

Laboratory Practice 1 :

ARTIFICIAL INTELLIGENCE AND ROBOTICS LAB

Grocery Basket Analysis and Prediction

Project ID: 5

Project created by :

Harsimran Singh Dhillon (17)

Himanshu Garud(18)

Ojas Ahire(19)

Aniruddha Kulkarni(20)

Project Guide: Prof. A. V. Taware

December 19, 2020

Contents

1 Problem Statement	1
2 Objectives	2
3 Algorithms	3
3.1 Apriori Algorithms	3
4 Block Diagram	5
5 Screenshot of Output	6
6 Outcomes	9
7 Conclusion	10

List of Figures

4.1 Block Diagram	5
5.1 Bar chart of items	6
5.2 Top 10 First Choice	7
5.3 Top 10 Second Choice.	7
5.4 Top 6 Third Choice	8

Chapter 1

Problem Statement

Grocery shopping could take a lot of time to find items and people might forget to buy items that they would require. This could affect the sales of supermarkets and can cause inconvenience to the user. So to increase sales frequently bought items could be predicted.

Chapter 2

Objectives

- To use different methods for grocery item sales data visualization
- To develop an application to analyse and predict frequently bought items to increase the sales of the supermarkets
- To use the most efficient algorithm for predicting the sales and mining frequent itemset

Chapter 3

Algorithm

3.1 Apriori Algorithm

- Apriori algorithm uses frequently bought item-sets to generate association rules
- It is built on the idea that the subset of a frequently bought items-set is also a frequently bought item-set
- Frequently bought item-sets are decided if their support value is above a minimum threshold support value
- How does the Apriori Algorithm work?
 - The Algorithm considers 3 important factors which are, Support, Confidence and Lift.
 - **Support** : It tells us about the combination of items bought together frequently. It gives the part of transactions that contain both A and B.

$$Support = \frac{freq(A, B)}{N}$$

- **Confidence**: It tells us how frequently the items A and B are bought together, for the no. of times A is bought.

$$Confidence = \frac{freq(A, B)}{freq(A)}$$

- **Lift** : It indicates the strength of a rule over the randomness of A and B being bought together. It basically measures the strength of any association rule.

$$Lift = \frac{Support}{Supp(A) \times Supp(B)}$$

More the lift more is the strength of the rule

- Steps for Apriori Algorithm

below are the steps for the algorithm :

- **Step-1** : Determine the support of itemsets in the transactional database, and select the minimum support and confidence.
- **Step-2** : Take all supports in the transaction with higher support value than the minimum or selected support value.
- **Step-3** : Find all the rules of these subsets that have higher confidence value than the threshold or minimum confidence.
- **Step-4** : Sort the rules as the decreasing order of lift.

