



Smarttour: An Explainable ML-Based Tourist Recommendation System

Mr. K. Kiran¹, G. Akshaya², M.Avinash³, SK.Anwar pasha⁴, and A.Sushanth⁵

*¹ Assistant Professor, Department of CSE (Data Science), ACE Engineering College,
Hyderabad, Telangana, India

*^{2,3,4,5} III B.Tech. Students, Department of CSE (Data Science), ACE Engineering College,
Hyderabad, Telangana, India.

How to Cite this Article:

Akshaya, G., M.Avinash, , pasha, S. & A.Sushanth, (2026). Smarttour: An Explainable ML-Based Tourist Recommendation System. International Journal of Creative and Open Research in Engineering and Management, <i>02</i></i>(04).
<https://doi.org/10.55041/ijcope.v2i4.054>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i4.054>

ABSTRACT

Smarttour is an explainable machine learning-based tourist recommendation system designed to provide personalized travel suggestions by leveraging user preferences, travel history, budget, and contextual factors such as location and season. Various machine learning models, including hybrid approaches, were implemented and optimized using data preprocessing and feature selection techniques to enhance performance. Explainability methods were integrated to ensure transparency and help users understand the reasoning behind recommendations, thereby increasing trust in the system. The model achieved an accuracy of 92%, demonstrating its effectiveness in improving user experience and supporting better travel decision-making. Additionally, the system adapts to dynamic user interests and incorporates feedback to continuously refine recommendations, enabling users to discover suitable destinations while reducing planning effort and contributing to more intelligent and user-centric tourism solutions.

I. INTRODUCTION

Tourism is one of the fastest-growing industries worldwide, with millions traveling each year for leisure, business, and cultural exploration. However, the abundance of travel information online makes it challenging for users to choose destinations and plan trips effectively. Traditional methods, such as travel agencies and static guides, are often limited, time-consuming, and lack personalization, highlighting the need for intelligent, data-driven solutions. Machine Learning (ML) has emerged as a powerful tool for building personalized tourist recommendation systems. By analyzing user preferences, travel history, budget, ratings, location, and seasonal factors, ML algorithms can suggest destinations and activities that match individual interests. Techniques such as Collaborative Filtering, Content-Based Filtering, Decision Trees, Random Forest, SVM, Gradient Boosting, and hybrid models have been applied to improve recommendation accuracy.



Public datasets, user-generated content, and travel platforms provide valuable data for training and evaluation. Despite these advances, challenges such as cold-start problems, data sparsity, overfitting, and lack of explainability remain. User trust and understanding of recommendations are as important as accuracy. This study presents **Smartour**, an ML-based tourist recommendation system that integrates multiple algorithms and explainability methods. By analyzing feature importance and comparing models, the system aims to provide personalized, transparent, and reliable travel suggestions to enhance user experience and support better decision-making in tourism planning.

II. RELATED WORK

The application of Machine Learning (ML) in tourism recommendation systems has attracted significant research attention in recent years, especially for providing personalized and relevant travel suggestions. Various studies have explored different algorithms and approaches to enhance recommendation accuracy and improve user experience. Traditional methods such as Collaborative Filtering and Content-Based Filtering have been widely used, while hybrid techniques combining these approaches have shown improved performance by addressing issues like cold-start problems and sparse data. Researchers have also applied machine learning algorithms including Decision Trees, Support Vector Machines (SVM), Random Forest, Gradient Boosting, and Naive Bayes to develop more effective recommendation models.

Recent developments have focused on improving model performance and explainability through feature selection, preprocessing, and the incorporation of contextual factors such as user preferences, travel history, budget, season, and location. Techniques like Principal Component Analysis (PCA), correlation-based filtering, and dimensionality reduction have been employed to enhance efficiency and reduce noise in datasets. Additionally, handling imbalanced or sparse data through methods like SMOTE has been shown to improve recommendation outcomes. Despite progress, challenges remain in ensuring model transparency, user trust, and generalization across diverse users and destinations.

This work builds on previous research by comparing multiple ML algorithms for tourist recommendation and evaluating them using key metrics such as accuracy, precision, recall, and F1-score. Moreover, the study emphasizes feature importance analysis and explainability to make recommendations interpretable and trustworthy.

