18th April Quiz

7 out of 7 correct

1. Which of the following is not a hyperparameter in Xgboost?

- ○ Number of trees
- ○ Learning rate
- ○ Max depth
- ◉ Number of features

Explanation: Hyperparameters are the parameters that are not learned by the model, but are set by the user before training the model. Xgboost hyperparameters include the number of trees, learning rate, max depth, and others, but the number of features is not a hyperparameter in Xgboost.

2. Which of the following is a common technique to prevent overfitting in Xgboost?

- ○ Increasing the learning rate
- ○ Increasing the max depth
- ○ Reducing the number of trees
- ◉ Using regularization

Explanation: Overfitting occurs when a model is too complex and fits the training data too closely, resulting in poor performance on unseen data. Regularization is a technique that adds a penalty to the loss function to discourage the model from overfitting the training data.

3. What is the default objective function in Xgboost for binary classification problems?

- ○ Binary logistic regression

- (●) Binary cross-entropy
- ( ) Mean squared error
- ( ) None of the above

Explanation: The objective function is the function that is minimized during training to optimize the model's performance. For binary classification problems, binary cross-entropy is commonly used as the objective function, and this is the default in Xgboost.

4. Which of the following is a limitation of Xgboost?

- ( ) It can handle missing data easily
- (●) It requires large amounts of training data
- ( ) It is not scalable to large datasets
- ( ) It cannot handle categorical features

Explanation: Xgboost is a powerful machine learning algorithm that is widely used in industry. However, one limitation of Xgboost is that it can require large amounts of training data to achieve good performance.

5. Which of the following is a parameter in Xgboost that controls the trade-off between overfitting and underfitting?

- ( ) Learning rate
- ( ) Max depth
- (●) Gamma
- ( ) Subsample

Explanation: Gamma is a regularization parameter in Xgboost that controls the minimum reduction in the loss required to make a further partition on a leaf node of the tree. A higher value of gamma leads to a more conservative model that avoids overfitting.

6. Which of the following is a parameter in Xgboost that controls the fraction of observations to be randomly sampled for each tree?

- ◯ Learning rate
- ◯ Max depth
- ◯ Gamma
- ◉ Subsample

Explanation: Subsample is a technique to introduce randomness in the training process and avoid overfitting. It controls the fraction of observations to be randomly sampled for each tree. A value of 1.0 means that all observations are used for each tree, while a value less than 1.0 means that a fraction of the observations is used.

7. Which of the following is a parameter in Xgboost that controls the weights of positive and negative examples in the loss function?

- ◉ Scale positive weight
- ◯ Gamma
- ◯ Min child weight
- ◯ Subsample

Explanation: Scale positive weight is a hyperparameter in Xgboost that controls the weights of positive and negative examples in the loss function for binary classification problems. This parameter can be used to balance the classes when the data is imbalanced.

Submit