XXXXXX STORE OTC RECOMMENDATION

Vasantha Raj Tejas

Business Objective

Image by Rina Piccolo

Client want to understand their customer transactional data and identify combinations of items that occur together frequently in store visits. This help to market related products to the customer based on prior purchase behavior patterns or what is currently in their cart .

Also, place products that co-occur together in the analysis in close proximity on the store floor to improve the shopping experience of the customer.

About Data

I. Client has provided the customer transactional data from different databases.
II. It consists of 21 columns for one month with 30,000 observations.
III. Consist of following columns,

Customer Details – Customer No, Customer Name

Contact Details – Address Line1,Address Line2, City, State, Postal Code, Fax, Phone Number , Country, Territory

Order Details – Order Create Date , Order Number, Quantity Ordered, Order Amount

Item Details – Item Code, Item Description, Item Category Number, Item Category Name

Date variables – Fiscal Year, Fiscal Month

Data Understanding

I. Clients wants to measure the similarity between the items that users add to their basket and other items.

II. Data consist of user details and items ordered in each transaction.

III. In order to achieve our business objective, we have identified the key 3 variables for next level of analysis.

a) Customer Number – Nominal, uniquely assigned to customers

b) Order Create Date – Date, 1 month period from 1st October 2018 to 31st October 2018)

c) Item Description – Product(item) name. Nominal. Total item count is 4898.

Data Preparations

I. Frist, we have sub setting what all the (3) columns needed.
II. Next, quick preparation was with respect to Order Create Date, by converting it to date format.
III. Further, as we read the data from different sources earlier, observed that some double entries in the dataset(0.001% of data), same has been removed.
IV. Unnecessary special character in the data set has been removed.
V. Now, for our objective we have current data set at each customer level against a day of purchase, i.e. transaction records.
VI. Thus, further such data has been transformed into one order (or) in one row (basket format) for moving further for analysis.
VII. As, transactions with 1 item are not useful for analysis, same has been removed and left with greater than or equal to two items.

Approach

- Idea is to determine what products customers purchase together in the basket during their grocery shopping.
- Such, business requirements fall under recommendation systems.
- Wherein given our data with respect to user and items, indicate collaborative filtering techniques are not a bad choice.
- Hence in order to recommend best products, study exploited one of the popular technique namely., association rules that fall under un supervised machine learning.
- That is, to find out association between each item with respect to others in the basket.
- Hence, proceeded with APRIORI for algorithm for defined businsess objective.

Model and Results

- APRIORI technique with support greater than or equal to 0.1% and confidence greater than or equal to 1% have been employed the prepared transactional data.
- In first instance, it generated 51 rules.
- However, based on lift statistic, we have provided top 5 rules which are present below.

Table 1: Glimpse of top 5 rules from APRIORI Technique

LHS	RHS	Support	Confidence	Lift
xxxxx Vitamin D Whole Milk - 1 gal	xxxxx Round Top White Bread - 20 Oz	0.001037037	0.934	900
xxxxx Fat free Whole Milk - 1 gal	Doritos Nacho cheese -9.75 oz	0.001259259	1	750
xxxxx Grade A Large Eggs - 18 ct	xxxxx Processed Cheese	0.001259259	0.945	750
Diet Pepsi Soda	Tortila Chips - 10 Oz	0.001037037	1	614
Yoplait Original Orange Cream Yogurt	xxxxx Vitamin D Whole Milk - 1 gal	0.001037037	0.637	613

Business Insights

Client can use this insights gained from results in a number of ways, including:

Cross Sell: place below products in close proximity on the store floor to improve the shopping experience of the customer.

I. Milk Products, Yogurt and Bread items
II. Eggs and Cheese
III. Beverages/Soda and Chips/Nachos

Item Recommendation: Recommend following associated products that are frequently bought together with high confidence(approximately100%) and Lift values in the POS devices in the billing counters.

I. xxxxx Vitamin D Whole Milk - 1 gal and xxxxx Round Top White Bread - 20 Oz
II. xxxxx Fat free Whole Milk - 1 gal and Doritos Nacho cheese -9.75 oz
III. xxxxx Grade A Large Eggs - 18 ct and xxxxx Processed Cheese
IV. Diet Pepsi Soda and Tortilla Chips - 10 Oz

Marketing Promotions: Target marketing campaigns to customers and entice them to purchase with promotions on the related products.

Deployment

Generated association scores (i.e. top rules) have been presented to the client by way of JSON format for his implementation purposes.

Thank you