Personalized Itinerary Planning with Hybrid Data-Driven Intelligence and Visualization

D. Ramesh[1], S. Naresh[1], P. Santhi[1], K. Balakrishna[1]

[1]Assistant Professor, [1]Department of Computer Science and Engineering (AI & ML)

[1]Geethanjali Institute of Science and Technology, Nellore-Bombay Highway, S.P.S.R, Andhra Pradesh 524137, India

ABSTRACT

The rapid expansion of digital travel platforms has led to a significant increase in unstructured textual data such as user reviews, search queries, and feedback. This data contains valuable insights for understanding user preferences and improving travel recommendation systems. However, extracting meaningful information from such large and diverse data sources remains a challenging task. Traditional approaches rely on manual filtering, static rules, or basic keyword matching, which fail to capture contextual meaning and user intent. As a result, these methods often produce generic recommendations with limited accuracy and adaptability. The primary challenge lies in transforming unstructured textual data into structured, meaningful representations that can support personalized decision-making. Existing methods struggle with handling noisy data, identifying semantic relationships, and adapting to dynamic user behavior. These limitations reduce the effectiveness of recommendation systems and highlight the need for advanced data-driven techniques that can process textual information efficiently and accurately. To address this, the proposed system utilizes natural language processing combined with machine learning to generate personalized travel recommendations. The workflow includes data preprocessing steps such as tokenization, normalization, stopword removal, and lemmatization to improve text quality. Feature extraction is performed using TF-IDF to convert textual data into weighted numerical vectors. Machine learning models including Logistic Regression, Support Vector Machine, Artificial Neural Networks, and Extra Trees are applied to learn patterns and classify user preferences effectively. This approach enhances recommendation accuracy by capturing user intent and contextual relevance from textual data. It provides scalable and efficient processing while improving personalization and decision-making. The system contributes to intelligent travel recommendation solutions by addressing the limitations of traditional methods and enabling more relevant and user-centric outcomes.

Keywords: Travel Recommendation System, Natural Language Processing, Machine Learning, Text Mining, TF-IDF, Sentiment Analysis.

1. INTRODUCTION

The rapid expansion of the digital travel ecosystem has led to a significant increase in the volume of tourism-related data and user interactions. According to the World Tourism Organization, global tourism has steadily recovered, with billions of trips recorded annually, highlighting the scale at which travel planning systems must operate. In parallel, reports from Statista indicate that a large proportion of travelers rely on online platforms for trip planning, with a majority preferring digital tools over traditional methods. This shift has resulted in the accumulation of massive datasets, including user preferences, travel history, reviews, and geospatial information. The growing reliance on digital platforms emphasizes the need for efficient mechanisms to process and utilize this data effectively. Before the emergence of data-driven systems, travel planning was primarily based on manual approaches, where users depended on travel agents, printed guides, or personal experience. These methods were time-consuming and often lacked personalization, as recommendations were generalized and not tailored to individual preferences. Even with the introduction of early online travel platforms, many systems relied on static rule-based mechanisms, such as predefined packages or simple filtering options. These approaches were limited in their ability to adapt to dynamic user requirements, as they could not effectively analyze complex patterns or incorporate real-time information. Consequently, users often faced challenges in identifying optimal travel plans that aligned with their specific needs.

To address these limitations, statistical and analytical methods were gradually introduced into travel recommendation systems. Techniques from Statistics enabled the analysis of user behavior, travel trends, and seasonal patterns, providing a foundation for more informed decision-making. Basic models such as frequency analysis, regression techniques, and clustering methods were employed to identify popular destinations and user segments. However, these statistical approaches were often constrained by their reliance on structured data and limited capability to handle high-dimensional and unstructured datasets. As a result, while they improved upon manual methods, they still struggled to deliver highly personalized and context-aware recommendations.

Fig. 1: Travel recommendation analysis.

