

Report

January 24, 2020

Table of Contents

- 1 Project Overview
- 1.1 Data and Input
- 1.2 Problem Statement
- 1.3 Expected Result
- 2 Metrics
- 3 Evaluation
- 4 Conclusions
- 5 Citation

0.1 Project Overview

This project revolves around the Starbucks dataset provided through Udacity. I chose this problem because I have some experience in predicting customer outcomes from self-reported data as it relates to product use and engagement. I was also interested in the extent that demographics will actually have in offer completion. Does income or age matter at all in terms of who completes offers or who simply walks in a buys something.

From the Udacity project outline:

Your task is to combine transaction, demographic and offer data to determine which demographic groups respond best to which offer type.

- As such this project will try to predict each of the four event types: transact, complete offer, view offer, and receive offer. With a model that can predict transaction, completion, view, and receive then we can see what types of features have predictive power for a given outcome.
- It machine learning terms, this is a multilabel classification task. Tree boosted algorithms such as XGB perform very well specially given the size of the dataset and relatively large number of features.

0.1.1 Data and Input

The datasets are provided by Udacity and Starbucks. They include: - profile.csv - portfolio.csv - transcript.csv

These will provide the basis for the multilabel classifier. No further inputs will be needed.

0.1.2 Problem Statement

I will build a tree boosted algorithm to predict who will complete and offer, who will transact, and who will view/recv but not complete.

- These are major events that can be tracked directly with the data. If a customer chooses to not opt into an offer and completes anyway, that is only a win for business as they have spent more money and received \$0 in reward. It also tells us something about their purchase habits. If they only transact, then again they have spent money with the need for reward. Lastly, if they complete an offer after viewing, then they are responding well to advertising and are likely spending more, or at least the same, amount of money to receive the incentive.
- Additionally, by knowing who is likely to complete, transact, view, or receive only then we can analyze the business practices around duration, marketing, and difficulty to determine changes in engagement for those who do not complete or transact.

Once a model is built and evaluated, we can view the feature importance to understand what features are important to a given outcome.

- Feature importance will be a great window into know which of the many data points collected hold the most predictive power for the each class label.

0.1.3 Expected Result

I would expect this model to be able to predict fairly well the transaction and completion classes as they are very distinct and likely correlate to very different habits of a buyer. Offer view and offer receive are highly similar in root behavior. Your phone pings you with a push notification, you look at the offer, you get distracted and forget completely. In that fraction of a second you have received and viewed the offer and gone about your day. This categorically different from opting into an offer, using the app to purchase towards that offer and then (possibly) cashing in on the reward.

- Expectations are low for predicting views/receives in terms of accuracy, precision, and recall
- Expectations are high for predicting completion and transaction in terms of accuracy, precision, and recall

0.2 Metrics

The primary metrics that will evaluate this model are accuracy, precision, and recall.

- Classification of views/receives of offer will likely have some form the precision/recall trade-off. Either the model will have high precision- it's percent correct, and low recall- it's ability to find cases, or the reverse.
- Classification of transaction and completion will likely be very good. The class labels are fairly well balanced so we would expect precision, accuracy, and recall to be well balanced if the model can predict classes well.

0.3 Evaluation

XGB multilabel classifier performance cited below. As expected, transactions and completions are well predicted on the hold-out set. For offer views and received, the model cannot predict these classes very well. Further work could be completed to determine a better solution for dissecting these groups.

	precision	recall	f1-score	support
0	0.88	0.17	0.28	7546
1	0.47	0.97	0.63	5689
2	1.00	1.00	1.00	3416
3	1.00	1.00	1.00	14003
accuracy			0.79	30654

Events are numeric as such:

offer received = 0, offer viewed = 1, offer completed = 2, transaction = 3

This is a great result for predicting something like a transaction or offer completion. If the goal is to find who will buy without offer and who will buy with an offer, then this is pretty good.

0.4 Conclusions

The top features that seem to dictate the likely hood a customer will transact, complete offer, view, or receive. One of the strongest predictors is reward amount, duration, the specific offer, and transaction amount. Secondary predictors appear to be offer type, marketing saturation, and difficulty. The second and third models evaluated tell a similar story. Since offer id carries a heavy predictive power, future work could be to determine the significance of each offer for completion. Reward amount was a top predictor as well, so that is also an avenue for further work. Which amount is optimal for completion.

Interestingly enough; age, gender, income, and offer type carry no (or little) predictive power. There is some literature support this idea re: Martens and Provost (2011); at a certain point sociodemographics add no additional predictive power when compared to transaction activity.

0.5 Citation

- Martens, D., & Provost, F. (2011). Pseudo-social network targeting from consumer transaction data. Working paper CeDER-11-05, New York University - Stern School of Business.