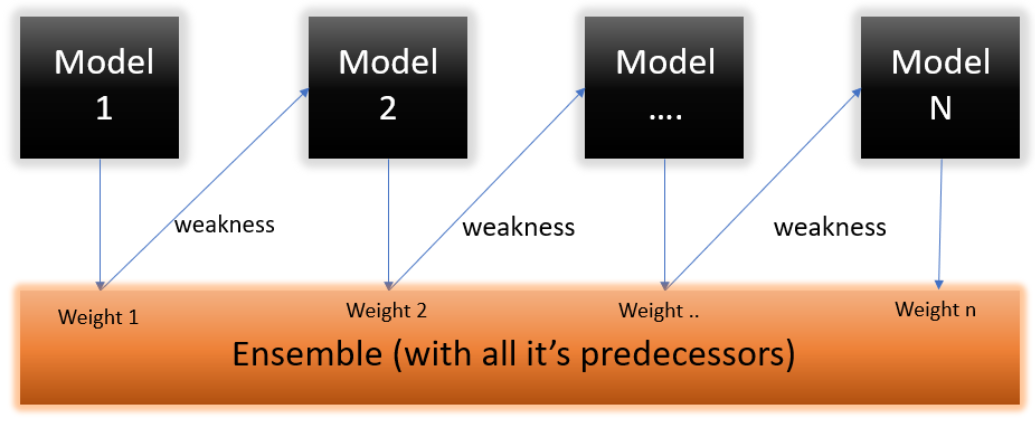
|  |  |
| --- | --- |
| SR | Title |
| Q1 | Explain Gradient Boost/ Any ML Algorithm? |
|  | * Supervised/ Unsupervised |
| * Parametric/ Non parametric/Assumptions |
| * Effect of outliers |
| * Effect of scaling |
| * Hyper parameter tunning |
| * Evaluation measures: r2/cnf\_matrix, ROC/AUC |
| Use above points to explain any algorithm if interviewer ask to explain | |

**Gradient Boost Algorithm**

Adaboost is a supervised machine learning algorithm and it is based on boosting ensemble technique. It is used for both regression and classification problem. In boosting, all the weak learners are connected in series. Each weak learner learns from the mistakes of his predecessor. The most common weak learner or base model used in gradient boost is Decision tree with more than one split. Usually split between 8 to 32.

Step1:

In gradient boost, we build a first model on training dataset and makes prediction. The prediction of first model is nothing but an average of all output points.

Step2:

In second step, we find residuals. Residual is nothing but the error between predicted and actual data points.

Step3:

In third step, we build a decision tree-based model for independent variables and residual. Note we are not taking output variable.

At this stage, the predictions are nothing but the sum of predictions of first model and second model.

Predictions = First model prediction + (leaning rate) \* (second model predictions)

It will keep adding decision and training tree based models until and unless a low error is received.

**Effect of Scaling on Gradient Boost Algorithm**

Algorithms Like Decision Trees, Random Forest, Adaboost, Gradient Boosting, etc. are not significantly affected by Feature Scaling Since the trees in these algorithms are constructed based on conditions and are not dependent on the Range of values. Algorithms like Linear Discriminant Analysis and Naive Bayes Also do not require feature scaling.

**Effect of Outlier on Gradient Boost Algorithm**

GB has great usability that can deal with missing values, outliers, and high cardinality categorical values on your features without any special treatment.

As we know, in gradient boost the decision tree are with more that one split, so it is less affected by outlier compared to Adaboost.

**Gradient Boost Hyper Parameters**

As we know, gradient boost uses decision tree as base estimator/model so it shares hyper parameters with decision tree.