



Jonathan Leung, Spencer Rose, Jacob Irvin, Cameron Kwieciak, Alberto Navarro Alaez, Matthew A. Lanham

leung28@purdue.edu; rose109@purdue.edu; irvin7@purdue.edu; ckwiecia@purdue.edu; navarr31@purdue.edu; lanhamm@purdue.edu

Abstract

This study develops an approach to predict customer loyalty towards a credit card firm. Our model can be used by merchants and credit card firms to target customers with marketing efforts that increase loyalty and repeat business. Indicators that predict loyal customers allow marketers to focus on retaining customers that spend more money. This provides additional revenue to companies. Our study provides an analytical approach to this problem by analyzing data from ELO, a payment company based in Brazil.

Introduction



Source: Aitegroup, February 2018:

Figure 1
% of Rewards Cardholders Who Reported Each Factor as 'Extremely Important'

Customer loyalty is key to a business' future success and is the backbone for long term profit. Repeat customers provide your organization with more accurate predictive data on customer behaviors and overall customer purchase patterns. Majority of transactions are made using credit or debit cards, and they are providing incentives to cardholders. Our predictive model and research is aimed to show how credit card rewards are able to generate loyal repeat customers.

Primary Research Questions

- Is a repeat customer going to spend more money in the long run or is this unrelated? We must assume this is true in order to have our main objective be true.
- Using the target variables we have identified, can we predict customers most inclined to be loyal, and the best targets for a marketing campaign aimed at retaining them?

Literature Review

The articles we based our study on focused on multiple aspects, such as customer loyalty, credit cards, merchants, benefits, and reward programs.

Study	Customer Loyalty	Credit Cards	Merchants	Benefits	Rewards
(Saad, 2013)	✓	✓			
(Bunyamin, 2015)	✓			✓	
(Chakravorti, 2007)	✓	✓	✓	✓	
(Carbo-Valverde, 2011)	✓	✓			✓
(Minser, 2010)	✓				
(Our Study)	✓	✓	✓	✓	✓

Figure 2

The authors of these articles focused on the relationship between two or three of these factors. However, our study encompasses all five of these aspects in order to identify the beneficial relationship between credit card companies, merchants, and consumers. Figure 2 displays this specific criteria for each category.

Methodology

Figure 3 outlines our project process. Including our thought process and breakdown of data.

