EEG Signal Classification using K-Nearest Neighbor Method to Measure Impulsivity Level

Arico Sempana Ginting^[1], Ruth Marsaulina Simanjuntak^[2], Nurima Lumbantoruan^[3], Delima Sitanggang^{[4]*}

Department of Sciences and Technology^{[1], [2], [3], [4]}

Universitas Prima Indonesia

Medan - Indonesia

delimasitanggang@unprimdn.ac.id^[1] xiomi123588@gmail.com^[2],ruth010402@gmail.com^[3],nurimasiltor@gmail.com^[4]

contexts[2].

Abstract— Impulsivity is the tendency to act without considering consequences or without careful planning. It involves a quick response to a stimulus without sufficient consideration of the consequences. Impulsivity needs to be measured and detected because it has a significant impact on various aspects of a person's life. The factors that influence the level of impulsivity include social environment, stress level, mental health, and genetic factors. Impulsivity can be divided into multiple components, such as reduced sensitivity to unfavorable behavioral outcomes, a disregard for long-term implications, and quick and spontaneous responses to stimuli. Electroencephalogram (EEG) studies can identify specific brain wave patterns such as, Alpha, Betha, Theta, and Gamma waves everything based on an individual brain's level of impulsivity. Signals from the brain are processed to extract specific features that reflect the user's intentions. EEG records brain activity without surgery, and this information is used for the diagnosis, monitoring, and treatment of neurological diseases, as well as scientific research on the brain and mind. K-Nearest Neighbor (KNN) is a classification algorithm that functions by utilizing several K nearest data values (its neighbors) as a reference to determine the class of new data. The K-Nearest Neighbors (KNN) algorithm is used for classification, clustering, and pattern recognition in EEG data where clustering is in 4 classifications (Impulsive, Not Impulsive, Potentially Impulsive, and Very Potentially Impulsive). This classification model shows high accuracy (Training Data: 94.7%, Testing: 91.3%, and Validation Data: 91.8%). This research shows that the KNN algorithm is effective for assessing the degree of impulsivity.

Keywords: Impulsivity, EEG Signal, K-Nearest Neighbors, Brainwaves, Classification.

I. INTRODUCTION

Assessing a person's level of impulsivity is important as it can significantly impact their quality of life in various areas. Impulsivity can be categorized into different components, such as a reduced sense of consequences for bad behavior, negligence towards long-term effects, and an instant, unplanned response to stimuli. However, given that impulsive behavior might lead to distractions from relevant conversations or swift reactions to irrelevant ones, It's feasible that involvement could not be necessary [1]. One of the most controversial areas of human technology is the identification of lying, which has significant implications in ethical, legal, and medical The scientific community has recently expressed a growing interest in this areaImpulsive people have several traits, including heightened distractibility, acting carelessly, having trouble reining in incorrect reactions, and overreacting to rewards that come in the short term[3]. An Electroencephalogram (EEG) test can be used to gauge an individual's impulsivity.

Electroencephalogram (EEG) captures the electrical activity of the brain, which aids medical professionals in determining whether there is a disruption in the brain region linked to impulsivity[4]. Because they exercise critical thinking, exercise self-control, consider the repercussions of their actions before acting, and are more alert, those with lower levels of impulsivity are typically more cautious. However, impulsivity is affected by several factors, including genetics, stress, mental health, and social and environmental influences.

Researchers frequently employ electroencephalogram (EEG) and brain-computer interface (BCI) to analyze and identify brain activity, including forecasting the emotions that occur [5]. The level of impulsivity can be measured through various methods and evaluation tools that have been developed by researchers and professionals in the field of psychology.

Based on the survey results, the level of consumer impulsivity measured through indicators of spontaneity, seeing buying immediately, acting without thinking, and buying now shows that consumers tend to be easily tempted to buy products without careful consideration [6]. Signals from the brain are processed to extract specific features that reflect user intentions[7].

Electroencephalogram (EEG) records brain activity without surgery, and this information is used for the diagnosis, monitoring, and treatment of neurological diseases, as well as scientific research on the brain and mind[8]. Measuring the brain's response to a stimulus using Electroencephalogram (EEG) has become one of the commonly used methods in cognitive neuroscience. This method relates physiological activity to information processing, sensory, perceptual, and cognitive activity[9].

p-ISSN 2301-7988, e-ISSN 2581-0588 DOI : 10.32736/sisfokom.v13i2.2154, Copyright ©2024 Submitted : April 30, 2024, Revised : May 21, 2024, Accepted : May 21, 2024, Published : June 15, 2024 Electroencephalogram (EEG) studies can identify specific brain wave patterns such as, Alpha, Betha, Tetha, and Gamma waves all depending on the level of impulsivity of a particular brain.

