Restaurant recommendation using hybrid collaborative filtering

Github link: https://github.com/VINEETHREDDYPANDALA/Restaurant-recomendation-system

Raviteja Chakilam
700740506
dept.Computer Science
University of Central Missouri
rxc05060@ucmo.edu

Pranay Paul Thirumala Reddy
700741744
dept.Computer Science
University of Central Missouri
Pxt17440@ucmo.edu

Vineeth Reddy Pandala
700743902
dept.Computer Science
University of Central Missouri
vxp39020@ucmo.edu

***Abstract*—Nowadays, ordering groceries, food, clothing, and other basics online is becoming more and more common. According to recent surveys, buyers spend a lot of time selecting the best products. Integrating recommendation algorithms on their platforms helps e-commerce and e-content application developers improve their sales. According to a survey, 40 percent of the items that buyers buy are based on recommendations. Content-based filtering recommendations are not personalized. The recommendations are generated based on popularity. This method benefits only new consumers who do not have a past search history. Collaborative filtering produces data sparsity this filtering can not produce recommendations to new users. In this research, we propose a matrix factorization-based hybrid filtering strategy for a restaurant recommendation system. The SVD (Singular Value Decomposition) method is used to implement this method. The user-item or item-user collaborative filtering methods produce data sparsity since users interact with a limited number of items. The matrix factorization method overcomes this problem by user-item matrix factorization by selecting a low-rank matrix. The recommendations for new users cannot be produced by collaborative filtering. In this case, matrix factorization deals with the cold start problem which uses the latent features of the users. The data for the experimental analysis of restaurants are gathered by Kaggle. Results from collaborative filtering and model-based content-based methods are also published to compare performance in addition to the hybrid method.**

[1] ***Index Terms*—Data sparsity, Matrix Factorization(MF), hybrid filtering, collaborative filtering, and SVD (Singular Value Decomposition)**

I. INTRODUCTION

The significance of the project is building multiple recommender systems content-based and hybrid collaborative filtering. Content-based filtering is used to recommend restaurants for new users and a hybrid collaborative filtering method using SVD (matrix decomposition) method. We avoid the data sparsity that will be created in general collaborative filtering using Matrix Factorization and decomposition methods.

Almost every online platform needs a recommendation system to engage customers and improvethe sales of the business. Traditional and existing methods recommend the products based on popularity and collaborative filtering raises the data sparsity. To mitigate these problems in this study a hybrid recommendation system is constructed to recommend the products based on the combination of traditional methods to improve the accuracy of recommendations.

Daily, a lot of courses are published to help the students in learning. Due to the high number of courses published, it has become a difficult taskto choose the best course for a particular student based on the prerequisites and learning level. In this system, a hybrid recommendation system and collaborative filtering system are implemented. To implement collaborative filtering KNN algorithm is used.

The main features of the project are: Controlling data sparsity, Diversity in recommendations, and being Well suited for demographic-based recommendations. Each method's performance is evaluatedand compared in a comparative analysis.

Limitations of existing systems: Content-based filtering recommendations are not personalized. The

[1]https://github.com/VINEETHREDDYPANDALA/Restaurant-recomendation-system

recommendations are generated based on popularity.

This method benefits only new consumers who do not have a past search history. Collaborative filtering produces data sparsity. Collaborative filtering can not produce recommendations for new users.

The main objectives of the project are: Building the model-based content-based, collaborative, and hybrid methods. Evaluating the results of each method Conducting experimental analysis

Different strategies are used to create recommendation systems while keeping in mind the target consumers and the amount of reach. In this section, we review the research on machine learning advancements in terms of recommendation algorithms.

The similarity-based concept was the first step in the evolution of recommendation systems. Based on the similarity score, recommendations for items or products are made. In this study, recommendations were based on user and item similarities. According to experimental findings, item similarity received the highest F1 score.

This paper proposed multiple model-based collaborative filtering methods to recommend restaurants to the user. While recommending the restaurants it considers various parameters like choice of food, location of the restaurant, ratings, climate, etc., This collaborative filtering method is used to predict the ratings of the restaurant considering a particular user. The SVM classification model outperformed the KNN classifier.

