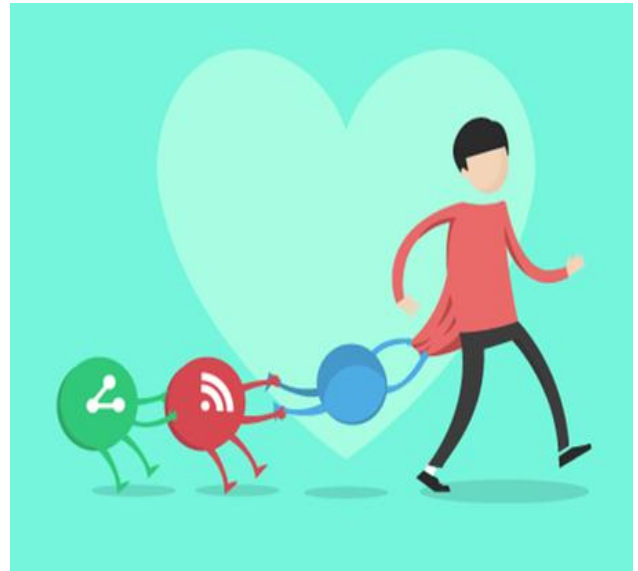# Churn Analysis In Telecom Industry

Problem Statement

Customer churn occurs when customers or subscribers stop doing business with a company or service, also known as customer attrition. It is also referred as loss of clients or customers. One industry in which churn rates are particularly useful is the telecommunications industry, because most customers have multiple options from which to choose within a geographic location.

Build classification models using decision tree algorithm to predict whether the customer be churned or not on the basis of its billing information and customer demographics.

Data Description:

- gender: This column describes about customers gender details.Value of this column is "Male" or "Female"

- SeniorCitizen: This column describes whether a customer is a senior citizen or not. Value of this column is "Yes" or "No"

- Partner: This column describes whether a customer has a partner or not. Value of this column is "Yes" or "No"

- Dependents: It describes whether a customer dependents or not. Value of this column is "Yes" or "No"

- Tenure: This column describes how many months a customer got a telephone connection. Value of this column is integer

- Call Service: This column describes whether a customer has a phone service with their connection or not. Value of this column is "Yes" or "No"

- Multiple Connections: This column describes whether a customer has single or multiple connection. Value of this column is "Yes" or "No"

- Internet Connection: This column describes whether a customer has an internet connection with their telephone connection or not.Value of this column is "Yes" or "No"

- Online Security: This column describes whether a customer has online security for internet connection or not. Value of this column is "Yes" or "No"

- Online Backup: Some customer wants to save their data in online storage. This column describes whether a customer has an online backup service or no. Value of this column is "Yes" or "No"

- Device Protection Service: Customer wants to protect their device from virus and other attack. So they might have device protection service which is got with telephone service. So this column describes whether a customer has device protection service or not.Value of this column is "Yes" or "No"

- Technical Help: If a customer wants technical help, they must have registered with their connection. So this column describes whether a customer has technical help service or not. Value of this column is "Yes" or "No"
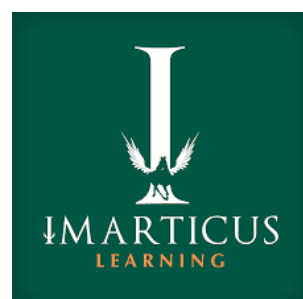
- Online TV: Some customer might watch a TV program online. This column describes whether a customer is subscribed with online TV service or not. Value of this column is "Yes" or "No"

- Online Movies: Some customer might watch online movies. This column describes whether a customer is subscribed with online movies service or not. Value of this column is "Yes" or "No"

- Agreement: A customer might have a month to month agreement or one-year connection or two-year connection and so on. This column describes a customer agreement type

- Value of this column is "Month-to-Month", "One year" and " Two years".
- Billing Method: This column describes whether a customer wants paperless billing or not. Value of this column is "Yes" or "No"

- Payment Method: This column describes in which payment way a customer will pay like Bank transfer(automatic), Credit card(automatic), Electronic check and Mailed check

- Value of this column is Bank transfer(automatic), Credit card(automatic), Electronic check and Mailed check.Monthly 18 - Service Charge: This column describes how much will be charged from a customer for their connection. Value of this column is numeric

- Total Amount: This column describes how much a customer has to pay totally for their connection. Value of this column is numeric

- Churn: This column describes a customer is churn or not. Value of this column is "Yes" or "No"

Evaluation Parameters

Evaluation will be based on:

- Model Comparison
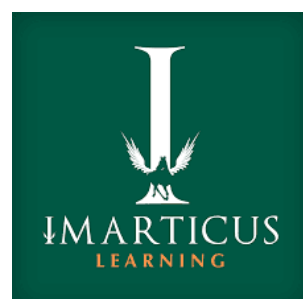
- Model Selection

- **Model Comparison**

Apply Decision tree algorithms using following parameters and compare results.

- Maxdepth

- Mindepth

- Minsplit

- Minbucket

- Splitting criterion

- Method

Get the result from the model and feed the model to pruning function if the result of the model leads to overfitting. When do pruning the tree, the complexity parameter value will be passed to pruning function. Still if the model gives low accuracy, change the above parameter value when build the model second time. Get the result from the second model and compare the result with the result of the first model.

- **Model Selection**

  Select the best model. Model selection to be based on Accuracy, Sensitivity, Specificity, Positive predictive value, Negative predictive value, Prevalence, Detection rate, Detection prevalence and balanced accuracy.

Expected Outcome:

Higher accuracy in predicting the outcome using test data.