**MACHINE LEARNING 7**

1. Which of the following in sk-learn library is used for hyper parameter tuning?

A) GridSearchCV() B) RandomizedCV()

C) K-fold Cross Validation D) All of the above

Answer: - D

2. In which of the below ensemble techniques trees are trained in parallel?

A) Random forest B) Adaboost

C) Gradient Boosting D) All of the above

Answer: - A

3. In machine learning, if in the below line of code:

*sklearn.svm.****SVC*** *(C=1.0, kernel='rbf', degree=3)*

we increasing the C hyper parameter, what will happen?

A) The regularization will increase B) The regularization will decrease

C) No effect on regularization D) kernel will be changed to linear

Answer: - B

4. Check the below line of code and answer the following questions:

*sklearn.tree.****DecisionTreeClassifier****(\*criterion='gini',splitter='best',max\_depth=None, min\_samples\_split=2)*

Which of the following is true regarding max\_depth hyper parameter?

A) It regularizes the decision tree by limiting the maximum depth up to which a tree can be grown.

B) It denotes the number of children a node can have.

C) both A & B

D) None of the above

Answer: - B

5. Which of the following is true regarding Random Forests?

A) It's an ensemble of weak learners.

B) The component trees are trained in series

C) In case of classification problem, the prediction is made by taking mode of the class labels predicted by the component trees.

D)None of the above

Answer: - D

6. What can be the disadvantage if the learning rate is very high in gradient descent?

A) Gradient Descent algorithm can diverge from the optimal solution.

B) Gradient Descent algorithm can keep oscillating around the optimal solution and may not settle.

C) Both of them

D) None of them

Answer: - C

7. As the model complexity increases, what will happen?

A) Bias will increase, Variance decrease B) Bias will decrease, Variance increase

C)both bias and variance increase D) Both bias and variance decrease.

Answer: - B

8. Suppose I have a linear regression model which is performing as follows:

Train accuracy=0.95 and Test accuracy=0.75

Which of the following is true regarding the model?

A) model is underfitting B) model is overfitting

C) model is performing good D) None of the above

Answer: - B

9. Suppose we have a dataset which have two classes A and B. The percentage of class A is 40% and percentage of class B is 60%. Calculate the Gini index and entropy of the dataset.

Answer: - The Gini index is 1 − (4/10) 2 -(6/10)2 = 0.48

10. What are the advantages of Random Forests over Decision Tree?

Answer: - A random forest is simply a collection of decision trees whose results are aggregated into one final result. Their ability to limit overfitting without substantially increasing error due to bias is why they are such powerful models. One way Random Forests reduce variance is by training on different samples of the data.

11. What is the need of scaling all numerical features in a dataset? Name any two techniques used for scaling.

Answer: - scaling is essential for machine learning algorithms that calculate distances between data. If not scale, the feature with a higher value range starts dominating when calculating distances.

Machine learning algorithm just sees number — if there is a vast difference in the range say few ranging in thousands and few ranging in the tens, and it makes the underlying assumption that higher ranging numbers have superiority of some sort. So these more significant number starts playing a more decisive role while training the model.

Scaling techniques - Nominal Scale , Ordinal Scale

12. Write down some advantages which scaling provides in optimization using gradient descent algorithm.

Answer: - We can speed up gradient descent by scaling. This is because θ will descend quickly on small ranges and slowly on large ranges, and so will oscillate inefficiently down to the optimum when the variables are very uneven.

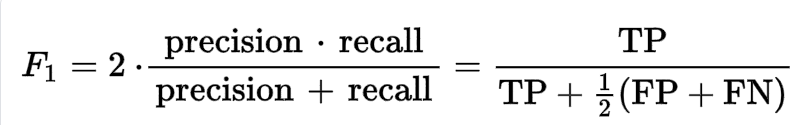
Another reason why feature scaling is applied is that gradient descent converges much faster with feature scaling than without it. It's also important to apply feature scaling if regularization is used as part of the loss function (so that coefficients are penalized appropriately).

13. In case of a highly imbalanced dataset for a classification problem, is accuracy a good metric to measure the performance of the model. If not, why?

Answer: - in the framework of imbalanced data-sets, accuracy is no longer a proper measure, since it does not distinguish between the numbers of correctly classified examples of different classes. Hence, it may lead to erroneous conclusions

14. What is “f-score" metric? Write its mathematical formula.

Answer: - The F-score, also called the F1-score, is a measure of a model’s accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into ‘positive’ or ‘negative’.

f-score formula 

15. What is the difference between fit(), transform() and fit\_transform()?

Answer: - The fit() function calculates the values of parameters. The transform function applies the values of the parameters on the actual data and gives the normalized value. The fit\_transform() function performs both in the same step. Note that the same value is got whether we perform in 2 steps or in a single step.