Chapter 4

Block Diagram

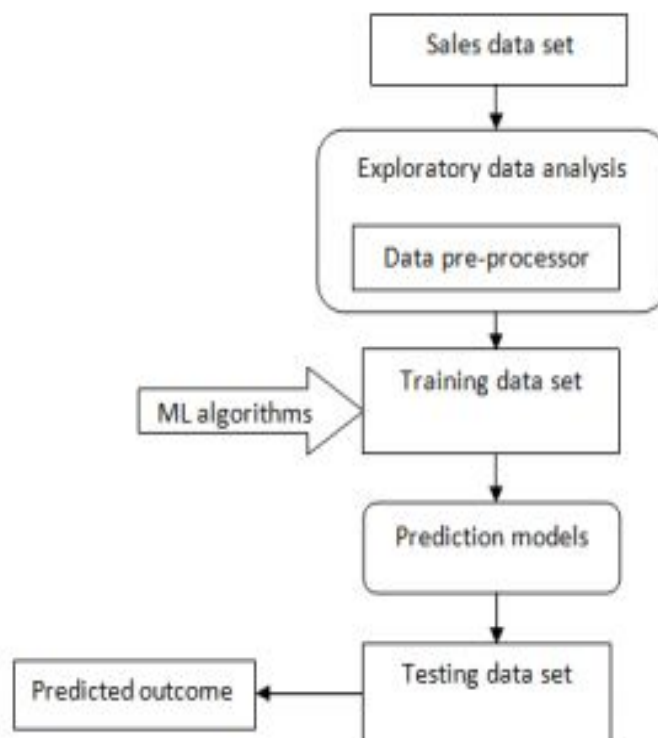


Figure 4.1: Block Diagram

Chapter 5

Screenshot of Output

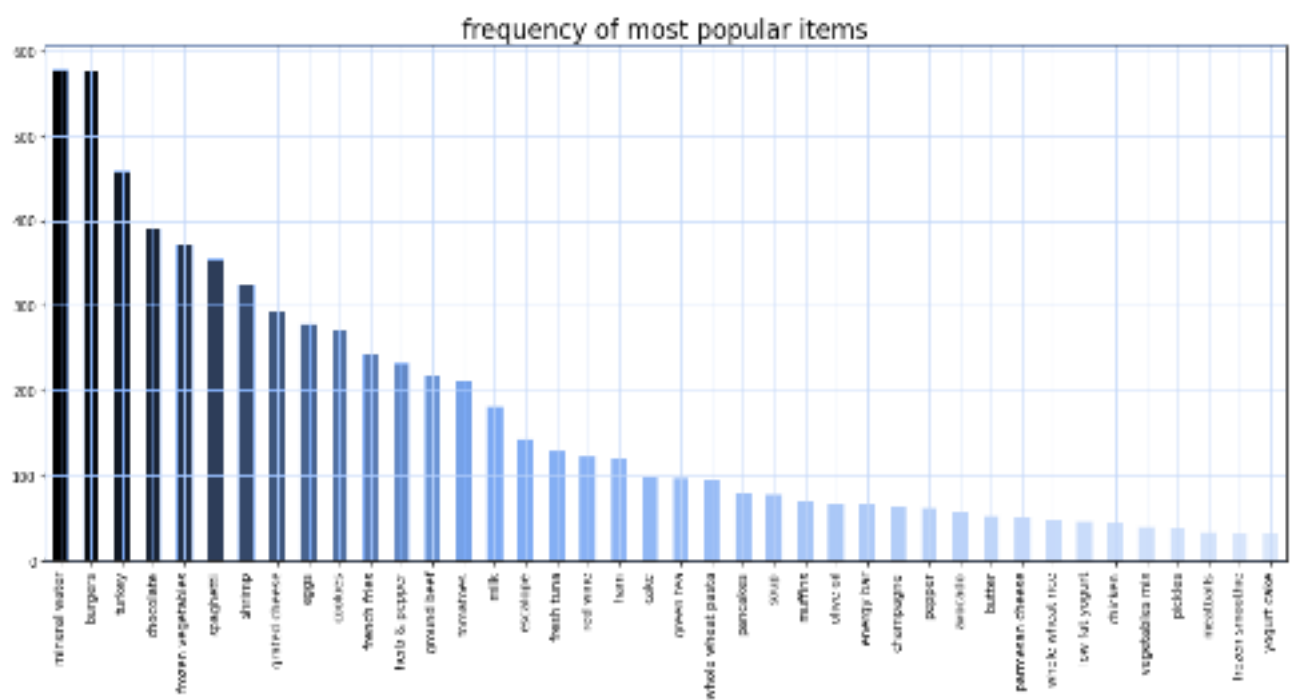


Figure 5.1: Output of the code

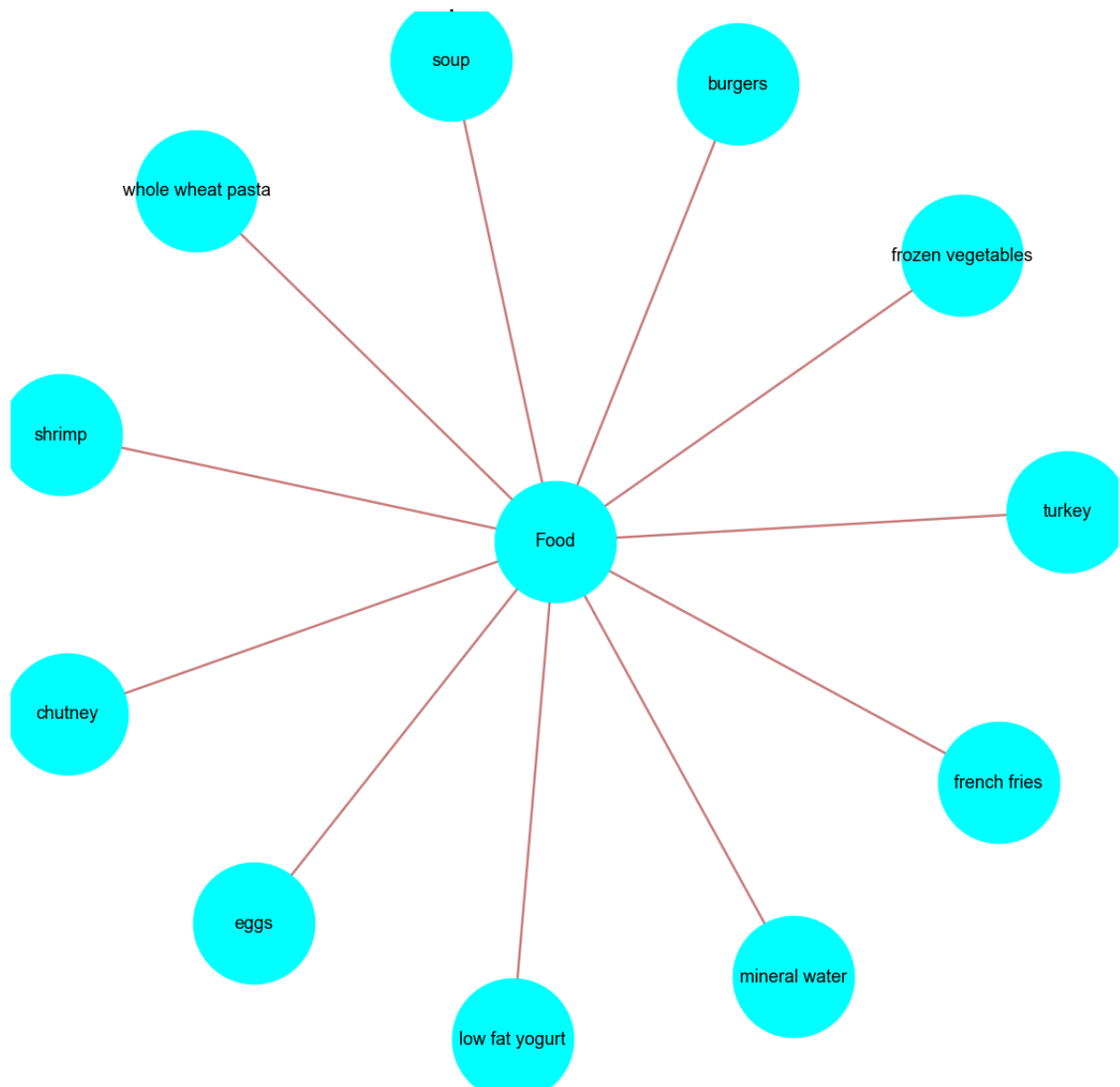


