

## **Domain Background:**

In 2020 Starbucks is estimated to have over 32,000 stores worldwide. In 2019, the Starbucks mobile app had over 17,000,000 downloads and orders through the app accounted for nearly 30% of Starbucks total orders. A unique way that Starbucks is leveraging their mobile platform is sending each user customized offers to generate more orders. Optimizing which customers receive which offers could drastically increase revenue and help customer retention.

## **Problem Statement:**

Starbucks wants to make sure each customer is receiving the offer that is most likely going to cause them to place an order. We can determine this from looking at how users in certain demographics respond when given different offers. We can test the models we develop by applying them to past data and assess their performance.

## **Datasets and Inputs:**

We are given three sets of data.

**Portfolio:** Contains the offers given and their type. Either a discount, a BOGO, or an Informational offer.

**Profile:** List of customers that used the app and their demographics

**Transcript:** Collection of all of the actions taken by the company and user. Included offers received by users and their transactions.

## **Solution Statement:**

My solution will involve determining which offer type is best for each customer. I will build a model for each offer type (discount, BOGO, Informational) and compare each model's results for a given customer.

## **Benchmark Model:**

We will calculate conversion rate for each offer and look to surpass that with our ensemble model.

## **Evaluation Metrics:**

**Precision/Recall/F1:** Calculated from confusion matrix, these metrics will be able to tell us how many of our conversions were improved from the overall conversions.

**Area under the ROC curve:** A measure of how well a parameter can distinguish between two diagnostic groups(converted/no converted).

## **Project Design:**

**Data Preparation and Preprocessing:** Get knowledge for the structure of our model and remove null values. Create conversion metric from transcripts.

**Data Exploration:** Explore relationships in our data. For example, our certain type of offers in certain demographics lead to high conversions.

**Build Our Models and Tune Hyperparameters:** We will build models for both BOGO and Discount since it is difficult to judge true conversion for informational offers. We will try random forest, support vector machines, and a neural network for each type of offer and tune each model.

**Compare models:** We will then compare each model to determine which model to use for each offer type.

**Conclusion:** Have a finalized model for each type of offer to determine when to give an offer and to what customer.