Final Project

Total points: 100

Please submit

1. R code
2. FutureFundraising.csv with the predicted values
3. Final project report -- The project template is provided in a separate file.

Direct-Mail Fundraising

Two data sets are provided—Fundraising.csv and FutureFundraising.csv

Please answer the following questions. There are two target variables, i.e., TARGET\_B (binary indicator) for response and TARGET\_D (donation amount). Please recommend one model for TARGET\_B and one model for TARGET\_D respectively

1. TARGET\_B

Please build a decision tree model to predict TARGET\_B using Fundraising.csv. Various data mining techniques can be used to mine the data. No on technique is universally better than another. You are also encouraged (optional) to build a neural network model to make prediction, and then compare it with the decision tree model. If you are interested in using neural network, you probably need to choose the input variables according to the decision tree model.

1. Please fully understand the data and variables by checking their values, types, distributions, etc.
2. Please create training and validation sets.
3. Please build the model
4. Please assess each solution’s performance (confusion matrix and gain chart)
5. Please use the model to make prediction on FutureFundraising.csv
6. TARGET\_D

Please using regression model to build a model to predict TARGET\_D for donor (TARGET\_B=1). Use Fundraising.csv file to build and test the model.

1. Please also create training and validation sets.
2. Please comment on the model performance.
3. Please use the model to make prediction on FutureFundraising.csv

(hint: some functions such as select observations and variable are available at <https://www.statmethods.net/management/subset.html>)