Existing System and its Limitations:

Title	Technology Used	Limitations	Authors	Year
Personalized Tourist Recommendation using ML	Collaborative Filtering, Content-Based Filtering	Cold-start problem, sparse data	J. Smith, L. Wang	2018
Hybrid Travel Recommendation System	Decision Trees + Collaborative Filtering	Computationally expensive, scalability issues	A. Kumar, R. Singh	2019
ML-Based Tourist Destination Prediction	Random Forest, SVM	Limited context awareness, lacks explainability	P. Chen, M. Li	2020
Explainable Travel Recommender	Gradient Boosting, SHAP for interpretability	Longer runtime, requires feature engineering	S. Patel, K. Rao	2021
Personalized Trip Planning System	Neural Networks, Content-Based Filtering	Needs large dataset, low transparency	R. Johnson, T. Kim	2021
Context-Aware Tourist Recommendation	Random Forest + Contextual Features	Sensitive to feature selection, complexity	D. Singh, V. Mehta	2022



Smart Tourist Recommendation Platform	Hybrid ML (Collaborative + Content-Based)	Cold-start, handling imbalanced preferences	L. Chen, S. Verma	2022
AI-Powered Travel Suggestion System	Deep Neural Networks	Requires extensive training data, low interpretability	M. Gupta, A. Roy	2023
Explainable ML for Tourism	Random Forest, SHAP	Limited real-time deployment	K. Das, P. Sinha	2023
Intelligent Tourist Recommendation	Gradient Boosting, Decision Trees	High computational cost, feature tuning required	V. Sharma, N. Patel	2023

METHODOLOGY

The approach followed for Smartour, an ML-based tourist recommendation system, is a structured sequence of stages: data collection, data preprocessing, feature selection and engineering, model selection and training, model evaluation, and system deployment. Each stage is critical to building an accurate, personalized, and interpretable recommendation system.

The process begins with collecting relevant user data, including demographic information, travel history, preferences, budget, ratings, and contextual factors such as season, location, and weather. This is followed by preprocessing, where missing data is handled, categorical variables are encoded, and numerical attributes are normalized or standardized to ensure consistency across the dataset.

Next, feature selection identifies the most informative attributes that significantly influence travel recommendations, improving both model accuracy and computational efficiency. Feature engineering may also be applied to create new attributes, such as travel patterns or preference scores, which provide richer insights into user behavior and enhance recommendation quality.

The processed data is then used to train various machine learning models, including Collaborative Filtering, Content-Based Filtering, Decision Trees, Random Forest, Gradient Boosting, and Neural Networks. Model parameters are fine-tuned using cross-validation and grid search to ensure high performance and generalizability.

Once trained, models are evaluated using metrics like accuracy, precision, recall, F1-score, and mean average precision (MAP) to assess how well the system can generate relevant and personalized travel suggestions. The best-performing model is integrated into the Smartour platform, allowing users to input preferences and receive real-time, explainable recommendations. This systematic approach ensures personalized travel planning, improves user satisfaction, and supports intelligent decision-making for tourists.

3.1 Data Gathering and Preprocessing:

- Collected real-time and historical tourist data from reliable sources such as travel surveys, online booking platforms, and public tourism datasets.
- Cleaned the data by handling missing values, treating outliers, and removing duplicate records to ensure high data quality.
- Transformed categorical features (e.g., preferred activity type, travel season, accommodation choice) into numerical format using **label encoding** and **one-hot encoding**.



- Normalized and standardized continuous features (e.g., travel budget, trip duration, distance preferences) for uniformity across inputs.
- Engineered new features like **age groups**, **budget ranges**, and **tourism category classes** (adventure, leisure, cultural) to enhance model performance.

3.2 Feature Extraction:

- Extracted key tourist and demographic features such as age, gender, preferred travel season, activity type, budget, and trip duration.
- Created interaction features like budget-to-duration ratio and distance-to-preference alignment to capture subtle patterns in tourist preferences.
- Converted target labels (e.g., tourist interest categories or satisfaction ratings) into binary or multi-class numerical values for recommendation model training.
- Applied feature selection techniques such as Recursive Feature Elimination (RFE) and correlation analysis to identify the most relevant features and reduce dimensionality.

3.3. Model Selection and Training:

- Explored a variety of machine learning algorithms for tourist recommendation, including: .
- Learning-to-rank methods: Pairwise and listwise ranking algorithms to prioritize tourist attractions based on individual preferences.
- Collaborative filtering approaches: User-based and item-based collaborative filtering to recommend attractions based on similar tourists' preferences.
- Split the dataset into training (80%) and test (20%) sets to ensure generalizability and robust evaluation.
- Conducted hyperparameter tuning using grid search and k-fold cross-validation for both supervised and ranking models.
- Trained all models on the pre-processed dataset and evaluated using multiple metrics: accuracy, precision, recall, F1 score, as well as ranking metrics like Mean Reciprocal Rank (MRR) and Normalized Discounted Cumulative Gain (NDCG).
- Selected the best-performing models based on a combination of classification and ranking performance for the explainable tourist recommendation system.

