# ANALYSIS OF DEPRESSION IN COLLEGE STUDENT DURING COVID-19 PANDEMIC USING EXTREAM GRADIENT BOOST

**Agung Prabowo[1]; Dharma Ajie Nur Rois[2]; Amar Luthfi[3]; Ultach Enri[4]**

Teknik Informatika [1, 2, 3, 4]
Universitas Singaperbangsa Karawang
www.unsika.ac.id
agung.prabowo18082@student.unsika.ac.id [1]; dharma.ajie18120@student.unsika.ac.id [2];
amar.luthfi18154@student.unsika.ac.id [3]; ultach@staff.unsika.ac.id [4]

***Abstract***—*The Covid-19 pandemic that spreads in Indonesia causes health, economic, and social problems in the community, including mental health. Of course, this mental health problem also hit students. Seeing these conditions, we conducted research on students of the Faculty of Computer Science, University of Singaperbangsa Karawang using the Patient Health Questionnaire-9 which measures a person's level of depression. In this study, we used Extreme Gradient Boost or XGBoost to classify students' depression tendencies. We break down the dataset into training data and testing data with 4 data sharing combinations, they are 80 : 20, 50 : 50, 90 : 10, 70 : 30. The combination of 90 : 10 data sharing has the best performance with accuracy, precision, recall, and F1-scores respectively 92.86%, 94.29%, 92.86% , and 92.06%. This method also has better performance than K-Nearest Neighbor, Random Forest, Multi Layer Perception, Support Vector Machine and Decision Tree .*

***Keywords***—*Depression, PHQ-9, XGBoost*

**Intisari**—Pandemi Covid-19 yang terjadi di Indonesia menyebabkan permasalahan kesehatan, ekonomi, dan sosial yang ada di masyarakat, termasuk kesehatan jiwa. Tentu saja permasalahan kesehatan jiwa ini juga melanda mahasiswa. Melihat kondisi tersebut kami melakukan penelitian kepada mahasiswa Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang menggunakan kuesioner *Patient Health Questionnaire*-9 yang mengukur tingkat depresi seseorang. Pada penelitian ini menggunakan *Extream Gradient Boost* atau XGBoost untuk melakukan klasifikasi kecenderungan depresi pada mahasiswa. Penulis memecah dataset menjadi data latih dan data test dengan 4 kombinasi pembagian data, yaitu 80 : 20, 50 : 50, 90 : 10, 70 : 30. Kombinasi pembagian data 90 : 10 memiliki performa yang paling baik dengan *accuracy*, *precission*, *recall*, dan *F1-score* masing-masing 92.86%, 94.29%, 92.86%, dan 92.06%. Metode ini juga memiliki performa lebih baik dibanding dengan *K-Nearest Neighbor, Random Forest*, *Multi Layer Perception*, *Support Vector Machine* dan *Decision Tree.*

***Kata Kunci****:* Depresi*,* PHQ-9*,* XGBoost*.*

## INTRODUCTION

The spread of Covid-19 in Indonesia as of August 25, 2021, shows that 257,677 positive cases have been confirmed, 3,639,867 cases have recovered and 129,293 people have died (Satuan Tugas Penanganan Covid-19, 2021). The growth of Covid-19 cases, which creates anxiety about when this pandemic will end, has a negative impact on the community. The impact of the Covid-19 pandemic includes health, economic, and social problems in the community, including mental health (Lempang et al., 2021).

Based on survey data regarding Psychological Problems in the Covid-19 Pandemic Era to self-check users from Perhimpunan Dokter Spesialis Kedokteran Jiwa Indonesia (Perhimpunan Dokter Spesialis Kedokteran Jiwa Indonesia, 2020), 68% experienced anxiety problems, 67% experienced depression and 77% experienced psychological trauma. In one study it was explained that people aged 18-29 were more susceptible to anxiety due to COVID-19 than other age groups. (Indra et al., 2020). In another study, 21.1% of 147 students had mild depression, 17% moderate depression and 3.4% had severe depression (Hasanah et al., 2020).

Seeing these conditions, we conducted research on students of the Faculty of Computer Science, University of Singaperbangsa Karawang to find out if they had problems related to mental health by providing a Patient Health Questionnaire-9 survey that measured a person's level of depression. Depression is a feeling disorder that is common to everyone. According to WHO data in 2018 the number of people with depression ranged from 300 million people (Kusuma et al., 2021).

