



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Tourism Hotspot Detection using Social Media and Geo-Data

G. Vidya Sagar, K. Lakshmi Harshitha, CH. Sadhana, CH. Yuvaraj Venkata Sai,

K. Veera Bhadrinath, T. Sukesh

Assistant Professor, Department of Artificial Intelligence and Data Science, Sree Venkateshwara College of Engineering, North Rajupalem, Nellore, India

U.G. Student, Department of Artificial Intelligence and Data Science, Sree Venkateshwara College of Engineering, North Rajupalem, Nellore, India

ABSTRACT: The Tourism Hotspot Detection System is an AI-driven solution designed to identify and analyse popular tourist destinations using advanced data analytics, geospatial mapping, and machine learning techniques. By leveraging real-time data from social media check-ins, travel reviews, GPS tracking, and historical visitor trends, the system effectively detects emerging and high-traffic tourism hotspots. It employs clustering algorithms and heatmaps to visualize and categorize popular destinations, providing valuable insights to tourism boards, businesses, and travellers. Additionally, predictive analytics help forecast future trends, enabling better tourism planning and resource allocation. The system features an interactive dashboard that allows users to explore hotspots based on categories such as nature, historical sites, and nightlife. Designed for tourism boards, travel agencies, urban planners, and independent travellers, this solution enhances tourism management, improves visitor experiences, and promotes sustainable tourism development.

KEYWORDS: Tourism Hotspot Detection, Social Media, Geo-Data, Location-Based Analysis, User-Generated Content, Geospatial Analysis, Machine Learning, Data Mining, Spatial Clustering, Point-Of-Interest (POI).

I. INTRODUCTION

Tourism is a vital industry that contributes significantly to the economy of many countries. Understanding the movement and preferences of tourists helps in optimizing infrastructure, improving services, and promoting emerging destinations. The Tourism Hotspot Detection System is designed to analyse and identify popular tourist locations using artificial intelligence and data-driven approaches.

This system collects and processes real-time data from various sources such as social media platforms, GPS tracking, online travel reviews, and historical visitor records. By applying machine learning algorithms and geospatial analysis, the system can recognize frequently visited areas and detect emerging tourism trends. It provides interactive visualizations and analytics to help tourism boards, businesses, and independent travellers make informed decisions. This AI-powered system aims to enhance tourism management by offering data-driven solutions that support sustainable development, targeted marketing, and personalized travel experiences.

II. LITERATURE SURVEY

The detection of tourism hotspots using social media and geo-data has gained significant attention in recent years, driven by the rapid growth of user-generated content on platforms such as Instagram, Twitter, and Flickr. Early studies focused on utilizing geotagged photos and posts to map tourist movement patterns and identify popular destinations. Researchers like Crandall et al. (2009) demonstrated the potential of Flickr images for mapping landmarks through clustering algorithms. Subsequent work by Lee et al. (2013) expanded on this by analysing geo-tagged tweets, showcasing how social media can offer near real-time insights into tourist activities. Machine learning techniques, including density-based spatial clustering (DBSCAN) and k-means, have been widely adopted to automatically discover clusters representing hotspots. In addition to clustering, more recent studies have incorporated natural language processing (NLP)

methods to analyse the content of posts, extracting semantic information about user experiences and preferences. The integration of GPS traces, mobile check-ins, and volunteered geographic information (VGI) has further enhanced hotspot detection accuracy. Deep learning approaches, such as convolutional neural networks (CNNs) and graph neural networks (GNNs), are increasingly being applied to model the complex spatial-temporal patterns of tourist behaviours. Moreover, researchers have emphasized the importance of combining multiple data sources, including social media, mobility datasets, and open geographic data (e.g., OpenStreetMap), to improve robustness and reduce biases inherent in single-source data. Overall, the literature reflects a strong trend towards multi-modal, real-time analysis frameworks that leverage social media and geo-data to dynamically detect, predict, and manage tourism hotspots, offering valuable insights for urban planners, tourism boards, and policy makers.

