

RESEARCH ARTICLE

The application of deep learning in economic analysis and marketing strategy formulation in the tourism industry

Jing Zhang, Ming Gao^{ID}*

Hospitality Management Department, Tourism College of Zhejiang, Hangzhou, China

* gaoming_edu@hotmail.com



Abstract

The tourism industry is ever-evolving in nature, as it operates in a global marketplace that has become progressively global and offers great potential due to technological advances. The tourism industry faces challenges in accurately forecasting economic impacts and understanding visitor patterns that rapid global changes. Motivated by these needs, this research introduces the Tourism Variational Recurrent Neural Network (TourVaRNN), aiming to enhance the tourism industry by predicting economic impacts and visitor behaviors for effective marketing strategies through advanced Deep Learning (DL) techniques. The research applies a variational recurrent neural network for enhancing tourism demands and model the complex temporal dependencies within tourism data. The proposed TourVaRNN integrates variational autoencoders to capture latent variables representing visitor preferences and spending habits, while recurrent neural networks model complex temporal dependencies in tourism data. Marketing campaigns in the tourism sector can be fine-tuned through visitor segmentation, which seeks to comprehend and classify visitors according to their demographics, preferences, and behaviors. The model employs robust forecasting of economic impacts, visitor spending patterns, and behavior while accounting for uncertainty through variational inference. The implementation uses Python language on a tourism dataset comprising necessary attributes like visitor numbers, days, spending patterns, employment, international tourism samples over a specific region, and a diverse age group analyzed over a year. The proposed method is evaluated in terms of performance metrics such as economic impact assessment, visitor segmentation efficiency, inference time analysis, and budget allocation utilization for effective economic and marketing strategy analysis in the tourism industry. TourVaRNN's improved segmentation efficiency of 15.7 percent allows for more targeted marketing, increasing engagement with visitors and income. Decisions may be made in real-time, improving operational efficiency in tourism management, thanks to a 17.5% reduction in inference time (to 40ms). The most efficient use of funds is guaranteed by a 13.4% rise in budget allocation utilization, leading to maximum economic benefits.

OPEN ACCESS

Citation: Zhang J, Gao M (2025) The application of deep learning in economic analysis and marketing strategy formulation in the tourism industry. PLoS One 20(6): e0321992. <https://doi.org/10.1371/journal.pone.0321992>

Editor: Vincenzo Basile, University of Naples Federico II: Università degli Studi di Napoli Federico II, ITALY

Received: July 18, 2024

Accepted: March 14, 2025

Published: June 6, 2025

Copyright: © 2025 Zhang, Gao. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All relevant data are within the paper.

Funding: This study is supported by National Social Science Foundation Art Project, "Research on the driving mechanism and improvement path of digital cultural consumption under the guidance of expanding domestic demand" (Project No: 23BH155).

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

The tourism industry can transform by adopting new strategies and innovations to meet evolving traveler/visitor demands. This includes liberalized interest rates, employment opportunities, varying earning levels to adapt to market changes and economic growth [1,2]. Embracing revolutionary developments in the travel industry has transformed worldwide activities, increased operational efficiency, and defined the future of marketing strategy [3]. Evolving tourist behavior and demand forecasting, impacted by tourist activities and visiting patterns, can lead to unpredictable demand patterns. This requires short-term operational decisions and long-term planning to manage their revenue and economic levels effectively [4]. For making strategic planning processes during tourism activities using branding campaigns, hotel feedback improves the efficiency of marketing communication with visitors [5]. The tourism industry emphasizes government response, technology innovation, and visitor importance, generating a global economic order [6].

ML techniques for real estate values analyze the property value changes inference by tourism [7]. Deep learning on tourism businesses refines their marketing strategies and improves services through review-based insights [8]. Research on tourism's impact on factors like tourist arrivals influence direct and indirect employment and identified seasonal impact without offering predictive solutions [9]. The advanced technologies change hotel elements' operational and marketing strategies for tourist destinations through demand forecasting of individuals in revenue generation [10]. Developing focus related to tourism elements like travellers utilized intelligent tourism services with ML algorithms to recognize region congestion rates maintained a narrow behavioral scope [11].