The Patient Health Questionnaire-9 is a nine-point questionnaire that helps in early diagnosis of depression and helps in selecting and monitoring treatment. The advantages of using PHQ-9 are that it can be done quickly, has good psychometric properties, can be given repeatedly, the results will show improvement or worsening of depression during the treatment period (P. D. Kusuma et al., 2018). In addition, in one study, the results of the PHQ-9 survey on 235 ITEKES Bali Nursing Students showed 45.5% had mild depression, 1.7% had severe depression and 29.4% did not experience depression (Kusuma et al., 2021).

In a study that analyzed several machine learning algorithms including Random Forest, Naive Bayes, Support Vector Machine, K-Nearest Neighbors to detect mental stress in students from Jaypee Institute of Information Technology with 206 data (Ahuja & Banga, 2019). The algorithm with the best accuracy is the Support Vector Machine with 85.71%. Then in another study that aims to analyze whether a person is depressed or not with a machine learning approach (Kumawat et al., n.d.). machine learning algorithms used include XGBoost Tree, Random Trees, Neural Network, SVM, Random Forest, C5.0, and BayNet. From these results, it is evident that the C5.0 classifier provides the highest accuracy with 83.94% and for each classifier, the results are derived based on no pre-processing.

To support this research, we use the XGBoost algorithm to model the data we have with the help of the Python programming language. XGBoost is a technique or method of data modeling to solve regression and classification problems based on the Gradient Boosting Decision Tree efficiently and operate in parallel (Karo, 2020). In research, the classification of breast tumor types using the Extreme Gradient Boost algorithm has an accuracy rate of 97.60% to 97.80% (Handayani et al., 2017).

## MATERIALS AND METHOD

The research methodology used is Knowledge Discovery in Database (KDD) which has 5 stages, namely Data Selection, Preprocessing, Transformation, Data Mining, and Evaluation (Gullo, 2015). The flow of the research carried out is as shown in Figure 1.

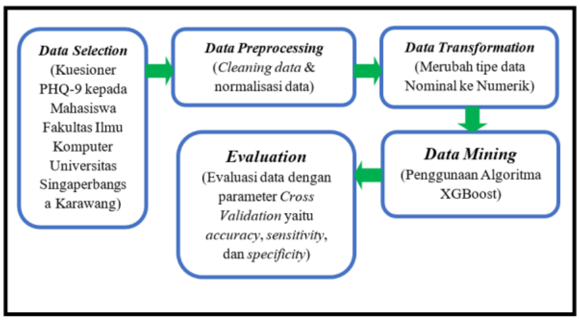


Figure 1. Research Flow

**Data Selection**

Based on data from Pangkalan Data Pendidikan Tinggi (Kemdikbud, 2021) the number of students at the Faculty of Computer Science, Singaperbangsa Karawang University is 1272. This data will be used to measure the number of samples to be used. The method of determining the sample in this study uses Simple Random Sampling. This method takes a random sample from the population regardless of the strata in the population and is commonly used in similar or homogeneous populations (Arieska & Herdiani, 2018). To determine the number of samples in this study, the author uses the Slovin formula with a population size of N = 1272 and an estimator error of d = 8%, so the number of samples is 140. The following is the Slovin formula:

$$n = \frac{n}{N \cdot d^2 + 1} \quad \text{.........................} (1)$$

Description :
n = sample size
N = population size
d = forecast error (tolerable error rate)

The dataset used by the study was obtained through a questionnaire based on PHQ-9 to students of the Faculty of Computers Sciences Unsika. The data that have been collected are 140 data with 9 questions as shown in table 1 while in table 2 are the types of depression tendencies which will be used as the target class.

