Analysis of Personalized Tourism Recommender Systems

Dipit Sharma, Eklavya Banwala, Imane Elouaghzani and Rahul Katarya

January 27, 2023

Analysis of Personalized tourism recommender systems

Dipit Sharma
Department of Computer Science and Engineering
ds.techin@gmail.com
Delhi Technological University
Delhi, India

Eklavya
Department Of Computer Science and Engineering
eklavyab3@gmail.com
Delhi Technological University
Delhi, India

Imane Elouaghzani
Department of Computer Science and Engineering
imanecodeblack@gmail.com
Delhi Technological University
Delhi, India

Rahul Katarya
Department of Computer Science and Engineering
rahuldtu@gmail.com
Delhi Technological University
Delhi, India

***Abstract*— Traveling is one of the most important leisure pursuits in life. While on vacation from one's regular routine, one experiences a new culture, new food, and new sights. As a result, recommending tourist destinations based on the user's interests may enable an individual to have incredible new experiences. Currently, different types of recommendation systems are used in the industry. Our primary goal is to compare all of the methods and determine which one is the best recommendation system for tourism. In this paper, we investigated various recommendation system methods used in tourist recommendation systems. The goal of this research is to identify all of the recommendation systems used in the industry. We tested which method would be the best for tourism recommendation by implementing them on a common dataset. Since content-based is item-based and collaborative filtering is user-based, these two alone are insufficient to produce an accurate result; therefore, a deep learning model is required in addition to these two. According to the parameters and factors considered for tourism, we discovered that the hybrid is the best option. Additionally, no recommendation system considered personality as a factor, so we are currently developing a recommendation system that does.**

***Index Terms*— Collaborative Filtering, Content Filtering, Cosine Similarity, Hybrid Method, TF-IDF Equation**

I. INTRODUCTION

In 2021-22, the total number of domestic visitors recorded a 98% year-on-year growth [1]. In order to find the best method to recommend tourist destinations for a user, we must find out about the likes and dislikes of the user. In Section 2 we will discuss what methods and mathematical formulae are used for creating recommendation systems - mainly TF-IDF and Cosine Similarity methods. Then in Section 3, we jump into discussing 3 main types of filtering methods. We have mentioned their work with dataset examples, advantages, and disadvantages. Fig. 1 represents the basic functioning of any production-level recommendation system. In Section 4, we finally talk about the conclusion and future work, we talk about how we plan to remove the current inefficiencies and try to implement a completely new method of recommendation system. But first, let's highlight the factors that each visitor

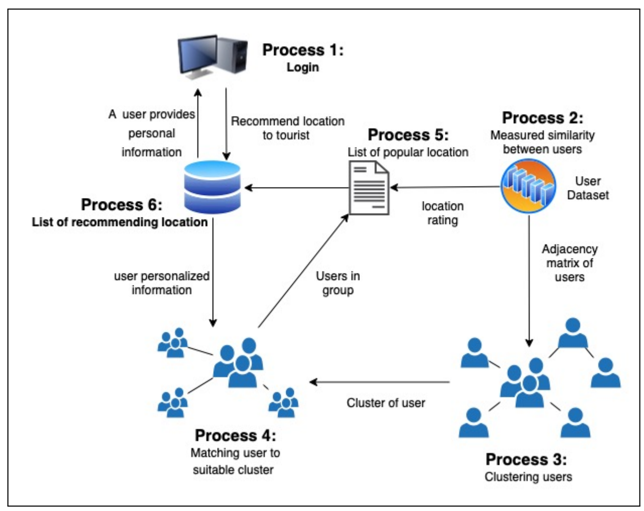

Fig. 1. Working flow of a tourist recommender system

should consider while arranging a trip. There are different types of people with different resources and personal characteristics. So we discuss the important aspects, such as budget, destination economy, weather, religious or cultural importance, tourist personality, number of people traveling, etc [2]. These influencing factors are collected from a variety of research publications, studies, and surveys. There are all varieties of personalities represented: upper-class, middle-class, disabled, and so on [3].

