



GlobeTrail

Submitted in partial fulfillment of the requirement
of the degree of the course **Innovative Product
Development-IV**

Year 3, Sem VI, Bachelor of Technology in
Department of Computer Engineering

By

Arham Shah	60004220110
Granth Jain	60004220193
Jainam Shah	60004220085
Divij Mehta	60004220053

Under the guidance of

Prof. Lynette D'mello



CERTIFICATE

This is to certify that the project entitled "**GlobeTrail**" is a bonafide work of
"Arham Shah, Granth Jain, Jainam Shah, Divij Mehta" (**60004220110, 60004220193, 60004220085, 60004220053**) submitted as a project work for the
course **Innovative Product Development-III, Year 3, Semester V, TY B. Tech**
Computer Engineering

Prof. Lynette D'mello
Internal Guide

Dr. Meera Narvekar

Head of the Department

Dr. Hari Vasudevan

Principal



IPD Project Report Approval for BTech Semester V

This project report entitled “**GlobeTrail**” by “Arham Shah, Granth Jain, Jainam Shah, Divij Mehta” is approved for the **Innovative Product Development-III, Year 3, Semester V, TY B. Tech Computer Engineering**

Examiners

1. _____

2. _____

Date:

Place:



DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included. We have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Arham Shah 60004220110

Granth Jain 600042200193

Jainam Shah 60004220085

Divij Mehta 60004220053

Date:



Acknowledgement

We are deeply grateful to everyone who contributed to the success of our research project.

First and foremost, we would like to extend our heartfelt thanks to Prof. Lynette D'mello, faculty of the Computer Engineering Department, and our project guide. Her mentorship, expertise, and insightful feedback were instrumental in shaping the direction of our work. We sincerely appreciate the time and dedication she invested to help elevate the quality of this project.

We also wish to express our appreciation to the faculty members and our Head of Department Dr. Meera Narvekar of the Computer Engineering Department at Dwarkadas J Sanghvi College of Engineering. Their continuous encouragement and commitment to fostering an environment of learning and innovation have been invaluable to our research journey.

Additionally, we would like to thank our Principal, Dr. Hari Vasudevan, for making the college library resources accessible to us, which greatly supported our research efforts by providing essential reference materials.

Name of the Students:

Arham Shah

Granth Jain

Jainam Shah

Divij Mehta



Table of Contents

Chapter	Contents	PageNo.
1	Introduction	
2	Need for the product	
	2.1 Need for the product	
	2.2 Drawbacks of Existing Platforms	
	2.3 Applications Of GlobeTrail	
3	Literature Review/Survey of Existing Systems	
	3.1 Literature Survey	
	3.2 Field Survey	
	3.3 Analysis of Literature Survey	
	3.4 Outcome of survey	
4	Problem Formulation	
	4.1 Problem Formulation	
	4.2 Product Objectives	
	4.3 Novelty	
	4.4 Scope of the project	
5	Proposed Design	
	5.1 Proposed Model	
	5.2 Database Design	
	5.3 Use Cases	
6	Implementation	
	6.1 GUI Design	
	6.2 Modules Implementation	
7	Experimentation And Results	
	7.1 Datasets	
	7.2 Parameter tuning experiments	
	7.3 Results	
8	Conclusion	
9	References	



List of Figures

Fig. No.	Figure Caption	Page No.
5.1	Proposed Design	22
6.11	HomePage of GlobeTrails Website	24
6.12	About Us of GlobeTrails Website	24
6.13	User Credentials Input UI of GlobeTrails Website	25
6.14	Footer of GlobeTrails Website	25
6.15	Comparison Table for GlobeTrails Chatbot with other Travel Itinerary Chatbots	26
7.31	K-Means clustering for adventure and scenic destinations in Manali based on weights	28
7.41	Output recommendation sample	30
7.42	Output recommendation Mappings	31



List of Tables

Table No.	Table Title	Page No.
3.3	Analysis of Literature Survey	15
7.11	Sample tuples in Scrapped Dataset	28



1. Introduction

1.1 Background

Travel has long been one of humanity's most cherished experiences, offering opportunities for exploration, connection, and relaxation. However, the process of planning a trip often presents significant challenges, particularly in today's fast-paced, digitally driven world. Travelers face an overwhelming array of options when it comes to accommodations, transportation, attractions, and activities. Navigating these choices while accounting for budget constraints, personal preferences, and time limitations can turn what should be an exciting process into a daunting task. Moreover, unforeseen variables such as weather changes, transport delays, or crowded destinations further complicate the travel experience.

The rise of digital technologies has revolutionized many aspects of travel planning, yet traditional approaches remain fragmented. Existing travel platforms often provide static recommendations that fail to account for real-time data, personalized user preferences, or dynamic changes during the journey. Travelers are left juggling multiple tools—booking engines, map services, review platforms—leading to inefficiencies and an incomplete sense of control. This gap highlights the need for an intelligent, integrated solution that can seamlessly handle the complexities of modern travel.

1.2 Motivation

The motivation for developing an AI-powered travel planning platform stems from the pressing need to address these inefficiencies while enhancing the overall travel experience. With the advent of machine learning, natural language processing (NLP), and real-time data analytics, there is an unprecedented opportunity to leverage technology to simplify and personalize trip planning. By integrating these advanced technologies, the platform can deliver value across multiple dimensions:

- Personalization at Scale:** Using machine learning algorithms, the platform can analyze traveler profiles, preferences, and behaviors to offer curated recommendations tailored to individual needs. This enables a shift from generic suggestions to deeply personalized itineraries that resonate with users.
- Real-Time Adaptability:** With access to live data streams, such as weather updates, traffic conditions, and event changes, the platform can dynamically adjust plans. This ensures that travelers can navigate disruptions effortlessly, minimizing stress and maximizing enjoyment.
- Seamless Integration:** Modern travelers expect convenience. By integrating APIs for bookings, maps, and local services, the platform provides a unified experience, eliminating the need to toggle between multiple tools.
- Data-Driven Insights:** Predictive analytics and sentiment analysis allow the platform to identify trends and anticipate traveler needs. For instance, it can suggest emerging destinations or recommend less crowded alternatives to popular spots, enhancing the experience while avoiding common pitfalls.
- Empowering Travelers:** Beyond simplifying logistics, the platform empowers users with insider tips, local knowledge, and optimized routes. This not only enriches the travel experience but also builds trust and engagement with the platform.



2. Literature Review/Survey of Existing Systems

2.1 Field Survey

A field survey for **GlobeTrail**, a travel platform offering curated itineraries and integrated booking, would aim to validate trends and gather in-depth user feedback using insights from multiple online data sources. The survey would focus on understanding key preferences, challenges, and expectations of travelers to align GlobeTrail's offerings with user needs and industry trends. The data for the survey would be informed by publicly available research from platforms like **TripAdvisor**, **Skyscanner**, and **Booking.com**, which provide comprehensive insights into user behavior, feature preferences, and pain points. Additionally, travel forums like **Reddit travel communities** and **Facebook travel groups** offer anecdotal & community-driven insights into what users expect from travel platforms. Finally, benchmarking reports from trusted research firms such as **Statista**, **McKinsey**, and **Deloitte** would offer high-level insights into emerging trends, such as the growing demand for personalization, eco-friendly travel options, and real-time updates.

