



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 6, June 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com



Coffee Recommendation Using Machine Learning

Ketki Bhirud¹, Sakshi Dange², Prajakta Hon³, Diksha Sonawane⁴, M. B. Gawali⁵

U.G.Student, Department of Information Technology Engineering, SRES Sanjivani College of, Engineering,
Kopargaon, India¹⁻⁴

Associate Professor, Department of Information Technology Engineering, SRES Sanjivani College of Engineering,
Kopargaon, India⁵

ABSTRACT: Now a days, implementation of different software applications is increasing user engagement. Social influence plays an important role in product marketing. Over the social-media online user groups are created and share their experiences, interests and views with each other. For providing better service to users and growing in business, there is a constant need to analyze user interest, need, preferences, and habits. This system is built on computational models for personalized Coffee recommendation using user reviews and choices. Coffee has become intrinsic part of our everyday lives. However, serving an excellent cup of coffee is given less significance as a coffee blend typically comprises rich aromas, indulgent and unique flavours and a lingering aftertaste. Our work addresses this by proposing a computational model which recommends optimal coffee beans powder resulting from the user's preferences. Specifically, given a set of coffee bean properties objective features, we apply different unsupervised learning techniques to predict user desired coffee.

KEYWORDS: Recommendation System, Collaborative Filtering, PCA, Cosine Similarity

I. INTRODUCTION

Recommender Systems are used as a computer-based intelligent technique, designed to help people with the problem of excess information and product overload. The purpose is to generate proficient personalized solutions in e-business domains, benefiting both the customer and the merchant. The basic entities in recommender system are the user and the item. Users are the one who utilize the Recommender System and are required to share their opinion about past items. The motive of the Recommender System is to generate suggestions about new items for particular users. It achieves that by applying the selected filtering algorithm on the input provided, which is usually expressed in the form of user ratings on items. The primary aim of the application is to suggest users the best coffee to have based on their coffee feature preferences. The application is targeting everyone who wishes to explore vast varieties of coffee. Coffee recommendation system works based on collaborative filtering. This application takes the coffee preferences and ratings into consideration to recommend to the users. In this collaborative filtering method uses item based and user-based filtering methodologies.

II. RELATED WORK

Mr. N. Varatharajan, J. Guruprasad, K. Mathumitha in their paper titled 'Restaurant Recommendation System Using Machine Learning Model' [1], introduced the primary aim of the application is to suggest users the best food to have on the given food preferences. The application is targeting everyone who wishes to eat food in restaurant. Restaurant recommendation system works based on collaborative filtering. This application takes the food preferences and ratings into consideration to recommend food to the users. In this collaborative filtering method uses



item based and user-based filtering methodologies. In this paper, according to the characteristics of restaurant recommender systems, the restaurant recommendation is done based on improved collaborative and content-based filtering method is proposed to be analyzing the end user's behaviors.

Praphula Kumar Jain, E. A. Yekun, R. Pamula, Gautam Srivastava in their paper titled 'Consumer Recommendation Prediction in Online Reviews using Machine Learning' [2], introduced the digital technology and social media have delivered many advantages in understanding human psychology, which is essential to industrial growth. In online reviews, consumer recommendations are critical indicators for service providers to improve their consumer policies as well as service quality. It is also helpful for future customers to help get information on future purchases prior to making them. Therefore, previous consumer recommendations based on online reviews play a vital role in airline recommendations. Our main goal is to use our proposed cuckoo optimized machine learning model to predict airline recommendations.

Rachiel Yera Toleda, Ahmad A. Alzahrani, Luis Martínez in their paper titled 'A Food Recommendation System Considering Nutritional Information and User Preference'[3], introduced that personalized nutrition emerges as a new research field for providing tailored food intake advices to individuals according to their physical, physiological data, and further personal information. Specifically, in the last few years, several types of research have proposed computational models for personalized food recommendation using nutritional knowledge and user data. This paper presents a general framework for daily meal plan recommendations, incorporating as main feature the simultaneous management of nutritional aware and preference-aware information, in contrast to the previous works which lack this global viewpoint. Furthermore, it incorporates an optimization-based stage for generating a daily meal plan whose goal is the recommendation of food highly preferred by the user, not consumed recently, and satisfying his/her daily nutritional requirements.

