**Abstract**

Biomedical datasets pose a unique challenge for machine learning and datamining techniques in order to extract accurate, comprehensible and hidden knowledge. This paper comprehensively investigates the role of a biomedical dataset such as hypothyroid dataset on the classification accuracy of an algorithm. The datasets are retrieved from UCI machine learning repository, Hypothyroid is a kind of disease, which occurs due to the insufficient production of thyroid hormones to the thyroid gland. This data classification is based on machine learning algorithms to provide very accurate predictions for real-world datasets and also to quantify the complexity of a biomedical dataset in terms of its missing values, imbalance ratio, and information gain from that dataset. The data mining classification algorithms in Weka tool is used to classify the data. The predicted values for the classifiers were evaluated using 10-fold cross validation and the results were compared.