By studying the characteristics of electroencephalogram (EEG) waves, researchers can understand various mental states, such as levels of focus, relaxation, and awareness. When we're awake, eyes closed, and calm, the brain is in charge by alpha waves (8–13 Hz). Beta waves, which range from 14 to 30 Hz, become more active while we concentrate. During light sleep, drowsiness, or stress, theta waves (4–7 Hz) appear. Deep sleep is accompanied by delta waves (0.5–3 Hz), while gamma waves (30–50 Hz) signify a fully conscious state [10]. The P300 component proved to be a reliable indicator of key brain states [11]. The P300 reflects the brain's response to an unexpected stimulus or requires additional cognitive processing. Some studies suggest that theta wave frequency in an electroencephalogram (EEG) may be related to impulse control.

The use of electroencephalograms (EEGs) began in 1924 when Hans Berger, a German physiologist and psychiatrist, recorded the electrical activity of the human brain for the first time. This discovery was later confirmed and developed by other scientists. In 1935, Gibbs, Davis, and Lennox discovered the interictal spike-wave and 3 Hz spikeand-wave complex in the absence of seizures. Gibbs and Jasper also found interictal spike waves as a clue to focal epilepsy [12]. Along with the advancement of technology, electroencephalogram (EEG) machines and examination techniques have progressed rapidly. Initially, electroencephalogram (EEG) machines used pen techniques that directly printed recordings on paper. These days, direct processing of brain electrical data is made feasible by computers. This allows the preparation of various montages in the same unit of time [13].

The tool developed is the electroencephalogram (EEG), which can capture the brain's electrical activity and inform the state of mind, such as emotion, alertness, concentration, and fatigue. In EEG-based BCI applications, the shift of correlated variables to independent variables is the most common case when incorporating different distributions of training and testing samples [14]. In this case, this research will focus on developing a classification method using the K-Nearest Neighbor (KNN) algorithm to analyze EEG signals and identify the level of Impulsivity. K-nearest neighbor (KNN) is a classification algorithm that works by taking several K nearest data values (its neighbors) as a reference to determine the class of new data [15].

The reason for why using this K-Nearest Neighbor (KNN) method, is that it has the advantage of being able to classify unknown employee candidate data in the presence of training data and test data [16]. K-Nearest Neighbor (KNN) can perform mathematically based procedures to evaluate the value of these criteria into a classification statement. This method can classify data accurately by first selecting the value of K-nearest neighbors appropriately. K-Nearest Neighbor

(KNN) can also sort out candidate data sets that are classifiable into good, best, and poor[17].

Previous researchers (Decoding Neural Signatures: Develop Drug Addiction Level through Electroencephalogram (EEG) Patterns based on Adaptive Neuro-Fuzzy Inference) have used the Adaptive Neuro-Fuzzy Inference System (ANFIS) method to classify drug addiction levels based on EEG patterns. The ANFIS method offers a high degree of accuracy and good adaptability[18]. However, the current research proposes using the K-Nearest Neighbor (KNN) method. This method has several advantages over ANFIS, especially in terms of effectiveness for large data, robustness to noise, and ease of implementation. KNN is more suitable for situations with large amounts of data and high levels of noise due to its simplicity and ease of implementation.



Fig 1. Flowchart EEG

A. Data Collecting

The method was conducted through data collection in one of the correctional institutions in North Sumatra in collaboration with Prima Indonesia University and Padjajaran University Bandung. The experiment involved 21 male subjects aged between 20 to 30 years old. The purpose of this Electroencephalogram (EEG) experiment is to group the subjects into 4 classifications based on their respective levels of impulsivity.

The human brain is a complex and mysterious organ. Its activity can be studied in various ways, one of which is by recording its electrical activity. This recording is done using tools and materials such as Win EEG software, amplifiers, gels, and electrocap. The electrocap, which is equipped with many electrodes, is attached to the scalp. These electrodes serve to capture the electrical signals generated by the brain. These signals are then amplified by an amplifier and recorded by the Win EEG software. Gel was used to increase the contact between the electrodes and the scalp, making the signals clearer.

