**Background**: As the spiking growth of medical data due to the rise of health information technology in the past decades, it would be a tragic lost if we do not develop an effective and application algorithm to utilize the sea of medical data. Decision tree such as ID3 is one of the most used classification method. C4.5 and C5.0 are improved versions of ID3. Random forest is an extended decision tree method. It was created to resolve the error propagation problem associated with traditional decision. This study employed the algorithm of C5.0 decision tree and random forest to analyze and predict health care problems.

**Methods**: Healthcare database was downloaded from the center for machine learning and intelligent system, University of California, Irvine. Original dataset has 16 different classifications. In the study, we combined all the cases and classify the data into only two categories: normal (245 cases) and abnormal (the remaining 207 cases). Both the C5.0 decision tree and the random forest were done in Python platform. Before the training of data, the database was shuffled before picking the first 350 data points as the training dataset. Totally, three decision trees based on C5.0 and three random forests were built. For the random forest, totally 500 trees were built and each tree consists of 18 attributes and 221 cases. The sampling process is done with replacement.

**Results**: The averages sensitivity for decision tree and random forest are 79.03% and 84.07% respectively. The average specificity for decision tree and random forest are 70.10% and 80.74% respectively. The attributes used in the decision tree splitting, only three attributes were used by all three decision trees. The attributes used to build random forest models were much more consistent than the decision tree (13 out of top 20 most important attributes were common).

**Conclusions**: The results showed that the decision tree in general is worse than random forest in the classification. However, the sensitivity and specificity are still not high enough compare to doctor’s diagnosis, even though we only classify them in two categories: normal and abnormal. In order to apply to clinics, the algorithms need to be improved. Possible direction might be to consult experienced doctors to combine expert system with machine learning. It can help to preselect attributes before the data training. Another direction to improve random forest is to examine each tree in the random forest to examine the heterogeneity of each tree, then giving weight in the final decision phase.