

## **Predicting the Products an Online Grocery Shopper Will Purchase Again**

Yi Li

Online grocery shopping is growing rapidly these years. According to the U.S. Online Grocery Survey 2020, released on May 6 2020, one half (52.0%) of all respondents had bought groceries online - more than double the shopper numbers from two years ago. The coronavirus pandemic is transforming consumers' needs and behaviors, and has encouraged more grocery shoppers to start buying or buying more online.

There are many grocery delivery apps in the market today such as Instacart, Shipt, Amazon prime now, and Walmart grocery delivery etc. Features that help customers' shopping experience more easy and efficient will make the app stand out from others. Correctly predicting customers' shopping behavior using machine learning, and incorporate it into the features of the apps will make their consumers' shopping experience more pleasant.

I am going to use a dataset from a Kaggle competition to predict the products a customer will buy again. This dataset is provided by Instacart. This dataset is anonymized and contains a sample of over 3 million grocery orders from more than 200,000 Instacart users. For each user, 4 to 100 of their orders are provided, with the sequence of products purchased in each order. The week and hour of day the order was placed are also provide, and a relative measure of time between orders.

Link: <https://www.kaggle.com/c/instacart-market-basket-analysis/data>

The deliverables will be code and a PowerPoint slide.