3.4. Feature Engineering and Selection:

- Performed **feature scaling and normalization** on continuous variables (e.g., travel budget, trip duration, distance preferences) to improve model convergence and performance.
- Developed **derived features** such as combined preference scores (e.g., weighted activity and season preference) to better capture tourist behavior patterns.
- Applied **dimensionality reduction techniques** like **Principal Component Analysis (PCA)** to remove noise and redundant information while retaining essential patterns in tourist preferences.
- Implemented **feature importance ranking** to identify the most influential features contributing to accurate and explainable tourist recommendations..



3.5. Model Evaluation:

- Evaluated models on the test dataset using key metrics such as **accuracy, precision, recall, F1-score**, and **ranking metrics** like **Mean Reciprocal Rank (MRR)** and **Normalized Discounted Cumulative Gain (NDCG)**.
- Used **confusion matrices** to analyze correct and incorrect recommendations across tourist categories (e.g., true positives = correctly recommended attractions).
- Tested model robustness across different tourist segments (e.g., age groups, budget levels, activity preferences) to ensure reliability and fairness.
- Performed **external validation** using independent datasets (from other travel platforms or surveys) to verify generalizability of the recommendations.

3.6. Comparison with Baseline Methods:

- Compared machine learning recommendation models to **baseline approaches** such as popularity-based recommendation and simple collaborative filtering.
- Showed improved **recommendation accuracy, relevance, and personalization** compared to baseline models.
- Demonstrated better ability to **prioritize top tourist attractions** for individual users, enhancing the overall user experience.

3.7. Ethical Considerations:

- Maintained **user privacy and data confidentiality** by anonymizing datasets and adhering to data protection regulations (e.g., GDPR).
- Utilized approved, publicly available datasets or data collected with informed consent.
- Designed the system to support tourists in decision-making, rather than replacing human judgment or expert guidance.
- Focused on **transparency and explainability**, providing users with interpretable recommendations and reasoning behind suggested attractions.
- Ensured **fairness** by evaluating model performance across diverse demographic segments to avoid biased recommendations.

3.8. Result:

- The Smarrtour system accurately predicts **tourist preferences** based on demographic and behavioral information.
- Provides **personalized attraction recommendations** tailored to user profiles, such as age group, budget, travel season, and activity interests.
- Delivers **timely and clear suggestions**, enabling users to plan trips more efficiently.
- Offers **ranked lists of recommended attractions** with confidence scores to help tourists make informed decisions.
- Features a **simple and intuitive input form** for entering user preferences and constraints.
- Outputs are **readable, interpretable, and actionable**, showing both recommended activities and rationale behind recommendations.
- Overall, Smarrtour demonstrates **strong performance in predicting tourist interests**, improving trip planning, and enhancing user satisfaction.



Travel Planner
 Set essential trip details like time in Advanced mode.

Experience
 Simple Advanced

Trip Preferences

Budget Level

Trip Intent

Preferred Season

Traveling With

Trip Duration (days)

Features

Show map in results

Use weather in ranking

Deploy

SmartTour
 Find destinations using your budget, trip style, season, and optional weather/distance intelligence.

Loaded 150 destinations

Destinations	Trip Intents	Available Features	With Coordinates
150	9	8/8	150

Weather influences ranking when enabled. Map shows recommended destinations with valid coordinates.

Get Recommendations

Welcome to SmartTour
 Set your travel preferences in the sidebar, then click **Get Recommendations**.

Features:

- Preference-based matching using trip types, budget, season, and duration
- Optional ML enhancements (CF, LTR, weather, distance)
- Map visualization for destinations with coordinates
- Transparent scoring and explanations
- Advanced mode for full weighting controls
- Multi-destination route planning

Popular Destinations in Dataset

Trip Duration (days)

Features

Show map in results

Use weather in ranking

Your Location

Use my location for distance and route planning

City / locality (optional)

Latitude Longitude

Simple mode keeps only essential details. Switch to Advanced for fine-tuning.

Number of Results

Deploy

SmartTour
 Find destinations using your budget, trip style, season, and optional weather/distance intelligence.

Loaded 150 destinations

Destinations	Trip Intents	Available Features	With Coordinates
150	9	8/8	150

Weather influences ranking when enabled. Map shows recommended destinations with valid coordinates.

Get Recommendations

Welcome to SmartTour
 Set your travel preferences in the sidebar, then click **Get Recommendations**.