A. Types Of Tourism

These are the following types of tourism, as briefed in Fig. 1.

1) Domestic Tourism: In this sort of tourism, travelers visit different parts of the same country. It is inexpensive and knowing more about his or her country. For example, many Delhi residents vacation in Manali.

2) Health and Wellness Tourism: This type of tourism aids in the recovery of travelers from psychological and physical stress. It is an essential type of tourism. This usually entails places that are tranquil and near nature. India is an excellent demonstration of medical tourism since the physicians are well-qualified and speak English, and the treatment is also less costly.

3) Dark Tourism: Some areas of the country have passed through dark and tough periods. Travelers visit to learn about the incidents and their historical importance. For example, Jallian Wallan Bagh is well-known for the genocide committed during colonial rule.

4) Countryside Tourism: People would like to travel to remote regions to witness traditional and old-fashioned lifestyles. Solo tourists are the most likely to do this. Chokhi Dhani in Jaipur, for example, provides a taste of real Indian Rajasthani culture.

5) Business tourism: This is mostly a business trip. for example, import/export, employment, and inspection reasons, among others.

6) Educational tourism: This sort of tourism entails traveling with the intention of learning and studying new things. This is mostly for students and researchers.

B. Factors Depending On Tourist

These are the factors that change from traveler to traveler[5], as briefed in fig. 2.

1) Physical Accessibility: This refers to the physical amenities that are available at a location. This is only for travelers with disabilities. This facility most likely provides more than you realize. These amenities can help older persons become more self-sufficient.

2) Convenience: It refers to overcrowded food courts and a lack of basic essentials, which can make users feel disadvantaged. However, customers who do not require luxury but low-budget services might make advantage of these establishments. As a result, it is dependent on the user's preferences.

3) Security: Females and families traveling with infants may be more apprehensive about male users.

4) Peer Knowledge: Peers' or coworkers' recommendations may influence one's decision to visit a certain location. In general, individuals prefer to visit sites that have previously been visited by their peers and have gotten positive feedback.

5) Service: It includes the services offered during the journey. Hotel, transportation, tour guides, and other services may be provided.

C. Factors Depending On Location

These are the factors that change from location to location [6], as briefed in fig. 2.

1) Natural attractions: Climates, natural landscapes, beaches, and so forth. These criteria can be utilized to match the user's travel preferences.

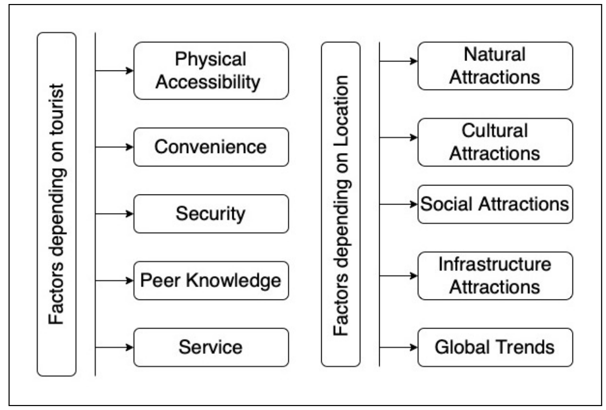

Fig. 2. Different parameters depending on tourists and locations

2) Cultural attractions: This represents the destination's culture. includes food, rituals, clothing, and so forth.

3) Social attractions: This consists of a location's hospitality as well as a welcoming population.

4) Infrastructure attractions: It focuses on luxury hotels, city lights, and breathtaking views. Essentially, the city's contemporary attractions and social life

5) Global Trends: Economic (low pricing), political (safe destination), environmental (environmental protection measures), demographic (health and leisure amenities), and technical aspects are all considered (internet connectivity).