Table 1. PHQ-9 Questions

| No | Variable | Descriptions |
|---|---|---|
| 1 | X1 | Lack of interest or enthusiasm in doing anything |
| 2 | X2 | Feeling down, sad, or hopeless |
| 3 | X3 | Difficulty sleeping / waking up easily, or sleeping too much |
| 4 | X4 | Feeling tired or lacking energy |
| 5 | X5 | Lack of appetite or eating too much |
| 6 | X6 | Lack of confidence or feel that you are a failure or have let yourself or your family down |
| 7 | X7 | Difficulty concentrating on something, for example reading the newspaper or watching television |
| 8 | X8 | Move or speak so slowly that others notice them. Or vice versa; feeling restless or restless so that you move more often than usual. |
| 9 | X9 | Feeling better off dead or wanting to hurt yourself in any way. |

Source: (Prabowo, et al, 2021)

Table 2. Depression Tendency Class

| No | Class | Type of Depressive Tendency |
|---|---|---|
| 1 | 0 | Symptoms of mild depression |
| 2 | 1 | Mild depression |
| 3 | 2 | Moderate depression |
| 4 | 3 | Severe depression |

Source: (Prabowo, et al, 2021)

### Data Validity Test

The validity test is useful to determine the suitability of the questionnaire in measuring and obtaining research data from the respondents. Validity is a measurement to prove the accuracy of research tools/instruments in measuring what you want to measure in research (Budiastuti & Bandur, 2018). The level of suitability of the data was assessed using the significance value. A data can be said to be valid if it obtains a significance value < 0.05. Tests were carried out using SPSS software using Bivariate Pearson correlation (Pearson Moment Product).

**Correlations**

| | | X01 | X02 | X03 | X04 | X05 | X06 | X07 | X08 | X09 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| X01 | Pearson Correlation | 1 | .616** | .464** | .548** | .275** | .424** | .393** | .295** | .327** | .678** |
| | Sig. (2-tailed) | | <,001 | <,001 | <,001 | .001 | <,001 | <,001 | <,001 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X02 | Pearson Correlation | .616** | 1 | .459** | .636** | .448** | .593** | .450** | .418** | .435** | .800** |
| | Sig. (2-tailed) | <,001 | | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X03 | Pearson Correlation | .464** | .459** | 1 | .604** | .476** | .386** | .343** | .250** | .341** | .712** |
| | Sig. (2-tailed) | <,001 | <,001 | | <,001 | <,001 | <,001 | <,001 | .003 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X04 | Pearson Correlation | .548** | .636** | .604** | 1 | .426** | .482** | .392** | .349** | .361** | .773** |
| | Sig. (2-tailed) | <,001 | <,001 | <,001 | | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X05 | Pearson Correlation | .275** | .448** | .476** | .426** | 1 | .474** | .371** | .363** | .312** | .687** |
| | Sig. (2-tailed) | .001 | <,001 | <,001 | <,001 | | <,001 | <,001 | <,001 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X06 | Pearson Correlation | .424** | .593** | .386** | .482** | .474** | 1 | .428** | .358** | .385** | .739** |
| | Sig. (2-tailed) | <,001 | <,001 | <,001 | <,001 | <,001 | | <,001 | <,001 | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X07 | Pearson Correlation | .393** | .450** | .343** | .392** | .371** | .428** | 1 | .387** | .271** | .648** |
| | Sig. (2-tailed) | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | | <,001 | .001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X08 | Pearson Correlation | .295** | .418** | .250** | .349** | .363** | .358** | .387** | 1 | .357** | .588** |
| | Sig. (2-tailed) | <,001 | <,001 | .003 | <,001 | <,001 | <,001 | <,001 | | <,001 | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| X09 | Pearson Correlation | .327** | .435** | .341** | .361** | .312** | .385** | .271** | .357** | 1 | .595** |
| | Sig. (2-tailed) | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | .001 | <,001 | | <,001 |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| Total | Pearson Correlation | .678** | .800** | .712** | .773** | .687** | .739** | .648** | .588** | .595** | 1 |
| | Sig. (2-tailed) | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | <,001 | |
| | N | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 | 140 |

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 2. Validation Test Result

It can be seen from Figure 2, all the numbers show numbers below 0.05, then the data is valid and feasible to use.

### Cronbach Alpha Reliability Test

Test Reliability or reliability is the consistency of a series of measurements made for research. Reliability is a benchmark for data to determine whether the data is relevant if it is done repeatedly. Reliability is essentially a tool to measure a questionnaire which is an indicator of a variable or construct (Ghozali, 2013).

**Case Processing Summary**

| | | N | % |
|---|---|---|---|
| Cases | Valid | 140 | 100.0 |
| | Excluded[a] | 0 | .0 |
| | Total | 140 | 100.0 |

a. Listwise deletion based on all variables in the procedure.