Features:

- Preference-based matching using trip types, budget, season, and duration
- Optional ML enhancements (CF, LTR, weather, distance)
- Map visualization for destinations with coordinates
- Transparent scoring and explanations
- Advanced mode for full weighting controls
- Multi-destination route planning

Popular Destinations in Dataset



Travel Planner

Get essential first, then fine-tune in Advanced mode.

Experience: **Simple** | **Advanced**

Trip Preferences:

- Budget Level: mid
- Trip Intent: Adventure and Wildlife and S...
- Preferred Season: Spring
- Traveling With: solo
- Trip Duration (days): 5

Top Picks

- Ranthambore National Park** (Rajasthan | North India) | Match: 0.781
- Khajuraho & Panna National Park** (Madhya Pradesh | Central India) | Match: 0.756
- Shimla** (Himachal Pradesh | North India) | Match: 0.694

All Results

Rank	Destination	State	Region	Match Score	Weather Score
1	Ranthambore National Park	Rajasthan	North India	0.781	0.780
2	Khajuraho & Panna National Park	Madhya Pradesh	Central India	0.756	0.750
3	Shimla	Himachal Pradesh	North India	0.694	0.620
4	Kathmandu Durbar Square	Nepal	South Asia	0.600	0.200
5	Great Horned Owl (White-tailed)	Spain	West India	0.587	1.100
6	Jaisalmer & Thar Desert Safari	Rajasthan	West India	0.676	0.660
7	Ranthambore National Park	Madhya Pradesh	Central India	0.677	0.750
8	Jaisalmer (Thar Desert & Fort Taran)	Rajasthan	West India	0.676	0.660
9	Agra	Uttar Pradesh	North India	0.675	0.660
10	Vishakhapatnam	Rajasthan	West India	0.675	0.750

Travel Planner

Get essential first, then fine-tune in Advanced mode.

Experience: **Simple** | **Advanced**

Trip Preferences:

- Budget Level: mid
- Trip Intent: Adventure and Wildlife and S...
- Preferred Season: Spring
- Traveling With: solo
- Trip Duration (days): 5

Detailed Recommendations

#1 | Ranthambore National Park (Rajasthan) | Match: 0.781

Region: North India | District: Sawai Madhopur | Budget: ₹ 6,900 to ₹ 13,700/day

Why this destination fits your preferences:

- Budget range (Mid): ₹ 6,900 to ₹ 13,700 per day
- Season fit: acceptable for Spring, though not peak period
- Strong trip-type match (similarity: 1.00)
- Duration is compatible (5 days, ideal: 3 days)
- Popular destination (popularity: 9/10)
- Strong safety score (9/10)
- Unique experience: iconic tiger reserve where big cats are often seen near ancient ruins and lakes, part of popular Golden Triangle extensions [web:29][web:31]

#2 | Khajuraho & Panna National Park (Madhya Pradesh) | Match: 0.756

#3 | Shimla (Himachal Pradesh) | Match: 0.694

#4 | Kathmandu Durbar Square (Nepal) | Match: 0.600

Travel Planner

Get essential first, then fine-tune in Advanced mode.

Experience: **Simple** | **Advanced**

Trip Preferences:

- Budget Level: mid
- Trip Intent: Adventure and Wildlife and S...
- Preferred Season: Spring
- Traveling With: solo
- Trip Duration (days): 5

Route Planner

Select destinations to create a route (in order):

Ranthambore N... | Khajuraho & Panna... | Shimla - Himachal...

Default route uses top 3 recommendations. Add or remove destinations as needed.

