

MACHINE LEARNING

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 :- 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

Answer :- A) Regression

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 :- C) Reinforcement learning

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 :-A) Cross Validation

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 :- A) True

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: - 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.

Answer: - A) We don't have to choose the learning rate. AND B) It becomes slow when the number of features is very large

ASSIGNMENT – 39 MACHINE LEARNING

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

13. Explain the term regularization?

Answer:- It is one of the most important concepts of machine learning. This technique prevents the model from overfitting by adding extra information to it. Regularization is a form of regression that shrinks the coefficient estimates towards zero. In the regularization technique we reduce the magnitude of the independent variables by keeping the same number of variables. Regularization works by adding a penalty or complexity term or shrinkage term with Residual Sum of Squares (RSS) to the complex model.

14. Which particular algorithms are used for regularization?

Answer:- Algorithms are used for regularization are-

1) Ridge Regression (L2 Norm) - In this technique, the cost function is altered by adding the penalty term (shrinkage term), which multiplies the lambda with the squared weight of each individual feature.

$$\text{Loss} = \sum_{j=1}^m \left(Y_i - W_0 - \sum_{i=1}^n W_i X_{ji} \right)^2 + \lambda \sum_{i=1}^n W_i^2$$

2) Lasso Regression (L1 Norm) – It stands for Least Absolute Shrinkage and Selection Operator. This technique is different from ridge regression as it uses absolute weight values for normalization. λ is again a tuning parameter and behaves in the same as it does when using ridge regression.

$$\text{Loss} = \sum_{j=1}^m \left(Y_i - W_0 - \sum_{i=1}^n W_i X_{ji} \right)^2 + \lambda \sum_{i=1}^n |W_i|$$

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

Answer:- The term error represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable.

Within a linear regression model tracking a stock's price over time, the term error is the difference between the expected price at a particular time and the price that was actually observed. In instances where the price is exactly what was anticipated at a particular time, the price will fall on the trend line and the term error will be zero. Points that do not fall directly on the trend line exhibit the fact that the dependent variable, in this case, the price, is influenced by more than just the independent variable, representing the passage of time. The term error stands for any influence being exerted on the price variable, such as changes in market sentiment. The two data points with the greatest distance from the trend line should be an equal distance from the trend line, representing the largest margin of error. If a model is heteroskedastic, a common problem in interpreting statistical models correctly, it refers to a condition in which the variance of the term error in a regression model varies widely.

-----*****-----