GBM

THEORY

1) Boosting is a method of converting weak learners into strong learners. In boosting, each new tree is a fit on a modified version of the original data set.
2) The objective is to minimize the loss function by adding weak learners using gradient descent.
3) MSE loss function will be $L = \frac{1}{n}\Sigma(y_i - \gamma_i)^2$ as now the target column is continuous. Make $\frac{dL}{d\gamma}$ zero & calculate γ.
4) Now calculate the psuedo residuals which are observed value - predicted value.
5) Then we find the output values for each leaf of our decision tree. The new prediction is given as previous prediction + learning rate * tree made on residuals.

QUIZ

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

ANSWER

1) It may cause overfitting as well as overemphasizing the outliers, also is a time-consuming and memory exhaustive algorithm.
2) Boosting algorithms follow ensemble learning which enables a model to give a more accurate prediction that cannot be trumped.