Tour Spot Recommendation System via Content-Based Filtering

Mishal Muneer
Department of Computer Science and IT
University of Lahore
Lahore, Pakistan
mishal.muneer@cs.uol.edu.pk

Uzair Rasheed
Department of Computer and Software Engineering
College of Electrical and Mechanical Engineering, NUST
Islamabad, Pakistan
urasheed.cse19ceme@ce.ceme.edu.pk

Sadia Khalid
Department of Computer and Software Engineering
College of Electrical and Mechanical Engineering, NUST
Islamabad, Pakistan
skhalid.cse19ceme@ce.ceme.edu.pk

Muhammad Ahmad
Department of Computer Engineering
University of Engineering and Technology
Lahore, Pakistan
Muhamadahmad1996pak@gmail.com

2022 16th International Conference on Open Source Systems and Technologies (ICOSST) | 978-1-6654-6477-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICOSST57195.2022.10016820

***Abstract*— Online recommendation systems have gained prominence. The online recommendation systems provide a better and faster way to choose Tour Spots for traveling and transactions. Recommender systems are powerful emerging technologies that help consumers to find the best places for touring according to their interests. The recommendation system recommends the most appropriate destinations for the end users. Internet travel portals also compete with each other regularly by taking into consideration several characteristics. Recommendation systems are one of the best methods to raise income and attract consumers. Established systems retrieved irrelevant information that resulted in low customer satisfaction and disappointment of customers. In this research, the Tour Spot recommendation system through content-based filtering is proposed and designed. It recommends the best picnic spot according to the budget and interest of the user. The proposed system improves the tourist experience and recommends the best place according to the things they would like; most significantly it is budget friendly as it suggests different places according to the interests and needs of the user. Data collection, including user information, consolidated user engagement reports, and tour spot data, is the primary challenge of implementing a travel recommendation system. User information originates mainly from information entered by the user during the registration process. When the system has the data of user likes and dislikes, then the system compares the user preferences to the dataset that the author makes for the tour recommendation systems in the dataset first identifies all the features that must be considered at the time of selection of travel destination. This Article also considers user likes and dislikes and the previous history of users to recommend a similar tour spot.**

Keywords*— *tour spot, tourism, recommender systems, collaborative filtering, content-based filtering,

I. INTRODUCTION

The aim of data mining is to use various data mining tools and techniques to analyze processes and derive information from data in the form of large datasets. In the information industry, a vast number of data are available; all of these data are of no value until they are turned into valuable or useful information. This helps developers to access data from several different aspects. That is why choosing the information a consumer is interested in often becomes a hard and complicated task. Therefore, in the center of a sea of information, highly interesting information will get lost.

Recommendation has been used in a range of applications such as film, songs, books, news, and products, etc. A recommender System [1] is a particular category of filter for adaptive information. It is a method that aims to provide information to the customers which are important for customer. The term "personalization" is the central axis of the recommender system. The aim of personalization is to give customers what they want without explicitly asking for it. The proposed system will predict what the user requests not only by taking the information that the user provide initially in the form , but also by matching his profile with other accounts with a similar one. A difficult challenge for recommendation systems is to produce recommendations according to user tastes. In the recommendation process, filtering factors play an important role [8]. Collaborative filtering, knowledge-based filtering, content-based filtering, and social filtering are the most widely used filtering methods [2]. Many experiments have already led to the development of different recommendation systems, such as film, tourism, books, online search, e-commerce, e-learning, and music [3].The primary purpose of using personalization methods is to produce personalized user recommendations according to the desire, user preference and user interest [4]. The aim of the recommendation system is to eliminate all the unnecessary information and provide relevant results for a specific user [5]. Recently, in order to promote information search and trip planning activities, recommendation systems have been introduced in the tourism sector to offer customized recommendations to travelers. In the tourism recommendation systems [6], the proposed model learns consumer expectations and produces attractions according to user interests.

