International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 09 | Sep 2024 www.irjet.net p-ISSN: 2395-0072

Gateway Guide – ML Powered Travel Itinerary Curation

P Jahnavi[1]

Velagapudi Ramakrishna Siddhartha Engineering College

Abstract - A travel itinerary generator is a tool that helps you plan your trips in a customized and efficient way. It effortlessly combines activities, accommodations, and routes to create a seamless and personalized travel experience just for you. Getaway Guide is an AI powered travel itinerary system for corporate professionals to take a break during long weekends or business trips in Karnataka, India. It uses deep learning to combine user preferences, a massive Points of Interest database and real time data to give personalized recommendations and optimize travel plans. The methodology involves collecting data from geographic databases and APIs, clustering Points of Interest using Gaussian Mixture Model and Ant Colony Optimization to find the best travel routes. Real time weather forecasts are used to fine tune the travel plans and HERE API is used for route generation. Evaluation showed high accuracy in personalized itineraries and efficiency in travel paths, Ant Colony Optimisation minimized travel distance and time and Gaussian Mixture Model ensured relevant Points of Interest. Getaway Guide changes the way travel planning happens in Karnataka by giving personalized and efficient solutions to corporate professionals, saving time and effort and giving unique and immersive travel experiences. This is a big leap in AI powered travel planning catering to modern business travelers.

Key Words: Optimal Route, Clustering, Ant Colony Optimisation, Folium, Gaussian Mixture Model, Points of Interest

1.INTRODUCTION

Planning trips online has become so much easier with the rise of travel planning tools. But still, many travelers find it challenging to create personalized itineraries that suit their interests and travel needs. This struggle is especially felt by corporate professionals who are looking for quick getaways during long weekends or business trips in Karnataka, India. Karnataka is not only known for its abundance of tourist attractions but also as a hub for technical conferences and corporate events, making it a popular destination for business travelers.

The traditional way of planning trips involves browsing through various websites and guidebooks, which can often lead to confusion and exhaustion. Existing travel platforms usually offer generic recommendations that fail to cater to individual preferences. That's where Getaway Guide comes in. It's an AI-powered travel itinerary system designed specifically for travelers in Karnataka. Getaway Guide is tailored to suggest destinations and activities that are perfect for corporate professionals seeking quick getaways. It uses advanced deep learning techniques to provide personalized recommendations and optimize travel plans. By considering a vast database of Points-of-Interest (POIs) within Karnataka, user preferences, and real-time data, Getaway Guide helps corporate travelers create immersive and customized experiences. Whether you're interested in exploring vibrant cities, cultural landmarks, or natural wonders, Getaway Guide adjusts to your preferences. Its aim is to enhance the travel experience for corporate professionals in Karnataka, whether they're attending conferences or simply taking a break from work.

In the following sections, we'll dive into the methodology behind Getaway Guide, explaining its core modules and algorithms. By understanding how it handles user input, generates itineraries, and refines recommendations, we'll gain insight into how Getaway Guide transforms travel planning for corporate professionals in Karnataka. In this age of technological advancement, Getaway Guide represents a fresh approach to travel planning, providing personalized and efficient solutions for exploring Karnataka. With its combination of AI algorithms and user-friendly design, Getaway Guide aims to offer corporate professionals enriching travel experiences within the state.

2.LITERATURE REVIEW

Divya Mereddy et al.'s [1] study is based on the Hamilton graph using HDBSCAN clustering, with tourist limitations taken into consideration. A Hamilton tree is created using the geospatial distance between the points and the closest neighbours. After that, HDBSCAN clustering is used to divide the tree into clusters. Up until the cluster time spent is roughly equivalent to K hours, the data is further separated into sub-clusters. Therefore designing an itinerary plan that is optimized. Chao Yu et al. [2] present a personalized travel itinerary for a travellers with similar needs. The following factors are used to compute the similarity: "hotel," "restaurant," "travel time," and "attractions." Later, tour itineraries are customized for each group separately using online information resources. In the end, the cost of the tour is decided, usually in direct proportion to the number of participants, and includes a specified itinerary and route. Remigijus Paulavičius et al.'s [3] research determines which attractions are the most appropriate based on a number of parameters, including the maximum tour length, required visits, start and finish locations, and amount of time spent at comparable attractions. Finding optimal or nearly optimal

