Jurnal Algoritma, Logika dan Komputasi *Vol.VIII (No.01) :762 - 771 Th.2025*
Versi Online: https://journal.ubm.ac.id/index.php/alu *p-ISSN: 2620-620X*
DOI: http://dx.doi.org/10.30813/j-alu.v2i2.8652 *e-ISSN: 2621-9840*

Classification Of Student Mental Health Based On Academic And Social Variables Using The Decision Tree Method

Desi Anggreani, desianggreani@unismuh.ac.id [1)], Chyquitha Danuputri, Chyquithadanuputri@unismuh.ac.id [2)]*, Muhyiddin A M Hayat, muhyiddin@unismuh.ac.id [3)], Dedi Setiawan, dedisetiawan@unimen.ac.id [4)]

[1)2)3)])Department of Informatics Engineering, Universitas Muhammadiyah Makassar, Indonesia
[4)]Faculty of Teacher Training and Education, Universitas Muhammadiyah Enrekang, Indonesia

Diterima 20 Maret 2025 / Disetujui 20 Juni 2025

ABSTRACT

Mental health problems are suffered by many people, including students who often have poor lifestyles. Depression and anxiety are widespread among students, with all universities reporting students with depression and 75.5% reporting students with severe anxiety. This research aims to determine the classification of student mental health based on academic and social by using the Decision Tree method so that early treatment can be carried out. The dataset used consists of 11 aspects concerning academic and social. The data that has been collected is processed through the preprocessing stage and analyzed using the Decision Tree classification method. The classification results showed that out of 973 students who did not suffer from depression, the method classified them correctly. In addition, of the 104 college students who were classified as suffering from major depression, all of them were actually suffering from major depression. The agreement between the classification results and the actual condition shows the reliability of this method, with an accuracy rate of 76.71%. This research underscores the importance of academic and social variables in influencing students' mental health. The findings confirm the reliability of the Decision Tree method in detecting students' mental state and point to the need for effective counseling services and mental health interventions in campus and social environments.
Keywords: *Mental Health, Academic, Social, Classification, Decision Tree*

ABSTRAK

Masalah kesehatan mental telah dialamai oleh kebanyakan orang, termasuk mahasiswa yang sering memiliki gaya hidup yang kurang baik. Depresi dan kecemasan tersebar luas di kalangan mahasiswa, dengan semua universitas melaporkan mahasiswa yang mengalami depresi dan 75,5% melaporkan mahasiswa dengan kecemasan berat. Penelitian ini bertujuan untuk menentukan klasifikasi kesehatan mental mahasiswa berdasarkan aspek akademik dan sosial dengan menggunakan metode Decision Tree sehingga dapat dilakukan penanganan dini. Dataset yang digunakan terdiri dari 11 aspek yang berkaitan dengan akademik dan sosial. Data yang telah dikumpulkan diproses melalui tahap prapemrosesan dan dianalisis menggunakan metode klasifikasi Decision Tree. Hasil klasifikasi menunjukkan bahwa dari 973 mahasiswa yang tidak menderita depresi, metode ini mengklasifikasikan mereka dengan benar. Selain itu, dari 104 mahasiswa yang diklasifikasikan menderita depresi berat, semuanya benar-benar menderita depresi berat. Kesepakatan antara hasil klasifikasi dan kondisi aktual menunjukkan keandalan metode ini, dengan tingkat akurasi sebesar 76,71%. Penelitian ini menekankan pentingnya variabel akademik dan sosial dalam mempengaruhi kesehatan mental mahasiswa. Temuan ini mengkonfirmasi keandalan metode Decision Tree dalam mendeteksi kondisi mental mahasiswa dan menunjukkan perlunya layanan konseling yang efektif serta intervensi kesehatan mental di lingkungan kampus dan sosial.
Kata Kunci: Kesehatan Mental, Akademik, Sosial, Klasifikasi, Decision Tree

*Korespondensi Penulis:
E-mail: **Chyquithadanuputri@unismuh.ac.id**

INTRODUCTION

Mental health is one of the most serious issues that needs to be considered, accounting for 15% of the diseases suffered by the entire world population[1]. The World Health Organization (WHO) reported that in 2019 there were 950 million people living with mental disorders, 52.4% of women and 47.6% of men[2]. In developing countries such as Indonesia, the 2018 Indonesian Basic Health Research, as many as 9.8 percent or around 26 million Indonesians over the age of 15 experience emotional and mental disorders[3]. Without a healthy mental state, the person will face difficulties in living their daily lives. With good mental health, a person feels an improved mood, builds resilience, and helps to enjoy life as a whole[4][5].