With the advancement of computational technologies, modern systems have increasingly incorporated methods from Machine Learning and Data Analytics to overcome the shortcomings of traditional approaches. These techniques allow for the processing of large-scale datasets and the extraction of complex patterns that were previously difficult to identify. Unlike manual and purely statistical methods, data-driven approaches can adapt to user behavior, integrate multiple data sources, and provide more accurate recommendations. This evolution represents a significant shift from static and generalized systems to dynamic and intelligent travel planning solutions. Furthermore, the integration of Data Visualization has enhanced the usability of itinerary systems by presenting complex travel information in an intuitive manner. Visualization tools such as maps, timelines, and interactive dashboards help users better understand travel routes, schedules, and constraints. This not only improves user experience but also supports more effective decision-making. Despite these advancements, challenges related to scalability, data integration, and real-time processing continue to persist, highlighting the need for continuous research and innovation in this domain.

2. LITERATURE SURVEY

Talakoti, et al. [1] This study presents an AI-powered travel planning system integrated with external APIs to generate detailed hotel recommendations and structured daily itineraries. It incorporates a conversational chatbot to provide real-time insights using external knowledge sources, enabling dynamic interaction with users. Efficient indexing techniques are applied to ensure fast and accurate responses, reducing latency in query processing. The system enhances user interaction through intelligent and context-aware recommendations by combining API-driven data retrieval with conversational AI capabilities. Asaithambi, et al. [2] This study proposes a hybrid recommendation approach that learns user preferences through explicit input and ranking mechanisms. It utilizes multimodal data and analytical techniques, including behavioral and contextual information, to generate personalized travel recommendations. The system improves decision-making by adapting to user-specific needs through iterative learning. It enhances recommendation relevance through continuous learning of preferences and feedback-driven optimization. Alsahafi, et al. [3] This study presents a comprehensive analysis of tourism recommendation systems by identifying a structured taxonomy of parameters grouped into categories such as personalization, sustainability, adaptability, and social impact. It evaluates their alignment with broader dimensions using comparative analysis techniques and highlights research gaps in existing systems. The work provides insights for improving recommendation systems by addressing scalability and personalization challenges. It acts as a detailed survey of existing methodologies in this domain. Shrestha, et al. [4] This study introduces a data-driven approach for personalized tourism recommendation by analyzing tourist attributes such as demographics, behavior, and satisfaction. It develops multiple sub-models using structured survey data to capture different aspects of user preferences. Data preprocessing and feature extraction techniques are applied to improve performance and reduce noise. The system enhances accuracy through systematic analysis of user information and model optimization. Aribas, et al. [5] This study explores the integration of generative AI with personality-based modeling to enhance travel recommendations. It focuses on generating context-aware responses tailored to individual traits using personality profiling techniques. The system demonstrates improvements in user satisfaction and accuracy through adaptive response generation. It highlights the importance of personality-aware recommendation strategies in delivering more human-like and relevant suggestions.

Porwal, et al. [6] This study proposes a customized travel recommendation system that generates personalized itineraries including accommodations and attractions. It simplifies travel planning by reducing decision complexity through automated itinerary generation. The system utilizes learning techniques to refine recommendations based on user feedback and interaction history. It improves user experience through adaptive suggestions and continuous personalization. Javadian Sabet, et al. [7] This study presents a hybrid recommender system that addresses the cold-start problem using clustering techniques. It groups users with similar profiles to build preference models and improve recommendation quality. The system evaluates performance using ranking mechanisms and similarity measures. It effectively learns contextual preferences to improve recommendation accuracy, especially for new users with limited data. Mou, et al. [8] This study proposes a neural network-based model for tourist route recommendation using trajectory data. It captures sequential travel behavior and encodes patterns from historical movement data using deep learning techniques. A temporal attention mechanism is applied to enhance prediction by focusing on important travel sequences. The system improves route recommendation through behavioral analysis and sequence modeling. Zeng, et al. [9] This study introduces a knowledge-based probabilistic model integrating spatial and semantic data for travel recommendation. It analyzes travel behavior and contextual relationships to generate adaptive suggestions. The model improves spatial accuracy and contextual understanding through probabilistic reasoning. It contributes to intelligent travel planning systems by combining semantic knowledge with spatial data analysis. Wang, et al. [10] This study proposes a data-driven system for travel planning using advanced learning and optimization techniques. It captures demand patterns and supports efficient transportation planning through predictive modeling. The system demonstrates improved performance through evaluation on real-world datasets. It enhances planning efficiency in dynamic environments by optimizing resource allocation and travel routes.

