BAB V
CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

Based on the results of the research and testing conducted on traffic violation data from Sistem Informasi Penelusuran Perkara (SIPP) of the Tulungagung District Court in 2023 using the *Gradient Boosted Trees* (GBT) algorithm, it can be concluded that:

1. The GBT algorithm was successfully applied to classify the severity levels of traffic violations into minor, moderate, and severe violations. The classification process utilized the attributes Article, Vehicle, and Evidence as predictor features, with the nominal fine amount serving as the basis for class labeling. Test results indicate that the model is capable of learning the patterns of relationships among attributes and generating accurate predictions regarding the severity level of traffic violations.
2. ased on all testing scenarios conducted, the model's performance is influenced by the data split ratio and the *hyperparameter* configuration used. Tests using data split schemes of 70:30, 75:25, and 80:20 showed that the 80:20 scheme produced the best performance. Additionally, *hyperparameter optimization* via GridSearch improved the model's performance compared to the GBT model without optimization. The best model was obtained using the parameter combination from GridSearch, with an *accuracy* of 64.54%, *precision* of 64.24%, *recall* of 46.24%, and an *F1-Score* of 47,53%. Based on these results, *hyperparameter optimization* was shown to have a positive impact on the model's ability to classify traffic violation levels.

5.2 Recommendations

Based on the results of this study, there are several recommendations that can be considered for future research, namely:

1. Future research is recommended to use a dataset with a larger volume of data and a longer observation period so that the model can produce better performance that is more representative of real-world conditions, such as traffic violation data spanning several years or from different regions.

Increasing the amount of data is expected to help the model learn violation patterns more comprehensively, thereby resulting in better classification performance.

2. Future research could develop the model by adding more diverse attributes or features, such as the location of the violation, the violator's age, or more detailed vehicle types. Additionally, comparisons can be made with other classification algorithms such as *Random Forest*, XGBoost, SVM, and LightGBM to identify the method that yields the best performance, thereby enabling its optimal use in the traffic violation data classification process. *Hyperparameter optimization* can also be performed using more efficient methods, such as *Bayesian Optimization* or *Optuna*, to generate the best parameter combinations with reduced computational time.