

ROLE OF MACHINE LEARNING IN THE TOURISM SECTOR

¹Nazeefa Kazi,

²Mandar Joshi

¹Student, ²Professor

¹Department Of Information Technology

¹Finolex Academy of Management and Technology, Ratnagiri, Maharashtra
, India

Abstract: In today's world, travel and tourism are greatly influenced by technology. Machine learning and artificial intelligence are successfully applied to this area. World's largest travel companies invest a fascinating amount over recommendation systems that in return help them to attract customers and make their services more user-friendly. In this paper, the recommendation system algorithms used in the travel domain are studied. The focus of the research is on how machine learning and artificial intelligence are being used to improve the sector. The paper also includes the study of different recommendation algorithms that can be used to generate personalized suggestions.

1.1 INTRODUCTION:

Tourism is one of the interesting and always emerging domains. Nowadays it becomes even more interesting due to advanced technology. It becomes so easy for people to find suitable destinations as well as the most attractive places of that destination. Moreover, the travel companies can suggest to users about hotels and traveling modes according to their taste. Additionally, having such functionalities engage users for a longer time and provide a satisfying experience. As a result, users prefer to use the same application and website again which increases the revenue of the company.

In any industry, success is not only a matter of services but also about how efficient they are. Because according to natural human tendency the human brain gets attracted to things that are similar to the person's preferences. And to find out that for every user is a challenging task. A tremendous amount of data is generated by the tourist industries so analyzing that data and find useful insights can only be performed by machine learning algorithms. Machine learning is the subset of artificial intelligence where the models are built to predict the results based on given data. There are lots of machine learning algorithms invented through which high accuracy predictions can be generated. They are powerful enough to handle large datasets and figure out important insights from them. This paper includes a thorough study of how machine learning and artificial intelligence are used in the tourism sector.

1.2 .LITERATURE SURVEY:

1.2.1-Tourism Recommender System using Machine Learning Based on User's Public Instagram Photos

Recommendation systems are used in many entertainment apps like Netflix, Youtube, and music apps. Using a user's public Instagram photos is a new approach to recommendation systems. This method aims to design a recommendation system without having to ask users a bunch of questions. It extracts useful data from the feed.

1.2.2- A machine learning-based tourist path prediction

Recommendation systems guiding people about where to go is a very helpful tool for personal as well as commercial travel. It is based upon tourist's historical visiting sequences and machine learning algorithms like LambdaMART and Random Forests. The comparison between five machine learning algorithms, ListNet, Ranking SVM, LambdaMart, Rank Boost, and Random Forests show that Random forest outperforms the other algorithm.

1.2.3- Management of Tourism Resources and Demand Based on Neural Networks

Predictions for inbound tourism demand are important for the development of inbound tourism strategies in China. The BP neural network is a common traditional machine learning method that is widely used but has limitations such as overfitting and difficulties in setting parameters. This paper combines the BP neural network with an advanced machine learning algorithm called ensemble learning to solve the problems.

2.REVIEWED APPLICATIONS:**2.1.Arinib:**

Arinib is an application used for finding accommodation by tourists. Arinib connects thousands of people who want to rent their house or property as accommodation for tourists. This application is accessible via mobile as well as website.

2.2. Tripadvisor;

Tripadvisor is an application where people can plan their trip based on the recommendations and reviews given by various tourists. It is the platform where suggestions are made for each user according to their preferences. This is found as the best example of machine learning in the tourism sector.

2.3. Tripit:

Tripit is a trip planning application but a little different than other applications. Here the customer sends their booking information to Tripit and then it will generate the plans for the customer. The hours of planning is done within few minutes through tripit

2.4.Trivago:

Trivago compares the hotel worldwide from 300 different websites. Customers can easily find the hotel as per their budget and requirement within a few clicks. The booking of the hotel can also be done with the respective hotel's website. Trivago only compares and helps to find out the best hotel for the user. For other functions such as booking, the user is forwarded to that particular website.