There are many poor mental conditions suffered by various groups, including students who have a bad lifestyle[6][7]. Students lack knowledge and awareness of the importance of mental health, so without realizing the various problems in their minds are the impact of the mental health they feel. Depression and anxiety are widespread among students with all universities reporting students with depression and 75.5% reporting students with severe anxiety[8][9]. A study in the UK found that levels of psychological distress increased upon entering university and recent evidence suggests that the prevalence of mental health problems among students, including self-harm and suicide, is increasing[10]. Many mental health in college students can be influenced by various factors including friendship environment, genetics, family, social environment, lifestyle, and others[11][12].

The classification process of diseases has been implemented by many researchers using different methods[13][14]. In 2023 research was conducted for the classification of heart disease using the K-Nearest Neighbor method. Based on this research, it is obtained that heart disease patients are more than healthy patients based on the symptoms felt[15]. In the latest research, classification of Electroencephalography (EEG) examinations was carried out to identify diseases such as seizures, Alzheimer's, and schizophrenia. In the study implemented the machine learning method and obtained 71.4% accuracy. This research will classify the mental health of students by focusing on academic and social students so that this research is expected to provide knowledge of the influence of academic and social variables on student mental health. The student mental health classification process is carried out using the Decision Tree method[16]. This research is expected to be the basis in students maintaining health, detecting early symptoms of mental health disease. By maintaining the mental health of students well, it will increase the resilience of the body and the feeling of calmness in undergoing lectures within the university.

RESEARCH METHODS

This research uses student mental health data obtained from the official Kaggle website. The data includes various important information related to the mental health conditions and depression levels of students as well as factors that affect their psychological well-being based on academic and social variables. The data that has been collected amounts to 7023 data which has 11 variables, namely Department, Age, Gender, GPA, Sleep Quality, Physical Activity, Social Environment Support, Marital Status, Extracurricular Activities, Total Semester Credit Load and Student Residence. The target variable in the research data is Student Depression Status which consists of

two classifications, namely whether the student suffers from severe depression and does not suffer from depression. Some details of data regarding student mental health can be seen in Table 1.

Table 1. Student Mental Health Dataset

Age	Course	Gender	IPK	Sleep Quality	Physical Activity	Social Support
25	Others	Male	3.56	Good	Moderate	Moderate
24	Engineering	Female	2.44	Average	Low	Low
19	Business	Female	3.74	Good	Low	Moderate
19	Computer Science	Male	3.65	Average	Low	Moderate
18	Business	Male	3.4	Good	Low	High
21	Medical	Female	3.35	Good	Moderate	High
18	Law	Male	3.65	Good	Moderate	Moderate
21	Business	Female	3.4	Average	Low	Low
24	Medical	Male	3.8	Poor	Low	Moderate
19	Engineering	Female	3.05	Average	Moderate	Low
23	Law	Female	3.74	Average	Low	Moderate
28	Engineering	Female	4	Average	Moderate	Moderate
22	Computer Science	Male	3.19	Average	Moderate	Moderate
27	Medical	Male	3.26	Average	Moderate	High
24	Medical	Female	3.2	Average	Low	Moderate
25	Law	Male	3.61	Good	Low	Moderate
18	Medical	Female	3.85	Good	Low	Moderate
19	Medical	Male	3.26	Good	Low	High
22	Computer Science	Male	3.46	Good	Moderate	Moderate