Tour spot Recommendation through content based filtering is purposed in this paper. Basically, it recommends the best picnic spot according to the budget and interest of

978-1-6654-6477-2/22/$31.00 ©2022 IEEE

the user. The proposed system will improve the tourist experience and recommends the best place according to the things they would like and most importantly it is budget friendly and according to the interest of the user. Previous recommender systems only recommend the spots based on only user preference for example they only consider one user preference either location or budget etc. but in purposed system authors consider all the preferences and interest of the user and also focus on the things they would like or not for example from where or which country he/she want to visit? Which type of place they would like? Which type of hotels they want to stay? Also consider the safety and transport issue and most importantly the climate of that place is also considered to recommend the best picnic spot to the user.

There are many approaches to recommend the best place but in this system. The content based filtering methods are cited because it best picnic to the user by taking similarity of spots and it also consider the user previous history in order to recommend a similar tour spots. In order to predict his preferences, the content-based recommendation system compares the features of products which have not been rated with the features of previously rated product.

This paper is organized as follows Section II discusses Literature Review in detail. Section III states the proposed framework that is used in this article. Section IV discusses the results of the study. Section V discusses the conclusion and future work of the study.

II. Literature Review

The tourist sector has a big influence on the economy nowadays. And so, with multi-layered expectations and demands, the tourism industry is getting more challenging and complex [7]. It is versatile, often encountered while traveling, with both perfection and variety in high demand. Consequently, there should be some high-quality alternatives to the tourist offers. In addition, for the services delivered, consumers need to improve agile service improvements with shorter lifetimes [8].

In this diverse environment, information communication technology abbreviated as ICTs are launching the tourism industry and taking to our world a whole modern phenomenon of 'smart tourism [9][10]. The Smart tourism services refer to the use of ICTs to establish novel techniques in the field of tourism. It encourages coordinated destination initiatives to recognize new ways of gathering and using data from the tourism and social connectivity industries and consumers in connection with advanced technology to enhance quality, stability and experience [11]. Internet of Things, cloud computing, mobile communication [12], and artificial intelligence are the ICT tools used by smart tourism [13][14].

The key goal of smartly tourism is to enhance the efficiency of the service [15]. A digital agency or web-based agency [1] is one of the utilities established in recent years. The main objective of these companies is to propose appropriate bookings (ticket, accommodation and tours package), while the minor objective is to establish an appropriate place for customers to interact and exchange information with the tourism industry [16].

In the e-commerce sector, recommendation systems have arisen and are being developed to improve [17]. Recommender systems is typically divided into four types: content-based filtering (CB), knowledge-based filtering (KBF), Collaborative filtering (CF) and hybrid filtering [18][19].

CF is a filtering method that incorporates a methodology based on information filtering methods on the prior judgments or history of previous sales by the consumer. This strategy, then, depends heavily on the knowledge of the other consumers. This methodology would not propose any recommendations without this knowledge, which is the main limitation of a CF-based system [20]. Two significant obstacles must be overcome by CF: the question of sparsity and the scalability issue [21].

Unlike Collaborative filtering, content-based filtering analyses data sets evaluated by every client and it takes into account both the ratings as well as the quality of those data, to establish a customer profile that is used to suggest additional products of interest [22].

Content-based recommendations recommend user things that are close to their previous experiences and preferences [23]. The suggestions or preferences of other parties are not involved. This is a benefit of serving the individual who has particular preferences and does not require similar consumer input. Recommendation structures focused on content begin with the recommendation method by collecting knowledge about item characteristics and the profile of the customer. Such knowledge is used to make sure that calculate the resemblance of the features of an item to the profile of the customer. The software then selects Products that are most close to the needs of the consumer [24]. Information retrieval or attribute-based filtering algorithms are used in the method's filtering. [25].

There are two methods of collecting profile information: through unique channels, such as a form that Users complete the service as they sign up, or from implicit channels, such as the user's The history of browsing when they interact with the recommendation systems [26][27].The feature weighting approach is used by content-based recommendation systems for filtering out Objects that have characteristics matching the specified preferences of the users and their previous profile history [28]. Term Frequency Inverse Document Frequency (TF-IDF) is the method that is most frequently used in the weighting method of features [27]. TF-IDF aims to see if the profile of the users has the similar variables (terms) as the keywords in the definition of the item [28].