3.IMPLICATIONS OF AI AND ML IN TOURISM:**3.1. Recommendation systems:**

Recommender systems are machine learning systems used to predict things as per user behavior or preferences and aim to recommend products in which they are interesting. Recommendation engines are made up of machine learning algorithms that analyze all information about users' likes and dislikes, ratings, reviews, and also search queries. Based on these factors the products are suggested by the machine to the user. Sometimes it also performs predictions of most likely items based on similar users. Mostly in travel sectors, tourists require the most attractive place of the destination, accommodation facility, and cultural food items, these can be easily predicted to them by using above-mentioned factors additionally, location wise recommendations are most popular to recommend nearby services.

Following are some recommendation algorithms most commonly being used in the industry:

A. Collaborative Filtering Recommendation:

Collaborative filtering recommender systems give suggestions based on the preferences of similar users. It needs only users' historical preferences on a certain item set. There are two techniques used in collaborative filtering viz. Model-based collaborative filtering and memory-based collaborative filtering. In Model-based collaborative filtering rating, the dataset is used as a source. The information from the rating dataset is retrieved and used as a model to make recommendations. Clustering techniques are used to cluster users into similar groups and make suggestions based on group preferences to which user belongs. Newly added information to the database is neglected by the system as there are no initial ratings available for those items.

In memory-based collaborative filtering, similar users which have similar ratings for similar items are found, and then the target user's rating for the item with which the target user has never interacted is predicted.

User-User:

User-user algorithms work on the concept of how users are similar to each other. Let's say user A is purchasing product P and product Q, the other user B purchases product P then this algorithm finds similarity between those A and B users and recommends product Q to the user B as well. These algorithms are most widely used by travel agencies to recommend destinations, hotels, and restaurants.

Item-item:

Item-item algorithms consider the products which are purchased together by several users. For example, the number of users has bought bread, butter and milk together then, any user purchases bread and butter then milk will get recommended to him because they are frequently bought together by several customers.

User-item:

It is the hybrid form of user-user and item-item. It is generally based on matrix factorization techniques.

B.Content-based filtering Recommendation:

Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback. Other users' feedback or preferences are not involved. Data for the recommendation process is obtained from item features and user profiles. The similarity between item feature and user profile is calculated and the item with the highest similarity with users preferences is selected. User profile data is obtained in two ways: through explicit sources and through implicit sources. Explicit data is the data which is given by the user to the system through the signup and implicit data is obtained from users

browsing history, downloads, subscriptions, etc. There are some limitations of content-based recommendation filtering. Recommendations are made only on the basis of only users' own profiles without considering others' feedback and preferences. Another limitation is the cold-start problem.

C.Knowledge-based recommendation:

A recommender system is knowledge-based when it makes recommendations based not on a user's rating history, but on specific queries made by the user. Knowledge-based recommendations are not dependent on the rating, nor do they have to gather information about a particular user to give recommendations. Knowledge base recommendation is a good choice for complex domains where no history or previous ratings are available. The data is collected explicitly when the user interacts with the system and stored in the form of statements, rules, or analogies. For this data storage, two methods are used case-based reasoning and rule-based reasoning.

3.2.CHATBOTS:

Chatbots are basically chatting windows that communicate with customers. Artificial intelligence and NLP are the most commonly used technologies for chatbots. There comes out to be a great source to answer customers' inquiries without delay. Travel companies use these chatbots to provide immediate information to the end-user so that the employee's time is saved and can be utilized for other tasks. The customers are using these chatbots to find out exotic places, restaurants that must be visited, how to go to a particular place, etc, even after working hours or when the information center is closed. Chatbots are more compatible with websites or the platform company is using that is why users need not install any extra application on their system. Hesitation for asking questions is removed as users can chat with these chatbots in their own words without concern about the language accuracy and fluency. On the other hand, chatbots are critical assets for research, collecting useful data, and knowing users' requirements and interests. Chatbots can store all their conversation that is used for research. And the important one is that 24/7 availability and immediate response to the doubts satisfy users and creates a happy and favorable environment.