Tourism demand forecasting establishes nonlinear relationships between tourist arrivals and destination search volumes [12]. In an earlier tourist trend, a trip planner application investigated the seasonal visitor behavior and evaluated strategies to encourage off-peak travel. The results demonstrate that Expenditure is higher during off-peak seasons and that events are more effective than applications [13]. The tourism industry's dynamic nature presents significant chances for applying DL techniques to support better economic analysis and strategic planning for market trends [14]. Earlier studies employed DL to predict sentiments in tourism reviews, improve customer satisfaction analysis, enhance economic insights, and optimize hospitality marketing strategies [15]. The application of DL techniques mainly enhances the accuracy of forecasting tourism volume, which was evaluated using Hong Kong inbound arrivals data using a bi-directional gated recurrent unit (Bi-GRU) [16]. The development of the international tourism market enhances the potential of the sustainable condition of the world economy and its associated national economies, promotes travellers coming from different regions of the world and provides opportunities related to employment [17]. The significance of DL techniques emerges to overlook the diverse solutions of segmenting the customers and delivering precise forecasts to real-time analysis of visitor satisfaction through spending patterns in marketing [18].

Problem statement

Traditional tourism forecasting models, such as deep neural networks, long-term term models, and others, struggle with capturing complex temporal dependencies, especially in highly seasonal and volatile demand environments. Existing forecasting models fail to adapt to sudden shifts like economic crises, leading to forecast errors exceeding 20% in some cases. Standard LSTMs overfit short-term patterns, limiting their generalization to long-term trends. Hence, the introduced TourVaRNN addresses these gaps using variational inference, enhancing adaptability and accuracy in dynamic tourism markets.

Purpose of the research and its novelty

The recent works of the tourism demand forecasting model's reliance on less sophisticated methods often misallocates marketing resources, leading to a failure to capitalize on market opportunities. By developing a more accurate predictive model like TourVaRNN, this research seeks to enable tourism businesses to make more informed decisions regarding budget allocation, campaign targeting, and service offerings. The research novelty lies in the TourVaRNN model, which integrates VAE and RNN. Enhancing visitor segmentation by analyzing demographics and spending patterns enables more precise economic impact assessments and optimized budget allocation. The model's real-time capabilities improve marketing campaign efficiency and support data-driven decision-making.

The innovations of this research include:

- i) To accurately forecast economic impacts and visitor behaviors using Tourism Variational Recurrent Neural Network (TourVaRNN), formulate marketing strategies.
- ii) To uncover visitors' complete range of economic value in tourism, including influence on employment and expenditures.
- iii) To provide a foundation for market strategy formulation with a detailed analysis of traveler activities and the dynamic nature of economics.

The following sections discuss the arrangement of the research article.

A concise summary of recent literature works in the field of the tourism industry is discussed in Section 2. Section 3 talks about the advanced DL algorithm employed in this research as the TourVaRNN algorithm for analyzing the visitor patterns and its impact on economic growth and marketing strategy formulation. Section 4 delves into the detailed analysis of tourism data employed in this research and the performance comparison study results implementation. Section 5 concludes the research work and the future scope.

2 Related works

2.1 Tourism experience and behavioral insights

The correlations between service quality parameters and MTE dimensions were analyzed by Polyakova et al. [19] using a structural equation modeling (SEM) technique. Eight hundred seventy-eight swimmers from 10 different Oceanman events across various countries contributed their responses to the dataset. Hedonic and local cultural characteristics were the most robust predictors of MTEs' effects on word-of-mouth behavior, as validated by the results. The inability to generalize to different event types was a significant weakness caused by the study's narrow geographic and event-specific emphasis.

In contrast, Neshat et al. [20] suggested a Deep Neural Network (DNN) model to predict long-term policies attracting tourists to developing nation's economies after the pandemic. The model promoted domestic tourism and disaster preparedness as the best action based on analyzing crisis features and historical data. The case study demonstrated that it can help with tourist management decision-making, concrete policy suggestions, and prediction of policy outcomes. The model's forecast accuracy and usefulness in quickly changing settings may be affected by its dependence on historical evidence and predictions about foreseeable crisis dynamics.

