# DATA ANALYTICS WITH COGNOS – GROUP 3

## CUSTOMER CHURN PREDICTION

### PROJECT DESCRIPTION:

The "Customer Churn Prediction" project is a model which can predict when a customer, user, subscriber or any kind of client stops or ends their relationship with the particular organization. We explore the dataset which includes all the parameters of the customers including the customers who had churned. The key goal of this project is to preprocess the given data, visualize the data, apply the model and check the accuracy based on it.

### ALGORITHM – RANDOM FOREST:

Random forest is an ensemble machine learning algorithm used for both classification and regression tasks. Its methodology is to combine the decision trees during training and combine their prediction during testing to improve accuracy. The reason why we choose random forest out of all other algorithm due to its ability in handling high dimensional data effectively, being robust to outliers, versatility, feature importance, reduction of variance, efficient on parallel processing and prevents overfitting.

### METHODOLOGY:

### STEP1: IMPORTING THE DATSET

Importing the provided dataset provided through pandas into our model.

### STEP2: EXPLORATORY DATA ANALYSIS

Visualizing the data to gain insights about the data such as trends, to understand how variables are distributed and noise.

**STEP3: PREPROCESSING THE DATA**

Preprocessing the data to assure that they are normally distributed and that they are on the same scale. It also assures the data type of the features is numeric.

**STEP4: TRAINING AND TESTING SET**

Splitting the dataset into training and testing set in order to train the model and test the model on the testing data.

**STEP5: IMPLEMENTING THE MODEL**

Importing the random forest classifier and training the training set using the model.

**STEP6: PREDICTION**

Predicting the test set using the classifier model.

**STEP7: ACCURACY**

In order to evaluate the performance of the model the accuracy of the model is calculated.

**CONCLUSION:**

This project equips Random Forest Machine Learning Algorithm to predict the churning of customer based on some particular parameters from the dataset. By implementing the above methodology we can obtain a prediction level with a considerably acceptable accuracy level. Further the implementation of the above project will be executed.