

Reporting Date: 15th March 2024

TABLE OF CONTENTS

[ABSTRACT 2](#_Toc162017460)

[INTRODUCTION 3](#_Toc162017461)

[DATASET AUGUMENTATION 3](#_Toc162017462)

[CONTENT BASED RECOMMENDER SYSTEM 5](#_Toc162017463)

[IDENTIFIED BLOCKERS 6](#_Toc162017464)

[CONCLUSION 7](#_Toc162017465)

[REFRENCES 8](#_Toc162017466)

[APPENDIX 9](#_Toc162017467)

# ABSTRACT

The target of this week's project was personalized dietary mixtures by the incorporation of dietary preferences and thus informing each user with their own meal plan. One of the major tasks was dietary preference data preprocessing that we included in the main dataset and used the NLTK to match the ingredients and updated ingredient prices. The architecture of our content-based recommender system ensured the analysis of the customer's preferences and a creation of personalized recommendations. With the contribution of TF/IDF analyzer and Vector Space Model, we made sure that our model displayed the results that matched perfectly with users' needs without having a big history in user's preferences All in all, our efforts have powered an optimized recommendation system which inspires the users to gain from the recommended personalized meal options based on their health desires.

# INTRODUCTION

AutoBasket is a company founded by Larry and Veronica Smiles with the purpose of simplifying the grocery shopping experience for individuals and families. Situated in the heart of Toronto, the company's headquarters pulsate with the energy of a city known for innovation and diversity. The company's app automates grocery and recipe lists, linking recipes to required products, and streamlining the shopping process. Today, AutoBasket assists households across the world to save time and energy on their weekly grocery runs, making it easier for busy families and individuals to get everything they need for a great meal.

AutoBasket has hired us as interns to focus on all areas of development for the organization and to help create solutions to common issues within the industry. Our role as interns is not only focused on the knowledge you have learned in school but also on developing your soft skills, including presentations, teamwork, and leadership. This internship will provide us with valuable real-world experience and an opportunity to contribute to the continued success of AutoBasket.

# DATASET AUGUMENTATION

To apply individual preferences based on their dietary needs which we have improved by adding an extra column to our originally collected dataset. This supplementary column can propose meals that suit the demands of every individual user. First, these columns were integrated into the main dataset using data preprocessing techniques.

We list down the vast majority of the dietary preferences followed by the users in users table, and store them in a list. Lastly, we filtered the main dataset columns that indicated dishes and their ingredients in the aggregated manner. Subsequently, we used NLTK to see if the list of ingredients exists in the dataset or not.

A screenshot of a computer

Description automatically generated

Figure-1: Updated user details data

A screen shot of a computer

Description automatically generated

Figure-2: Count of dietary preferences

If the ingredients are found listed in the main dataset, they are flagged by using binary indicators in separate columns such as egg, meat, peanut, halal. The symbolic binary columns shows the main dietary preferences included in the recommendation.

A screenshot of a computer program

Description automatically generated

Figure-3: Egg presence recipe dataset

We have also updated the price of each ingredients in our dataset to have a better representation of those ingredients according to the Category they belong too.

A screenshot of a menu

Description automatically generated

Figure-4: Updated ingredients price

# CONTENT BASED RECOMMENDER SYSTEM

A specific kind of recommendation engine that works by examining an item's intrinsic qualities and properties in order to provide recommendations to users is called a content-based recommender system. These systems create profiles for persons and things by analyzing variables such as genre, keywords, metadata, and other descriptive elements, instead of depending solely on previous user interactions. By mapping user preferences to objects with similar content qualities, the system may make personalized recommendations.

Users provide data in two ways: directly by rating items and indirectly by clicking on links. After that, a user profile is constructed using this data, and that profile acts as the basis for the creation of tailored suggestions. The technology improves the user experience by gradually improving its accuracy as users interact with these recommendations.

Important ideas like Term Frequency (TF) and Inverse Document Frequency (IDF) are used by content-based recommender systems. Word frequency within a document is measured by TF, whereas term rarity throughout the collection of documents is determined by IDF. These methods can more efficiently ascertain the value of terms in documents by merging TF and IDF into TF-IDF scores, which use logarithmic functions to lessen the influence of high-frequency terms.

The algorithm evaluates the similarity between items or their relevance to a user profile after calculating TF-IDF scores. The Vector Space Model does this by using the angle between the vectors that represent the items and the user profiles to determine closeness. This method makes recommendations that are more accurate and useful by making sure they are based on the user's preferences as well as content similarity.

Moreover, content-based recommender systems are capable of managing situations in which user history is either absent or restricted. They are especially helpful in scenarios when users are unfamiliar or have a variety of preferences that are not adequately reflected in previous interactions because of their independence from previous user interactions. These systems can get around some of the drawbacks of collaborative filtering techniques by emphasizing the inherent qualities of objects and the expressed preferences of users, providing a more customized and individualized recommendation experience.

Furthermore, content-based systems may handle a wide variety of content, including items, locations, and movies in addition to articles. Their capacity to evaluate and comprehend content attributes enables them to offer suggestions in a variety of fields, resulting in a smooth and captivating user experience. These technologies have the potential to significantly improve user interactions and increase engagement across a broad spectrum of applications and industries as machine learning and natural language processing techniques continue to grow.

In addition to that, the content-based systems have been seen to be involving complex content such as objects, places and clips as well as numerous articles. Adding to their capability to analyze and understand content features, their suggestions are very influential in various fields and results in a great and interesting experience of a user. Those technologies could provide a huge boost to audience interaction about the range of applications and industries due to increase of machine learning and natural language processing techniques.

# IDENTIFIED BLOCKERS

A black screen with white text

Description automatically generated

Figure 5: Specific user details

However, the most important task we have completed is learning about the architecture of our content based recommender system. This is an image of the architecture module that makes up the recommender system. Based on the system shown in the figure above, we have proceeded with the research and determined which relevant steps on our system will be applicable to the user interaction templates that we will implement in our website. By it, we will not only develop custom and correct AutoBasket users data but also will form the data of their interaction with website exclusively. We are thinking of different methodologies to predict the user demand and help our system develop a content whatchable content recommendation system.

# CONCLUSION

In a nutshell, there is a bright future for a content-based recommender system since it is considered a significant step to fit the users’ preferences and dietary demands. The insertion of an extra column to the dataset specifically dealing with dieting from our side has upgraded the recommendation process and at that users receive dishes with their dietary preferences in mind. Trying out the component such as NLTK for similar ingredient and prices update provides the pertinence and relevance of the rating which eventually improves the user satisfaction.

Furthermore, the architecture of such a system creates a stable base for conducting user interactions and performing a recommendation process. The vectors space model is used in analyzing the user's preferences and therefore the system can provide with personalized suggestions even when user's history is inadequate. By removing the obstacles and always improving the system based on user reactions and conversations, we propose to design an environment that is both engaging and friendly, one that attracts more clients and brings higher satisfaction than before. These pursuits help us build a solid recommendation system that will be both effective and efficient in order to deliver both useful and profitable recommendations to our users.

# REFRENCES

1. (2023, August 31). Retrieved from analyticsvidhya: https://www.analyticsvidhya.com/blog/2015/08/beginners-guide-learn-content-based-recommender-systems/#:~:text=Content%2Dbased%20recommender%20systems%20are,mapping%20it%20to%20users'%20preferences
2. *introduction-to-recommender-systems*. (2020, July 29). Retrieved from towardsdatascience: https://towardsdatascience.com/introduction-to-recommender-systems-1-971bd274f421

# APPENDIX