**Logistic Regression – Classification Model**

**Problem Statement** – Telecom Churn Case Study

With 21 predictor variables needed to predict whether a particular customer will switch to another telecom provider or not. In telecom terminology, this is referred to as churning and not churning, respectively.

**Building Model steps** -

1. Data cleaning and preparation
   1. Combining three data frames
   2. Handling categorical variables
      1. Mapping categorical variables to integers
      2. Dummy variable creation
   3. Handling missing values
2. Splitting the Data into Training and Testing Sets
3. Model Building
   1. Feature elimination based on correlations
   2. Feature Selection Using RFE (Recursive Feature Elimination) coarse tuning
   3. Manual feature elimination (using p-values and VIFs) Fine Tuning
4. Model Evaluation
   1. Accuracy
   2. Sensitivity and Specificity (Accuracy is not enough)
   3. Optimal cut-off using ROC curve
   4. Precision and Recall (another evaluation method)
5. Making predictions on the test set