MACHINE LEARNING

The answers are highlighted in red

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.

MACHINE LEARNING

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

13. Explain the term regularization?

Answer: Regularization is a technique which helps us to tackle the problem of overfitting (it occurs when the model's training data picks up noise or random fluctuations and learns it as concepts) and rectify it.In regularization the co-efficient of features is basically reduced or regularised towards Zero.

14. Which particular algorithms are used for regularization?

Answer: There are mainly 2 types of regularization:

i L2 regularization (Ridge)

Cost function equation

$$\sum_{i=1}^{M}(y_i-\hat{y}_i)^2=\sum_{i=1}^{M}\left(y_i-\sum_{j=0}^{p}w_j\times x_{ij}\right)^2+\lambda\sum_{j=0}^{p}w_j^2$$

In this technique a penalty term(lambda) is added to the linear regression cost function equation. This lambda regularizes the co-efficient and if the coefficient takes large values, then the optimised function is penalised. As lambda tends to 0 the more the equation will become similar to a linear regression equation

ii L1 regularization (Lasso)

Cost function equation

$$\sum_{i=1}^{M}(y_i-\hat{y}_i)^2=\sum_{i=1}^{M}\left(y_i-\sum_{j=0}^{p}w_j\times x_{ij}\right)^2+\lambda\sum_{j=0}^{p}|w_j|$$

This technique also works like ridge, the only difference is that in the above equation lambda =0, and instead of taking squares of the coefficient, magnitudes are taken into account. Lasso technique not only reduces over fitting but also helps in feature selection.

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

Answer: The equation for linear regression is $y=\beta_0+\beta_1 x+\varepsilon$

Where

β_1= the slope

β_0=the intercept (value of y when x is 0)

ε= the error term

x is the independent variable

y is the dependent variable

the error term will capture the errors in a measurement of y and the effect on y of any variable missing from the equation that would contribute to explaining variations in y