B. Prepocessing Data

This experiment was conducted for two minutes with eyes closed. The electrodes were placed right on the subject's head

using a fastening device in the form of a belt. To collect data and life history information from each subject, the researcher used the interview technique can be seen in Figure 1 below:



Fig 2. Electrical Activity in the brain

The Electroencephalogram (EEG) recorder must be controlled by a computer in order to analyze the collected data. The condition of the room—which is devoid of outside distractions like light and sound—also affects the outcomes. Appropriate documentation sessions were employed when the subjects' behavior during the experiment was being observed.

Electroencephalogram (EEG) software is used to analyze and interpret recorded data and brain activity.Mitsar provided the Electroencephalogram (EEG) data used in this experiment. Two of the twenty-one electrocap channels—A1 and A2—were auricle electrodes. The remaining electrocap channels were distributed around the scalp in designated places. The left ear area is where electrodes A1 and A2 were positioned. A2 was positioned in the area of the right ear. For the electrodes in the ear or auricle, a composite reference electrode is utilized can be see in figure 2 :



Fig 3. Mitsar EEG

In the field of data processing, the K-Nearest Neighbor (KNN) algorithm is one of the classification techniques that classifies a set of data based on classification information or labels from previous learning data. KNN belongs to the category of supervised learning, where new data is classified based on most similarities with the existing categories around it in the KNN group.

The steps to calculate the K- Nearest Neighbor method are as follows: Determining the K parameter, Calculating the distance between training data and testing data, Sort the distance formed, Determining the closest distance to the order K, Pairing the corresponding classes, Find the number of classes from the closest neighbour and set that class as the data class to be evaluated.

The simplest data mining technique, one of which is KNN. It is commonly referred to as memory-based classification, i.e. the training data needs to be in memory at run-time. When dealing with continuous attributes, the difference between the attributes is calculated using Euclidean distance. If the first instance is (a1, a2, a3, ..., an) and the second instance is (b1, b2, b3, ..., bn)[19], then the distance between them is calculated using the equation:

$$d = \sqrt{(a_1 + b_1)^2 + (a_2 + b_2)^2 \dots (a_n + b_n)^2}$$

where: a: test data, used to test the modelling generated from the training data, b: training data, used to find the right modelling, n: nth data.



Fig 4. Block Diagram (Brain Computer Interface)

Stimuli were administered to 21 subjects at the beginning of the Brain-Computer Interface (BCI) process, Mitsar 201 EEG Amplifier, which has 21 channels, was used to record brain signals with high accuracy after the stimuli were administered. Preprocessing was done to remove noise. Preprocessing is done using filters to filter out the required signals[20].

The K-Nearest Neighbors (KNN) algorithm is used to perform the classification process, this algorithm clusters and finds patterns in the data created by the brain. Next, the classification results pass through an impulsivity level interface. This interface allows the system and the user to interact with the BCI system and allows an impulse level interpretation or response to the classification results. As the final element of the block diagram, the user can provide additional instructions or receive feedback that is based on the impulsive response of the system. Overall, this procedure results in a system that can link brain activity affected by stimuli with impulsive level responses. This is achieved through the processes of data acquisition, filtration, feature extraction, classification, and user interface located within the Computer Brain Interface structure.

III. RESULTS AND DISCUSSION

A. Data Classification

In the initial stage of analysis, the raw Electroencephalogram (EEG) data will be classified based on the brainwave type for each subject. Furthermore, the data will be grouped into 4 categories of impulsivity, namely, Impulsive (Range 16.35-21.83), Potentially Impulsive (Range 10.9-

16.35), Potentially Impulsive (Range 5.45-10.9), Not Impulsive (Range 0-5.45). This grouping is done to facilitate the analysis and interpretation of EEG data related to the level of individual impulsivity can be see in Figure 5:



B. Data Filtering

This research began by collecting initial data from the research subjects. The data obtained is still raw and contains high noise. Therefore, the researchers used a BPF (Band Pass Filter) filter with a frequency range of 0.5 to 50 Hz to improve the data quality. BPF is designed to select signals in a specific frequency range, such as alpha, theta, beta, and other waves. The purpose of applying a BPF filter is to remove unwanted noise and focus the analysis on specific brainwaves. The displayed image shows the impact of BPF on the data. The BPF successfully removed unwanted noise, making the data cleaner and easier to analyze. This result allowed the researchers to focus on analyzing the desired brainwaves and produce more accurate findings can be seen in Figures 6 and 7:



