**Assumptions:**

**KNN:** class of a neighbour is similar to it’s class.

**Naïve-Bayes**: Features are conditionally independent.

**Logistic Regression:** Dataset is linearly seperable (If not do it via feature eng)

**Correlation matrix vs Covariance matrix:**

If variables are on similar scale then use covariance matrix,

If variable are on different scale then use correlation matrix.

**Multicollinearity test:**

Perbutation test can be done to test multicollinearity within features. Refer Notes 12 of Naïve bayes for description.

**Error Metrics:** which error metrics to use when?

**Classification:**

1. Avg Accuracy
2. Log loss
3. AUC

**Regression:**

1. RMSE (Root Mean Squared Error)

**Ranking:**

1. Precission and Recall