Asia Pacific Academy of Science Pte. Ltd.

Smart Tourism 2024, 5(1), 2526.
https://doi.org/10.54517/st.v5i1.2526

Article

A seven-year analysis of the tourism sector of India with a prediction model—A case study of smart tourism

Swati Lipsa*, Ranjan Kumar Dash

School of Computer Sciences, Odisha University of Technology and Research, Bhubaneswar 751029, India
*** Corresponding author:** Swati Lipsa, slipsait@outr.ac.in

CITATION

Lipsa S, Dash RK. A seven-year analysis of the tourism sector of India with a prediction model—A case study of smart tourism. Smart Tourism. 2024; 5(1): 2526. https://doi.org/10.54517/st.v5i1.2526

ARTICLE INFO

Received: 29 January 2024
Accepted: 18 March 2024
Available online: 30 April 2024

Abstract: The tourism sector requires in-depth analysis and forecasting to provide a clear picture of various factors that affect the visits of foreign tourists to certain countries. In this context, the work carried out in this paper provides an in-depth analysis of the number of tourists to India and the revenue generated from them from the years 2014 to 2020. Furthermore, the analysis of the different states of India to which the tourists visited the most and the quarterly analysis of the tourists to India are also presented. The impact of the corona pandemic on the tourism sector of India is also shown by comparing the number of tourists in 2019 and 2020. Support vector regression (SVR) is trained with historical data on the number of tourists from 2001 to 2016 and validated for 2017 to 2019. This trained model is used to forecast the number of tourists from 2020 to 2023 to study the impact of corona pandemic on the number of foreign tourists to India. Similarly, historical data on foreign exchange fees from 2001 to 2016 is used to train the model, which is validated with data from 2017 to 2019. This train model is used to predict the Foreign Exchange Earning (FEE) for the years 2020 to 2022. The actual FEE is compared with the predicted FEE to show the impact of the coronavirus pandemic on the revenue generated from tourism in India.

Keywords: smart tourism; foreign exchange money; machine learning; support vector regression; foreign tourists

1. Introduction

Smart tourism refers to the use of technology and data to enhance the overall tourism experience for travelers. It's a term that describes the use of information and communication technologies (ICT) to improve the accessibility, convenience, and sustainability of tourism by offering real-time information, personalized recommendations, seamless booking processes, and interactive experiences. Smart tourism can benefit both tourists and tourism destinations by enhancing their convenience, satisfaction, safety, and environmental awareness. Smart tourism encompasses several key features, including cutting-edge infrastructure that ensures sustainable development and equal access for all, as well as initiatives to promote more sustainable tourism by minimizing waste, energy consumption, pollution, and congestion. It offers free Wi-Fi access in public areas and streets and adopts electric mobility as a viable substitute for conventional transportation methods. Additionally, it provides real-time updates on public transportation delays or incidents, as well as cultural and interactive events that utilize virtual reality (VR) and augmented reality (AR) technology to offer tourists immersive and captivating experiences.

These features can help tourism destinations increase their competitiveness by offering unique value propositions to tourists based on data-driven insights, enhancing their attractiveness by creating memorable experiences for tourists that meet their

needs and preferences and fostering their sustainability by promoting social responsibility, environmental protection, and community engagement. As smart tourism is a big business in the travel and tourism industry, it also requires a lot of innovation, collaboration, and leadership from various stakeholders.

Considering that the field of smart tourism is burgeoning and evolving, there is a dire necessity to analyze smart tourism in various aspects to understand the current situation, challenges, opportunities, and trends of smart tourism development and management. The various aspects to be analyzed include the number of visitors, place of visit, revenue generated, effect of the pandemic, time of visit, etc. These aspects can help to measure the performance, impact, and potential of smart tourism destinations and services. Some possible methods to analyze smart tourism are data analysis that can help to identify patterns, trends, correlations, and causal relationships among different variables related to smart tourism [1], survey analysis that helps to evaluate the effectiveness and quality of smart tourism offerings and activities [2], case study analysis that can be utilized to learn from best practices and lessons learned from real-world scenarios [2,3] and benchmarking analysis to identify strengths and weaknesses of smart tourism development and management [4].

Even though the above-mentioned methods render insightful analysis, they can be refined further by the use of machine learning algorithms. Some possible ways to use machine learning for smart tourism analysis are classification used to identify tourists and residents based on their photo content [4], or to predict the optimal pricing strategy for smart tourism products and services based on demand and supply factors [5], clustering used to segment tourists based on their preferences, behaviour, or feedback [6], or to identify hotspots or trends in tourist demand [7], and regression used to forecast the revenue generated by smart tourism activities based on historical data and external factors [7], or to measure the impact of the pandemic on tourist satisfaction and loyalty [6].

Despite the available analytical methodologies, research in this domain is still in its nascent stages. For instance, a key component for theoretical development, i.e., the definition of smart tourism, appears to not be universally agreed upon. Furthermore, it appears that practitioners exhibit inconsistency in their interpretation of what defines a smart property, such as a smart hotel. Due to the rapid progress of intelligent technologies and their applications, it is crucial to conduct a thorough evaluation of the present literature to lay a solid foundation and offer direction for future research. This study intends to offer a comprehensive analysis of the smart tourism literature, shedding light on research trends, topics, locations, theories, methodologies, and industry applications.

