Roni Kaakaty

DSC630-T302

07/04/2021

Week 4 Exercise

Target PA Case Study

Overview

The goal of every successful retail company is to find ways to maximize their earning potential. Whether that is by cleverly filling up end caps with popular items or identifying a consumer spend trend to capitalize on, every dollar is important. Companies need to find ways to stay ahead of the competition and back in 2010, Target identified a way to do just that. The global pregnancy care market cap size is estimated to reach 33.22 million by the year 2025, according to Grand View Research. Target knew that if they could accurately predict when their customers were pregnant, that they could tap into that market cap and increase their sales within that population. They knew that if they were able to create a predictive analytics model based on various factors that could predict a patron’s pregnancy, they would be able to target them with marketing and coupons to entice them to shop at their store.

Business Understanding

The first course of action would be to identify a way to confirm the women that are indeed pregnant. I agree with the Target approach that located the customers that signed up for their baby registry and notated them with great certainty that those individuals were pregnant. With that information, we will then be able to track their purchases along with at what time intervals to try and get a gauge of what period of their pregnancy they are at. By doing this, we can send targeted ads based on the changing needs of the mother-to-be.

Data Understanding

This section will focus on identifying which purchases are being made and the quantities of each item. I will attempt to identify a trend that the customers who created the registry were buying regularly. With that information, I will then be able to target the segment of the population that may have signed up for a baby registry elsewhere or have chosen not to create a registry. I will need to make sure not to include customers that made one or two purchases as they could just have been buying those products as gifts for people who made a registry. I would create a minimum threshold that I would go off of and would eliminate products based on those criteria. Clothing would be another factor to consider as soon to be mothers will tend to buy maternity clothes or larger than usual clothes sizes to offset their growing body.

Data Preparation

Data preparation would involve identifying trends in product sales from customers that have been identified as pregnant. Eliminate data that does not qualify for the minimum threshold. Identify different consumer shopping habits that do not align with typical spending. Assign a variable to indicate whether or not the model has identified the woman as pregnant. Compare that output with non-registered patrons to see if they are purchasing similar products at similar quantities. Eliminate consumers based on age. Identify the features that contain a positive correlation to pregnancy.

Modeling

Create a classification predictive model using input variables to capture discrete output variables. Since we are not looking to predict a continuous variable, it would make more sense to go with a classification model instead of the usual regression model. I will train my model on a data set with a 70/30 split. I would use Python as my language of choice and R for the visualizations. The dataset will contain information from consumer spending habits of individuals who have signed up for a baby registry at Target and those who have not signed up, but have comparable spending trends to those that have.

Deployment

This stage involves “what-if” analysis. This section will factor in ethical considerations that I would investigate to possibly exclude certain data from being used. For example, I would aim to omit data captured from individuals under the age of 18. A reason for this, other than the obvious under-age aspect, is that these patrons would not have enough historical spending data to go off of. Individuals at that age might not have the financial means to get all the necessary products that are typically purchased during a mother’s pregnancy journey.

Conclusion

This predictive analytics model can be used to accurately predict when a consumer is pregnant. Effectively identifying these individuals early on can help the marketing team steer that customer to the retail store to create a registry. This registry will then force other individuals to purchase products at the same company for the individual who has that registry. This model can effectively increase sales. A limitation of this model is being able to predict pregnancies other than the first. Parents that already have children are less likely to buy as many items for the upcoming baby if they already have the items in stock. New parents typically need everything, which makes accurately predicting whether or not the patron is pregnant based on their spending habits easier. Strollers, car seats, cribs are items that are saved and passed down, so those products should not be used as features in the model creation.