II. RELATED WORKS

For related works, please refer to Table 1. In the table, we have discussed different papers that describe several recommendation methods for tourism. We have listed the inefficiency or scope of improvement in the respective techniques used. The main disadvantage was the methods were not personalized for a person. They are pretty general, for example, luxury, and high/low budgets are relative terms. They cannot be the same for every user. We need to take the status of the user in the account. The "status" means financial, health, emotional, etc.

The disadvantages can be summarised as follows:

- **limited dataset** - in the papers we considered the dataset used is not sufficient for consistent results, the data is neither generalized nor elaborate
- **subjective criteria for calculation** - the criteria used for the calculation of data in the dataset present is subjective to an individual user and not valid for a variety of people
- **factors** - the dataset also has a lack of factors taken into account that can influence a tourist
- **sentimental analysis** - none of the papers considered the personality of the users, which has a major contribution to the decision-making of any human

TABLE I

A SYSTEMATIC REPRESENTATION OF CURRENT WORKS AND THEIR SHORTCOMINGS

Author	Methods/Approaches Used	Limitations
Jinglei et al.[7] 2022	Based on Personalized Recommendation Algorithms (hybrid method based on social reviews)	Due to the limitations of research methods and research conditions, The variables selected in this study have not been fully considered, and the source of the questionnaire samples is relatively limited
Luis et al.[8] 2022	Based on User based Content Filtering and demographic information approach	the user's rating to tourist site does not exist.
Joseph et al.[9] 2021	context-based filtering	Images are classified into only two categories happy and sad while there can be meant emotions.
Praditya et al.[10] 2021	hybrid method	The sub-criteria used for data calculation is subjective as 50,000 Rp may be cheap for some but expensive for others.
Garipelly et al.[11] 2021	collaborative filtering	Fewer data and less accurate
Chiwong et. al.[12] 2020	Personalized travel planning system (PTPS)	Provides only recommendations based on other users and there is no cost or time estimates.
Z. Duan et. al.[13] 2020	The hybrid method used with deep learning	not pay attention to the efficiency of the OP orientation problem
Alnogaithan et. al.[14] 2019	input-output information recommendation technique, sentiment analysis, and rating review	Proposed tourism recommendation system for hotels based on user reviews, ratings, and sentiment. Limited to the local area and previous dataset
Ye J et. al.[15] 2019	Completely based on deep learning using social media data (LSTM)	The model includes only complement data and negative news which may affect the user influence.
Saputra et. al.[16] 2019	collecting tourist contextual data and making a database to support the recommendation system	Showed a correlation between user and tour data results and how condition change can affect decision-making for users.

III. TECHNIQUES USED IN TOURISM RECOMMENDATION SYSTEM

A. *TF-IDF Equation*

Similarities are calculated based on an overview of the place and location category. Vectors are constructed in order to create a similarity matrix, refer to Fig. 3 for example, it has few sentences with certain similarities. The cosine similarity of all vectors is then calculated[17]. So we considered a dataset that had 1000 destinations with their names, categories, and an overview of the location. So now, we've created a vector array using the TF-IDF Vectorizer. Here is a brief description of how the TF-IDF library works:

$$TD-IF = TF(t,d) * IDF(t) \tag{1}$$

In eq.(1), t is the number of times the term t appears in document d.

$$TF(t,d) = \sum_{x \in d} FR(x,t) \tag{2}$$

$$FR(x,t) = \begin{cases} 1 & \text{if } x = t \\ 0 & \text{if } otherwise \end{cases} \tag{3}$$

Here the TF(t,d) returns how many times the term t present in document d.

$$IDF(t) = \log \frac{|D|}{1 + |d : t \in d|} \tag{4}$$

In eq.(4)

$$|d : t \in d|$$

is the number of documents where the t term appears when the term frequency satisfies

$$TF(t,d) \neq 0$$

we are only adding 1 to the formula to avoid zero division.

