**Problem Description**

A financial institute is planning to roll out a stock market trading facility for their existing account holders. The service costs significant amount of capital for the institute in terms of infra, licensing and people cost. To make the service offering profitable, they charge a percentage of base commission on every transaction being made. However, this is not a service offered only by them, many of their other competitors are offering the same service and quite a few at a lower commission rates. To retain or attract people who trade heavily in stock market and in turn generate a good commission for institute, they are planning to offer discounts as they roll out the service to entire customer base.

Problem is, that this discount, hampers profits coming from the customers who do not trade in large quantities. To tackle this issue, the institute wants to offer discounts selectively to customers who trade more often than the rest. To be able to do so, they need to know which of their customers are going to be heavy traders or money makers for them.

To be able to do this, they decided to do a beta run of their service to a small chunk of their customer base. For these customers, they have manually divided them into categories i.e. category 1 and category 2. Category 1 customers are the ones who are heavy traders and are money makers for the institute. The institute wishes to offer discounts to this group. Category 2 customers are the ones who need to be kept away from the discount offers.

The objective of this project is to build a predictive model, which given information regarding the customers and their trade behavior can identify customers to whom the institute can extend discount offers.

**Dataset**

The dataset provided comes from the beta run conducted by the institute. The data has information for ~11000 customers with 36 features comprising of personal, financial and portfolio details.

**Project Flow**

* Problem definition
* Data description
* Data Preprocessing pipeline
* Model building
* Model Evaluation
* Hyper Parameter Tuning
* Experimentation with other complex models
* Hyper Parameter tuning and Evaluation
* Final conclusion