The survey would specifically examine user priorities by asking about their travel behaviors, such as preferred travel experiences (adventure, cultural, or leisure), and evaluate the importance of features like AI-driven personalization, seamless booking for flights and activities, and real-time updates. The survey would also explore common pain points users face with current platforms, such as the inability to personalize itineraries, fragmented booking experiences across multiple sites, and the lack of live updates on traffic, weather, or POI availability. By asking users to rank these challenges and desired features, the survey would provide GlobeTrail with actionable insights into how to address these gaps.

Methodology

To gather diverse and comprehensive feedback, the survey combines primary and secondary data collection methods:

- Primary Sources:
 - Online Surveys: Distributed via platforms like Google Forms, targeting travel enthusiasts on social media and community groups.
 - In-Person Surveys: Conducted at travel expos and tourist destinations to collect input from frequent travelers.
 - Interactive Polls: Shared on platforms like Instagram and Twitter for quick insights from active users.



- Secondary Sources:

- Public Research: Data from platforms like TripAdvisor and Skyscanner on user behaviors and trends.
- Travel Forums: Insights from discussions on Lonely Planet and Reddit.
- Industry Reports: Market trends and insights from sources like McKinsey and Statista.

The survey targets frequent, occasional, and first-time travelers, aiming for 200–300 responses to ensure reliable and diverse feedback.

Expected Outcomes

- User Insights: Key features users prioritize, such as interactive planners and real-time updates.
- Pain Points: Challenges like the need for integrated platforms and outdated data.
- Feature Prioritization: Focus areas like AI-powered recommendations and itinerary customization.
- Market Positioning: Opportunities to differentiate from competitors by offering a seamless and unified platform.

2.2 Literature Survey

This section provides an overview of the current landscape surrounding **Travel Itinerary Planning, Real-Time Recommendation Systems, and Comprehensive Travel Platforms**, with a focus on advancements and challenges relevant to **GlobeTrail**. The aim is to examine progress in these areas, highlight key developments, and identify gaps that **GlobeTrail** can address as a full-fledged travel solution. Each subsection delves into different aspects of these categories, offering a deeper understanding of the current state and future opportunities.

Travel itinerary planning has significantly evolved with the adoption of AI and data-driven techniques, enabling platforms to provide tailored recommendations based on user preferences. However, gaps persist in **real-time adaptability**, such as incorporating weather updates, traffic conditions, and last-minute POI closures. Real-time recommendation systems, while effective, often fail to offer integrated booking and user feedback mechanisms, leading to fragmented user experiences.

Additionally, existing travel platforms are limited in combining **curated recommendations, interactive itinerary customization, and real-time ticket booking**. These challenges align with increasing user demands for automation, personalization, and sustainability. By addressing these gaps, **GlobeTrail** aims to provide a seamless, personalized, and comprehensive travel planning experience that caters to modern travelers' needs for flexibility and convenience.



● Analysis of Literature Survey

No.	Authors	Title	Source	Methodology
1.	Michel Dumas, Jacques Desrosiers, François Soumis	An Optimal Algorithm for the Traveling Salesman Problem with Time Windows	Technical paper (1995)	Developed an optimal algorithm that integrates time-window constraints into solving the Traveling Salesman Problem. It incorporates dynamic programming techniques.
2.	Syed Md. Mukit Rashid, Mohammed Eunus Ali, Muhammad Aamir Cheema	DeepAltTrip: Top-k Alternative Itineraries for Trip Recommendation	IEEE Transactions on Knowledge and Data Engineering (Preprint, 2021)	Proposed DeepAltTrip using Itinerary Net (ITRNet) with forward and backward LSTMs and graph autoencoders to recommend diverse, popular itineraries
3.	Mohamed Badouch, Mehdi Boutaoune	Personalized Travel Recommendation Systems: A Study of Machine Learning Approaches in Tourism	Journal of Artificial Intelligence, Machine Learning, and Neural Network (2023)	Presented a hybrid recommender system combining collaborative and content-based filtering, enhanced through machine learning for personalized tourism recommendations.
4.	Bhargava H C, Yashas V, Vijay, Prof. Sunanda H G, Suhas M	Personalized Travel Recommendation System Using Machine Learning	IJARIIE (2023)	Used a hybrid recommendation system employing cosine similarity and Singular Value Decomposition (SVD) for tailored travel suggestions.
5.	Kristina P. Sinaga, Miin-Shen Yang	Unsupervised K-Means Clustering Algorithm	IEEE Access (2020)	Proposed an unsupervised k-means (U-k-means) algorithm to automatically determine the optimal number of clusters without prior initialization
6.	Peilin Chen	Design of Travel Itinerary Planning System Based on Artificial Intelligence	Journal of Physics: Conference Series (2020)	Developed a travel planning system utilizing association rule mining, intelligent data processing, and multi-threaded performance optimization.



7.	Yu Chao, Yang Mengli	Co-existing in Differences: Tailoring Travel Itineraries for Tourists with Similar Interests	IEEE Access (2023)	Applied clustering algorithms to group tourists with similar preferences, optimizing itinerary customization based on heterogeneous needs
8.	Divya Mereddy	Travel Itinerary Creation Based on Data Techniques	International Journal of Data Mining Management Systems (IJDMS), Vol. 1, No. 4, October 2022	Utilized Hamilton graphs and HDBSCAN clustering to create optimized multi-day travel itineraries, addressing tourist constraints like time and distance.
9.	Sajal Halder, Kwan Hui Lim, Jeffrey Chan, Xiuzhen Zhang	Efficient Itinerary Recommendation via Personalized POI Selection and Pruning	Knowledge and Information Systems (2022)	Proposed the EffiTourRec algorithm using Monte Carlo Tree Search (MCTS)-based reinforcement learning with heuristic pruning to optimize POI selection.
10.	José Ruiz-Meza, Jairo R. Montoya-Torres	A Systematic Literature Review for the Tourist Trip Design Problem	Operations Research Perspectives (2022)	Conducted a systematic review of the Tourist Trip Design Problem (TTDP), analyzing solution techniques and future research opportunities.
11.	Kwan Hui Lim et al.	PersTour: A Personalized Tour Recommendation and Planning System	Conference Paper (2016)	Introduced PersTour, leveraging ant colony optimization to customize itineraries based on user preferences and time constraints.
12.	Luis Castillo, Eva Onaindia, Laura Sebastia, et al.	SAMAP: A User-Oriented Adaptive System for Planning Tourist Visits	Expert Systems with Applications (2013)	Developed SAMAP, which uses machine learning, user modeling, and case-based reasoning to plan and adapt personalized tourist itineraries.



