G-Tech : Jurnal Teknologi Terapan
Volume 7, No. 3, Juli 2023, hal. 958-966
E-ISSN: 2623-064x | P-ISSN: 2580-8737

Prediction of Alleged Stress Symptoms based on Indonesian Sentiment Lexicon using Multilayer Perceptron

Monika Evelin Johan[1]✉, Sheyla Aurelia Azka[2]

[1,2] Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara, Indonesia

Informasi Artikel

Riwayat Artikel

Diserahkan : 24-05-2023
Direvisi : 29-05-2023
Diterima : 05-06-2023

ABSTRAK

Stres adalah fenomena mental atau fisik yang terbentuk melalui penilaian kognitif seseorang terhadap rangsangan dan hasil interaksi seseorang dengan lingkungan. Layaknya penyakit lainnya, stres harus segera ditangani, agar tidak mengganggu keseharian hidup seseorang. Namun, diagnosis mandiri perlu dihindari untuk mencegah terjadinya penanganan yang keliru. Metode terbaru mendeteksi stres melalui media sosial, dimana psikiater dapat mengenali gejala stres seseorang berdasarkan postingan yang terus-menerus di media sosial. Penelitian ini menggunakan dataset dari Twitter, dengan kuesioner DASS-42, model algoritma Multilayer Perceptron dan Indonesian Sentiment Lexicon. Analisis terhadap tweet dapat membentuk model prediktif yang diterapkan untuk mendeteksi sentimen serupa di tweet lain. Percobaan dua kasus uji, yaitu dengan parameter *Adam solver* menghasilkan akurasi 86%, sedangkan dengan parameter *SGD solver* menghasilkan akurasi 72%. Hal itu dikarenakan *Adam solver* bekerja lebih baik dari segi waktu pelatihan dan skor validasi pada kumpulan data yang relatif besar.

Kata Kunci:

Analisis sentimen, gejala stress, *Indonesian Sentiment Lexicon, Multilayer Perceptron*, Twitter.

ABSTRACT

Stress is a mental or physical phenomenon that is formed through a person's cognitive assessment of stimuli and the results of one's interaction with the environment. Just like other diseases, stress must be treated immediately, so it does not interfere with one's daily life. However, self-diagnosis needs to be avoided to prevent erroneous handling. The newest method of detecting stress is through social media, where a psychiatrist can recognize a person's symptoms of stress based on persistent postings on social media. This study uses a dataset from Twitter, with the DASS-42 questionnaire, the Multilayer Perceptron algorithm model, and the Indonesian Sentiment Lexicon. Analysis of tweets can form predictive models that are applied to detect similar sentiments in other tweets. Experiment with two test cases, namely the Adam solver parameter produces an accuracy of 86%, while the SGD solver parameter produces an accuracy of 72%. That's because Adam's solver works better in terms of training time and validation scores on a relatively large data set.

***Keywords* :**

Indonesian Sentiment Lexicon, Multilayer Perceptron, sentiment analysis, stress, Twitter.

***Corresponding Author*:**
Monika Evelin Johan
Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara
Jalan Scientia Boulevard Gading, Curug Sangereng, Serpong, Kabupaten Tangerang, Banten 15810
Email: monika.evelin@umn.ac.id

INTRODUCTION

The use of social media and online news sites is increasing from year to year, giving rise to new events. Everyone is free to express anything through their social media accounts. They can easily share on social media accounts and anyone can comment on it (Juditha, 2017). So is Twitter, which has 17 million active users (Aslam, 2022). User-generated tweets convey feelings, core personalities, emotions, and even anxiety. For example, venting feelings online (Jalonen, 2014). This can be used to do early detection of stress which in the long run will really help to avoid serious problems. Since messages via social media have become a popular medium of communication, the ability to detect depression in content is very useful (Lin, et al., 2017).