2. Related works

The study of Ghorbani et al. [8] marks the first endeavour to formulate a smart tourism organization (STO) and delineate its various dimensions. The insights gleaned from this exploration offer valuable assistance to tourism organization managers and fellow researchers in enhancing their understanding and fostering intelligence within tourism organizations, particularly in situations where a dearth of specialized research exists on tourism organization and management. The analysis of current smart tourism

research [9] revealed 11 prominent themes. Additionally, the analyses of co-citations also provided visual representations of four areas of knowledge that strengthen the current literature and enhance the insights gained from theme analysis. Akdu [10] explains the term "smart," its associated technology, and how it affects the travel industry. It also covers smart tourist locations. The paper by Wang et al. [11] suggests harnessing the capabilities of 5G and AI to tackle challenges related to implementing the Internet of Things (IoT) in smart tourism. Subsequently, the effectiveness of 5G and AI is evaluated by unlocking the full potential of IoT within the realm of smart tourism. Cavalheiro et al. [12] introduce a model, termed the Smart Tourism Destination Development Model, designed to outline a strategic trajectory for a tourism destination aspiring to embrace smart features. The study delves into the concept of STDs and asserts that, in addition to elevating destination competitiveness, an STD initiative should be rooted in a sustainable paradigm to generate public value for the host community. Pai et al. [13] focused on how visitors' satisfaction and likelihood of returning were affected by their experience with smart tourism technologies. Among the factors influencing the smart tourism technology experience, accessibility is the most important component, while personalization is not of much significance. The paper further finds a strong connection between the use of smart tourism technology and travel satisfaction. This travel satisfaction had a beneficial impact on both the happiness of tourists and their inclination to visit again.

Bastidas-Manzano et al. [14] use a bibliometric method that offers significant findings for public organizations and tourist entities, assisting them in identifying critical aspects to prioritize during the evolution of smart tourism. Corrêa et al. [15] put forth a comprehensive theory elucidating the genesis of smart tourism experiences (STEs) within the framework of smart tourism destinations (STDs), as perceived by travelers. The STE is conceived through the dynamic interplay between travelers and various stakeholders within the STD framework, facilitated by the utilization of technology for information sharing and personalized experiences. Gretzel [16] suggests adopting a utopian perspective when observing the growth of smart tourism, stating that a critical evaluation of the existing condition and a collaborative imagining of better tourism and destinations are required. This approach could potentially elevate smart tourism initiatives beyond their current instrumental, short-term, and fragmented nature. Additionally, the paper puts forth the idea of a smart tourism mindset, stating that the idealistic goal of smart tourism ought to be moulded by unique traits and principles that impact people on every level. Muniz et al. [17] illustrate that destination management organisations (DMOs) can benefit from customer knowledge management (CKM) in smartly managing the visitor experience. The objective is to exert a positive impact on the development of innovative solutions and the progress of STDs. In order to achieve this goal, a conceptual structure for CKM, consisting of eight processes, is formulated, providing guidance to managers in implementing this approach. Errichiello et al. [18] create a framework for the governance process that is specific to smart tourist destinations by integrating the modern smart approach with the notion of destination governance. This framework serves as a versatile tool for policymakers and destination managers since it shows how to tailor the creation of governance systems and procedures to each specific location. It also outlines the steps to make use of "smart dimensions" for development, with a progressive logic that

moves through successive, interconnected phases.

Borges-Tiago et al. [19] carry out a scientometric assessment of smart tourism that utilizes a classification framework to identify new trends and potential areas for further research. While the majority of works predominantly revolve around smart ecosystems and technologies, this underscores the imperative to enhance our understanding of other thematic areas. In order to increase tourism, Novera et al. [20] apply bibliometric analysis based on citations and text mining to review previous studies that addressed the impact of IoT implementation on the growth of the tourism industry. The study uses log-likelihood analysis to find nine relevant thematic models; each model includes a summary of previous works, a list of main authors, and posterior probability derived from latent Dirichlet allocation. Kontogianni et al. [21] develop an innovative cloud-based architecture that leverages image labelling through deep learning and neural network-based collaborative filtering models with the aim of providing personalized suggestions within the realm of smart tourism. Furthermore, the study outlines the architecture/topology of artificial neural network (ANN) models employed for predicting tourists' preferences and the experimental results, highlighting the configuration that yielded the most accurate forecasts. Balakrishnan et al. [22] propose a conceptual model by synthesizing essential principles obtained from the elaboration likelihood model (ELM) and flow theory. The findings revealed the crucial role of smart tourism technology (STT) in shaping tourists' experiences and influencing their intention to revisit.

