collection from Instagram comments is done using the Instagram API with the keywords #MentalHealth, #Depression. The system will ask users to log in using their Instagram account and password in order to access and collect relevant comments.

III. RESULT AND DISCUSSION

A. Data Crawling

Web Crawling is a method in web mining that functions to retrieve data that exists or is stored on certain web pages or websites [15]. Collecting comment data about this research is done by web crawling using a module provided by the Python library. The system will crawl the URL link that has been entered by the user so that the system can retrieve comment data from the object to be studied. The comments obtained are raw data that still needs further processing. The dataset obtained is a raw dataset that has not gone through the next process, for example the word "bundirr" is still a word with inappropriate spelling because it has not gone through the word cleaning process. The crawling results can be seen in Table 2.

TABLE 2. CRAWLING DATA

No	Comment
1	*Kehilangan diri sendiri itu lebih sedih* 😥
2	*Fisik ku masih di bumi tapi jiwa ku sudah di pemakaman*
3	*"Kenapa masih gak di jemput? Ini berapa banyak lagi masalah yg harus aku hadapi?ini aku udah*

	beneran nyerah sama hidup aku ya Allah, mati muda juga aku gakppa 🥹🥹🥹"
4	*Andai bundirr gak dosa…..*

B. Data Pre-Processing

After the data is processed through web crawling, the next step is to preprocess the data to prepare the data before analysis. This process consists of several stages. First, data cleaning is used to filter out unnecessary characters or symbols, such as punctuation marks and irrelevant numbers. Second, case folding is performed to convert all text into lowercase letters, eliminating capitalization differences. Then, tokenization to break the text into individual words or *tokens* to facilitate subsequent analysis. Then, stopword is done to remove common words that do not have specific meaning, stemming to convert words into basic objects so that words that have the same root can be recognized as one entity. The process can be seen in Table 3.

TABLE 3. PRE-PROCESSING DATA

Pre-Processing Data	Result
Dataset	*Capek nggak sih, pura baik-baik saja (?)*
Data Cleaning	*capek nggak sih pura baik baik saja*
Case Folding	*capek nggak sih pura baik baik saja*
Tokenization	*"capek" "nggak" "sih" "pura" "baik" "baik" "saja"*
Stopword	*"capek" "nggak" "sih" "pura"*
Stemming	*"capek" "nggak" "sih" "pura"*

C. Data Labelling

After the data has been processed through preprocessing, the next step is to perform labelling to prepare the data before sentiment analysis. At this stage, the system uses a lexicon-based method to assign weights to certain words or phrases. This weight determines whether the comment falls into the positive, negative, or neutral category. In the labelling process, the positive sentiment results are 535 reviews, negative sentiment 485 reviews, and neutral sentiment 265 reviews. Table 4 displays the results of the labelling process.

TABLE 4. LABELLING DATA

Teks Clean	Score	Sentiment
Ehehe	0	Neutral
Alay fomo	-4	Negative

Kuat kuat kuat	6	Positive

D. Evaluation Data

The last process is evaluation. After the data is processed, the next step is to evaluate it to prepare the analysis results in visual form. At this stage, the system evaluates the classification results with several methods. First, TF-IDF extraction is used to calculate the weight or importance of each word in the comment. Then, the confusion matrix is generated to visualize the classification performance, by showing the number of correct and incorrect predictions in each sentiment category. Finally, the analysis results are visualized in the form of diagrams to facilitate interpretation. This process resulted in 1300 Instagram user comment data that had been converted into clean data after pre-processing and labelling, so that it could be analyzed using lexicon-based techniques and NB to classify the data set. Using text polarity, NB generates positive, negative and neutral values. To guarantee a more accurate categorization, the data used in this study is in Indonesian as it matches the lexicon used.

