Meeting Point Recommendation System Using Collaborative Filtering Method

Himma Filangga Sutopo,Tita Karlita, Entin Martiana Kusumaningtyas

Department of Informatics and Computer
Electronic Engineering Polytechnic Institute of Surabaya
Surabaya, Indonesia

himma29@gmail.com, tita@pens.ac.id, entin@pens.ac.id

Abstract. Recommendation systems aim to provide item recommendations to users. By offering recommendations, users can consider the suggested items. One method for building a recommendation system is Collaborative Filtering. Collaborative filtering works by utilizing user ratings to predict their ratings for specific items. Cafés are popular places for social gatherings with friends or family. Currently, numerous cafés and restaurants are operating with diverse menus and facilities. The variety of options makes it challenging for individuals to choose a suitable place to visit. To simplify the selection process, a recommendation system is needed. In this study, the author develops a system recommending cafés to consumers. The system uses the user-based collaborative filtering algorithm and employs cosine similarity. The dataset used in this research consists of 284 cafes with over 25,000 reviews obtained by scraping Google reviews. The dataset includes usernames, given ratings, cafe names, and cafe ambiance. Currently, with 284 cafes, the system yields cosine similarity values below 0.1. However, experimental findings indicate that these cosine similarity values are expected to increase as the dataset size grows. This observation suggests the system's recommendations will grow more robust and precise with an expanding dataset. The system's efficiency was underscored by its performance even with smaller datasets, achieving an average response time of 120ms for smaller and 1600ms for larger datasets. This research lays the groundwork for further exploration of factors shaping user preferences and sets a benchmark for cafe recommendation system performance. Interestingly, even with smaller datasets, the recommended cafes remain consistent despite variations in larger datasets. This finding suggests that the system can provide reliable recommendations even with a smaller dataset, showcasing its effectiveness in capturing user preferences. This research also establishes a foundation for further studies in identifying other factors influencing user preferences when selecting cafes.

Keywords. Recommendation system, Collaborative filtering, Place gathering, User-based collaborative filtering.

M. U. H. Al Rasyid and M. R. Mufid (eds.), *Proceedings of the International Conference on Applied Science and Technology on Engineering Science 2023 (iCAST-ES 2023)*, Advances in Engineering Research 230,
https://doi.org/10.2991/978-94-6463-364-1_43

1. INTRODUCTION

Location serves as the basis for determining the position of human activities. Sociologist Ray Oldenburg explains that locations can generally be divided into "The First Place" and "The Second Place" [1]. "The First Place" refers to the location used as a residence, while "The Second Place" refers to the location used for work or daily activities.

In human development, there is a need for alternative spaces beyond the home, a place to engage in serious discussions or pass the time. As inherently social beings, humans always require a venue for interacting with others. These alternative spaces, which provide opportunities for interaction with family members, friends, or others, are called the "Third Place" [1]. The "Third Place" can take various forms, including parks or cafes [2].

Cafes, one of the representations of the "Third Place," have witnessed a rapid surge in numbers across Indonesia, particularly in major cities like Jakarta, Bandung, Yogyakarta, and Surabaya, with an estimated count exceeding 10,000 cafes [3].
Furthermore, based on consumer behavior in choosing a cafe, a study has revealed three significant findings. Firstly, menu pricing is paramount among nine criteria when customers choose a restaurant. Secondly, the occasion for dining out greatly influences restaurant selection criteria, with menu pricing crucial for quick meals and social gatherings, brand reputation for business needs, and word-of-mouth for celebrations. Lastly, menu pricing is a top consideration across full-service, quick-casual, and quick-service restaurants, emphasizing its significance when faced with similar dining options [4].

Considering consumers' preferences for different styles and experiences, the author aims to facilitate consumers in choosing cafes as gathering places. In this research, the author plans to develop a recommendation application that utilizes the Collaborative Filtering decision-making method. This application will encompass four different types of cafes: traditional, industrial, modern, and retro.

The selection of these cafe types is based on strong considerations. Traditional for affordable hangouts, industrial to explore contemporary design preferences, modern for insights into technology and trends, and retro to capture nostalgia and retro lifestyle allure.