According to Chen Jian the solution to the cold-start problem [29] is an automatic content-based recommendation using the clustering technique. Although that user has no prior history or background, a content-based recommendation system will make suggestions to them by clustering.

Users are recommended by knowledge-based recommendation systems according to both the choice of the consumer and the specifications of the item needed [23]. Aids of navigation are used to get information about the needs of the customer, which is not saved for further use. It is important to evaluate the attributes of products currently desired by the consumer and the product descriptions to make correct recommendations for a knowledge-based recommendation framework [30].

Every time user interacts with the system, knowledge-based filtering directly gathers data from the System. Data

collected shall be store in the structure of sentences, laws or ontologies [25] using the data collected either of the two approaches: case-based and rule-based reasoning [31].

Both of the recommended systems we have talked about have their limitations and strengths, but it is reasonable to attempt to increase their revenues by hybridizing two or three recommended strategies to reduce their weaknesses. Hybrid systems of content-based and collaboration-based systems save user interests and merge them with the greatest In order to make the recommendations, related things. There are also hybrid arrangements between collaborative and social networks; content-based and knowledge-based systems; and also between knowledge-based and collaboration-based systems.

Prem Melville and Vikas Sindhwani have purposed a methodology on "Recommender System" in they describe several recommendations' methods and strategies. They also tried to explore the frequent problems and restrictions in the system of recommendations [32].

Many applications in the area for recommendation systems have been published within the tourism domain literature. Authors can cite iTravel, recommendation system in the mobile peer-to-peer environment [33], consensus-driven community recommendation system [34], tourism tour recommendation [35], social-based tourism recommendation system [36] or e-Government tourism recommendation system [37] with a focus only on the most significant and recent works, citing only some of the most interesting works on the issue. A recent study on tourism recommendation systems [18] highlights the growing value of mobile apps for advanced recommenders, while pre-visit approaches do typically include web-based models. Latest studies dedicated to mobile recommendation systems [38] also confirm this claim. Referring to the work on tourism recommendation systems [18], the authors also stress that tourism recommendation systems incorporate videos, images and digital charts. The best conclusion in the work, though, may be one that underlines that the advice process is a critical feature of tourism advisory systems.

Personalized recommendation, taking into account unique user profiles or characteristics by using a probabilistic Bayesian learning approach (e.g. gender, age, race) [39][40]. Based on the GPS history info, the users' comments were inserted into the travel recommendation system. Along with potential things that can be undertaken there, the system can explore interesting sites. Taking into account the location correlation [40] a personalized location recommendation framework focused on user-generated GPS trajectories was performed, taking into consideration the travel skills and challenges of a user.

Majid et al. [41] considering the temporal and weather sense of tourism, a context-aware customized travel suggestion framework. By building the travel similarity between users from their travel history, the device learned unique travel interests from users. The personalized travel destinations in unvisited regions were subsequently suggested to visitors.

III. PROPOSED FRAMEWORK

In the proposed approach researchers used the content-based filtering method to recommend the best picnic spot to the user. It recommends the best picnic spot to a user by taking the similarity of spots. It also considers the user previous history in order to recommend the similar spot. In purposed system, researchers used python to implement the recommendation system through content-based filtering. Figure 1 show the overview of the proposed Firstly researcher make the dataset of all the tour spot .data set contain all the features that are mentioned in the table 1 .after that some pre-processing is done on the text and convert that text into vectors using TF-IDF. Then after finding the similarity recommend the best picnic spot to the user.

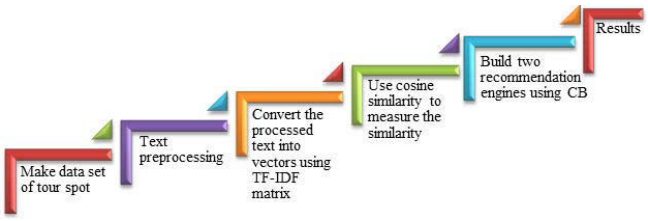

Figure 1. Overview of methodology

A. Dataset

To recommend the best picnic spot researcher make the data set of tour spot. And identify all the best picnic spot from different sources and identify the different feature through literature and used in tour spot data set of tourism that must be consider at the time of selection of travel destination. Table 1 shows the features and their values that are used in tour spot recommendation system data set.

