

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

Both A and B are correct answers.

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

Option A states that "Linear regression is sensitive to outliers", which correctly reflects the impact outliers can have on the regression analysis. Thus, option A is the true statement about outliers in linear regression.

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

A) Positive B) Negative

C) Zero D) Undefined

Slope is 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

C option is correct.

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

C option is correct.

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

Option B is correct.

7) Lasso and Ridge regression techniques belong to _____?

- A) Cross validation B) Removing outliers
- C) SMOTE D) Regularization

Option D is correct.

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

- A) Cross validation B) Regularization
- C) Kernel D) SMOTE

Option D is correct answer.

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

Option A is correct.

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

- A) True B) False

The statement is 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

Option A is correct.

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.

Option B is correct.

13) Regularization is a technique used in machine learning and statistics to prevent overfitting and improve the generalization of models. It involves adding a penalty term to the loss function during training, which encourages the model to prefer simpler models (with smaller coefficients or fewer parameters) that are less likely to overfit the training data.

14) Regularization techniques can be applied to

various machine learning algorithms to prevent overfitting and improve generalization. Here are some common algorithms where regularization is typically used:

1)Linear Regression

2)Logistic Regression

3)Support Vector Machine

4)Neural Network

5)Decision Tree

6)Ensemble Method

15)In the context of linear regression, the term "error" typically refers to the difference between the observed values of the dependent variable (target variable) and the values predicted by the linear regression model. These errors are also known as residuals.