

Market Basket Insights

INTRODUCTION

The retailer wants to target customers with suggestions on itemset that a customer is most likely to purchase. I was given a dataset containing data of a retailer; the transaction data provides data around all the transactions that have happened over a period of time. Retailer will use result to grow in his industry and provide for customer suggestions on itemset, we be able increase customer engagement and improve customer experience and identify customer behavior

DATASET

The data is obtained from <https://www.Kaggle.com/data>

COLUMNS USED

From Market_Basket.csv data the following columns are used

- Country name
- Item name
- Quantity
- Price
- Count
- Sum Price
-

LIBRARIES USED

The Python 3 environment comes with many helpful analytics libraries installed and several helpful packages to load.

The essential libraries used in this project are :

- Importing OS (for kaggle inputs)
- Numpy and Pandas libraries
- Matplotlib
- Seaborn

TRAIN AND TEST

Training the dataset by `isnull().sum()`, `drop()`, `show()`, and by using apriori algorithm we train the data.

Testing the data by importing `mlxtend.frequent_patterns` from apriori with ensuring the plot merging and sorting the values by displaying 3D plot by `mpld3` package.

REST OF THE EXPLANATIONS

Association Rule

The Association Rule is a tool used to establish associations between objects in a set, particularly in identifying frequent patterns in a transaction database. It helps retailers identify common customer purchases and relationships between items.

Data Preprocessing

The transaction dataset displays a matrix of items bought together, but no rules or purchase frequency is visible. To check the number of transactions, the R function `read.transactions` of the `arules` package is used to load the data into an object of the transaction class, with the data frame format being `basket`.

Strategy and Data description

- 1)Data Import
- 2)Data Understanding and Exploration
- 3)Transformation of the data
- 4)Running association rules
- 5)Exploring the rules generated
- 6)Filtering the generated rules
- 7)Visualization of Rule