Relationship	Extracurricular Involvement	Semester Credit Load	Residence Type	Depression
Single	Moderate	17	On-Campus	Tidak Depresi
Married	Low	27	On-Campus	Tidak Depresi
Single	High	15	On-Campus	Tidak Depresi
In a Relationship	Moderate	20	Off-Campus	Tidak Depresi
Single	High	23	On-Campus	Tidak Depresi
Married	Moderate	19	Off-Campus	Depresi Berat
Single	Moderate	20	Off-Campus	Tidak Depresi
Single	Moderate	23	On-Campus	Tidak Depresi
Married	Low	28	On-Campus	Tidak Depresi
Single	Moderate	27	Off-Campus	Depresi Berat
In a Relationship	Moderate	22	Off-Campus	Tidak Depresi
In a Relationship	Moderate	20	Off-Campus	Tidak Depresi
In a Relationship	High	17	On-Campus	Tidak Depresi
In a Relationship	Moderate	17	With Family	Tidak Depresi
In a Relationship	Moderate	18	On-Campus	Tidak Depresi
Single	Moderate	25	Off-Campus	Tidak Depresi
In a Relationship	Moderate	23	Off-Campus	Tidak Depresi
Married	Moderate	29	With Family	Tidak Depresi
Single	High	29	Off-Campus	Tidak Depresi

The data that has been collected will be processed at the preprocessing stage, which is data on classification involving several important steps to ensure the data used in the model has quality and is ready to be analyzed. The following are general steps in the data preprocessing process, namely data cleaning, data transformation, feature extraction and selection, and division of datasets into training data and testing data. After the data has gone through the preprocessing process, the data is ready to enter the classification method stage, namely the Decision Tree.

The method applied to create a Decision Tree is the CART (Classification and Regression Tree) technique. CART is a classification method that uses historical data to build a decision tree. Each terminal node in this tree is given a class label or predicted value, thus forming a tree structure called a decision tree where the tree

Vol.VIII (No.01) :762 - 771 Th.2025
p-ISSN: 2620-620X
e-ISSN: 2621-9840

structure is made based on weight entropy. The greater the depth of the tree, the higher the accuracy of the results.The modeling structure of the decision tree method can be seen in Figure 1.

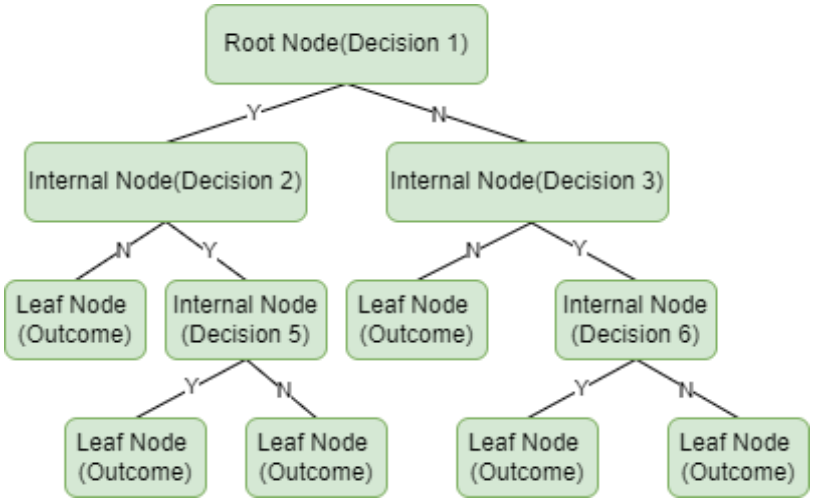

Figure 1. ***Modeling the Decision Tree Method***

In the decision tree method, the root node will be obtained from the selected attribute based on the calculation of the gain value and entropy value. The highest gain value is the variable that will become the root note. The sequence in performing the calculation is to calculate the entropy value first, after obtaining the entropy value it will be used to calculate the Gain. The equations used in the calculation of the decision tree method are found in equations (1) and (2).

$$Entropy(S) = \sum_{i=1}^{n} -pi.\log_2 pi \quad (1)$$

Description:

S = set of cases

n = number of partitions of S

pi = proportion of Si to S

$$Gain(S,A) = \sum_{i=1}^{n} \frac{|Si|}{|S|} * Entropy(Si) \quad (2)$$

Description:

