Q.1 = Least Square Error

Q.2 = Linear regression is sensitive to outliers

Q.3 = Positive

Q.4 = Correlation

Q.5 = High Variance and Low Bias

Q.6 = Predictive Model

Q.7 = Regularization

Q.8 = Cross Validation

Q.9 = TPR and FPR

Q.10 = False

Q.11 = Construction bag-of-words from a email

Q.12 = A - We don't have to choose the learning rate.

B – It becomes slow when number of features is very large.

Q.13 = When we use regression models to train some data, there is a good chance that the model will overfit the given training data set. Regularization helps to sort the overfitting problem by restricting the degrees of freedom of a given equation i.e. simply reducing the number 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.

Q.14 = There are 3 algorithms used for Regularization:

I. LASSO

II. RIDGE

III. ELASTICNET (LESS POPULAR)

Q.15 = We can say the error term as Residual. The residual is the difference between the actual Y and predicted Y.