TABLE I. OVERVIEW OF FEATURES AND THEIR VALUE

Feature Name	Values	Description
Tour Spot	Names of best picnic spots	This column have the names of best picnic spot that recommend to the user according to their preferences
Country	Country Names	This column have names of different countries, having tour spots
Region	South Asia, South America, Africa, North America, Europe, Antarctica, and Australia.	This column have the names of different regions
Geography	Mountains, Lakes, plain, Rivers, Valleys, coastal belt, desert	This column have the names of different types of place
Budget	Cheap, moderate, expensive	In this the cost of that place is mention.
Hotels	Star Rating	Hotels available near at Tour Spot
Safety	Yes/No	Is that tour spot safe or not?
Transport	Yes/No	Is transport available at that spot
Climate	Koopen Classification method	Climate of that spot also consider
Description and Attractions	Simple words, and popular places	Describe the place in words and tell which type of activity tourist can perform at that place and what are the other attractions near at that place

B. *Text Preprocessing*

The following are the steps for text pre-processing:

- Firstly remove all the stop words in the data. Stop words are words in English that do not bring any meaning to a sentence. Without losing the context of the phrase, they can safely be ignored. For example the vocabulary likes, for instance, the, he, have etc.
- Remove all the NON-ASCII characters. Python 2 uses ASCII as the default encoding for source files, which means that to use NON-ASCII Unicode characters in literals, need to choose another encoding at the top of the file. For source files, Python 3 uses utf-8 as the default encoding, so this is less of a concern.
- Convert all the text into lower case letter.
- Remove punctuations and HTML tags from the data.

C. *TF-IDF*

- Convert the processed text into vectors using TF-IDF matrix for better recommendations. For its content-based recommendation scheme, this filtering technique employs the feature weighting approach known as Word Frequency-Inverse Document Frequency (TF-IDF). This method searches the data collection for things whose descriptions contain a keyword(s) that correspond to the user's profile attributes. In the tour information database, the TF-IDF content-based recommender system can compare user preferences with keyword definitions and recommend those that match them.
- And find out the word frequency.

D. *Similarity Measures*

The goal is to find a similar tour spot to a given spot o feature of that spot and recommend those similar tour spots to the user. Now the question is how to find whether the given spot is similar or dissimilar? For this a similarity measure is used to find out the similar spot. The similarity measure quantifies how similar two data objects are. A distance whose dimensions correspond to the properties of the objects is a measure of similarity in a data mining or machine learning context. The similarities between the features are highest when the distance is small. There are different similarity measures such as Euclidean distance, Manhattan distance, Minkowski distance, Cosine Similarity and Jaccard Similarity. The proposed recommendation system uses the cosine similarity measure because this metric helps us to measure how similar the documents are irrespective of their size.

E. *Recommendation Engine*

- authors build two recommendation engines one is using tour spot as input because if a user visits any spot in past ,then it takes the user previous history according to the spot he/she visit in past then recommend the similar spot to the user immediately because system know the user preferences already from the user previous history.
- And the other recommendation engine is if system have new user and user do not visit any place in the past then the system ask the user to fill the form in which user tell about their likes and dislikes and then from that form system can get the user preferences and can match that preferences in data set.

IV. RESULTS

Following are the results of two purposed recommendation systems. Both recommendations systems recommend the best picnic spot to the user according to their interest by considering all their preferences. Figure 2 shows the result for the existing users who like one spot in past then it recommend the top 5 tour spot according to their previous history. Figure 3 shows the result if the system have a new user and the system does not have the information related to the user then it take all the user preferences and recommend the Top 5 places to the user.

```
recommend("Rawal Lake Islamabad")
```

	Tour Spot
1	Saif-ul-Muluk Lake Naran
2	Upper Kachura Lake Skardu
3	Upper Kachura Lake Kaghan
4	Hawks Bay Karachi
5	Satpara Lake Skardu