S = set of cases A = features
n = number of partitions of attribute A

|Si| = proportion of Si to S

|S| = count of cases in S

Classification performance measurement is done by comparing the number of all correctly classified test data to the total number of test data. Equation (3) is a confusion matrix that can be used as a test model to measure classification performance. The use of this equation allows an objective evaluation of how well the Decision Tree model is able to classify student mental health conditions based on existing academic and social variables. True Positive (TP) is the amount of data that is Positive and correctly classified as Positive. False Positive (FP): The

amount of data that is Negative but classified as Positive. False Negative (FN): The amount of data that is Positive but classified as Negative. True Negative (TN): The amount of data that is Negative and classified in the system correctly as Negative[17][18]. The equation can be found in equations (3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (3)$$

RESULT

This research aims to identify and classify the mental health conditions of university students by using academic and social variables as the main factors in the classification process. The student mental health classification process has been implemented using the Decision Tree method. In this research, the data used is divided into two sets, namely training data and testing data. A total of 80% of the overall data is allocated as training data to train the model, while the remaining 20% is used as testing data to evaluate the performance of the model. This data division aims to ensure that the model built has good classification capabilities when applied to data that has never been processed before, so that the results of the classification of student mental health become more accurate

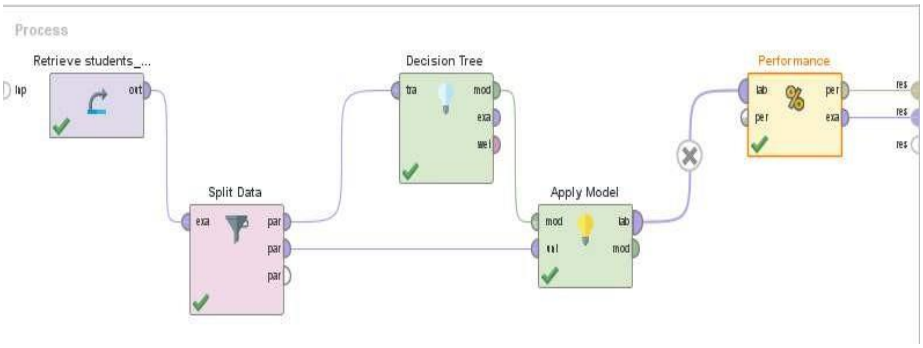

Figure 2. ***Classification Process with Decision Tree Method***

The results of entropy and gain calculations have been carried out thoroughly to obtain Confidence values for both categories, namely not depressed and severely depressed. Through this process, each attribute is analyzed to determine how much it contributes in distinguishing between the two categories. The entropy values help measure the uncertainty or irregularity in the data, while the gain measures the decrease in uncertainty after dividing the data by an attribute. Confidence values for each category are calculated, which represent how confident the model is in performing the classification for each data sample. The classification result is then determined based on the highest Confidence value, where the data sample will be classified into the category that has the highest Confidence value.

From the testing data, which includes 20% of the total dataset, the classification results show that there are 1,059 students who do not suffer from depression. This data illustrates that most of the student population in the sample are in a healthy mental state and do not experience depressive disorders. On the other hand, the classification results also showed that there were 345 students who suffered from major depression. This highlights a number of students who need more attention and support in dealing

with their mental health conditions. This finding underscores the importance of effective counseling services and mental health interventions on campus to help students with major depression.

Tabel 2. Classification of Student Mental Health

No	Depression	Classification Results of Decision Tree Method (Depression)	Confidence (Tidak Depresi)	Confidence (Severe Depression)
1	Tidak Depresi	Tidak Depresi	0.7962128043282236	0.2037871956717764
2	Depresi Berat	Tidak Depresi	0.7962128043282236	0.2037871956717764
3	Tidak Depresi	Depresi Berat	0.4848116646415553	0.5151883353584447
4	Tidak Depresi	Tidak Depresi	0.7962616822429907	0.20373831775700935
5	Depresi Berat	Tidak Depresi	0.8077611940298507	0.19223880597014925
6	Depresi Berat	Tidak Depresi	0.7962128043282236	0.2037871956717764
7	Tidak Depresi	Depresi Berat	0.4848116646415553	0.5151883353584447
8	Depresi Berat	Tidak Depresi	0.7962128043282236	0.2037871956717764
9	Tidak Depresi	Tidak Depresi	0.7962128043282236	0.2037871956717764
10	Depresi Berat	Depresi Berat	0.4848116646415553	0.5151883353584447

The results of the decision tree method will be subjected to a series of tests to evaluate the performance of this method in performing classification. The testing process is carried out by comparing the classification result data generated by the decision tree with the actual data on student conditions. This test aims to measure how accurate the decision tree method is in identifying student conditions. The classification result data is compared directly with the actual data to see the level of conformity between prediction and reality. Through this analysis, insight can be gained into the effectiveness of the decision tree method in performing classification.