Mental health has become a very important issue for the well-being of people in the world community, in 2021 there will be around 47.6 million people affected by mental illness (Mental Health Statistics, 2021). Stress can lead to poor health (Pillai, Reshmi, Thelwall, & Orasan, 2018). Stress is a mental or physical phenomenon that is formed through a person's cognitive assessment of stimuli and is the result of a person's interaction with the environment (Bhargava & Trivedi, 2018). A person who is under pressure may feel emotionally, physically, or even mentally uncomfortable, and will feel that he or she is not going to succeed at something. (Lin, et al., 2017).

Research is ongoing, to identify people with suspected symptoms of stress. There are many literature studies that study stress detectors. Several traditional and scientific methods are used to identify people who are under stress, such as questionnaires, sensor measurements, and social media (Pillai, Reshmi, Thelwall, & Orasan, 2018). Psychiatrists administer large questionnaires and use the answers to determine whether a person has symptoms of stress or not. This method has its own limitations and drawbacks because often the answers are not factual, checking is wrong because the expert has to check all the answers in large numbers which results in fatigue and negligence in making decisions. Another method is sensor measurement. This method has the limitation that it is time-consuming and a bit expensive.

Another newest method for detecting stress is social media (Pillai, Reshmi, Thelwall, & Orasan, 2018). Stress detection is possible through social media, based on one's writing on social media, one's reaction to certain problems on social media, and one's likes and dislikes. Through continuous writing or posting on social media, a psychiatrist can find people under pressure, stress, and madness with a typical subject. The use of social networking sites analyses a person's state of mind and thinking (De Choudhury, Gamon, Counts, & Horvitz, 2021). The analysis of social media is a basic approach to studying social relationships to understand the relationship between people in social groups (Alamsyah & Indraswari, 2017). Social media analysis is described as the study of human relations through graph theory. Social media studies are defined as a collection of nodes or network members and are linked by different types of relationships called links. Early attempts to detect stress in Twitter data have been promising. In (Jones, Wojcik, Sweeting, & Silver, 2016), determined that Twitter users who suffer from stress tend to post tweets containing more negative emotional sentiments compared to healthy users. Sentiment refers to the attitude of the speaker or writer towards a topic (Yuan, Wang, Yin, & Wang, 2021).

Multilayer Perceptron is a class of feedforward neural networks constructed by layered acyclic graphs. Multilayer Perceptron consists of at least three layers, namely, input layer, hidden layer, output layer, and non-linear activity. The hidden unit processes the input, calculates the result of the activation function for each neuron, and passes the result to the next layer. The output layer from the input layer is received as input to the hidden layer. Similarly, the hidden layer sends the result to the output layer which is called feedforward. A commonly used algorithm for Multilayer Perceptron training is backpropagation. The word backpropagation has the meaning of parameters (synaptic weights) gradually based on errors (output compared to the desired output). Most importantly modify the synaptic weights from the output layer to the hidden layer. This error is then propagated to the previous layer. That is, changes in synaptic weight in one layer are affected by changes in synaptic weight in the next layer. Backpropagation is nothing but a gradient-based optimization method applied to ANN (Putra, 2019).

There is a study on the use of sentiment analysis to analyze sentiment using a lexicon-based approach by using InSetLexicon as an Indonesian opinion dictionary (Desi, Khaira, Utomo, Utomo, & Suratno, 2021). Another study discusses sentiment analysis on social media Twitter by combining the lexicon and machine learning methods. Identify the most common types of social media sites to extract information for sentiment analysis (Drus & Khalid, 2019). Other research (Livingston, Selvi, & Jenifer, 2019) discusses the comparison of machine learning algorithms such as the SVM algorithm, Naïve Bayes, which has accuracy results using Multilayer Perceptron is analyzed to be 93% more accurate and precise compared to the Naïve Bayes algorithm and Support Vector Machine. Besides, research (Stephen & Prabu, 2019) discusses the use of algorithms as an adjutant for medical observers to understand patients, using the calculated sentiment scores can be combined with different emotions to provide a better method for calculating depression scores. Research (Tirtopangarsa & Maharani, 2021) discusses building a system that predicts depression levels in Twitter users using the K closest neighbor classification, using the DASS42 questionnaire score, there are 15 accounts experiencing major depression, very severe depression, and moderate depression. The depression prediction results have different values. The higher the percentage rate, the more likely the account will be negative. The study (Mara, Sediyono, & Purnomo, 2021) discusses sentiment analysis from Facebook comments and the pre-processing method which is divided into positive and negative using the K-nearest Neighbours algorithm and Rapid Miner software to process and find the accuracy, precision, and recall value.

