Deep Learning based Tourism recommendation system

Ismat Fathima[1a*]
Bonthu Kotaiah[2]

[1a*] *Research Scholar,*
Department of Computer Science & Information Technology,
Maulana Azad National Urdu University, Hyderabad. India.
[2a]*Assistant Professor,*
Department of Computer Science & Information Technology,
Maulana Azad National Urdu University, Hyderabad. India

[1 a)*] *ismat@manuu.edu.in*
[2a)]*kotaiah.bonthu@manuu.edu.in*

Abstract **-** Many different fields make use of intelligent computer-based recommendation systems (RS). Social media sites produce a lot of posts, likes, and other interactions because of their massive user bases, and these interactions operate as a database for various recommendation algorithms. This study's core area of interest is the tourism sector, and RS is a crucial tool for trip planning for travelers. The research in this broad area has been studied extensively, and this paper provides a current overview of that research, taking into account the various interface types, the range of recommender algorithms, the features that such models provide, and their application of artificial intelligence techniques. This survey also provides some insights on the development of the industry's most promising research areas for the following years.

Keywords: *Social media, Points of Interest (POI), Recommender System (RS), Deep Learning*

1. Introduction

Traveling is an experience in itself a combination of the journey, the destination, time spent traveling, the stay, the season, and the activities undertaken. Tourists typically seek assistance while organizing a tour to make the most of their travel outings. Modern travelers want to make their judgments about their trip, select specific options, complete bookings, and make direct payments to avoid using intermediaries. For this, they turn to the internet. However, there is an abundance of information on the web about many tourism-related topics, making it time-consuming to find a suitable trip package, product, or service. Some tourists plan their trip, right from deciding which country to visit and the activities/events they will be participating in, while others may be more spontaneous in first reaching a destination and then planning their itinerary as the trip progresses. In both cases, the user is flooded with a plethora of options and may not be sure what to pick, which is where an automated recommender system (RS) comes into the picture.

A recommender system can help with a variety of tour-related questions, such as the best vacation spots in a particular season, the best weather to visit a place of interest (POI), the quickest mode of transportation, or advice on where to find good spots for photography enthusiasts. For example, Umanets et al developed a service that could be run on mobile phones and the web, targeted at finding previously unseen points of

interest [2]. This would help returning tourists to visit places they might have missed before. To improve the effectiveness of the RS, researchers have employed various techniques. The primary sources of user data are blogs, geo-tags, check-in information, GPS records, and so forth. Every method used to improve the effectiveness of the travel recommendation systems has some benefits and drawbacks. In the next section, related work is discussed. The most popular implementation strategies in the field of travel recommendation to date are covered in Section 3. The last section concludes with future scope of travel RS.

2. Related Work

Previous publications that are briefly covered in this section illustrate how recommendation systems are applied in the tourism industry. Many systems employ machine learning, while some apply Deep Neural Networks for their working.

Lucas et al in their study, apply classification based on association in a recommender system for tourism, implementing a hybrid recommendation technique. It is named Personalized Sightseeing Planning System [8]. A. Umanets and others developed an application called GuideMe and its primary features include its ability to integrate with social networks. This prototype works on both Android and iOS in helping suggest unexplored tourist destinations based on user ratings [2]. Employing the Amazon Reviews dataset,. Kulkarni et al rank tourist places based on the positive and negative reviews about them. They use a deep learning algorithm in arranging the sequence of POIs [1]. Proposing an RS using Social Network analysis for the city of Seoul in South Korea, Jeong et al argue that the personality type of the tourist plays a major role in selecting a tourism destination [3].

Wang processed users' information like demographic variables viz age, gender, profession or occupation and city, and review data to provide a system to recommend personalized travel products. The researcher used a total of 1,283,715 reviews for his study [13]. G and H. Verma et al focussed on rural tourism in India using opinion mining. They employed a supervised machine learning method to categorize the sentiments from a variety of travel-related companies and websites, reviews of hotels, and tourism agencies left by foreign travelers. "Term Frequency - Inverse Document Frequency (TF-IDF)" metrics are calculated to find the relevance of the words identified for sentiment analysis, and a robust model is proposed [5]. Offering a combined Lexicon-based and rule-based NB sentiment analysis approach to extract tourist characteristics from mobile apps, Muthukrishnan and others in their paper collect different reviews on Twitter connected to travel and processes them to identify the polarity and categorize each into different sentiments [6].

