# ABSTRACT

Regional development disparities pose significant challenges to achieving equitable growth in Indonesia. The Ministry of Villages, Development of Disadvantaged Regions, and Transmigration utilizes the Village Development Index to assess village progress and classify villages into five categories, which include Independent, Advanced, Developing, Disadvantaged, and Severely Disadvantaged. This study compares the performance of the Random Forest and Gradient Boosting algorithms in classifying village status in Central Java based on 2024 data. The class imbalance problem is addressed using Random Oversampling technique. The dataset comprises 7,810 villages and includes three predictor variables: Social Resilience Index, Economic Resilience Index, and Environmental Resilience Index. Model performance was evaluated using three data split scenarios: 90:10, 80:20, and 70:30. Random Forest algorithm achieved the best results under the 90:10 split, with an accuracy of 98.97% and G-Mean of 98.91%. In comparison, Gradient Boosting algorithm under the 80:20 scenario reached an accuracy of 96.69% and a G-Mean of 74.38%. Random Forest also outperformed Gradient Boosting in other metrics, achieving 98.78% precision, 98.91% recall, and 98.85% F1-score. These findings demonstrate the effectiveness of Random Forest in handling multiclass classification on imbalanced data and support the application of machine learning models to inform more objective and targeted village development policies.

**Keywords**: Regional Development Disparities, Random Forest, Gradient Boosting, Random Oversampling, Class Imbalance, Multiclass Classification