Figure 2. Results using user previous history

```
recommend("Natural beauty, horse riding,boating,
```

	Tour Spot
1	Saif-ul-Muluk Lake Naran
4	Hawks Bay Karachi
30	Lake View Park islmabad
5	Satpara Lake Skardu
12	Ratti Gali Lake Azad Kashmir

Figure 3. Results for new user

V. CONCLUSION AND FUTURE WORK

Tour spot Recommendation through content-based filtering is implemented in this paper. It recommends the best picnic spot according to the budget and interest of the user. This system improves the tourist experience and recommends the best place according to the things they would like and most importantly it is budget friendly and according to the interest of the user. For this Authors make the dataset of different tour spots that are in different countries and different regions. And also find out the features that help in the recommendation of tour spots according to user preferences. Authors build two recommendation engines that give the best results and recommend the top 5 places according to their interest to the user.

In the future, system is improved by development of ontology of this framework for more excellent results and recommendations.

REFERENCES

[1] Schafer, J.B., J.A. Konstan, and J. Riedl, E-commerce recommendation applications. Data mining and knowledge discovery, 2001. 5(1): p. 115-153.

[2] Thorat, P.B., R. Goudar, and S. Barve, Survey on collaborative filtering, content-based filtering and hybrid recommendation system. International Journal of Computer Applications, 2015. 110(4): p. 31-36.

[3] Lu, J., et al., Recommender system application developments: a survey. Decision Support Systems, 2015. 74: p. 12-32.

[4] Rahman, M.M. and N.A. Abdullah, A personalized group-based recommendation approach for Web search in E-learning. IEEE Access, 2018. 6: p. 34166-34178.

[5] Sahoo, A.K., et al., Intelligence-based health recommendation system using big data analytics, in Big data analytics for intelligent healthcare management. 2019, Elsevier. p. 227-246.

[6] Kesorn, K., W. Juraphanthong, and A. Salaiwarakul, Personalized attraction recommendation system for tourists through check-in data. IEEE Access, 2017. 5: p. 26703-26721.

[7] Aldebert, B., R.J. Dang, and C. Longhi, Innovation in the tourism industry: The case of Tourism@. Tourism management, 2011. 32(5): p. 1204-1213.

[8] Hassannia, R., The multipurpose application of radio frequency identification (RFID) in the tourism industry: On a requirement analysis for employing RFID technology in the hotel sector. 2014, Eastern Mediterranean University (EMU)-Doğu Akdeniz Üniversitesi (DAÜ).

[9] Gretzel, U., et al., Smart tourism: foundations and developments. Electronic Markets, 2015. 25(3): p. 179-188.

[10] Jovicic, D.Z., From the traditional understanding of tourism destination to the smart tourism destination. Current Issues in Tourism, 2019. 22(3): p. 276-282.

[11] Gretzel, U., L. Zhong, and C. Koo, Application of smart tourism to cities. International Journal of Tourism Cities, 2016.

[12] Palos-Sanchez, P., et al., Users acceptance of location-based marketing apps in tourism sector: an exploratory analysis. Journal of Spatial and Organizational Dynamics, 2018. 6(3): p. 258-270.

[13] Amato, F., et al., Exploiting cloud technologies and context information for recommending touristic paths, in Intelligent Distributed Computing VII. 2014, Springer. p. 281-287.

[14] Reyes-Menendez, A., J.R. Saura, and C. Alvarez-Alonso, Understanding# WorldEnvironmentDay user opinions in Twitter: A topic-based sentiment analysis approach. International journal of environmental research and public health, 2018. 15(11): p. 2537.

[15] Neuhofer, B., D. Buhalis, and A. Ladkin, A typology of technology ‐ enhanced tourism experiences. International journal of tourism research, 2014. 16(4): p. 340-350.

[16] Gavalas, D., et al., A survey on algorithmic approaches for solving tourist trip design problems. Journal of Heuristics, 2014. 20(3): p. 291-328.

[17] Park, D.H., et al., A literature review and classification of recommender systems research. Expert systems with applications, 2012. 39(11): p. 10059-10072.

