Daily activity recommender system

André Couto
uc2019215056@student.uc.pt

1 Introduction

In the fast-paced world of today, where technology permeates every aspect of our everyday lives, the relevance of recommender systems has never been more significant. These systems, powered by sophisticated algorithms and machine learning models, leads us through a myriad of choices in e-commerce, entertainment, and even our social interactions.

However, one area that remains relatively unexplored is the recommendation of daily activities. In an era where time is a luxury and the paradox of choice can lead to decision paralysis, many of us often find ourselves at a loss on how to best spend our time. This is why this project is so important.

This project aims to bridge this gap by providing personalized recommendations on how to spend one's day based on individual preferences, historical data, and contextual factors such as weather and time of day. Whether it's suggesting a new book to read on a rainy afternoon or recommending a local hiking trail on a sunny weekend, this system will serve as a personal concierge, helping users make the most of their time.

2 State of the art

Recommender systems are trained to comprehend the preferences, decisions, and characteristics of individuals or products, utilizing data collected with the intent of understanding these interactions to suggest new actions [Carol McDonald and Gabriel Moreira,] . Despite the plethora of articles published on recommender systems, few recommend daily activities for everyday practice. Some systems integrate social network data and location-based services to offer personalized suggestions to users [Khoshnood, 2012], considering spatial factors and personal information to recommend restaurants, museums, and concerts in a specific area. Other systems offer more personalized services, such as Murshid, a service-based mobile tourist advisor application [Echtibi *et al.*, 2009]. Murshid provides personalized, context-aware information to tourists in a new country. It operates on a client-server model where the client application runs on the user's mobile device and communicates with the server for information. Murshid offers various services like special event notifications, sightseeing recommendations, weather forecasts, currency exchange, language translation, and location bookmarking and sharing. In more specific cases [Majeed *et al.*, 2020], a recommender system integrated into a mobile app for tourists in alpine destinations is discussed. The system provides personalized recommendations for daily activities based on view data, enhancing the tourist experience. It addresses the challenge of choosing from numerous options in an unfamiliar location by offering suggestions tailored to the user's interests. The system uses implicit user feedback, specifically view duration, to generate these recommendations. Finally, LARS is a recommender system that primarily relies on the location of users and items to recommend, in this case, movies [Sarwat *et al.*, 2014]. This method is highly accurate and scalable; however, it has not been used for this purpose, like those previously mentioned. Hence the need to address this topic.

3 Objectives

The objectives for this project are as follows:

- **Design and Development:** Develop a recommender system that suggests daily activities to users based on their preferences, location, and other relevant factors. The system should be user-friendly and provide accurate and personalized recommendations
- **Personalization:** Ensure that the recommendations provided by the system are personalized and relevant to each user. The system should take into account the user's interests, preferences, and past activities
- **Contribution to the Field:** Contribute to the field of recommender systems by addressing a gap in the literature for daily activity recommendation. The findings from this project could provide valuable insights for future research in this area.

4 Methodology

- **Data Construction:** Given the scarcity of data pertaining to daily activities, the solution was to construct a dataset that would enable the development of a daily recommendation system.
- **Data Preprocessing:** Clean and preprocess the data to ensure it is in a suitable format for the recommender system. This could involve handling missing values, normalizing data, and transforming categorical variables into numerical ones.

- **Feature Extraction:** Identify important features that will influence the recommendations. These could include user location, time of day, weather conditions, user interests, and past activities.

- **Model Building:** Use a suitable machine learning algorithm to build the recommender system. According to the [Carol McDonald and Gabriel Moreira,]), matrix factorization and deep learning models are commonly used in recommender systems. We may need to experiment with different models to find the one that provides the most accurate and personalized recommendations.The exploration of fundamental techniques for recommendation systems, namely collaborative, content-based, and knowledge-based approaches, is undertaken to discern the advantages and disadvantages inherent to each method.

- **Training:** Train the model using the collected data. This involves feeding the data into the model so it can learn the relationships between different variables and how they influence the recommendations.

