Upgrade with Uplift: Gaining More Accurate Predictive Models with Uplift Modeling

Josh Chan, Katarina McGuckin, Norah Alwalan, Andrew Moolenaar, Daniel Yang, Matthew A. Lanham

Purdue University Krannert School of Management

chan151@purdue.edu; kmcgucki@purdue.edu; nalwalan@purdue.edu; amoolena@purdue.edu; yang817@purdue.edu; lanhamm@purdue.edu

Abstract

This study focuses on understanding how marketers can leverage uplift models to better predict consumer behaviors compared to traditional response models. Our focus is to find the best uplift modeling method based on specific characteristics for each dataset. The motivation for this study is to apply uplift modeling techniques to any business problem. Uplift modeling has been shown to outperform conventional methods in terms of identifying specific consumer behaviors. We use various techniques such as: reducing the number of features, creating branches for decision trees, and ending with a variety of models. After each model performs, we compare the performance by discussing the uplift score, implementation difficulty, and time complexity.

Introduction

Standard Predictive Modeling fails because it predictively scores for expected responses without accounting for the drivers of that response. If we look at the nature of purchasing behavior, we would see that there are four main categories a customer could fall into, as shown in Figure 1. The group that function as a good target are highlighted in green, and the groups that function as bad targets are highlighted in red.

Buy if DO receive a treatment	Buy if DON'T receive a treatment: Yes	Buy if DON'T receive a treatment: No
No	Do-Not-Disturb	Lost Causes
Yes	Sure Things	Persuadable

Figure 1: Four types of marketing targets.
Source: Siegel (2011) and Radcliffe (2007b).

The standard predictive model fails to capture the negative effects of giving treatment to the bad targets. The standard model also attributes the positive results to marketing efforts without distinguishing the responses from "Sure Things" or "Persuadables". The uplift modeling technique, which identifies positive and negative lift, is capable of capturing the difference from these treatment groups in a way that the standard predictive model simply does not. We aim to showcase how the uplift model differentiates between customer groups and informs researchers to make better decisions.

Literature Review

The majority of the previous studies used a variety of methods to have an accurate model for uplift modeling. Certain studies used more advanced techniques to have an in-depth analysis on the data. We believe each method is complementary to the respective dataset so we used the uplift package in R Studio to illustrate the efficiency of using uplift techniques in comparison to traditional methods.

Study	Logistic Regression	Decision Tree	Two Model Approach	Treatment Dummy Approach	Probabilistic w/o weighting	Probablistic w weighting	Churn Uplift Modeling	Random Forest	Neural Network	Forward Slection	Backward Selection
(Siegel 2011)		✓					✓				
(Kane, 2014)	✓		✓	✓	✓	✓					
(Rantzer, 2016)								✓	✓		
Our Study								✓		✓	✓