Accuracy = (TP+TN)/(TP+TN+FP+FN)*100%
Accuracy = (973+104)/(973+104+241+86)*100%
Accuracy = 1077/1404*100% Accuracy =
76,71%

DISCUSSION

The Impact of Academic Variables on Students' Mental Health

Academic variables used in the process of classifying student mental health in this study include several important factors, namely the grade point average (GPA), the number of credits taken, participation in extracurriculars, and the major chosen by the student. These variables were chosen because they are considered to provide a deeper insight into the relationship between academic achievement and psychological well-being. Based on the results of the classification process, it was found that students who experienced severe depression tended to have GPA scores in the range of 3.1 to 3.7 and students who did not experience depression in the range of 3.2 to 4. This finding indicates that students with relatively high GPA achievement but may be under

pressure to maintain or improve their grades, are at higher risk of developing major depression. Figure 2 shows the results of the classification process based on GPA scores.

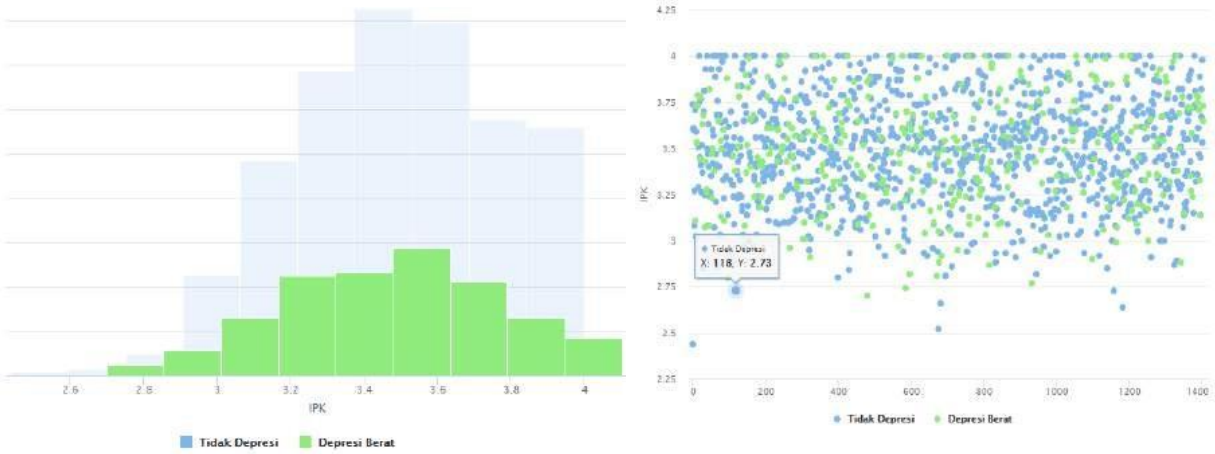

Figure 4. ***Classification of Student Mental Health Based on GPA Score***

The Impact of Social Variables on Student Mental Health

The social aspect of college students' mental health has a very significant influence. In this research classification, various social variables were taken into account to understand how these factors contribute to students' psychological well-being. These variables include external social support, such as support received from family, friends, and community, which can provide a sense of security and help reduce stress and anxiety. Students' relationship status is also an important variable to consider, as whether students are In a Relationship or Single can affect their emotional stability and psychological well-being. Students' sleep quality is a major focus in this study, given that adequate and high-quality sleep is essential for maintaining mental and physical health.

The last social variable is where students live, whether they live with family, in dormitories, or in boarding houses, is also included as a relevant social variable. This factor can affect the level of comfort and stress that students experience in their daily lives. Social support variables for university students in this study were divided into three main categories based on the level of support they received, namely low, moderate, and high support. Based on the classification results, severe depression is felt in students who get moderate support. Of the 345 students who became test data, 154 students with moderate social support suffered from severe depression.