Fig.6 Data before filtering



Fig.7 Data after filtering

This research focuses on five types of brainwaves: alpha, theta, delta, beta 1, and beta 2. Particularly in the eyes-closed condition, the dominant wave type will change according to the state of the subject. When the subject is very sleepy and

experiencing light sleep, theta waves will be the most active. When the subject starts to relax or enters the resting phase, alpha waves will increase. In contrast, beta waves dominate when the subject is performing daily activities, while delta waves appear when the subject is asleep or meditating. By observing the changes in these dominant wave types, researchers can understand the subject's mental state and activity level.

To obtain classification results, 50% of training data, 25% of testing data, and 25% of validation data were used to measure the accuracy of the data. The interval value is used in determining the summation result of each classification class. The data that has been classified is then tested again using the KNN method to determine the level of accuracy.

Average Electroencephalogram (EEG) data from various subjects, measured with different frequencies (Delta, Theta, Alpha, Beta 1, and Beta 2). Each subject was labelled with a specific class, which contributed to the level of impulsivity associated with neurological conditions, cognitive conditions such as S2, S6, and S9 that displayed higher values in some waves, indicating potential differences in neural patterns. The different categories that occur with each subject are labelled in classes in the table. On analysing the Electroencephalogram (EEG) data alongside these class labels can relate to specific conditions, potentially revealing underlying patterns. The results showed that subjects S2, S6, and S9 had very high Electroencephalogram (EEG) wave values. This proves that each subject has a unique or different neural pattern compared to other subjects, which can be seen in table 1.

TABLE I. FEATURE COMPARISON RESULTS OF EACH SUBJECT

Subject	Delta	Theta	Alpha	Beta1	Beta2	Total
S1	2.715	2.699	4.832	1.364	1.311	12.922
S2	3.788	3.156	5.481	1.728	1.423	15.788
S3	2.549	1.898	3.011	1.084	1.064	9.606
S4	2.152	1.653	3.577	1.726	1.2.69	10.378
S5	2.466	1.954	4.753	1.671	1.748	12.576
S6	5.757	3.871	5.034	2.038	2.087	18.788
S7	2.101	1.808	3.906	1.500	1.534	10.85
S8	1.607	1.793	2.944	1.321	1.763	9.403
S9	5.022	2.967	6.446	1.978	2.532	18.947
S10	3.05	2.197	2.754	1.392	1.356	10.751
S11	3.222	2.083	2.892	1.064	1.346	10.61
S12	4.42	2.601	4.84	1.568	1.753	15.185
S13	4.432	2.017	2.029	1.195	1.106	10.782
S14	3.257	2.369	2.404	1.69	1.395	11.118
S15	2.843	2.064	1.961	0.928	1.051	8.849
S16	3.246	2.266	3.78	1.834	1.613	12.745
S17	4.169	2.923	3.206	1.221	1.172	12.69
S18	2.5	1.754	3.094	1.263	1.304	9.918
S19	2.503	1.707	2.734	1.16	1.251	9.356
S20	3.52	2.29	3.014	1.707	1.743	12.276
S21	3.073	2.329	3.495	1.728	1.941	12.569

C. Data Visualisation

The results of Electroencephalogram (EEG) data visualisation using the KNN method that can help us understand the patterns and graphs found by KNN in EEG data can be seen in the figure 8.



D. Confusion Matrix Graph Model

- Confusion Matrix is a table that shows the performance of the classification model in predicting data labels.
- In the first column (True Class), 94.9% of the data belongs to the Potentially Impulsive class, 5.1% of the data belongs to the Highly Potentially Impulsive class, and 100% of the data belongs to the Not Impulsive class.
- The Predicted Class shows the prediction results of Potentially Impulsive data, 94.9% are categorised correctly, and 5.8% are categorised as potentially impulsive. Of the potentially impulsive data, 95.0% were categorised correctly, and 5.0% were categorised as potentially impulsive.
- True Positive Rate (TPR) is the proportion of positive data that is correctly categorised. For the potentially impulsive class, the TPR was 94.2% and for the highly potentially impulsive class, the TPR was 95.0%.
- False Negative Rate (FNR) is the proportion of positive data categorised as negative. For the potentially impulsive class, the FNR is 5.8%, and for the very potentially impulsive class, the FNR is 5.0%. Based on Figure a and Figure b, both models show good performance in classifying the data. We can see the confusion matrix model in Figure 9.



