

ABSTRACT

May Linda Tri Alfianti.5040709

COMPARISON OF THE PERFORMANCE OF THE RANDOM FOREST, ADABOOST, AND XGBOOST ALGORITHM IN PREDICTING THE RISK OF OSTEOPOROSIS DISEASE.

Undergraduate Thesis, Informatics, Faculty of Industrial Technology, Gunadarma University, 2024.

Keywords: *Osteoporosis, Random Forest, Adaboost, Xgboost, Machine Learning, Risk Prediction.*

(xi + 68 + Attachment)

Osteoporosis is a bone disease characterized by a decrease in bone mass and an increase in bone fragility, which raises the risk of fractures, particularly in the hip, spine and wrist. This disease can be suffered by both men and women, especially in older age. Risk factors for osteoporosis include being female, older age, lack of physical activity, smoking and alcohol consumption. This disease often does not show symptoms in the early stages, making early prediction very important for prevention and proper management. This research aims to compare the performance of three machine learning algorithms, namely Random Forest, Adaboost and Xgboost, in predicting the risk of osteoporosis. The model's performance evaluation is conducted using metrics from the confusion matrix. The dataset used is "Osteoporosis Risk Prediction" with 1781 data points, which are divided into three data splitting schemes: 80:20, 70:30, and 60:40. In the Random Forest algorithm, the 80:20 scheme yields the best results with an accuracy of 87.11% and a precision of 89.09%. Meanwhile, the Adaboost algorithm shows the best performance in the 60:40 scheme with an accuracy of 92.01% and a precision of 93.13%. The XGBoost algorithm demonstrates the best results in the 80:20 scheme with an accuracy of 90.20% and a precision of 90.77%. The research findings indicate that the Adaboost algorithm has the best performance among the three tested algorithms. After that, implementing the osteoporosis risk prediction model into the website, which allows users to make predictions based on 14 variables: age, gender, family history, ethnicity, medication use, history of fractures, hormonal changes, weight, calcium intake, vitamin D intake, physical activity, smoking, alcohol consumption and medical conditions.

References (2020-2024)