- **Evaluation:** Evaluate the performance of the model using appropriate metrics such as precision, recall and F1 score. This will help us to understand how well the model is performing and where improvements can be made. Conduct thorough testing of the recommender system under different scenarios and conditions. This includes testing the system's performance with different types of users, varying numbers of users, and different levels of user activity.

5 Materials

Just as there is a gap in the literature regarding daily activity recommendation systems, there is also a gap in the necessary data to develop such a system. The aforementioned articles utilize observational data that is not readily available for use. Other tools that provide datasets also lack the necessary data to create this type of system. Therefore, a temporary solution was implemented. With the assistance of Chat GPT, synthetic data was generated for the purpose of training the recommendation system. Figure 1 shows a sample of this dataset.

	Name	Likes	Activity	Local	Date	Hour	Weather
0	Jane	artistic_activities	Pottery	Lake Richard	2023-10-14	08:58	Sunny
1	Jane	artistic_activities	Writing	East Jonathan	2023-10-05	13:50	Rainy
2	Alice	outdoor_activities	Chess	Tinafort	2023-10-08	15:10	Rainy
3	Doe	night_activities	Stargazing	Stephenchester	2023-10-10	13:09	Snowy
4	Charlie	sports_activities	Basketball	West Rodney	2023-10-15	11:43	Cloudy
...	...	...	...	...	...	...	...
995	Charlie	sports_activities	Baseball	Clarkborough	2023-09-24	18:09	Cloudy
996	John	physical_activities	Rugby	Yatesbury	2023-09-22	07:01	Windy
997	Doe	night_activities	Stargazing	New Markmouth	2023-10-06	13:11	Windy
998	Bob	water_activities	Stargazing	South Jamesport	2023-10-02	00:03	Rainy
999	Mallory	extreme_activities	Kayaking	Garcialand	2023-10-08	20:57	Rainy

Figure 1: Sample of dataset

6 Results and Discussion

6.1 Collaborative Filtering

The initial technique employed was collaborative filtering, a common method for generating recommendation systems. This technique is widely used by major commercial e-commerce sites, is well-understood, and has numerous algorithms and variations available for implementation. While it is typically used for book and movie recommendations, this project sought to adapt this technique to daily activity recommendation systems.

User-based nearest-neighbor collaborative filtering

The first strategy involved leveraging the "wisdom of the crowd". The initial step was to identify the active user and the target activity. The second step, termed "Find the peers", involved identifying users who have similar preferences to the active user using the Pearson correlation coefficient, a common measure in such scenarios. The final step, which is repeated for all unrated activities, involves estimating the active user's rating for the target activity through a weighted average of the peers' ratings for the target activity.

The outcomes of this strategy did not meet the anticipated standards, with specific results delineated in Table 1.

As can be observed in the table, the results do not align with the types of activities each user prefers. Consequently, a new approach was employed.

Item-based collaborative filtering

This time, an item-based collaborative filtering approach was employed, which entails studying the similarity between items (activities) rather than between users. However, this approach presupposes the existence of ratings for the activities, a feature not present in our dataset, hence the relatively poor results yielded by this method (see Table 2)

As discernible from the table, the results appeared to be identical for all users, albeit in a different order. Even after introducing randomness into the recommendations, the issue persisted. Thus, the problem was hypothesized to lie within the dataset. Upon further analysis, it was concluded that the dataset is sparse, given that it contains 9797 duplicates.

Instead of considering each occurrence of an activity as a distinct interaction, the solution was aggregate them and treat each unique user-activity pair as a single interaction. Employing this strategy yielded intriguing results, indeed the most promising within the scope of this project, as can be observed in Table 3.

6.2 Content-based recommendation

Another technique explored in this project was content-based recommendation, which enables the recommendation of activities based on those previously engaged in. In other words, it recommends activities that are similar to the user's preferences.

Simple Approach

A simple approach of this technique is to compute the similarity of an unseen activity with the user profile based on the keyword overlap using the Cosine Similarity.