Ivens Portugal, Donald Cowan, Paulo Alencar in their paper titled 'Use of Machine Learning Algorithms in Recommender Systems :A Systematic Review'[4], introduced that recommender systems use algorithms to provide users with product or service recommendations. Recently, these systems have been using machine learning algorithms from the field of artificial intelligence. However, choosing a suitable machine learning algorithm for a recommender system is difficult because of the number of algorithms described in the literature. This paper presents a systematic review of the literature that analyzes the use of machine learning algorithms in recommender systems and identifies new research opportunities. The goals of this study are to (i) identify trends in the use or research of machine learning algorithms in recommender systems; (ii) identify open questions in the use or research of machine learning algorithms; and (iii) assist new researchers to position new research activity in this domain appropriately.

III. METHODOLOGY

Recommender systems are information filtering systems that provide a solution to the problem of information explosion. We address the challenge of tailoring the choice of coffee to the user's preferences. To do so, we propose a computational model which relies on unsupervised learning techniques. Given that the coffee review process is expensive and requires expert domain knowledge, it is essential to consider that a portion of certain coffee beans have been reviewed. The unavailability of official reviews is indeed a practical concern in the coffee industry, as it occurs every time a novel blend is introduced or in case of local and niche products. It should not happen that, while providing the user with the best experience, a recommender system should also be able to suggest unreviewed coffee in case the expected coffee qualities match the user's preferences. We aim for a system supporting this functionality in the context of an informative dataset that we can use for coffee recommendation. We implemented this system on Windows 10 or later versions with i3 processor, 1GB RAM and the programming language used is python on Jupiter notebook.



The data collection phase is an important step for the accuracy of the recommendation engine. It is helpful to generate models for making recommendations. For proper construction of recommendations that are responsive of a user's preferences, a recommendation engine relies on different types of inputs such as explicit user input, which is directly specified by a user and represents a user's interest for a particular item or implicit user input, which is captured from a user's interaction with the system. We rely on an existing open-source dataset from coffeereviews.com. By inspecting the correlation between the selected features, we ended up with 17 attributes. These are species, country of origin, region, variety, color, category one defects, category two defects, processing method, broken down into objective coffee properties and aroma, flavour, body, sweetness, acidity, balance, uniformity, aftertaste for subjective features such as coffee qualities. The terms to convey subjective features are self-explanatory. A collaborative filtering system collects and analyzes a user's behavior based on a user's preferences given in the form of feedback, ratings, and other interactions. It is a domain-independent prediction technique. This technique can be used in any domain where content cannot be easily described by metadata. To predict and recommend items for active users, collaborative filtering techniques use data other than a user's behavior within a system. More specifically, a user-item rating matrix of preferences for items by users is constructed. From this matrix, user matches are made based on similar preferences and interests by calculating similarities between user profiles. While similarities between profiles can be calculated in many different ways, the fundamental principle of collaborative filtering is to aggregate user preferences in such a way as to provide predictions for a user based on his or her unique preferences. Our collaborative filtering model utilizes information related to user ratings based on the coffee properties.

A Principal Computing Analysis (PCA) algorithm has been applied to get the recommendations. PCA is a technique of matrix factorization that is used to reduce the number of features in the data set. The matrix factorization is done on a matrix which is generated by the user's feedback in the form of ratings on different coffee properties. PCA is a technique used to detect latent relationships between users and items. PCA generates a low dimensional representation of the original matrix space to calculate a neighborhood in the reduced space. It uses orthogonal transformation to convert correlated variable into uncorrelated variables. Algorithm for PCA using eigenvalue and eigenvectors is:

Step 1: Get the data from $m \times n$ matrix B

Step 2: Find the covariance matrix

Step 3: Find the eigenvectors and eigenvalues of the covariance matrix

Step 4: Choosing principal components and then forming a feature vector

Step 5: Deriving the new data set and form the clusters

Most collaborative filtering based recommendation systems have used similarity functions to compute the similarity between two users, such as the Pearson correlation coefficient, cosine similarity, or distance-based similarity. Here we are following the approach with cosine similarity. The similarity between coffees and the properties in the dataset is calculated using Cosine Similarity. Cosine Similarity is the Cosine of the angle between two vectors. Cosine Similarity discovers the direction between two vectors if it is the same or not. Further, the average of the Cosine Similarity scores is generated using features, types, etc. The resultant profiles are listed in descending order of similarity score to show more relevant recipes at the top. Calculations in this method are done by calculating the Cosine value between two vectors. The proximity of user characteristics can be found by cosine similarity formula,

$$\text{similarity} = \cos \theta = \frac{A \cdot B}{|A||B|} = \frac{\sqrt{\sum_{i=1}^n A_i \times B_i}}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

The below approach was followed to get recommendations using PCA. The 5 steps in this algorithm can be summarized as follows:

1. Desired coffee and the coffees of that type are extracted in such a way that it represents a unique relationship via its rating.
2. A matrix of features is constructed such that each row represents a coffee name and columns represent its associated features. A single cell represents a rating for a coffee given by a user.
3. A sparse matrix is generated. Sparse matrices are efficient in performing mathematical operations on large datasets. This sparse matrix is constructed using the 'csr matrix' function found within python's 'scipy.sparse' library. The generated sparse matrix is sent as input to the PCA algorithm.
4. The Principal Component Analysis (PCA) is implemented against a user's rating matrix, which returns factors of the rating matrix. The product of these factors is a rating matrix with predicted ratings. Using cosine similarity we calculate the most close coffee according to the ratings.
5. Recommendations are presented back to the user.

IV. EXPERIMENTAL RESULTS

We present the results related to model and evaluate the collaborative filtering model, ratio of relevant items from recommended items to the number of all relevant items, precision, or the ratio of the relevant items from recommended items to the number of all recommended item, and accuracy, which is the fraction of correct predictions predicted by a system. The most popular measure of effectiveness is accuracy metrics. They are used to evaluate either the prediction accuracy of estimating the ratings of specific user-item combinations or the accuracy of the top-k ranking predicted by a recommender system.

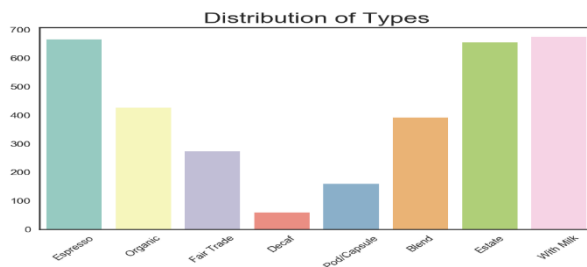


Fig1: Coffee types that are mostly consumed

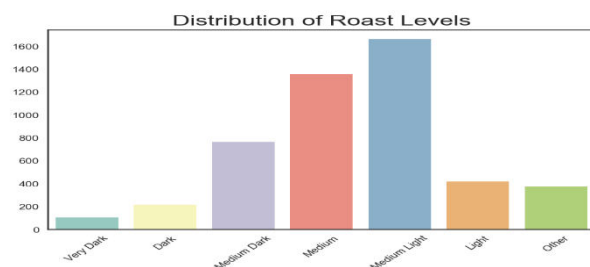


Fig2: Roast levels of coffee

Recommendation is calculated based on the similarity between items using people's ratings on those items. Here, each item is represented by a coffee. For example, in order to calculate the similarity between Chokocha Espresso and other espresso, we created two vectors with all the reviews for each coffee and then calculated the vector's cosine similarity. The vector contains several 0 values to fill in null values when a movie was not rated by a user. We calculated the cosine similarity between all coffee data. We've compared Chokocha Espresso with the other espressos in the database. If active user or customer liked eating Chokocha Espresso, the information above shows that Organic cacaoEspresso ,Mocha Espresso , Kenya Fair Trade , Espresso Republic , are the most similar coffees with 0.99, 0.72, 0.04 and 0.86 similarity scores. Ranked last, Kenya Fair Traid with similarity score 0.04. The similarity scores relies on over thousands user ratings, which is a sample for creating a robust recommender system.