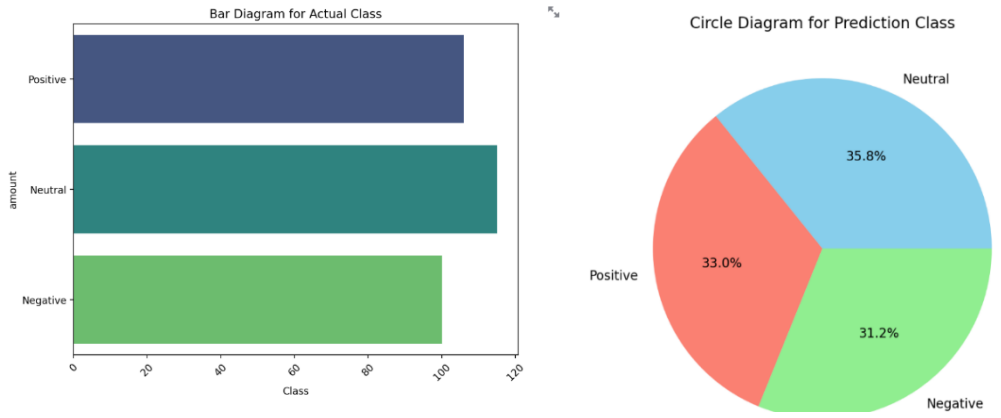

Figure 2. Visualization results evaluation

The visualization in the figure above shows that the polarity of positive sentiment is 33.0%, the polarity of neutral is 35.8%, and the polarity of negative sentiment is 31.2%. After completing all the analysis steps, the last stage is evaluation using a 3x3 confusion matrix. This model can classify the data into positive, negative, and neutral classifications. Thus, we can evaluate the performance and accuracy of the sentiment analysis model and identify areas where the model needs to be further improved. Confusion matrix is a method that is often used in deep learning methods to calculate the accuracy rate.[13]. In table 5 are the results of accuracy, precision and recall of data that has gone through several processes until the evaluation process.

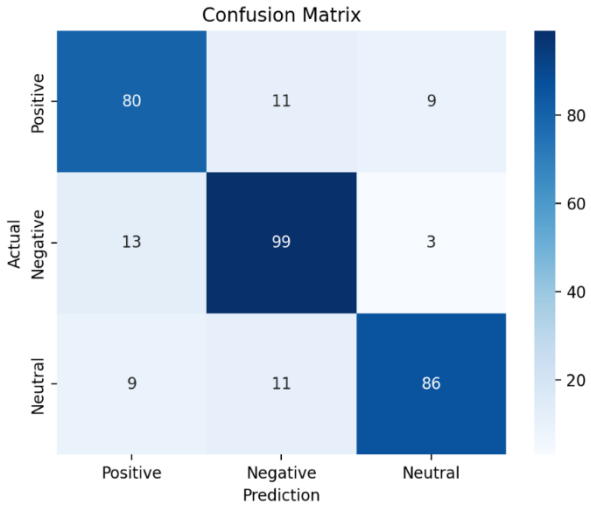

Figure 2. Confusion matrix results in 3x3

TABLE 5. NAÏVE BAYES RESULT

Matrix	Presentation
Accuracy	82.55%
Precision	82.72%
Recall	82.55%

E. TF - IDF

TF - IDF calculation is an approach method that combines two weighting concepts, namely the frequency of occurrence of a word in a particular document and the inverse frequency of documents containing that word [10]. This method is useful for identifying important terms and increasing the accuracy of text analysis. This method is also useful for identifying key terms and improving accuracy in text analysis. Term Frequency (TF) works by calculating how often a word appears in a particular document [16]. The TF formula is quite simple, by calculating the frequency of occurrence of a particular word in a document and dividing it by the number of words in the same document. TF calculation can be seen in Table 5:

TABLE 5. TF WEIGHTING

Stemming	Term Frequency (TF)	
	Word	TF Value
['aku','juga','mau','dirayakan','sebelum','aku','bundir']	*aku*	2/7
	juga	1/7
	mau	1/7
	dirayakan	1/7
	sebelum	1/7
	bundir	1/7

After calculating Term Frequency (TF), the next process is to calculate Inverse Document Frequency (IDF). The IDF process aims to determine how important a word is in a text document, which is closely related to how often the word appears in various documents. The simple formula used to calculate IDF is IDF(t) = log(total number of documents / number of documents containing word t) [16]. The IDF calculation can be seen in Table 6.

TABLE 6. IDF WEIGHTING

Word	N	DF	IDF
Aku	2	1	0.301029
juga	1	1	0
mau	1	1	0
dirayakan	1	1	0
sebelum	1	1	0
bundir	1	1	0

Next, calculate the Term Frequency - Inverse Document Frequency (TF - IDF) value. This calculation has the result of a word value which is a combination of the TF value and IDF value in the weight calculation. The TF - IDF calculation is quite simple by multiplying TF with IDF. This value indicates the importance of the word in the document compared to the entire corpus, helping to identify important words in text analysis [16]. The calculation can be seen in Table 7.

TABLE 7. TF-IDF WEIGHTING

Word	N	IDF	TF – IDF
Aku	2	0.301029	0.100342
juga	0	0	0
mau	0	0	0
dirayakan	0	0	0
sebelum	0	0	0
bundir	0	0	0

IV. CONCLUSION

The results of this study show that sentiment analysis on Instagram comments related to depression issues can provide insight into people's views on mental health issues. Neutral sentiment is the dominant sentiment (35.8%), followed by positive (33.0%) and negative (31.2%) sentiments with an accuracy rate of 82.55%. These results can be used by mental health professionals to understand public perception, analyze the stigma around depression, and design more appropriate interventions. With this better understanding, more targeted mental health education can be disseminated to help people experiencing depression feel supported to seek help without fear or shame. This study also has the potential to help reduce the stigma surrounding depression and encourage more positive public awareness about the importance of mental health. However, this study also has some limitations. Firstly, the size of the dataset used may not be large enough to illustrate public sentiment more broadly. Also, the Naïve Bayes (NB) method used has limitations in handling more complex sentiments, especially in distinguishing emotional moods in text. The NB accuracy rate of 82.55%, while high, may still fall short in capturing ambiguous or branching sentiments that often appear in mental health discussions. This limitation needs to be addressed in future research that may involve larger data sets and more adaptive methods for complex sentiment analysis.