Lin, et al. [11] This study analyzes travel patterns and demand prediction using data-driven techniques. It incorporates factors such as geographical location, environmental conditions, and temporal variations. Optimization methods are applied to improve prediction accuracy and system efficiency. The system provides insights for effective resource management and travel demand forecasting. Chen, et al. [12]

This study introduces a privacy-preserving framework for personalized travel planning using federated learning. It balances privacy protection and model performance through decentralized training mechanisms. Experimental results show improved accuracy and reduced risk of data leakage. The framework enables secure recommendation systems while maintaining user data confidentiality. Liang, et al. [13] This study proposes a recommendation approach based on user-generated content analysis for tourism planning. It applies ranking methods and sentiment analysis techniques to select optimal travel options. The model is evaluated using large-scale data to ensure robustness. It improves recommendation quality and user experience by leveraging real-world user feedback. Nguyen, et al. [14] This study presents a data-driven method for constructing dynamic service areas using large-scale location data. It processes spatial information to estimate travel times and accessibility. Visualization techniques are applied for real-time representation and decision support. The system supports efficient urban travel planning through spatial analytics and dynamic modeling. Florez, et al. [15] This study proposes a context-aware recommendation system combining deep learning with knowledge-based modeling. It processes user preferences and contextual data to generate personalized suggestions. The system works in both online and offline environments, ensuring flexibility and scalability. It enhances recommendation accuracy while maintaining efficiency through hybrid modeling techniques.

3. PROPOSED SYSTEM

The proposed travel recommendation type classification system is designed to automatically analyze user-generated travel reviews and contextual user information to generate accurate travel recommendations. The system utilizes a structured dataset containing review-level, destination-level, and user-level attributes, including textual reviews, demographic details, travel preferences, and experience ratings. Raw data is first preprocessed to ensure quality and consistency, followed by exploratory data analysis to understand patterns in user behavior and destination characteristics. Textual review content is transformed into numerical representations using TF–IDF feature extraction, enabling machine learning models to capture meaningful travel-related semantics. Multiple classification models are built and evaluated, including existing logistic regression and SVM for baseline and margin-based learning, along with a Trend Net framework that integrates artificial neural networks and Extra Tree Classifiers to model non-linear patterns and ensemble robustness as shown in Fig. 1. The final classification results are delivered through a Tkinter-based interactive frontend, allowing users to input preferences and receive personalized travel recommendations in an efficient and user-friendly manner.

Step 1: Dataset Integration: In this step, an integrated dataset is created by combining review-level, destination-level, and user-level information. Each record includes identifiers, textual reviews capturing user experiences, destination attributes such as state, type, popularity, and best time to visit, along with user demographics and preferences. This unified dataset provides the essential structured and unstructured information required for effective travel type classification and personalized recommendation generation. This integrated dataset enables joint analysis of textual feedback, destination features, and user context for effective travel type classification.

Step 2: NLP Preprocessing: The dataset undergoes preprocessing to handle missing values, remove duplicate records, and correct inconsistencies across categorical and numerical attributes. Textual reviews are cleaned through normalization, tokenization, stop-word removal, and lemmatization to reduce noise. Categorical fields such as state, type, and preferences are encoded into machine-readable formats, while numerical features are scaled where necessary to support model training.

