Business Analytics: Homework 3

## Chengjun Liu, 7116099552

* Understand situation and need: We try to distinguish influence on delinquent credit accounts with different attributes paying debt by 13 different actions and find the suitable action for each delinquent credit accounts.
* Data
  + The data are collected from internal resources and credit bureau records. We only select accounts which are delinquent and taken 1 of 13 actions.
* General approach: We try to roll and slice customer records by months and select suitable records. We divide records according to 13 actions. For each action, we build individual statistical models to judge whether delinquent accounts with different attributes pay debt under one specific action. We input attributes of delinquent accounts into different action models to decide which action to take.
* Procedures
  + Select influential attributes
    - We select and focus on several main characteristics of customers. Here we suggest purchase/payment frequency, single/total/average payment amount, Value Growth, Gender, Billing, Age, Channel, Region, Race, Browsing Time. If possible, we can add more variables.
  + Find influential variables
    - Let variables in previous time period (like 6 months) be the influential variables and “current” customer values be the influenced result we want to learn. By comparing whether different factors make different value distributions in each variable, we can select the variables that indeed have influence. We can select those variables as the characteristics of our best customers.
  + Find more high-value customers
    - We match the characteristics of best customers with other current customers. It’s better for us to build a model to convert characteristics into predicted values. If previous characteristics make high values, we assume they are potential high-value customers.
    - We gather variables and customer values of different customers and try to build a proper statistical model. We can slice the time period of each customer into multiple samples to enlarge sample base.
    - For statistical model, we suggest neuron-network statistical model. After training the statistical model with high-value customers we observed and testing its reliability, we can input the characteristics of current customers and sort predicted customer values to find potential best customers.
    - We can observe the leverage of variables in the statistical model. If new customers and registered visitors match the characteristics that have a bigger leverage and score higher predicted values, we can target them as high-value customers.