Based on several previous studies, this research will use the Indonesian Sentiment Lexicon as an automatic dictionary of sentiment labelling and use the Multilayer Perceptron (MLP) method. InSetLexicon (Indonesian Sentiment Lexicon) consists of 3,609 positive words and 6,609 negative Indonesian words, each with a polarity value ranging from -5 to +5. The Lexicon InSet is designed to identify existing tweets and classify them into positive, neutral, or negative opinions that can be used to analyze public opinion about a particular topic, event, or product (Desi, Khaira, Utomo, Utomo, & Suratno, 2021). Keyword phrases extracted and defined from each DASS-42 statement to find tweets and can help predict the use of expressing negative sentiments or experiencing symptoms of stress disorder. By analyzing these tweets, we can form predictive models that can be applied to detect similar sentiments in other tweets. Using the Multilayer Perceptron algorithm and the Indonesian Sentiment Lexicon will generate patterns, providing a deeper understanding of the relationship between the semantics used and user sentiment.

METHODOLOGIES

Data was collected by interviewing experts aged 25-35 years, with at least a Master's degree in Psychology, and knowing the DASS-42 measurement tool. The DASS-42 is described in Table 1. Statements based on the DASS-42 are below. In addition, questionnaires were distributed to respondents aged 21 -25 years, currently or have graduated with a bachelor's degree in psychology, have a minimum grade of B or 80 in language exams or courses, and have been Twitter users for at least 2 years. Determination of keywords by determining 3 words from sources and 2 words from questionnaire respondents according to their respective statements based on DASS-42. After the keywords are collected, the data will be scraped on Twitter social media according to the keyword sentences. The results of the keyword sentences will be scraped as many as 1000 tweets in each keyword sentence. After finishing scraping, it will be combined into a dataset. Once collected, the data is aggregated, removing special characters like (!"#$%&()*+-/:;<=>?@[\\]^_`{|}~) or numbers (0 -9) other than [a-z] and [A-Z], removing double spaces removes sentence and lower case matching.

The next stage of text pre-processing tweet data is done using Jupyter Notebook with Python programming language. The process is as follows:

1. Removing mentions or Twitter usernames in deleted tweets using the 'replace' function in the 'string' library.
2. Removes hashtags from tweets.
3. Remove words that contain hyperlinks in tweets.

4. Remove the word RT in tweets.
5. Eliminate the special characters to be removed from each tweet are in the punctuation set in the 'string' library: !"#$%&'()*+, -./:;<=>?@[\]^_`{|} ~.
6. Eliminate all numbers in the tweet using the 'replace' function in the 'string' library.
7. Eliminate all double spaces in individual tweets.
8. Remove all spaces at the beginning of the tweet.
9. Remove all spaces at the end of the tweet.
10. Eliminate all words with repeated letters and the repeating letters are only 1 letter.
11. Eliminate the same sentence and only one of them will be selected.