When recommending products, content-based and collaborative filtering have their limitations. To get over the problems with content and collaborative filtering in this study, hybrid filtering is used. By comparing user ratings and restaurant characteristics, content-based results are produced. Utilizing weighted, cluster-based, and similarity-based techniques, collaborative filtering is carried out. Both the collaboration and content outcomes are inte-grated into the hybrid method results. According to experimental findings, the hybrid filtering method performs better than the current methods.

The hybrid filtering method overcomes the challenges of the traditional method. There are a handful of methods to perform hybrid filtering matrix factorization. In this method matrices of the combined features are reduced to a lower rank and latent features are observed.

II. Motivation

Bengaluru is known as the IT capital of India. Presently there are more than 12,000 restaurants present in this city and these restaurants serve food all over the country with multiple cuisines.Since most of the people are working professionals who mainly depend on restaurant food there is an increasing demand for restaurant food. But finding the restaurant based on the cuisines and reviews is difficult. So in this project, we are building a recommendation system using a reviews list and a hybrid recommender system. The significance of the project is building multiple recommender systems with content-based and hybrid collaborative filtering. Content-based filtering is used to recommend restaurants for new users and a hybrid collaborative filtering method using SVD (matrix decomposition) method. We avoid the data sparsity that will be created in general collaborative filtering using Matrix Factorization and decomposition methods.

In this project model based recommendation systems are implemented which gives the advantage of easy evaluation. The user-item or item-user collaborative filtering methods produce data sparsity since users interact with a limited number of items. The matrix factorization method overcomes this problem by user-item matrix factorization by selecting a low-rank matrix. The recommendations for new users cannot be produced by collaborative filtering. In this case, matrix factorization deals with the cold start problem which uses the latent features of the users.

III. Objectives

The main objectives of the restaurant recommendation system are:

- Exploring the data using visualizations
- Building a user-interactive Bengaluru restaurant recommendation system using content-based recommendation and hybrid filtering methods using demographic and restaurant-specific features.
- Building content-based recommendation systems for new users using similar reviews restaurants based on cosine similarity of the reviews
- Building hybrid recommendation system us-ing Matrix factorization method by using SVD(Singular Value Decomposition) algorithm

- Mapping the recommended restaurant locality on the maps using geopy and folium Python modules

IV. Related work

Recommendation systems (RSs) suggest products to consumers based on their purchasing behavior or search history to deduce users' search scope. Although there are advancements in detecting changes in users' preferences with time, it is challenging to always compute the users' preference variations. To address the issue, we propose collaborative filter- ing algorithms. We shall quantify user preference change in this paper. Furthermore, suggesting in the reverse direction, that is, we will recommend things that different users have rated to the indicated user. The experimental analysis is done on MovieLens 1M and MovieLen's newest small shows dataset[18].

This study provides a computational memory solution (CMS), a cost-effective and scalable memory solution recommendation system. CMS provides training and scales as far as the high-speed serial interface allows, in contrast to other cutting-edge near-memory processing accelerators that only give inference and have scaling constraints. Our evaluation results show that CMS outperforms state-of-the-art accelerators in terms of throughput and power efficiency by up to 7.5 times [7].

A major challenge in academic areas is finding the right materials and information related to the articles. At the moment the internet is overloaded with information and it has become more challengingto find the right information. In this study, we are proposing knowledge recommendation algorithmsto find the right article recommendations and a brief overview of the articles. To achieve this intelligent algorithms and deep learning-based algorithms are proposed [17].

The current study aims to develop and evaluate a sightseeing site suggestion system for assisting tourists while taking into account their visit frequency. This system to recommend tourist sites to users was created by combining a social networking service (SNS), a Web-geographic information system (GIS), and a recommendation system. The system was used for 6 weeks in Kamakura City with a total of 61 users. According to the web questionnaire survey, the recommendation function of sightseeing sites received mostly positive feedback [9] [].

are that. As the users of online news portals are interested in specific domains and the hottest topics of the day. Recommendation systems help people make decisions and take action in these situations. Recommendation systems with content-based filtering techniques help users to select the best products. In this study, we are implementing unsupervised learning algorithms to produce recommendations that make clusters of similar groups [11] [16].