In employing this technique, we encountered one of its major challenges. Given that the dataset contains only eight

User	Likes	Recommended Activities
Alice	Outdoor activities	’Rafting’, ’Surfing’, ’Scuba Diving’, ’Facial Treatment’, ’Massage’
John	Physical activities	’Kite Flying’, ’Skiing’, ’Kayaking’, ’Rafting’, ’Surfing’
Jane	Artistic activities	’Beach Volleyball’, ’Kite Flying’, ’Skiing’, ’Kayaking’, ’Surfing’
Doe	Night activities	’Meditation Retreat’, ’Beach Volleyball’, ’Kite Flying’, ’Skiing’, ’Kayaking’
Bob	Water activities	’Kite Flying’, ’Skiing’, ’Beach Volleyball’, ’Meditation Retreat’, ’Snowboarding’
Charlie	Sports activities	’Meditation Retreat’, ’Kayaking’, ’Rock Climbing’, ’Ice Skating’, ’Running’
Eve	Home activities	’Meditation Retreat’, ’Beach Volleyball’, ’Kite Flying’, ’Skiing’, ’Kayaking’
Mallory	Extreme activities	’Beach Volleyball’, ’Meditation Retreat’, ’Sailing’, ’Fishing’, ’Bowling’

Table 1: Recommended activities for each user using User-based collaborative filtering

User	Likes	Recommended Activities
Alice	Outdoor activities	’Stargazing’, ’Hiking ’, ’Walking the dog’, ’Movie Watching’, ’Potluck Dinner’
John	Physical activities	’Walking the dog’, ’Movie Watching’, ’Potluck Dinner’, ’Stargazing’, ’Hiking ’
Jane	Artistic activities	’Movie Watching’, ’Potluck Dinner’, ’Stargazing’, ’Hiking ’, ’Walking the dog’
Doe	Night activities	’Movie Watching’, ’Walking the dog’, ’Potluck Dinner’, ’Hiking ’, ’Stargazing’
Bob	Water activities	’Stargazing’, ’Hiking ’, ’Walking the dog’, ’Movie Watching’, ’Potluck Dinner’
Charlie	Sports activities	’Hiking ’, ’Walking the dog’, ’Stargazing’, ’Potluck Dinner’, ’Movie Watching’
Eve	Home activities	’Potluck Dinner’, ’Hiking ’, ’Movie Watching’, ’Stargazing’, ’Walking the dog’
Mallory	Extreme activities	’Movie Watching’, ’Scuba Diving ’, ’Walking the dog’, ’Hiking ’, ’Stargazing’

Table 2: Recommended activities for each user using Item-based collaborative filtering

distinct individuals, there was an overlap with the user profile. To address this issue, we disregarded the input user in the similarity calculation. Consequently, the results obtained were akin to those depicted in Table 4 .

As can be observed, the activity suggested for Alice appears to be incongruous given her preference for outdoor activities, yet a movie was recommended. Therefore, this strategy may require further refinement.

Term-Frequency - Inverse Document Frequency (TF-IDF)

One method to enhance this simple approach is to employ the standard measure: Term Frequency-Inverse Document Frequency (TF-IDF). This technique encodes the data as a weighted term vector, utilizing Term Frequency to quantify how frequently a term appears, and Inverse Document Frequency to diminish the weights of terms that occur more frequently. This refined strategy could potentially yield more accurate and relevant recommendations.

However, as can be observed in Table 5, the outcomes did not surpass those achieved by the simplest strategy, thereby reinforcing the notion that there is considerable room for improvement.

Term-Frequency - Inverse Document Frequency (TF-IDF) with nearest neighbors

The previous approach was enhanced using the nearest neighbors method. However, the results remained unchanged, even when different metrics were employed in place of cosine similarity. This phenomenon is likely attributable to the limitations of the dataset. The users exhibit similar activity preferences, leading to identical recommendations despite variations in the parameters of the same approach (see table 5).