solutions is an NP-hard problem that is solved with a greedy genetic algorithm. The model is a useful tool for travellers as it offers highly customized travel experiences. Peilin Chen et al. [4] propose an artificial intelligence-based travel planning system. This study investigated the multi-threaded speed of the route search submodule and the webpage exploration tendency of the vertical search crawler. In addition, a stress test was run on this section to see how the system would function with more users logged in at once. The test findings show that the system fundamentally meets user requests and operates effectively in terms of efficiency and stability. Qiang Gao et al.'s [5] study presents a model which utilises the human mobility patterns to develop an efficient route. It consists of a trip encoder and a trip decoder. The Adversarial Net to create a rendition of a query. Aayushi Bhansali [6] aims to generate a real-time travel itinerary by considering travel time between sites of attraction and traffic congestion levels. Kristina P et al. [7] propose an unsupervised learning structure for the K-Means Algorithm which automatically finds an ideal value for K with no pioneer parameter definition. Anita Gehlot et al.'s [8] The study proposes a smart tourist system which provides travel recommendations in Sri Lanka, taking into consideration various factors like travel time, food, transport and boarding options, employing the Random Forest classifier to give an accuracy of 90%. Suresh Babu Dasari et al's [9] study proposes a personalised smart travel planner for vacation planning. Personalising the plan according to every user's individual preferences complies the use of a ybridd model. Scraping data from various official travel websites, the Gaussian Mixture Model and K-Means algorithm are used to cluster the tourist spots.

3.DATASET

In this section, we describe the datasets used in the development and testing of the Getaway Guide system. These are core datasets for the recommendation systems.

To enrich the geographical and tourism-based information available to Getaway Guide, we scraped data from a leading website that houses information on cities and tourist places in the region of Karnataka. That dataset contains basic information on the city name, its geographical coordinates, famous landmarks, and various places of interest in those cities. Additionally, the dataset contains information on the best time to visit those places and the average climatic conditions, which allow Getaway Guide to provide more informed recommendations for travel.

The scraping process entailed data extraction for each city webpage, parsing relevant HTML elements using Beautiful-Soup Library in Python, and storing the information in a structured format suitable for feeding into the Getaway Guide system.

4.METHODOLOGY

The methodology employed in this research encompasses a comprehensive approach to curating personalized travel itineraries through the amalgamation of data-driven techniques and machine learning algorithms. The process unfolds through a systematic series of steps, each meticulously designed to optimize the travel experience for users. Fig. 1 represents a flow of the methodology undertaken.

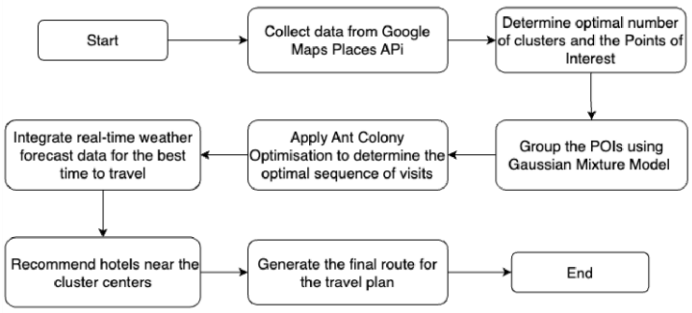

Fig. 1. Proposed pipeline for the methodology.

4.1 Data Collection

Reliable, broad-based data forms the core to the success of the proposed automated travel itinerary generation system. Geographical coordinates of Points of Interest are obtained from reputable sources like geographic databases or APIs, such as the Google Maps Places API. Real-time weather forecasts or historical weather data are also collected from reliable services in order to deliver accurate weather predictions for the travel period. Table 1 gives a glimpse of the curated dataset.

Table 1. Karnataka dataset curation

Index	cityName	latitude	longitude	bestTime	idealDuration
0	Coorg	12.32502	75.80545	Sep-Jun	2-3 days
1	Badami	15.91838	75.68033	Oct-Apr	1-2 days
2	Mysore	12.30902	76.65306	Throughout the year	1-2 days
3	Hampi	15.33306	76.45447	Oct-Mar	2-4 days
4	Gokarna	14.54887	74.32324	Oct-Mar	2-3 days

4.2 Clustering Phase

This is the clustering phase, which means segmenting the points of interest into meaningful clusters to enable itineraries to be organized. First, the determination of the optimal number of clusters is attained. This is done iteratively, using Geopy for distance calculations, Shapley values for feature importance, geometric principles, and efficient numerical computations using NumPy. By iteratively evaluating the average area covered by potential

clusters, the algorithm will converge towards identifying the optimal number of clusters. Subsequently, Gaussian Mixture Model (GMM) clustering is applied to group points of interest into clusters based on their geographical proximity and relevance. The mathematical formula given in equation (1) calculates the cluster density for each value of n as follows:

$$\text{cluster-density}[n] = \frac{\sum_{i=1}^{n}\left(\frac{\text{area-of-cluster}[i]-n}{\text{points-in-cluster}[i]-n}\right)}{n} \quad (1)$$

Where

- n - The number of clusters, ranging from 2 to 8.
- i - The index representing each cluster, ranging from 1 to n.
- area-of-cluster[i]n- The area of the i^{th} cluster
- points-in-cluster[i]n- The number of points in the i^{th} cluster.
- cluster-density[n] - The list storing the calculated cluster density for each n.