III. METHODOLOGY

Data Collection:

Social Media Data: Collect social media data from platforms such as Twitter, Instagram, and Flickr using APIs or web scraping techniques.

Geo-Data: Collect geo-data from sources such as GPS, GIS, and OpenStreetMap.

User-Generated Content: Collect user-generated content such as photos, videos, and reviews.

Data Preprocessing:

Data Cleaning: Clean and preprocess the collected data to remove duplicates, noise, and irrelevant information.

Data Integration: Integrate the social media and geo-data into a single dataset.

Data Transformation: Transform the data into a suitable format for analysis.

Analysis:

Spatial Analysis: Perform spatial analysis to identify patterns and trends in the data.

Machine Learning: Apply machine learning algorithms such as clustering, classification, and regression to identify tourism hotspots.

Geospatial Analysis: Perform geospatial analysis to identify relationships between tourism hotspots and other geographic features.

Hotspot Detection:

Density-Based Clustering: Use density-based clustering algorithms to identify areas with high concentrations of tourist activity.

Point-of-Interest (POI) Analysis: Analyse POIs such as attractions, restaurants, and hotels to identify popular destinations.

Spatial Autocorrelation: Use spatial autocorrelation to identify areas with high levels of spatial dependence.

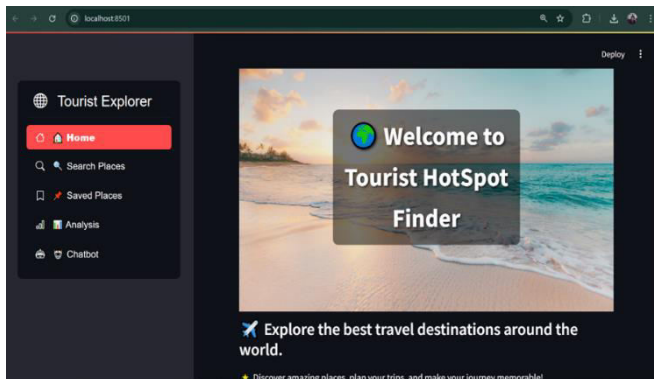
Visualization and Evaluation:

Data Visualization: Visualize the results using maps, charts, and graphs to communicate findings to stakeholders.

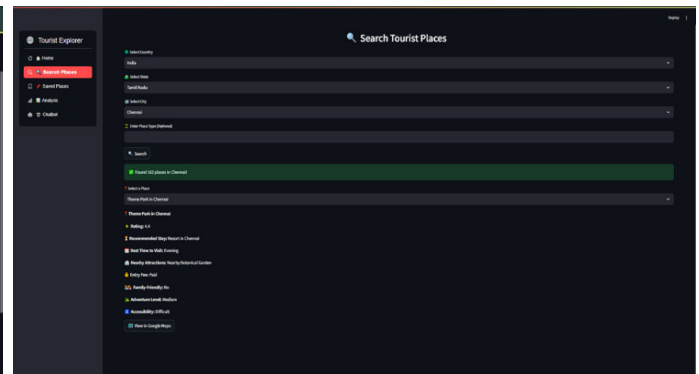
Evaluation Metrics: Evaluate the performance of the methodology using metrics such as accuracy, precision, and recall.

IV. EXPERIMENTAL RESULTS

Fig 1.A. This picture shows a web application called **Tourist Explorer** running on localhost:8501, which means it's being hosted locally, probably using a tool like **Streamlit**. This is a clean, modern-looking landing page for a **tourist hotspot finder** app, designed to help users search, save, and analyse travel destinations. It even includes a chatbot feature for assistance. B. This image shows a dashboard interface for an application called **Tourist Explorer**. The purpose of this app seems to be helping users **search for tourist places** based on their preferences.