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| .861 | 9 |

Figure 3. Reliability Test Result

**Item-Total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|
| X01 | 8.7214 | 27.368 | .596 | .847 |
| X02 | 8.8929 | 25.909 | .740 | .834 |
| X03 | 8.2143 | 25.076 | .599 | .846 |
| X04 | 8.6214 | 25.230 | .692 | .836 |
| X05 | 8.7714 | 25.516 | .569 | .849 |
| X06 | 8.6500 | 24.833 | .635 | .842 |
| X07 | 8.9786 | 26.654 | .537 | .851 |
| X08 | 9.3571 | 28.015 | .487 | .855 |
| X09 | 9.5071 | 27.734 | .489 | .855 |

Figure 4. Overall Statistic Result

In Figure 3 shows the value of Cronbach's Alpha = 0.861, it can be said that the data is reliable. Then for details in Figure 4, the data sample with the lowest reliability is in the X02 sample with 0.834. The factors that influence the level of validity and reliability of a measuring instrument other than the instrument are the user of the measuring instrument who takes the measurement and the subject being measured (Sugiyono, 2014). The results of this reliability measurement can be used as a benchmark in the future if you want to conduct a similar research basis.

### Data Preprocessing

In data mining research, the process of data collection and pre-processing of data consumes most of the energy required for research. There are even some studies that say that 80% of data mining is data collection and pre-processing of data (Karo, 2016).

Pre-processing on the data is carried out to overcome conditions that are not ideal in the dataset such as the custom of missing data or scaling on data that is not the same. Pre-processing on the dataset is done by equalizing the existing scale in the dataset. This can be done by normalizing the data. Data normalization is done to standardize the data in the dataset, so that every value in the dataset variable has the same scale. This study uses the z-score normalization method, with the following equation:

$$Z = \frac{x - \mu}{\sigma} \text{.......................................................................} (2)$$

Description:
Z = normalized value
x = amount of data
μ = average value
σ = standard deviation

**Data Transformation**

At this stage the categorical data in the form of answers from respondents is converted into numerical data as can be seen in table 3. In each question the respondents choose one of the 4 available options, namely: Never symbolized 0, Several days is symbolized 1, More than half The time in question is denoted 2, and Every day is symbolized 3. Then an example of a dataset that has been transformed can be seen in table 4.

Table 3. Questionnaire Rating

| No | Value | Description |
|---|---|---|
| 1 | 0 | Never |
| 2 | 1 | A few days |
| 3 | 2 | More than half the time in question |
| 4 | 3 | Every day |

Source: (Prabowo, et al, 2021)

Table 4. Example of Transformation Result

| X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | Class |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 3 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2 | 1 | 3 | 1 | 2 | 3 | 0 | 0 | 3 | 2 |
| 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 1 |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 2 | 3 | 3 | 3 | 1 | 2 | 1 | 3 | 2 |
| 1 | 1 | 2 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 |

Source: (Prabowo, et al, 2021)

**Data Mining**

After normalizing the dataset, the dataset already has a value with the same scale. The next step is to break the data into two parts, namely the training data used to build the classification model and the test data used to test the model obtained from the training data. In this study, we tried to divide the data with several combinations. The first combination is 80% training data and 20% test data. The second combination is 50% training data and 50% test data, the third combination is 90% training data and 10% test data, the fourth combination is 70% training data and 30% test data. The classification modeling process in this study was carried out using the Extreme Gradient Boost (XGBoost) method.

Extreme Gradient Boost is a classification algorithm that is enhanced by gradient boosting decision trees and can also build boosted trees efficiently and operate in parallel (Chen, 2016). XGBoost is part of a family of trees (Decision tree, Random Forest, bagging, boosting, gradient boosting). XGBoost is an ensemble method with the aim of reducing bias as well as variance. In the interior of the regression tree, inside nodes represent values for the test attribute and leaf nodes represent decisions. The prediction result is the number of scores predicted by the tree (Cherif, 2019).

$$\hat{y}l = \sum_{k}^{K} f_k\,(x_i), f_k\ \varepsilon\ F \text{..............................................} (3)$$

$$obj(\theta) = \sum_{i=1}^{n} l\big(y_i, \widehat{yl}\big) + \sum_{k}^{K} \Omega\ (f_k) \text{..................} (4)$$

Where a is $\sum_{i=1}^{n} l\big(y_i, \widehat{yl}\big)$ differentiable loss function to measure whether the model is suitable for the training data set and $\sum_{k}^{K} \Omega\ (f_k)$ is an item that determines the complexity of the model. As the complexity of the model increases the corresponding score is reduced.