In this study, a recommendation system that helps travelers called TravelMate is proposed. This system helps tourists choose appropriate tours ortrip companions. In certain situations like users who know what they're looking for, content-based filtering works best. However, for users who are still unsure about the requirements, the conversational recommendation system (CRS) assists in clarifying the demands during the chatting process safely and engagingly [10].

Nowadays digital marketing is at its peak. The challenging task is to improve the sales of the business. This study aims to design a recommendation system for current intelligent cinema and television advertisements based on digital media and a decision tree. The model automatically takes the characteristics from the description. This is used in the recommendation of smart data and multimedia systems [20] [1] [2].

The employment demand of companies is increasing exponentially. There is demand for employment amongst students and employers. In this study, we are proposing a recommendation system for employment. This system is using the collaborative filtering recommendation framework Mahout built in this study [4].

Traditional recommendation systems in e-commerce provide recommendations such as a single personalized suggestion approach and no further data analysis. In this study, we propose a hybrid personalized recommendation system for both offline mining, real-time mining, and deep learning technology recommendation algorithm and its application, and analysis and summary of many e-commerce user behavior data. The central

idea is to collect and pre-process data in real-time from numerous e-commerce platforms that generate user data, as well as to collect all of the users' personalized data [15] [12] [3].

A good recommendation system promotes products on an e-commerce platform to a user based on their preferences and history. This enhances user engagement on the site. The recommendation system works by establishing relationships betweenitems or users. Every online platform needs a good recommendation system. Hybrid Recommendation combines the predictions of two or more recommendation systems. This system combines predictions from product-based predictions, done using a Graph-based recommendation system, and predictions from users, implemented using a Collaborative Filtering recommendation system [8] [6].

A recommendation system is required to aid users in filtering the copious lodging information. This research uses the content-based filtering method to deliver good accommodation recommendations by offering relevant recommendations by extracting content features from the new users. The new users can get the proper recommendations. The test results showed that the recommendation system average accuracy of 76.67 percent [14] [19].

An effective recommendation system on online platforms provides creators by allowing them to receive appropriate compensation and consumers by allowing them to experience something relevant, entertaining, and personalized. In this research offers a mechanism for implementing the liquid democracy principle in a content recommendation system. This work ranking system helps to encourage suggestions based on personal preferences to improve the accuracy and diversity of recommendation outcomes on Bitcoin news on Twitter to identify the opinion leader. This study focuses on tier-2 implementation [13] [5].

V. Data Description

Zomato Bangalore Restaurants is collected from Kaggle. The dataset contains 51717 records and17 features. The features include demographic fea- tures like city, address, and location, and restaurant-specific features type of restaurant, online order, and approximate cost of food.

S.No.	Column name	Description
1.	URL	URL of the business
2.	address	Address of the restaurant
3.	name	Name of the restaurant
4.	online order	Yes/No
5.	book table	Yes/No
6.	rate	Ratings
7.	votes	Number of votes
8.	phone	Phone number
9.	location	Location or city of the restaurant
10.	rest type	Type of restaurant
11.	dish liked	Dishes liked by the customer
12.	cuisines	cuisines
13.	approx$_c ost$	cost for two people
14.	reviews list	List of reviews
15.	menu item	Menu items
16.	Listed in type	dine out or delivery,pubs or bars
17.	listed in(city)	Table I City name

DATASET FEATURE DESCRIPTION

VI. Proposed Framework

The project implementation is divided into 4 stages:

A. Data preparation

- In any machine learning project the efficiency of results depends on the quality of the data collected and how effectively the preprocessing is done.
- In this step, raw data is collected from the Kaggle repository and cleaned.
- Dropping the unnecessary columns URLs, phone, dish liked columns and finding the duplicates in the data and removing them
- Observing the duplicates and dropping them
- Cleaning the ratings and cost columns and converting them to float values
- Text preprocessing of review columns and cuisines.Removing the punctuation marks, removing the stop words and removing the URLs.