Vol.VIII (No.01) :762 - 771 Th.2025
p-ISSN: 2620-620X
e-ISSN: 2621-9840

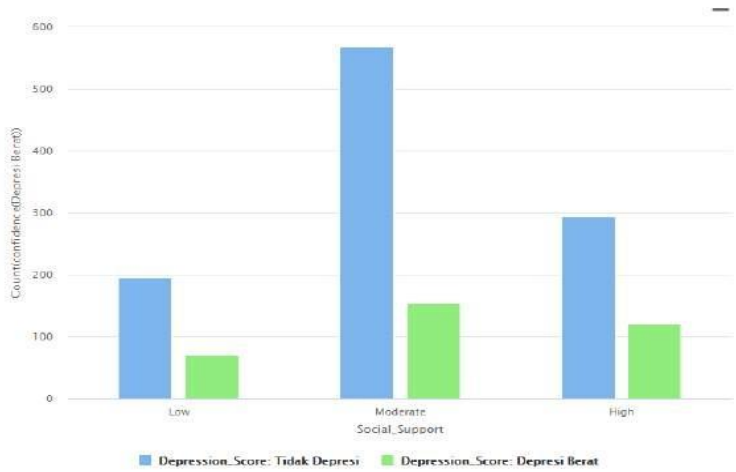

Figure 5. ***Classification of Student Mental Health Based on Social Support***

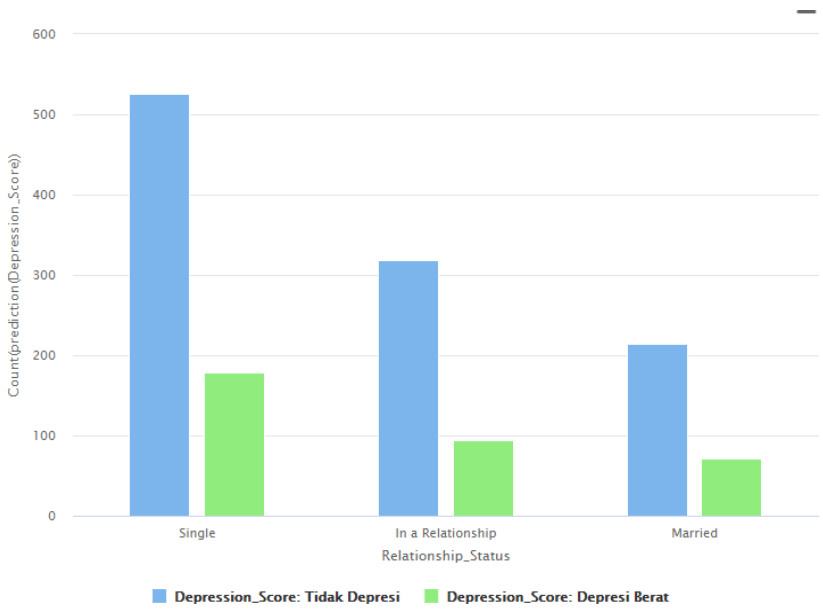

Figure 6. ***Classification of Students' Mental Health by Relationship Status***

Based on the 345 data of students who suffered from major depression, it was found that the relationship status "single" had the highest number in terms of prevalence of depression. A total of 178 single students experienced major depression, indicating that this group is the most vulnerable compared to other groups of students. This indicates a significant correlation between relationship status and depression levels, which requires special attention in mental health intervention and support efforts for university students.

CONCLUSION AND SUGGESTIONS

This research aims to classify students' mental health conditions based on academic and social variables using the Decision Tree method. The results show that the Decision Tree method is able to classify the mental state of college students with a high level of accuracy. From the overall testing data, which covers 20% of the total dataset, it was found that this method identified 1,059 students who did not suffer from

depression and 345 students who suffered from severe depression. From the aspect of GPA (Grade Point Average) value, the analysis shows that students with higher GPA value tend to have a lower risk of suffering from depression. Conversely, students with lower GPA scores are more prone to depression, both mild and severe. This shows that academic achievement has a significant correlation with students' mental health condition. Social variables also play an important role in students' mental health conditions. Students who receive high social support from family, friends, and the surrounding environment tend to have better mental health and a lower risk of depression. Conversely, moderate levels to lack of social support can increase the risk of depression in college students.