Though numerous strategies have been put forth to differentiate between visitors and natives, such as supervised machine learning (ML) algorithms and heuristic techniques (those based on stay durations), these methods have been shown to have a couple of limitations. First, it is not feasible to assess and validate the results of heuristic techniques. Furthermore, there is disagreement on the best method for establishing the criteria (like the minimal duration of stay) and providing the rationale for their values. Two, when using ML-based algorithms, researchers overlook aspects like weather conditions, population size, and blog posts that could influence the differentiation between tourists and locals.

In an effort to overcome these limitations, the current study develops a regression-based machine-learning method for differentiating between tourists and natives.

3. Materials and methods

The data for this analysis was collected from these studies [23,24]. The data [24] is real-time, authentic data prepared by the Government of India and made available on their website. The important collected data for a seven-year duration, i.e., from 2014 to 2020, includes:

1) Number of foreign tourists from different countries to India.
2) Revenue generated by India from tourists as foreign exchange fees (FEE).
3) The effects of the coronavirus epidemic on India's travel industry.
4) Forecasting the number of tourists and revenue using a machine-learning approach.

The above-mentioned data are analyzed in depth, which is indeed helpful to the

tourists for their seamless visit to India, while the analysts from India can get a comprehensive idea of the number of inbound tourists as well as the revenue generated from tourism. Further, it also helps India take the necessary remedial steps to attract more tourists and increase its economic growth.

3.1. Analysis of the number of foreign tourists from different countries to India

3.1.1. Number of tourists to India between 2014 to 2020

The study on the count of foreign tourists to India is important as this directly affects the country's economy. The main source of revenue, **Figure 1**, depicts the number of tourists to India over a period from 2014 to 2020. This number rises annually, but it experienced a significant decline in 2020, primarily attributed to the worldwide repercussions of the COVID-19 pandemic.

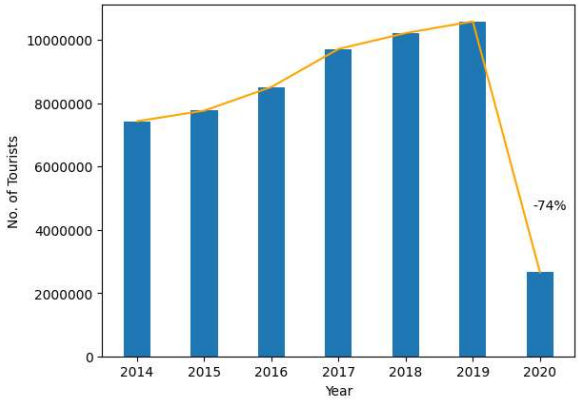

Figure 1. Number of inbound tourists to India from 2014 to 2020.

3.1.2. Number of inbound tourists from different countries

India attracts tourists due to many factors, including the weather, environment, and historical heritage. So, it is indeed important to analyze the inbound tourists from different countries. **Figure 2** provides an analysis of inbound tourists from the top ten countries. The years 2014 and 2015 record the United States as the top country from which the most tourists migrate to India. For the rest of the years, Bangladesh replaced the United States at the top. The United Kingdom enjoys a consistent third place on this list. The rest of the places are mainly filled out by countries like Canada, Sri Lanka, Australia, Malaysia, Russia, Germany, Japan, etc.

3.1.3. The inbound tourists to different states of India

It is equally important to analyze the preferences of inbound tourists to different states of India (**Figure 3**). Most tourists prefer to visit either Tamil Nadu or Uttar Pradesh, while Andhra Pradesh, Karnataka, and Maharashtra become their next choices interchangeably. Other Indian states that are preferred by tourists are Telangana, Madhya Pradesh, West Bengal, Rajasthan, and Gujarat during this period of analysis.