In order to create the vector array we need to pass our data into the function (pseudo code): TF-IDF(dataset, name of the column). Here, the function takes in two arguments, the dataset is the main dataset that we have, and the name of the column is the name of the column in the dataset that we want the vector array to be based on.

B. *Cosine Similarity*

The cosine similarity method determines how similar two vectors are to one another. If the angle between them is less then vectors are similar and vice-versa. The equation for the cosine similarity between vector A and vector B is as follows:

$$S(A,B) = cos(\theta) = \frac{A.B}{|A|.|B|} \tag{5}$$

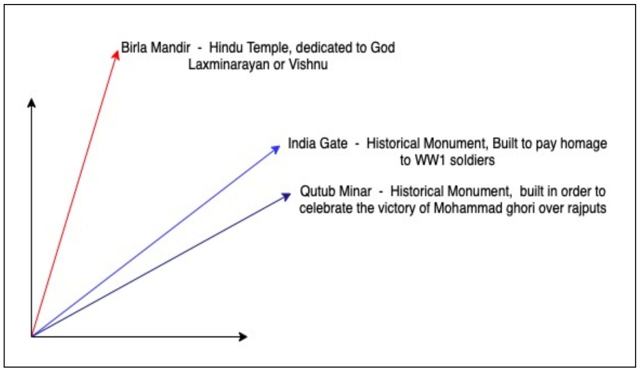

Fig. 3. Different sentences represented in the form of vectors

TABLE II
LOCATIONS AND THEIR CATEGORY - A SAMPLE DATASET

Locations	**Categories**
Jaipur Zoo	Wildlife
K.V. Garden	Park
Anokhi Museum	Museum
Akshardham Temple	Pilgrimage
Albert Hall Museum	Museum
Hawa Mahal	Heritage
Birla Mandir	Pilgrimage
Nahargarh Fort	Heritage

IV. DIFFERENT TYPE OF FILTERING METHODS

A. *Content-based recommendation system*

It is advised in content-based filtering to take into account the data provided about the item or the item's past history. The entire concept is built on the type of thing that the user loves or may enjoy. As a result, item-based matching. [18] . Consider a list of places and their respective classifications. To implement a content-based recommendation system, a vector matrix is built using TF-IDF formulas. Using vector points and cosine similarity methods, a similarity matrix is created as shown in Fig.4. In Fig.4, the blue regions among orange regions represent similar locations.

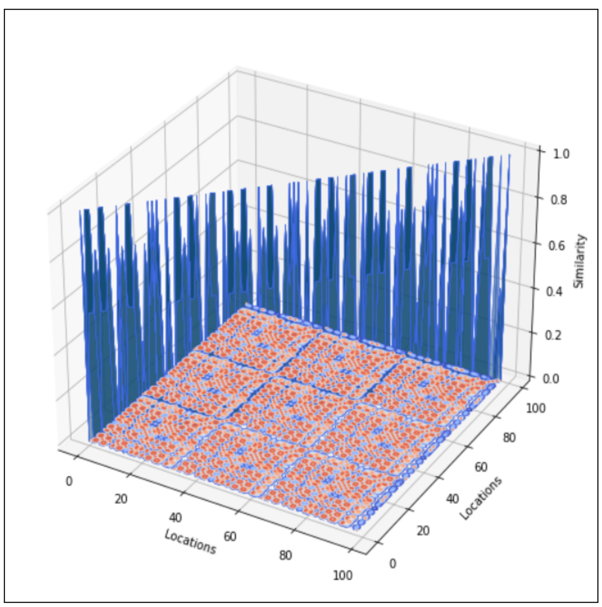

Fig. 4. Graph representing the similarity matrix among different locations

Advantages

- All the data is personalized to the user; they only get what they want. So anything that may not be of interest to them is not recommended.
- There is no need for any other user's data.

Disadvantages

- It is based on the actual interests of the users, and if those interests change or increase, it may expand accordingly.
- It will not be able to provide new items.