13.	Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, Shanika Karunasekera	Personalized Trip Recommendation for Tourists Based on User Interests, Points of Interest Visit Durations, and Visit Recency	Knowledge and Information Systems (2018)	Proposed the PersTour algorithm to recommend personalized tours using POI popularity, user interest preferences, and trip constraints. Modeled using the Orienteering Problem.
14.	Aayushi Bhansali, Niharika Premkumar, Parshav Pagariya	Trip Itinerary Planner	IJRASET (2023)	Developed a real-time itinerary planner using the A-star algorithm, optimizing travel time, traffic, and POI ratings.
15.	Sahil Ambade, Ajinkya Gaikwad, et al.	Survey Paper on Automated Tour Planner	International Journal of Advanced Research in Science, Communication, and Technology (IJARSCT), Volume 3, Issue 3, May 2023	Reviewed existing tour planners and proposed automated solutions for personalized itineraries.
16.	Weimin Zheng, Mengling Li, et al.	Leveraging Tourist Trajectory Data for Effective Destination Planning	Tourism Management, Volume 89, 2022	Utilized tourist trajectory data to optimize destination planning with heuristic approaches
17.	Suresh Babu Dasari, V. Vandana, A. Bharathee	Smart Travel Planner Using Hybrid Model	2023 International Conference on IDCIoT	Presented a hybrid model combining optimization algorithms for intelligent travel planning systems
18.	Anita Gehlot, Rajesh Singh	Research Guide for ML-Based Smart Tourist System	IEEE IIHC-2022, November 2022	Discussed various machine learning techniques applied to develop smart tourist systems.



19.	Yu, C., Yang, M.	Co-existing in Differences: Tailoring Travel Itineraries	IEEE Access (2023)	Clusters tourists by preferences and optimizes itineraries for each group.
20.	Ho, N. L., Lee, R. K.-W., Lim, K. H.	SBTREC: A Transformer Framework for Personalized Tour Recommendation	IEEE Access (2023)	Uses transformer-based sequential modeling to recommend personalized tours.



2.3 Outcome of survey

The survey provided deep insights into the preferences, pain points, and expectations of travelers, offering valuable direction for the development of GlobeTrail as a comprehensive travel platform. Below are the key findings and conclusions derived from the survey:

Key Findings

1. User Preferences:
 - Personalized Recommendations: Travelers overwhelmingly expressed the need for AI-driven personalized itineraries, tailored to their interests, time, and budget constraints.
 - Real-Time Adaptability: Features like live updates on weather, traffic, and POI availability emerged as critical for a seamless and reliable travel experience.
 - Ease of Use: The majority of users favored a unified platform that integrates itinerary planning, ticket booking, and accommodation arrangements into a single solution.
2. Challenges Faced:
 - Fragmented Travel Planning: Users reported frustration with the need to switch between multiple platforms for itinerary creation, bookings, and trip updates.
 - Outdated Information: Frequent complaints about outdated POI details, unavailable attractions, and non-dynamic itinerary tools highlighted a gap in real-time data integration.
 - Sustainability Gaps: There is an unmet demand for eco-conscious travel options, such as routes optimized for walking or public transport and accommodations rated for sustainability.
3. Insights on Competitors:
 - Existing platforms like TripAdvisor and Booking.com were seen as functional but lacking personalization, real-time responsiveness, and a holistic user experience.
 - This provides GlobeTrail an opportunity to differentiate itself by addressing these shortcomings.

Conclusions

The survey results highlight GlobeTrail's potential to lead the market by addressing key traveler pain points through personalization, real-time adaptability, and seamless integration. Emphasizing sustainable travel options and interactive tools for real-time itinerary adjustments further aligns the platform with user demands. By leveraging AI-driven personalization, GlobeTrail can deliver a unified, cutting-edge solution that exceeds expectations.



3. Need For the Product

3.1 Need for the Product

GlobeTrail, a travel website offering curated itinerary recommendations and integrated ticket booking, addresses key gaps in existing travel platforms and caters to the growing demand for seamless, personalized, and dynamic travel planning. Current travel platforms often fall short in providing a holistic, user-friendly experience, leading to fragmented and time-consuming processes for travelers. With GlobeTrail, users can plan their trips, receive AI-curated itineraries, and book tickets all in one place, ensuring a hassle-free travel experience.

3.2 Drawbacks of Existing Platforms

TripAdvisor: Focuses on POI popularity and user reviews but lacks real-time adaptability and personalized itineraries. Static recommendations fail to reflect dynamic user needs.

How GlobeTrail Solves It: Incorporates AI and real-time updates, offering tailored and dynamic recommendations.

Booking.com: Specializes in booking hotels and activities but lacks curated or AI-driven itineraries. Offers little flexibility for customization or real-time itinerary adjustments.

How GlobeTrail Solves It: Combines booking with personalized itinerary planning, ensuring a seamless user experience.

GoExplore: Relies on basic route optimization and POI popularity but does not cater to offbeat destinations or provide personalized itineraries.

How GlobeTrail Solves It: Uses sentiment analysis and AI-powered recommendations to deliver diverse and user-centric travel experiences.

TourMate: Offers fixed itineraries with minimal input but lacks real-time data for traffic, weather, and POI closures. Recommendations are static and fail to adjust dynamically.

How GlobeTrail Solves It: Leverages real-time adaptability and user-specific interest modeling to deliver flexible and accurate travel plans.

3.3 Applications of GlobeTrail

1. Personal Travel Planning:

GlobeTrail provides travelers with AI-curated itineraries based on their preferences and constraints. Users can plan multi-day trips, select POIs, and customize itineraries in real-time.

2. Integrated Ticket Booking:

The platform allows seamless booking of flights, accommodations, and attractions, ensuring users don't have to switch between multiple platforms.

3. Group Travel Management:

GlobeTrail offers tools for collaborative itinerary planning, catering to families, friends, or corporate groups. It provides group-specific recommendations and allows shared access for easy coordination.



4. Sustainable Tourism:

The platform promotes eco-conscious choices by recommending sustainable accommodations, public transport, and low-impact activities, aligning with global sustainability goals.

5. Dynamic Itinerary Adjustments:

With real-time updates, GlobeTrail dynamically adjusts itineraries to account for weather changes, traffic delays, or POI closures, ensuring a smooth travel experience.

6. Destination Marketing and Local Tourism:

GlobeTrail can partner with local businesses and tourism boards to highlight hidden gems and lesser-known destinations, supporting local economies and promoting offbeat travel.

Conclusion

GlobeTrail is an essential innovation for modern travelers, addressing the fragmented, impersonal, and outdated nature of existing travel platforms. By combining AI-powered personalization, real-time adaptability, and integrated booking capabilities, GlobeTrail offers a seamless, sustainable, and user-friendly travel planning experience. Its applications extend to personal, group, and sustainable travel, ensuring it caters to diverse user needs while standing out as a comprehensive and competitive solution in the travel industry.



4. Problem Formulation

4.1 Problem Formulation

Travel planning is often fragmented and inefficient, with existing platforms failing to provide a seamless and dynamic experience. Current systems are either too focused on individual services like ticket booking or rely heavily on static data, offering generic recommendations that do not cater to real-time scenarios or diverse user needs. Additionally, managing group preferences and providing eco-friendly options remain significant challenges in the travel industry. GlobeTrail aims to bridge these gaps by introducing a unified platform that leverages real-time adaptability, sentiment analysis, and advanced AI techniques to offer personalized and flexible travel solutions.