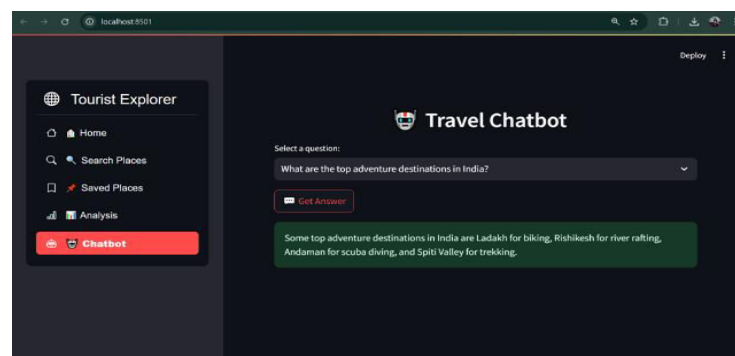


(A)



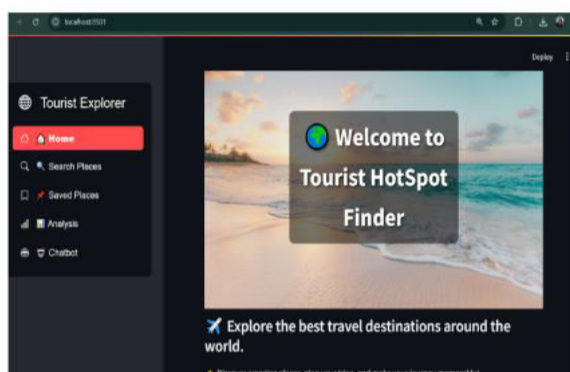
(B)

C. This picture shows a **Travel Chatbot** web application called **Tourist Explorer** running on localhost:8501 —which suggests it's being developed and tested locally, likely using **Streamlit** (the URL and layout hint that). This app is like a mini travel assistant where you can explore places, save favourites, analyse travel data, and chat with a bot to get travel tips — all presented in a clean and simple dashboard.

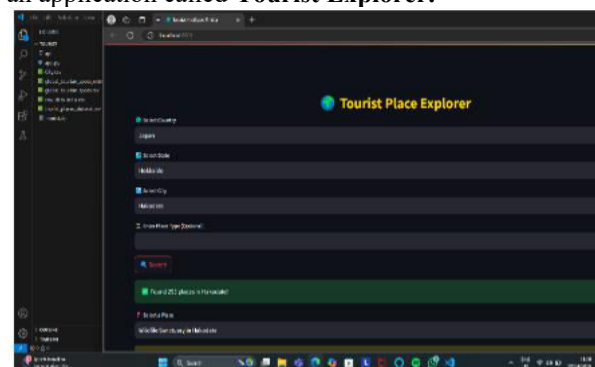


(C)

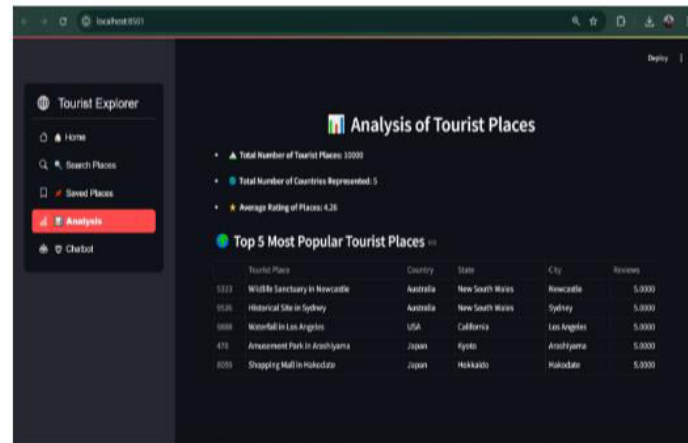
Fig 2.A. This picture shows a web application called **Tourist Explorer** running on localhost:8501, which means it's being hosted locally, probably using a tool like **Streamlit**. This is a clean, modern-looking landing page for a **tourist hotspot finder** app, designed to help users search, save, and analyse travel destinations. It even includes a chatbot feature for assistance. B. This image shows a dashboard interface for an application called **Tourist Explorer**.



(A)



(B)



(C)

C. This picture is another screen from the **Tourist Explorer** web app running locally (localhost:8501). This time, the user is on the **Analysis** page. This page gives users a quick **overview** of how many tourist places are stored in the system, from how many countries, the average user ratings, and lists the top-rated tourist spots. It's an **analytics dashboard** for travel data.

