

## **ABSTRACT**

### ***AI Credit Scoring to Predict Creditworthiness Test Results Using the XGBoost Algorithm***

*Author:*

Muhammad Hilmi Kafabi

19102115

*Lending institutions represent a highly potential business. Loan businesses can yield significant profits for companies. However, such businesses also expose companies to substantial financial losses. Therefore, lending institutions need to conduct comprehensive analyses of basic information and credit histories of applicants to estimate the likelihood of credit repayment and then decide whether the applications can be approved. Credit data, as financial data, has its own characteristics such as credit data imbalance due to default applications, credit features, and complex feature relationships. The Gradient Boosting algorithm, specifically the eXtreme Gradient Boosting (XGBoost) algorithm, has demonstrated superior performance in Kaggle data competitions. XGBoost allows specialized objective functions and evaluation metrics to guide the model in various situations, such as handling imbalance using the AUC metric as the objective function. It learns how to handle missing values by computing gains in left and right subtrees to determine the default direction of the missing value. XGBoost is well-suited for credit data's missing value handling mechanisms. XGBoost is used to predict bankruptcy. The dataset used was obtained from the Kaggle platform, containing 150,000 data points. However, before modeling, data pruning was performed to achieve data balance, resulting in 19,250 remaining data points. From this data, results obtained with `n_estimators` set at 50 and `max_depth` at 24 achieved an accuracy of 80.94% and an F1-Score of 81.10%. Meanwhile, for the worst-case scenario, using `n_estimators` of 100 and `max_depth` of 12 resulted in an accuracy of 78.89% and an F1-Score of 78.93%.*

***Keywords : Credit Scoring, XGBoost, Machine Learning***