**Evaluation**

The model that has been made of course must be checked with several parameters to determine the quality of the model in classifying according to the truth. Accuracy alone is not enough to validate a classification model (Karo, 2020), so that other parameters are needed to validate the model. In this study, the authors use the accuracy, Precision, Recall, and F1-Score values to be included in the confusion matrix. TP is the class that is recorded as true and the model results are correct. FN is a class that is recorded as correct, but the model predicts wrongly, FP is a class in the data incorrectly, while the model results predict correctly and TN is a wrong class and the model also predicts incorrectly as shown in table 5.

Table 5. Model Validation

| Class | Class 1 (Prediction) | Class 2 (Prediction) |
|---|---|---|
| Class 1 (Actual) | TP (True Positive) | FN (False Negative) |
| Class 2 (Actual) | FP (False Positive) | TN (True Negative) |

Source: (Prabowo, et al, 2021)

Accuracy is the ratio of 'True' (positive and negative) predictions to the overall data. Accuracy in answering the question "What percentage of depression categories are correctly predicted and not depressed from the overall trend of depression data", Precision is the ratio of positive correct predictions compared to the overall positive predicted results. Precision answers the question "What percentage of students have a tendency to be depressed out of all students who are predicted to have a tendency to depression". Recall/Sensitivity is a ratio of correctly positive predictions compared to the overall data that are correctly positive (Martin & Nilawati, 2019). Recall answers the question "What percentage of students are predicted to have a tendency to depression from all students who actually have a tendency to depression". and F1 score is a compares weighted mean precision and recall. The parameters used for model evaluation can be seen in the following equation.

$$accuracy(y,\hat{y}) = \frac{1}{n_{sample}} \sum_{i=0}^{n_{sample}-1} 1(\hat{y}i = y_i) \quad (5)$$

$$precission = \frac{tp}{tp+fp} \quad ......(6)$$

$$recall = \frac{tp}{tp+fn} \quad ......(7)$$

$$F_\beta = (1 = \beta^2)\ \frac{precission\ x\ recall}{\beta^2 precission+recall} \quad ......(8)$$

## RESULT AND DISCUSSIONS

### XGBoost

The trend of depression dataset using the PHQ-9 questionnaire used in this study has 9 variables. As an initial reference, the authors divide the dataset into 80% training data and 20% test data to be tested using the XGBoost classification method. Next, the writer will divide the dataset with other combinations and also compare it with other classification methods to choose the best data combination and classification method to be tested with 4 parameters, namely accuracy, precision, recall/sensitivity, and F1-score. The results are as shown in table 6.

Table 6. Evaluation Result of XGBoost (80 : 20)

| Method | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| XGBoost | 83,33 | 85,92 | 83,33 | 83,02 |

Source: (Prabowo, et al, 2021)

This information is then continued by measuring the model using a confusion matrix with 3 parameters, namely precision, recall, and F1-score for each class. Then we get the results listed in table 7.

Table 7. Confussion Matrix (80 : 20)

| Validation Method | Class | Value (%) |
|---|---|---|
| Precission | Class 0 | 91 |
| | Class 1 | 75 |
| | Class 2 | 100 |
| | Class 3 | 50 |
| Recall | Class 0 | 95 |
| | Class 1 | 75 |
| | Class 2 | 50 |
| | Class 3 | 100 |
| F1-Score | Class 0 | 93 |
| | Class 1 | 75 |
| | Class 2 | 67 |
| | Class 3 | 67 |

Source: (Prabowo, et al, 2021)

Based on the evaluation of the model, it shows that the model does not yet have accurate results, then the author tries to use other combinations of data sharing, namely the distribution of 50% training data and 50% test data, then 90% training data and 10% test data, and the last is a combination distribution. 70% training data and 30% test data. From the data sharing experiment, the results are shown in table 8 and table 9.