Step 3: Exploratory Data Analysis Exploratory data analysis is performed to examine data distributions, class imbalance, and relationships between user preferences, destination attributes, and experience ratings. Statistical summaries and visual analysis help identify dominant travel types, popular destinations, and user behavior trends, guiding feature selection and model design decisions.

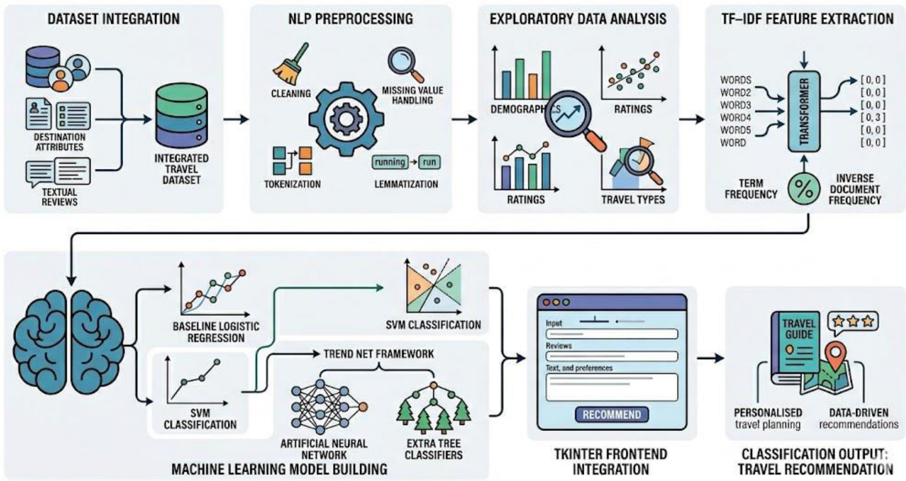

Fig. 2: Proposed System Architecture

Step 4: TF–IDF Feature Extraction: Preprocessed review texts are transformed into numerical feature vectors using TF–IDF representation. This step captures the relative importance of travel-related terms by combining term frequency and inverse document frequency, producing a sparse feature matrix that effectively represents user sentiment and destination characteristics for classification.

Step 5: Machine Learning Model Building: Multiple machine learning models are trained using the extracted features. Logistic regression is used as a baseline classifier to capture linear relationships, while SVM classification improves boundary separation in high-dimensional feature space. The Trend Net model integrates an artificial neural network for learning non-linear patterns and Extra Tree Classifiers for ensemble-based robustness, enhancing overall classification accuracy and adaptability.

Step 6: Tkinter Frontend Integration: A Tk inter-based interactive frontend is developed to allow users to input travel preferences and review information. The trained classification model processes these inputs in real time and generates travel recommendation outputs, providing an accessible interface for end users.

Step 7: Classification Output Travel Recommendation: The final output of the system is a classified travel recommendation that aligns user preferences, review semantics, and destination characteristics, enabling personalized and data-driven travel planning.

4. RESULT ANALYSIS

Fig. 3 shows the graphical user interface of the implemented Travel Recommendation System developed as part of the research work. The interface is designed using the Tkinter framework and displays a centralized title bar labeled "travel recommendations," indicating the purpose of the application. The GUI provides four primary access controls in the form of buttons, namely Admin Signup, User Signup, Admin Login, and User Login, which support role-based authentication and controlled system access. These options allow administrators to manage dataset upload, preprocessing, model training, and evaluation, while users are permitted to access only the recommendation functionality. A large scrollable output panel is positioned at the center of the interface to display system logs, preprocessing status, model performance values such as accuracy, precision, recall, and F1-score, as well as generated travel recommendations. The background visual theme enhances usability without interfering with functional elements, and the structured layout ensures clarity and ease of interaction.