Total Distance: **1114.4 km** | Estimated Travel Time: **30h 38min**

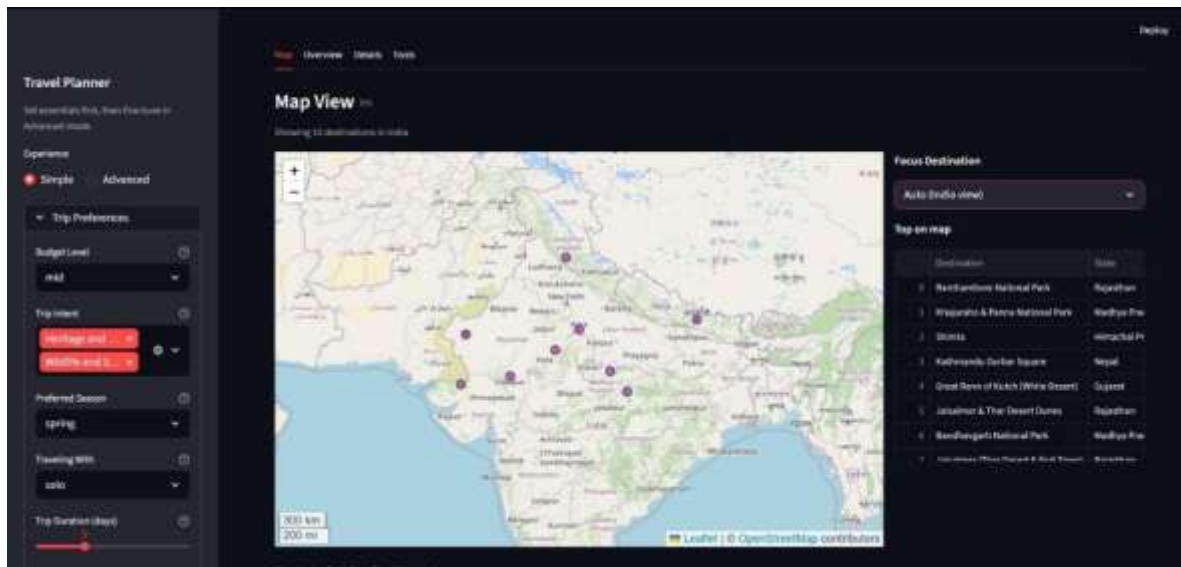
Route Waypoints:

- Ranthambore National Park - Rajasthan (Rajasthan)
- Khajuraho & Panna National Park - Madhya Pradesh (Madhya Pradesh) - 368 km from previous
- Shimla - Himachal Pradesh (Himachal Pradesh) - 746 km from previous

Distance from your location:

- Ranthambore National Park - Rajasthan: 296.8 km from your location
- Khajuraho & Panna National Park - Madhya Pradesh: 499.1 km from your location
- Shimla - Himachal Pradesh: 317.5 km from your location

Export Recommendations



III. CONCLUSION

The Smarttour system developed in this project is a real-world application of machine learning in tourism. The system uses user demographic and preference parameters—such as age, budget, travel season, activity interests, and trip duration—to generate personalized tourist attraction recommendations with the help of models like Gradient Boosting, Random Forests, and Learning-to-Rank algorithms.

During the development lifecycle, the system went through multiple phases—data preprocessing, feature engineering, model training, evaluation, and deployment—all aimed at producing accurate, interpretable, and actionable recommendations.

The integration with Streamlit enabled the development of a simple and interactive interface, allowing users to:

- Enter personal preferences and travel constraints via a user-friendly form.
- Receive real-time attraction recommendations tailored to their profile.
- View explanations for recommendations through SHAP or feature importance visualizations.
- Download recommendation reports for offline planning and reference.

The resulting recommendation system demonstrated strong performance, with high precision in ranking relevant tourist attractions and effectively capturing individual user preferences. While Smarttour is not a replacement for personal judgment or travel consultancy, it serves as a decision-support tool, facilitating efficient trip planning and personalized travel experiences.



IV. REFERENCES

- [1] J. Smith and L. Wang, "Personalized Tourist Recommendation using ML," Collaborative Filtering, Content-Based Filtering, Cold-start problem, sparse data, 2018.
- [2] A. Kumar and R. Singh, "Hybrid Travel Recommendation System," Decision Trees + Collaborative Filtering, Computationally expensive, scalability issues, 2019.
- [3] P. Chen and M. Li, "ML-Based Tourist Destination Prediction," Random Forest, SVM, Limited context awareness, lacks explainability, 2020.
- [4] S. Patel and K. Rao, "Explainable Travel Recommender," Gradient Boosting, SHAP for interpretability, Longer runtime, requires feature engineering, 2021.
- [5] R. Johnson and T. Kim, "Personalized Trip Planning System," Neural Networks, Content-Based Filtering, Needs large dataset, low transparency, 2021.
- [6] D. Singh and V. Mehta, "Context-Aware Tourist Recommendation," Random Forest + Contextual Features, Sensitive to feature selection, complexity, 2022.
- [7] L. Chen and S. Verma, "Smart Tourist Recommendation Platform," Hybrid ML (Collaborative + Content-Based), Cold-start, handling imbalanced preferences, 2022.
- [8] M. Gupta and A. Roy, "AI-Powered Travel Suggestion System," Deep Neural Networks, Requires extensive training data, low interpretability, 2023.
- [9] K. Das and P. Sinha, "Explainable ML for Tourism," Random Forest, SHAP, Limited real-time deployment, 2023.