Through the developed recommendation application, consumers are expected to easily find cafes that align with their preferences for gathering places. By including these four types of cafes, this research will provide more varied recommendations that can adapt to each individual's unique preferences.

The researchers chose Surabaya as the research scope because of its vibrant cafe culture, reflective of evolving societal trends and preferences. The city's diverse population, including students, professionals, and a burgeoning middle class, has fueled the growth of cafes as social hubs and venues for various activities. This lively cafe culture offers a rich dataset that allows for the analysis of consumption patterns, preferences, and trends within the context of urban lifestyle. Surabaya's unique blend of traditional coffee shops, modern cafes, and international coffee chains provides a comprehensive perspective, enabling a deeper understanding of the city's coffee culture and its intertwining with urban living.

This study develops a cafe recommendation system for gathering places. The challenge is the abundance of cafe options, making it hard for users to find suitable ones. Proximity and user preferences are crucial. The aim is to create an efficient recommendation system using Collaborative Filtering and user ratings.

This project aims to design and develop a recommendation system application focused explicitly on cafes, using Collaborative Filtering. The system will utilize user rating data from Google Maps and identify preference similarities between the target user and others. The aim is to provide personalized recommendations for cafes that effectively align with users' preferences.

This project benefits the public through time-saving cafe selection. It also advances recommendation system technology by combining collaborative filtering and user rating analysis, offering insights to researchers. It contributes to computer science knowledge and is a foundation for further research in recommendation systems.

2. RELATED WORKS

This study draws inspiration from various sources that explore the use of Collaborative Filtering in recommendation systems. Ching-Seh [9] focuses on building a film recommendation system through Collaborative Filtering and the Apache Mahout framework. The study employs user ratings from Yahoo! Movies and includes details like titles, genres, directors, and actors. Results show that while the item-based approach demands extensive similarity computations, its efficiency benefits from static data. Conversely, the user-based system utilizes dynamic data from multiple users, necessitating storage for user data and dedicated processing.

In another study, Missi Hikmatyar and Ruuhwan [10] developed a recommendation system for a library's Book Search System. This system addressed the challenge users encounter when trying to choose books aligning with their preferences from an abundance of search results. The study proposed a desktop-based book search recommendation system using Python and MySQL. The primary goal was to systematically suggest books based on users' past searches, reducing errors in obtaining the necessary reference books. The system employed a user-based collaborative filtering method, considering lending patterns and book subjects to establish similarity between members. The resulting book search recommendations were presented in a user-oriented order, providing suitable book titles based on their profiles.

Another relevant study was conducted by Elham Asani, Hamed Vahdat-Nejad, and Javad Sadri [11]. Their research delves into sentiment analysis-based restaurant recommender systems, emphasizing the utilization of sentiment analysis to derive individuals' food preferences from their comments. The proposed context-aware recommender system employs a semantic approach to cluster food names extracted from users' comments and analyze sentiments associated with them. Subsequently, the system recommends nearby open restaurants based on their similarity to user preferences.

This study makes a unique contribution by employing the user-based collaborative filtering algorithm and cosine similarity to recommend cafes to consumers. While previous research has predominantly focused on item-based collaborative filtering or hybrid techniques, this study adopts explicitly the user-based approach. By leveraging user preferences and similarities, the system generates recommendations that align closely with individual preferences and enhance the cafe-finding experience.

By emphasizing the utilization of user-based collaborative filtering and the incorporation of cosine similarity, this study presents a novel and alternative approach to cafe recommendation systems. These methodological differences provide a fresh perspective and offer potential improvements in the accuracy and relevance of the recommendations generated.

Overall, this study's unique contributions lie in its adoption of the user-based collaborative filtering algorithm, the utilization of cosine similarity, and the exploration of alternative methodologies in the field of cafe recommendation systems. These distinctive elements contribute to advancing recommendation systems and offer valuable insights for future research and development in this domain.

