Machine Learning Linear Regression Questions & Answers

1. What is Linear Regression?  
   Ans. It is a statistical method used for predictive analysis that estimates the relationship between independent & dependent variable.
2. How can we calculate error in Linear Regression?  
   Ans. Linear Regression uses mean-square error (MSE) method to calculate the error of the model.
3. What is error in Linear Regression?  
   Ans. The difference between the true value and the measured value.
4. What is difference between Loss & Cost function?  
   Ans. the loss function is to capture the difference between the actual and predicted values for a single record whereas cost functions aggregate the difference for the entire training dataset.
5. What is the difference between MSE, MAE & RMSE?  
   Ans. MSE is highly biased for higher values. MAE is less biased for higher values. It may not adequately reflect the performance when dealing with large error values.
6. How Gradient descent works in Linear Regression?  
   Ans. Gradient descent (GD) is an iterative first-order optimisation algorithm used to find a local minimum/maximum of a given function. This method is commonly used in machine learning (ML) and deep learning (DL) to minimise a cost/loss.
7. Explain what the Intercept term is?  
   Ans. The intercept (often labelled as constant) is the point where the function crosses the y-axis.
8. Write all the assumption for Linear Regression?  
   Ans. Linear relationship: There exists a linear relationship between the independent variable, x, and the dependent variable, y.   
   Independence: The residuals are independent. There is no correlation between consecutive residuals in time series data.  
   Homoscedasticity: The residuals have constant variance at every level of x.  
   Normality: The residuals of the model are normally distributed.
9. How is hypothesis testing used in Linear Regression?  
   Ans. While training the model we are given: x input training data (univariate – one input variable(parameter)) y: labels to data (supervised learning) When training the model – it fits the best line to predict the value of y for a given value of x
10. How would you decide the importance of Variable for the multivariate regression?  
    Ans.