

## 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: D) none of these**

**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: 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: B) Predictive modal**

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

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

**13. Explain the term regularization?**

**Answer:** Regularization is one of the most important principles of machine learning. It is a way to prevent the model from overfitting through adding more facts to it. Sometimes the system gets to know the model performs nicely with the education statistics, however, it does not carry out nicely with the test data. It means the model isn't always able to expect the output while dealing with unseen records through introducing noise in the output, and hence the model is referred to as overfitted. This trouble may be dealt with the help of a regularization technique.

This approach may be utilized in such a way that it's going to allow to preserve all variables or features inside the model with the aid of decreasing the value of the variables. Hence, it keeps accuracy as well as a generalization of the version. It in particular regularizes or reduces the coefficient of features toward zero. In simple words, in regularization method, we reduce the magnitude of the coefficients by using maintaining the same variety of capabilities.

#### **14. Which particular algorithms are used for regularization?**

**Answer:**

- Ridge Regression
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression
- Elastic-Net Regression

#### **Ridge Regression**

Ridge regression is a method for analyzing data that suffer from multi-collinearity.

#### **LASSO Regression**

LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

#### **Elastic-Net Regression**

Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

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

**Answer:** Within a linear regression model tracking a stock's price over time, the error term is the difference between the expected price at a particular time and the price that was actually observed.