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.

Ans. 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

Ans. Linea regression is sensitive to outliers.

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

A) Positive   B) Negative

C) Zero   D) Undefined

Ans. 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

Ans. 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

Ans. 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

Ans. Predictive modal

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

Ans. 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

Ans. 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

Ans. 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

Ans. 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.

Ans. 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?

Ans. Regularization is one of the basic and most important concept in the world of Machine Learning. Regularization is a technique used to reduce by the errors by fitting the function appropriately on the given training set and avoid overfitting.

The commonly used regularization techniques are :-

1. L1 regularization

2. L2 regularization

3. Dropout regularization

1. <u>L1 Regularization</u> technique is called LASSO (Least Absolute Shrinkage and Selection Operator) regression.

2. <u>L2 regularization</u> technique is called Ridge regression**.**

Lasso Regularization Techniques

There are two main regularization techniques, namely Ridge Regression and Lasso Regression. They both differ in the way they assign a penalty to the coefficients. In this blog, we will try to understand more about Lasso Regularization technique.

14. Which particular algorithms are used for regularization?

Ans. There are three main regularization techniques, namely :- LASSO (L1 Norm) , Ridge Regression (L2 Norm) & Dropout.

Lasso Regression (L1 Regularization)

This technique is different form ridge regression as it uses absolute weigh values for normalization. Λ is again a tuning parameter and behaves in the same as it does when using ridge regression.

As loss function only considers absolute weights, optimization algorithms penalize higher weight values. In ridge regression, loss function along with the optimization algorithm brings parameters near to zero but not actually zero, while lasso eliminates less important features and sets respective weight values to zero. Thus, lasso also performs feature selection along with regularization.

Ridge Regression (L2 Regularization)

When you use $\lambda = 0$, it returns the residual sum of square as loss function which you chose initially. For a very high value of λ, loss will ignore vore loss function and minimize weight's square and will end up taking the parameters' value as zero.

Now the parameters are learned using a modifies loss function. To minimize the above function, parameters need to be as small as possible. Thus, L2 norm prevents weights form rising too high.

Dropout

Dropout is a regularization technique used in neural networks. It prevents complex co- adaptations form other neurons.

In neural nets, fully connected layers are more prone to overfit on training data. Using dropout, you can drop connections with 1-p probability for each of the specified layers. Where p is called keep probability parameter and which needs to be tuned.

Along with Dropout, neural networks can be regularized also using L1 and L2 norms.

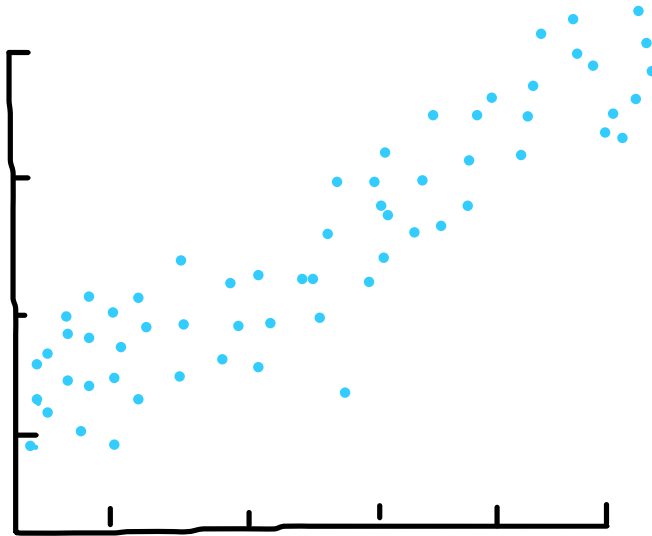**15. Explain the term error present in linear regression equation?**

Ans. In a linear regression model tracking a stock's price over time, the error term is the difference between the expected prove at a particular time and the price that was actually observed. The error term stand for any influence being exerted on the price variable, such as changes in market sentiment.

Error is the difference between the actual value and Predicted value and the goal is to reduce this difference.

Why is this important?

A regression line always has an error term because, in real life, independent variables are never perfect predictors of the dependent variables. Rather the line is an estimate based on the available data.
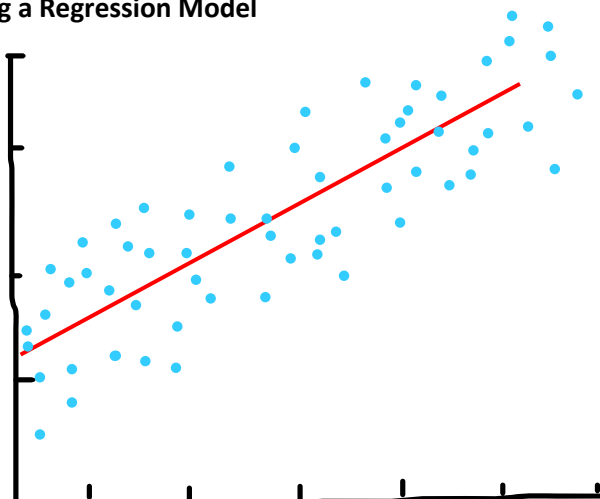
Relationship between these two variables:-

The y-axis is the amount of sales (the dependent variable, the thing you're interested in, is always on the y-axis) and the x-axis is the total rainfall. Each blue dot represents one month's data how much it rained that month and how many sales you made that same month.

Now imagine drawing a line through the chart above, one that runs roughly thorugh the middle of all the data point. This line will help you answer, with some degree of certainty, how much you typically sell when it rains a certain amount.

Building a Regression Model

This is called the regression line and it's drawn to show the line that best fits the data. "the red line is the best explanation of the relationship between the independent variable and dependent variable."