Fig.9. Evaluation Matrix.

E. Scatter Plot

Matlab scatter plot is a graph used to visualise the relationship between two variables. In figure b the X-axis is the value predicted by the model for each data and the Y-axis is the true value of each data. Green colour indicates potentially impulsive results, the more green dots the better or more accurate the model performance. Red colour is potentially impulsive and yellow colour is not impulsive. We can see the confusion matrix model in Figure 10.



Fig. 10. Scatter Plot.

- F. ROC (Receiver Operating Characteristic).
 - Receiver Operating Characteristic (ROC) is a visual representation to assess the performance of a classification model.
 - ROC displays the relationship between TPR and FPR across different model classifications, TPR indicates the proportion of positive data that is correctly predicted, while FPR indicates the proportion of negative data that is incorrectly predicted as positive.
 - The higher the ROC curve, the better the model's ability to distinguish between positive and negative.
 - In figure c the area under the ROC curve (AUC) summarises the performance of the model, with AUC = 0.99 indicating that the classification model is better, the higher the AUC value, the better the performance of the classification model. We can see the ROC Model in Figure 11.



Fig. 11. Receiver Operating Characteristic (ROC).

IV. CONCLUSSION

This study classifies subjects into four levels namely, not impulsive, impulsive, very potentially impulsive, and potentially impulsive. This classification model showed high accuracy (Training Data: 94.7%, Testing: 91.3%, and Validation Data: 91.8%). To improve accuracy, this study must be expanded to include a bigger and more diverse sample and to incorporate additional data. The findings are potentially useful for improving impulsive behaviour intervention, educational program development, and further impulsivity research.

ACKNOWLEDGMENT

The author would like to thank the Faculty of Science and Technology of Prima Indonesia University, and also to all friends (seniors and juniors) who have participated. In most cases, recognition of sponsorship and financial support.

REFERENCES

- Lee, J. Y., Park, S. M., Kim, Y. J., Kim, D. J., Choi, S. W., Kwon, J. S., & Choi, J. S. (2017). Resting-state EEG activity related to impulsivity in gambling disorder. Journal of Behavioral Addictions, 6(3), 387–395. https://doi.org/10.1556/2006.6.2017.055
- [2] MEYE, F. J., & DIGIROLAMO, G. J. (2007). Handbook of Psychophysiology (3rd edn). By J. T. Cacioppo, L. G. Tassinary and G. G. Berntson. (Pp. 908; \$175.00; ISBN 0521844711 hb.) Cambridge University Press. 2007. Psychological Medicine, 37(12), 1818–1819. https://doi.org/10.1017/s0033291707001201
- [3] Bongers, P., van de Giessen, E., Roefs, A., Nederkoorn, C., Booij, J., van den Brink, W., & Jansen, A. (2015). Being impulsive and obese increases susceptibility to speeded detection of high-calorie foods. Health Psychology, 34(6), 677–685. https://doi.org/10.1037/hea0000167
- [4] Candra, D., & Novitasari, R. (n.d.). Klasifikasi Sinyal EEG Menggunakan Metode Fuzzy C-Means Clustering (FCM) Dan Adaptive Neighborhood Modified Backpropagation (ANMBP). http://sccn.ucsd.edu/~arno/fam2data/publicly_
- [5] Fitriawati Zakiyyah, A. (2021). Klasifikasi Emosi Untuk Mengetahui Pengalaman Emosional Melalui Sinyal EEG Menggunakan Algoritma

Artificial Neural Network (Vol. 3, Issue 2).

