Market Basket Analysis

Final Project Report

Advanced Database Topics

Group Members: The Brogrammers

Henishkumar Kiritkumar Patel (110073489)

Jaydeep Pritesh Dharamsey (110088507)

Yash Somaiya (110087733)

Veera Venkata Bharat Kumar Vayitla (110088432)

Submitted to:

Dr. Olena Syrotkina

1. Abstract

The culinary sector of the food and beverage industry is expanding extremely quickly. Restaurant owners must always be ready to innovate and remain able to meet consumer needs through products that can attract customers as well as decide on strategies and promotions that can boost sales to stay in a very competitive industry. This forces food and beverage business owners, especially restaurants, to make the right decision to stay in business. Information from stored transaction data, such as client purchase patterns and sales trends, can be extracted using data mining techniques. The owner of the restaurant must carefully choose the dishes to put on the menu in order to draw in more customers. Consequently, a platform that allows users to forecast foods based on previously entered dishes would be helpful for both restaurant owners and the general public.

2. Introduction and Motivation

The business environment is changing quickly in the restaurant industry. The advent of numerous restaurants with unique features or novel concepts serves as a sign of this. Numerous countries throughout the world continue to achieve high growth rates for the food and beverage industries. Due to the increasing expansion of the restaurant industry, there is currently competition among restaurant business owners. As a result, restaurant owners must be well-prepared for this competition, which spans everything from restaurant concepts to product innovation and operational systems. Therefore, in order to manage restaurant operations, restaurant owners must think carefully about creating an operational system that will enable them to thrive and profit in line with their restaurant's goals.

One of the reasons for the growth of the restaurant industry is that the current food and beverage industry, particularly this restaurant, promises enormous profits and views people's need to eat and drink as a necessity that must be met in order to exist. This is a chance that can be used to start a successful business. It makes sense that there are so many restaurants, each with a unique concept, set of operating procedures, creation with a unique product innovation, or concept that can be demonstrated to clients.

The ability to conduct operations while achieving operational goals in the business world competitiveness involves a strategy and system that may support and enhance restaurant aims. Understanding the value of preparation or developing a good system. Restaurant owners still need to consider several things while developing this plan, such as the items that will be sold, product advancements, and consumer interests in marketing and product promotion. It is difficult to increase consumer interest in a product before they decide to purchase it.

The intent of this software is to recommend a various dishes that is popular and is trending now based on the order history of people. The goal is to provide a tailored list of dishes based on previously entered dishes in order to meet people's cravings and eateries' fierce competitiveness. Small eateries can use the system's findings to boost their productivity and chances of succeeding in their industry to provide combo of items that most of the people can eat together.

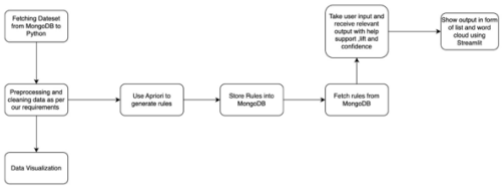
3. Related work

In one of the research-papers, we found during our search, the author created an app that suggests groceries based on user preferences and previous purchases. Also, the app suggests meals based on the time of the day.[1]

Much research has been done that describes the application of data mining techniques for recommending products to users based on their preferences [2].

One of the major drawbacks we found is that the author used SQL as a database to build the system. Nowadays, SQL can become a significant barrier for a business owner looking for a time- saving system. According to our point of view, NoSQL databases like MongoDB, HBase can be an effective substitute for SQL. We shall attempt to get around SQL's restriction.

4. Proposed Model



We have described our steps in the figure for our proposed model. We gather data from two different sources that are closely related but not completely identical, so we load both the dataset from our MongoDB cluster. We then clean the data as per our requirement like having identical column names, datatypes etc.

After cleaning the data, we create some visualizations which can be used in future work such as sales prediction or inventory forecasting. Now we are using apriori algorithm to generate rules according to the support, lift, and confidence from the itemset. We store the rules in MongoDB.

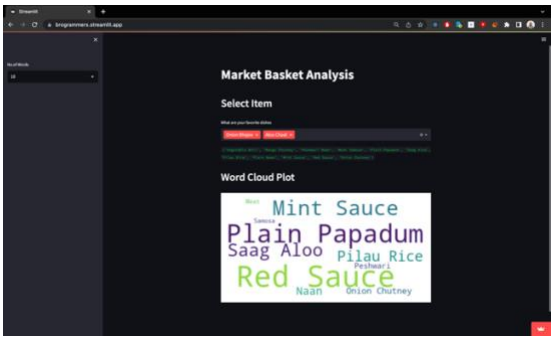
Now when the user asks for recommendations, we fetch the rules from mongoDB and give output which is the best result and mostly bought together with the input item. We transfer all this data to our streamlit app. And show results in form of a wordcloud and a list. The wordcloud shows the items most likely to be bought in bigger size than the ones unlikely.