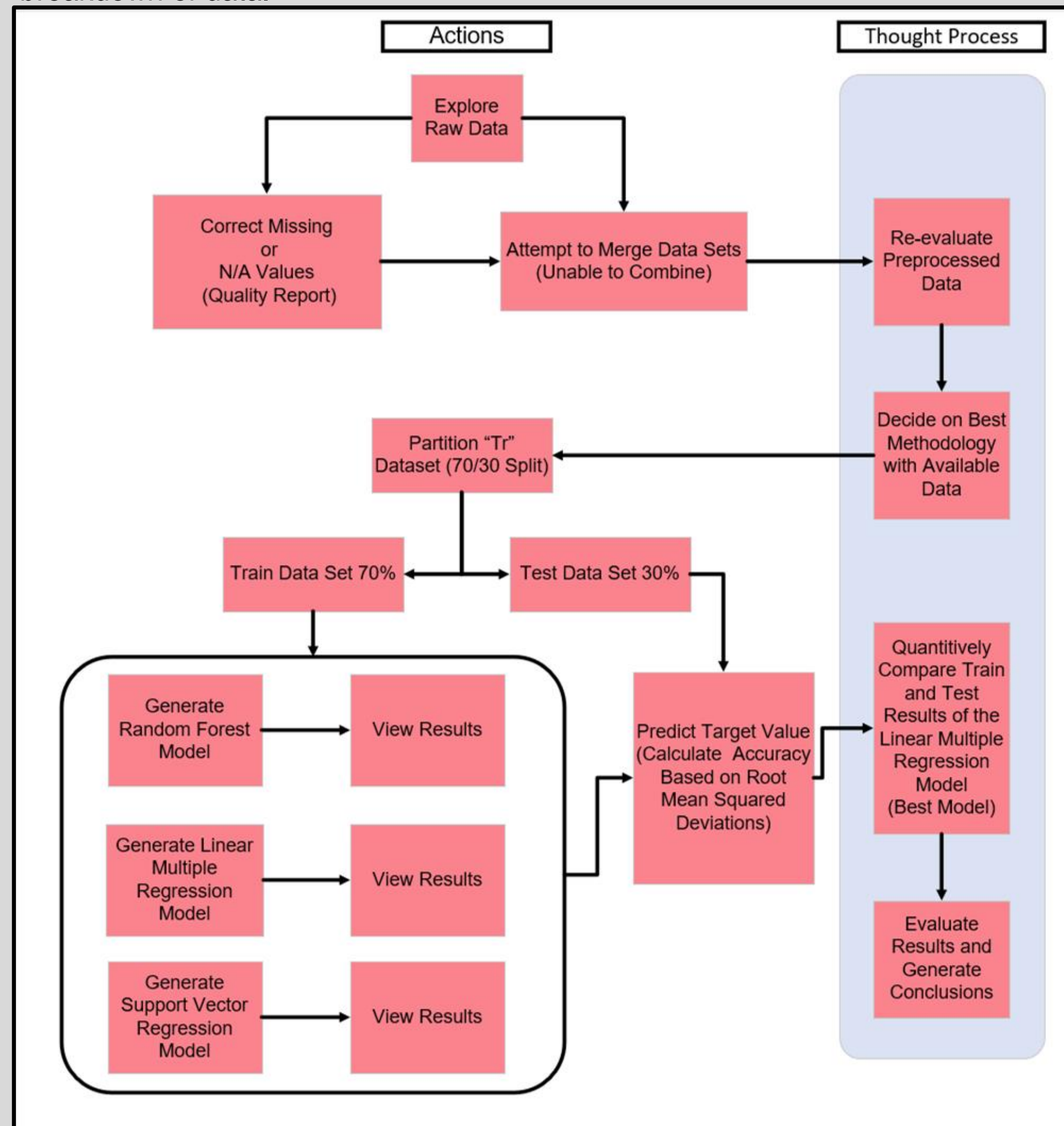


Figure 3

Data

Our data contains information about merchant transactions, customer IDs, first active month, and several numeric features that describe the data presented. Giving us an overview of customer interactions and purchase data.

Data Cleaning & Pre-Processing

- We attempted to merge several data files, but could not fully utilize all the data 90% did not have common matches.
- Resulting in us focusing in on the "TR" dataset.
- Using quality reports we found zero missing values in this dataset, making this the optimal dataset for our analysis.
- We then tested the variety of the data to see if there was a sufficient amount of different values. We removed the features with limited variation.
- Preprocess the data by standardizing these numeric values using a z-score standardization. This generated a normalized distribution of data.

Results

We evaluated three models:

- Linear Multiple Regression: a simple, basic model. RMSE:..
- Random Forest: supervised learning model with decision trees; we used 500 trees (different trials) and selected the mean result. RMSE: .
- Support Vector Regression: artificial neural network model. We used the eps regression type and the radial kernel to deal with data. RMSE:..

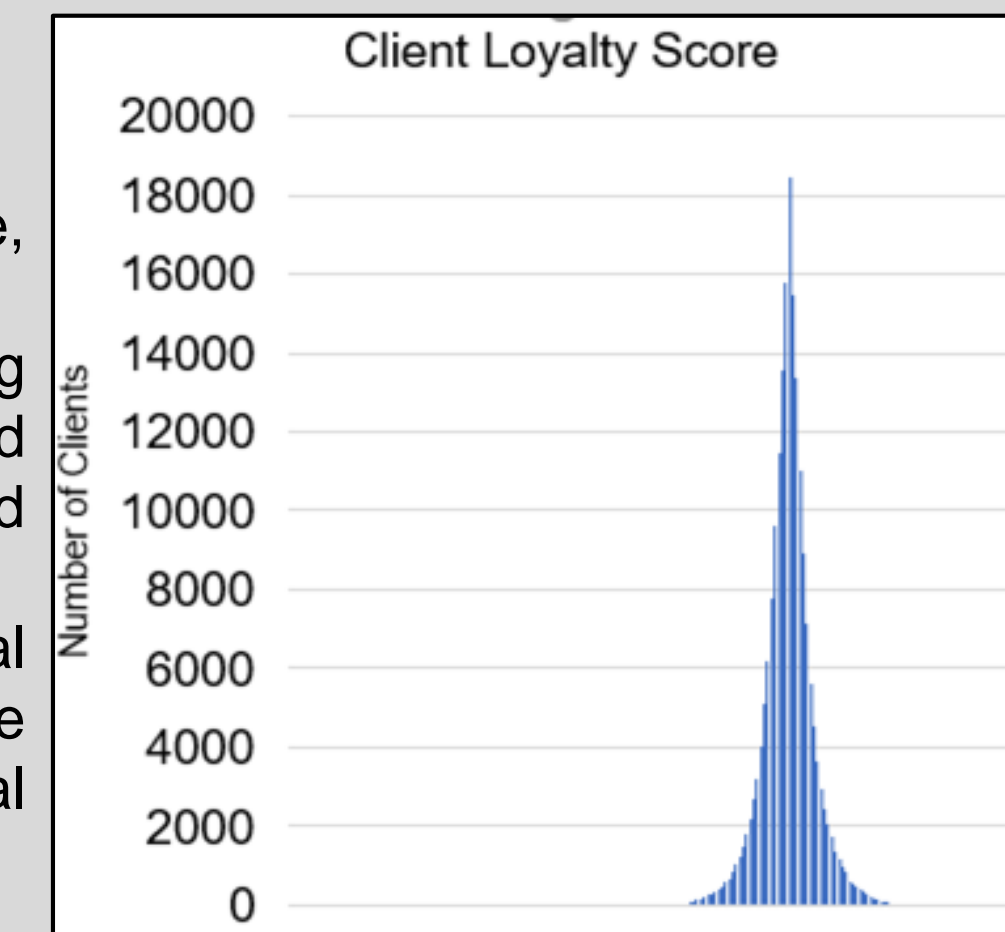


Figure 4

Figure 4 represents the variation of Client Loyalty Scores. Only few clients are extremely loyal, making them very valuable.

