**Weekly report 6**

**Group 14: Learners.**

**Heart Disease Prediction.**

**25-03-23**

Members:

1. Varun Parekh - AU2040011
2. Vanshit Shah - AU2040098
3. Madhvendra Jhala - AU2040162
4. Ushmay Patel - AU2040253

As we had mentioned in the previous weekly report that after studying and using Logistic Regression for heart disease prediction, we were able to achieve accuracy of around 70%, which could still be increased. Also as heart disease being a topic which requires very high accuracy as it being a matter of life and hence we had decided to go through “XGBoost”.

**XGBoost:**

What is XGBoost?

It is an optimized gradient boosting machine learning algorithms under the framework of Gradient Boosting.

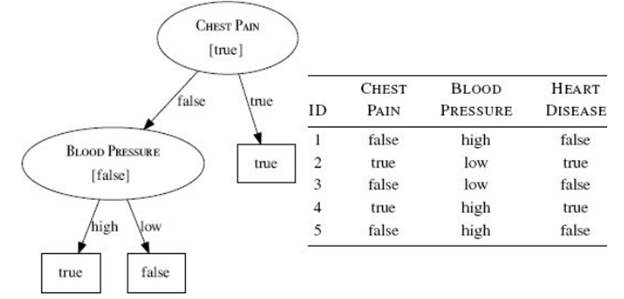
XGBoost builds upon the following methods: Supervised machine learning, Decision Tree, Ensemble Learning and Gradient Boosting.

The concept of supervised machine learning algorithm is that the after using the algorithm for training the model and to find patterns with the given labels and features. And then it uses the trained model to predict the labels on a new(unseen) dataset’s features.

Decision Tree:

It predicts the output label by evaluating a tree of “if-then-else” true/false feature question and then estimating the minimum number of questions needed to assess the probability of making a correct decision.

For e.g.: In our problem of prediction of heart disease following is an example of a simple decision tree:



Boosting: The principle of boosting algorithms is that we firstly built a model and use it on training dataset, then based upon the results obtained we build another model to rectify the errors present in the first model. So the main concept behind the algorithms is to build models sequentially, such that the subsequent models reduce the errors of the previous model.

XGBoost stands for Extreme Gradient Boosting, which is a scalable and distributed gradient-boosted decision tree machine learning library.

XGBoost is largely built for maximizing machine learning model performance and computational speed.

In the following week we plan to implement XGBoost algorithm on our dataset.