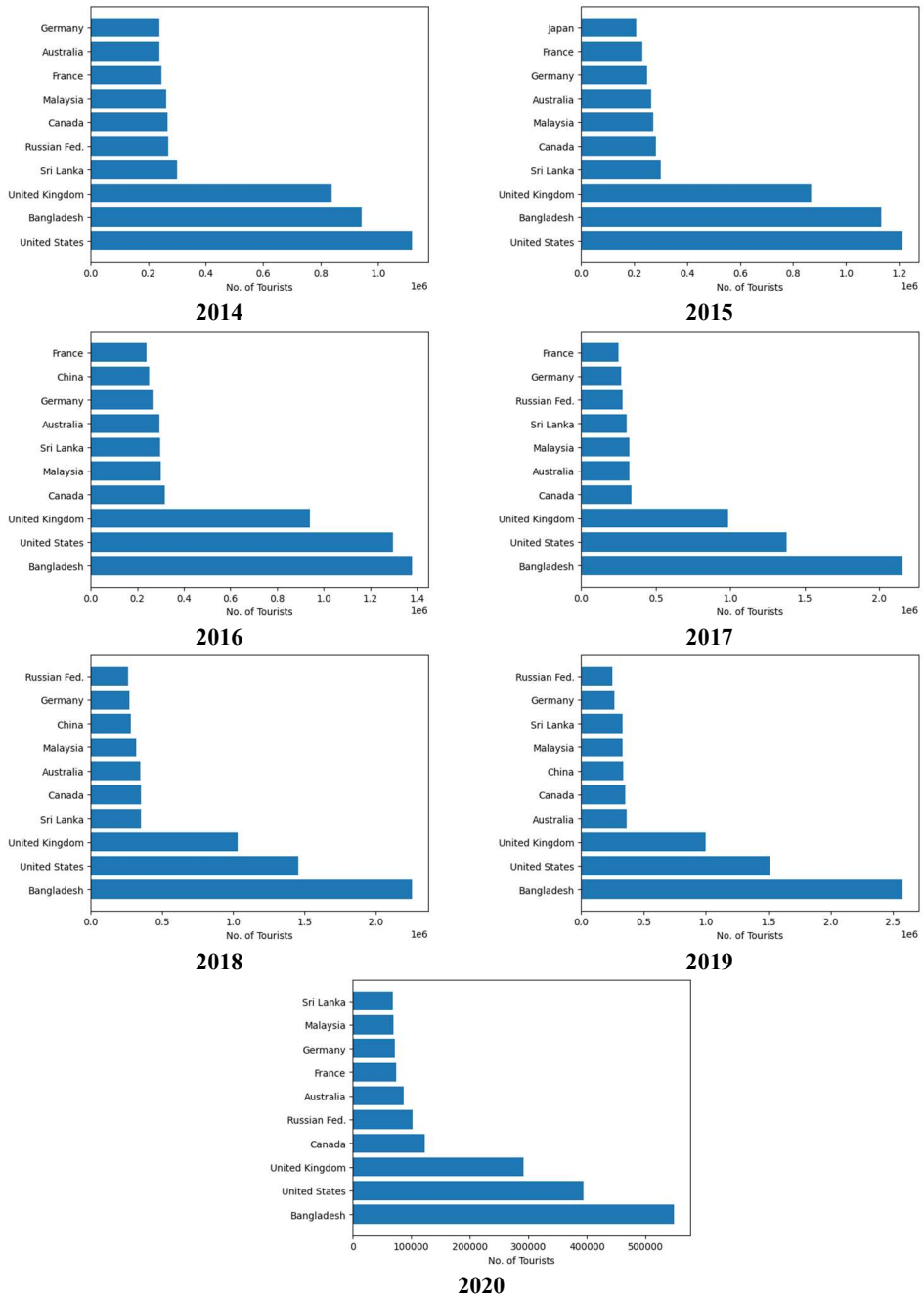

Figure 2. Number of inbound tourists from the top ten countries from 2014 to 2020.

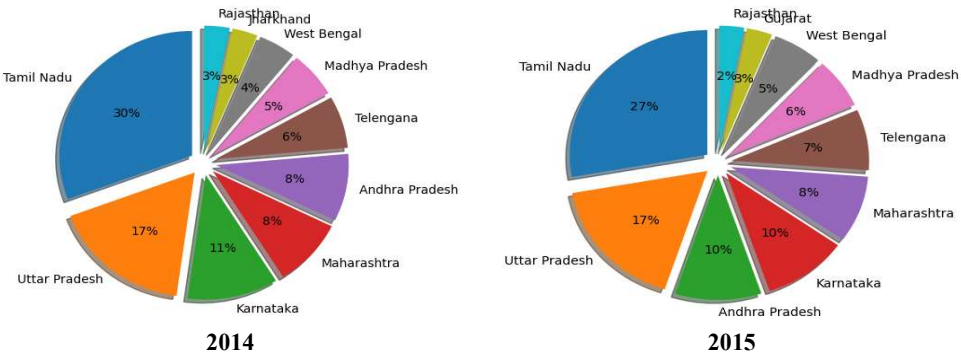

Figure 3. (*Continued*).