Table 1. Statements based on DASS-42

No	STATEMENTS	No	STATEMENTS
1	I feel that I get angry over trivial things.	22	I find it hard to rest.
2	I feel that my lips are often dry.	23	I have difficulty swallowing.
3	I couldn't feel any positive feelings at all.	24	I cannot feel pleasure from the things I do.
4	I have difficulty breathing (for example: often panting or unable to breathe even though I did no physical activity before).	25	I am aware of heart activity, even though I have not been physically active (eg feeling my heart rate increase or decrease).
5	I don't seem to have the strength to do an activity anymore.	26	I feel hopeless and sad.
6	I tend to overreact to situations.	27	I feel that I am very irritable.
7	I feel unsteady (for example, my legs feel like they are going to fall off).	28	I feel I'm almost panicking.
8	I find it hard to relax.	29	I find it hard to calm down after something has upset me.
9	I found myself in a situation that made me feel very anxious and I would feel very relieved when this was all over.	30	I fear that I will be 'bogged down' by trivial tasks that I am not used to doing.
10	I feel that there is nothing to look forward to in the future.	31	I don't feel enthusiastic in any way.
11	I find myself easily irritated.	32	I have a hard time being patient with interruptions in what I am doing.
12	I feel like I've wasted a lot of energy worrying.	33	I'm feeling restless.
13	I feel sad and depressed.	34	I feel that I am worthless.
14	I find myself getting impatient when I encounter delays (eg: traffic jams, waiting for something).	35	I can't afford anything to stop me from finishing what I'm doing.
15	I feel weak like I'm going to pass out.	36	I feel very scared.
16	I feel I have lost interest in everything.	37	I see no hope for the future.
17	I feel that I am worthless as a human being.	38	I feel that life is meaningless.
18	I feel that I am easily offended.	39	I find myself easily agitated.
19	I sweat excessively (for example: sweaty hands), even though the temperature is not hot or I haven't done any physical activity before.	40	I feel worried about situations where I might panic and embarrass myself.
20	I feel scared for no apparent reason.	41	I feel shaking (eg: in my hands).
21	I feel that life is not worthwhile.	42	I find it difficult to increase initiative in doing things.

The next stage of pre-processing tweet data is done using Jupyter Notebook with Python programming language. The process is as follows:

1. Changing Capitalization: All characters or letters in tweets are converted to lowercase, using the 'pandas' library.

2. Tokenizing: The tokenizing process to separate sentences into individual words.
3. Removal of Stop words: Stop words are words that have no meaning. Stop words need to be removed to avoid sentiment analysis inaccuracies. Example: deleting the words "for", "when", "at", "can", "with". Words including stop words in Indonesian are taken from the NLTK library.
4. Stemming: Stemming is the process of converting words into basic words and parsing the form of words.
5. TF-IDF Word Weighting: The feature extraction process is carried out using the Term Frequency-Inverse Document Frequency (TF-IDF) for weighting in a collection or corpus.

For the modeling stage, the extraction results were divided into training data and test data with a ratio of 70:30. Then, the data will run on the model used for the classification process, namely the Multilayer Perceptron. Finally, Evaluation of the classification results is assessed based on the values of accuracy, precision, recall, and F1-score.

RESULT AND DISCUSSION

Collecting Keywords

Keywords were collected by conducting expert interviews and distributing online questionnaires to 7 respondents according to the criteria. Interviews were conducted with 2 experts which were held on March 1st, 2022 by the first expert and March 31st, 2022 by the second expert. The results of collecting keywords from experts were 84 sentences while the respondents were 294 sentences. All keyword sentences are collected for text processing. Furthermore, the respondent's and expert's keywords will be used as a dataset before cleansing, such as deleting special characters, deleting double spaces, deleting the same sentence, and changing sentences to lowercase. The stored dataset will be selected by eliminating the same keyword sentences. Keywords that have been removed with special characters will be deleted with double spaces. After cleansing, remove the double spaces. The next step of cleansing the keyword sentence is eliminating the same keyword sentence. The initial keyword sentences numbered 294 to 210. There are 84 similar keyword sentences. For the same keywords, only one will be selected. In Table 2 are the top 20 keyword sentences out of a total of 294 keyword sentences that have been selected by the respondents.

Table 2. Keyword Collection Results

No	Keywords	No	Keywords
1	angry, trifle	11	often, dry
2	I, angry	12	I, often
3	angry, trifle	13	lips, dry (empty)
4	angry, trifle	14	lips, often
5	feeling, angry	15	no, positive
6	angry, stuff	16	no, feelings
7	angry, trifle	17	no, sense
8	lips, dry (empty)	18	no, sense
9	I, feel	19	feelings, positive
10	feel, often	20	feelings, positive

Scrapping Data

The process of scraping keyword sentences is carried out 210 times, according to the keyword sentences that have been processed previously. Using TwitterSearchScraper with the Twitter package for scraping. Each scraping result is saved in a file in CSV format. After each keyword is scraped, it will be used as a dataset, a total of 170,460 tweets resulting from scraping.