4.2 Product Objectives

1. Real-Time Context Awareness: Provide dynamic itinerary adjustments based on live traffic, weather, and POI crowd levels, ensuring users have the most feasible plans during their trip.
2. Sentiment-Based Optimization: Use sentiment analysis of reviews and social media data to refine POI rankings, ensuring recommendations align with current user satisfaction and trends.
3. Group Preference Balancing: Dynamically handle group travel preferences using advanced multi-agent reinforcement learning and collaborative filtering.
4. Eco-Friendly Travel Recommendations: Promote sustainable tourism by suggesting public transport routes, green accommodations, and low-impact activities.
5. Seamless Integration: Offer a unified platform combining itinerary planning, ticket booking, and customization tools, eliminating the need for multiple apps or websites.

4.3 Novelty

GlobeTrail introduces cutting-edge solutions to overcome the limitations of existing travel platforms, setting it apart through the following innovations:

1. Real-Time Context Awareness:
 - Proposed Solution: Integrates APIs such as Google Maps, OpenWeather, and crowd data sources to provide live updates for traffic, weather, and POI availability.
 - Why It's Novel: Unlike existing approaches like Monte Carlo Tree Search (MCTS)-based



algorithms (Lim et al., 2020), which incorporate traffic and queuing times but lack weather integration, GlobeTrail dynamically adjusts itineraries for real-time conditions.

2. Sentiment-Based Optimization:

- Proposed Solution: Leverages advanced NLP models like BERT and RoBERTa to analyze user reviews and social media comments, refining POI rankings based on real-time sentiments.
- Why It's Novel: Builds on frameworks like SBTREC (Ho et al., 2023), which uses sentiment analysis, and improves upon DeepAltTrip (2021), which relied heavily on historical data. GlobeTrail introduces real-time sentiment-driven ranking, making recommendations more accurate and relevant.

3. Group Preference Balancing:

- Proposed Solution: Implements multi-agent reinforcement learning (MARL) to treat each group member as an agent, dynamically balancing individual preferences while using weighted collaborative filtering to resolve conflicts.
- Why It's Novel: Moves beyond individual-focused systems like PersTour (2018) and SAMAP (2008), which did not address group dynamics. GlobeTrail ensures group travel planning is collaborative and inclusive.

4. Eco-Friendly Integration:

- Proposed Solution: Provides recommendations for sustainable accommodations, public transport, and eco-conscious activities.
- Why It's Novel: Few platforms focus on sustainability as a core feature. GlobeTrail aligns with modern travel trends by actively promoting green travel options.

5. Interactive Customization:

- Proposed Solution: Offers tools for real-time itinerary customization, allowing users to add, remove, or rearrange POIs dynamically.
- Why It's Novel: Most platforms provide fixed itineraries with minimal flexibility. GlobeTrail enhances user control, making the experience more engaging and adaptable.



4.4 Scope of the Project

Immediate Scope:

1. **Develop AI-Powered Recommendation Systems:** Design algorithms to generate highly personalized itineraries based on user profiles, preferences, and past behaviors.
2. **Real-Time Data Integration:** Leverage APIs for live updates on weather, traffic, crowd density, and local events to ensure adaptable and up-to-date travel plans.
3. **Sentiment Analysis for POI Optimization:** Build machine learning models to analyze user reviews and feedback for dynamically ranking points of interest (POIs) based on popularity, satisfaction, and relevance.
4. **Interactive Customization Tools:** Implement user-friendly interfaces for itinerary adjustments, enabling travelers to modify plans seamlessly with real-time feedback.
5. **Targeted Platform Launch:** Roll out the platform for individual travelers, prioritizing frequent and occasional travelers to ensure early adoption and user engagement.

Future Scope:

1. **Group Travel Support:** Expand capabilities to accommodate collaborative travel planning for families, friends, and corporate teams with shared itineraries and synchronization features.
2. **Enhance Partnerships with Local Businesses:** Collaborate with local service providers to promote hidden gems, cultural experiences, and offbeat destinations, fostering unique travel opportunities.
3. **Global Scaling:** Enable multilingual and multi-currency support to cater to an international audience, ensuring accessibility and inclusivity.
4. **Eco-Conscious Features:** Introduce advanced sustainability tools, such as carbon footprint calculators and recommendations for eco-friendly accommodations and activities.
5. **Seamless Platform Evolution:** Continue iterating based on user feedback and emerging trends, ensuring the platform remains innovative, relevant, and user-focused.