In their paper, Zelenka et al. present a conceptual framework for offering validated destination assessments and verified ratings of tourism services. Tripadvisor and Booking.com are the two case studies used in the main research. A trust model is created after conducting a SWOT analysis of the review and verification processes of the aforementioned websites [7]. To define the geographical, temporal, and demographic tourist flows throughout a tourist region, the research by Paolanti and others presents and evaluates a deep learning geodata framework. The dataset included

Tweets in both Italian and English, and four DNNs were used to determine the sentiment of the Tweets. By manually estimating with the help of human annotators in addition to automatically judging by the surrounding sentences or hashtags, the authors generated a multilingual dataset that was more accurate than the ones already available. The paper employs data visualization techniques such as statistical graphics, charts, and information graphics [12]. The study by Li and another analyzes various well-known tourist destinations in China using sensor detection technologies, taking into account factors like tourist type, gender, and location. It also combines the survey findings of tourists' satisfaction with comprehension with online comments on Ctrip [20].

Fig. 1 Word cloud of the titles of the selected papers

Santamaria-Granados and others suggest a recommender system based on physiological information gathered from wearable devices worn by travelers. The engine also creates a list of the best POIs based on the user's profile, emotional state, and portfolio of travel experiences [10]. The creation of an RS based on big data technologies, AI, and operational research to promote tourism in the Daraa-Tafilalet region of Morocco is covered in the work by Farani et al. The RS acts as a travel planner, creating a thorough itinerary that takes into account a variety of tourism resources for a set amount of time [9]. Based on a deep learning analysis of social media, the study by Mishra and others offers a tactical strategy for surviving the Covid-19 pandemic. It examines the hospitality and healthcare industries as two tourism sub-domains [14].

A unique bagging-based multivariate ensemble deep learning strategy that combines stacked auto encoders and kernel-based intense learning machines is the B-SAKE model, which Sun et al. propose. This model increases data volume and improves feature extraction efficiency to address the over-fitting issue and improve forecasting accuracy. [17]. The work by Yuan creates a Hadoop Distributed File System (HDFS) based on the Map-Reduce concept and presents an enhanced support vector machine (SVM) method based on travel consumer sentiment analysis. [22]. Zhang et al in their paper employ text mining, sentiment analysis, and social media analytics and explore original reviews about 160 Chinese cultural theme parks. Here, 9 cultural aspects are used and sentiment classification is done on this basis [18].

The paper as part of the CRUMBS project, classifies their CRUMBS recommender system as “a Web-based, pull-based, user constraint-based TRS” and “Mobile Dynamic Social Recommender System”. The recommendation mechanism depends on data about tourist destinations and the categories to which they belong, demographic data, explicit tourist preferences regarding POI categories, tags about POIs created by visitors and POI ratings. This is the information used as input by the system's suggested algorithms for user modeling and recommendation [4]. Farokhi et al propose a recommender

using multi-criteria using a collaborative filtering technique, which will output user-based and item-based recommendations. They used offline as well as online training of data [11]. Khomash et al use big data analytics to extract the opinion of tourists by the reviews they leave on Google Maps. Focusing on 8 tourist destinations in Indonesia, they concluded that there is a high correlation between sentiment and reality, and that user reviews help a great deal for the tourist to decide about traveling to a point of interest [16].

Baharmian and others propose an ontology-based tourism recommendation system that employs the spreading activation model. Their technique uses three steps to address the issue of information overload - "ontology-based content analyzer, ontology-based profile learner, and ontology-based filtering component". Considering 1573 POIs in Tehran, Iran, they limited the recommendations to only those which are at a predefined distance from the user [21]. Bhaidkar et al created a prototype in which they used the triangulation method with GPS to track the location of the user and recommend "nearest as well as categorized tourist visiting places with rating". This will give the tourist a list of recommendations to choose from on a user-friendly mobile app [19].

3. Methods Used

There are many factors which tourists take into consideration before planning a trip - the popularity of the destination, user's interest in the place, traveling time to reach the place, the mode of transport and the traffic, and finally the cost involved in the entire trip. The popularity of a recommender system depends on the optimisation techniques used by it to present the best results to the users. This section will cover the popular strategies implemented and the algorithms used by researchers for implementing tourism recommender systems.

The recommendation framework by Lucas et al [8] consists of using "CBA-Fuzzy", an algorithm developed specifically for this purpose. This algorithm "mines the fuzzy association rules that constitute the associative classification model used for recommendation". Their method constitutes two stages - The method's initial phase entails creating user groups with comparable interests and traits. The active user is categorized into one or more of these groups in later phases so that recommendations can be made to him based on his profile. A clustering algorithm is used to create the user groupings using attributes holding demographic information about users (such as age, postal code, and degree of education), as well as attributes of goods to be recommended, which users have rated or purchased. The Personalized Sightseeing Planning System (PSiS) is used which was designed to help tourists in Oporto, Portugal, obtain a customized tour itinerary. The approach starts by using data from the 241 sites of interest that are already in the PSiS, the dataset of which was provided by the user community to assess the procedure.