B. Exploratory Data Analysis

In this step, the relationship of features is analyzed using various graphs using seaborn and matplotlib. In any machine learning lifecycle, this step is important to understand the data distribution and apply appropriate algorithms to recommend/predict

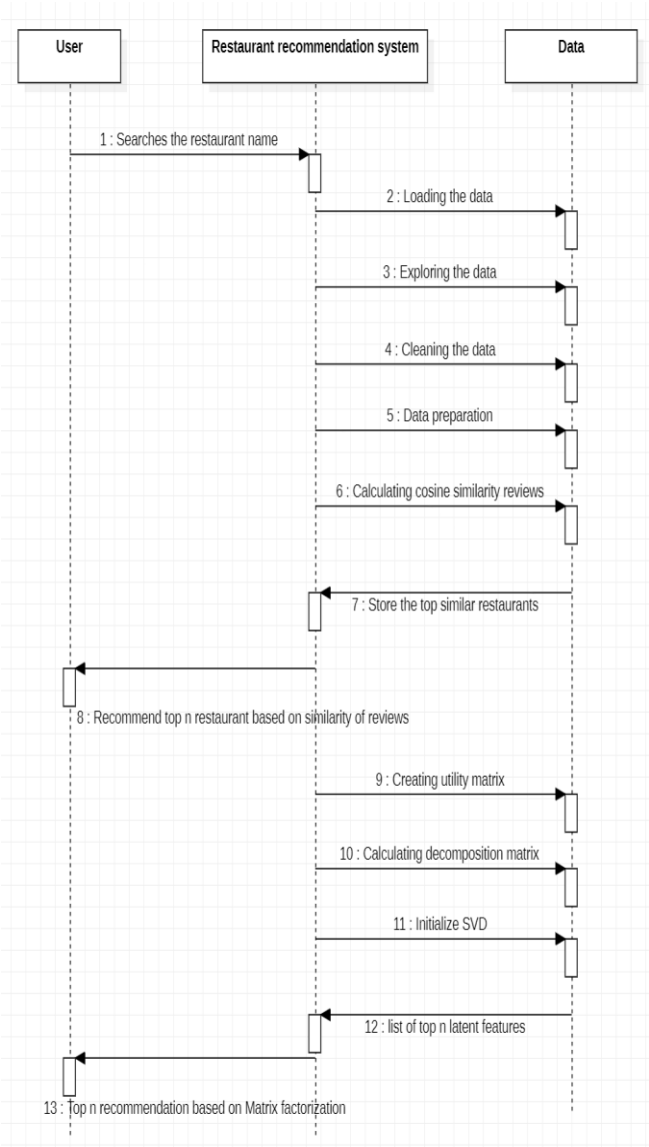

Figure 1. Sequence diagram

the data. Some of the observations are conclusions drawn from EDA are:

- Top 6 cities that have the highest number of votes are Sarjapur Road, Bellandur, Indiranagar, Koramangala 7th Block, and Koramangala 6th Block.
- The distribution of restaurant ratings follow a normal distribution
- In the dataset there are the highest number of samples are dine-out category and the least number of samples are in pubs and bars
- The visualization of cost vs restaurant category reveals that there are presence of outliers in the data and high number of outliers in the dine-out category.
- Average rating is high for table booking classes compared to the non booking classes

C. Building recommendation system

In this first step of building a recommendation system review lists are vectorized using TFIDF vectorizer and then cosine similarity is calculated to choose the recommendations based on similar reviews. Content-based filtering: In content-based filtering cosine similarity of the reviews is calculated and then top n recommendations are displayed to the user when he searches the restaurant name. In the singular value decomposition method rows are taken as ratings and columns are taken as restaurant names and an orthogonal matrix is constructed. By transposing the Utility matrix SVD is trained. Based on Pearson's correlation coefficient recommendations are generated.