[18] Borràs, J., A. Moreno, and A. Valls, Intelligent tourism recommender systems: A survey. Expert systems with applications, 2014. 41(16): p. 7370-7389.

[19] Hassannia, R., et al., Web-based recommendation system for smart tourism: Multiagent technology. Sustainability, 2019. 11(2): p. 323.

[20] Herlocker, J.L., et al., Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS), 2004. 22(1): p. 5-53.

[21] Su, X. and T.M. Khoshgoftaar, A survey of collaborative filtering techniques. Advances in artificial intelligence, 2009. 2009.

[22] Lops, P., M. De Gemmis, and G. Semeraro, Content-based recommender systems: State of the art and trends. Recommender systems handbook, 2011: p. 73-105.

[23] Hinze, A. and S. Junmanee, Travel recommendations in a mobile tourist information system. Information systems technology and its applications, ISTA'2005, 2005.

[24] Barranco, M.J. and L. Martínez. A method for weighting multi-valued features in content-based filtering. in International conference on industrial, engineering and other applications of applied intelligent systems. 2010: Springer.

[25] Berka, T. and M. Plößnig. Designing recommender systems for tourism. 2004: ENTER 2004: 11th International Conference on Information Technology in ….

[26] Biuk-Aghai, R.P., S. Fong, and Y.-W. Si. Design of a recommender system for mobile tourism multimedia selection. in 2008 2nd International Conference on Internet Multimedia Services Architecture and Applications. 2008: IEEE.

[27] Brusilovski, P., A. Kobsa, and W. Nejdl, The adaptive web: methods and strategies of web personalization. Vol. 4321. 2007: Springer Science & Business Media.

[28] Zhu, X., H. Ye, and S. Gong. A personalized recommendation system combining case-based reasoning and user-based collaborative filtering. in 2009 chinese control and decision conference. 2009: IEEE.

[29] Jian, C., Y. Jian, and H. Jin. Automatic content-based recommendation in e-commerce. in 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service. 2005: IEEE.

[30] Li, X. and T. Murata. Customizing knowledge-based recommender system by tracking analysis of user behavior. in 2010 IEEE 17Th International Conference on Industrial Engineering and Engineering Management. 2010: IEEE.

[31] Jadhav, A. and R. Sonar. An integrated rule-based and case-based reasoning approach for selection of the software packages. in International Conference on Information Systems, Technology and Management. 2009: Springer.

[32] Mathew, P., B. Kuriakose, and V. Hegde. Book

Recommendation System through content based and collaborative filtering method. in 2016 International conference on data mining and advanced computing (SAPIENCE). 2016: IEEE.

[33] Yang, W.-S. and S.-Y. Hwang, iTravel: A recommender system in mobile peer-to-peer environment. Journal of Systems and Software, 2013. 86(1): p. 12-20.

[34] Castro, J., et al., A consensus ‐ driven group recommender system. International Journal of Intelligent Systems, 2015. 30(8): p. 887-906.

[35] García-Magariño, I., ABSTUR: an agent-based simulator for tourist urban routes. Expert systems with applications, 2015. 42(12): p. 5287-5302.

[36] García-Crespo, A., et al., SPETA: Social pervasive e-Tourism advisor. Telematics and informatics, 2009. 26(3): p. 306-315.

[37] Al-Hassan, M., H. Lu, and J. Lu, A semantic enhanced hybrid recommendation approach: A case study of e-Government tourism service recommendation system. Decision Support Systems, 2015. 72: p. 97-109.

[38] Gavalas, D., et al., Mobile recommender systems in tourism. Journal of network and computer applications, 2014. 39: p. 319-333.

[39] Cheng, A.-J., et al. Personalized travel recommendation by mining people attributes from community-contributed photos. in Proceedings of the 19th ACM international conference on Multimedia. 2011.

[40] Zheng, Y. and X. Xie. Learning location correlation from gps trajectories. in 2010 Eleventh International Conference on Mobile Data Management. 2010: IEEE.

[41] Majid, A., et al., A context-aware personalized travel recommendation system based on geotagged social media data mining. International Journal of Geographical Information Science, 2013. 27(4): p. 662-684.