Major functionalities of our project:

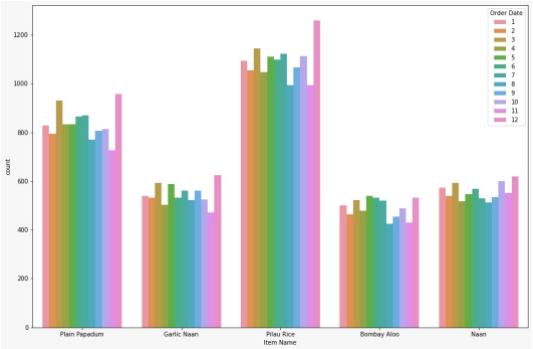
- **User Interface: A** web page will serve as the user interface. Users need to enter the dishes name they want to check.
- **Proper Forecasting:** According to the user's input, the algorithm must accurately anticipate the related food dishes that can be ordered together.
- **Database:** The rules generated by apriori are stored in MongoDB.

5. Figure/diagram or result

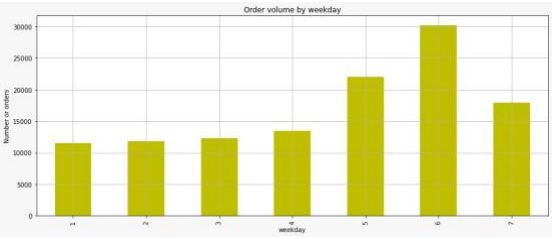
Result: We deployed our association rules using Streamlit. Following snapshots show the final version of our product. Where user just need to enter their favorite dishes and our model using association rules will give word cloud plot of recommendation of dishes based on input from users.



Figures: Top sold items in every month (Fig 1) and Number of orders based on weekdays where 1 is Monday and so on (Fig 2), this data can be helpful for the useful for the businesses to manage their inventory.



(Fig 1)



(Fig 2)

6. Limitations or Challenges

Limitations:

- Firstly, Scalibility can be issue in near project. Currently, the dataset is not large enough and we can able to host database in mongoDB Atlas. But in near future, if storage requirements will grow and as we have to incorporate more datasets, then we have to looks for some paid cloud storage. The commercial versions of powerful hosting platforms like AWS and Heroku can help with this, though.
- Second, Dishes is influenced by current market trends. It varies in accordance with the market and follows it, and our project is unable to examine it because of how time-dependent it is.

Challenges:

- We attempted to create an association with the dataset with a support level of 0.02 but got no results. But to get the required result, we have to adjust the support rule in the query to 0.003.

7. Conclusions or Future Work:

Conclusion: Based on our model, users and restaurant operators can forecast the dishes based according to the highly co-related selected dishes.. However, several additional characteristics, such vegetarian, non-vegetarian, and vegan, are significant but are not included in any dataset. This kind of data will be gathered and applied to the system more frequently as the industry develops, setting it apart from previous prediction systems.

Future Work: By extending the dataset, the study can be expanded in the future. Deep neural networks may be effective with greater datasets, allowing for a more thorough analysis of dishes forecast. To provide studies of how a certain dish will perform in the near future, it is also possible to add a feature called sales forecasting. Additionally, a tool called Inventory Predection can be added to Forecast to assist Businesses with Efficient Inventory Management.

8. References

[1] F.-F. Kuo, C.-T. Li, M.-K. Shan, and S.-Y. Lee, "Intelligent Menu Planning," *Proceedings of the ACM multimedia 2012 workshop on Multimedia for cooking and eating activities - CEA '12*, 2012.

[2] "Automated menu recommendation system based on past preferences - thesai.org." [Online]. Available: https://thesai.org/Downloads/Volume5No7/Paper_11-Automated_menu_recommendation_system_based_on_past_preferences.pdf. [Accessed: 27-Nov-2022].

[3] "Apriori algorithm," *GeeksforGeeks*, 13-Jan-2022. [Online]. Available: https://www.geeksforgeeks.org/apriori-algorithm/. [Accessed: 27-Nov-2022].

9. Contribution Table

Contribution	Details	Names	Student ID
25 %	Mongo Connection (Cloud) Data Cleaning	Henishkumar Kiritkumar Patel	110073489
25 %	UI (Streamlit) + Hosting	Jaydeep Pritesh Dharamsey	110088507
25 %	Associations Rules	Veera Venkata Bharat Kumar Vayitla	110088432
25 %	Data Cleaning Data Visualizations	Yash Somaiya	110087733