Based on the results of the study on the classification of students' mental health using the Decision Tree method, there are several suggestions that can be considered for further research, namely the use of Other Machine Learning Methods to be able to be a comparison of the results of the current study. In addition, collaborating with Psychology Experts in interpreting the results and preparing more appropriate recommendations. A multidisciplinary approach will enrich the analysis and implementation of research results

BIBLIOGRAPHY

[1] K. M. Aldiabat, N. A. Matani, and C. Le Navenec, "Mental Health among Undergraduate University Students : A Background Paper for Administrators , Educators and Healthcare Providers," no. December 2020, 2014, doi: 10.13189/ujph.2014.020801.

[2] W. H. Organitation, "Mental disorders," *World Health Organitation*. https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

[3] M. Health, F. Student, I. Facing, and A. New, "al-Tazkiah :," vol. 12, no. June, pp. 15–30, 2023.

[4] N. K. Iyortsuun, S. Kim, M. Jhon, and H. Yang, "A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis," pp. 1–27, 2023.

[5] M. Defrianto, "Mental Health of College Students and Its Relation to Life Satisfaction and Social Media Abuse," vol. 15, no. 2, pp. 193–218, 2020, doi: 10.21580/sa.v15i2.6950.

[6] R. Readi, "Bibliometric Analysis : Student Mental Health," vol. 1, no. 3, pp. 253–261, 2023.

[7] P. Pedrelli, "HHS Public Access," vol. 39, no. 5, pp. 503–511, 2015, doi: 10.1007/s40596-014-0205-9.College.

[8] G. U. Mukminin *et al.*, "Inovasi Kurikulum," vol. 21, no. 1, pp. 443–454, 2024.

[9] Y. R. Shim, R. Eaker, and J. Park, "Mental Health Education , Awareness and Stigma Regarding Mental Illness Among College Students," vol. 6, pp. 6–15, 2022.

[10] C. D. Dr. Stephen Pereira, Katie Reay, Jo Bottell, Lucy Walker, "Mental Health Survey 2018," no. March, 2019.

[11] F. Campbell, L. Blank, A. Cantrell, S. Baxter, C. Blackmore, and J. Dixon, "Factors that influence mental health of university and college students in the UK : a systematic review," *BMC Public Health*, pp. 1–22, 2022, doi: 10.1186/s12889-022-13943-x.

[12] M. K. Sari and E. A. Susmiatin, "Deteksi Dini Kesehatan Mental Emosional pada

Mahasiswa," vol. XIII, no. 1, pp. 10–17, 2023.

[13] R. Habenicht *et al.*, "Machine Learning Based Linking of Patient Reported Outcome Measures to WHO International Classification of Functioning , Disability , and Health Activity / Participation Categories," 2023.

[14] J. Behavioral *et al.*, "Mental health professionals ' use of the ICD-11 classi fi cation of impulse control disorders and behavioral addictions : An international fi eld study," 2024, doi: 10.1556/2006.2023.00083.

[15] E. Sulistio, Y. Maulita, and M. Simanjuntak, "Classification For Predicting Heart Disease Using The K Nearest Neighbor Method Sylvani General Hospital Binjai City," vol. 3, no. 1, 2023.

[16] S. Bhat and S. R. Reeja, "EAI Endorsed Transactions A Comprehensive Study on Mental Illness Through Speech and EEG Using Artificial Intelligence," vol. 10, 2024, doi: 10.4108/eetpht.10.5328.

[17] L. Desi Anggreani, "IMPLEMENTASI METODE CERTAINTY FACTOR (CF) PADA APLIKASI SEHAT ORGANIK DALAM MENDIAGNOSA," vol. 6, no. 1, pp. 61–70.

[18] T. K. Reska NNA, "Comparison of KNN, Naive Bayes, and Decision Tree Methods in Predicting the accuracy of classification of Immunotherapy Dataset."