# CHAPTER V
# CONCLUSIONS AND RECOMMENDATION

This chapter presents the conclusions based on the research results and recommendations for future research development. The conclusions are formulated as a summary of the main findings and as answers to the research objectives. The recommendations are provided as suggestions for improving future studies so that the results can be more optimal and beneficial.

## 5.1. Conclusions

Based on the research results regarding the comparison of XGBoost and LightGBM methods with Random Oversampling (ROS) for obesity level classification using the Dhaka Obesity dataset, the conclusions are as follows.

1. The implementation of XGBoost and LightGBM models with the application of Random Oversampling (ROS) to the training data was successfully carried out for obesity level classification. The Dhaka Obesity dataset, which initially consisted of 2,182 respondent data and 17 attributes, underwent several preprocessing stages, including adult respondent selection, target adjustment based on the Asia Pacific BMI classification adopted by the Indonesian Ministry of Health, data cleaning, categorical variable encoding, data splitting using stratified train test split, and the application of ROS only to the training data. After preprocessing and duplicate data removal, the final dataset consisted of 2,125 data with five target classes, namely Underweight, Normal, Overweight, Obesitas_I, and Obesitas_II. This implementation process produced a measurable modeling workflow that could be used to build and evaluate the XGBoost and LightGBM models.
2. Variations in data split ratios and stepwise manual hyperparameter testing affected the performance of the XGBoost + ROS and LightGBM + ROS models. Based on the data split ratio testing, the 80:20 ratio produced the best results for both models because it achieved higher performance than the 70:30 and 60:40 ratios. Furthermore, stepwise manual hyperparameter

testing produced different best configurations for each model. XGBoost + ROS obtained the best configuration with learning_rate = 0.10, max_depth = 7, n_estimators = 200, subsample = 0.90, and colsample_bytree = 0.90. Meanwhile, LightGBM + ROS obtained the best configuration with learning_rate = 0.10, max_depth = 7

3. , n_estimators = 200, subsample = 0.80, and colsample_bytree = 0.80. These results show that selecting appropriate split ratios and hyperparameters can improve classification performance and help the models achieve more balanced performance across all target classes.
4. The application of Random Oversampling (ROS) improved the performance of both XGBoost and LightGBM compared to the scenarios without ROS. In XGBoost, ROS increased accuracy from 0.955294 to 0.962353, F1 macro from 0.929277 to 0.942597, and balanced accuracy from 0.921420 to 0.937209. In LightGBM, ROS increased accuracy from 0.955294 to 0.967059, F1 macro from 0.926915 to 0.948016, and balanced accuracy from 0.921230 to 0.942773. In addition, ROS helped improve recognition of the Overweight class, which was previously more difficult to predict. Based on the overall evaluation results, the best model in this study was LightGBM + ROS with an 80:20 split ratio and the best hyperparameter configuration, achieving an accuracy of 0.967059, precision macro of 0.955319, recall macro of 0.942773, F1 macro of 0.948016, and balanced accuracy of 0.942773. Therefore, LightGBM with Random Oversampling was the most optimal model for obesity level classification in this study.

## 5.2. Recommendations

Based on the research results, several recommendations for future studies are as follows.

1. This study only used one dataset, namely the Dhaka Obesity dataset. Therefore, future research is recommended to use more diverse datasets in terms of data size, respondent characteristics, and data sources. This is important to test the model’s generalization ability on different data.

2. This study only applied Random Oversampling (ROS) as the class imbalance handling technique. Therefore, future research is recommended to compare other balancing methods, such as SMOTE, ADASYN, and SMOTETomek, to identify the most effective method for improving performance and balancing classification results across classes.
3. The models used in this study were limited to XGBoost and LightGBM. Future research may add comparisons with other algorithms, such as CatBoost, Random Forest, Support Vector Machine, or deep learning based approaches, to obtain a broader understanding of classification model performance.
4. This study has not explored model interpretability analysis in depth, such as feature importance or explainable AI (XAI). Therefore, future research is recommended to add model interpretation analysis to understand the factors that have the greatest influence on obesity level classification.
5. The testing in this study was still limited to certain data split ratios and hyperparameter variations. Future research may explore broader hyperparameter ranges or use other optimization approaches, such as Grid Search or AutoML, to obtain more optimal model performance.

*This page is intentionally left blank.*