**XGBoost**

## Aaryan

## CO21BTECH11001

XGBoost (eXtreme Gradient Boost) is an algorithm which follows the principle of gradient boosting. The major difference with GBM is that XGBoost used a more regularized model formalization to control over-fitting, which gives it better performance. It uses boosted trees for regression or classification problems.

**Input:** , a loss function , max iterations = M, a learning rate .

1. Initialize the model with a constant value  
   which we can calculate by
2. Iterate k =1 to M:
   1. Compute the gradients () and hessians ()  
      **for** i =1 **to** m:   
         
         
      **end**
   2. Fit a base learner (tree) with input by solving:
   3. Update the model:
3. Output function: .

**Questions:**

1. XGBoost is a/an:  
   (a) Supervised learning algorithm  
   (b) Unsupervised learning algorithm  
   **Ans.** (a)
2. XGBoost can be used for:  
   (a) Regression  
   (b) Classification  
   (c) Both of these  
   (d) None of these  
   **Ans.** (c)
3. Which of the following is an important assumption in XGBoost algorithm?  
   (a) The value of every feature of data should be between 0 and 1.  
   (b) Each sample model ( ) is a stump.  
   (c) Loss function can be approximated by a second order approximation.  
   (d) None of the above  
   **Ans.** (c)
4. Why is regularization used in XGBoost?  
   **Ans.** Regularization prevents overfitting of the model by reducing the dependency of hypothesis function on one particular training data.
5. What are the advantages of using XGBoost?  
   **Ans.** It does not need normalized features and work with non-linear data also.
6. What are the disadvantages of using XGBoost?  
   **Ans.** It is sensitive to outliers. It does not perform so well on sparse and unstructured data.