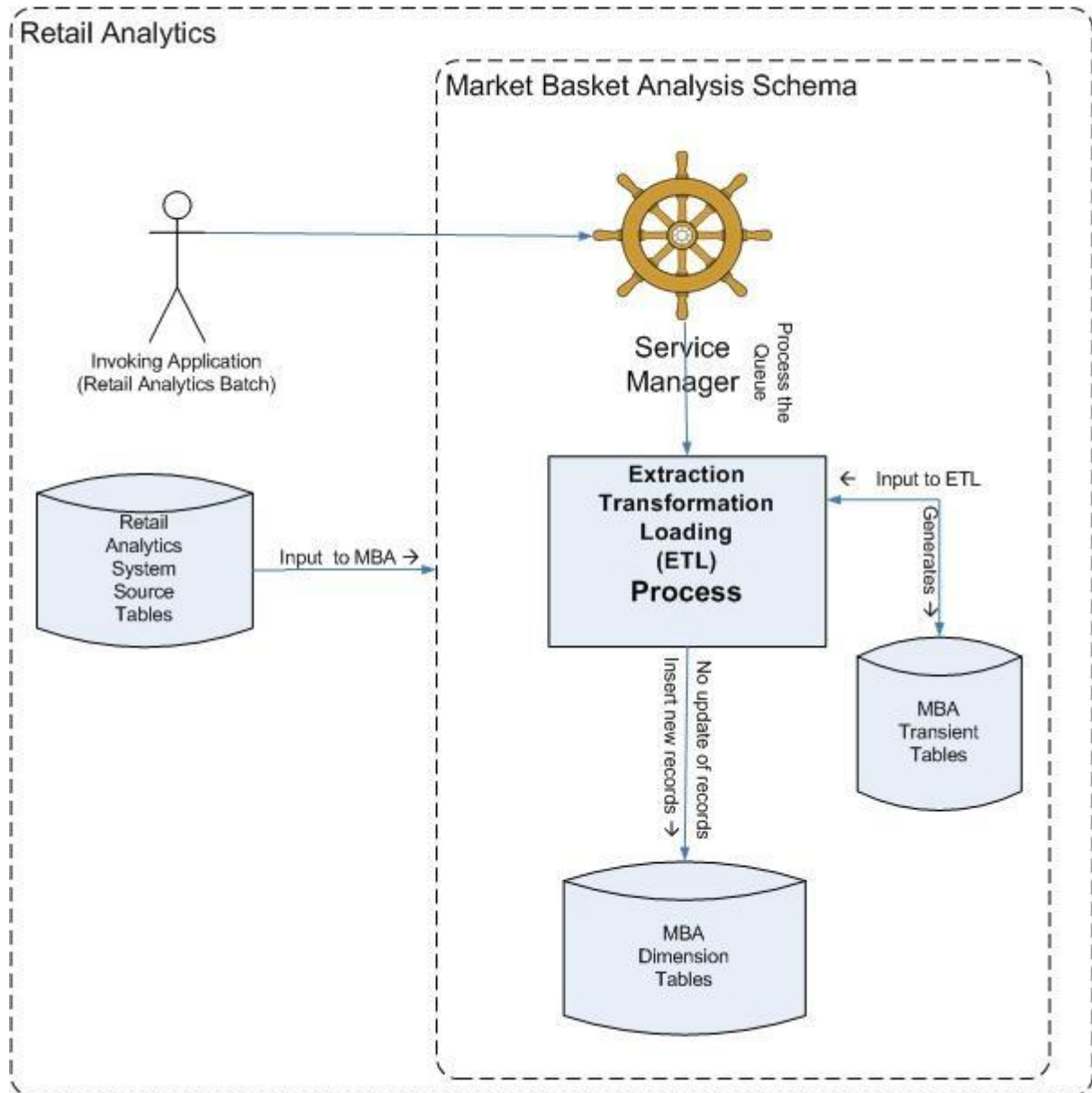
ALGORITHMS USED

Apriori algorithm.

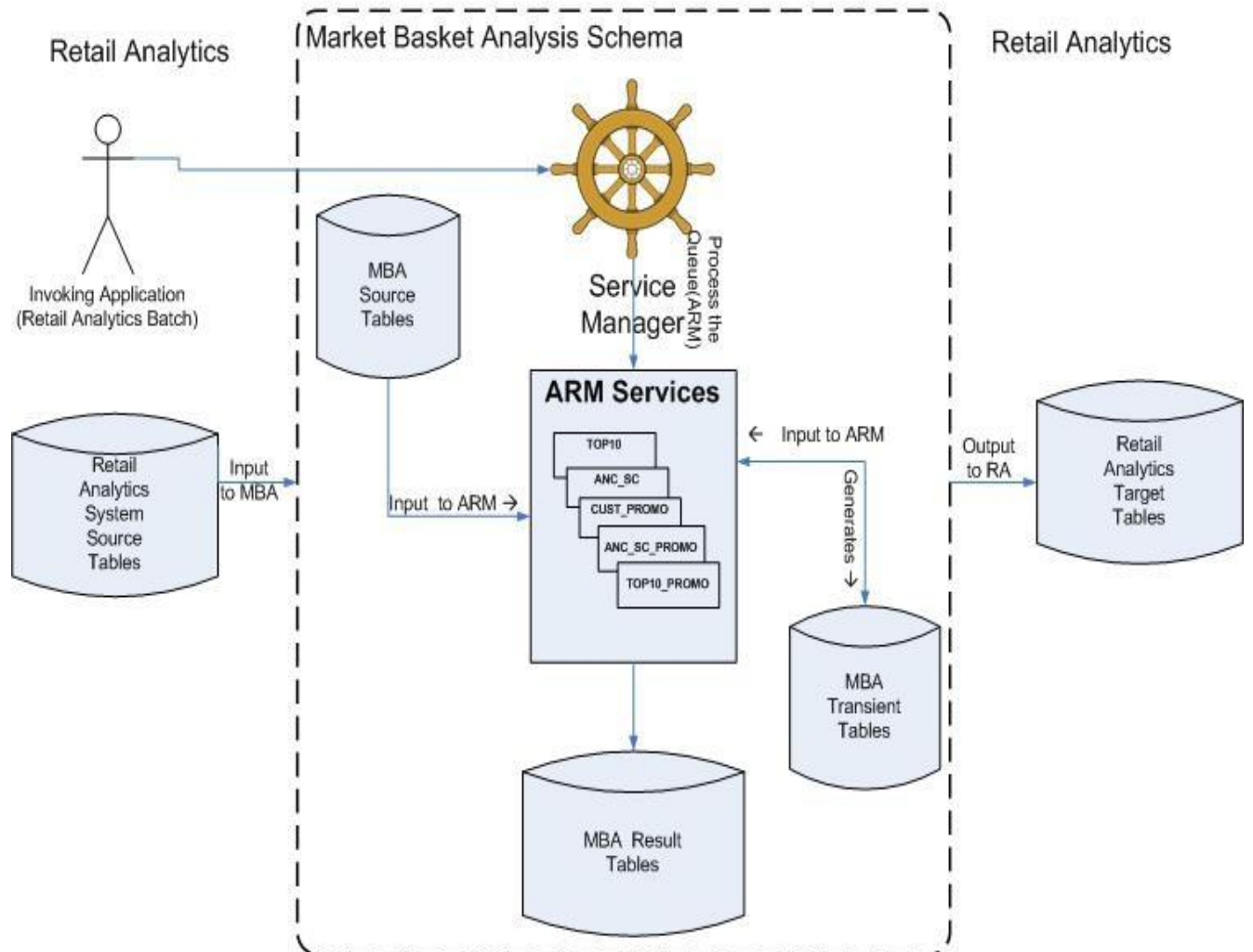
Libraries in are :

- `arules`
- `arulesViz`
- `tidyverse`
- `readxi`
- `plyr`
- `ggplot2`
- `knitr`

MBA ETL Process flow Diagram



MBA ARM Services Process flow Diagram



Retail Analytics DATABASE

