WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A TRIP ADVISOR SYSTEM USING DECISION TREE

Mrs. Pratibha Mishra[1], Gagana P Giridhar[2], J Sudha[3], J Swathi[4], M Shireesha[5]

[1]Asst. Professor, Department of CSE, Ballari Institute of Technology & Management, Ballari. prathibhamishra@bitm.edu.in
[2]Department of CSE, Ballari Institute of Technology & Management, Ballari pgaganag@gmail.com
[3]Department of CSE, Ballari Institute of Technology & Management, Ballari joladharashisudha@gmail.com
[4]Department of CSE, Ballari Institute of Technology & Management, Ballari. swathigouda98@gmail.com
[5]Department of CSE, Ballari Institute of Technology & Management, Ballari shireeshamu24@gmail.com

ABSTRACT:

Selecting the ideal tourist spot from lot of available information online and from other sources constitute a outstanding challenge for travellers during the planning phase and even during their journey. While prior Trip Advisor System have aimed to encounter this mess, they often fall short in respect of precision and user satisfaction. This neglects crucial technical aspects and practical usability factors. To overcome these shortcomings, a deep knowledge of tourists' decision-making processes and innovative models for information search are essential. This study introduces an innovative TAS that prioritizes the requirements and inclination of travellers, particularly in unfamiliar cities. This system considers both real and scientific characteristics in detail to ensure the best recommendations possible, unlike conventional methods. It employs a unique two-step method for selecting features to streamline the input data and utilizes decision tree to generate personalized recommendation. The experimental findings demonstrate the productiveness of this novel TAS in delivering tailored recommendations that meets the varied needs of travellers to enhance their overall satisfaction and travel experience.

Keywords: Random Forest, Gradient Boosting, and XG Boost classifiers.

INTRODUCTION:

Tourism plays a crucial part in shaping Karnataka's economic landscape, making a substantial contribution to its GSDP and employment opportunities. Lately, Karnataka has emerged rapidly expanding tourist destinations globally, showcasing its abundant aesthetic legacy and vibrant traditions. According to statistics, tourism contributed approximately 14.8% to Karnataka's GSDP in 2024, with a steady growth trajectory expected in the years ahead. The Karnataka government has identified tourism as an important source for economic growth and development, implementing different inventiveness to enhance infrastructure, promote sustainable tourism practices, and attract international visitors. With iconic landmarks like Hampi,Mysuru,Jogfalls,Bengaluru,LalbaghBotanicalGarden,Udupi.Karnataka offers a plethora of attractions catering to diverse interests and preferences. In particular, regions like Coorg, Mysore Palace, Gokarna Beach have gained and identified as compelling tourism offerings, drawing millions of domestic and international tourists annually. The Karnataka tourism sector has witnessed a significant surge in both inbound and domestic tourism, reflecting a growing interest among travellers to traverse the country's rich tapestry of experiences. In modern period, the internet has revolutionized the way tourists gather information, plan their itineraries, and make travel-related decisions. With the rapid expansion of digital platforms, tourists now are able to gather a vast repository of information, enabling them to discover destinations, compare options, and customize their travel experiences according to their preferences. However, amidst this abundance of information, travellers often face challenges in navigating through the plethora of choices, making informed decisions, and make sure a great travel experience. This satisfies the demands for advanced recommendation systems that leverage data mining techniques to provide personalized recommendations customized to individual preferences, interests, and constraints. This project aims to address the aforementioned challenges by proposing a novel human-centric Trip Advisor System (TAS) customized for Karnataka tourism landscape. By utilizing the potential of data mining methodologies, the proposed TAS seeks to streamline the tourist decision-making process, enhance user satisfaction, and optimize the complete travel adventure.

LITERATURE SURVEY:

In the paper [1], titled "Smart tourism products and services The paper by Huimin Hu and Chunhong Li (2023) delves into incorporated the structure of smart tourism, utilizing big data and AI techniques. They prioritize user experience, offering personalized recommendations and real-time

decision-making. By leveraging the Decision Tree algorithm and analysing user feedback, they aim to optimize facilities and enhance client satisfaction. Overall, the study underscores the revolutionary influence of smart tourism in time of large data and AI integration.