REFERENCES

[1] V. P. Ramadhan and G. M. Namung, "Klasterisasi Komentar Cyberbullying Masyarakat di Instagram berdasarkan K-Means Clustering," *J-Intech*, vol. 11, no. 1, pp. 32–39, 2023, doi: 10.32664/j-intech.v11i1.846.

[2] M. Fachriza, F. I. Komputer, and U. E. Unggul, "Analisis Sentimen Kalimat Depresi Pada Pengguna Twitter Dengan Naive Bayes , Support Vector Machine , Random Forest," pp. 49–58, 2023.

[3] A. Primadhani Tirtopangarsa and W. Maharani, "Sentiment Analysis of Depression Detection on Twitter Social Media Users Using the K-Nearest Neighbor Method Analisis Sentimen Deteksi Depresi pada Pengguna Media Sosial Twitter dengan Menggunakan Metode K-Nearest Neighbor," *Semnasif*, vol. 1, no. 1, pp. 247–258, 2021.

[4] F. Darmawan, M. Joe, Y. I. Kurniawan, and L. Afuan, "Analisis Sentimen Kemungkinan Depresi dan Kecemasan pada Twitter Menggunakan Support Vector Machine," pp. 24–36, 2023, doi: 10.30864/eksplora.v13i1.854.

[5] A. Rizki and Y. Sibaroni, "Analisis Sentimen Untuk Pengukuran Tingkat Depresi Pengguna Twitter Menggunakan Deep Learning," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 11367–11375, 2021.

[6] M. L. Wicaksono, R. Rusdah, and D. Apriana, "Sentiment Analysis Of Mental Health Using K-Nearest Neighbors On Social Media Twitter," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 19, no. 2, p. 98, 2022, doi: 10.36080/bit.v19i2.2042.

[7] K. Yan, D. Arisandi, and T. Tony, "Analisis Sentimen Komentar Netizen Twitter Terhadap Kesehatan Mental Masyarakat Indonesia," *J. Ilmu Komput. dan Sist. Inf.*, vol. 10, no. 1, 2022, doi: 10.24912/jiksi.v10i1.17865.

[8] M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," *Procedia Comput. Sci.*, vol. 87, pp. 44–49, 2016, doi: 10.1016/j.procs.2016.05.124.

[9] I. R. Hendrawan and E. Utami, *Natural Language Processing: Eksplorasi Sentimen Masyarakat dalam Evaluasi Produk Lokal Indonesia menggunakan Algoritma Bag of Words, TF-IDF, Word2Vec, dan Doc2Vec*. 2023. [Online]. Available: https://books.google.co.id/books?id=caPgEAAAQBAJ

[10] N. Wiliani, N. L. Chusna, and P. B. Ramadhan, *Analisis Sentimen terhadap Pro Kontra Aksi*

Unjuk Rasa Mahasiswa dengan Naïve Bayes dan Information Gain. Penerbit NEM, 2023. [Online]. Available: https://books.google.co.id/books?id=RvTUEAAAQBAJ

[11] T. Sanubari, C. Prianto, and N. Riza, *Odol (one desa one product unggulan online) penerapan metode Naive Bayes pada pengembangan aplikasi e-commerce menggunakan Codeigniter*. in 1. Kreatif, 2020. [Online]. Available: https://books.google.co.id/books?id=s4j_DwAAQBAJ

[12] A. S. Huda, R. M. Awangga, and R. N. S. Fathonah, *Prediksi Penerimaan Pegawai Baru Dengan Metode Naive Bayes*. in Data Science. Kreatif, 2020. [Online]. Available: https://books.google.co.id/books?id=aawGEAAAQBAJ

[13] *Konsep Dasar Natural Language Processing (NLP)*. CV Jejak (Jejak Publisher), 2023. [Online]. Available: https://books.google.co.id/books?id=FTnhEAAAQBAJ

[14] S. E. M. M. Akbar Bahtiar *et al.*, *MEMAHAMI METODOLOGI PENELITIAN*. Cendikia Mulia Mandiri, 2024. [Online]. Available: https://books.google.co.id/books?id=xJTsEAAAQBAJ

[15] B. Christiansen, L. Willyanto, and H. Juwiantho, "Pembuatan Website untuk Rekomendasi Smartphone," *J. Infra*, no. 031, 2020, [Online]. Available: http://publication.petra.ac.id/index.php/teknik-informatika/article/view/9996%0Ahttp://publication.petra.ac.id/index.php/teknik-informatika/article/viewFile/9996/8955

[16] S. T. M. Dr. Adi Nugroho, *DATA SCIENCE MENGGUNAKAN BAHASA R: Analisis Data, Visualisasi, serta Pemodelan*. Penerbit Andi, 2024. [Online]. Available: https://books.google.co.id/books?id=CdUHEQAAQBAJ