

Starbucks targeted marketing Capstone Proposal

Background:

Marketing has been a major focus of organizations for a long time. With the growing access to technology and the accompanying creation of data, corporations can better market their products and services (Kolathayil). Marketing is defined as “the process of determining customer wants and needs and then providing customers with goods and services that meet or exceed their expectations” (A Brief History). Part of this is being able to determine what kind of promotions consumers are likely to respond to. A company has many tools available to it for marketing. With a variety of consumers, it is essential that a company figure out what tool will be effective with which consumer. This way the corporation will be able to cost-effectively spend its marketing money and consumers will be able to enjoy their desired products at a preferable price.

My personal interest in this domain comes from the wealth of data available on consumer behavior. I believe marketing is a matured and prevalent application of machine learning which all of us have likely been affected by, whether that be through YouTube, Google, or Facebook. I want to dive into this data set to get an understanding of how predictable consumer behavior really is and have the opportunity to work with a domain I have no prior experience in.

Problem Statement:

Starbucks utilizes a rewards mobile app to send offers to customers on Starbucks products. Starbucks does this with the goal of boosting its revenue. Starbucks has simulated customer behavior for a set time for a single product from members of its rewards mobile app. Starbucks was not always successful in influencing its customers to buy its product using its promotions. It would be helpful if Starbucks could send promotions to users who will likely be influenced by it. Influence here is defined by whether the customer will view the promotion and consequently, within the validity of that promotion, shop at Starbucks.

Our goal then is to use demographic information about Starbucks customers and their response to past promotions to determine whether a given customer will be influenced by a certain Starbucks promotion.

Datasets and Inputs:

Starbucks has provided simulated data for this project through Udacity. The dataset has three tables, portfolio, profile, and transcript.

Portfolio includes the ten promotions that Starbucks gave to customers during a set period. We are given information on the kind of promotion such as whether it was a discount, a buy one get one (bogo), or an informational.

The profile table has data on each customer. Specifically, when they joined the Starbucks mobile app, their gender, income, and age. It is notable that here we have 17000 customers in total; however, 2175 of them have no data other than join date. Consequently we will be discarding this data as we

hope to utilize demographics as a means of predicting whether a customer will be influenced by the promotion.

The transcript table has entries on how customers interacted with Starbucks. This includes when and how much they bought and when they received, viewed, and completed a promotion. This table was used to construct what we are calling “influence”. This is an indicator on whether a specific promotion was able to cause a customer to shop at Starbucks. The way we are computing this is by seeing that after receiving a promotion, did the customer view it and consequently shop at Starbucks in the effective period of the promotion.

Solution Statement:

To figure out which promotions were successful in influencing customers, we are going to use the provided information about customers and promotions to predict if the given customer will be influenced by the given promotion. We are going to do this by pairing all the times a customer received a promotion with all the times they were influenced by one. If they received but were not influenced that’s an indication of the promotion failing to influence said customer. If they received and were influenced, then that’s an indication that the promotion works. With this binary set up and the provided data we can construct models that predict which customer and promotion pairs will be successful.

The models I plan on using are XGBoost due to its adaptability to potentially non-linear data and Linear Learner as Sagemaker allows optimizing it for precision/recall.

Benchmark Model and Evaluation Metrics:

As a benchmark we ran the multinomial naïve-bayes. This model is an efficient algorithm that is often used in binary classification tasks (Ortner). The Naïve-Bayes classifier achieved an F1 score of 0.523. This will be our primary evaluation metric. This is the case because a corporation would want to send promotions to as many users as would be influenced by it while not incurring the costs of such advertisement if it will be unsuccessful. To identify such users the ideal model would have a high recall score, and to minimize advertisement costs the model should have a high precision. The f1 score is a compromise between these two.

Project Design:

As part of the data exploration and benchmark creation steps I have already created visualizations that demonstrate relationships between a promotions nature and its success in influencing potential customers. Consequently, the next step is finding the superior model to predict success of promotions. In this direction I will utilize sagemaker’s Linear Learner and XGBoost algorithms to achieve a model with superior accuracy. In doing so I will try hyper-parameter tuning as well as utilizing linear learner’s ability to optimize for recall/precision.

In the interest of space I have excluded visualization I had done in exploring the data, please refer to data_exploration1.ipynb and data_exploration2.ipynb for these. The benchmark is in feature_creation.ipynb.

Bibliography

A BRIEF HISTORY OF MARKETING. Valencia College

<http://faculty.valenciacollege.edu/srusso/chapter13.htm>. Accessed 10 Mar. 2021

Kolathayil, Yasim. "Machine Learning for Marketers." *Medium*, Towards Data Science, 28 Feb. 2019,

<https://towardsdatascience.com/machine-learning-for-marketers-78bff070cbd6>.

Ortner, Alex. "Top 10 Binary Classification Algorithms [a Beginner's Guide]". *Medium*, 28 May. 2020,

<https://medium.com/@alex.ortner.1982/top-10-binary-classification-algorithms-a-beginners-guide-feeacbd7a3e2>