Abstract

Diabetes is a chronic disease that occurs when the pancreas does not produce enough insulin or when the body cannot use the insulin it produces effectively. Insulin is a hormone that regulates blood glucose. Type 1 diabetes mellitus (T1DM) occurs in 90% of children and adolescents. Data containing gene expression profiles in children with T1D and T2D, measurements were taken at the time of initial diagnosis and repeated 4 months later, and also after receiving treatment, then the gene expression matrix was then transposed and three demographic features were considered important, namely age, gender, and race. After carrying out the GSA process, the dataset will be classified. Using the main method namely Adaptive Boosting (AdaBoost), then added 2 ensemble method as a compiration,namely KNeighbors (KNN) and Multi-Layer Perceptron (MLP), Then hyperparameter tuning is carried out with the aim of finding the most optimal value by improving model performance. Parameter scanning in the tuning process is carried out using search cross validation (grid search CV). This will be the benchmark for evaluating the three models used to obtain the most optimal results, namely AdaBoost with an accuracy of 0.666 and an F1-Score of 0.769.

Keywords: Diabetes mellitus, Gravitational Search Algorithm, Multi-Layer Perceptron, Adaptive Boosting, KNeighbors