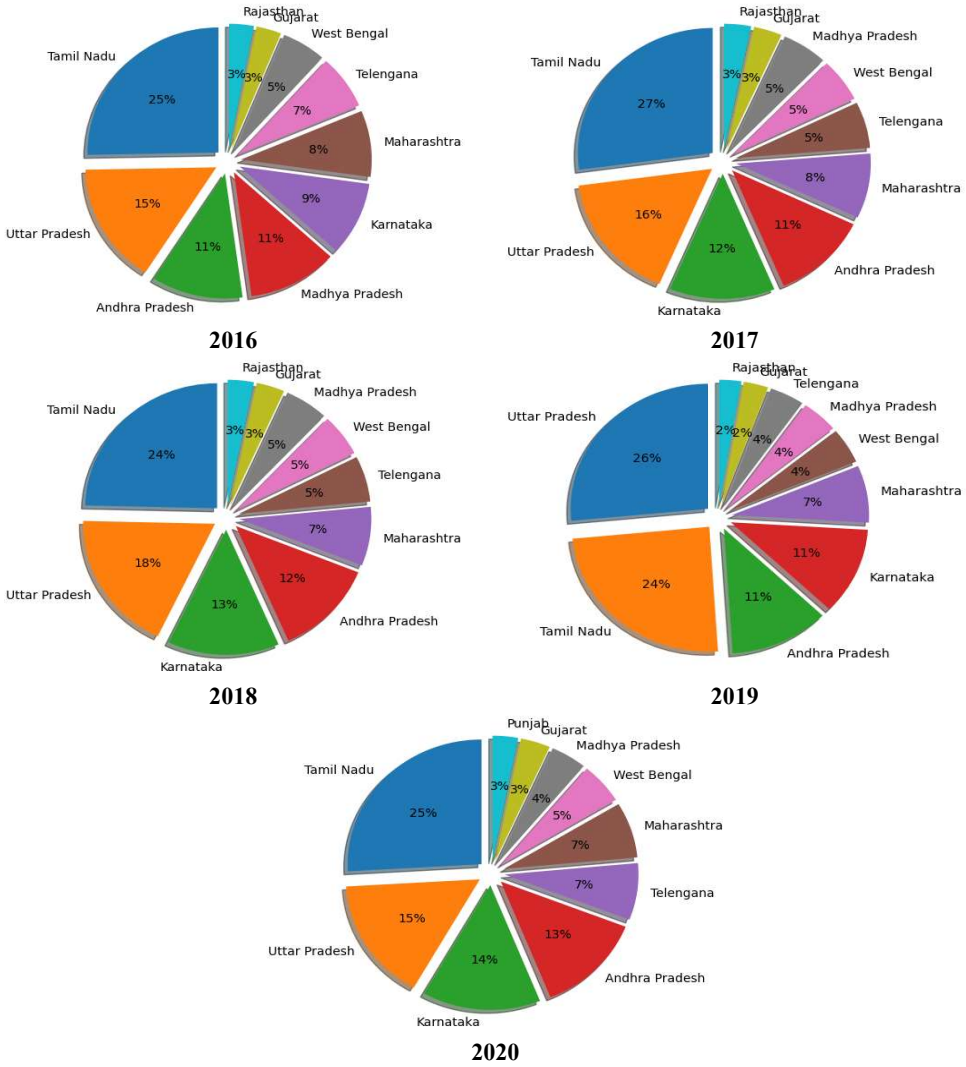

Figure 3. Number of inbound tourists to different states of India from 2014 to 2020.

3.1.4. Quarterly analysis of the number of tourists to India

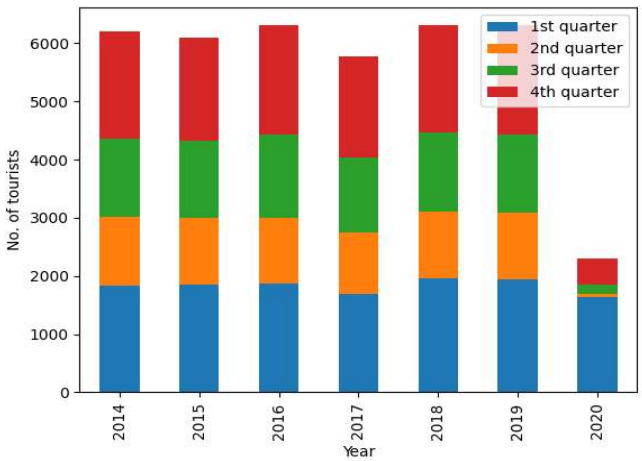

Figure 4. Quarterly analysis of the number of tourists to India.

Quarterly analysis of the number of tourists to India is useful in determining the impact of specific months on the number of tourists (**Figure 4**). The first quarter

includes January to March; the second and third quarters are from April to June and July to September, respectively, and October to December is the fourth quarter. **Figure 4** shows that the preferred time to have the most tourists is during the first quarter, followed by the fourth one. The third quarter and second quarter can be placed in the third and fourth positions on this list, respectively. The climatic conditions, such as extreme heat, maybe the reason for the minimal movement of tourists to India during the second quarter.

3.2. Analysis of revenue generated by India from tourists as foreign exchange fee (FEE)

Month-wise revenue generated by India from tourists as foreign exchange fee (FEE)

Tourism plays a critical role in generating revenue and thus increasing the GDP of a country. The main source of income is the foreign exchange fee. A month-wise analysis of revenue generated by India in US dollars from inbound tourists for a period from 2014 to 2022 is presented in **Figure 5**. This figure reveals the increase in revenue over the course of each year except 2020, which shows a sharp decrease of even zero due to the corona pandemic. This figure also assists in determining the highest and lowest revenue-generating months. This could be attributed to climate-related factors such as excessively high temperatures (about 40 °C) in several places in India. Therefore, the appropriate steps should be taken to draw in more visitors during these months.