VII. Results Summary

GeoPandas is an open-source module that works with geographical data in Python. GeoPandas expands pandas' datatypes to provide spatial operations on geometric types. Shapely handles geometric operations. Geopandas also requires Fiona for file access and matplotlib for plotting.Singular Value Decomposition is a mathematical technique used to decompose a matrix into three separate components: a left singular matrix, a diagonal matrix of singular values, and a right singular matrix. It is a way to break down a matrix into three simpler parts. These parts provide useful information about the original matrix and can be used for various purposes, such as data compression, dimensionality reduction, and solving linear equations.

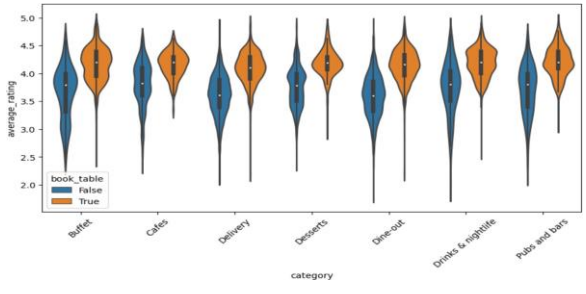

Figure 2. average rating categorywise

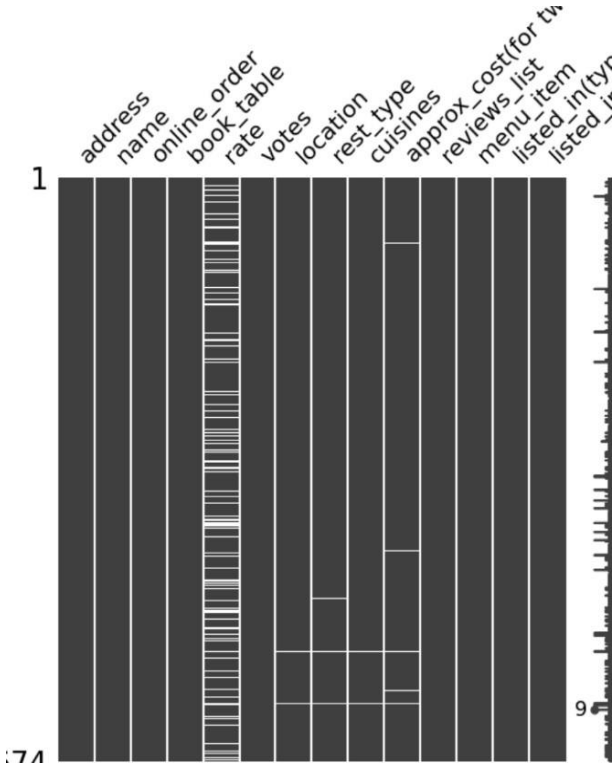

Figure 3. Null values distribution

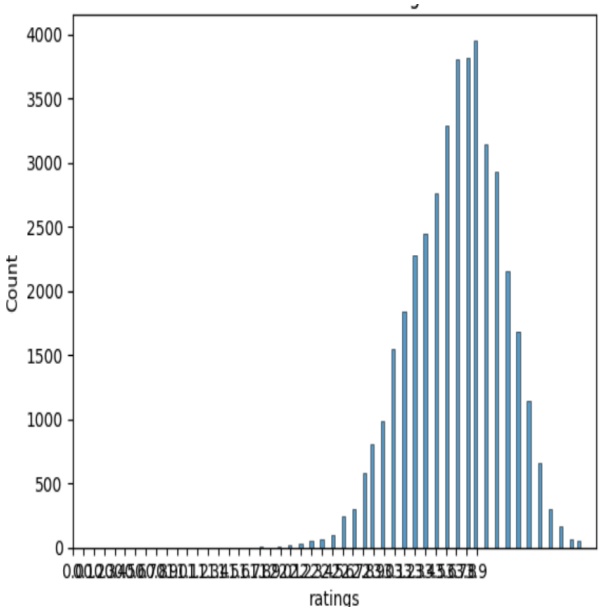

Figure 5. distribution of ratings

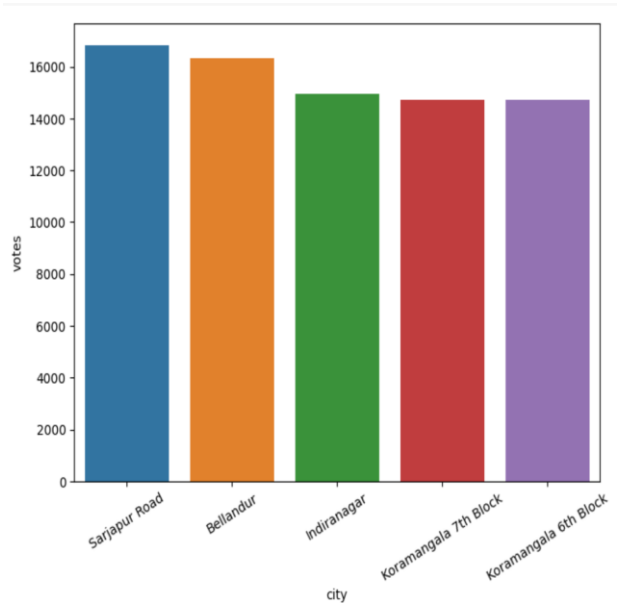

Figure 4. Citywise votings

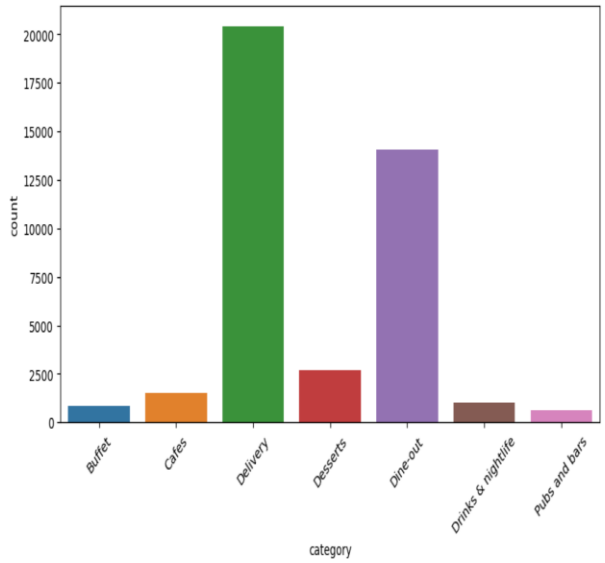

Figure 6. countplot of categories

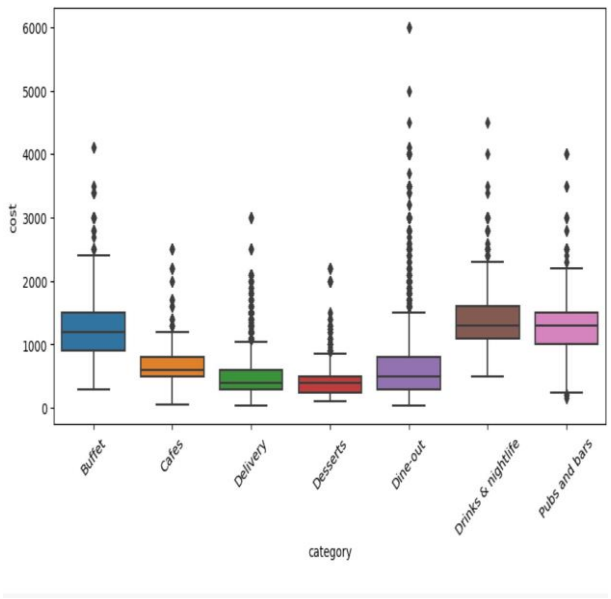

Figure 7. Categorywisecost

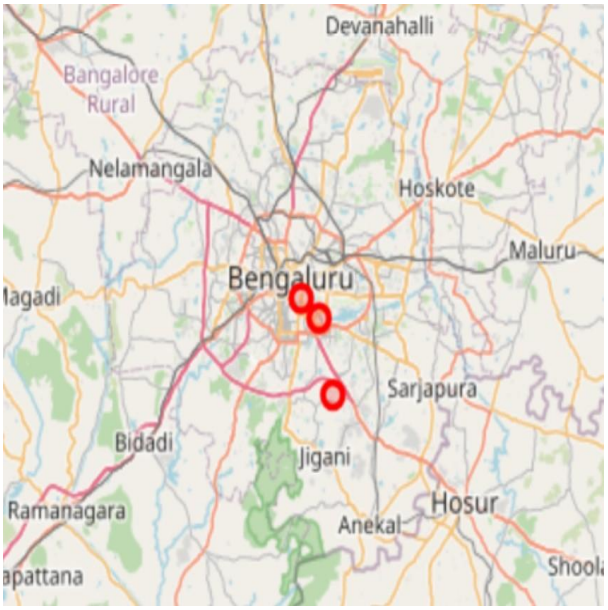

Figure 9. mapped results of content-based filtering

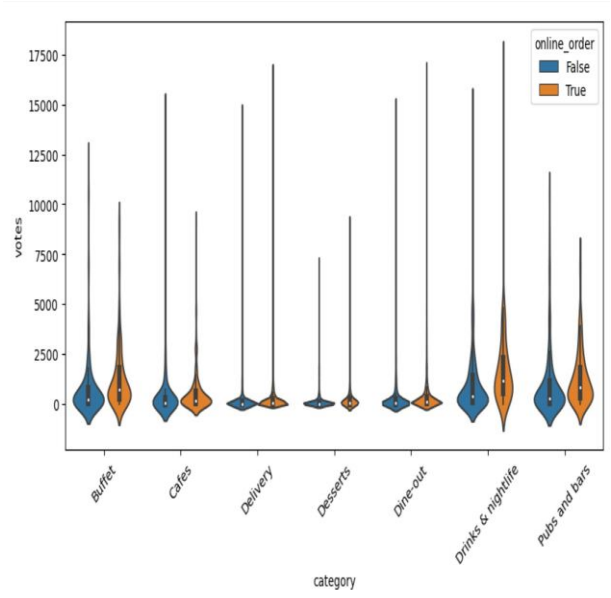

Figure 8. votes categorywise

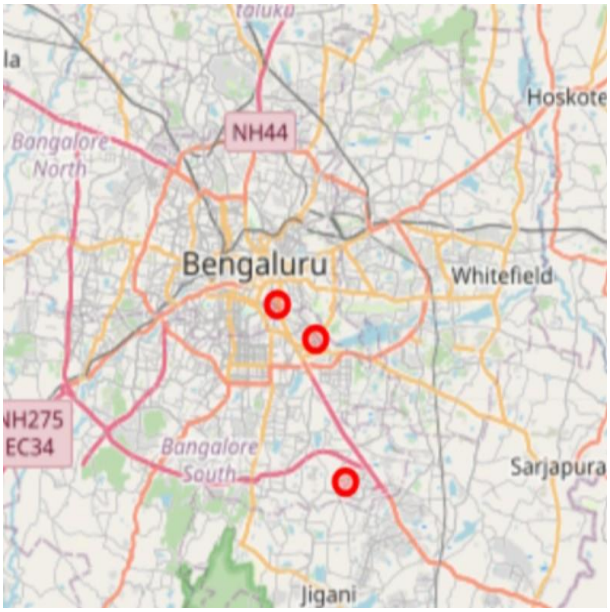

Figure 10. mapped results of hybrid filtering

REFERENCES

[1] Jayashree Agarkhed and Geetha Pawar. Recommendation-based security model for ubiquitous system using deep learning technique. In *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1–6, 2022.

[2] V. Anupama and M. Sudheep Elayidom. Course recommendation system: Collaborative filtering, machine learning and topic modelling. In *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 1459–1462, 2022.

[3] Hongxia Cai, Cheng Wang, and Zhishu Liu. An improved pearson coefficient based literature recommendation system for small research teams. In *2021 13th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pages 208–211, 2021.

