

MACHINE LEARNING ASSIGNMENT

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**

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**

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

- A) Positive
- B) Negative
- C) Zero
- D) Undefined

ANS: **B**

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**

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**

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

- A) Descriptive model
- C) Reinforcement learning

B) Predictive model

D) All of the above

ANS: **B**

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation

C) SMOTE

B) Removing outliers

D) Regularization

ANS: **D**

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

A) Cross validation

C) Kernel

B) Regularization

D) SMOTE

ANS: **D**

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

C) Sensitivity and Specificity

B) Sensitivity and precision

D) Recall and precision

ANS: **A**

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**

11. Pick the feature extraction from below:

A) Construction bag of words from an email

B) Apply PCA to project high dimensional data

C) Removing stop words

D) Forward selection

ANS: **B**

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**

SUBJECTIVE TYPE QUESTIONS

13. Explain the term regularization?

ANS: When we use regression models to train some data, there is a good chance that the model will overfit the given training dataset. Regularization helps sort this overfitting problem by restricting the degrees of freedom of a given equation, that is simply reducing the number of degrees of a polynomial function by reducing their corresponding weights/ coefficients.

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. So, Regularization constraints the weight of such features to avoid Overfitting.

14. Which particular algorithms are used for regularization?

ANS: The two common algorithms used for regularization in Machine Learning are:

1) **LASSO**(*Least Absolute Shrinkage & Shrinkage Operator*)
REGRESSION (L1 Form):

LASSO regression penalizes the model based on the sum of magnitude of the coefficients. The regularization term is given by:

$$\text{Regularization} = \lambda * \sum |\beta_j|$$

Where, λ is shrinkage factor.

2) **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:

$$\text{Regularization} = \lambda * |\beta_j|^2$$

Where, λ is shrinkage factor.

ELASTICNET is less popular.

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

ANS: The term “Error”, refers to the difference between the observed (actual) values of the dependent variable and the values predicted by the Linear Regression Model. These differences are also known as residuals.

We know that the general form is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + E$$

Here,

- Y is the dependent variable.
- X_1, X_2, \dots, X_n are the independent variables.
- $\beta_0, \beta_1, \dots, \beta_n$ are the coefficients of the model.
- E represents the error term.

The error term captures the unobserved factors or random variability in the dependent variable that the model does not account for. It represents the discrepancy between the predicted values from the regression equation and the actual observed values. In a well-fitted model, the goal is to minimize the errors.

The sum of squared error (SSE) is often used to assess how well the linear regression model fits the data. Also, Root mean squared error (RMSE) is recommended to be used as the primary metric to interpret the model.