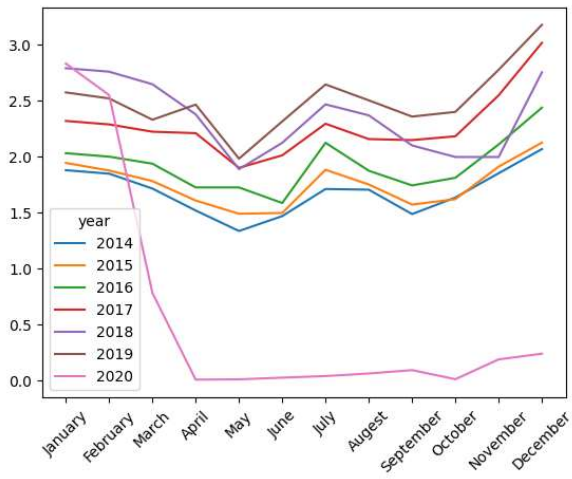

Figure 5. Month-wise revenue generated by India from tourists as foreign exchange fee (FEE).

3.3. The impact of the coronavirus pandemic on the tourism sector of India

Due to the sudden surge in COVID cases, countries around the globe preferred to shield themselves within the geographical barriers and thus imposed shutdowns and lockdowns. The tourism sector across the world has suffered due to very limited movements by human beings. This effect can also be observed in India, where there were almost 74% fewer tourists traveling to India in 2020 as compared to previous

years (**Table 1**). The comparison of the number of inbound tourists from different countries to India for 2019 and 2020 is shown in **Figure 6**. A sharp decline in the number of inbound tourists can be observed from this figure for 2020.

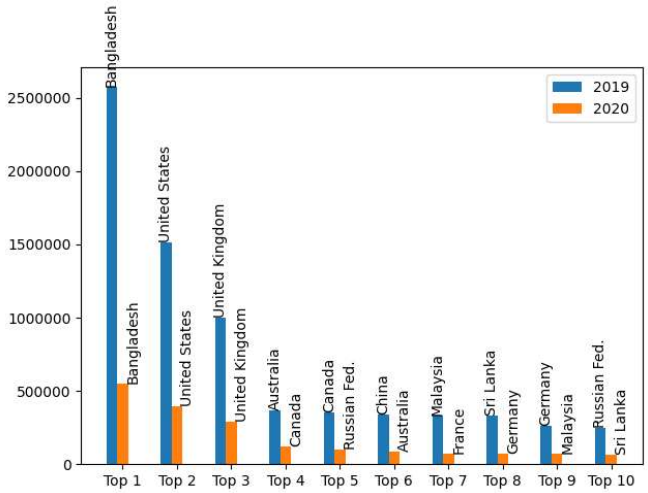

Figure 6. The comparison of the number of tourists to India for the years 2019 and 2020.

Table 1. Comparison of the number of tourists to the top ten states of India in the years 2019 and 2020.

Top-1	Uttar Pradesh	Tamil Nadu	535,855,162	140,651,241	0.74
Top-2	Tamil Nadu	Uttar Pradesh	494,865,257	86,122,293	0.83
Top-3	Andhra Pradesh	Karnataka	237,051,508	77,453,339	0.67
Top-4	Karnataka	Andhra Pradesh	227,934,714	70,828,590	0.69
Top-5	Maharashtra	Telengana	149,294,703	39,997,001	0.73
Top-6	West Bengal	Maharashtra	92,366,025	39,234,591	0.58
Top-7	Madhya Pradesh	West Bengal	88,707,139	28,841,732	0.67
Top-8	Telengana	Madhya Pradesh	83,035,894	23,519,632	0.72
Top-9	Gujarat	Gujarat	58,864,661	19,464,517	0.67
Top-10	Rajasthan	Punjab	52,220,431	16,692,197	0.68

3.4. Forecasting the number of tourists and revenue using a machine-learning approach

3.4.1. Prediction of the number of tourists from historical data

Prediction of the number of tourists from historical data from 2001 to 2019 is a regression problem [25]. In this problem, the variable year is considered the independent feature or predictor variable, while the number of tourists is the dependent variable. To judge the suitability of a machine learning model for such prediction, a comparison of different machine learning models is performed in terms of the coefficient of determination (R^2) and root mean square error (RMSE), and the result is presented in **Table 2**. This comparison entails using the support vector regressor to fit the data very well, as it has the maximum R^2 value and the least RMSE.

Table 2. Comparison of different machine learning models.

Model	R^2	RMSE
Support vector regressor	0.995	0.19
Linear regression	0.975	0.41
Random forest regressor	0.988	0.27
Lasso regression	0.97	0.41
Elastic net regression	0.972	0.43
Ridge regression	0.975	0.41
Decision tree-based regression	0.992	0.21

the support vector regressor [25] is trained over the yearly historical data in terms of the number of tourists visiting India starting from 2001 to 2016, and its validation is performed over the year from 2017 to 2019 (**Figure 7**). The figure reveals the SVR to well fit the data. This trained model is used to predict the number of tourists from 2021 to 2023 (**Figure 8**). Further, **Figure 8** also shows the actual number of tourists who visited India during this period. The actual number is far lower than the predicted one due to the corona pandemic.