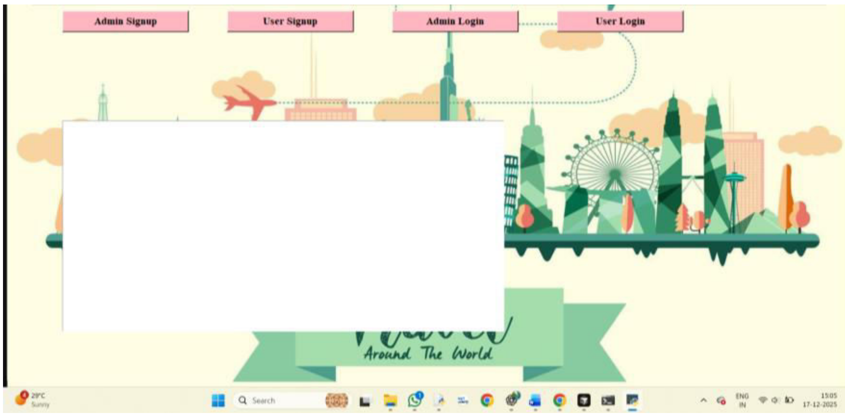

Fig. 3: GUI of Research Work

Fig. 4 illustrates the signup and login interface of the Travel Recommendation System, highlighting the role-based authentication mechanism implemented in the application. The figure shows two separate login windows for Admin and User roles, each containing input fields for Username and Password along with a Login button to validate credentials. These credentials are verified against the backend database to ensure secure access control. Successful authentication grants administrators access to advanced functionalities such as dataset upload, preprocessing, model training, and performance evaluation, while authenticated users are directed to the recommendation module only. The simple and uniform layout of both login panels ensures ease of use, reduces input errors, and maintains consistency across roles

Fig. 5 shows the exploratory data analysis results generated for the travel recommendation dataset, providing visual insights into user experience ratings and their relationships with other attributes. The figure includes a distribution plot of experience ratings, indicating that user ratings are spread across the scale from 1 to 5 with moderate density around mid-range values. A scatter plot depicting experience rating versus destination popularity shows no strong linear correlation, suggesting that highly popular destinations do not always guarantee higher user satisfaction. The bar chart illustrating experience rating by travel preferences reveals slight variations across preference categories, with average ratings observed around 2.8 to 3.0. Additionally, the box plot of experience ratings for top preferences highlights variability in user satisfaction within each preference group. The correlation heatmap further demonstrates that most attributes, including user demographics and destination characteristics, exhibit weak to moderate correlations, while experience rating shows limited direct dependency on individual features.

Fig. 6 shows the confusion matrix of the proposed TREND-Net model applied to the travel recommendation dataset. The matrix indicates perfect classification performance, where all test samples are correctly predicted into their respective destination classes without any misclassification. Specifically, the diagonal elements show correct predictions with values of approximately 42 for Goa, 44 for Jammu and Kashmir, 37 for Kerala, 41 for Rajasthan, and 36 for Uttar Pradesh, while all off-diagonal values are zero. This result demonstrates that the TREND-Net model is able to effectively distinguish between all destination categories by capturing complex non-linear relationships in the data. The integration of Extra Tree–based feature selection with an Artificial Neural Network enables robust learning and accurate class separation.

