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

    Answer – 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

    Answer – 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

    Answer – A) Postitive

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

    Answer – 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

    Answer- 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

    Answer – D) All of the above

7. Lasso and Ridge regression techniques belong to _____?
    A) Cross validation B) Removing outliers
    C) SMOTE   D) Regularization

    Answer – D) Regularization

8. To overcome with imbalance dataset which technique can be used?
    A) Cross validation B) Regularization
    C) Kernel   D) SMOTE

    Answer – 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

Answer- A) TPR and FPR

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

A) True   B) False

Answer – 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

Answer - A) Construction bag of words from a email

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.

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

FLIP ROBO   ASSIGNMENT – 39

MACHINE LEARNING

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

13. Explain the term regularization?

Answer -

When we use regression models to train some data, there is a good chance that the model will overfit the given training data set. Regularizationhelps sort this overfitting problem by restricting the degrees of freedom of a given equation i.e. simply reducing the number of degrees of apolynomial function by reducing their corresponding weights. In a linear equation, we do not want huge weights/coefficients as a small change in weight can make a large difference for the dependent variable(Y). So, regularization constraints the weights of such features to avoid overfitting.

To regularize the model, a Shrinkage penalty is added to the cost function. Let's see different types of regularizations in regression:

LASSO
RIDGE
ELASTICNET (Less popular)

14. Which particular algorithms are used for regularization?

Answer -

To regularize the model, a Shrinkage penalty is added to the cost function. Let's see different types of algorithms used for regularizations :

LASSO
RIDGE
ELASTICNET (Less popular)

LASSO(Least Absolute Shrinkage and Selection Operator) Regression (L1 Form)
LASSO regression penalizes the model based on the sum of magnitude of the coefficients. The regularization term is given by
regularization=
Where, λ is the shrinkage factor.

Ridge Regression (L2 Form)
Ridge regression penalizes the model based on the sum of squares of magnitude of the coefficients. The regularization term is given by
regularization=
Where, λ is the shrinkage factor. $\lambda * \Sigma \| \beta_j$

Difference between Ridge and Lasso
Ridge regression shrinks the coefficients for those predictors which contribute very less in the model but have huge weights, very close to zero. Butit never makes them exactly zero. Thus, the final model will still contain all those predictors, though with less weights. This doesn't help ininterpreting the model very well. This is where Lasso regression differs with Ridge regression. In Lasso, the L1 penalty does reduce somecoefficients exactly to zero when we use a sufficiently large tuning parameter λ. So, in addition to regularizing, lasso also performs feature selection.

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

Perfect prediction on the basis of regression equation seems to be difficult. What is needed is a measure that would indicate how precise the prediction of y is based on x. The term error in a regression model is a standard error of estimate. This standard error of estimate has the same concept as the standard deviation.

The standard deviation measures dispersion of the observations about the mean of distribution while the standard error of estimate measures dispersion about the line of regression.

In other words standard error measures the accuracy of estimated value. The smaller the value of standard error, lesser the dispersion around the regression.If the error value is zero it means there is no deviation about the line. That shows that the correlation is perfect.