The optimal value of n for which the cluster density is minimum k can be determined using: Fig. 2 visualises the clusters generated for the collected data taking into consideration the city's latitude and longitude using Gaussian Mixture Model. The formula in equation (2) helps determine the optimal number of clusters based on the calculated cluster density.

$$k = \arg\min_n(\text{cluster-density}[n]) \quad (2)$$

Where k is the optimal number of clusters.

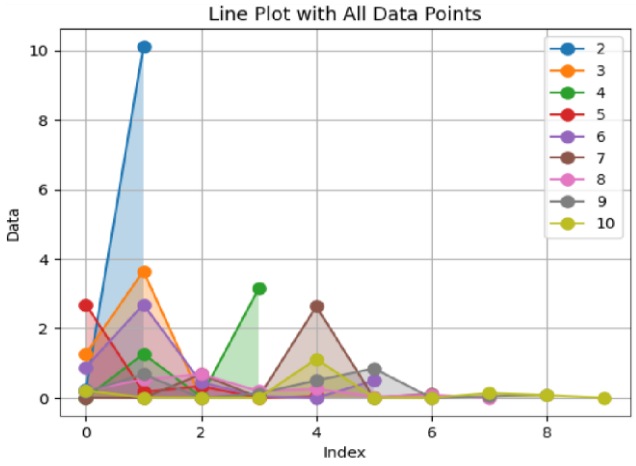

Fig. 2. Visual representation for selection of optimal number of clusters.

4.3 Ant Colony Optimisation

Ant Colony Optimization (ACO) is a bio-inspired algorithm employed in Getaway Guide to efficiently plan travel routes between clusters of Points-of-Interest (POIs) in Karnataka, India. Inspired by ant foraging behavior, ACO initializes artificial ants to explore potential paths, with each path representing a potential travel itinerary. As ants traverse paths between POIs, they deposit pheromones, influencing future path choices; higher pheromone levels indicate shorter, more efficient routes. Iteratively, the algorithm updates pheromone levels based on route success, reinforcing optimal paths and allowing less favorable paths to fade, preventing premature convergence. ACO integrates user preferences, such as travel dates and specific POIs, ensuring personalized itineraries that minimize travel time and distance. This approach enhances the travel experience for corporate professionals, providing seamless and tailored travel plans within Karnataka's diverse landscape of cultural, historical, and natural attractions. The Optimal Path generated using ACO is clusters 3, 1, 4, 0, 7, 2, 5, 6, 3 which gives the optimal distance of 53.8589 kilometers.

4.4 Weather Forecasting

Integrating weather forecasts into the itinerary planning process helps optimize travel plans by taking weather conditions into account. You can use either historical weather data or real-time forecasts to predict the weather during your trip. By including weather forecasts in the itinerary algorithm, you'll be able to anticipate and prepare for any weather-related disruptions or inconveniences. This proactive approach enhances your overall travel experience by minimizing the impact of bad weather on your plans, making it easier to transition between destinations and reducing the chances of unexpected delays.

Additionally, it gives you the information you need to make informed choices about outdoor activities, accommodations, and transportation, resulting in a smoother and more enjoyable journey. Ultimately, planning your itinerary with weather in mind allows for greater flexibility and resilience, so you can adapt to sudden changes and be prepared for any unexpected weather challenges that may come your way.

4.5 Hotel Recommendation

Cluster centers, which are basically the geographic centers of the clusters, are calculated to help recommend hotels. We use both geographic proximity and user preference to suggest hotels around these cluster centers to travelers. This recommendation strategy makes it easier to offer accommodation options that are closer to these popular points of interest. This not only adds convenience but also minimizes travel logistics. By considering factors such as price, amenities, and user reviews, our hotel

recommendation feature further personalizes your travel itinerary based on your preferences and requirements. This ensures that you have access to strategically located accommodations near key attractions, enhancing your overall travel experience. It also saves you time and resources by reducing unnecessary travel between different points of interest. By incorporating real-time data on hotel availability and pricing, our feature helps you find the best deals and options that match your budget and preferences. It provides a comprehensive and user-centered approach to travel planning. Moreover, it promotes sustainable travel by encouraging stays in central locations, reducing the carbon footprint associated with extensive commuting.