Figure 5.2: Top 10 First Choices

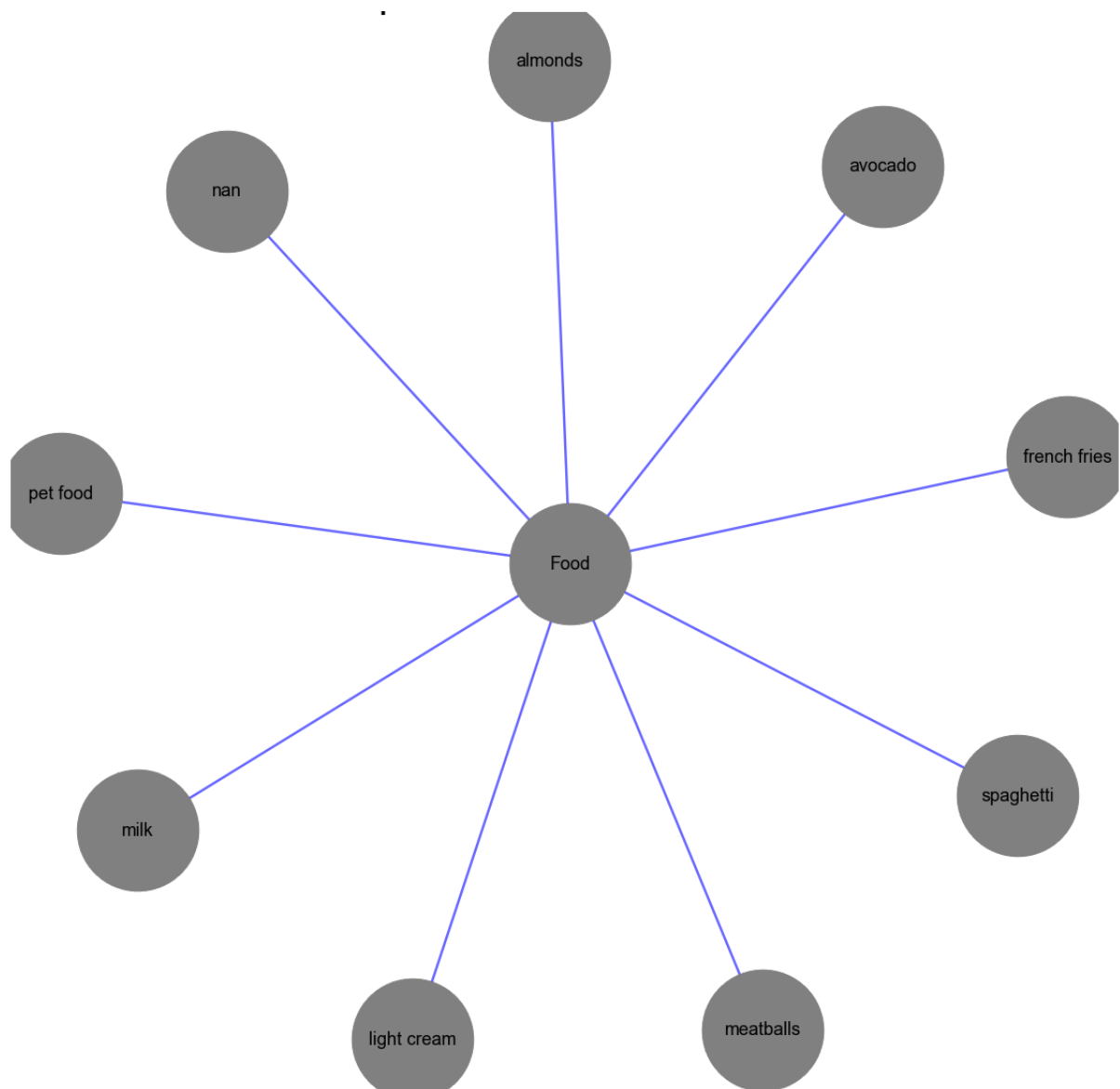


Figure 5.3: Top 10 Second Choice

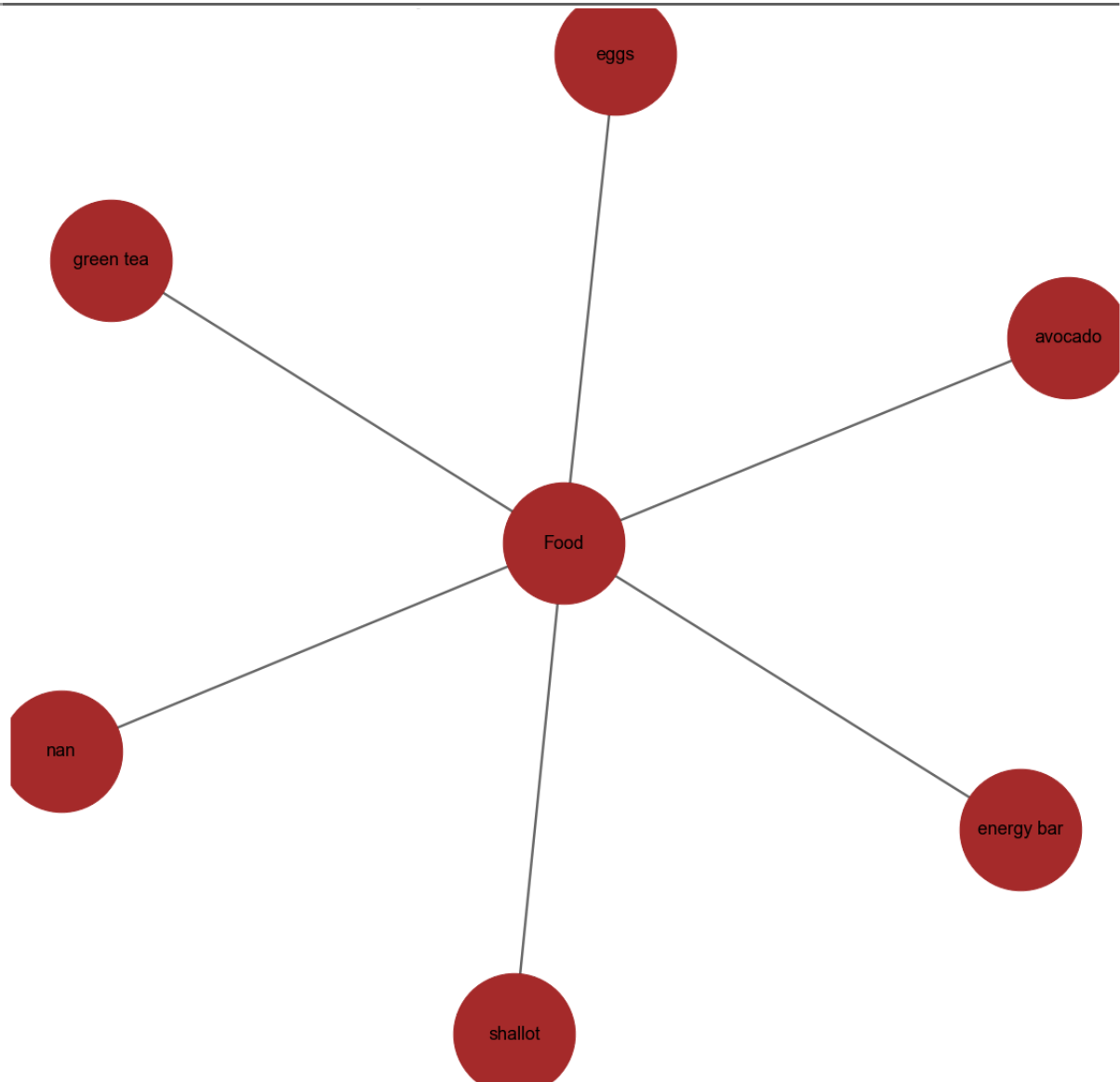


Figure 5.4: Top 6 Third Choice

Chapter 6

Outcomes

- Grocery sales data is analyzed and items are visualized using various graphs
- Grocery items sales are predicted using Apriori Algorithm
- This can help supermarkets to stack products accordingly to increase their sales
- This can help in various different business strategies

Chapter 7

Conclusion

Grocery sales prediction is an important part of the business strategic planning process. For predicting the basket items we applied Apriori Algorithm. Apriori algorithm gave the confidence values for the items which could be bought together.