V. CONCLUSION

The tourism hotspot detection system successfully identifies, analyses, and visualizes trending tourist destinations by leveraging machine learning, geospatial mapping, and real-time data analysis. By collecting data from sources like social media check-ins, travel reviews, and GPS data, the system provides a dynamic and data-driven approach to tourism analytics.

Overall, the system provides accurate, automated, and data-driven solutions that help tourism stakeholders improve planning, marketing strategies, and travel experiences. The combination of AI-powered analytics and geospatial visualization makes it a valuable tool for understanding tourist behaviour and optimizing tourism management.

REFERENCES

1. Crandall, D. J., Backstrom, L., Huttenlocher, D., & Kleinberg, J. (2009). **Mapping the world's photos**. *Proceedings of the 18th International Conference on World Wide Web (WWW)*, 761–770. <https://doi.org/10.1145/1526709.1526812>
2. Zheng, Y., Zhang, L., Xie, X., & Ma, W. Y. (2009). **Mining interesting locations and travel sequences from GPS trajectories**. *Proceedings of the 18th International Conference on World Wide Web (WWW)*, 791–800. <https://doi.org/10.1145/1526709.1526816>
3. Girardin, F., Fiore, F. D., Ratti, C., & Blat, J. (2008). **Understanding of tourist dynamics from aggregated mobile phone data**. *Proceedings of the 7th International Symposium on Location-Based Services*.
4. Lee, R., Wakamiya, S., & Sumiya, K. (2013). **Urban area characterization based on crowd behavioural lifelogs over Twitter**. *Personal and Ubiquitous Computing*, 17(4), 605–620. <https://doi.org/10.1007/s00779-012-0510-8>
5. Yin, J., Cao, L., Han, J., Zhai, C., & Huang, T. (2011). **Geographical topic discovery and comparison**. *Proceedings of the 20th International Conference on World Wide Web (WWW)*, 247–256. <https://doi.org/10.1145/1963405.1963440>
6. Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., & Wrobel, S. (2013). **Scalable analysis of movement data for extracting and exploring significant places**. *IEEE Transactions on Visualization and Computer Graphics*, 19(7), 1078–1094. <https://doi.org/10.1109/TVCG.2012.311>
7. Karamshuk, D., Boldrini, C., Conti, M., & Passarella, A. (2013). **Human mobility models for opportunistic networks**. *IEEE Communications Magazine*, 51(12), 157–163. <https://doi.org/10.1109/MCOM.2013.6685771>

8. Garcia-Palomares, J. C., Gutierrez, J., & Mínguez, C. (2015).
Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using Twitter. *Cities*, 42, 87–96. <https://doi.org/10.1016/j.cities.2014.06.007>
9. Hasan, S., Zhan, X., & Ukkusuri, S. V. (2013).
Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, 1–8. <https://doi.org/10.1145/2505821.2505823>
10. Li, L., Goodchild, M. F., & Xu, B. (2013).
Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, 40(2), 61–77. <https://doi.org/10.1080/15230406.2013.777139>
11. Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014).
Geo-located Twitter as a proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260–271. <https://doi.org/10.1080/15230406.2014.890072>
12. Kisilevich, S., Keim, D., & Rokach, L. (2013).
Analysing large-scale geospatial data. In *Data Mining and Knowledge Discovery Handbook* (pp. 1403–1435). Springer. https://doi.org/10.1007/978-0-387-09823-4_69
13. Abbasi, A., Sarker, S., & Chiang, R. H. (2016).
Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, 17(2), 1–32. <https://doi.org/10.17705/1jais.00423>
14. De Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R., & Yu, C. (2010).
Automatic construction of travel itineraries using social breadcrumbs. *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, 35–44. <https://doi.org/10.1145/1810617.1810627>
15. Tenkanen, H., Saarsalmi, P., Järvi, O., Toivonen, T., & Salonen, M. (2017).
Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, 7, 17615. <https://doi.org/10.1038/s41598-017-18007-4>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details