The main focus of the mobile application developed by Umanets and others [2] is "the users' current location, which is represented by the geographic coordinates obtained using the positioning service available on the user's device, provided by GPS or Wi-Fi connections". The app provides local points of interest and their respective distances based on the users' current location. It enables users to

examine points of interest without showing the user's distance from the nearby tourist attraction when positioning or location services are unavailable.

206,387 users, 7,300 scenic locations, and 1,283,715 reviews were included in the dataset for the model created by Wang [13]. Data was acquired using the distributed web crawler system made possible by Scrapy+redis+MongoDB, which collects network data and converts it into JSON format. A few of the user's details include their gender, age, occupation, city, previous review items, and review information. The user's behaviour features are retrieved in order to analyse user interests and build the feature model. The information regarding the service includes the name, address,, and review details of the tourism service products. Data processing was done using word embedding technology. The review text data of users and objects were processed using Chinese word segmentation technology. Abeysinghe et al present in this paper that the social media analytics platform for the tourism industry can be utilized to provide economic value and support decision-making. Their suggested system has the capacity to recognise conversation threads and topics covered in microblogs, as well as to examine sentiments and emotions that go along with them. Only text data is gathered and used in the insight generating process for the system's current iteration [15].

Granados and others use Emotion Recognition (ER) for their prototype. Tourists use a wearable smart phone/band that gathers information about their emotions and registers it on the band app. It extracts the emotion from the dataset, and a hybrid Deep Learning (DL) approach from “CNN and Long Short-Term Memory (LSTM) networks” are implemented. After the emotion is identified, the “Tourist Experience Recommendation System” is set to work. To obtain the contextual data, an interface is employed using the “Tourist Traceability Ontology (OntoTouTra)”. This recommender system works with TE and MongoDB datasets, and following the analysis, a list of travel suggestions is presented [10]

4. Conclusion

All the systems reviewed are based on enough data collected for the model suggested or prototype created. However, most systems are specific to a particular place, generally to a city. The interfaces used are either web-based or mobile device based. Some are built both for Android and iOS. Most people looking for a flawless itinerary generally use web-based recommender systems because of the ease of using them, without having to download or install extra apps. The common factor in all these systems is the ‘intelligence’ used to return suggestions to the users. Research papers and publications from the last 8 years were analyzed to conclude that recommender systems in tourism have increased in leaps and bounds for the modern tourist. The paces at which these systems are enhancing with each passing season say that many people are turning to AI for their trip planning. Though most authors focus on a particular tourist destination with multiple points of interest for the tourist, these studies can be generalized for wider audiences and other tourist places. Tourists are looking for more personalized recommendations for their travel, and would want a comprehensive service for all their travel needs, which would include the best season to visit a place, recent change in visiting timings of any POI,

