We have used **Python** to estimate the performance of models on the provided data 'Churn-train.csv'. Now that you have an idea about the project, we can look at the implementation of various models through the code and look at the outputs of the same.

This folder contains the train data (churn-train.csv) as well as the holdout data (churn-holdout.csv) at the path '\Datasets'.

To understand and compare the performance of data through various models (Forest Random Classifier (RFC), KNN Classifier, Generalised Liner Model (GLM) and Support Vector Machines (SVM)) that are used in the project, we have combined the code in a single file 'modelAnalysis.py'. This file has step by step commented code to implement any or all of the four chosen models.

The output of the 'modelAnalysis.py'' file is a graph which is stored at the path '\Results\TestResults\comparisonGraph'. The graph contains AUC score and ROC curve for all four models, thus, making it easy to compare and decide on the final approach.

As the graph depicts that RFC has the best performance for the given data, we have used RFC model to predict the unseen data '\Datasets\churn-holdout.csv'. The code to predict the holdout data is in the file 'main.py' which is well commented to provide step by step process of generating final predictions for the data.

The final result is a csv file containing the identifier (Phone Number), probability of each class and the prediction based on the probability. This output is stored at the path '\Results\FinalPredictions\ 118221161_118221388_118220570.csv'.