Text Preprocessing

At this stage, several text preprocessing will be carried out, namely eliminating mentions, hashtags, hyperlinks, RT, special characters, numbers, double spaces, sentence beginning spaces, sentence ending spaces, repeated letters, and the same tweet. because the tweet data is still in the form of raw data. The text processing stage is carried out using the help of a library in the Python3 programming language. Here are some examples of the implementation of each text processing. The last stage in text processing removes the duplicated tweet. Tweets that initially numbered 170,465 to 133,035, there are 37,430 of the same tweet. The same tweet will be retrieved from one tweet. The results of the last stage of text processing are in Table 4, stored in a CSV format dataset.

Preprocessing

At this stage, the sentences will be changed to lowercase, tokenizing, removing stop words, and stemming. Steps to change the lowercase sentence using lambda x.lower(). The tokenization stage is the stage of separating words, symbols, phrases, and other important units (called tokens) from the text. the tokenization stage is carried out using a library called NLTK. The remove stop word stage is a step to remove words that have no meaning. The process of removing stopwords from the data is done by first defining the words listed as stopwords. Using the NLTK library with set(stopwords.words('indonesian')) which is used as a list of stop words in Indonesian. The stemming stage is a step to change the word into its basic form by removing the pre and post-affixes. The stemming stage is carried out using the Literature library by utilizing an Indonesian-language stemmer.

Indonesian Sentiment Lexicon

The stored dataset resulting from stemming is then labeled using the Indonesia Sentiment Lexicon automatically to distinguish positive, negative, and neutral sentiments. The final result of this study is a class consisting of 3 categories, namely positive, neutral, and negative. Negative class means that the tweet is detected to contain words that can be predicted to have symptoms of stress. The positive class means that the tweet contains words that can be predicted as early symptoms of stress. While the neutral class means that there are absolutely no words in the tweet that describe the alleged symptoms of stress. Sentiment that has a polarity of more than 0 will be categorized into positive sentiment. Sentiments whose polarity results are less than 0 will be categorized into negative sentiment. If not, both will be categorized as neutral.

The results of the stage of determining the polarity of sentiment from labeling are 65% or 86433 negative tweets, 28.6% or 38081 positive tweets, and 6.4% or 8521 neutral tweets. Based on the results of stemming carried out, it is necessary to know the frequency of words that are often used by Twitter social media users in the dataset. Visualization results in word cloud form in Figure 1.

Figure 1. Word Cloud Results from the Dataset

TF-IDF

The vector value is based on the presence of words in the standard words in the document. A value of 1 or the number of words found to appear is given if there are words related to the document, if not, a value of 0 is given. The data that has become a word vector is then calculated TFID to produce a weighted word vector. TF (Term Frequency) is the frequency of occurrence of terms in the document, and IDF (Inverse Document Frequency) is the result of the calculation of the spread of terms in a collection of documents.

Multilayer Perceptron Implementation

After calculating the TF-IDF weights, the data is divided into training data and test data using the train_test_split library. The comparison used in the distribution of data is 70:30 for practice and test. there are 93,124 for training data and 39,911 for testing data using random_state = 42. Then, the creation and selection of test scenarios aimed at using Multilayer Perceptron (MLP) as a classification model will be carried out. Before selecting the test scenario, this research was conducted by conducting trials using 3 different solvers, namely LBFGS, Adam, and SGD. It turns out that the results of the accuracy of LBFGS and SGD are the same. In the end, using only 2 test scenarios with 2 different parameters.

