



# **PROJECT REPORT**

**Comparison of Random Forest Algorithm  
Accuracy with XGBoost Using Hyperparameters**

**KEVIN STEFANUS**

**19.K1.0009**



**Faculty of Computer Science  
Soegijapranata Catholic University  
2023**

## ABSTRACT

*Diabetes is one of the most dangerous diseases in the world and many people do not realize that they have diabetes in them. So many factors affect the occurrence of diabetes such as pregnancies, glucose, blood pressure, skinthickness, insulin, BMI, diabetes pedigree function, and age. so diabetes threatens silently and will appear suddenly. Therefore, this study will make a diabetes prediction using Random Forest and XGBoost algorithms. The model will be evaluated with accuracy, F1-Score, recall ,and precision. for randomization or random state will use random state 0 and 45. The results obtained from the comparison of these two algorithms are the highest accuracy of the random forest algorithm has a value of 88,98% while the highest accuracy of XGBoost gets an accuracy value of 87,00% at random state 45 and data division 90/10, while at random state 0 random forest has the highest accuracy value also with a value of 78,43% with data division 90/10 while XGBoost gets the highest accuracy value of 76,47% at data division 90/10. It can be concluded that random forest is better at predicting diabetes data than the XGBoost algorithm.*

*Keyword: Prediction, Random Forest, XGBoost, Accuracy, Hyperparameters*