B. *Collaborative recommendation system*

This method of filter uses another similar user to recommend the content. Briefly, using the likes and dislikes of one user to filter content for another user is not restricted to one similar user; it can contain several users, perspectives, and data sources .

It has two types of inputs:

- **Implicit**: the system automatically recognizes the likes and dislikes of the user
- **Explicit**: The system needs input, whether the user likes the content or not.

So let's consider some data (table 2). Now, in collaborative filtering, we can process it in two ways:

- **1-D**: Creating a single line that indicates users' interest in the type of tourist location according to their age. Now if we get where user-4 (the new user) lies on this line, we can easily provide him with some relevant recommendations as shown in Fig. 5.

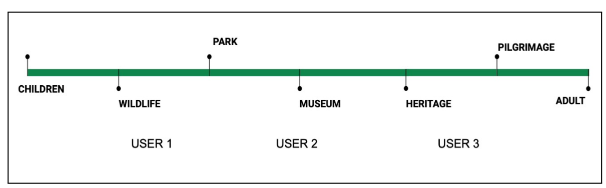

Fig. 5. 1-D Method of Of Categorizing Locations

- **2-D**: When one aspect of a location is insufficient for a recommendation, we integrate extra features, broadening the scope of our method. For example, we may include the range of the budget of the user for travel and tourism. We can map users as indicated in Fig. 6. We can determine a new user's area of interest and hence present him with better suggestions.

In its implementation, a user similarity matrix is created as depicted in Fig. 7, which is why this approach is known

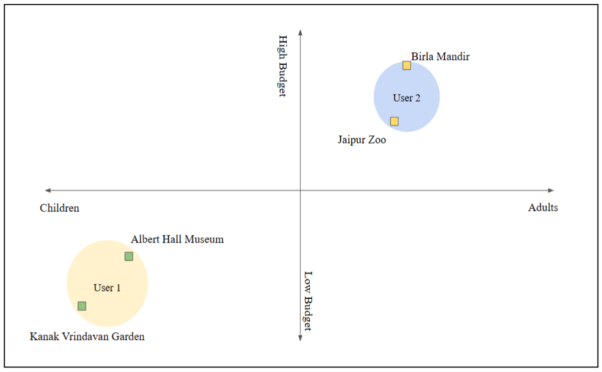

Fig. 6. 2-D Method of Of Categorizing Locations

as user-based filtering. The resemblance might be due to the user's employment, age, or other factors. [19].

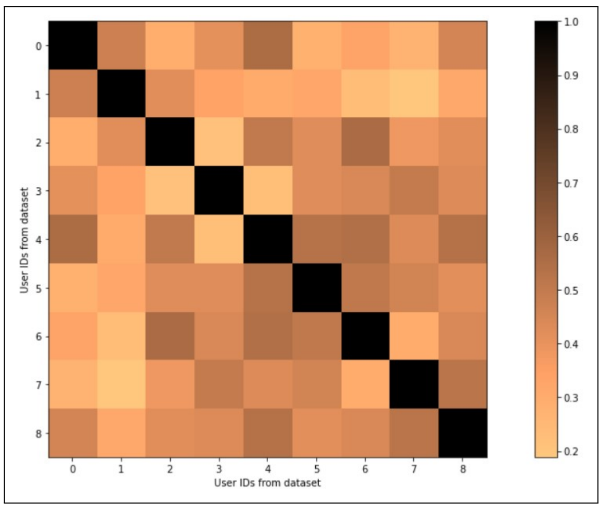

Fig. 7. Image of the similarity matrix among different users

Advantages

- It can recognize the likes and dislikes of users without requiring feedback.
- This model can present users with content that they may like and allow them to explore new content.
- It can be used by newly created systems with no existing data.

Disadvantages

- The prediction of the model for every given user-item combination is the dot product of the corresponding embedding. If an item isn't there during training, the system won't be able to generate a score for it. This is also known as the "cold start difficulty".
- Any characteristics that exist in addition to the query or item ID are referred to as side features. Country or age may be optional options for movie choices. Including usable features boosts the model's quality.

