**Predictive Analytics**

**Data set Description**

Refer to tab “LikelytoChurn” in the attached excel file. Like any other industry, A common phenomenon in Fund services business is that the Advisors change their service providers for better service, which is synonymous to classic customer churn data science problem.

The data set consists of all Advisors (Identified by unique identity “ID”) that changed a service provider and all Advisors that did not change a service provider indicated by column “Switched” as “Yes” or “No”

respectively. There are various factors that describe the nature of the advisors business and are indicated as “Factor#1” to #Factor#38”. The purpose of Predictive analytics exercise has two objectives.

**QUESTIONS:**

1) Determine which Factors influence the outcome of changing (“switching”) a service

provider the most and describe why using various statistical exploratory analysis

techniques. In the interest of time, You can limit your description to Top3 factors that

influence the most.

2) Build a Classification predictive model to predict the likelihood of an Advisor changing

(“Switching”) a service provider. You could choose to split the data into 70% Train and

30% Test data sets with the aim of training the model with “Train” data set and testing

the model on the “Test” data set. The final result should include “Confusion matrix” on

the “Test” data set and any model validation metrics that you chose to include. Write a

report on your approach and the summary of your observations. The aim of this exercise

is to convince a business user of the predictive value and take an appropriate action to

avoid losing business from Advisors