MACHINE LEARNING

Q1 -> Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans-> Least Square Error

Q2 -> Which of the following statement is true about outliers in linear regression?

Ans-> Linear regression is sensitive to outliers

Q3-> A line falls from left to right if a slope is _____?

Ans-> Negative

Q4-> Which of the following will have symmetric relation between dependent variable and independent variable?

Ans-> Correlation

Q5-> Which of the following is the reason for over fitting condition?

Ans-> Low bias and high variance

Q6-> If output involves label then that model is called as:

Ans-> Predictive modal

Q7-> Lasso and Ridge regression techniques belong to _____?

Ans-> Regularization

Q8-> To overcome with imbalance dataset which technique can be used?

Ans->SMOTE

Q9-> The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans-> TPR and FPR

Q10-> In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

Ans->True

Q11-> Pick the feature extraction from below:

Ans->Apply PCA to project high dimensional data

Q12-> Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans-> 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

Q13->. Explain the term regularization?

Ans-> Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting this

Is a form of regression,that regulize the cofficient estimate zero ,this technique discourage learning a more complax or flexible model ,so as to avoid the risk of overfitting ,

Regularixation , significantly reduce variance of the model , without substantial increase in bias So the tuning parameter λ, used in the regularization techniques described controls the impact on bias and variance. As the value of λ rises, it reduces the value of coefficients and thu reducing the variance. ***Till a point, this increase in λ is beneficial as it is only reducing the variance(hence avoiding overfitting), without loosing any important properties in the data.*** But after certain value, the model starts loosing important properties, giving rise to bias in the model and thus underfitting.

Q14-> Which particular algorithms are used for regularization ?

Ans->there are two main algorithms used in regularization is ridge regression and lasso regression or we can L2 and L1 regularization and one more algorithm called elasticnet(less popular)

Lasso --> (least absolute shrinking and selection operator) regression L1 form

Lasso regression penalize the model based on the sum of magnitude of the coefficients the regularization term is give by where λ * Σ !B1!

Where λ is the shrinkage factor

Ridge--> (L2 form) ridge regression penalizes the model based on the sum of squares of magnitude of the cofficients the regularization term is given by regularization = λ * Σİβ2İ

Where λ is the shrinkage factor

Q15-> Explain the term error present in linear regression equation?

Ans-> 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. So the error term tells you how certain you can be about the formula