5. Proposed Design

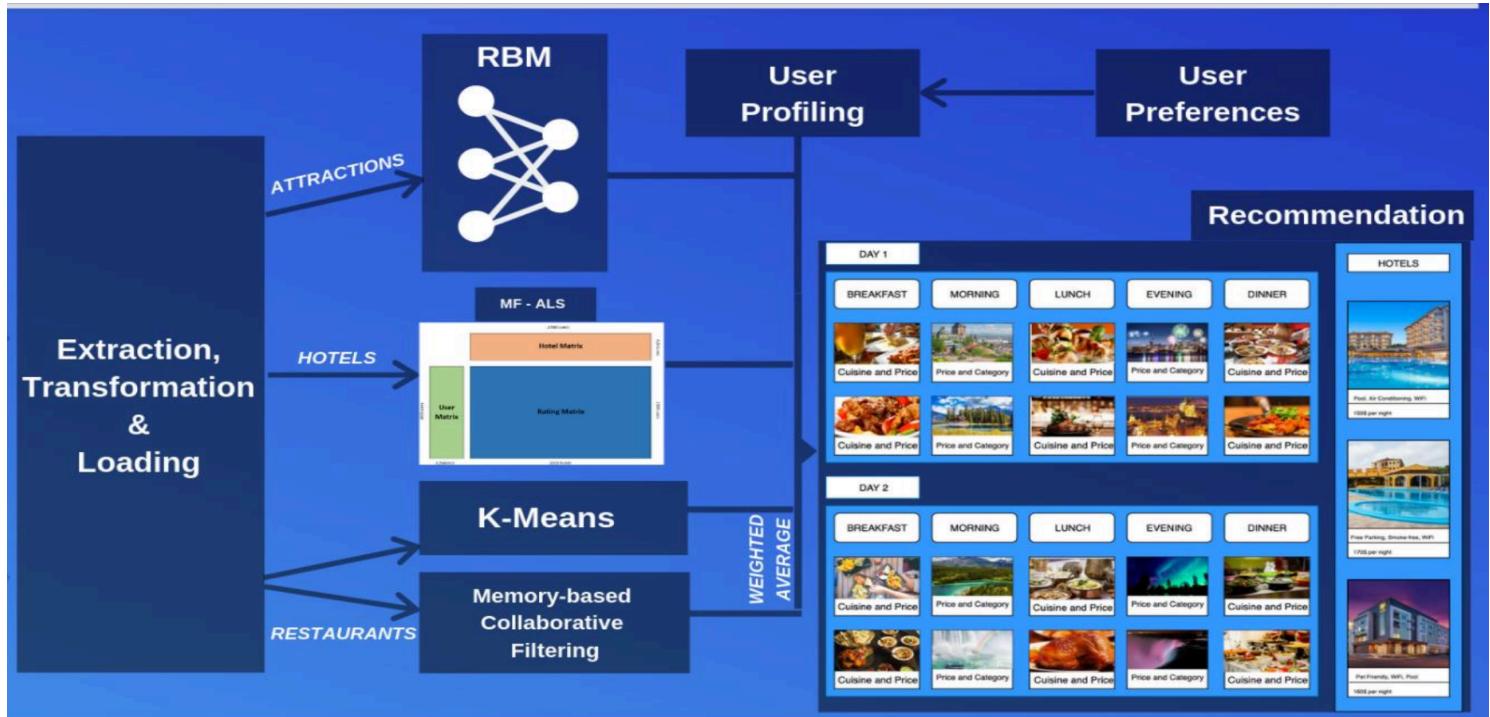


Figure 5.1: System architecture illustrating data extraction, transformation, and loading pipeline integrated with RBM, matrix factorization, and K-Means clustering for generating personalized travel recommendations.

1. **Extraction, Transformation, and Loading (ETL):** Raw data related to attractions, hotels, and restaurants is collected from various sources. The data undergoes a transformation process, where it is cleaned, standardized, and formatted to ensure consistency and usability. Finally, the processed data is loaded into a structured database or system, ready for analysis and use by the recommendation models.

2. Recommendation Techniques:

- **LLM-Driven User Profiling:**

The chatbot uses a large language model to interpret user input (e.g., source location, budget, trip duration, group size), even when incomplete or ambiguous. Based on this, it dynamically builds a user profile and fills in missing preferences with intelligent defaults or common travel patterns.

- **Matrix Factorization (MF-ALS):**

For optimizing destination and accommodation choices, collaborative filtering via MF-ALS is applied. It analyzes past user behavior patterns (if available) or anonymized historical trends to suggest similar travel destinations and stays.



- **K-Means Clustering:**

Destinations and activities are grouped based on features such as type (hill station, cultural site, adventure), cost, popularity, and review sentiment. This clustering helps the system identify which options align best with user interests and travel context.

- **Travel Mode & Cost Inference Engine:**

A custom engine estimates viable transportation options (e.g., flight, train, car) based on the user's source and preferred budget/time constraints. It combines geographic data with real-world travel cost APIs to recommend efficient travel modes and realistic costs.

- **Hybrid Weighted Scoring:**

Outputs from recommendation models and heuristics (e.g., proximity, affordability, popularity) are aggregated using a weighted scoring system. The weights are dynamically adjusted based on user preferences (e.g., budget-focused vs. experience-focused travelers).

3. **Weighted Average:** The system combines the outputs from the various recommendation techniques using a weighted average approach. Each model's output is assigned a weight based on its relevance and accuracy for the specific user or context. This step ensures that the final list of recommendations is both comprehensive and prioritized, reflecting the user's preferences accurately.

4. **User Preferences and Profiling:** The chatbot creates a composite profile of the traveler(s) using inputs like budget, duration, group size, and trip goals. It then aligns this profile with the structured data (attractions, transport options, accommodations, etc.) to generate a personalized travel plan. The system makes context-aware decisions—e.g., recommending family-friendly places for a group with kids or budget options for college students.

5. **Output:** The final output is a day-wise, personalized itinerary that includes:

- Suggested destinations and the best modes of transport (with costs and durations)
- Accommodation options tailored to group size and budget
- Daily plans covering key attractions, optimized routes, and estimated expenses
- Adaptive fallback recommendations in case of incomplete user input



