Cake Recommendation Using Machine Learning

Yash Adulkar[1*], Parth Doshi[2]

[1,2]*Department of Information Technology, Atharva College of Engineering, Mumbai, India*

Abstract: **This paper presents implementation of the Cake Recommendation System using Machine Learning.**

Keywords: **cake recommendation system, machine learning.**

1. Introduction

A recommendation system helps users search and choose products from the huge data available on the internet. Similarly, a cake recommendation system provides a level of comfort and personalization that helps the user interact better with the system and order cakes that cater to their needs. Providing this level of comfort to the user is the primary motivation in opting for the cake recommendation system as our Project. The chief purpose of our system is to recommend cakes to its users based on their order history and the feedback that they provide. Collaborative filtering and content-based filtering are the principal approaches used in recommendation systems. They both have advantages in a certain situation with inclusive disadvantages due to which we have proposed a hybrid system consisting of both the systems contradicting their disadvantages and thus improving the performance and accuracy of our system

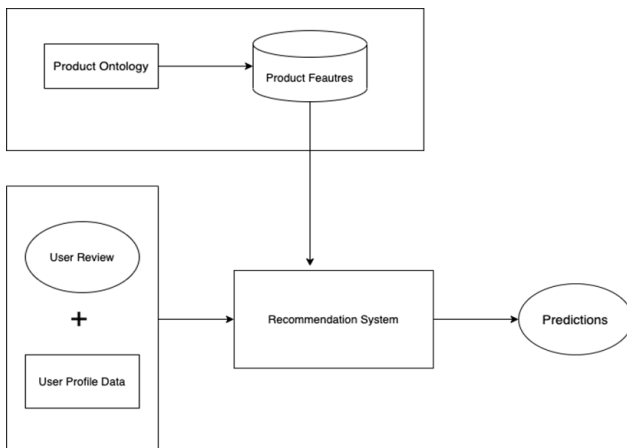## 2. Filtering System

A. Content-based Filtering

Fig. 1. Content-based Filtering

In content-based filtering, data are recommended based on user profile history and data profile. A user profile is a set of assigned keywords got from the algorithm from products by the user. A set of keywords of data is the data profile. For example, consider a person go to buy his favorite cake 'X'.

Unfortunately, cake 'X' is not available, so the owner recommends the buyer to try cake 'Y', which is made up of similar ingredients to cake 'X'.

Fig. 1. displays the flow of a content-based recommendation system. Similar entities of data are kept together as input. The main content of data is extracted from data ontology. The contents of data are applied to the content-based recommender system.

Advantages of content-based filtering:
a) They capable of recommending unrated data.
b) It gets the user's interests and recommends products to specific users.

Disadvantages of content-based filtering:
a) It does not work for new users, who do not have any history yet.
b) No recommendation of serendipitous data.
c) The system processes the users rating and reviews.

B. Collaborative Filtering

A collaborative filtering system recommends data based on similarity measures between users and their needs. The system recommends data preferred by similar users. This is based on the situation, where a person asks his friends, who have similar tastes, to recommend him some cakes.
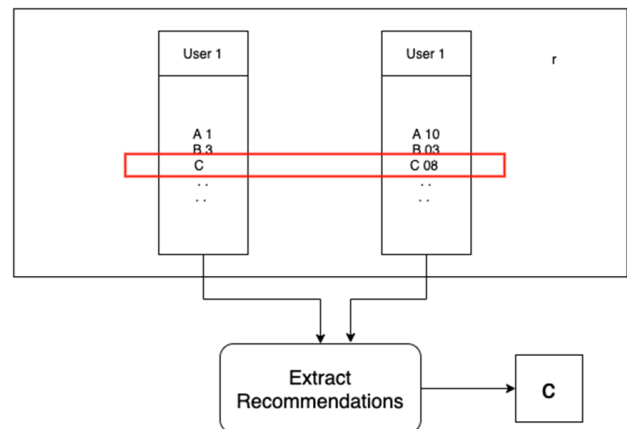

Fig. 2. Collaborative Filtering

Fig. 2. It displays the flow of the Collaborative filtering recommendation system. It shows how collaborative filtering considers only numerical feedback given by users and recommends new data. The user reviews are stored in a backend database to make further references and predictions. In the

---

*Corresponding author: yash151099@gmail.com

figure, User 1 and User 2 show similar interests. Using this similarity, a review about a product not rated by User 2 can be predicted using the reviews of User 1. Thus, a prediction regarding product B's review by User 2 is made using the available data. From these results, recommendations are extracted and suggested to the user.

Active filtering is separated from passive filtering because, in Active filtering, it is required for the user to dedicate time to review the products, while in passive filtering, the user provides data by only accessing the product.

1) Active Filtering

Active Filtering is a method for collaborative filtering because of its peer-to-peer similarity approach. Various profiles from peers are checked to extract similar interests. This approach is based on the fact that people exchange information such as appreciation of specific products.