4.6 Route Generation

The HERE API is used to generate routes, which is an incredibly strong and reliable platform for routing and naviga- tion services. Within every cluster, we have optimized travel routes, taking into consideration the capabilities presented by the HERE API. The API allows users to access a variety of ge- ographic information, including realtime data, road networks, points of interest, and much more, in generating specific efficient routes in response to a traveler's needs. This system employs HERE API to dynamically compute the best routes based on several factors: distance, current traffic conditions, and user preference. Then it connects the generated routes seamlessly with other travel itinerary details, giving users the chance to move with much ease while traveling around new places. Each traveler who is situated within any clustered area would be in a position to efficiently and effectively view.

5. RESULTS

To evaluate the effectiveness of the Getaway Guide system, we conducted a series of tests focusing on the accuracy of the personalized itineraries generated and the efficiency of the travel paths.

5.1 Path Optimsation using ACO

The Ant Colony Optimization (ACO) algorithm was em- ployed to determine the optimal sequence of visits to the clus- tered points of interest (POIs) in Karnataka. The ACO's ability to consider geographical distances, user-specified dates, and other constraints resulted in highly efficient travel itineraries. Below is a detailed example of the path optimization process using ACO and the HERE API for routing. In this figure, the green markers represent the cluster center of each of the clusters consisting of the POIs, and the blue line indicates the optimized travel path connecting these points. The algorithm effectively minimizes travel distance and time, ensuring a seamless travel experience.

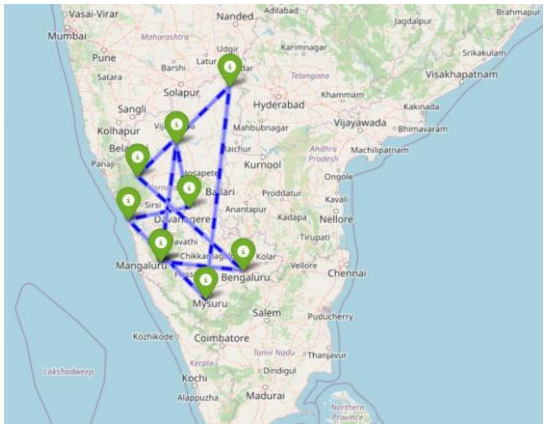

Fig. 3. Optimal Path

5.2 Clustering Analysis

The clustering phase involved segmenting POIs into meaningful clusters. The Gaussian Mixture Model (GMM) was applied to group POIs based on their geographical proximity and relevance. Below is a visual representation of the clustering results.

Table 2. Assigning cluster labels

Index	cityName	latitude	longitude	bestTime	idealDuration	cluster
0	Coorg	12.32502	75.80545	Sep-Jun	2-3 days	3
1	Badami	15.91838	75.68033	Oct-Apr	1-2 days	2
2	Mysore	12.30902	76.65306	Throughout the year	1-2 days	1
3	Hampi	15.33306	76.45447	Oct-Mar	2-4 days	7
4	Gokarna	14.54887	74.32324	Oct-Mar	2-3 days	0
...	...	...	...	...	...	...
62	Mekedatu	12.26646	77.44507	Sep-Dec	1 day	1
63	Gulbarga	17.33282	76.84033	Oct-Mar	1 day	6
64	Sakleshpur	12.94085	75.78161	Oct-Mar	2-3 days	3
65	Kasarkod-beach	14.25615	74.44368	Nov-Mar	1 day	0
66	Padubidri-beach	13.12437	74.77436	Nov-Mar	1 day	3

6. CONCLUSION

Regarding contemporary travel planning, the Getaway Guide system stands out because of its effectiveness and convenience. The approach saves travelers a lot of time and effort because it uses reliable data sources as well as advanced algorithms such as GMM and ACO to streamline itinerary building process. It is a very systematic way of planning an itinerary by incorporating hotel recommendations, route development, weather predictions, clustering and optimization into one. It is convenient in the sense that each travel itinerary made using this method is tailored to meet the specific needs and interests of the

individual traveler. This enhances convenience through providing HERE API's comprehensive navigation instructions which easily enables any tourists reach new places effortlessly. Taking everything into consideration, the recommended methodology offers a basic and effective means for planners to organize their trips thus enabling them to have unique trips with ease. That makes it a priceless asset for presentday tourists.