Table 8. Evaluation Result of XGBoost

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | Data Split(%) |
|---|---|---|---|---|---|
| XGBoost | 75,71 | 77,65 | 75,71 | 76,09 | 50 : 50 |
| XGBoost | 92,86 | 94,29 | 92,86 | 92,06 | 90 : 10 |
| XGBoost | 83,33 | 85,92 | 83,33 | 83,02 | 70 : 30 |

Source: (Prabowo, et al, 2021)

Table 9. Confusion Matrix

| Data Split(%) | Validation Method | Class | Value (%) |
|---|---|---|---|
| 50 : 50 | Precission | Class 0 | 94 |
| | | Class 1 | 58 |
| | | Class 2 | 50 |
| | | Class 3 | 100 |
| | Recall | Class 0 | 86 |
| | | Class 1 | 75 |
| | | Class 2 | 40 |
| | | Class 3 | 75 |
| | F1-Score | Class 0 | 90 |
| | | Class 1 | 65 |
| | | Class 2 | 44 |
| | | Class 3 | 86 |
| 90 : 10 | Precission | Class 0 | 100 |
| | | Class 1 | 80 |
| | | Class 2 | 100 |
| | | Class 3 | 100 |
| | Recall | Class 0 | 100 |
| | | Class 1 | 100 |
| | | Class 2 | 50 |
| | | Class 3 | 100 |
| | F1-Score | Class 0 | 100 |
| | | Class 1 | 89 |
| | | Class 2 | 67 |
| | | Class 3 | 100 |
| 70 : 30 | Precission | Class 0 | 91 |
| | | Class 1 | 75 |

| | | |
|---|---|---|
| | Class 2 | 100 |
| | Class 3 | 50 |
| Recall | Class 0 | 95 |
| | Class 1 | 75 |
| | Class 2 | 50 |
| | Class 3 | 100 |
| F1-Score | Class 0 | 93 |
| | Class 1 | 75 |
| | Class 2 | 67 |
| | Class 3 | 67 |

Source: (Prabowo, et al, 2021)

Based on the results of the evaluation of the model from the 4 combinations of data sharing that have been carried out. Then it can be seen in table 8 that the combination of data sharing with 90% training data and 10% test data has the best model evaluation results compared to other data sharing combinations with 92.86% accuracy, 94.29% precision, 92.86% recall/sensitivity and F1- score 92.06%.

In the next step, the writer predicts the class on the test data which amounts to 10% using the model with the best performance, namely the model with a combination of training data and test data of 90: 10 each. Then the results are as shown in table 10.

Table 10. Prediction Result

| Index | Actual Class | Prediction Class |
|---|---|---|
| 4 | 1 | 1 |
| 11 | 0 | 0 |
| 99 | 0 | 0 |
| 75 | 2 | 2 |
| 1 | 1 | 1 |
| 129 | 0 | 0 |
| 109 | 0 | 0 |
| 134 | 0 | 0 |
| 80 | 3 | 3 |
| 72 | 0 | 0 |
| 82 | 2 | 1 |
| 24 | 1 | 1 |
| 28 | 0 | 0 |
| 125 | 1 | 1 |

Source: (Prabowo, et al, 2021)

In table 10 it can be seen that the model managed to predict almost all classes correctly except for data 82, the model incorrectly predicted the actual class that should be 2 but the XGBoost model made predicts the class in data 82, namely 1. The XGBoost model that was made was able to become a tool in conduct early diagnosis and screening of students who have a tendency to depression. However, development is needed in the future to get better performance on the XGBoost Model made.

The next step is to create a visualization on the XGBoost model with the best performance. Because XGBoost is part of a tree family (Decision tree, Random Forest, bagging, boosting, gradient boosting), the final XGBoost result is visualized in the form of a tree, as shown in Figure 5.

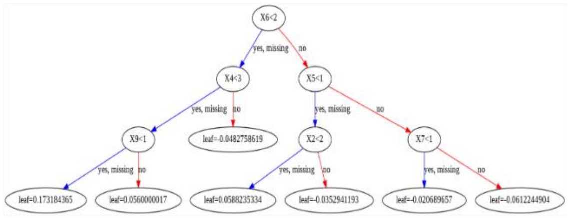


Figure 5. XGBoost Tree

**Comparison with Other Algorithms**

In this section, the author also performs a performance comparison with several other classification models. The algorithm used by the author to compare is K-Nearest Neighbor, Random Forest, Multi Layer Perception, Support Vector Machine and Decision Tree. Then the author also conducted several tests with the 4 combinations of data sharing described previously. The results of the model evaluation are described in table 11.

