

ASSIGNMENT – 39

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

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

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

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

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

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

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

7. Lasso and Ridge regression techniques belong to _____?

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

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

- A) **Cross validation** B) Regularization
C) Kernel 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

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

- A) True 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

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.

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

13. Explain the term regularization?

Ans - Regularized regression is a type of regression where the coefficient estimates are constrained to zero. The magnitude (size) of coefficients, as well as the magnitude of the error term, are penalized. Complex models are discouraged, primarily to avoid overfitting.

14. Which particular algorithms are used for regularization?

Ans - There are three main regularization techniques, namely:

1. Ridge Regression (L2 Norm) – It is a method of estimating the coefficients of multiple-regression models in scenarios where independent variables are highly correlated. It has uses in fields including econometrics, chemistry, and engineering.
2. Lasso (L1 Norm) – It is a type of linear regression that uses shrinkage. Shrinkage is where data values are shrunk towards a central point, like the mean. The lasso procedure encourages simple, sparse models (i.e. models with fewer parameters). This particular type of regression is well-suited for models showing high levels of multicollinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.
3. Dropout - It is a technique where randomly selected neurons are ignored during training. They are “dropped-out” randomly. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass.

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

Ans - An **error term** is a value which represents how observed data differs from actual population data. It can also be a variable which represents how a given statistical model differs from reality. The error term is often written ϵ .

An error term essentially means that the model is not completely accurate and results in differing results during real-world applications. For example, assume there is a linear regression function that takes the following form:

$$Y = \alpha X + \beta p + \epsilon$$

where:

α, β = Constant parameters

X, p = Independent variables

ϵ = Error term

When the actual Y differs from the expected or predicted Y in the model during an empirical test, then the error term does not equal 0, which means there are other factors that influence Y .