3. PROPOSED METHOD

3.1. Solution Description

This section outlines the development of a Collaborative Filtering-based cafe and restaurant recommendation system aimed at aiding users in choosing meeting point places aligned with their preferences. It leverages Collaborative Filtering techniques to assess user preferences and item similarities, utilizing a dataset of user ratings from Google Maps for tailored suggestions. The system incorporates proximity-based recommendations through the Haversine formula for distance calculation [14]. Integrated with the Google Maps API, it furnishes users with comprehensive details about recommended venues, including location, reviews, ratings, and more. The user interface, crafted using ReactJS and Ant Design, ensures an intuitive and user-friendly experience, amalgamating Collaborative Filtering, proximity-based selection, API integration, and an accessible interface for precise and personalized recommendations, simplifying decision-making for users.

3.2. System Design

This chapter will provide an overview of the material and method used in this research as a design system. This paper proposes a system design, as shown in Figure 1.

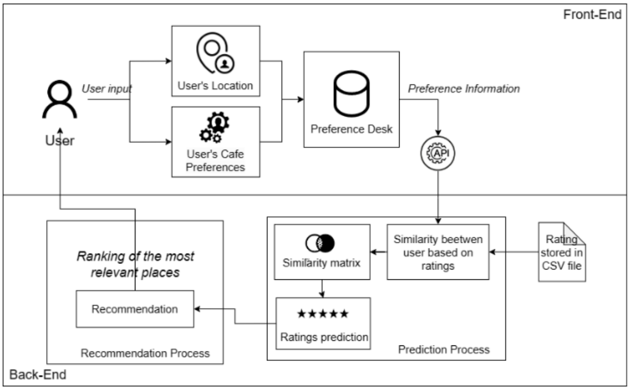

Fig 1. System Design

The proposed system design in figure 1 follows a sequential process. Initially, users input their location and cafe history. This information is then compared with the cafe dataset in the Surabaya database. Using the Collaborative Filtering approach, the system calculates similarity scores to gauge alignment between user preferences and existing cafes. Based on these scores, the system identifies and presents the top cafe recommendations with the highest similarity values to the user interface.

The dataset is obtained from Google Maps, a service that provides various cafe reviews using web scraping. The collected dataset will be used as a data collection for comparison using the Collaborative Filtering method. This is done using a service-based architecture, also known as a client-server architecture.

In this architecture, the backend and frontend are separate entities communicating through an Application Programming Interface (API). By using this architecture, the back-end and front-end can be developed independently. If the API remains consistent, the back-end can be updated or replaced without disrupting the front-end.

3.3. Data Collection

The dataset was gathered by conducting web scraping of cafe reviews publicly available on Google Review. The collected data comprises reviews of cafes from the past year, capturing essential information such as the cafe's name, menu, address, pricing, operating hours, and seating capacity. The collection process emphasized extracting the most recent reviews from consumers, ensuring that the dataset is up to date with current cafe trends and preferences. Specifically, the reviews were extracted within a defined time interval, focusing on the past year. This approach was adopted to maintain the dataset's relevancy and accuracy by considering recent reviews. An example of the collected data can be seen in Figure 2

Fig 2. Google Review

The review data will be classified to form a user-rating matrix, which will be used together with the user-item matrix for processing using Collaborative Filtering.

3.4. Process And Results

This research involved several important stages in using the Collaborative Filtering method to recommend cafes to users. From the user-item matrix and the previously created dataset collection, the user-item matrix will be used to calculate the similarity between users within the data collection. Figure 3 shows an example of the user-rating matrix, consisting of users who have ratings for cafes.

Fig 3. User-rating Matrix

In Figure 3, black numerical entries represent user ratings, while blue entries represent database ratings. These values are crucial for constructing the user-item matrix, a foundational element for Collaborative Filtering. In the Processing phase, user similarity is computed using this matrix. Higher similarity indicates closer user preferences. The Result phase yields a list of top-rated cafes based on similarity values, along with directions.

3.5. Analysis AND Discussion

The research employed a systematic approach to apply Collaborative Filtering in cafe recommendations. Data was gathered through web scraping of Google Reviews, capturing cafe details like menus, addresses, and pricing. This data was then structured into user-rating and user-item matrices to facilitate the Collaborative Filtering process.

