Carlos He He, Troy Kirinhakone

CPTS 575

Project Proposal

Consumer Behavior Analytics: Ecommerce

Our purpose of this project is to predict consumer behavior, by analyzing the ecommerce industry where there is volume and richness in the datasets. We shall use various data science techniques to optimize business practices in order to maximize efficiency and profit. Our data is from a Brazilian ecommerce dataset from Olist Store which is from an actual business with sensitive information substituted. The dataset includes information of over 100,000 orders from 2016 to 2018 purchased at multiple marketplaces in Brazil. Some of the initial techniques we intend to employ would be the use of clustering and linear regression in order to gain insight to optimize upon the following business practices:

* Sales Prediction: Using purchase date information we will predict future sales.
* Delivery Performance: Utilizing models we will employ potential alternatives to enhance delivery efficiency.
* Product Quality: Using the customer satisfaction metrics along side the product reviews we will determine product trajectory growth.

As we discover more information about the business we also plan to implement feature engineering to develop new features from the existing data. In addition, we have found quarterly ecommerce sales data from both the U.S Census and the UK. After cleaning the data, we intend on exploring these datasets and cross reference them to our Brazilian datasets. We foresee potential similarities and differences in behavior among the consumers in the three regions thereby making further predictions and discoveries.