Machine Learning Engineer Nanodegree Program

Starbucks Customer Behavior Prediction with Ensemble Machine Learning Methods

Carlos André Antunes

2021

Introduction

Domain background

Starbucks operates more than 31,000 stores worldwide, become one of the biggest coffee giants in operation at now. In 2017, Starbucks launched an initiative called “Deep Brew”. As said Kevin Johnson, CEO of Starbucks, *the vision of the company allowed that Starbucks recruit some of the best talents that, in another times, would like to work in the big companies of technology*. The recommendation platform was built to reach customers using multiple channels, including the Starbucks ordering app.

This new platform could delivers highly customized offers to the members of My Starbucks Rewards loyalty program. It helps Starbucks to automating offer assembly and management, reward fulfillment, and KPI measurement and tracking, at enterprise scale and deploy individualized offers across channels.

The industry-leading Starbucks Rewards Program has continued to grow. According to their website, “Membership has grown more than 25% over the past two years alone, climbing to 16 million active members as of December 2018, a 14% increase over the prior year since its introduction in 2007.

For Starbucks, is essential to use AI technology to enhance the customer experience, making improvements and customize more and more the attendance.

Problem statement

he goal that I have to achieve here is to best. Not all users receive the same offer, and that is the challenge to solve using the data set that is provided by Starbucks, which was captured over 30 days. I’ll also build a machine learning model that will predict the response of a customer to an offer.

Since it’s important to provide customized experience from ordering to offer on loyalty app, determining which kind of offer to send to each user based on their response to the previously sent offers, the problem becomes how to correctly and precisely locate the target customer group.

The main goal is maximize the conversion rate across deciding what the best-personalized offer to send to users. To do this, it’s critical to know which groups of people are most responsive to each type of offer, and how best to present each type of offer.

Datasets and inputs

There are three data files in json format in this project including demographic, offer and transaction data from the Starbucks rewards mobile app. They are:

portfolio.json

- id (string) - offer id
- offer_type (string) - type of offer ie BOGO, discount, informational
- difficulty (int) - minimum required spend to complete an offer
- reward (int) - reward given for completing an offer
- duration (int) - time for offer to be open, in days
- channels (list of strings)

profile.json

- age (int) - age of the customer
- became_member_on (int) - date when customer created an app account
- gender (str) - gender of the customer (note some entries contain 'O' for other rather than M or F)
- id (str) - customer id
- income (float) - customer's income

transcript.json

- event (str) - record description (ie transaction, offer received, offer viewed, etc.)
- person (str) - customer id
- time (int) - time in hours since start of test. The data begins at time t=0
- value - (dict of strings) - either an offer id or transaction amount depending on the record

We can see this files graphically in the mind map below:

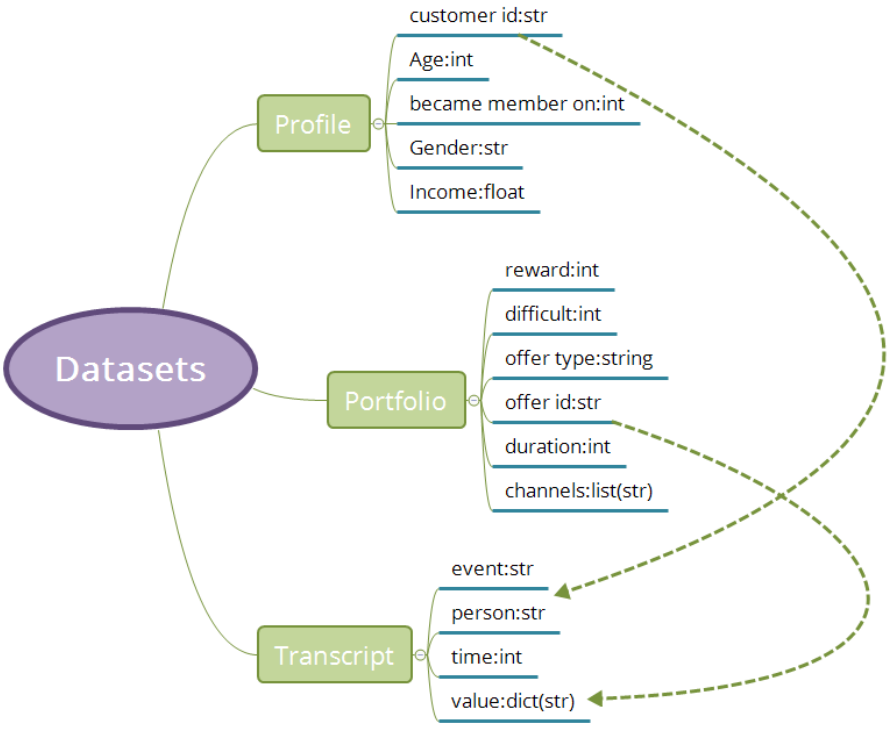

Figure 01: Datasets Mind Map

Solution statement

This project aims to build a machine learning model to decide which is the best type of offer to send to each customer. The dataset will be separated into three types of offer and fit the data into three supervised classification models.

Using this way, the model will predict whether the offer will be responded by customer or not when sent to them. Also, by investigating the feature importance of the model, it can help answer the question that what factors mainly affect people to make the decision and finally complete the transaction.

Therefore, with the solution above, it'll be easier and more accurate to identify which groups of people are most responsive to each type of offer, and how best to present each type of offer.

Benchmark model

In this project, we'll use some models to capture the the feature importance. Then we will compare this models ne each other. Finally we will select at least three models to make our final ensemble model.

Evaluation metrics

Since the project is building classification model, we choose both accuracy and F1 score as the model evaluation metric. The reason of choosing both metrics is sometimes when the dataset is imbalanced, the accuracy only couldn't objectively show how the model is performing on the dataset, while F1 score provides a better sense of model performance compared to purely accuracy as takes both false positives and false negatives in the calculation.

With an imbalanced class distribution, F1 may be more useful than accuracy. Also, since the F1 score is based on the harmonic mean of precision and recall and focuses on positive cases.

Project design

The project is designed with the following steps:

- Prepare and clean data: combine the three datasets. Understand the connection between columns and dataset.

- Data exploration: In order to analyze the problem better in next sections, first need to explore the datasets which includes checking the missing value, visualizing the data distribution, etc.

- Data preprocessing: In order to find out what mainly affect the finish of the transaction by sending the offer, in the data processing process, also need to process the data to merge the events of each specific offer sent so as to find out which offer were received, viewed and finally completed with a transaction.

Feature engineering

After basic processing, the next step will look if there are any columns that can be used to create new features. For example, generating a new column for length of customer's membership, the count of offer received for each user, calculate the time lap between offers, etc. Here I will use PCA methods trying to discover another relationships between data that could be used on the models.

Building model

The next step is to build the model using response flag generated in previous steps to predict whether the customer will respond to the offer or not.

Here we will choose the basic tree model as a baseline which will help explain the feature importance better so that we can get some insight into what factors affect customer's behavior most.

Model tuning

Compare the model using metrics selected above and tune the parameters of initial model using GridSearch method to get higher performance.

Conclusion and further improvement

Compare the final selected model to benchmark to see if the solution provide a better personalized offer. Also, review the built process and see if there's any opportunities to enhance the model in the future.