By harnessing user similarities and preferences, Collaborative Filtering generated personalized cafe suggestions rooted in similar user ratings and reviews. The approach's efficacy and accuracy were validated using metrics like precision, recall, and mean average precision, highlighting the quality of recommendations.

Overall, the study demonstrated Collaborative Filtering's efficiency and reliability in cafe recommendations, emphasizing personalized suggestions based on user data and similarities. The integration of a comprehensive dataset, thoughtful classification, and effective evaluation techniques culminated in a successful recommendation system, enhancing the cafe selection experience for users.

3.6. Experimental Results And Discussion

The system to be built will use a web-based client-server architecture. The system will perform similarity calculations between users and make rating predictions. The user interface will allow users to input their data and interact with the plan to obtain rating predictions and recommendations for cafes or restaurants with the highest similarity scores.

3.7 Experiment Parameters

The parameters to be used for experimentation and analysis in the recommendation system website that has been developed to observe the process of the application system are as follows:

1) Functionality Success: Assessing the success of the application's functionalities.
2) Flask and ReactJs Framework Utilization Success: Evaluating the successful utilization of Flask and ReactJs frameworks as the back-end and front-end, respectively.
3) User Input Handling Success: Evaluating the application's success in handling user inputs.
4) System's Success in Providing Recommendations According to User Preferences.
5) Speed and Responsiveness: Evaluating the system's performance regarding response speed to user requests. This may include the application's response time to search queries or recommendation requests.
6) Availability and Scalability: Evaluating the system's availability in handling high user loads or traffic spikes. Checking if the system can operate stably and efficiently when faced with a large volume of users.

7) Usability: Evaluating the user experience when using the system, including clarity of the user interface, easy navigation, and user satisfaction in using the system.
8) A/B Testing: Conducting experiments by testing two or more system variants to determine which yields better results. For example, comparing different recommendation methods to determine which provides more accurate or relevant recommendations.

3.8. Data Characteristics

The researcher accessed cafe data from an external source via a third-party API, encompassing essential details like location, ratings, and reviews. The dataset consisted of textual (cafe names, reviews) and numerical (ratings) data. It incorporated attributes such as cafe names, geographic locations, user ratings, and preferences from reviews. After the researcher cleansed the data by removing irrelevant reviews and validating missing or incomplete entries, the collected dataset involved approximately 252 cafes with over 25,000 user reviews.

3.9. Testing Location, Time, and Equipment Specifications

The testing for this research occurs at the author's residence, allowing flexibility in timing for implementing new algorithms. The hardware comprises an AMD Ryzen™ 5-3550H Processor with 16.00 GB RAM. The software setup includes Windows 10 OS, JavaScript, Python, and VS Code application.

3.9. Experimental Results

For this research, sample data was collected via web scraping from Google Maps, utilizing the Google Maps API. The study employed numerical ratings on a scale of 1 to 5. The distribution of the café dataset is presented in Table 1.

Table 1. Distribution of the dataset

Dataset	Total Data
Kafe	252
Rating Kafe	25759

A structured dataset was employed, as displayed in Table 2. This table outlines the cafe names and corresponding user ratings, forming the foundation for analysis.

Table 2. Dataset structure table

Name	King Cafe	Favorit Cafe	Boncafe Raya Gubeng	Urban Latte Gubeng	Glowstick Coffe	...
Patrick Tan	5	-	3	-	-	...
Bharin Rizqi	-	2	-	3	-	...
Andi Andi	-	5	-	-	-	...
Moad Muhank	-	-	4	-	5	...
...	...	...	...	...	...	...

The data standardization process was performed using the min-max normalization method. In this stage, each value in the dataset was compared to its mean and then normalized by dividing it by the difference between the maximum and minimum values. This process ensures the data has a consistent scale and facilitates comparison between different values.

After standardizing the dataset, the next step is to compute the similarity or closeness between users using the Cosine Similarity method. This involves calculating the average rating, squaring it, and taking its square root. The results of the cosine similarity calculation can be seen in the table below.