Table 11. Algorithm Comparison Result

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | Data Split (%) |
|---|---|---|---|---|---|
| KNN | 76,19 | 78,32 | 76,19 | 73,32 | 80 : 20 |
| SVM | 88,10 | 89,04 | 88,10 | 87,25 | |
| Random Forest | 85,71 | 86,23 | 85,71 | 84,99 | |
| XGBoost | 83,33 | 85,92 | 83,33 | 83,02 | |
| Multi Layer Percepton | 88,10 | 89,04 | 88,10 | 87,25 | |
| Decission Tree | 83,33 | 83,33 | 83,33 | 81,99 | |
| KNN | 77,14 | 76,37 | 77,14 | 75,32 | 50 : 50 |
| SVM | 80,00 | 80,30 | 80,00 | 79,64 | |
| Random Forest | 82,86 | 83,84 | 82,86 | 82,70 | |
| XGBoost | 75,71 | 77,65 | 75,71 | 76,09 | |
| Multi Layer Percepton | 85,71 | 88,60 | 85,71 | 84,86 | |
| Decission Tree | 78,57 | 78,98 | 78,57 | 78,61 | |
| KNN | 64,29 | 52,86 | 78,57 | 55,46 | 90 : 10 |
| SVM | 85,71 | 86,61 | 85,71 | 84,76 | |
| Random Forest | 85,71 | 86,61 | 85,71 | 84,76 | |
| XGBoost | 92,86 | 94,29 | 92,86 | 92,06 | |
| Multi Layer Percepton | 85,71 | 86,61 | 85,71 | 84,76 | |
| Decission Tree | 85,71 | 86,61 | 85,71 | 84,76 | |
| KNN | 76,19 | 78,32 | 76,19 | 72,32 | 70 : 30 |
| SVM | 88,10 | 89,04 | 88,10 | 87,25 | |
| Random Forest | 80,95 | 82,24 | 80,95 | 80,02 | |
| XGBoost | 83,33 | 85,92 | 83,33 | 83,02 | |
| Multi Layer Percepton | 85,71 | 84,92 | 85,71 | 84,87 | |
| Decission Tree | 83,33 | 83,30 | 83,33 | 81,99 | |

Source: (Prabowo, et al, 2021)

## CONCLUSION

Depression trend dataset using the PHQ-9 questionnaire on computer science students at the University of Singapore with 9 variables and 140 total data. In making the classification model, the distribution of the dataset affects the process of making the Extream Gradient Boost classification model. The best combination of data distribution is to divide it into 90% training data and 10% test data so as to produce an evaluation of the accuracy, precision, recall/sensitivity model, and the F1-scores are 92.86%, 94.29%, 92, 86% and 92.06%. This method is also still better than other methods, both in terms of model evaluation results and overall performance in each combination of dataset distribution. This shows that the model will predict more accurately if it is formed with many datasets. The more data that is trained, the easier it is for the model to predict new data. The dataset used in this study can still be added to the number of samples to improve model performance and also minimize generalizations on the model. The Extreme Gradient Boost model that was made, although it could not be a standard goal for students' depression tendencies, was able to become a tool for early diagnosis and screening of students with depression tendencies. The Extream Gradient Boost model that is made can still be developed in the future to get better performance.

## REFERENCE


Ahuja, R., & Banga, A. (2019). Mental stress detection in university students using machine learning algorithms. *Procedia Computer Science*, *152*, 349–353. https://doi.org/10.1016/j.procs.2019.05.007

Arieska, P. K., & Herdiani, N. (2018). Pemilihan teknik sampling berdasarkan perhitungan efisiensi relatif. *Jurnal Statistika*, *6*(2), 166–171. https://jurnal.unimus.ac.id/index.php/statistik/article/view/4322/4001

Budiastuti, Dyah & Bandur, A. (2018). *Validitas Dan Reliabilitas Penelitian*. Mitra Wacana Media. https://www.mitrawacanamedia.com/validitas-dan-reliabilitas-penelitian

Chen, T. et al. (2016). Xgboost: A scalable tree boosting system. *In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/https://doi.org/10.1145/2939672.2939785

Cherif, I. L. et al. (2019). On using eXtreme Gradient Boosting (XGBoost) Machine Learning algorithm for Home Network Traffic Classification. *2019 Wireless Days (WD)*, 1–6. https://doi.org/10.1109/WD.2019.8734193