[4] Zhigao Chen, Wenqian Liang, Xiong Gao, Zhijian Zhou, and Meng Wu. Research on the accurate recommendation management system for employment of college graduates on hadoop. In *2019 5th International Conference on Big Data and Information Analytics (BigDIA)*, pages 19–22, 2019.

[5] Muhammad Hassan Farooq Butt, Jian Ping Li, and Tehreem Saboor. A tunable and explainable attributes (tea) for recommendation system. In *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pages 39–43, 2020.

[6] Rishi Gulati and Mohnish Bhagat. Content and knowledgebased food recommending intelligence enabled system (fries). In *2021 5th International Conference on Computing Method- ologies and Communication (ICCMC)*, pages 1743–1749, 2021.

[7] Minho Ha, Joonseop Sim, Donguk Moon, Myunghyun Rhee, Jungmin Choi, Byungil Koh, Euicheol Lim, and Kyoung Park. Cms: A computational memory solution for high-performance and power-efficient recommendation system. In *2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, pages 491–494, 2022.

[8] Ilhamsyah, Syahru Rahmayudha, Izhan Fakhruzi, Ferdy Febrianto, Ahmad Cahyono Adi, and Veri Gunawan. Recommendation system for the improvement of e-government services in the tourism sector of pontianak city. In *2022 International Conference on ICT for Smart Society (ICISS)*, pages 1–4, 2022.

[9] Yudai Kato and Kayoko Yamamoto. Development of sightseeing spot recommendation system considering users' visit frequency. In *2020 Joint 11th International Conference onSoft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS)*, pages 1–8, 2020.

[10] Yii-Lin Lu, Yih-Jiun Lee, and Kai-Wen Lien. A conversational recommendation system has better usability? a case-study of travelmate. In *2022 IEEE 11th Global Conference on Consumer Electronics (GCCE)*, pages 450–452, 2022.

[11] Mara-Renata Petrusel. An unsupervised topic-driven new york times recommendation system. In *2022 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 1–6, 2022.

[12] Abhishek Saxena and Anton Kolonin. Application of liquid rank reputation system for content recommendation. In *2022 Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, pages 196–199, 2022.

[13] Bin Sun, Hua Tan, Dongxiao Yang, Yu Zhang, Yefei Wu, and Xinyu Jin. The system of personalized learning resource recommendation and experimental teaching based on collaborative filtering. In *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 862–866, 2022.

[14] Haroldo Shigueaki Teruya, Ingrid Marçal, Ronaldo Celso Messias Correia, Rogério Eduardo Garcia, Danilo Medeiros Eler, and João Osvaldo Rodrigues Nunes. Urecommender: An api for recommendation systems. In *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6, 2020.

[15] Zhen Wang, Allam Maalla, and Mingbiao Liang. Research on e-commerce personalized recommendation system based on big data technology. In *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, volume 2, pages 909–913, 2021.

[16] Agnes Werner-Stark and Zsuzsanna Nagy. A heuristic method to recommendation systems. In *2020 6th International Conference on Information Management (ICIM)*, pages 200–204, 2020.

[17] Huaiyuan Yang, Hua Zhou, and Yucheng Li. A review of academic recommendation systems based on intelligent recommendation algorithms. In *2022 7th International Conference on Image, Vision and Computing (ICIVC)*, pages 958–962, 2022.

[18] Jing Yu, Jinaiing Shi, Yunwen Chen, Daqi Ji, Wenhai Liu, Zhijun Xie, Kai Liu, and Xue Feng. Collaborative filtering recommendation with fluctuations of user' preference. In *2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE)*, pages 222–226, 2021.

[19] Jing Yu, Jingjing Shi, Yunwen Chen, Wenhai Liu, Kai Liu, and Zhijun Xie. Enhancing collaborative filtering recommendation by user interest probability. In *2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, pages 525–529, 2021.

[20] Huawen Zhao, Yanhua Liu, and Xiaohong Ling. Design of intelligent film and television advertisements recommendation system based on digital media and decision tree. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, pages 968–971, 2021.