MACHINE LEARNING

In Q1 to Q11, only one option is correct, choose the correct option:

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

A) Least Square Error B) Maximum Likelihood

C) Logarithmic Loss D) Both A and B

Ans- A) Least Square Error

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

A) Linear regression is sensitive to outliers B) linear regression is not sensitive to outliers

C) Can't say D) none of these

Ans- A) Linear regression is sensitive to outliers

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

A) Positive B) Negative

C) Zero D) Undefined

Ans- B) Negative

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

A) Regression B) Correlation

C) Both of them D) None of these

Ans- B) Correlation

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

A) High bias and high variance B) Low bias and low variance

C) Low bias and high variance D) none of these

Ans- C) Low bias and high variance

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

A) Descriptive model B) Predictive modal

C) Reinforcement learning D) All of the above

Ans- B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation B) Removing outliers

C) SMOTE D) Regularization

Ans- D) Regularization

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

A) Cross validation B) Regularization

C) Kernel D) SMOTE

Ans- D) SMOTE

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

A) TPR and FPR B) Sensitivity and precision

C) Sensitivity and Specificity D) Recall and precision

Ans- C) Sensitivity and Specificity.

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

A) True B) False

Ans- B) False

11. Pick the feature extraction from below:

A) Construction bag of words from a email

B) Apply PCA to project high dimensional data

C) Removing stop words

D) Forward selection

Ans- B) Apply PCA to project high dimensional data

In Q12, more than one options are correct, choose all the correct options:

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

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.

D) It does not make use of dependent variable.

Ans- A, B and C

ASSIGNMENT – 39

MACHINE LEARNING

Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Ans- Regularization is the process used to reduce the errors by overfitting function.

Lets understand what is overfitting and how regularization works. When a model performs well on the training data and does not perform well on the testing data, then the model is said to have high generalization error. In other words, in such a scenario, the model has low bias and high variance and is too complex. This is called overfitting.

Regularization refers to the modifications that can be made to a learning algorithm that helps to reduce this generalization error and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing the complexity of a model.

14. Which particular algorithms are used for regularization?

Ans- There are basically two algorithms used for regularization.

1-Ridge Regression- Ridge regression is useful in solving problems where you have less than one hundred thousand samples or when you have more parameters than samples. Ridge regression is almost identical to linear regression except we introduce a small amount of bias. In return, we get a significant drop in variance. In other words, by starting with a slightly worse fit, Ridge Regression can provide better long-term predictions.

2-LASSO Regression (Least absolute shrinkage and selection operation)- Lasso Regression is almost identical to Ridge Regression, the only difference being that we take the absolute value as opposed to the squaring the weights when computing the ridge regression penalty. As a result of taking the absolute value, Lasso Regression can shrink the slope all the way down to 0. whereas Ridge Regression can only shrink the slope asymptotically close to 0.

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

Ans- What your model estimates may not match what is actually observed. If your model specification, your decision to use linear regression, is a good one then the difference between your model estimate and what is actually observed is small. This difference is the error. It is used to account for the difference between what is observed and what your model estimates.

In almost all the data mining techniques, we try to find out the closer estimates of a variable we want to predict. In regression models, we assume that the relation between the response variable and predictors to be linear and we find out a straight line that fits the relation well.

The error term in regression is a catch-all for what we miss out with this model, because in reality

-The true relation may not be linear

-There may be other variables not included in the model that cause variation in response variable

-There may be measurement errors in the observations