Advantages of Active Filtering:
 a) Rating is collected from a real person.
 b) Highly relevant information is provided

Disadvantages of Active Filtering:
 a) The data is expensive and rare.
 b) Chances of a review being biased.

2) Passive Filtering

Passive filtering is a process to collect information implicitly from the user. Few examples are: purchasing a product, commenting repeatedly on a product, Referring to a website.

3) Product-based Filtering

In this approach, products are reviewed and used as parameters for similarity rather than users. These products are separated and then grouped and are proposed to users. Users can then compare and review them. The preferences allow grouping users by similarity.

Advantages of collaborative filtering based systems:
 a) It is content-independent.
 b) Real quality assessment of products by considering other people's experiences.

Disadvantages of collaborative filtering are:
 a) These systems cannot provide recommendations for new products since there are no user ratings on which to base a prediction.
 b) For the system to work, a group with similar characteristics is needed.
 c) The amount of products exceeds the number of users.

3. System Description

Due to the disadvantages of content-based and collaborative filtering, we have to use a mixture of both i.e. hybrid filtering also known as a content-based collaborative filtering system. In the hybrid system, the use of the content, as well as its relations and similarities among users, are used. The instinct behind this technique is to use the content predictor and the user rating matrix. The dataset is collected by the different pastry centers and then filtered according to the model and stored in the backend database. The dataset consists of user reviews about different cakes and the location of the cakes along with their ingredients.

A. Hybrid Algorithm

Step 1: Use content-based data to calculate the pseudo-user-rating vector 'v' for every user 'u' in the database.

$v_{u,\,i} = r_{u,i}$: is user u rated product i

$v_u, = r_{u,i}$: otherwise

Step 2: Check the similarity of all users with the active users.

1. The similarity between users is measured as the Pearson correlation between their rating vectors.

Step 3: Select 'X number of users based on similarity.

1. These users form a group.

Step 4: Predict from the weighted combination of the selected group ratings.

In step 2, the similarity between two users is computed using the Pearson correlation coefficient:

$$P_{a,u} = \frac{\sum_{i=1}^{m} (r_{a,i} - \bar{r}_a) \; X \; (r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i=1}^{m} (r_{u,i} - \bar{r}_a)^2 \; X \; \sum_{i=1}^{m} (r_{u,i} - \bar{r}_a)^2}}$$

Where,

r_a, i is the rating given to product I by the user a,

$\bar{r}_a$ is the mean rating given by user a, m is the total number of items.

In step 4, predictions are computed as the weighted averages of deviations from the group mean:

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{n} (r_{u,i} - \bar{r}_u) \; X \; P_{a,u}}{\sum_{u=1}^{n} P_{a,u}}$$

Where,

$p_{a,i}$ is the prediction for the active user a for item I,

$P_{a,u}$ is the similarity between users a, and u, n is the number of users in the group.
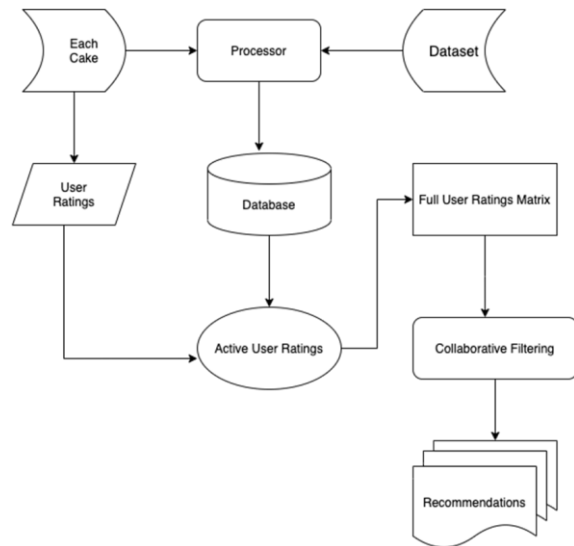

Fig. 3. System Overview

4. Conclusion and Future Work

This approach overcomes the drawbacks of Content-based and collaborative filtering. Individually both have their

disadvantages which contradicts the use of the proper recommendation system at the industry level. With this approach, this system could be used in the industrial market by the cakes startup and pastry centers to provide users with their needful products.

References

[1] James Bennett, Stan Lanning, "The Netflix Prize", In KDD Cup and Workshop in conjunction with KDD, 2007.

[2] Mirza Illahami, Suharjito, "Film Recommendation System using Matrix Factorization and Collaborative Filtering."

[3] Marwa Hussien Mohamed, Mohamed Helmy Khafagy, Mohamed Hasan Ibrahim, "Recommender Systems Challenges and Solutions Survey" presented in 2019 International Conference on Innovative Trends in Computer Engineering, Aswan, Egypt, 2-4 February 2019.

[4] Taiping Lai, Xianghan Zheng, "Machine Learning-Based Social Media Recommendation" presented in the College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China.

[5] Celeste Murnal, Deepika "Movie Recommendation System using AI and ML", In International Journal of Research in Engineering, Science and Management Volume 2, Issue 11, November 2019.