Abstract

Diabetes is a disease that causes death no. 3 largest in Indonesia, which is estimated in 2015 there were as many as 10 million diabetics in Indonesia with a death rate of 6.7%. The high mortality rate from this disease requires a system that can classify patients with possible diabetes. This study implements the Modified Balanced Random Forest (MBRF) method as a classification method. The MBRF algorithm is a development of the Random Forest algorithm that can handle large amounts of data. This method is then validated with K-Fold Cross Validation to find the optimal value of the MBRF model. The number of trees will affect the accuracy of Random Forest predictions. MBRF is a development from Random Forest which is able to handle imbalanced data which is a common problem in machine learning classifications due to the disproportionate spread of ratios in each class. Tests show that the MBRF system model has an accuracy of 97.8%, this presentation is increased from the Random Forest model which produces an accuracy of 92.8%.