Q13. Explain the term Regularization?

Ans.- **Regularization**:- When we use regression models to train some data, there is a good chance that the model will overfit the given training dataset.

Regularization helps sort overfitting problem by restricting the degrees of freedom of a given equation i.e. simply reducing the no. of degrees of a polynomial function by reducing their corresponding weights.

In a linear equation, we do not want huge weights/coefficients as a small change in weight can make a large difference for the dependent variable(Y). So, regularization constraints the weights of such features to avoid overfitting. To regularize the model, a shrinkage penalty is added to the cost function.

Q14. Which particular algorithms are used for regularization?

Ans.- There are three main regularization techniques, namely:

1. LASSO Regression (L1 norm)
2. RIDGE Regression (L2 norm)
3. ELASTIC NET Regression (less popular)

1. **<u>Lasso (Least Absolute Shrinkage $ Selection)</u>** :- It penalizes the model based on the sum of magnitude of the coefficients . The regularization term is given by

   Regularization = $\lambda * \Sigma |\beta j|$, Where λ is the shrinkage factor

2. **Ridge Regression** :- It penalizes the model based on the sum of squares of magnitude of the coefficients. The regularization term is given by

   Regularization = $\lambda * \Sigma |\beta j|^2$, Where λ is the shrinkage factor

3. **Elastic-Net Regression**:- It is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

Q15. Explain the term error present in Linear Regression equation?

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

The distance between each point and the linear graph (shown as black arrows on the above graph) is our error term. So we can write our function as $R^B=\beta_0 + \beta_1 E^x + \varepsilon$, where $\beta 0$ and $\beta 1$ are constants and ε is an (non constant) error term.