XGBoost

THEORY

1) XgBoost stands for Extreme Gradient Boosting, which was proposed by the researchers at the University of Washington. It is a library written in C++ which optimizes the training for Gradient Boosting.
2) In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost. Weights are assigned to all the independent variables which are then fed into the decision tree which predicts results.
3) The weight of variables predicted wrong by the tree is increased and these variables are then fed to the second decision tree. These individual classifiers/predictors then ensemble to give a strong and more precise model.
4) It can work on regression, classification, ranking, and user-defined prediction problems.

QUIZ

1) What are the disadvantages of XGBoost?
2) What are the advantages of XGBoost?

ANSWER

1) XGBoost requires manual label encoding for categorical features before feeding them into the models and is again sensitive to outliers.
2) XGBoost consists of a number of hyper-parameters that can be tuned, also has an inbuilt capacity to handle missing values.