Predicting the consumer’s next purchase order

Instacart Future Order Analysis by User

Serhan Sari



# Instacart

Instacart provides grocery delivery services from selected grocery stores to consumers’ door.

# Problem Statement

Predicting which item(s) will be in a user’s next order by looking at over 3 Million historical purchase orders from more than 200,000 anonymized Instacart users.

# Dataset

There are 6 different files that consist of a relation between each other which describe consumers’ orders over time. There are more than 3 million grocery orders and from more than 200,000 anonymized Instacart users which contains each users’ past purchased products from 4 to 100 in each order by weekly, daily and hourly that the order was placed and also the interval time between the orders.

1. **aisles.csv**

* Aisle numbers and description;

Headers: aisle\_id, aisle

1. **deparments.csv**

* Department numbers and description;

Headers: department\_id, department

1. **order\_products\_prior.csv & order\_products\_train.csv**

* These files provide us in each order which product placed an order along reordered rate;

Headers: order\_id, product\_id, add\_to\_cart, reordered

1. **orders.csv**

* This file contains each and every user’s order history along with the order times (hour & day), reorder time interval between the last time ordered and which set (prior, train) an order belongs.

Headers: order\_id, user\_id, eval\_set, order\_number, order\_dow, order\_hour\_of\_day, days\_since\_prior\_order

1. **products.csv**

* This file describes which products belong to which aisle and department

Headers: product\_id, product\_name, aisle\_id, department\_id