Table 3. Result *cosine similarity*

Name	King Cafe	Favorit Cafe	Boncafe Raya Gubeng	Urban Latte Gubeng	Glowstick Coffe	...
King Cafe	1.000	-0.021	-0.0063	-0.011	0.0152	...
Favorit Cafe	-0.02	1.000	-0.0084	-0.019	-0.005	...
Boncafe Raya Gubeng	-0.01	-0.008	1.0000	-0.006	-0.003	...
Urban Latte Gubeng	-0.011	-0.019	-0.0057	1.000	-0.005	...
Glowstick Coffe	0.015	-0.005	-0.0034	-0.005	1.000	...
...	...	...	...	...	....	...

The data above represents the outcomes derived from calculations using the cosine similarity method. Cosine similarity is employed to identify items that are highly similar to the item under consideration. In this context, if a user has provided ratings for various items, cosine similarity can be leveraged to recommend other items with a high similarity based on the given ratings.

After using the cosine similarity calculation method, it is further utilized to measure the extent to which users in the dataset have similar preferences. In this case, each user is represented as a rating vector in the rating matrix. The similarity calculation results in a similarity matrix among users.

3.10. Web Application Development

In this experiment, testing was conducted on the front-end and back-end of the system that utilizes a dataset of cafes and restaurants obtained from text-mining Google reviews and collaborative filtering.

Fig 4. Homepage Web Application

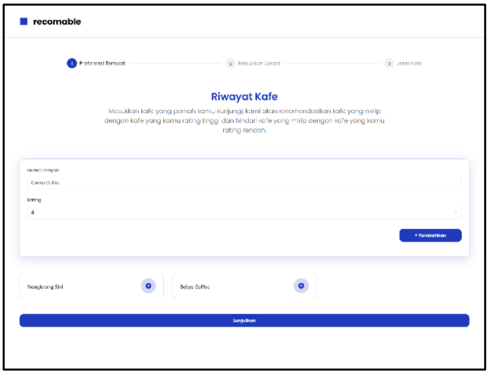

Fig 5. Visited Cafes

In this phase, users are directed to input their historical cafe visits, a pivotal step as collaborative filtering necessitates user data for comparison against the existing dataset, leading to tailored cafe recommendations based on cafes favored by the user and avoiding those with lower ratings, as illustrated in Fig 5.

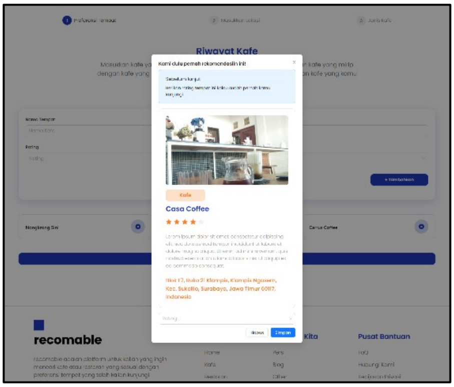

Fig 6. Specific Conditions Visited Cafes

Fig 6 showcases a user who has previously received recommendations and visits a cafe. Here, they can rate their experience, contributing to their cafe history data.

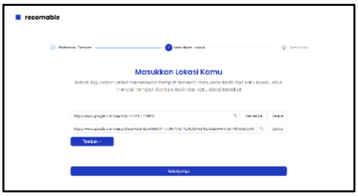

Fig 7. Location Selection Page

The location selection page, as displayed in Fig 7, enables users to manually input or auto-generate their location via GPS activation. This chosen location shapes the recommendation area, confining suggestions to cafes within a 10 km radius. If multiple locations are entered, the system calculates a midpoint for consideration.

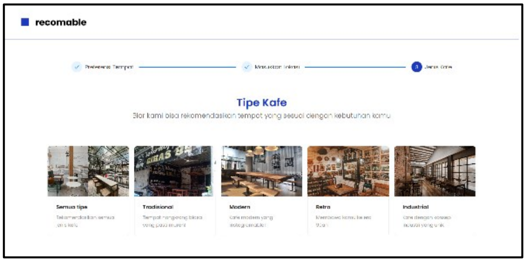

Fig 8. Cafe Theme Page

Users are presented with diverse cafe types, including traditional, modern, retro, and industrial, within the cafe theme page (Fig 8). This variety ensures alignment between recommended cafes and user preferences.