In the paper [2]"To Analyse the Various Machine Learning Algorithms That Can Effectively Process Large Volumes of Data and Extract Relevant Information for Personalized Travel Recommendations" by Karthiyayini, J., & Anandhi, R. J. (2024), it concentrate on evolving recommendation systems, particularly in tourism. It reviews recommendation types, AI and ML techniques, data processing, and algorithmic methodologies. The study summarizes current research, highlighting challenges and future directions in developing proficient tourism recommendation systems

In the paper [3] titled "Exploiting the similar issues using decision tree for recommendation" by Kolankar, P., Patel, R., Dangi, N., Sharma, S., & Jain, S. (2020) introduces an algorithm designed to filter a large database of earthquake cases efficiently for recommendation purposes. Showcased during the REDSET 2019 conference, the algorithm matches input cases with stored ones, generating recommendations based on similarity. By concentrating on the earthquake disaster domain, they utilize the ID3 algorithm to generate a decision tree, streamlining computations for matching similar cases. Overall, the paper offers a method to leverage decision trees for earthquake recommendations, efficiency in fetching best case.

In the paper [4]"Sustainability of Tourism Destinations and the significance of Certification" by Costa, J., Rodrigues, D., & Gomes, J. (2019), explores tourism destination sustainability and the certification title. It highlights the significance of environmental certification in aiding consumer choices and strategic marketing for companies. By examining literature on tourism impacts and sustainable development, the paper underscores certification's significance in fostering environmentally responsible tourism. It offers original insights into destination certification's benefits, often overlooked in the tourism sector. Ultimately, the research enriches our understanding of sustainable tourism practices and emphasizes certification's crucial role in promoting eco-friendly tourism development.

In the paper [5] "A decision tree based recommendation system for tourists" authored by Thiengburanathum, P., Cang, S., & Yu, H. (2015) introduces a human-centric Travel Recommendation System (TRS) to assist tourists in selecting destinations. It addresses previous TRS limitations by employing real-world data and a two-step feature selection method. The system utilizes the algorithm to offer personalized recommendations. Early consequences indicate its productiveness in providing tailored suggestions aligned with tourists' preferences and needs.

PROPOSED METHODOLOGY:

1. **Data Acquisition:** Start by identifying sources of data relevant to tourist preferences, destinations, attractions, and reviews. These sources can include tourism websites, review platforms like TripAdvisor, digital platforms, tourism portfolio, and government tourism portals. Utilize web scraping techniques or APIs provided by these platforms to gather data. This data may include destination names, descriptions, user reviews, ratings, and other relevant information. Ensure observance with terms of usage of data sources.

2. **Data Pre-processing:** Following data acquisition, pre-processing streamlines subsequent analysis with key steps. Cleaning removes duplicates, irrelevant entries, and inconsistencies. Text pre-processing refines textual data, employing tokenization, stop word removal, and lemmatization. Categorical variables are encoded into numerical representations using techniques like one-hot encoding. Unavailable data handled through imputation or removal. This prepares facts and figures for data analysis, setting the stage for the empowerment of an effective recommendation system.

3. **Data Analysis:** With pre-processed data analysis involves exploring patterns and relationships crucial for a trip advisor using decision tree. This contains Exploratory Data Analysis (EDA) to unveil trends and correlations, selecting attributes for pertinent variables, and training the trip advisor model using decision tree with algorithms. Model evaluation using metrics such as accuracy and precision assesses the system's efficacy in providing accurate suggestions to tourists.

4. **Result Interpretation:** Following model training and evaluation, interpret the results to derive actionable insights. This involves understanding decision rules provided by decision trees to grasp how recommendations are made. Identifying key features reveals which variables significantly influence recommendations, aiding in prioritizing areas for refinement. Utilizing visualization techniques such as tree diagrams or feature importance plots effectively communicates the model's findings. Iteratively refine the recommendation system based on result interpretation by fine-tuning the model, updating data sources ensuring continuous improvement and enhanced performance.

SIGNIFICANCE OF "A TRIP ADVISOR SYSTEM USING DECISION TREE":

The significance of this study lies in the utilization of a decision tree-based trip advisor system aimed on a single dataset. Despite its limitations, such as being confined to the dataset's scope, this proposal provides valuable perspectives and benefits. By efficiently providing personalized recommendations without external data sources, the system streamlines the user experience. It fosters trust and engagement by tailoring suggestions to the dataset's characteristics, enhancing user satisfaction. Moreover, while recommendations may be restricted, they encourage exploration within

the dataset's domain, potentially unveiling hidden attractions. Additionally, this study lays the groundwork for future enhancements and iterations of the recommendation system, contributing to the advancement of tourist planning tools.