6. Implementation

6.1 GUI Design

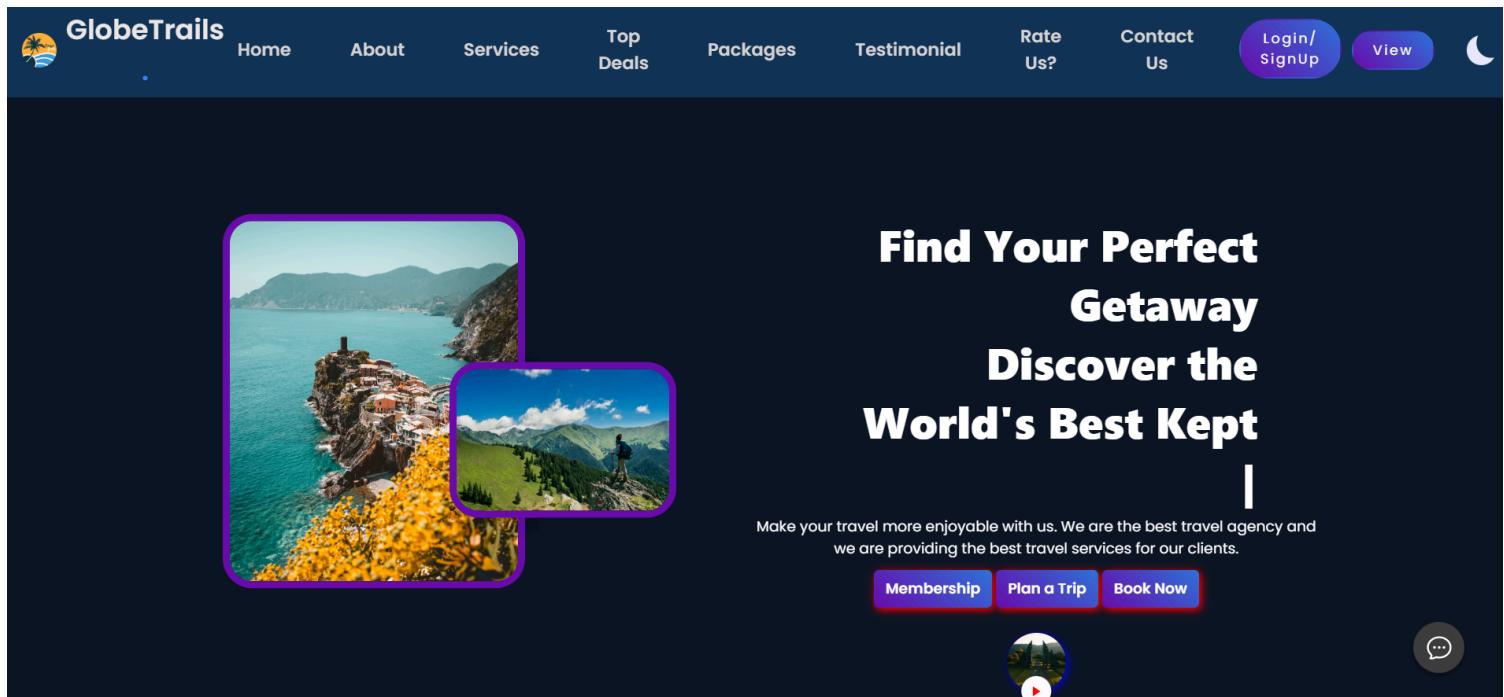


Figure 6.11 HomePage of GlobeTrails Website

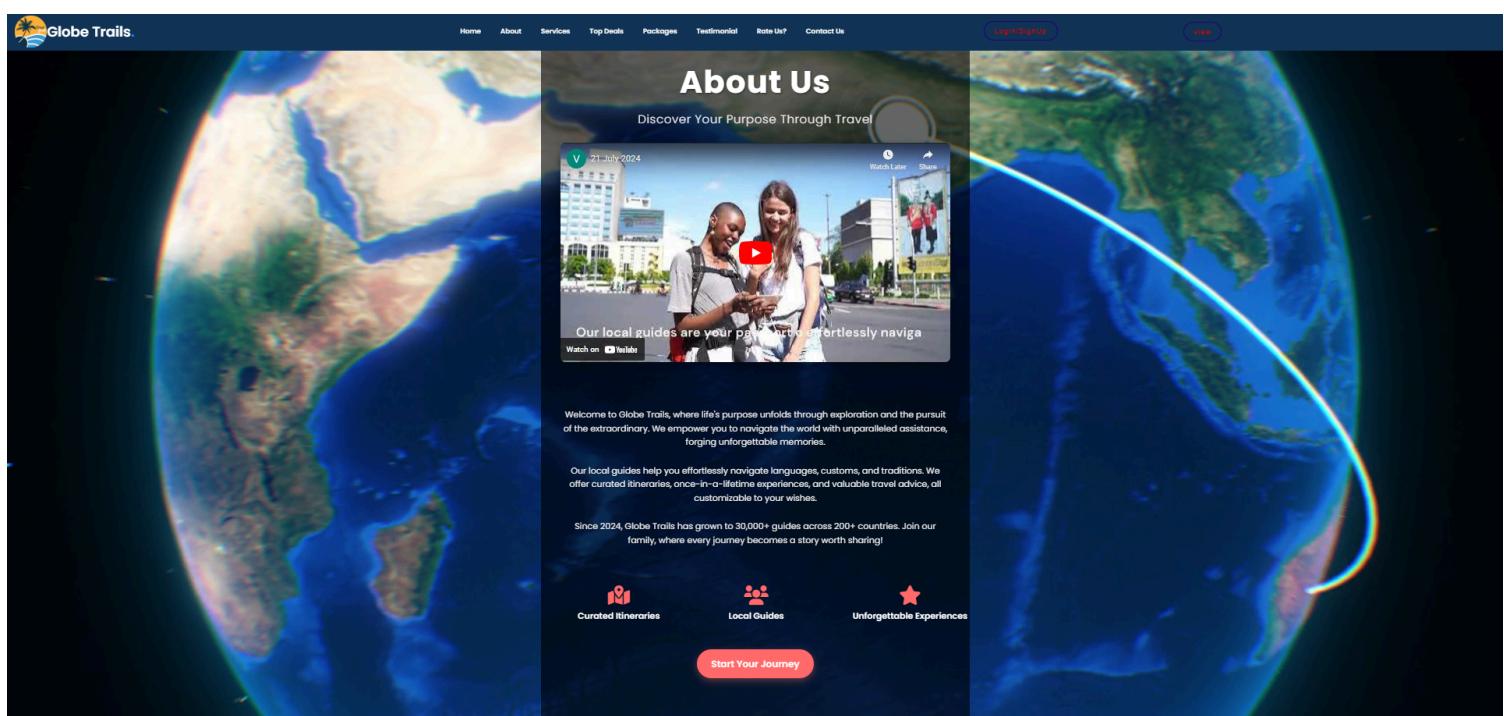


Figure 6.12 About Us of GlobeTrails Website



The screenshot shows the 'Plan a Trip' section of the GlobeTrails website. It features a form with fields for Destination, Start Date, End Date, Budget, and Accommodation (set to Hotel). Below the form is a 'Plan Trip' button. To the right, there's a 'Trip Suggestions' section with four cards: Manali (showing a river scene), Jaipur (showing a night view of a city), Meghalaya (showing a bridge over water), and Maldives (showing a beach resort). The background has a travel-themed illustration with a plane, binoculars, a compass, and various icons.

Figure 6.13 User Credentials Input UI of GlobeTrails Website

The screenshot shows the footer of the GlobeTrails website. It includes a 'GlobeTrails' logo and a navigation bar with links to Home, About, Services, Top Deals, Packages, Testimonial, Rate Us?, Contact Us, Login/SignUp, and View. Below this is a 'Subscribe to get special prize' section with a 'Subscribe' button and an input field for email. The footer is divided into several sections: 'GlobeTrails' (with a message about unforgettable adventures), 'Places You May Want to Visit' (listing Cultural Experiences, Honeymoon Places, and Adventure travel), 'Support' (listing Our Blogs, FAQs, Terms & Conditions, Privacy Policy, Licensing, Contact Us, Our Contributors, and SiteMap), 'Address' (a map showing the location of Wellington Marketing Pvt Ltd), 'Contact Us' (with email, phone number, social media links, and visitor count), and 'Select Language' (a box powered by Google Translate).

Figure 6.14 Footer of GlobeTrails Website



Feature	Globetraffic	Google Gemini	Expedia's Romie	GuideGeek	KLM's BlueBot	Trip.com Assistant
ML-Powered Personalization	✓	✗	✗	✓	✗	✗
Real-Time API Integration	✓	✓	✓	✓	✓	✓
Multilingual Support	✓	✓	✓	✓	✓	✓
Conversational UI	✓	✓	✓	✓	✓	✓
Destination Insights	✓	✓	✓	✓	✓	✓
Cost Optimization	✓	✗	✗	✗	✗	✗
Seamless Booking	✓	✓	✓	✓	✓	✓
Proactive Alerts	✓	✓	✓	✓	✓	✓
Customizable Preferences	✓	✗	✗	✗	✗	✗
Continuous Learning	✓	✗	✗	✗	✗	✗

Figure 6.15 Comparison Table for GlobeTrails Chatbot with other Travel Itinerary Chatbots

6.2 Technologies Used

6.2.1 Backend

Python (Flask), pandas, scikit-learn, LLM integration, Selenium for dynamic data extraction.

6.2.2 Frontend

HTML, CSS, JavaScript (AJAX, Leaflet.js), Axios for HTTP requests.

6.2.3 Database

Utilizes CSV files for storing and processing data related to travel destinations and user preferences.

6.3 Algorithms Used

6.3.1 Web Scraping

- **Selenium:** Used for scraping Google Maps and other travel data sources.

6.3.2 Natural Language Processing (NLP)

- **LLM and basic NLP:** Used for interpreting user queries, inferring missing inputs, and analyzing reviews/descriptions.

6.3.3 Clustering

- **K-Means Clustering:** Categorizes destinations based on location, type, and popularity for better trip grouping.

6.3.4 Itinerary Optimization

- **Heuristics + Distance-based Sorting:** Optimizes daily routes and activities to ensure efficiency and relevance.

6.3.5 Real-Time Data Integration

- **API Calls:** Fetches live transport costs, travel modes, weather and traffic data adapt the itinerary dynamically.



