Product Recommendation System Using Association Rule Mining

Team 11:- Dhrumil Makwana, Abhinav Venkatesh, Tejas Kaushik

Project Description

- Through this project we intend to identify customer purchasing pattern by analyzing the items purchased in sequence to determine cross-sell. For this project we will consider data from multiple platform / store to get relevant interesting patterns for that store / platform.

- By Identifying this interesting pattern, we can increase the sales by optimizing the product placement, offer special deals and creating product bundle to encourage further sales of these combination.

- Popularly used in Amazon, Walmart, Target and many more.

Prior Work

- Market basket analysis

No	Research	Description
1	R. Agrawal, T. Imielinski, A. Swami, Mining association rules between sets of items in large databases, Proceedings of the ACM SIGMOD International Conference on Management of Data, Washington, D.C., 1993, pp. 207 – 216.	Explores the problem of mining association rules. Presents the significance of rules with minimum transactional support and minimum confidence.
2	R. Agrawal, R. Srikant, Fast algorithms for mining association rules, Proceedings of the 20th VLDB Conference, Santiago, Chile, 1994, pp. 478 – 499.	Presents the apriori algorithm for identifying association rules between items in a large database of transactions.

Prior Work

No	Research	Description
3	Raorane AA, Kulkarni RV, Jitkar BD. Association Rule – Extracting Knowledge Using Market Basket Analysis.Research Journal of Recent Sciences 2012:1(2):19-27.	Explores the use of market basket analysis to facilitate the arrangement of the products in a supermarket to increase its profits.
4	I. Bose, R.K. Mahapatra, Business data mining—a machine learning perspective, Information and Management 39 (2001) 211 – 225.	Presents the business application of data mining techniques in different fields such as marketing, retailing, finance etc.

Datasets

- Market Basket Analysis Kaggle (https://www.kaggle.com/datasets/aslanahmedov/market-basket-analysis)
- Online Retail 2 UCI Kaggle (https://www.kaggle.com/datasets/mashlyn/online-retail-ii-uci)
- All members have downloaded the datasets.

Proposed Work

- Data Cleaning and Preprocessing:
  - Remove irrelevant data(country, quantity, description).
  - Remove redundant data items.
  - Deal with missing data(Removing Data entries as replacing value will make unnecessary noise).
  - Fix structural errors (Aliasing item name to stock code).
  - Reduce the dimension of dataset (Reducing to only important parameters after combining datasets).

Proposed Work

- Data transformation:
  - Combining items in one basket based on bill number and Invoice number.
  - Combining both datasets once cleaning and preprocessing is done.
  - Grouping product name based on category (e.g.:- Milk and Skimmed Milk can be categorized as Milk).
- Data mining:
  - Association rule mining.
  - Recommendation System.

Tools

- Python
- Pandas
- NumPy
- Matplotlib
- SciKit-learn

Evaluation

- Product Recommendation System:-
  - Accuracy
  - Precision
  - Recall