MOTIVATION FOR PROJECT:

The imaginative idea for this project arises from addressing a remarkable niche market in the realm of travel planning. Despite the lot of travel information available online, tourists frequently encounter challenges in finding recommendations that precisely match their individual choices and constraints. This project seeks to leverage decision tree algorithm as an innovative solution to this problem. These algorithms are well-suited for handling complex decision-making processes, allowing for tailored recommendations that accommodate traveller needs. By adopting a user-centric approach, this system aims to prioritize the specific requirements of tourists, ultimately enhancing their travel planning experience and overall satisfaction.

OBJECTIVES OF THE PAPER:

1. To provide tailored travel suggestions based on user input.
2. To enhance recommendation accuracy, decision tree algorithm to process and analyse user data, enabling more precise and tailored suggestions based on individuals.
3. To streamline travel planning, the system provides quick recommendations based on predefined criteria, helping users find suitable destinations and activities without extensive research.
4. To leverage user data in refining and enhancing the recommendation process, our focus lies in utilizing existing predefined datasets to derive insights and optimize recommendations, ensuring adaptability within the constraints of available data.

ALGORITHMS:

1. RANDOM FOREST:

Random Forest is a machine learning method used for prediction. It leverages ensemble learning, combining many decision trees to solve complex tasks. Through bagging, or bootstrap aggregating, the algorithm trains a 'forest' of decision trees. Predictions are made by averaging outputs from these trees, enhancing precision. Random Forest overcomes limitations of decision tree by reducing overfitting and providing higher accuracy without extensive configuration. It handles missing data effectively and doesn't require extensive hyperparameter tuning. Each tree selects a randomly selected subset of features at node splitting points. Decision trees, forming the basis of Random Forest, consist of decision, leaf, and root nodes, recursively partitioning the data until reaching terminal leaf nodes.

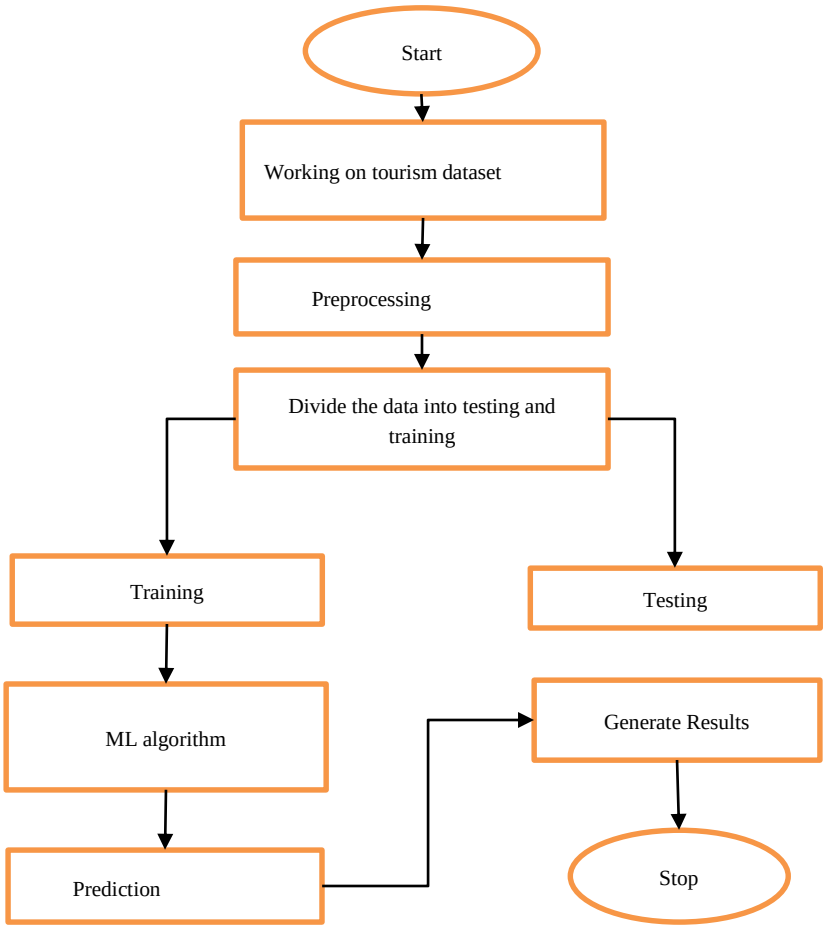
2. XGBOOST:

XG Boost, an acronym for "Extreme Gradient Boosting," is a fundamental aspect in ML, celebrated for its unparalleled efficiency, flexibility, and portability. As a streamlined distributed gradient boosting library it has gained broad acclaim for its capability to implement machine learning methods within the Gradient Boosting framework. By offering parallel tree boosting, XG Boost facilitates the swift and accurate resolution of a different data science problems. Unlike bagging algorithms, XG Boost boasts the unique capability of controlling both bias and variance, rendering it exceptionally effective in model optimization. Its popularity is not unfounded, as it has proven itself in Kaggle competitions, particularly excelling in handling structured data. Central to its success is its ingenious extension of gradient boosted decision trees (GBM) and its adoption of column subsampling, approaches that not only deter overfitting but also expedite computations. Respectively, XG Boost stands as an indispensable tool for data scientists and ML enthusiasts seeking to push the boundaries of their models and achieve superior performance.

3. GRADIENT BOOSTING:

The gradient boosting technique is a redoubtable force in machine learning, renowned for its ability to minimize bias error within models. Unlike Ad boosting, where we can specify the base estimator, gradient boosting mandates by utilizing a fixed Decision Stump. However, we can modify the n-estimator parameter, defaulting to 100 if unspecified. Its versatility extends to predicting both continuous and categorical variables, serving as a classifier with Mean Square Error (MSE) and Log Loss as respective cost functions. Consider an example where Age act as a dependent variable and Likes Exercising, go to Gym, and Drives Car are predictors, employing the Gradient Boosting Regress. To elucidate the idea of estimator-2, unlike Ada Boost, residues from the first estimator form root nodes.

DATA FLOW DIAGRAM



DISCUSSION:

The study's results showcase the efficacy of the advanced methodology, combining supervised and unsupervised machine learning. Overall, the methodology presents a real way of recommendations.

CONCLUSION:

The current difficulty of the destination TRS has been labelled in this work by presenting a decision tree-based tourist recommendation system. Using relevant tourist domain knowledge, the facts and figures divides into two sub datasets. This was done to enhance classification perfection while simultaneously decreasing the decision tree's complexity. Finally, the experiment suggests that the proposed TAS is applicable. The planned TRS meets the needs of visitors who are planning to come or are already in dataset.

REFERENCES:

1. Hu, Huimin, and Chunhong Li. "Smart tourism products and services design based on user experience under the background of big data." Soft Computing 27.17 (2023): 12711-12724.

2. Thiengburanathum, P., Cang, S., & Yu, H. (2015, September). A decision tree based recommendation system for tourists. In 2015 21st international conference on automation and computing (ICAC) (pp. 1-7). IEEE.

3. Costa, J., Rodrigues, D., & Gomes, J. (2019). Sustainability of tourism destinations and the importance of certification. Worldwide Hospitality and Tourism Themes, 11(6), 677-684.

4. 4.Khadar, A. A., Hayath, T. M., & Nagraj, M. S. (2017, August). Design and implementation of ZigBee based smart grid system for power management. In 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon) (pp. 1339-1343). IEEE.

5. Karthiyayini, J., & Anandhi, R. J. (2024). To Analyse the Various Machine Learning Algorithms That Can Effectively Process Large Volumes of Data and Extract Relevant Information for Personalized Travel Recommendations. SN Computer Science, 5(4), 336.

6. Kolankar, P., Patel, R., Dangi, N., Sharma, S., & Jain, S. (2020). Exploiting the most similar cases using decision tree to render recommendation. In Data Science and Analytics: 5th International Conference on Recent Developments in Science, Engineering and Technology, REDSET 2019, Gurugram, India, November 15–16.

7. Gao, Z., Cheah, J. H., Lim, X. J., Ng, S. I., Cham, T. H., & Yee, C. L. (2023). Can travel apps improve tourists' intentions? Investigating the drivers of Chinese gen Y users' experience. Journal of Vacation Marketing, 13567667231152938.

8. Garipelly, V., Adusumalli, P. T., & Singh, P. (2021, July). Travel Recommendation System Using Content and Collaborative Filtering-A Hybrid Approach. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-4). IEEE.

9. Bhaidkar, Y., Bhagwat, P., Bhalere, P., Gujar, R., & Walunj, S. (2015). Tourist place recommendation system. Int. J. Adv. Res. Innov. Ideas Educ.(IJARIIE), 2, 3032-3037.

10. Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: a survey. Decision support systems, 74, 12-32.

11. L. Kzaz, D. Dakhchoune and D. Dahab, "Tourism Recommender Systems: An Overview of Recommendation Approaches", Int. J. Comput. Appl., vol. 180, no. 20, pp. 975-8887, 2018.