Machine Learning (3 Marks Qn)

1.

I) In GD and SGD you update set of parameters in an iterative manner

II) In SGD you have to run for all the samples for a single update of parameter in each iteration

III) In GD you either use entire data or a subset of training data update a parameter in each iteration.

A) I is true

2. Which of the following hyper parameter when increased may cause random forest to over fit the data.

(a) no. of trees
(b) depth of tree
(c) learning rate.

A) depth of tree

3. Imagine you are working with analytics vidya and you want to develop a machine learning algorithm which predict no. of views on the article. your analysis is based on features name, author name, no. of article written by same author on the analytics vidya platform in past etc.

which of the following evaluation matrix would you choose in that case.

i) Min square error
ii) Accuracy
iii) SSE

A) minsquare error.

4. let's say that you are using activation funx in hidden layer of neural network at a particular neuron for given input. you get the output as -0.0001. which of the following function should x represent activation

(a) R.e.L.U = Rectified linear unit
(b) tan 8

(c) sigmoid
(d) None of these.

A) tan8

5. Of which of the following are one of the important step to preprocess the text in N.L.P based project.

s1. scemming
s2. stopword removal
s3. object standardizatio

A) All of these.

6. Adding a non important feature to a linear regression method. may result in

(i) increase in R^2
(ii) decrease in R^2

A) increase in R^2

7. in KNN model it is very likely to overfit due to the cause of dimensionality. which of the following option would you consider to handle this problem

(a) dimensality reduction
(b) feature selection.

A (a) & (b) both.

8. which of the following is true about the gradient boosting trees.

(i) in each stage. introduce a new regression tree. to compensate the shortcoming of the existing model.

(ii) we can use gradient distance (GD) method to minimize the loss function

(iii)

A (i) and (ii)

9. To apply bagging to regression trees which of the following are done in that case.

i) we build the n regression with n bootstrap sample.

ii) we take the average of n regression tree

iii) each tree has a high variance with low biased

A) All are true.

10. Is when you find noise in data which of following option would you considered in K.N.N

(i) g will increase the value of K
(ii) g will decrease the val of K
(iii) Noise can't be dependent on the value of K
(iv) None of these

A) (i) g will increase the value of K.

11. suppose you want to predict the class of new data point. $x=1$ and $y=1$ using Euclidean distance in 3NN. in which class these data points belongs to

(a) + class
(b) − class
(c) cannot be predicted.
(d) none of these

A) (a) + class

12. which of the following will be Euclidean distance betn the two data points A(1,3) and B(2,3)

A) 1

13. suppose ~~you~~ you are working on a binary classification problem with three input features and you choose to apply a bagging algorithm X on this data. you choose max_features = 2 and the n_estimators = 23 assume that each estimator has no. of samples. note that algorithm X is aggregating the results of

individual estimators based on maximum voting. what will be the maximum accuracy you can get

A) 100%.

14. in random forest or gradient boosting algorithm features can be of any type for example it can be a continuous features or categorical features. which of the following option is true when you consider these types of feature.

A) both the algorithm can handle real valued attributes by discretizing them

15. Which of the following is true about training and testing error. in the case described below.

suppose you want to apply ~~AdaBoost~~ ADaboost algorithm on data D. which has T observation. you have set half of the data for training and half for testing initially. Now you want to increase the no. of data points for training;

$T_1, T_2, \ldots$ upto T_n.
where $T_1 < T_2 < T_3 < \ldots < T_n$

A: The diff. betn training error and testing error decreases as the no. of observation increases.

16. Suppose you are given 3 variable X, Y, Z. The Pearson correlation coefficient for (X,Y), (Y,Z) & (X,Z) are C_1, C_2 and C_3 respectively. Now you have added (2) two in all values of X, subtract 2 from all values of Y and Z remains the same. The new coefficient (X,Y), (Y,Z) and (X,Z) are given by D_1, D_2, D_3 respectively. How do the values of D_1, D_2, D relate to C_1, C_2, C_3

A: $D_1 = C_1$, $D_2 = C_2$, $D_3 = C_3$