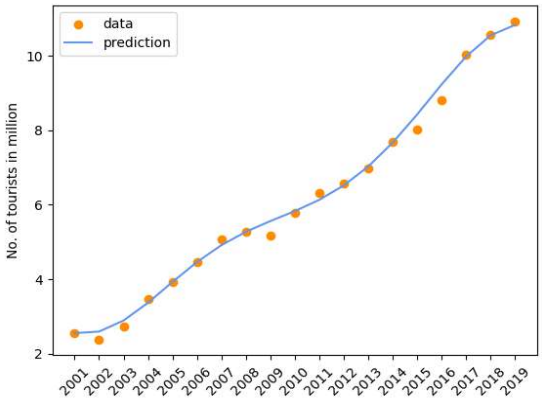

Figure 7. Training and validation of SVR over the number of foreign tourists.

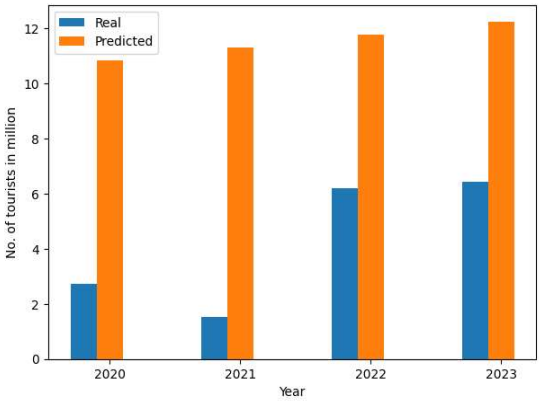

Figure 8. Actual vs. predicted number of foreign tourists to India in million.

3.4.2. Prediction of foreign exchange fee from historical data

Historical data from 2001 to 2016 is used to train the support vector regressor, while it is validated for the years 2017 to 2019 (**Figure 9**). This trained model is then

used to predict the foreign exchange fee for the years from 2020 to 2022 (**Figure 10**). This figure shows the amount of loss in the collection of FEE due to the corona pandemic.

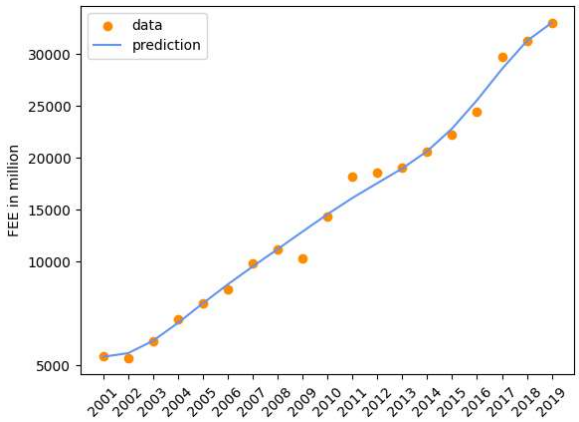

Figure 9. Training and validation of SVR foreign exchange fee.

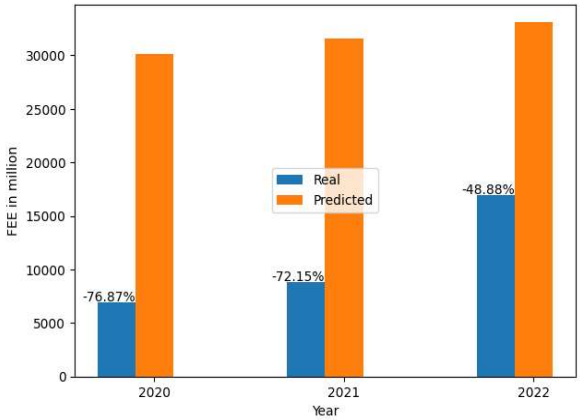

Figure 10. Actual vs predicted amount of FEE in $ million.

4. Conclusion

This study provides a comprehensive analysis of the Indian tourism industry, spanning the years 2014 to 2020. Further, the analysis is also augmented by machine learning predictive models to provide a smart analysis. The results presented in this paper will be very helpful to tourists, analysts, and stakeholders in the tourism sector.

The factors affecting the tourism sector in India include environmental factors, weather, and geographical locations. However, these factors, if moulded to the betterment may attract more tourists to India, and thus, its revenue from tourism will see a further surge.

The work carried out in this paper can be extended to provide an in-depth analysis of tourism across the world to study its trends and traits.

Author contributions: Conceptualization, SL and RKD; methodology, SL; software, RKD; validation, SL and RKD; formal analysis, SL; investigation, RKD; resources, SL; data curation, RKD; writing—original draft preparation, SL; writing—review and

editing, SL; visualization, RKD; supervision, RKD; project administration, SL. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