2.2 Tourism demand forecasting and regional trends

Similarly, Pan et al. [21] analyzed tourism-related data from a rural county using an attention-based Long Short-Term Memory (Att-LSTM) model. The proposed study incorporated rural governance towards tourism development. The model improves rural tourist

initiatives and bolsters rural rejuvenation, which makes essential contributions. The model examines data on customer satisfaction and finds that personal recommendations led to more promotion of visitors. According to the results, there are a few limitations to assessing and promoting rural tourism. These include the need for precise data and the model's flexibility to handle various rural environments.

Expanding on tourism demand forecasting, Bi et al. [22] combined LSTM networks with Autoregressive (Autoreg) models to estimate tourism demand, using data from 77 sites in Beijing. The model increases forecast stability with better accuracy and less overfitting while capturing the spatial correlation among attractions. While the hybrid technique and new predictor combination are key contributions, there are certain constraints, such as potential concerns with data fidelity and the model's adaptation to different scenarios.

Sulong et al. [23] used Bagged Classification and Regression Trees (BC-RT) to project the demand and economic efficiency of halal tourism demand in Malaysia. It analyses 338,233 tweets with emotion labels, 11 Google trend search terms, business-specific factors, and economic data covering 2009–2020. With an improved accuracy of 93.71% in demand prediction and 80.12% in financial performance prediction. Difficulty integrating different data sources and the possibility of bias are two limitations.

Turning to marketing strategies, Elgarhy et al. [24] applied the 7Ps of marketing services looks at how they impact customer happiness and loyalty using SEM. One significant contribution is its evidence for the partial mediation of customer happiness and loyalty in the relationship between business profitability and customer purchase intentions during travel. This information is beneficial for managers of tourism services. The most significant benefit is a thorough comprehension of the ability to use the 7Ps to improve connections with customers and increase profits. However, a possible drawback is the difficulty of precisely measuring and using all 7P components in varied and ever-changing tourist markets.

2.3 DL applications in tourism management and recommendations

Kyrylov et al. [25] examined the present situation and future possibilities of international tourism, paying special attention to Ukraine's role in this industry. Significant results include a comprehensive analysis of world tourist indices, which place the United States, France, and Spain at the top. Research shows that tourism is a major driver of national economies, increasing GDP and bringing in much-needed foreign currency. Timely risk mitigation and decision-making rely on robust forecast models developed to anticipate future travel trends. While there is a need for improved risk mitigation techniques, the research also recognizes constraints to managing information support for economic entities in the tourism sector.

On the technological frontier, Wei [26] enhanced tourist attraction recommendation systems by extracting advanced image features, incorporating multimodal data fusion, and user portrait analysis using neural collaborative filtering in conjunction with deep convolutional neural networks (CNNs). Using a dataset including photos of popular tourist spots and information on user interactions allowed for an 89% success rate in picture recognition and an 85% success rate in recommendation-making. However, there are also certain restrictions, such as the difficulty of dealing with massive datasets and the need to adapt to a wide range of user preferences.

Siddik et al. [27] investigated the effect of artificial intelligence (AI) on the expansion of ecotourism using a mixed-method strategy that included Artificial Neural Network (ANN) analysis in addition to more conventional econometric models. From 2010 to 2022, the dataset encompassed the top ten tourist destinations worldwide, considering variables like GDP, FDI, inflation, and urbanization. While economic variables did have a role, the results

demonstrated that AI was a major driver of tourism efficiency and sustainability. One restriction was the difficulty in applying the findings to various sites worldwide.

Srinivasan et al. [28] employed an adaptive attention mechanism and a transformer-based encoder to improve the sentiment classification of travel reviews. A diversified smart tourism dataset with user-generated information in multiple languages was used to train the model, which achieved an accuracy of 96.74%, higher than CNN-BiLSTM standards. Further optimization is needed for real-time applications due to limitations like computing complexity and the possibility of overfitting to domain-specific language.

The summary of existing research publications discussed in Table 1 that were evaluated include a wide range of approaches and applications that strive to improve tourist management and forecasting in various situations. Among these, you can find AI-driven marketing strategies for policy prediction after a pandemic, research using attention-based LSTM, models for forecasting tourism demand using hybrid LSTM-Autoreg, models for predicting demand in the halal tourism industry, models for understanding the impact of marketing on customer loyalty using SEM, and extensive analyses of tourism indices and economic impacts on a global scale. While these studies shed light on opportunities to enhance tourism plans and economic forecasts, they also shed light on obstacles, such as the need for more flexible models, better data integration, and better information management.

