End-to-End Machine Learning Project:
Recommender System using Association Rules (Predictive Analytics)

Qiuhong Wei

BUSINESS QUESTION

- **Challenge**: The dataset is big, and each customer has unique preference.
  - Customers make online grocery purchases from the Instacart grocery delivery service. When a customer buys an item, what are the related or complementary items that can be presented to them to promote cross-selling?
- **Business Questions:**
  - How can we mine association rules between various grocery items?

BACKGROUND INFORMATION

- Generate frequent item sets and association rules for a recommender system using ***Association Rule Mining***: an unsupervised machine learning technique
- **Methods:** Market Basket Analysis: **Apriori algorithm**
  - A frequent itemset is a set of items that frequently appear together and satisfies both minimum support threshold and minimum confidence threshold
  - **Support** is a measure how frequently the itemset appears in the dataset
  - **Confidence** is a measure of how often the rule has been found to be true
  - **Lift** of greater than 1.0 indicates some usefulness of the rule

BUSINESS RECOMMENDATIONS

- Based on the **maximum in the support** for boosting the usefulness of the rule.
- The insight is the same **food type**. Lime and lemon are citrus, and the raspberries and strawberries are berry.
- The store should recommend **similar type products** to clients when they buy some products.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift
21	(Limes)	(Large Lemon)	0.045980	0.062000	0.012156	0.264379	4.264159
20	(Large Lemon)	(Limes)	0.062000	0.045980	0.012156	0.196066	4.264159
31	(Organic Raspberries)	(Organic Strawberries)	0.042268	0.083028	0.012728	0.301118	3.626710
30	(Organic Strawberries)	(Organic Raspberries)	0.083028	0.042268	0.012728	0.153295	3.626710

BUSINESS RECOMMENDATIONS

- Based on the **maximum in the lift** for finding the frequent pairs in the historical data, the values are low.
- The common result is the **organic food** are usually connected.
- The store should recommend **organic food** to clients when they buy organic products.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift
6	(Organic Strawberries)	(Bag of Organic Bananas)	0.083028	0.117980	0.023428	0.282174	2.391714
7	(Bag of Organic Bananas)	(Organic Strawberries)	0.117980	0.083028	0.023428	0.198579	2.391714
2	(Organic Hass Avocado)	(Bag of Organic Bananas)	0.055583	0.117980	0.018444	0.331825	2.812560
3	(Bag of Organic Bananas)	(Organic Hass Avocado)	0.117980	0.055583	0.018444	0.156331	2.812560
0	(Bag of Organic Bananas)	(Organic Baby Spinach)	0.117980	0.074568	0.017042	0.144444	1.937082
1	(Organic Baby Spinach)	(Bag of Organic Bananas)	0.074568	0.117980	0.017042	0.228536	1.937082

PART B

	Business Question	Data	Predictive Analytics	Business Action	Business Outcome
1	How can we improve personalized recommendations for online shoppers?	Customer transaction history, browsing behavior, product details.	Recommender System (association rules)	Generate product recommendations based on frequently co-occurring items.	Increased customer satisfaction, higher conversion rates, and improved cross-selling. By leveraging Association Rules, the recommender system can identify patterns of items frequently purchased or viewed together by customers. This enables the system to make personalized recommendations, suggesting additional products that are likely to be of interest to the shopper.
2	How can we enhance content recommendations for streaming platforms?	User viewing history, content attributes, user feedback.	Recommender System (association rules)	Recommend related content based on associations between viewership patterns.	Improved user engagement, increased content consumption, and reduced churn. Association Rules can help identify patterns in user viewing history, enabling the recommender system to suggest similar or related content based on co-occurrence and association between different shows or movies.
3	How can we optimize product placement in a grocery store?	Point of sale data, product inventory, customer purchasing history.	Recommender System (association rules)	Determine optimal product placement based on item associations.	Increased sales, improved customer experience, and enhanced store layout. By analyzing transactional data and customer purchasing history, Association Rules can identify patterns of items frequently purchased together. This information can be utilized to strategically position related or complementary products in close proximity to increase cross-selling opportunities.

THANK YOU.