C. *Hybrid recommendation system*

Because one type of approach may not always yield correct results, we combine the two above methods to develop a new method known as the hybrid method. Deep learning is frequently used to improve the dynamic and accuracy of suggestions. [20].

If we understand it in the context of a tourist recommendation system, we will have factors like:

- Target Variable — Ratings can be explicit (i.e., the user leaves feedback) or implicit (i.e., assuming positive feedback if the user visits a location); either way, they are necessary.
- Product features —tags and descriptions of the items (i.e., location category), mostly used in the content-based methods.
- User Profile — Users' descriptive information can be demographic (such as gender and age) or behavioral (such as preferences, average time on screen, and most frequent time of usage).

We need a dataset for training and several data sets for testing because we are utilizing deep learning. Our dataset was separated into two parts: training and testing.

For the deep learning part, there are basically three primary layers: an input layer, a neural network layer, and an output layer. In Fig. 8 a sample layer model is depicted.

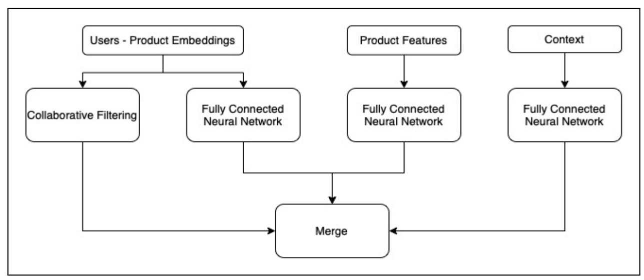

Fig. 8. A sample model for hybrid recommendation system

Advantages

- It covers all the limitations of content filtering and collaborative filtering

Disadvantages

- It has only the disadvantage that it is more difficult to implement than the other two methods.

V. Discussion

We analyzed multiple types of tourism and factors which take part in making recommendations. Moreover, we looked into various similarity detection techniques namely - the TF-IDF technique, which is based on finding creating vectors of the given sentences, and Cosine Similarity, which is basically used for creating a similarity matrix using dot products of the matrix. Then we listed various recommendation methods.

After studying different methods we tested all of them against a common dataset, and we deduced that content-based and collaborative filtering are good but do not fulfill personalized recommendation parameters. The content-based method involves item-item similarity and collaborative filtering involves user-user similarity, so

TABLE III
COMPARISON OF DIFFERENT METHODS ON THE BASIS OF PARAMETERS CONSIDERED

Method	Self Learning	No Dataset Required	User Based		
			Data Specific	Behaviour Specific	Personality Specific
Context-Based			✔		
Collaborative Filter	✔	✔		✔	
Hybrid Filter	✔	✔	✔	✔	

they each miss some parameters. The most appropriate is the hybrid method used with deep learning[21].

VI. CONCLUSION FUTURE WORK

In this paper, the techniques discussed have one or more drawbacks as depicted in Table 3. In order to overcome these drawbacks and integrate the best elements of both approaches, we need to create a hybrid recommendation system with a deep learning model which includes personal data. The fundamental goal of such hybridization is to increase rate prediction accuracy by overcoming the disadvantages of specific approaches when used independently.

In the future, we will introduce semantic analysis for each user along with existing data. Taking into account the preferences of multiple users with different personalities to create an evolved recommendation system.