Table 1. Summary of Research Gap Analysis in Existing Works.

Reference	Algorithm	Contribution	Results	Limitation	How TourVaRNN Addresses Gaps
Polyakova et al. [19]	Structural Equation Modeling (SEM)	Analyzed correlations between service quality and MTE dimensions.	The hedonic and local cultures were the strongest predictors of word-of-mouth behavior.	Downside due to event-specific and geographic focus.	TourVaRNN integrates VAE-RNN to model diverse segments and regional variations, enhancing generalizability.
Neshat et al. [20]	Deep Neural Network (DNN)	Policy-based tourism forecasting	Effective in policy suggestion and crisis handling	Limited adaptability to rapidly evolving situations	Uses recurrent structures to improve adaptability to dynamic tourism trends
Pan et al. [21]	Attention-based LSTM (Att-LSTM)	Enhances rural tourism governance	Boosts rural rejuvenation & tourist engagement	Requires highly precise tourism demand data.	TourVaRNN improves flexibility and can handle varying data precision across regions.
Bi et al. [22]	LSTM + Autoregressive Model	Tourism demand prediction with spatial correlation	More stable and accurate forecasting	Issues with data fidelity and generalization	Introduces variational autoencoders for better data representation and stability
Sulong et al. [23]	Bagged Classification & Regression Trees (BC-RT)	Demand and Economic Forecasting for Halal Tourism	93.71% accuracy in demand prediction	Bias risk and difficulty in integrating multiple data sources	Uses a variational framework to enhance data fusion and reduce bias
Elgarhy et al. [24]	Structural Equation Modeling (SEM)	Impact of 7Ps marketing on customer loyalty	Provides business insights for tourism managers	Lack of diverse market analysis	TourVaRNN integrates real-time analytics to adapt marketing strategies dynamically.
Kyrylov et al. [25]	Forecast Models	Global Tourism Trend Analysis	Identifies top tourism economies	Limited risk mitigation and data-driven decision-making	Enhances decision-making using predictive analytics for risk forecasting
Wei [26]	CNN + Neural Collaborative Filtering	Image-based attraction recommendations	89% accuracy in image recognition	Struggles with scalability and varied user preferences	Uses advanced embeddings to improve personalization and adaptability
Siddik et al. [27]	ANN + Econometric Models	AI's impact on ecotourism	Demonstrates AI-driven tourism efficiency	Lacks applicability to diverse locations	TourVaRNN adapts to various tourism settings with regionalized training
Srinivasan et al. [28]	HTAM (Hybrid Transformer-Attention Model)	Integrates transformer and attention mechanisms for enhanced sentiment classification in smart tourism.	Achieved 96.74% accuracy, outperforming CNN-BiLSTM.	Prone to overfitting with domain-specific jargon; high computational cost.	TourVaRNN handles temporal patterns and uncertainty in tourist behavior, complementing sentiment insights with economic predictions.

<https://doi.org/10.1371/journal.pone.0321992.t001>

3 Research methodology

The research aims to improve economic effect forecasts and give marketing strategy planners practical insights by using the proposed TourVaRNN model. Targeted marketing campaigns and policy decisions can be informed by in-depth analyses of tourist trends, employment effects, and expenditure patterns across various visitor types and locations, improving the local economy even further. To conduct an all-encompassing analysis of tourism data, the proposed Tourism Variational Recurrent Neural Network (TourVaRNN) system design incorporates variational autoencoders and recurrent neural networks. It includes layers to analyze inputs, encode hidden factors, anticipate requests, evaluate economic implications, and optimize to enable decisions based on information in tourism administration and marketing strategies. Analyze the direct and indirect effects of employment on tourism's economic contribution. The exact methodology and appealing message of Tourism Economics highlight the crucial role of the visitor economy. Expenditure in marketing and infrastructure to assist tourist growth can be justified with an understanding of the larger contribution of tourism to the economy provided by economic impact data. In order to plan accommodation (places to stay), schedules and marketing to encourage longer stays, it is necessary to understand the duration of guests' visits, and these indicators help with that. The goal of a visit can inform a marketing strategy's personalization. While vacationers are searching for fun things to do, business travelers may be more interested in meeting spaces. As a prime instance, it is possible to divide marketing campaigns into separate ones for local and foreign visitors.