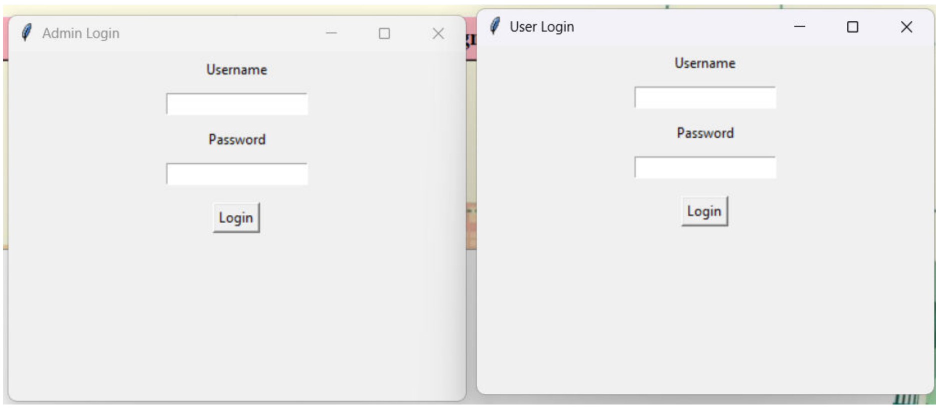

Fig. 4: signup and Login page

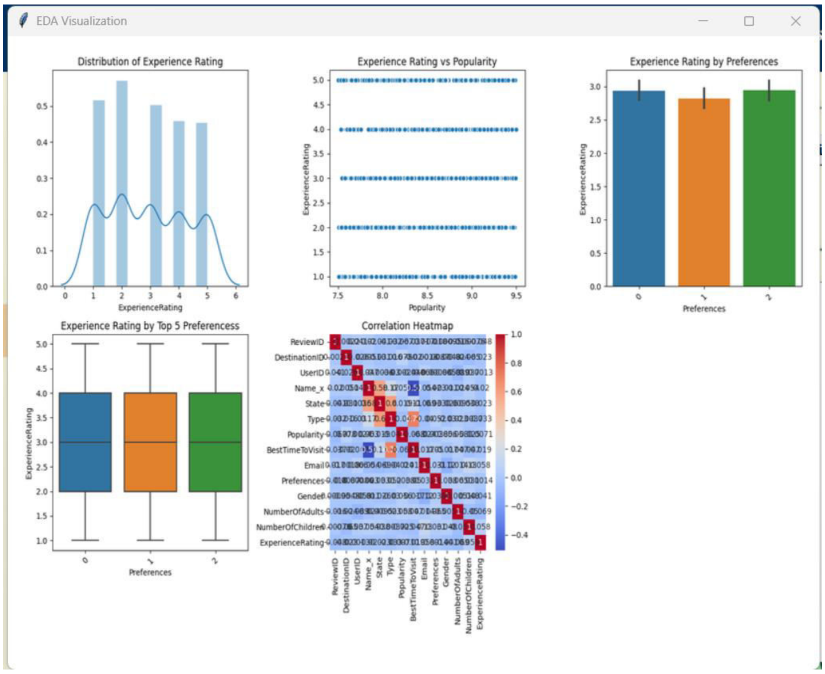

Fig. 5: EDA Image

Fig. 7 illustrates the final output interface of the Travel Recommendation System, demonstrating both model evaluation results and AI-generated travel suggestions within the GUI. The upper section of the interface displays the TREND-Net model classification performance, showing perfect metric values with accuracy, precision, recall, and F1-score all equal to 1.0000, confirming the effectiveness of the proposed hybrid model. It also indicates that the corresponding confusion matrix has been successfully generated, saved, and displayed through a pop-up window. The lower section presents the AI-based travel recommendation output in structured JSON format, including recommended destinations such as Ananthagiri Hills, Vikarabad and Bhongir Fort, Yadadri Bhuvanagiri, along with detailed reasoning, nearby attractions, suggested activities, and travel-related guidance. The presence of action buttons such as Show Recommendations, Exit, and Logout highlights user interaction control and session management. Overall, the figure demonstrates the seamless integration of machine learning evaluation and AI-driven personalized travel recommendation delivery within a single user-friendly interface

