ABSTRACT

Implementation of the C4.5 Decision Tree Algorithm in Diagnosing Kidney Stone Disease Based on Clinical Data

By

Safina Octaviana Putri

21110026

Kidney stones are a disease caused by the formation of hard material in the urine due to the interaction between salts and minerals. This disease can cause recurrent urinary tract infections, impaired kidney function, hematuria (blood in the urine), to the risk of causing kidney cancer. Kidney stone disease has been suffered by around 1,499,400 people in Indonesia so that prediction of kidney stone disease is needed. This study aims to build a classification model to predict kidney stone disease based on patient clinical data using the C4.5 Decision Tree algorithm because it is efficient in handling numerical data and producing models that are easy to interpret. The data used comes from RSUD Cideres Majalengka with 7 clinical variables used. The model was evaluated using the cross-validation method with 5 folds. The results showed that the model was able to achieve an accuracy of 91.39% on training data and 89.88% on test data. The model was also tested on 929 test data and successfully predicted 835 data correctly, with an overall accuracy of 90%. The cross-validation test showed an average accuracy of 87.44%, indicating consistent model performance and no overfitting or underfitting.

Keywords: Kidney stones, classification, diagnosis, early detection, Decision Tree C4.5