7. Experimentation & Results

7.1 Dataset formation

This module utilizes web scraping techniques and APIs to gather comprehensive data about tourist destinations, including their names, coordinates, ratings, reviews, and categories. The collected data undergoes preprocessing to clean and structure it for further analysis.

- The system gathers data from Google Maps using web scraping and API calls.
- Reviews, ratings, coordinates, and other relevant information are collected and stored in the database.
- Collected data is cleaned, normalized, and structured for use in machine learning models.
- Missing values are handled, and weights are calculated based on ratings and votes.

City	Place	Ratings	Distance	Place_desc
Manali	1. Capture the Sceneries of Old Manali	3.9	2 km from city center	On the other side of the Manalsu river is a part of Manali, time left behind...
Manali	2. Engage in the Adventures of Solang Valley	4.6	8 km from city center	Solang Valley is one of the most popular tourist destinations in Himachal Pradesh...
Manali	3. Jogini Waterfall	4.6	4 km from city center	Jogini Waterfall is located about 3 kilometres away from the bustling town of Manali...
...
Goa	1. Fun time in Calangute Beach	4.4	10 km from city center	Goa, Fun time in Calangute Beach, goa's beaches...

Table 7.11 Sample tuples in scraped dataset



7.2 Performing NLP and Extracting weights

- Text Cleaning and Tokenization is used to preprocess and analyze reviews and descriptions.
- Weighted Average: Calculates the weight of each place using ratings and votes, normalized to a 2.5 to 3.5 scale for consistency.
- K-Means Clustering + Memory based Collaborative Filtering Hybrid technique: Groups places based on geographical coordinates and user-specified categories to ensure convenient travel plans.

7.3 Clustering Phase

Being a Content Based Filtering Technique, K-Means does not suffer from cold start problem, but is highly sensitive to outliers in the dataset and dimensionality reduction techniques can help make the algorithm more efficient. Thus, we carefully performed EDA to analyze the restaurant's dataset. We computed the Positive-Negative sentiment ratio per restaurant based on user reviews. This helped weigh down the fact that the dataset was highly skewed towards positive reviews and tips. Also, we performed Principal Component Analysis to reduce the dimensionality, however, we did not scale our features as our input features did not have a linear relationship-scaling these would result in losing important semantic similarities in words. In-order to find the number of components apt for our dataset using PCA, we plotted the variance in features with respect to the number of components and chose the smallest number of components for which the variance was not less than 95%.

For KNN algorithm, the restaurants dataset was used to create the user-item interaction matrix, also known as rating matrix. The rating matrix contains n rows for users and m columns for each restaurant present in the dataset. The ratings provided by the user i for attraction j was filled at (i,j) position of the matrix. The positions without any user-item interaction were filled using a 0 rating. An inference data was created using the user preferences aka item attributes provided to us. The ratings used here is not the actual user rating, it is a weighted sum of user ratings and user defined labels for the item.

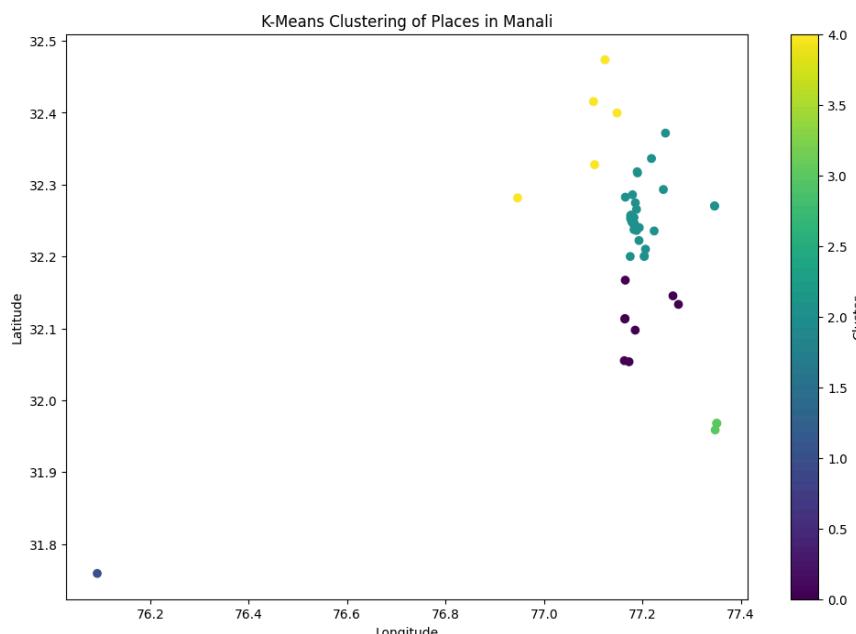


Table 7.31 K-Means clustering for adventure and scenic destinations in Manali based on weights



7.4 Results

Clustering Analysis-

RECOMMENDATION: For predictions on new points (new users), we use K-means to assign a cluster to the new point based on the user preference. Then, we find the points within the assigned cluster that are closest to our new point using ‘Cosine Similarity’. We go for cosine over euclidean as the input features in our case are word embeddings obtained from Word2Vec, whose semantics will be better captured by using cosine similarity (euclidean will calculate the mere difference in letters of the words). The final recommendation list is a weighted average (based on rating) of the recommendations obtained from the content-based (K-Means) technique and the Memory-Based Collaborative Filtering technique(K-Nearest Neighbors).

📍sattal

⦿Categories: Nature and Wildlife, Scenic, Beach Holidays

⦿Description: Sattal is a group of seven freshwater lakes located in the lower regions of Bhimtal. Home to many migratory birds, it is popular amongst nature buffs and bird watchers.

📍bhimtal

⦿Categories: Scenic, Adventure, Offbeat and Eco-Tourism

⦿Description: Bhimtal is an idyllic and less-crowded version of Nainital, 23 km away. It is a scenic hill station, the charm of which lies in its off-beat, tranquil atmosphere.

📍neem karoli baba ashram

⦿Categories: Beach Holidays, Offbeat and Eco-Tourism, Scenic

⦿Description: Dedicated to Shri Neem Karoli Baba, who was a Hindu saint and Guru and had a lot of devotees from all over the world. Steve Jobs had come to India in 1974 to become his devotee, though sadly he had passed away by that time.

Table 7.41 Output recommendation sample

Chatbot integration-

The user enters a query (e.g., "Plan a 5-day trip from Delhi under ₹50,000 for 3 people"). The chatbot uses natural language processing (via an LLM) to extract key parameters like source location, budget, duration, and group size. If any detail is missing, it intelligently fills in defaults. Extracted inputs are matched against a structured database containing destinations, reviews, coordinates, travel modes, and cost data. The system evaluates viable transport options and accommodations based on constraints like budget and time. A detailed, personalized itinerary is generated, including destination info, travel routes, cost estimates, accommodation options, and suggested activities—adjusted in real time using live weather and traffic data when available.



GlobeTrail

How can I assist you today?