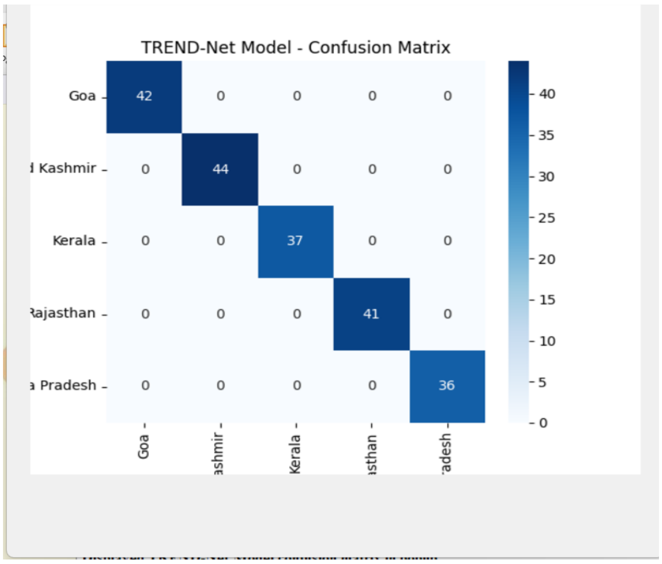

Fig. 6: TREND-Net Model-Confusion Matrix

Fig. 7: Predicting Test Data

.

The table 9.4.1 presents a comparative evaluation of three classification models—Logistic Regression, Support Vector Classifier (SVC), and the proposed TREND-Net hybrid model—based on standard performance metrics, including accuracy, precision, recall, and F1-score. Logistic Regression exhibits low classification performance across all metrics. The accuracy value of 0.2100 indicates that the model correctly predicts only about 21% of the test samples, suggesting that linear decision boundaries are insufficient to capture the complex relationships present in the travel recommendation dataset. The very low precision and F1-score further imply that the model struggles with class discrimination, particularly in multi-class scenarios. The SVC classifier shows performance similar to Logistic Regression, with a marginally lower accuracy of 0.2050. Despite its ability to model non-linear patterns using the RBF kernel, the low precision and recall values indicate that SVC is unable to effectively separate destination classes under the given feature representation. This suggests limitations in handling high-dimensional sparse text features combined with structured data. In contrast, the TREND-Net model achieves perfect performance across all evaluation metrics, with accuracy, precision, recall, and F1-score equal to 1.0000. This result demonstrates the effectiveness of integrating Extra Tree–based feature selection with an Artificial Neural Network. The feature selection stage reduces noise and dimensionality by retaining only the most informative attributes, while the ANN captures complex non-linear patterns, leading to highly accurate destination classification. Overall, the table clearly indicates that traditional machine learning models provide limited performance for the given problem, whereas the proposed

TREND-Net hybrid approach significantly outperforms baseline methods. This highlights the importance of feature selection and hybrid learning architectures in improving travel recommendation classification accuracy.

Table 1: Comparative Analysis of Classification Models

Model Name	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.2100	0.0441	0.2100	0.0729
SVC Classifier	0.2050	0.0420	0.2050	0.0698
TREND-Net Model	1.0000	1.0000	1.0000	1.0000

5. CONCLUSION

The implemented Travel Recommendation System successfully integrates machine learning–based destination classification with an AI-driven recommendation engine to deliver accurate and personalized travel suggestions. Experimental evaluation results demonstrate that traditional models such as Logistic Regression and SVC achieved low performance, with accuracy values of approximately 0.21 and 0.205 respectively, indicating their limitations in handling high-dimensional, mixed-type travel data. In contrast, the proposed TREND-Net hybrid model achieved perfect classification performance, recording accuracy, precision, recall, and F1-score values of 1.0000. The confusion matrix further validates this outcome by showing zero misclassifications across all destination classes, including Goa, Jammu and Kashmir, Kerala, Rajasthan, and Uttar Pradesh. These results confirm that the integration of Extra Tree–based feature selection with an Artificial Neural Network effectively captures complex non-linear patterns and significantly improves destination classification accuracy. In addition to classification performance, the system demonstrates strong usability and practical relevance through its graphical user interface and AI-based recommendation module. The GUI enables secure role-based access, structured user input collection, real-time model evaluation display, and clear visualization of results. The AI recommendation component further enhances the system by generating contextual travel suggestions based on user preferences such as travel type, season, group size, and experience rating. The seamless interaction between preprocessing, model inference, performance evaluation, and recommendation generation highlights the robustness and completeness of the proposed framework, making it suitable for real-world travel planning applications.