1. Kumar S, Kumar V, Bhatt IK, Kumar S. Mapping Research Trends on Smart Tourism: A Bibliometric Analysis. Digital Transformation of the Hotel Industry: Theories, Practices, and Global Challenges. 2023; 87-109.
2. Ye BH, Ye H, Law R. Systematic Review of Smart Tourism Research. Sustainability. 2020; 12(8): 3401. doi: 10.3390/su12083401
3. Shafiee S, Ghatari AR, Hasanzadeh A, et al. Developing a model for smart tourism destinations: an interpretive structural modelling approach. Information Technology & Tourism. 2022; 24(4): 511-546. doi: 10.1007/s40558-022-00236-7
4. Derdouri A, Osaragi T. A machine learning-based approach for classifying tourists and locals using geotagged photos: the case of Tokyo. Information Technology & Tourism. 2021; 23(4): 575-609. doi: 10.1007/s40558-021-00208-3
5. Peng R, Lou Y, Kadoch M, et al. A Human-Guided Machine Learning Approach for 5G Smart Tourism IoT. Electronics. 2020; 9(6): 947. doi: 10.3390/electronics9060947
6. Correia A, Dolnicar S. Machine learning applied to tourism—Contributions by Célia MQ Ramos. Women's voices in tourism research. 2021.
7. Ma H. Development of a smart tourism service system based on the Internet of Things and machine learning. The Journal of Supercomputing. 2023; 1-21.
8. Ghorbani A, Danaei A, Barzegar SM, Hemmatian H. Post modernism and designing smart tourism organization (STO) for tourism management. Journal of Tourism Planning and Development. 2019; 8(28): 50-69.
9. Mehraliyev F, Chan ICC, Choi Y, et al. A state-of-the-art review of smart tourism research. Journal of Travel & Tourism Marketing. 2020; 37(1): 78-91. doi: 10.1080/10548408.2020.1712309
10. Akdu U. Smart Tourism: Issues, Challenges and Opportunities. The Emerald Handbook of ICT in Tourism and Hospitality. Published online November 30, 2020: 291-308. doi: 10.1108/978-1-83982-688-720201018
11. Wang W, Kumar N, Chen J, et al. Realizing the Potential of the Internet of Things for Smart Tourism with 5G and AI. IEEE Network. 2020; 34(6): 295-301. doi: 10.1109/mnet.011.2000250
12. Cavalheiro MB, Joia LA, Cavalheiro GMC. Towards a Smart Tourism Destination Development Model: Promoting Environmental, Economic, Socio-cultural and Political Values. Tourism Planning & Development. 2019; 17(3): 237-259. doi: 10.1080/21568316.2019.1597763
13. Pai CK, Liu Y, Kang S, et al. The Role of Perceived Smart Tourism Technology Experience for Tourist Satisfaction, Happiness and Revisit Intention. Sustainability. 2020; 12(16): 6592. doi: 10.3390/su12166592
14. Bastidas-Manzano AB, Sánchez-Fernández J, Casado-Aranda LA. The Past, Present, and Future of Smart Tourism Destinations: A Bibliometric Analysis. Journal of Hospitality & Tourism Research. 2020; 45(3): 529-552. doi: 10.1177/1096348020967062
15. Corrêa SCH, Gosling MS. Travelers' Perception of Smart Tourism Experiences in Smart Tourism Destinations. Tourism Planning & Development. 2020; 18(4): 415-434. doi: 10.1080/21568316.2020.1798689
16. Gretzel U. Conceptualizing the smart tourism mindset: Fostering utopian thinking in smart tourism development. Journal of Smart Tourism. 2021; 1(1): 3-8.
17. Muniz ECL, Dandolini GA, Biz AA, et al. Customer knowledge management and smart tourism destinations: a framework for the smart management of the tourist experience – SMARTUR. Journal of Knowledge Management. 2020; 25(5): 1336-1361. doi: 10.1108/jkm-07-2020-0529
18. Errichiello L, Micera R. A process-based perspective of smart tourism destination governance. European Journal of Tourism Research. 2021; 29: 2909. doi: 10.54055/ejtr.v29i.2436
19. Borges-Tiago T, Veríssimo J, Tiago F. Smart tourism: a scientometric review (2008-2020). European Journal of Tourism Research. 2021; 30: 3006. doi: 10.54055/ejtr.v30i.2593
20. Novera CN, Ahmed Z, Kushol R, et al. Internet of Things (IoT) in smart tourism: a literature review. Spanish Journal of Marketing—ESIC. 2022; 26(3): 325-344. doi: 10.1108/sjme-03-2022-0035
21. Kontogianni A, Alepis E, Patsakis C. Promoting smart tourism personalised services via a combination of deep learning

techniques. Expert Systems with Applications. 2022; 187: 115964. doi: 10.1016/j.eswa.2021.115964
22. Balakrishnan J, Dwivedi YK, Malik FT, et al. Role of smart tourism technology in heritage tourism development. Journal of Sustainable Tourism. 2021; 31(11): 2506-2525. doi: 10.1080/09669582.2021.1995398
23. India Tourism Statistics 2022. Available online: https://tourism.gov.in/sites/default/files/2022-09/India%20Tourism%20Statistics%202022%20%28English%29.pdf (accessed on 6 March 2023).
24. India Tourism 2014-2020. Available online: https://www.kaggle.com/datasets/rajkachhadiya/india-tourism-20142020 (accessed on 6 March 2023).
25. Lipsa S, Dash RK. GASVR—A Model to Predict and Analyze Crude Oil Price. 2022 2nd Asian Conference on Innovation in Technology (ASIANCON). doi: 10.1109/asiancon55314.2022.9908764