WORKSHEET 07: MACHINE LEARNING

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

ANS: A) GridSearchCV()

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
ANS: C) Gradient Boosting

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
ANS: B) The regularization will decrease

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
ANS:

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
ANS: D) None of the above

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

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.
ANS: B) Bias will decrease, Variance increase

Q9 to Q15 are subjective answer type questions, Answer them briefly.

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.

10. What are the advantages of Random Forests over Decision Tree?
ANS: Advantages of Random Forests:
1. Powerful and highly accurate
2. No need to normalizing
3. Can handle several features at once
4. Run trees in parallel ways
5. Can perform both regression and classification tasks.
6. Produces good prediction that is easily understandable.

Advantages of Decision Tree:
1. Easy
2. Transparent process
3. Handle both numerical and categorical data
4. Larger the data, the better the result
5. Speed
6. Can generate understandable rules.
7. Has the ability to perform classification without the need for much computation.
8. Gives a clear indication of the most important fields for classification or prediction.

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

ANS: Scaling technique is a method of placing respondents in continuation of gradual change in the pre-assigned values, symbols or numbers based on the features of a particular object as per the defined rules. All the scaling techniques are based on four pillars, i.e., order, description, distance and origin.
Various Machine Learning algorithms are sensitive when the data is not scaled.
Data normalization and data standardization are two techniques used for scaling.

Data Normalization:
Normalization is the method of rescaling data where we try to fit all the data points between the range of 0 to 1 so that the data points can become closer to each other.
It is a very common approach to scaling the data. In this method of scaling the data, the minimum value of any feature gets converted into 0 and the maximum value of the feature gets converted into 1**.**

Data Standardization:
standardization is also required in some forms of machine learning when the input data points are scaled in different scales. Standardization can be a common scale for these data points.

The basic concept behind the standardization function is to make data points centred about the mean of all the data points presented in a feature with a unit standard deviation. This means the mean of the data point will be zero and the standard deviation will be 1.

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

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?
ANS:

14. What is “f-score" metric? Write its mathematical formula.
ANS: The F-score (also known as the F1 score or F-measure) is a metric used to evaluate the performance of a Machine Learning model. It combines precision and recall into a single score. F-measure formula:
F-score = 2 * (precision * recall) / (precision + recall)

15. What is the difference between fit(), transform() and fit_transform()?
ANS: fit()
In the **fit()** method, where we use the required formula and perform the calculation on the feature values of input data and fit this calculation to the transformer. For applying the fit() method (fit transform in python), we have to use **fit()** in frontof the transformer object.

transform()
For changing the data, we probably do transform in the transform() method, where we apply the calculations that we have calculated in fit() to every data point in feature F. We have to use **.transform()** in front of a fit object because we transform the fit calculations.

fit_transform() or fit transform sklearn
The fit_transform() method is basically the combination of the fit method and the transform method. This method simultaneously performs fit and transform operations on the input data and converts the data points.Using fit and transform separately when we need them both decreases the efficiency of the model. Instead, fit_transform() is used to get both works done.

THE END