Table1 :Cosine Similarity Between Chokocha Espresso and different espresso coffees

	Coffee	Cosine_similarity
1	Organic cacao Espresso	0.99
2	Mocha Espresso	0.72
3	Kenya Fair Trade	0.04
4	Espresso Republic	0.86

One of the most popular methods of computing accuracy metric is the root mean squared error, or RMSE. Here, we reviewed collaborative filtering methods and evaluation metrics to estimate effectiveness of recommender systems. We have prepared a theoretical basis for the implementation of collaborative filtering techniques for a coffee recommender system. We plan to vary different parameters such as similarity measure, scoring function to improve the effectiveness of recommender systems. The first results obtained fixing the parameter q and α . A correct setting of the parameter q is critically important. The best result was obtained for non-trivial $\alpha=0.15$ and $q=5$ in the user-based case.

V. CONCLUSION

Recommender systems are a powerful new technology for extracting additional value for a business from its user data-bases. These systems help users find items they want. Recommender system benefits users by enabling them to find items they like. Conversely they help the business by generating more sales. Here we presented and experimentally evaluated an algorithm or collaborative filtering based recommender systems. Our results show that user-based techniques hold the promise of allowing collaborative filtering based algorithms to scale to large data sets and at the same time produce high quality recommendations.

REFERENCES

- [1] Praphula Kumar Jain, E. A. Yekun, R. Pamula, Gautam Srivastava (2021), "Consumer Recommendation Prediction in Online Reviews using Machine Learning Models", Department of Computers and Electrical Engineering, volume 95 Elsevier.
- [2] R. Y. Toledo *et al.*: Food Recommender System Considering Nutritional Information and User Preferences. Part of the Spanish National Research Project TIN2015-66524-P, IEEE, date of publication July 17, 2019.
- [3] Ge, M., Elahi, M., Fernández-Tobías, I., Ricci, F., & Massimo, D., "Using tags and latent factors in a food recommender system," in Proc. of the 5th International Conference on Digital Health, pp. 105-112, ACM., May 2015
- [4] Freyne, J., & Berkovsky, S., "Evaluating recommender systems for supportive technologies," User Modeling and Adaptation for Daily Routines, pp. 195-217, Springer London, 2013.
- [5] Phanich, M., Pholkul, P., & Phimoltares, S., "Food recommendation system using clustering analysis for diabetic patients," in Proc. of International Conference on Information Science and Applications, pp. 1-8, IEEE, April 2010. Article
- [6] S. Isaacman, R. Becker, R. Caceres, S. Kobourov, M. Martonosi, J. Rowland, and A. Varshavsky, "Identifying important places in people's lives from cellular network data", Proceedings of Int'l Conf. on Pervasive Computing, 2011, pp. 133-151
- [7] J. Bowen and C. Morosan, "Beware hospitality industry: the robots are coming," Worldwide Hospitality and Tourism Themes, 2018.
- [8] L. N. Cain, J. H. Thomas, and M. Alonso Jr, "From sci-fi to sci-fact: the state of robotics and ai in the hospitality industry," Journal of Hospitality and Tourism Technology, 2019.



- [9] G.-Z. Yang, B. J. Nelson, R. R. Murphy, H. Choset, H. Christensen, S. H. Collins, P. Dario, K. Goldberg, K. Ikuta, N. Jacobstein et al., “Combating covid-19the role of robotics in managing public health and infectious diseases,” 2020.
- [10] C. Fu, Y. Yoshikawa, T. Iio, and H. Ishiguro, “Sharing experiences to help a robot present its mind and sociability,” International Journal of Social Robotics, pp. 1–12, 2020.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details