Based on the results of the existing scenario, the accuracy value is obtained with the first parameter with solver 'Adam', hidden_layer_sizes=(150,100,50), max_iter = 300 using activation = 'relu', by 86%. While the accuracy value obtained with the second parameter with solver 'SGD' with hidden_layer_sizes = (1000,500,200), learning_rate = 'adaptive', using activation = 'relu' and random_state = 42, by 72%.

There are several evaluations to see the performance of the classification method for each class, which can be seen through the values of precision, recall, and F1-score. The results of precision, recall, and F1-score values have a value of 0-1. The closer to 1, the better the value. The following is in Table 2 the results of the evaluation.

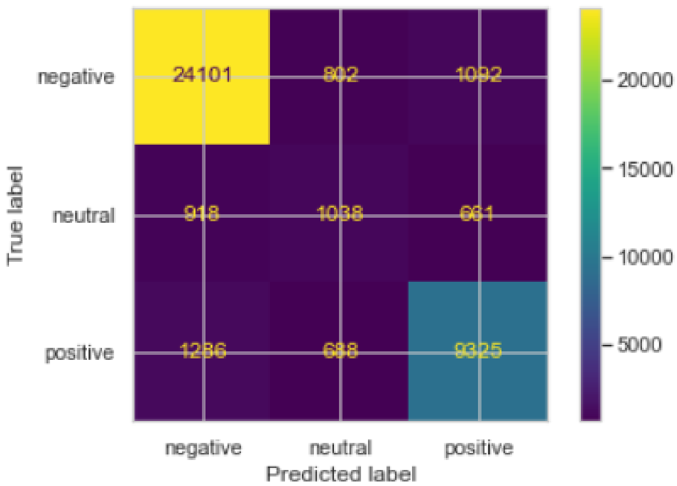

Figure 2. Confusion Matrix Results of the First Parameter Model Test (Adam)

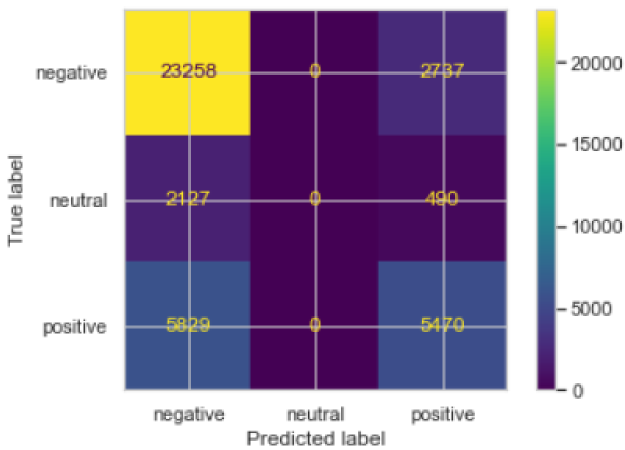

Figure 3. Confusion Matrix Results of the Second Parameter Model Test (SGD)

Table 3. Model Evaluation Results

Adam solver	**Precision**	**Recall**	**F1-score**
Negative	0.92	0.93	0.92
Neutral	0.41	0.40	0.40
Positive	0.84	0.83	0.83
SGD solver	**Precision**	**Recall**	**F1-score**
Negative	0.75	0.89	0.81
Neutral	0.0	0.0	0.0
Positive	0.63	0.48	0.55

Evaluation results are not better than previous studies. There are several factors, such as the weighting of the TF-IDF word in parameters using min_df. The use of min_df in this study aims to minimize the use of limited hardware memory. When not using the min_df parameter, it cannot perform TF-IDF steps.

CONCLUSION

Based on the results of testing using the Multilayer Perceptron algorithm that has been carried out, there are several conclusions. The use of social media Twitter is one way for psychologists to find out the personality of their patients through what is said in their accounts. The class results in this classification are the opposite of the meaning, the negative class means that the tweet is positively detected as a predictor of alleged stress symptoms. While the positive class means that negative tweets are detected as a predictor of early presumptive symptoms of stress. The use of the Multilayer Perceptron algorithm produces a good accuracy value for the first parameter using Adam solver rather than the second parameter using SGD solver. The results of accuracy, precision, recall, and F1-score are influenced by solver parameters. Because using the Adam solver will work well on relatively large data sets (with thousands of training samples or more) in terms of both training time and validation scores..

