

## ABSTRACT

This study presents an application of machine learning methods to classify the Human Development Index (HDI) levels of districts/cities in Java, Indonesia, in 2024 using Support Vector Machine (SVM), Random Forest, AdaBoost, and Stacking Ensemble. The data used are secondary data obtained from the Central Statistics Agency, covering economic, education, and health variables, with a data split of 70% for training and 30% for testing. Class imbalance is addressed using the Synthetic Minority Over-sampling Technique (SMOTE), while hyperparameter tuning is performed using GridSearchCV. Model performance is evaluated using accuracy and F1-Score. The results show that SMOTE improves class distribution and enhances the model's ability to recognize minority classes. The SVM model achieves the best performance with an accuracy of 0.91 and an F1-Score of 0.87, followed by Random Forest and Stacking Ensemble with an accuracy of 0.88 and an F1-Score of 0.85, and AdaBoost with an accuracy of 0.85 and an F1-Score of 0.84. SHAP analysis indicates that minimum wage, per capita expenditure, and access to proper sanitation are the most influential variables in HDI classification. The findings demonstrate that machine learning methods are effective for HDI classification.

**Keywords:** Human Development Index (HDI), Classification, Ensemble Learning, Stacking, Random Forest, SVM, AdaBoost, SMOTE, GridSearchCV, SHAP.