6.3 Knowledge-based recommendation

Lastly, the final technique employed was a knowledge-based recommendation, which, as the name suggests, relies on prior knowledge of the user’s past actions and thus requires such information. A straightforward approach yielded quite satisfactory results, given the acceptance that the weather conditions and activities performed are not personalized in the dataset used for this project (see Table 6).

However, the results were not particularly favorable for certain types of activities. The case of extreme activities is the most striking, with recommendations such as watching a movie, walking the dog, or having a potluck dinner. This may be due to the fact that the weather was snowy and the likes and weather carried equal weight in the search.

Knowledge-based Recommendation using similarity functions

An enhancement to the previous strategy was achieved through the use of similarity functions, specifically a technique known as fuzzy matching. This method calculates the similarity between the query and each item based on a certain metric. The results showed a slight improvement, particularly in the realm of extreme activities, where more appropriate recommendations were made compared to the previous approach, including activities such as walking the dog, stargazing, skiing, surfing, and having a potluck dinner.

Utility-based recommender system (MAUT)

Another strategy to enhance this approach is by employing a utility-based method, specifically the Multi-Attribute Utility Theory (MAUT). In this context, each item is evaluated according to a predefined set of dimensions that offer an aggregated perspective on the fundamental properties of the item. Following the implementation, the results were indeed quite

User	Likes	Recommended Activities
Sunny	Outdoor activities	'Beach Volleyball', 'Kayaking', 'Kite Flying', 'Meditation Retreat ', 'Skiing '
Rainy	Physical activities	'Beach Volleyball', 'Meditation Retreat ', 'Dancing', 'Gym Workout', 'Hair Salon Visit '
Snowy	Artistic activities	'Meditation Retreat ', 'Art Class', 'Art Gallery Visit', 'Ballet Attendance', 'Book Club Participation'
Doe	Night activities	'Hiking ', 'Movie Watching', 'Potluck Dinner', 'Stargazing', 'Walking the dog'
Bob	Water activities	'Kayaking', 'Rafting', 'Scuba Diving ', 'Surfing', 'Canoeing'
Charlie	Sports activities	'Beach Volleyball', 'Kite Flying', 'Rafting', 'Scuba Diving ', 'Skiing '
Eve	Home activities	'Baking ', 'Cooking ', 'Gardening ', 'Hiking ', 'Meditation '
Mallory	Extreme activities	'Kayaking', 'Kite Flying', 'Rafting', 'Scuba Diving ', 'Skiing '

Table 3: Recommended activities for each user using Item-based collaborative filtering with aggregation

User	Likes	Recommended Activities
Alice	Outdoor activities	Movie Watching
John	Physical activities	Hiking
Jane	Artistic activities	Movie Watching
Doe	Night activities	Hiking
Bob	Water activities	Hot Spring Bath
Charlie	Sports activities	Kite Flying
Eve	Home activities	Facial Treatment
Mallory	Extreme activities	Rock Climbing

Table 4: Recommended activities for each user using Content-based (Simple approach)

favorable when a weight of 0.7 was assigned to "Likes" and 0.3 to "Weather". This weight assignment is logical considering that the dataset randomly generates the weather at the time the activity was performed. If it were possible to experiment with another dataset, the weights could be adjusted according to their significance. Nevertheless, as we focused on previous reviews of extreme activities, the results showed significant improvement. The system suggested activities such as Archery, Rock Climbing, Hiking, Kayaking, and Ice Skating.

7 Conclusion

This project was conceived with the aim of developing a daily activity recommendation system for users, based on their preferences, locations, weather conditions, and past activities. Given the scarcity of attempts to address this subject and the consequent lack of necessary data to construct the intended system, a simulated dataset was created to realize this project. This dataset had considerable limitations concerning the weather, time, and locations where the activities were performed, as these were randomly generated and thus inconsistent. Bearing this aspect in mind, the project's results were better when only activities and preferences were used (although including weather also yielded satisfactory results most of the time). Initially, the exploration of machine learning models, specifically Support Vector Machines, was employed, but the results were too poor due to the dataset's limitations and hence are not discussed in this paper.