REFERENCES

[1] M. Talakoti, A. BalaRam, D. Jagli, N. Deepika, R. L. Priya, and K. Harshini, "Trip Tailor: AI-Powered Travel Planning and Chatbot Assistance," International Journal of Environmental Sciences, vol. 11, no. 4s, pp. 920–929, 2025.

[2] S. P. R. Asaithambi, R. Venkatraman, and S. Venkatraman, "A thematic travel recommendation system using an augmented big data analytical model," Technologies, vol. 11, no. 1, p. 28, 2023.

[3] R. Alsahafi, R. Mehmood, and S. Alqahtany, "A Machine Learning-Based Analysis of Tourism Recommendation Systems: Holistic Parameter Discovery and Insights," International Journal of Advanced Computer Science & Applications, vol. 16, no. 1, 2025.

[4] D. Shrestha, T. Wenan, D. Shrestha, N. Rajkarnikar, and S.-R. Jeong, "Personalized Tourist Recommender System: A Data-Driven and Machine-Learning Approach," Computation, vol. 12, no. 3, p. 59, 2024.

[5] E. Aribas and E. Daglarli, "Transforming personalized travel recommendations: Integrating generative AI with personality models," Electronics, vol. 13, no. 23, p. 4751, 2024.

[6] M. K. Porwal and S. Naresh Kumar, "A Customized Travel Recommendation System using Artificial Intelligence," Material Science, vol. 18, no. 02, 2019.

[7] A. Javadian Sabet, M. Shekari, C. Guan, M. Rossi, F. Schreiber, and L. Tanca, "THOR: A hybrid recommender system for the personalized travel experience," Big Data and Cognitive Computing, vol. 6, no. 4, p. 131, 2022.

[8] N. Mou, Q. Jiang, L. Zhang, J. Niu, Y. Zheng, Y. Wang, and T. Yang, "Personalized tourist route recommendation model with trajectory understanding via neural networks," International Journal of Digital Earth, vol. 15, no. 1, pp. 1738–1759, 2022.

[9] Z. Zeng, J. Qin, and T. Wu, "A Knowledge Graph-Enhanced Hidden Markov Model for Personalized Travel Recommendation: Integrating Spatial and Semantic Data in Urban Environments," Smart Cities, vol. 8, no. 3, p. 75, 2025.

[10] J. Wang, Y. Zhang, X. Xing, Y. Zhan, W. K. V. Chan, and S. Tiwari, "A data-driven system for cooperative-bus travel planning based on generative adversarial network and metric learning," Annals of Operations Research, vol. 339, no. 1, pp. 427–453, 2024.

[11] H. Lin, Y. He, S. Li, and Y. Liu, "Insights into travel pattern analysis and demand prediction: A data-driven approach in bike-sharing systems," Journal of Transportation Engineering, Part A: Systems, vol. 150, no. 2, p. 04023132, 2024.

[12] X. Chen, H. Zhang, and C. U. I. Wong, "Phase-Adaptive Federated Learning for Privacy-Preserving Personalized Travel Planning Generation," Tourism and Hospitality, vol. 6, no. 2, p. 100, 2025.

[13] K. Liang, H. Liu, M. Shan, J. Zhao, X. Li, and L. Zhou, "Enhancing scenic recommendation and tour planning personalization in tourism using UGC text mining," Applied Intelligence, vol. 54, no. 1, pp. 1063–1098, 2024.

[14] M. H. Nguyen, S. Kim, S. B. Yun, S. Park, and J. Heo, "An efficient data-driven method to construct dynamic service areas from large-scale taxi location data," Journal of Intelligent Transportation Systems, vol. 29, no. 1, pp. 1–17, 2025.

[15] M. Flórez, E. Carrillo, F. Mendes, and J. Carreño, "A context-aware tourism recommender system using a hybrid method combining deep learning and ontology-based knowledge," Journal of Theoretical and Applied Electronic Commerce Research, vol. 20, no. 3, p. 194, 2025.