

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

The following conclusions can be made based on the study's findings comparing how well the XGBoost and CatBoost algorithms classified students' levels of depression :

1. The XGBoost and CatBoost algorithms can categorize students into four depression categories: Normal, Mild Depression, Moderate Depression, and Severe Depression, according to the results and performance of these algorithms. Three data splitting scenarios—60:40, 70:30, and 80:20—were used for the experiment. Based on the test results, the performance of the XGBoost and CatBoost algorithms differed across each data split scenario. XGBoost achieved an accuracy of 0.80 in the 60:40 scenario, 0.86 in the 70:30 scenario, and 0.89 in the 80:20 scenario. Meanwhile, CatBoost achieved an accuracy of 0.85 in the 60:40 scenario, 0.80 in the 70:30 scenario, and 0.85 in the 80:20 scenario. These findings show that the data split proportions utilized during training and testing had an impact on both algorithms' performance.
2. Under an 80:20 data split, XGBoost was found to be the best model in this investigation. This model obtained an accuracy of 0.89, precision of 0.91, recall of 0.81, F1-score of 0.84, and AUC of 0.97. Instead of depending just on one metric, the model was chosen by concurrently taking into account the outcomes of multiple evaluation metrics. XGBoost with an 80:20 split showed the best accuracy, precision, and F1-score values when compared to other scenarios and algorithms, indicating good classification performance. Although CatBoost achieved an AUC of 0.98 in the 60:40 scenario, this value merely indicated the model's higher discriminative ability in that specific scenario and did not directly establish CatBoost as the best model overall. The optimal model for categorizing students' depression levels was found to be XGBoost with an 80:20 data split, taking into account the

balance of performance across all evaluation criteria, particularly accuracy, precision, recall, and F1-score.

5.2 Recommendations

The study's findings allow for the following suggestions for additional investigation :

1. A greater range of algorithms, including Random Forest, LightGBM, Support Vector Machine, and other ensemble techniques, may be compared in future studies. A more complete picture of the best method for categorizing students' depression levels might be obtained from such a comparison.
2. The optimal model might be included into a web or mobile application with follow-up suggestions in future study.