D212 Market Basket Project

Jacob Polomsky

July 28, 2024

A1. What products are commonly bought together so we know when to discount an item?

A2. One goal is to use Market Basket Analysis to see what items are commonly bought together so we can discount those products to customers who are buying those products.

B1. Market Basket Analysis groups particular items of interest, allowing you to see and count the number of transactions between those groups and find patterns between buying those two or more products. Market Basket Analysis makes it easy to compare a wide variety of items to see how buying one item influences buying another, allowing you to make connections between buying those items and possibly finding ways to recommend to people those commonly bought together to increase sales.

B2. An example is the chart below. You can see that when bread is bought, milk is also bought ¾ times, making an association between buying bread and milk.

A screenshot of a menu

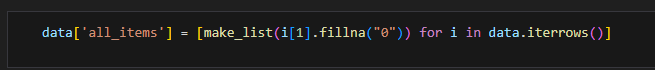
Description automatically generated

B3. One assumption is that for any infrequent item, paring all subsets of the pairing must also be infrequent. With the example above, you can see if we were to mix bread and cat toys, not in the dataset, that if you were to add bread, cat toys, beer, or anything else there, there would not be much association because bread and cat toys are not commonly bought together.

C1. First, I put all the items bought into one column by making a function that removes the null values

A screen shot of a computer code

Description automatically generated

Then, a for loop iterates over each row and passes all the columns into the function that removes null values.

Then I used an encoder to make a one-shot encoding data frame that shows items bought in each transaction. The data frame is true or false based on if the item was part of the transaction.A screen shot of a computer code

Description automatically generated

C2.

A screenshot of a computer

Description automatically generated

C3.

A screen shot of a computer

Description automatically generated

C4. The top three values are filtered based on support of 0.001 and a confidence of .9. Then the data is sorted by lift. Then, I included an extra point since positions 2 and 3 are similar enough to be grouped. The support shows how often the item has been purchased in all possible transactions that need to include the antecedent. Confidence is how often the item has been purchased when the antecedent items have been purchased. Lift is how much more common the item has been purchased with the other items compared to it being purchased alone.A screenshot of a computer

Description automatically generated

D1. With the top three, we see that a Logitech wireless mouse with an apple pencil and dust-off compressed pack has been purchased in 1 out of every 1000 transactions. In those transactions, Logitech, mouse has been purchased in 90 % of those transactions and is 12 times more likely to be purchased in a pack with an apple pencil and compressed gas than being purchased by itself. The iPhone case, nylon braid, and USB to HP 63XL ink have been purchased 3 out of 1000 purchases, and the ink has about a 97% chance to be purchased with those prerequisites, which increases the chance of ink being purchased by around 12 times. With the final of the three, we can see that Sandisk 128 and a dust-off compressed gas leads to 64 GB cards being purchased in 1 out of 1000 transactions. 64 GB card has a 90 % chance of being purchased when the prerequisites are purchased. It leads to 64 GB cards being 9 times more likely to be purchased.

D2. We found that there are quite a few items that are commonly bought together, which leads to a very high chance that another item will be purchased. Like Spark Nylon and Braided USB, plus some other items, HP 63XL Ink will be 12 times more likely to be purchased compared to it being purchased independently. This allows users to know what items to discount.

D3. We can use this algorithm to discount items that might interest them based on what they have already purchased. For a stylus pen for an iPad, we can offer a discount on the dust compression pack. This will help recommend discounts to these customers, reducing the chance they will churn by offering them.