Ghozali, I. (2013). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 21 Update PLS Regresi*. Badan Penerbit Universitas Diponegoro. https://mikroskil.ac.id/pustaka/index.php?p=show_detail&id=7026&keywords

Gullo, F. (2015). From patterns in data to knowledge discovery: What data mining can do. *Physics Procedia*, *62*, 18–22. https://doi.org/10.1016/j.phpro.2015.02.005

Handayani, A., Jamal, A., & Septiandri, A. A. (2017). Evaluasi Tiga Jenis Algoritme Berbasis Pembelajaran Mesin untuk Klasifikasi Jenis Tumor Payudara. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, *6*(4). https://doi.org/10.22146/jnteti.v6i4.350

Hasanah, U., Fitri, N. L., Supardi, S., & PH, L. (2020). Depression Among College Students Due to the COVID-19 Pandemic. *Jurnal Keperawatan Jiwa*, *8*(4), 421. https://doi.org/10.26714/jkj.8.4.2020.421-424

Karo, I. M. K. et al. (2020). A Hybrid Classification Based on Machine Learning Classifiers to Predict Smart Indonesia Program. *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, 1–5. https://doi.org/10.1109/ICVEE50212.2020.9243195

Karo, I. M. K. et al. (2016). Spatial clustering for determining rescue shelter of flood disaster in South Bandung using CLARANS Algorithm with Polygon Dissimilarity Function. *In 2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA)*, 70–75. https://doi.org/10.1109/ICMSA.2016.7954311

Kemdikbud, Pdd. (2021). *Profil Perguruan Tinggi Universitas Singaperbangsa Karawang*. https://pddikti.kemdikbud.go.id/data_pt/QzRENDE1QjMtOTgyRS00MDM0LUIyOUEtRUU5NDA4MDY4MUEw

Kumawat, K., Kumawat, G., Chakrabarti, P., Poddar, S., Chakrabarti, T., Hussaine, J., Kamali, A., Bolsev, V., & Kateb, B. (n.d.). A Machine Learning Technique to Analyze Depressive Disorders. *Research Square*, 1–10.

Kusuma, M. D. S. et al. (2021). TINGKAT GEJALA DEPRESI PADA MAHASISWA PROGRAM STUDI SARJANA KEPERAWATAN DI

INSTITUTE TEKNOLOGI DAN KESEHATAN (ITEKES) BALI. *Jurnal Riset Kesehatan Nasional*, *5*(1), 29–34. https://doi.org/http://dx.doi.org/10.37294/jrkn.v5i1.310

Kusuma, P. D., Marchira, C. R., & Prawitasari, S. (2018). Patient health questionnaire-9 (PHQ-9) efektif Untuk mendeteksi risiko depresi postpartum. *Jurnal Keperawatan Respati Yogyakarta*, *5*(3), 428–433.

Lempang, G. F., Walenta, W., Rahma, K. A., Retalista, N., Maluegha, F. J., & Utomo, F. I. P. (2021). Depresi Menghadapi Pandemi Covid-19 pada Masyarakat Perkotaan (Studi Literatur). *Pamator Journal*, *14*(1), 66–71. https://doi.org/10.21107/pamator.v14i1.9854

Martin, M., & Nilawati, L. (2019). Recall dan Precision Pada Sistem Temu Kembali Informasi Online Public Access Catalogue (OPAC) di Perpustakaan. *Paradigma - Jurnal Komputer Dan Informatika*, *21*(1), 77–84. https://doi.org/10.31294/p.v21i1.5064

Muslim, I., & Karo, K. (2020). Implementasi Metode XGBoost dan Feature Importance untuk Klasifikasi pada Kebakaran Hutan dan Lahan. *Journal of Software Engineering, Information and Communication Technology*, *1*(1), 10–16.

Perhimpunan Dokter Spesialis Kedokteran Jiwa Indonesia. (2020). *Masalah Psikologis Di Era Pandemi Covid-19*. http://www.pdskji.org/home

Satuan Tugas Penanganan Covid-19. (2021). *Peta Sebaran Kasus Covid-19 Di Indonesia*. https://covid19.go.id/peta-sebaran

Sugiyono. (2014). *Statistik untuk Penelitian* (E. Mulyatiningsih (ed.)). Alfabeta. https://adoc.pub/statistik-untuk-penelitian.html