- Our RMSE results support our recommendation to use the SVR model in this type of predictive analytic model.

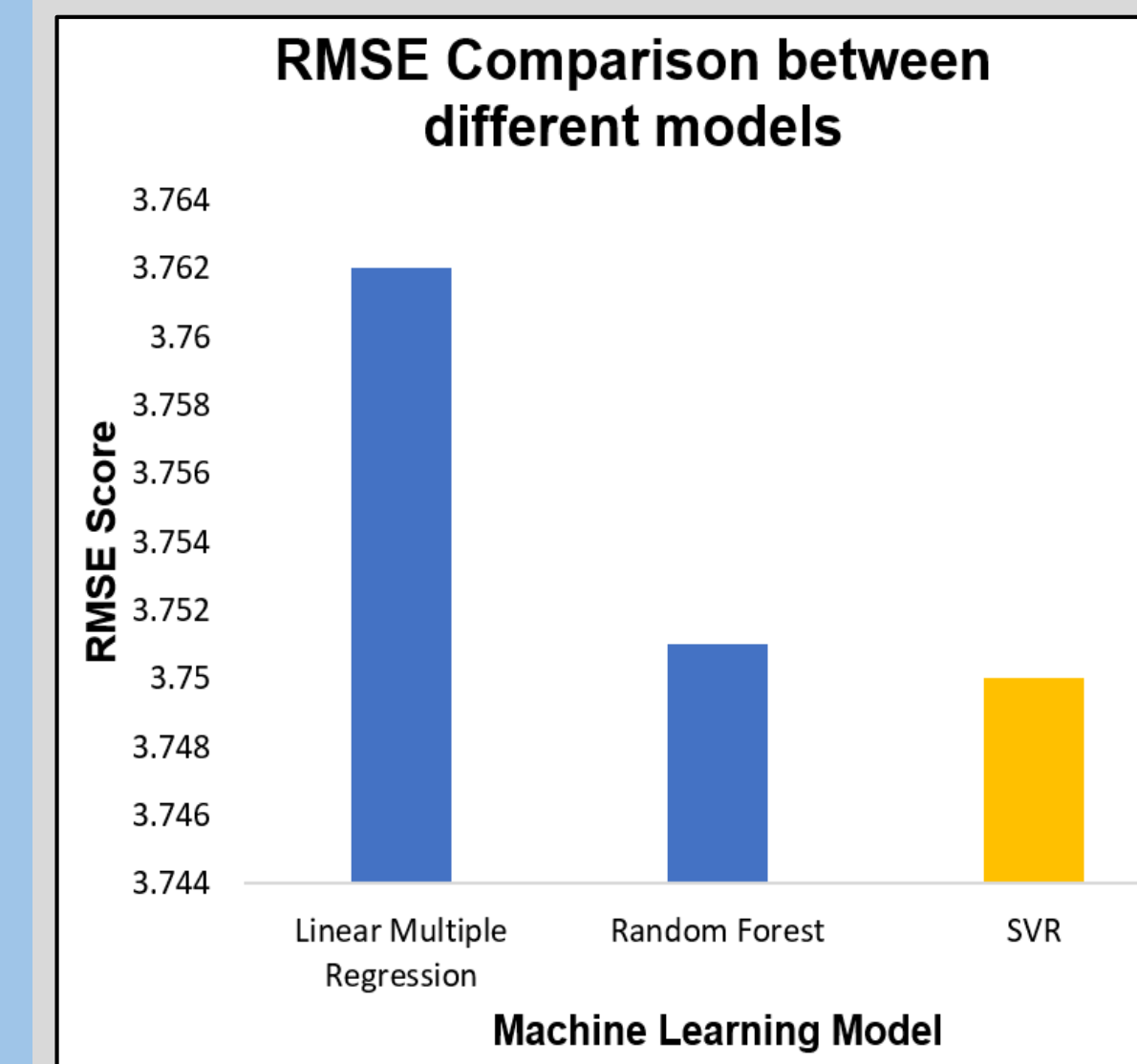


Figure 5

Figure 5 compares the three final models we chose, with SVR (highlighted) giving us the best results.

	LMR	RM	SVR
RMSE	3.7629	3.7512	3.7502

Conclusions

1. We found several useful models including SVR, giving us the best results.
2. If our algorithm is given features like average of purchase, customer active months, and city, we will be able to predict loyalty to our brand and design customer focus marketing strategies.
3. The performance of the model can be tuned frequently and can be rerun monthly providing frequent relevant feedback.

Acknowledgements

We thank Professor Matthew Lanham and Mehul Zavar for constant guidance. This project would not be possible without them.