The results of the research carried out still have many shortcomings, therefore hope this research can be developed. Determining effective keywords and not using too much training data, when weighting TFIDF words, cannot be used as a whole, due to the limited memory of the hardware used. Using a combination of other parameters and adding k-fold steps to increase the accuracy of the results. In this research, it is hoped that a website platform can be created that can run this model, starting from automatic scrapping so that later it can become big data to run on the model that has been created.

REFERENCES

Alamsyah, A., & Indraswari, A. A. (2017, April). Social Network and Sentiment Analysis for Social Customer Relationship Management in Indonesia Banking Sector. pp. 3808-3812.

Aslam, S. (2022). *Twitter by the Numbers: Stats, Demographics & Fun Facts*. Retrieved from https://www.omnicoreagency.com/: https://www.omnicoreagency.com/twitter-statistics/#:~:text=On average%2C 500 million tweets,200 billion tweets every year

Bhargava, D., & Trivedi, H. (2018). A study of causes of stress and stress management among youth. *IRA-International Journal of Management & Social Sciences*, 108-117.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2021). Predicting depression via social media. *Proceedings of the international AAAI conference on web and social media*, (pp. 128-137).

Desi, M., Khaira, U., Utomo, P., Utomo, P. E., & Suratno, T. (2021). nalisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon: Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 24-33.

Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, 707-714.

Jalonen, H. (2014). Social media–an arena for venting negative emotions. *Online Journal of Communication and Media Technologies*, 53-70.

Jones, N. M., Wojcik, S. P., Sweeting, J., & Silver, R. C. (2016). Tweeting negative emotion: An investigation of Twitter data in the aftermath of violence on college campuses. *Psychological methods*, 526.

Juditha, C. (2017). Hatespeech in Online Media: Jakarta on Election 2017. *Jurnal Penelitian Komunikasi Dan Opini Publik, 2*, 21.

Lin, H., Jia, J., Qiu, J., Zhang, Y., Shen, G., Xie, L., . . . Chua, T.-S. (2017). Detecting stress based on social interactions in social networks. *IEEE Transactions on Knowledge and Data Engineering*, 1820-1833.

Livingston, J. S., Selvi, T., & Jenifer, C. S. (2019). A neural network based approach for sentimental analysis on amazon product reviews. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 469-473.

Mara, A. A., Sediyono, E., & Purnomo, H. (2021). K-Nearest Neighbors Algorithm to Student Opinion of the Online Learning Method at Wira Wacana Sumba Christian University. *Ultima Infosys: Jurnal Ilmu Sistem Informasi*, 87-93.

Mental Health Statistics. (2021). Retrieved from https://www.thezebra.com/: https://www.thezebra.com/resources/research/mental-health-statistics/

Pillai, G., Reshmi, Thelwall, M., & Orasan, C. (2018). Detection of stress and relaxation magnitudes for tweets. *Companion Proceedings of the The Web Conference 2018* (pp. 1677-1684). International World Wide Web Conferences Steering Committee.

Putra, J. W. (2019). Pengenalan konsep pembelajaran mesin dan deep learning. *Comput. Linguist. Nat. Lang. Process. Lab*, 1–235.

Stephen, J. J., & Prabu, P. (2019). Detecting the magnitude of depression in Twitter users using sentiment analysis. *International Journal of Electrical and Computer Engineering*, 3247.

Tirtopangarsa, A. P., & Maharani, W. (2021). Sentiment Analysis of Depression Detection on Twitter Social Media Users Using the K-Nearest Neighbor Method. *Seminar Nasional Informatika (SEMNASIF)*, (pp. 247-258).

Yuan, X., Wang, L., Yin, X., & Wang, H. (2021). How text sentiment moderates the impact of motivational cues on crowdfunding campaigns. *Financial Innovation*, 46.