current traffic updates, and the local conditions

References

1. A. Kulkarni, M.R., P. Barve, and A. Phade, "A Machine Learning Approach to Building a Tourism Recommendation System using Sentiment Analysis". *International Journal of Computer Applications*, vol 178. pp 48-51, 2019. 10.5120/ijca2019919031.
2. A. Umanets, A. Ferreira, and N. Leite, "GuideMe – A Tourist Guide with a Recommender System and Social Interaction", *Procedia Technology*, vol. 17, pp 407-414, 2014. doi:10.1016/j.protcy.2014.10.248
3. C.-S. Jeong, K.-H. Ryu, J.-Y. Lee, and K.-D. Jung, "Deep Learning-based Tourism Recommendation System using Social Network Analysis," International Journal of Internet, Broadcasting and Communication , vol. 12, no. 2, pp. 113–119, May 2020.
4. CRUMBS Project, Ministry of Industry, Energy and Tourism, Spain
5. G. Verma, and H. Verma, "Opinion Mining on Rural Tourism In India-Qualitative Perspective," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 7, Apr. 2021, pp. 717–722.
6. H. Muthukrishnan, C. P. T. Selvi, M. Deivakani, V. Subashini, N. J. Savitha, and S. G. Kumar, "Aspect-Based Sentiment Analysis for Tourist Reviews," A*nnals of R.S.C.B*, vol. 25, no. 3, Mar. 2021, pp. 5183–5194.
7. J. Zelenka, T. Azubuike, and M. Pásková, "Trust Model for Online Reviews of Tourism Services and Evaluation of Destinations," *Administrative Sciences*, vol. 11, no. 2, p. 34, Mar. 2021, doi: 10.3390/admsci11020034.
8. J.P. Lucas., B.E.S. Coelho, M.N.M García, A.M.A Figueiredo, C.L Martins, "Applying Recommender Methodologies in Tourism Sector", In *Highlights in Practical Applications of Agents and Multiagent Systems. Advances in Intelligent and Soft Computing*, vol 89. Springer, Berlin, Heidelberg, 2011. https://doi.org/10.1007/978-3-642-19917-2_13
9. K.A. Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, "Hybrid Recommender System for Tourism Based on Big Data and AI: A Conceptual Framework", *IEEE, Big Data Mining and Analytics*, vol 4, no. 1, Mar 2021.
10. L. Santamaria-Granados, J.F. Mendoza-Moreno, A.Chantre-Astaiza, M. Munoz-Organero and G. Ramirez-Gonzalez, "Tourist Experiences Recommender System Based on Emotion Recognition with Wearable Data", *Sensors*, vol. 2, Nov. 2021, doi.org/10.3390/s21237854
11. M. Farokhi, M. Vahid, M. Nilashi, and O. b. Ibrahim, "A Multi-Criteria Recommender System for Tourism Using Fuzzy Approach", *Journal of Soft Computing and Decision Support Systems*, May 2016, pp 19-29.
12. M. Paolanti, A. Mancini, E. Frontoni, A. Felicetti, L. Marinelli, E. Marcheggiani, and R. Pierdicca, "Tourism destination management using sentiment analysis and geo-location information: a deep learning approach," *Information Technology & Tourism*, vol. 23, no. 2, pp. 241–264. Feb. 2021, doi:10.1007/s40558-021-00196-4ed
13. M. Wang, "Applying Internet information technology combined

with deep learning to tourism collaborative recommendation system", *PLoS ONE, vol* 15, no. 12:e0240656, Dec 2020. https://doi. org/10.1371/journal.pone.0240656
14. R.K. Mishra, S.Urolagin, J.A. Jothi, A.S. Neogi, and N. Nawaz, "Deep Learning-based Sentiment Analysis andTopic Modeling on Tourism During Covid-19 Pandemic", *Front. Comput. Sci*. 3:775368. doi:10.3389/fcomp.2021.775368
15. S. Abeysinghe, I. Manchanayake, C. Samarajeewa, P. Rathnayaka, M. Walpola, R. Nawaratne, T. Bandaragoda, D. Alahakoon, "Enhancing Decision Making Capacity in Tourism Domain Using Social Media Analytics" In Social Media Sentiment Analysis - UOM CSE, 2018.
16. S. Khomsah, R. D. Ramadhani and S. Wijayanto, "Big Data Analytics to Analyze Sentiment, Emotions, and Perceptions of Travelers (Case Study: Tourism Destination in Purwokerto Indonesia)", *Jurnal E-KOMTEK (Elektro-Komputer-Teknik),* Vol. 5, No. 2 (2021) pp. 284-297 doi.org/10.37339/e-komtek.v5i2.791
17. S. Sun, Yanzhao Li, J. Guo, and S. Wang. "Tourism Demand Forecasting: An Ensemble Deep Learning Approach." *Tourism Economics*, Jul 2021. https://doi.org/10.1177/13548166211025160.
18. T. Zhang, B. Li, and N. Hua, "Chinese cultural theme parks: text mining and sentiment analysis", *Journal of Tourism and Cultural Change*, pp. 37-57, Jan. 2021 doi: 10.1080/14766825.2021.1876077
19. Y. Bhaidkar, P. Bhagwat, P. Bhalere, R. Gujar, S. Walunj. "Tourist Place Recommendation System."*International Journal of Advance Research, Ideas and Innovations in Technology,* vol 1, no. 5, 2015.
20. Y. Li, and H. Gan, "Tourism Information Data Processing Method Based on Multi-Source Data Fusion," J*ournal of Sensors*, vol. 2021, Article ID 7047119 pp. 1–12, Jul. 2021, doi: 10.1155/2021/7047119.
21. Z. Bahramian, and R.A. Abbaspour, "An Ontology-Based Tourism Recommender System Based On Spreading Activation Model", in SMPR, Iran, 2015.
22. Z. Yuan, "Big Data Recommendation Research Based on Travel Consumer Sentiment Analysis", *Front Psychol*, Feb 2022. doi: 10.3389/fpsyg.2022.857292.