TARGETED MARKETING COMPAIN OBJECTIVES

EXECUTIVE SUMMARY

Telephone marketing campaigns can provide a means to reach customers and inform them of various services offered by a company. The financial services industry, in particular the retail banking industry, is ripe with opportunities to take advantage of this medium if the ROI (return on investment) is attractive for the institution. One way to increase the ROI for a given campaign is to specifically target the customers that have a higher likelihood of participation in a given program, rather than a simple random dialing of its customer base.

RESEARCH DESIGN

For this targeted marketing campaign, we can draw upon the databank of previously executed campaigns which track detailed information about the client demographics, such as age, job type, marital status and education. Along with basic demographics, we also have available to us contact times, call durations, previous participation, and financial information about the customer like average account balance and outstanding loans to the institution.

With this dataset, we can quickly ascertain that around 11.5% of all customers are interested in these kinds of term deposit offers. Looking at age, balance and call duration distributions of customers who are not interested in such offers and those who are, we see that customers who have an average account balance between 5,000 and 10,000 have a 15% chance of being interested in such a term deposit, a 75% increase in hit-rate.

There are two other economic indicators available to us in the previous campaigns dataset, the binary variables which indicate if the client has a loan with the bank and if the customer has a house. Using these economic indicators will provide our model with the data it needs to properly determine the likelihood of customer participation.

TECHNICAL OVERVIEW

The analysis of the previous campaign data was performed using traditional exploratory data analysis techniques in Python, a general-purpose programming language that is exceptional at performing exploratory data analysis and data modeling. Through exploratory data analysis of the previous campaigns and their outcomes, there were several interesting observations.

Given the relatively low overall interest in term deposits, the first logical step is look for patterns of interest in target demographics. The average customer who was interested in term deposits is relatively identical to one who was not, with the only noticeable metrics that could provide us with leads on narrowing down features of interest were average balance, 1,403 for those not interested compared to 1,572 to those who were. This would lead us to believe that the largest factors for a client being interested in such a program would be their economic metrics: balance, housing and outstanding loan.

Since we have the results of previously executed campaigns, we can use these features of a client and build two separate supervise learning models for classification. For this task we chose to use Naïve Bayes and Logistic Regression binary classifiers as they are robust and well suited for the task at hand.

The technical implementation will consist of two identical data feature configurations. The only difference here will be the classifier used, Naïve Bayes and Logistic Regression, respectively. We will measure the performance of the models using an industry standard performance metric, the ROC curve, or Receiver Operating Characteristic curve. This metric gives us a sense for how accurate our predictions will be relative to the TPR (true positive rate), as if we dial a customer on a false-positive, the worst possible outcome will be that they reject the offer (although, they might not).

CONCLUSION

It is our recommendation to management that we proceed with a rollout of the Logistic Regression version of the model. Under performance metrics scrutiny, the Logistic Regression model out-performed the Naïve Bayes 61% compared to the 64% performance of the Logistic Regression under cross-validated stress testing.

It is worth noting that by implementing the model to make suggestions on which customers to call you will be gaining a significant ROI on your campaign, as noted earlier there is a relatively small acceptance rate at 11.5%; the improved suggestions on who to call will both narrow the scope of the campaign and produce a higher call to deal ratio, yielding higher revenues per campaign and lowering the operational expenses of executing one.

For further information and details on this study was conducted, please visit the Colabratory Notebook.