भारतीय प्रौद्योगिकी संस्थान जम्मू
INDIAN INSTITUTE OF TECHNOLOGY JAMMU
विद्याधनं सर्वधन प्रधानम्

Summer Project 2020
ISI Delhi

—

Under Guidance of: Prof. S K Neogy(ISI Delhi Centre)

By: Ayushi Verma(CSE, Sophomore, IIT JAMMU)

10th June 2020

OBJECTIVE: Market basket Analysis or Recommender System

A market basket analysis or recommendation engine is what is behind all these advices we get when shopping online or when we receive targeted advertising. The underlying engine collects information about people's habits and then provides the most suitable advices as recommendations to increase the sales.

For example: Let in a store, if people buy pasta and wine, they are usually also interested in pasta sauces. So, the next time you go to the supermarket and buy pasta and wine, be ready to get a recommendation for some pasta sauce!

So basically we have,

IF{pasta, wine, garlic} THEN{pasta-sauce}

The first part of the rule is called "antecedent", the second part is called "consequent". A few measures, such as support, confidence, and lift, define how reliable each rule is. The most famous algorithm generating these rules is Apriori Algorithm[1].

[1]. Agrawal and R. Srikant, *Proc. 20th Int. Conf. on Very Large Databases (VLDB 1994, Santiago de Chile)*, 487-499, Morgan Kaufmann, San Mateo, CA, USA 1994

RELEVANT RESOURCES INVOLVED

- Dataset extracted either from https://cseweb.ucsd.edu/~jmcauley/datasets.html or the ones available on Github Repositories.
- Using the Jupyter Notebooks to work on the project with the implementation in the Python language.
- Implementation of algorithms under the domain of Machine Learning as well as the theoretical concepts of Market Basket Analysis.

METHODOLOGY OF WORKING

- First there should be a thorough reading of the basic algorithms involved and then about the manipulation and handling of data.
- Then the dataset involved will has to be filtered such that only the useful information is left to proceed with.
- The relevant data will be filtered first and then by the Association rules[1] the code will be implemented.
- At last all the compilation along with the required visual plots will be done for the ease of understanding.

[1]. http://en.wikipedia.org/wiki/Association_rule_learning

CODING & IMPLEMENTATION PART

- The language I will be using in this project will be Python3.
- Python libraries like pandas, numpy etc. will also be involved as they help in processing the data smoothly.
- I am aiming to work with the data either publicly provided by InstaCart or the Amazon.com.
- The data contains the purchase history from the organization like the products bought together, product IDs, date of purchase etc.
- The relevant data will be filtered first and then by the Association rules[1] the code will be implemented.

BENEFITS OF THIS PROJECT

As we all know that the growth of the “online world” has been really rapid in the past few decades so to cope up with such large amount of data along with no compromising on the quality of the services delivered, it becomes really important that we put genuine efforts to help the population with better experience of technology.

Consumerism has played an important role in this growth as well and thus this project based on “Market Basket Analysis” is meant to deal with the growth of this convenient culture.

FURTHER VALUE ADDED

This project involves the motivation from the Apriori Algorithm and the success of the recommendation systems such as the one incorporated with Amazon.com[1] has been very evident in the recent times. These techniques can be applied to various other firms not only related to product based industries, the service based industries like the Netflix and Spotify also implement these recommender systems to increase their consumerism.

We humans are under the consistent need of advices so recommender systems are a great way to deal this situation. In Fact we can have such recommender engines for:

1. Suggestions for job profiles by recommendations from your previous interests and work history
2. Recommendations for educational institutes based on the interest of study courses and faculty and college/institute review.
3. Implementation of such systems by Tour and travel industries based on the customers search history and rating on the past trips.
4. By the stock investment firms by advising the favourable units to invest the money wisely.

[1]. http://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf

SOME MORE REFERENCES

1. Gentle introduction to Market basket analysis- Association rules
https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce

2. Market basket analysis of heterogeneous data sources for recommendation system improvement
https://www.sciencedirect.com/science/article/pii/S1877050918315680

3.Recommender system on Purchase Data
https://medium.com/datadriveninvestor/how-to-build-a-recommendation-system-for-purchase-data-step-by-step-d6d7a78800b6

4.Amazon recommendations secret to selling more online
https://medium.com/datadriveninvestor/how-to-build-a-recommendation-system-for-purchase-data-step-by-step-d6d7a78800b6


thank
YOU:-)