7. FUTURE SCOPE

The Getaway Guide platform is looking into using Localised Content and Language Support to better cater to a wider user base. Expanding the availability of Getaway Guide's content to include more languages and dialects, particularly regional languages spoken in Karnataka and other target areas, could help serve more people. By collaborating with linguists, cultural experts, and local translators, Getaway Guide aims to ensure accurate localization of information and recommendations that respect cultural differences, ultimately enhancing user engagement and satisfaction.

To simplify the process of booking accommodations, travel, and experiences directly through Getaway Guide, the platform plans to efficiently integrate third-party booking platforms such as hotel reservation systems, airline booking engines, and activity booking websites. By establishing robust APIs and forging partnerships with leading reservation companies, Getaway Guide aims to streamline transactions and enhance user convenience.

REFERENCES

[1] Jayaram Nori*1 *1Broadcom Inc, USA. DOI :- https://www.doi.org/10.56726/IRJMETS53503

[2] P. Prasad and A. Rao, "A survey on various challenges and issues in implementing AI for enterprise monitoring," Journal of Network and Computer Applications, vol. 116, pp. 42-55, 2018, doi: 10.1016/j.jnca.2018.05.005.

[3] Y. Dang, Q. Lin, and P. Huang, "AIOps: real-world challenges and research innovations," in 2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion), 2019, pp. 4-5, doi: 10.1109/ICSE-Companion.2019.00023.

[4] D. Xu et al., "Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications," in Proceedings of the 2018 World Wide Web Conference, 2018, pp. 187-196, doi: 10.1145/3178876.3185996.

[5] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Computing Surveys (CSUR), vol. 41, no. 3, pp. 1-58, 2009, doi: 10.1145/1541880.1541882.

[6] M. Chen et al., "Big data: A survey," Mobile Networks and Applications, vol. 19, no. 2, pp. 171-209, 2014, doi: 10.1007/s11036-013-0489-0.

[7] Y. Li et al., "Deep learning for anomaly detection in cloud native systems," in 2020 IEEE International Conference on Cloud Engineering (IC2E), 2020, pp. 106-116, doi: 10.1109/IC2E48712.2020.00022.

[8] F. Salfner, M. Lenk, and M. Malek, "A survey of online failure prediction methods," ACM Computing Surveys (CSUR), vol. 42, no. 3, pp. 1-42, 2010, doi: 10.1145/1670679.1670680.

[9] F. Jiang et al., "Artificial intelligence in healthcare: Past, present and future," Stroke and Vascular Neurology, vol. 5, no. 2, 2020, doi: 10.1136/svn-2020-000443.

[10] F. Salfner, M. Lenk, and M. Malek, "A survey of online failure prediction methods," ACM Computing Surveys (CSUR), vol. 42, no. 3, pp. 1-42, 2010, doi: 10.1145/1670679.1670680.

[11] X. Liu et al., "PANDA: Facilitating usable AI development," arXiv preprint arXiv:2003.04070, 2020.

[12] G. A. Susto, A. Beghi, and C. De Luca, "A predictive maintenance system for epitaxy processes based on filtering and prediction techniques," IEEE Transactions on Semiconductor Manufacturing, vol. 25, no. 4, pp. 638-649, 2012, doi: 10.1109/TSM.2012.2209131.

[13] X. Liu et al., "PANDA: Facilitating usable AI development," arXiv preprint arXiv:2003.04070, 2020.

[14] E. Cortez et al., "Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms," in Proceedings of the 26th Symposium on Operating Systems Principles (SOSP), 2017, pp. 153-167, doi: 10.1145/3132747.3132772.

[15] Z. Yin et al., "An empirical study on configuration errors in commercial and open source systems," in Proceedings of the 26th Symposium on Operating Systems Principles (SOSP), 2017, pp. 159-176, doi: 10.1145/3132747.3132773.

[16] D. Wang et al., "Failure prediction using machine learning in a virtualised HPC system and application," Cluster Computing, vol. 20, no. 1, pp. 103-115, 2017, doi: 10.1007/s10586-016-0668-4.

[17] J. Gao, "Machine learning applications for data center optimization," Google White Paper, 2014. [Online]. Available: https://research.google/pubs/pub42542/

[18] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in 2010 IEEE Symposium on Security and Privacy, 2010, pp. 305-316, doi: 10.1109/SP.2010.25.

[19] R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," Journal of Internet Services and Applications, vol. 9, no. 1, pp. 1-99, 2018, doi: 10.1186/s13174-018-0087-

[20] A. Mestres et al., "Knowledge-defined networking," ACM SIGCOMM Computer Communication Review, vol. 47, no. 3, pp. 2-10, 2017, doi: 10.1145/3138808.3138810.