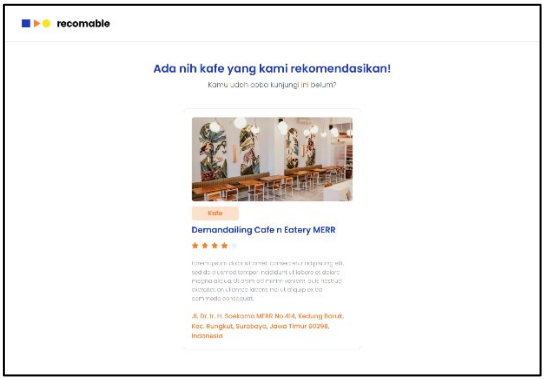

Fig 9. Recommendation Results Page

In the final stage, the recommendation results are displayed in Fig 9 based on the three cafes previously selected. The recommendations are generated based on the location, and only cafes within a maximum distance of 10 km from the user's location are recommended.

3.11. Experiment on Dataset Size

The experiment focused on understanding the impact of dataset size on recommendation outcomes. Specifically, the researcher aimed to provide cafe recommendations within the Sukolilo area. Two experiments were performed: one using a Surabaya-wide dataset with 25,000+ reviews, and another using a localized Sukolilo dataset with 3,500 reviews.

Table 4. User Preferences

Cafe Name	Rating
Tropikal Coffee	5
D'Greene	5
4K Coffee Shop	5

As depicted in Table 4, user preferences were captured, including cafe names and their corresponding ratings, such as Tropikal Coffee, D'Greene, and 4K Coffee Shop. To assess the effectiveness of the recommendation system, the researcher selected three cafes situated in the Sukolilo district

Table 5. Results of Dataset from One District

Cafe Name	Rating Similarity
Warung Sultan	0.04938
Havana Coffee	0.00862
Dara Jingga	-0.03263

Table 5 exhibits the outcomes achieved using the review dataset exclusively from the Sukolilo area. Notably, the results demonstrated cafe names along with their associated rating similarity scores. For instance, Warung Sultan exhibited a rating similarity of 0.10436, indicating a strong match.

Table 6. Results of Dataset from Entire Surabaya

Cafe Name	Rating Similarity
Warung Sultan	0.10436
Havana Coffee	0.07978
Goeboeg 99 Coffe Shop	0.06574

Conversely, Table 6 showcases the results obtained from using the review dataset encompassing the entire Surabaya. This comparison showcased consistent top-ranking recommendations, with some variance in the third position. This divergence could be attributed to the broader dataset's diversity across the entire city, as opposed to a localized dataset. Factors such as cafe availability, popularity, and diverse user preferences within the city played a role in these ranking differences.

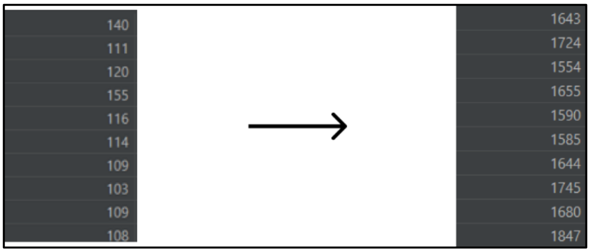

Fig 10. Latency from Small to Large Dataset

Furthermore, Fig 10 illustrates the latency in data processing between smaller and larger datasets. Evidently, processing smaller-scale data required less time (average of 120ms) compared to processing larger-scale data (average of 1600ms). This observation underscores the impact of dataset size on processing efficiency.

4. CONCLUSION

This research is highly relevant to cafe recommendations and significantly addresses the problem. The developed system successfully assists users in finding cafes that match their preferences and benefits cafe owners in increasing their popularity. The proposed collaborative filtering method offers relevant recommendations based on user preferences, introducing a novel approach to cafe recommendation systems. The designed system, with its web-based client-server architecture using Flask and ReactJS, ensures a smooth user experience and fast response speed.