Methodology

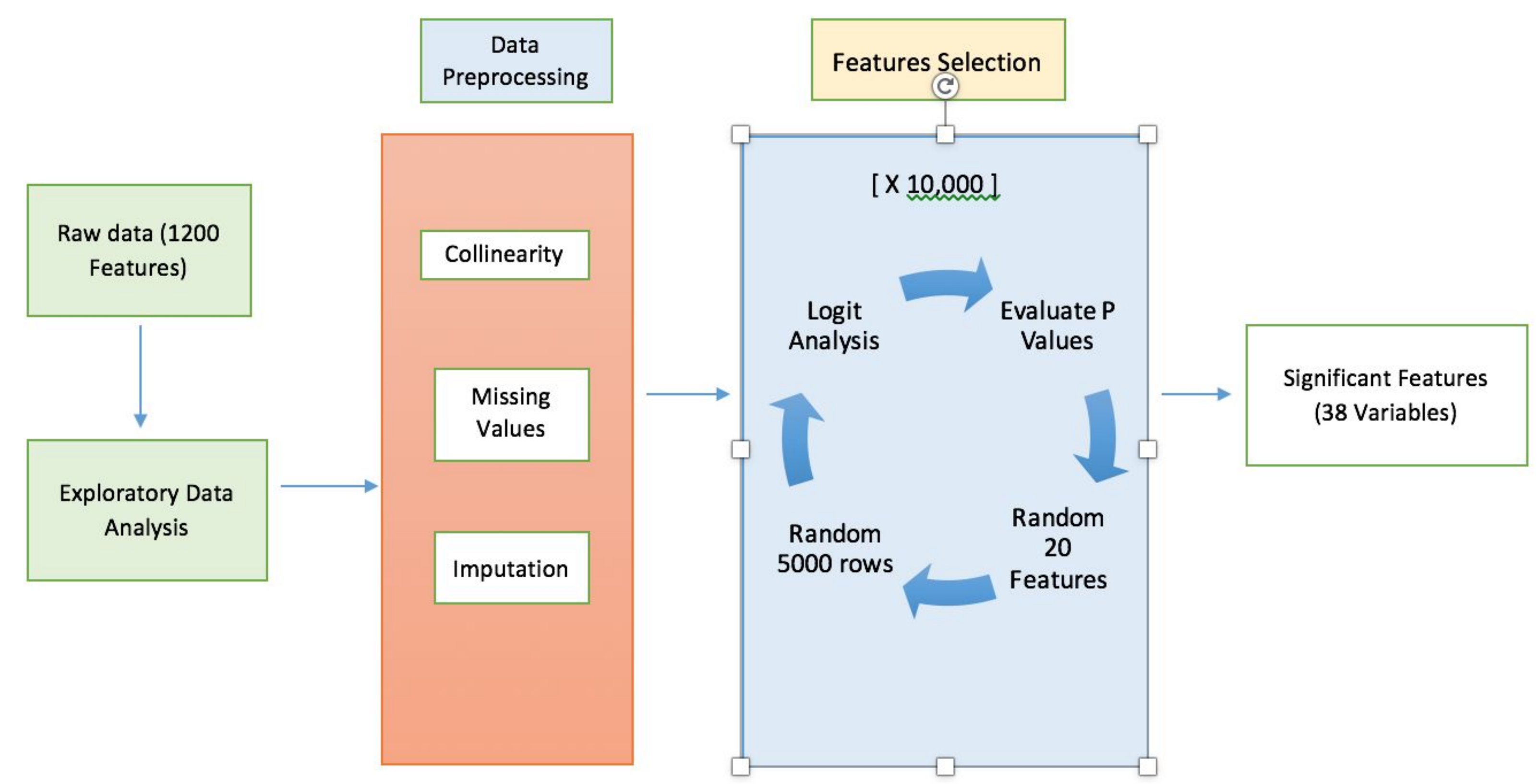

Figure 2. Feature Selection Process

Data
Raw: 1200 columns with no specified treatment variable
Clean: 37 variables and 1 treatment variable

Data Cleaning & Pre-Processing
Our raw data set is originally composed of nearly 1200 features, which is not beneficial for our business problem. We utilized various data cleaning techniques to identify noise in our dataset such as collinearity, non zero variances, and missing values.

Feature Selection
We created a loop to run for 10,000 times. For each cycle, our model randomly chooses 20 features and 5000 corresponding rows and run a logit analysis on the selected data. The p-values of all the logit models are captured, ranked, and evaluated. Based on p-values, insignificant features are removed and we are left with 38 features.

Model Design
We used a Causal Conditional Inference Forest to identify TR and CN as they are the only categories possibly comprised of the "Persuadables" target. We portioned our data into an 85% train-15% test split and ran several random forest and CCIF models to compare.

Responded	Treated: Yes	Treated: No
Yes	Treated Response TR	Control Response CR
No	Treated Non-Response TN	Control Non-Response CN

Methodology (Approach) Selection
We chose the Causal Conditional Inference Forest model because this model does not require pruning to avoid overfitting. This model is also unbiased in comparison to traditional tree models, which are biased towards variables with more potential splits.