Thus, three different techniques were utilized: Collaborative Filtering, Content-Based Recommendation, and Knowledge-Based Recommendation. Two different methods were tested with Collaborative Filtering: User-Based Nearest-Neighbor Collaborative Filtering, which showed results below the desired level, and Item-Based Collaborative Filtering, which after aggregation showed one of the best results obtained in this project. With the Content-Based Recommendation strategy, a simple approach was tested, which yielded somewhat poor results, as well as Term Frequency-Inverse Document Frequency (TF-IDF), which yielded similar results, and TF-IDF with Nearest Neighbors, which also showed similar results.

Finally, a Knowledge-Based Recommendation was implemented, which for this project and with this dataset turned out to be the most reliable, recommending activities that made the most sense, whether with similarity functions (fuzzy) or with MAUT.

In conclusion, this project reveals that recommendation system techniques have their advantages and disadvantages and should be used according to the type of problem at hand, as well as the dataset. It is also important to mention that as future work, surveys can be conducted to understand what activities people perform for fun, where and at what times, in order to implement a more personalized system, allowing all the techniques used in this project to evolve.

References

[Carol McDonald and Gabriel Moreira,] Carol McDonald and Gabriel Moreira. How to build a winning recommendation system. https://developer.nvidia.com/blog/how-to-build-a-winning-recommendation-system-part-1/.

[Echtibi *et al.*, 2009] Ahmed Echtibi, Jamal Zemerly, and Jawad Berri. A service-based mobile tourist advisor. *International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM)*, 1:177–187, 03 2009.

User	Likes	Recommended Activities
Alice	Outdoor activities	Movie Watching
John	Physical activities	Hiking
Jane	Artistic activities	Movie Watching
Doe	Night activities	Hiking
Bob	Water activities	Hot Spring Bath
Charlie	Sports activities	Kite Flying
Eve	Home activities	Facial Treatment
Mallory	Extreme activities	Rock Climbing

Table 5: Recommended activities for each user using Content-based (TF-IDF)

Weather	Likes	Recommended Activities
Sunny	Outdoor activities	’Movie Watching’, ’Golfing’, ’Hiking’, ’Potluck Dinner ’, ’Horseback Riding ’
Rainy	Physical activities	’Basketball’, ’Basketball ’, ’Walking the dog’, ’Snowboarding’, ’Gym Workout ’
Snowy	Artistic activities	’Sewing ’, ’Stargazing’, ’Art Gallery Visit’, ’Potluck Dinner’, ’Potluck Dinner’
Cloudy	Night activities	’Potluck Dinner’, ’Hiking ’, ’Potluck Dinner’, ’Movie Watching’, ’Hiking ’
Windy	Water activities	’Hot spring bath’, ’Surfing’, ’Rafting’,’Kayaking’, ’Movie Watching’
Sunny	Sports activities	’Movie Watching’ , ’Scuba Diving ’, ’Basketball’, ’Kite Flying’,’Beach Volleyball’
Rainy	Home activities	’Facial Treatment’, ’Reading ’, ’Hiking ’, ’Barbecue’, ’Farming Activities’
Snowy	Extreme activities	’Rock Climbing’, ’Walking the dog’ ,’Hiking’, ’ Walking the dog’, ’Potluck Dinner’

Table 6: Recommended activities for different likes and weather using knowledge-based

[Khoshnood, 2012] Fatemeh Khoshnood. Designing a recommender system based on social networks and location based services. *International Journal of Managing Information Technology*, 4(4):41–47, 2012.

[Majeed *et al.*, 2020] Tahir Majeed, Aline Stämpfli, Andreas Liebrich, and René Meier. What is of interest for tourists in an alpine destination: personalized recommendations for daily activities based on view data. *Journal of Ambient Intelligence and Humanized Computing*, 11, 11 2020.

[Sarwat *et al.*, 2014] Mohamed Sarwat, Justin Levandoski, Ahmed Eldawy, and Mohamed Mokbel. Lars*: An efficient and scalable location-aware recommender system. *IEEE Transactions on Knowledge and Data Engineering*, 26:1–1, 06 2014.