The system's experimental evaluation showed promising results, demonstrating its ability to offer accurate cafe recommendations for a substantial user base and extensive rating data. Initial tests with 284 cafes displayed cosine similarity values below 0.1, expected to improve with larger datasets. This implies better recommendation accuracy as the dataset grows. The system maintained quick response times, averaging 120ms for smaller datasets and 1600ms for larger ones, highlighting its adaptability to varying data scales.
This research has successfully devised a pertinent, efficient, innovative, and well-crafted cafe recommendation system and illuminated areas warranting enhancement for optimal system performance. This study lays a robust groundwork for forthcoming advancements in the domain of cafe recommendation systems and plays a pivotal role in advancing the knowledge landscape within this field.

In future work, further exploration of advanced recommendation algorithms, such as collaborative filtering or hybrid approaches, is recommended to enhance the system's accuracy and recommendation quality. Additionally, incorporating user feedback and implementing a feedback loop mechanism would contribute to continuous improvement and user satisfaction. Further research can also focus on integrating additional data sources and exploring the application of machine learning techniques for more sophisticated recommendation capabilities.

This research yields substantial contributions to recommendation systems and casts light on developing and assessing web-based recommendation systems. The unveiled findings chart a trajectory for future strides in personalized recommendation technologies, proving advantageous for end-users and enterprises across diverse domains.

REFERENCES

[1] Yulia Nurliani Lukito and Anneli Puspita Xenia, "Café as third place and the creation of a unique space of interaction in UI Campus", University of Indonesia, (2017).

[2] Bernhardt, Andrew, and S. Linda. "Creating Third Places: Places Where Communities Gather", Downtown Economics, University of Wisconsin-Extension, (2010).

[3] E. Dwi, M. Wijaya, and P. Marbun. "The Influence of Lifestyle, Physical Environment, and Menu Variety on Customer Loyalty through Customer Satisfaction in the Coffee Shop", International Journal of Research and Review, (2020).

[4] Chua BL, Karim S, Lee S, Han H. "Customer Restaurant Choice: An Empirical Analysis of Restaurant Types and Eating-out Occasions", Int J Environ Res Public Health, (2020).

[5] Singh, P. K., Pramanik, P. K. D., Dey, A. K., & Choudhury, P., "Recommender Systems: An Overview, Research Trends, and Future Directions", National Institute of Technology Durgapur, (2021).

[6] Prasetya, and Chandra, "Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor", Universitas Gadjah Mada, (2017).

[7] Heramawan, A., & Evan, "Hotel Recommendation System Using SAW (Simple Additive Weighting) and TOPSIS (The Technique for Order of Preference by Similarity to Ideal Solution) Method." Buddhi Dharma University, (2021).

[8] Rui-sheng Zhang, Qi-dong Liu, Chun-Gui, Jia-Xuan Wei, Huiyi-Ma, "Collaborative Filtering for Recommender Systems", Lanzhou University, (2014).

[9] Ching-Seh, Deepti, and Bhandary, "Movie Recommendation System Using Collaborative Filtering". San Jose State University, (2018).

[10] Hikmatyar, M., & Ruuhwan, "Book recommendation system development using user-based collaborative filtering." Universitas Perjuangan Tasikmalaya, (2022).

[11] Asani, E., Vahdat-Nejad, H., & Sadri, J, "Restaurant recommender system based on sentiment analysis", Journal of Artificial Intelligence and Data Mining, (2021).

[12] S. Najafi, Z. Salam, "Evaluating Prediction Accuracy for Collaborative Filtering Algorithms in Recommender Systems", KTH Royal Institute Of Technology School Of Computer Science And Communication, (2016).

[13] Z. Bao, "Web Scraping", University of Washington Seattle, (2017).

[14] Vinayak Hegde, T S Aswathi, R Sidharth, "Student Residential Distance Calculation using Haversine Formulation and Visualization through GoogleMap for Admission Analysis", Amrita University, (2016).