Model Evaluation / Statistical & Business Performance Measures
The predictive models were evaluated using the Qini coefficient and average uplift by decile. The average uplift by decile is an important business indicator because it separates consumers by how well they respond to the treatment, breaking them into specific groups the marketer should target.

Results

Figure 3. Model Evaluation

The above graph showcases the several models run, with a basic random forest resulting in negative results regardless of split method. The most consistent model found was the CCIF with Euclidean Distance as the split method.

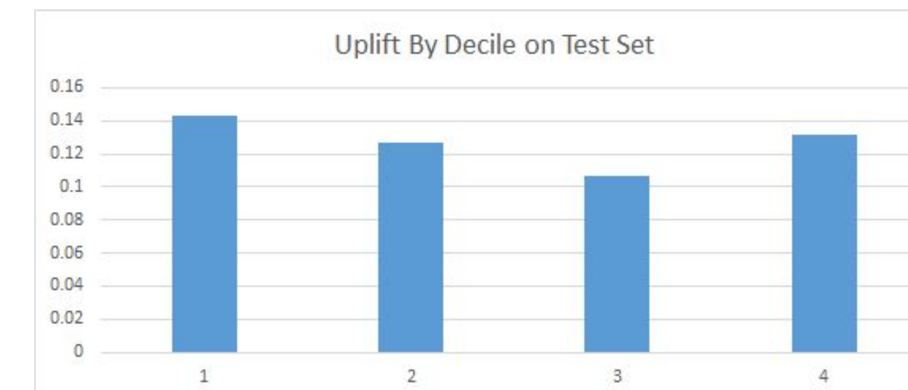

Figure 4. Decile Uplift Plot using ED Model

The above uplift plot shows how consumers in decile 1 are likely to generate the most uplift if targeted, essentially showing they are in the "Persuadables" target group.

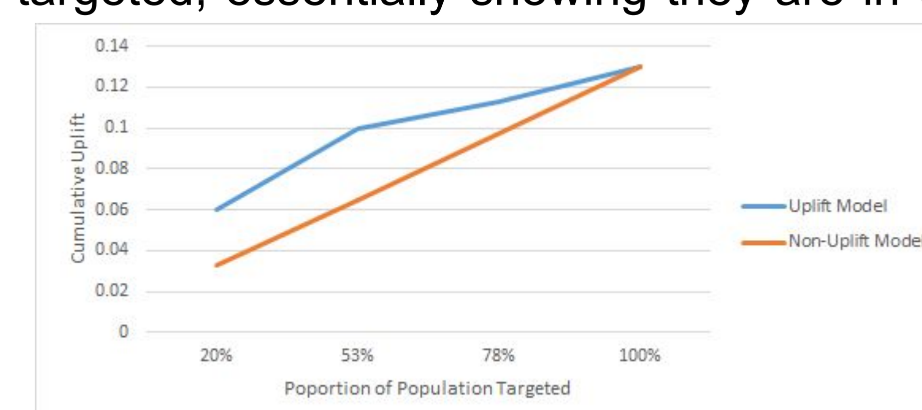

Figure 5. Qini Plot for ED Model

The above Qini plot showcases how much the uplift model predicts better than a standard random model. The Qini coefficient can be calculated using the area between the Uplift curve and the random line.

Conclusions

Our findings emphasize the importance of uplift modeling because they show how our modeling methods identify specific consumer behaviors that would not have been noticed in traditional modeling methods. We were able to showcase this by using the average uplift by decile as an indicator of how well specific groups of customers responded to outreach from a company. Understanding the detailed responses from each group helps marketers and businesses to reach their goal of effective consumer response.

Acknowledgements

We thank Professor Matthew Lanham for constant guidance on this project.