

MACHINE LEARNING

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans: least square error

2. Which of the following statement is true about outliers in linear regression?

Ans: none of these

3. A line falls from left to right if a slope is _____?

Ans: negative

4. Which of the following will have symmetric relation between dependent variable and independent variable?

Ans: correlation



5. Which of the following is the reason for over fitting condition?

Ans: none of these

6. If output involves label then that model is called as

Ans: predictive model

7. Lasso and Ridge regression techniques belong to _____?

Ans: regularization

8. To overcome with imbalance dataset which technique can be used

Ans: SMOTE

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans: TPR and FPR ; sensitivity and specificity

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less

Ans: false

11. Pick the feature extraction from below:

Ans: construction bag of words from email

MACHINE LEARNING

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans: A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

C) We need to iterate.

13. Explain the term regularization?

Ans: The word itself says to make things regular and it is used for reducing the errors by fitting the function correct manner.

We need to avoid overfitting.

14. Which particular algorithms are used for regularization?

Ans: Ridge regression

lasso

Dropout

15. Explain the term error present in linear regression equation?

Ans: suppose you are predicting the weight of a human being by seeing the height of that person then however height is related to weight so taller people tend to be more weight but we could be wrong because there is a range of weights at each height so the error is that difference between the predicted value and the actual value.