Project Proposal

Course:Data Mining

Professor: Dr. Alice Smith

Teacher Assistant: Jingwei Liu

Student: Mohammad Maydanchi

People who bought … also bought …

In a store, all vegetables are placed in the same aisle or all kinds of meat are together. Selecting a place for each product is not arbitrary. For example, diapers are far from beers, why? Based on data analytics, stores know many families who bought diapers bought beers too. So, if they are far from each other, customers who bought one of them and want to buy another one, will consume more time in the store, so it is more likely that they buy other products. Thus, if stores know all relations between products they can sell more. Association Rules help data scientists to find these relations.

My data is the **Market Basket Optimization** dataset which has almost 7500 rows. Each row contains different products of a grocery store which is bought by different customers at different times. The products are such as almond, chicken, olive oil, spaghetti, mint, … .The numbers of products in each row are not the same because it is obvious that different customers do not buy the same number of products.

The goal of this project is uncovering the probability relations between the products of the store. For this aim, I will use **Apriori** and **Eclat** as methods of Association Rules for analyzing the dataset. I completely apply these two methods to my data and compare their results as report.