REFERENCES

[1] A. Thananchana, K. Noinan and S. Wicha, "The Designing of Cultural-Based Tourism Recommendation System with Community Collaboration," 2022 Joint International Conference on Digital Arts, Media, and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT NCON), Chiang Rai, Thailand, 2022, pp. 510-513

[2] L. Febre, P. Valdiviezo-Diaz and R. Reátegui, "Collaborative Filtering Enhanced by Demographic Information for Tourist Sites Recommendations," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), Madrid, Spain, 2022, pp. 1-6

[3] L. Gang, "Personalized Recommendation of Tourist Attractions Based on Collaborative Filtering," 2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA), Xi'an, China, 2020, pp. 144-147

[4] M. L. Bai, R. Pamula and P. K. Jain, "Tourist Recommender System using Hybrid Filtering," 2019 4th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2019, pp. 746-749

[5] M. Li and J. Sasaki, "Recommendation Method for Attractive Destinations for Individual Tourists Using Profile Data," 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), Shanghai, China, 2021, pp. 248-252

[6] M. Kamal and I. Chatzigiannakis, "Influential Factors for Tourist Profiling for Personalized Tourism Recommendation Systems– A Compact Survey," 2021 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 2021, pp. 1-6

[7] N. Chiwong and P. Rojanavasu, "Schedule Travel Planning System for Phayao Information Recommendation," 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT NCON), Pattaya, Thailand, 2020, pp. 118-122

[8] N. Wayan Priscila Yuni Praditya, A. Erna Permanasari and I. Hidayah, "Designing a tourism recommendation system using a hybrid method (Collaborative Filtering and Content-Based Filtering)," 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Purwokerto, Indonesia, 2021, pp. 298-305

[9] O. Alnogaithan, S. Algazlan, A. Aljuraiban and A. A. Shargabi, "Tourism Recommendation System Based on User Reviews," 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhier, Bahrain, 2019, pp. 1-5

[10] P. Padia, K. H. Lim, J. Cha and A. Harwood, "Sentiment-Aware and Personalized Tour Recommendation," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 900-909

[11] Q. Qi, J. Cao, Y. Tan and Q. Xiao, "Cross-Domain Recommendation Method in Tourism," 2018 IEEE International Conference on Progress in Informatics and Computing (PIC), Suzhou, China, 2018, pp. 106-112

[12] R. A. Wahurwagh and P. M. Chouragade, "Personalized POI Travel Recommendation With Multiple Tourist Information," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-3

[13] S. Y. Shu, W. Hussain Wan Ishak and F. M. Yamin, "A Web-based Application for Interesting Place Recommendation," 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 2022, pp. 215-220

[14] T. Wenan, D. Shrestha, D. Shrestha, B. Gaudel and S. R. Jeong, "Analysis and Design of Tourism Recommender System for Religious Destinations of Nepal," 2020 IEEE International Conference on Sustainable Engineering and Creative Computing (ICSECC), Indonesia, 2020, pp. 214-220

[15] V. Garipelly, P. T. Adusumalli and P. Singh, "Travel Recommendation System Using Content and Collaborative Filtering - A Hybrid Approach," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-4

[16] X. Chen, Q. Liu and X. Qiao, "Approaching Another, Tourism Recommender," 2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C), Macau, China, 2020, pp. 556-562

[17] Y. A. Kusumawati and D. G. Renanda, "Architecture Design of Travel Guide Application for Tourism Recovery Post-Pandemic Covid-19," 2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS), Makasar, Indonesia, 2021, pp . 1-7

[18] Y. Cui, C. Huang and Y. Wang, "Research on Personalized Tourist Attraction Recommendation based on Tag and Collaborative Filtering," 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 2019, pp. 4362-4366

[19] Y. Xu, "Research on Travel Personalized Recommendation Information Retrieval Based on Big Data Improvement Algorithm," 2021 5th Annual International Conference on Data Science and Business Analytics (ICDSBA), Changsha, China, 2021, pp. 111-114

[20] Z. Duan, Y. Gao, J. Feng, X. Zhang and J. Wang, "Personalized tourism route recommendation based on user's active interests," 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 2020, pp. 729-734

[21] Z. Jia, Y. Yang, W. Gao and X. Chen, "User-Based Collaborative Filtering for Tourist Attraction Recommendations," 2015 IEEE International Conference on Computational Intelligence Communication Technology, Ghaziabad, India, 2015, pp. 22-25