Plan a 3-day trip from Mumbai to Delhi under ₹10,000 for 1 person

Here's a 3-day trip plan from Mumbai to Delhi:

Day 1:

1. Take an overnight train from Mumbai to Delhi (12 hours, ₹1,500-2,000)
2. Check-in at a budget hotel (₹500-800)
3. Visit:
 - * Qutub Minar
 - * Humayun's Tomb
4. Enjoy street food at Chor Bazaar for

Message

SEND

END CONVERSATION

Figure 7.42 Output itinerary recommendation



Here's a 3-day trip plan from Mumbai to Delhi:

Day 1:

1. Take an overnight train from Mumbai to Delhi (12 hours, ₹1,500-2,000)
2. Check-in at a budget hotel (₹500-800)
3. Visit:
 - * Qutub Minar
 - * Humayun's Tomb
4. Enjoy street food at Chor Bizarre for dinner (₹200-300)

Day 2:

1. Visit:
 - * India Gate
 - * President House
 - * Parliament House
2. Explore Chandni Chowk market
3. Enjoy local dinner at a dhaba (₹300-500)

Day 3:

1. Visit:
 - * Red Fort
 - * Jama Masjid
2. Explore Gandhi Memorial and Raj Ghat
3. Return to the train station for the journey back to Mumbai

Budget Breakdown:

1. Transportation: ₹1,500-2,000 (train ticket)
2. Accommodation: ₹1,500-2,400 (budget hotel for 2 nights)
3. Food and activities: ₹2,500-3,500 (street food, local dinner, entrance fees)
4. Total: ₹6,500-8,900

Figure 7.43 Complete itinerary response



8. Future Scope & Conclusion

1. **Collaborative group travel planning:** Expand GlobeTrail to support collaborative itinerary creation, enabling real-time synchronization, shared editing, and seamless coordination among group members.
2. **Multilingual and multi-currency integration:** Incorporate support for multiple languages and currencies to cater to a global user base, ensuring accessibility and convenience for international travelers.
3. **Sustainability-focused features:** Develop advanced eco-conscious tools, including carbon footprint calculators, sustainability ratings for accommodations and activities, and recommendations prioritizing low-impact travel options.
4. **Augmented reality (AR) integration:** Explore the use of AR technologies to provide immersive, interactive sightseeing experiences, enhancing user engagement and destination exploration.
5. **Deep learning-based personalization:** Leverage state-of-the-art deep learning models to refine personalization algorithms, delivering even more tailored recommendations based on nuanced user behavior and preferences.
6. **Partnerships with local businesses:** Establish collaborations with local vendors, tour operators, and cultural organizations to promote lesser-known attractions, authentic experiences, and community-based tourism.
7. **Predictive analytics for travel demand:** Incorporate predictive models to forecast trends in travel demand, seasonal variations, and emerging destinations, improving recommendation timeliness and relevance.
8. **Adaptive learning algorithms:** Implement adaptive learning systems that continuously refine recommendations by incorporating user feedback, implicit signals, and real-time behavioral data.
9. **Seamless integration of real-time data:** Enhance GlobeTrail's architecture to incorporate dynamic data streams (e.g., live events, weather, local advisories) to provide up-to-date, context-aware recommendations.
10. **Commitment to sustainable and responsible tourism:** Position GlobeTrail as a leader in promoting sustainable tourism practices, balancing personalization and convenience with environmental and cultural stewardship.
11. **Industry impact:** This project demonstrates the transformative potential of combining AI, real-time data, and sustainability in travel planning, setting a new benchmark for intelligent, eco-friendly, and user-centric travel solutions.



9. REFERENCES / BIBLIOGRAPHY

1. Sinaga, K. P., & Yang, M. S. (2020). *Unsupervised K-Means Clustering Algorithm*. IEEE Access.
2. Chen, P. (2020). *Design of Travel Itinerary Planning System Based on Artificial Intelligence*. Journal of Physics: Conference Series, 1533(3), 032078.
3. Yu, C., & Yang, M. (2023). *Co-existing in Differences: Tailoring Travel Itineraries for Tourists with Similar Interests*. IEEE Access.
4. Mereddy, D. (2023). *Travel Itinerary Creation Based on Data Techniques*. University of Cincinnati.
5. Dumas et al. (1995). *An Optimal Algorithm for the Traveling Salesman Problem with Time Windows*. Technical Report.
6. Rashid, S. M. M., Ali, M. E., & Cheema, M. A. (2021). *DeepAltTrip: Top-k Alternative Itineraries for Trip Recommendation*. IEEE Transactions on Knowledge and Data Engineering, under review.
7. Halder, S., Lim, K. H., Chan, J., & Zhang, X. (2022). *Efficient Itinerary Recommendation via Personalized POI Selection and Pruning*. Knowledge and Information Systems, 64, 963–993.
8. Onaindia, E., Sebastia, L., & Anaya, A. R. (2013). *SAMAP: An User-Oriented Adaptive System for Planning Tourist Visits*. Expert Systems with Applications.
9. Ruiz-Meza, J., & Montoya-Torres, J. R. (2022). *A Systematic Literature Review for the Tourist Trip Design Problem: Extensions, Solution Techniques, and Future Research Lines*. Operations Research Perspectives.
10. Bhansali, A., Premkumar, N., Pagariya, P., Jain, V., Mahansaria, V., & Varma, S. (2023). *Trip Itinerary Planner*. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 11(8), 918–920.
11. Lim, K. H., Chan, J., Karunasekera, S., & Leckie, C. (2018). *Personalized Trip Recommendation for Tourists Based on User Interests, Points of Interest Visit Durations, and Visit Recency*. Knowledge and Information Systems, 54(2), 375–406.
12. Ambade, S., Gaikwad, A., Bodare, S., Jadhav, P., & Deore, N. (2023). *Survey Paper on Automated Tour Planner*. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT).
13. Badouch, M., & Boutaoune, M. (2023). *Personalized Travel Recommendation Systems: A Study of Machine Learning Approaches in Tourism*. Journal of Artificial Intelligence, Machine Learning, and Neural Network.
14. Bhargava, H. C., Yashas, V., Vijay, S., Suhas, M., & Sunanda, H. G. (2023). *Personalized Travel Recommendation System Using Machine Learning*. International Journal of Advanced Research in Innovative Ideas in Education (IJARIIE), 9(3), 1612–1620.
15. Lim, K. H., Chan, J., Karunasekera, S., & Leckie, C. (2016). *PersTour: A Personalized Tour Recommendation and Planning System*. Proceedings of the 27th ACM Conference on Hypertext and Social Media (HT).
16. Ho, N. L., Lee, R. K.-W., & Lim, K. H. (2023). *SBTREC: A Transformer Framework for Personalized Tour Recommendation Problem with Sentiment Analysis*. Singapore University of Technology and Design.
17. Lim, K. H., Chan, J., Karunasekera, S., & Leckie, C. (2018). *Tour Recommendation and Trip Planning Using Location-Based Social Media: A Survey*. Knowledge and Information Systems, 61(2), 375–406.
18. Ho, N. L., Lee, R. K.-W., & Lim, K. H. (2023). *SBTREC: A Transformer Framework for Personalized Tour Recommendation Problem with Sentiment Analysis*. Singapore University of Technology and Design.



Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)