3.3. Predictions:

People are making several queries like when will the flights reopen for certain countries as now there are restrictions due to covid-19, when the ticket fare will be dropped, to answer such questions prediction needs to be performed first. users are asking such queries because they trust whatever the prediction is made so the accuracy must be as high as possible. These predictions are done with the Machine learning algorithms such as:

A. Linear Regression:

Linear regression is a supervised machine learning algorithm that provides a relationship between two continuous variables by generating the best fit line. One among the two values is dependent variables and the other one is an independent variable. The following equation represents linear regression:

The model takes :

X: Input data from the training set

Y: labels to the data

In the training phase of the model, it fits the best line for predicting the value of Y for the input value of X. The best fit line is the line for which total prediction errors are small and that is obtained by finding the best values for θ_1 and θ_2 .

θ_1 : intercept

θ_2 : coefficient of x

B. Decision Tree:

In decision tree the data is split continuously according to certain parameters. Leaves and decision nodes are the two main components of the decision tree. The leaves are nothing but the final outcome and nodes are the points where data get split.

Finding attributes for the nodes at each level is the most challenging part of the Decision Tree. And this is called attribute selection. attribute selection has two methods as follows:

B.1. Information gain:

Information gain is nothing but a change in entropy that occurred while splitting training data into small subsets using decision tree nodes.

B.2. Gini Index:

Gini Index is used to find the frequency of how many times the randomly chosen elements are incorrectly identified. So the attribute with a lower Gini Index is preferred.

C. Support Vector Machine:

It is an unsupervised machine learning algorithm and is used for both classification and regression. It used Hyperplane to represent different classes in multidimensional space. Support vector points are the data points that are nearest to the hyperplane and the distance between hyperplane and support vector points is called margin. The formula for this is

$$f(x) = B0 + \sum(ai * (x,xi))$$

D. K nearest neighbor KNN:

It is again a supervised machine learning algorithm used for both classification and regression algorithms. KNN algorithms group similar data and classify the new coming data into the appropriate group. The distance between the new point and the existing point is calculated by using Euclidean distance and Manhattan Distance.

3.4. Marketing:

Marketing plays an important role in business. Ai and ML techniques take the marketing domain to a much higher level where marketers can easily target the correct customers by analyzing cookies and device Ids.

3.5. Fraud Detection:

ML algorithms have the capability to learn from historical fraud patterns and detect them in future transactions. ML assesses the customer behavior individually and as soon as any abnormal behavior is detected the transaction gets blocked for that customer.

Travel companies tackle frauds like:

- Credit card frauds
- Fake accounts
- Abusive content
- Loyalty fraud

4. References:

- [1] Jake VanderPlas.: Python Data Science Handbook Essential Tools for Working with Data.
- [2] Danny Varghese., Comparative Study on Classic Machine learning Algorithms. <https://bit.ly/36cl0vt>, last accessed 18/08/2020.
- [3] G. Linden., B. Smith., J. York.: Amazon.com recommendations: item-to-item collaborative filtering, IEEE Internet Computing (Volume: 7, Issue: 1,Jan.-Feb. 2003).
- [4] Marwa Hussien Mohamed., Mohamed Helmy Khafagy., Mohamed Hasan Ibrahim.: Recommender Systems Challenges and Solutions Survey, 2019 International Conference on Innovative Trends in Computer Engineering (ITCE), Aswan, Egypt, Egypt.
- [5] Kunal Shah. Akshaykumar Salunke., Saurabh Dongare., Kisandas Antala.: Recommender systems: An overview of different approaches to recommendations, 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India.
- [6] researchgate.net/publication/344729288_Machine_Learning_in_Tourism
- [7]<https://www.emerald.com/insight/content/doi/10.1108/JTF-09-2019-0083/full/html#:~:text=Originality%2FValue,current%20trend%20of%20tourist%20demand.>