

# A Generalized and Automated Data Mining Approach for Customer Acquisition

---

**Introduction:** In today's world banking sector is growing rapidly. Because of lot of choices, customers are the kings for banks. Banks has only few profitable customers to source while many banks competing. This process of CA (Customer Acquisition) works in both ways, CA for already registered customers with banks and identifying new potential customers. Bank faces several losses while existing customers decide to leave and this might affects the further market campaigning or might cause reputation loss and future customer relationships with bank. In this model we are looking for the solution for above two problems. This model can be applied to any kind of customer acquisition of bank as long as sufficient data is available.

**Objective:** Based on available customer data, this model has 3 main objectives, 1<sup>st</sup> collect all the available data related to customers who switched the other banks and based on that train a model and predict future churning and finally predicting total loss, 2<sup>nd</sup> collect data related to customers who switch to our bank and train a model and predict total profit. 1<sup>st</sup> objective will give the information related to defective current market strategies and 2<sup>nd</sup> objective will reveal effective current market strategies bank possess. 3<sup>rd</sup> objective is to use collective data of above two and predict new market strategies and future potential customers. After all done, predict final PNL report for future strategy.

## Step Involved:

1. Data of all the customers who switched the bank will be collected, data preprocessing will be done to transform data into standard format. Data preprocessing may include but not limited to- handling missing data, data reduction, modification of volatile data, prioritization of data etc. (This is needed for automation, random data to standard one)
2. After getting standard data, about 50plus of machine learning algorithms will be used to train 50 models, among half of them will be used for classification model (these half will have better performance, accuracy and coverage etc., selection criteria can be fixed for automation and can be varied to get best performance).
3. Classification model will be used on the customer data base (data of old customers) to predict the future churn and will generate PNL report for future strategies. Same thing will be done for new coming customers and this will help to track down the productive and effective strategies.
4. Now with combined data of above two, we will use same approach to find new potential customers and effective regions for market campaigning and final PNL report will be forecasted.

## API Used:

1. From bank API, All customers and account API, + other API will be used for different-different strategies. Use of other API may vary with respect to type of customer acquisition. ML algorithms will be used extensively.
2. Based on type of data bank has (oracle or object oriented), SQL and python technologies will be used. And R-language will also be used. All the work will be automated thought python scripts.

## Benefits:

1. Use of multiple ML algorithms increases the accuracy and effectiveness of the model. Since its main frame uses python (can be implemented in bank's personal platform) which is easier to modify and effective. It has less security risk since it resolve around the data and don't use complicated open source library. It can complete work of many analytics single handedly as long as data is available. It can also be used in collection and compliance since its core idea resolve around data. It has ease of deployment as well as low maintenance cost (since standard data pattern might go under changes).
2. It not only come up with effective solution but can also tell the future trends of current strategy. For future strategies, it can predict cost and final PNL report which is hugely advantageous (Important while making decisions for future strategies).
3. For further information, please reach out to my own project  
<https://drive.google.com/open?id=0B7FkHAXjadg1NIVsVndSYXNEMVE> and case study on forecasting  
<https://drive.google.com/open?id=0B7FkHAXjadg1QkJZSkp1OHpGWGM> .

## Personal Details

**Name:** Maneesh Kumar Meena

**Contact No:** 7752894592

**Email ID:** [maneeshmeena1997@gmail.com](mailto:maneeshmeena1997@gmail.com)

**City:** Mumbai, Maharashtra