- [6] Lestari, S. I. P. (2018). Pengaruh Price Discount dan Bonus Pack terhadap Impulse Buying melalui Nilai Hedonik di Carrefour Surakarta. Jurnal Maksipreneur: Manajemen, Koperasi, Dan Entrepreneurship, 7(2), 129. https://doi.org/10.30588/jmp.v7i2.362
- [7] Turnip, A., Soetraprawata, D., & Kusumandari, D. E. (2013). A Comparison of EEG Processing Methods to Improve the Performance of BCI. International Journal of Signal Processing Systems, 63–67. https://doi.org/10.12720/ijsps.1.1.63-67
- [8] Hisyamudin, N., Nor, M., Emran, A., Khalid, I. A., Rasidi, M., Mohd, I., Mohideen Batcha, F., Mubarak, A., Arifin, T., Zaman, I., Sofian, B., Mohd, M., Sani, A., Waluyo, M. J., Siswanto, A., Ahmad, A., Rahmahwati, A., Akmal, A., Mohammed, N., ... Rahim, A. (2017). INTERNATIONAL CONFERENCE ON MECHANICAL AND MANUFACTURING ENGINEERING 2017 SCIENTIFIC EDITORS.
- [9] Mira Yulia, Anita, dan Cantika Miranda "Klasifikasi Sinyal EEG dengan Stimuli Aromatik Menggunakan Metode Support Vector Machine", Universitas Prima Indonesia, 2019.
- [10] Turnip, A., & Hong, K.-S. (2012). CLASSIFYING MENTAL ACTIVITIES FROM EEG-P300 SIGNALS USING ADAPTIVE NEURAL NETWORKS. In International Journal of Innovative Computing, Information and Control ICIC International c (Vol. 8, Issue 9).
- [11] Aris Catur Bintoro "Pemeriksaan EEG untuk Diagnosis dan Monitoring pada Kelainan Neurologi", Clinical Practice", 2012.
- [12] Aris Catur Bintoro "Pemeriksaan EEG untuk Diagnosis dan Monitoring pada Kelainan Neurologi", Clinical Practice", 2012.
- [13] Shimodaira H. Meningkatkan inferensi prediktif dalam pergeseran kovariat dengan memberi bobot pada fungsi log-likelihood. *Inferensi Rencana Stat J.* 2000, 90(2): 227–244.
- [14] Nafi, N., & Darma Setiawan, B. (2017). Penerapan Metode K-Nearest Neighbor (KNN) dan Metode Weighted Product (WP) Dalam Penerimaan Calon Guru Dan Karyawan Tata Usaha Baru Berwawasan Teknologi (Studi Kasus : Sekolah Menengah Kejuruan Muhammadiyah 2 Kediri) (Vol. 1, Issue 5). http://j-ptiik.ub.ac.id
- [15] Nafi, N., & Darma Setiawan, B. (2017). Penerapan Metode K-Nearest Neighbor (KNN) dan Metode Weighted Product (WP) Dalam Penerimaan Calon Guru Dan Karyawan Tata Usaha Baru Berwawasan Teknologi (Studi Kasus : Sekolah Menengah Kejuruan Muhammadiyah 2 Kediri) (Vol. 1, Issue 5). http://j-ptiik.ub.ac.id
- [16] Nafi, N., & Darma Setiawan, B. (2017). Penerapan Metode K-Nearest Neighbor (KNN) dan Metode Weighted Product (WP) Dalam Penerimaan Calon Guru Dan Karyawan Tata Usaha Baru Berwawasan Teknologi (Studi Kasus : Sekolah Menengah Kejuruan Muhammadiyah 2 Kediri) (Vol. 1, Issue 5). http://j-ptiik.ub.ac.id
- [17] Suci Aulia, Sugondo Hadiyoso, Dadar Nur Ramadan "Analisis Perbandingan KNN dengan SVM untuk Klasifikasi Penyakit Diabetes Retinopati berdasarkan Citra Eksudat dan Mikroaneurisma", 2015.
- [18] Mardi Turnip, Prada W. Siregar "Decoding Neural Signatures: Develop Drugs Addiction Level through EEG Patterns based Adaptive Neuro-Fuzzy Inference", 2024.
- [19] Diantoro, T., Alif Fiolana, F., & Arie Widhining Kusumastutie, D. (n.d.). ALINIER JURNAL VOL 4 NO 2 NOVEMBER 2023 Klasifikasi Sinyal Delta, Theta, Alpha, Beta, Gamma Pada Electroencephalography (EEG). www.elektro.itn.ac.id
- [20] Puspitawuri, A., Santoso, E., & Dewi, C. (2019). Diagnosis Tingkat Risiko Penyakit Stroke Menggunakan Metode K-Nearest Neighbor dan Naïve Bayes (Vol. 3, Issue 4). http://j-ptiik.ub.ac.id