This research idea depicted in Fig 1 adds to the knowledge of tourist analytics and economic prediction by introducing various novel ideas. The first thing you'll notice about TourVaRNN is how it uses recurrent neural networks and variational autoencoders to pick up on uncertainty and delays in tourist data. Additionally, it breaks new ground in economic effect modeling by breaking down the contributions of visitors to employment and spending by age group. Thirdly, it offers an information-driven structure for developing marketing strategies by capitalizing on in-depth knowledge of consumer habits and economic trends. Fourthly, using latent variables, TourVaRNN improves visitor segmentation and provides detailed information on tourist preferences. Finally, the model suggests a new

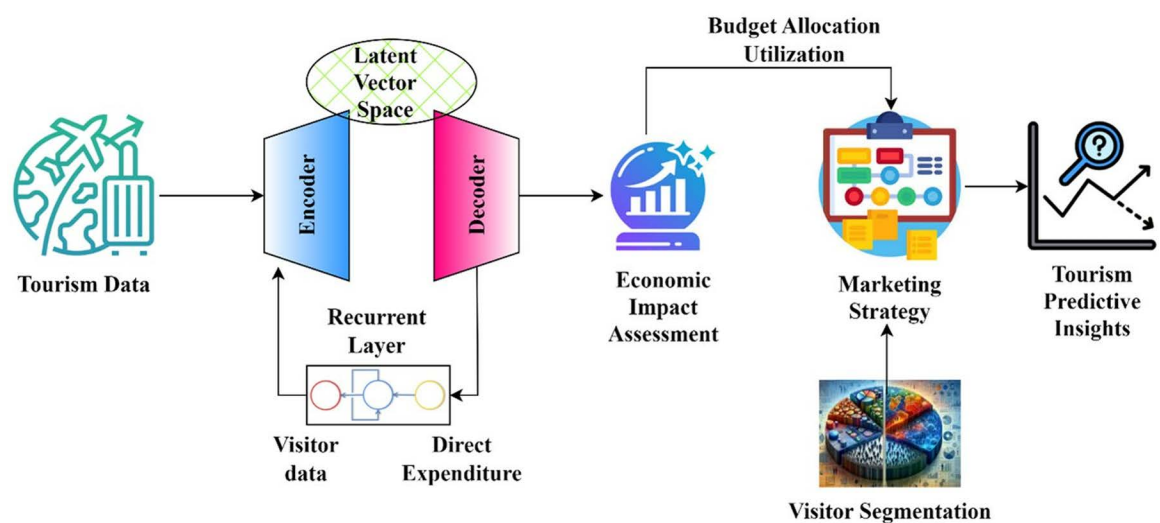


Fig 1. Schematic Representation of Proposed TourVaRNN Model.

<https://doi.org/10.1371/journal.pone.0321992.g001>

way to optimize the distribution of marketing funds by considering future economic effects and spending habits. Finally, TourVaRNN's capacity to adjust to changing times suggests that it will be useful in the tourist industry for a long time to come, filling in gaps in current approaches and leading to improvements in tourism management and strategy. This study's goal in doing visitor segmentation is to improve economic effect projections and marketing tactics by classifying tourists according to their demographics, interests, and behavioral tendencies.

3.1 Input tourism dataset description

The data employed for this research were sourced from the <https://datamillnorth.org/download/visitor-economy-data/9fb31f0b-8c0c-4215-ac09-91f49adafbe2/Visitor%2520economy%-2520data.csv> [29], that encompasses a tourists visitor numbers, visitor days, average length of stay, expenditures, economic impact and employment figures across different years from 2017 to 2021 has been collected from the data source <https://data.europa.eu/en/publications/datastories/high-value-datasets-tourism-eu> [30] for performance analysis since it contains the necessary values for all attributes. This data includes domestic overnight visitors, day visitors, and overseas inbound visitors, with detailed information on spending patterns and international visit statistics including purpose of visit can be segmented by various regions. The dataset includes information on the volume of events, number of delegates, and gross value of business tourism. The hotel performance data includes hotel occupancy rates, and average room rates.

[Table 2](#) discusses the features support the objectives of the study on marketing strategies and economic consequences by providing an all-encompassing view of tourist actions, revenue, and significant indicators pertaining to the tourist industry.

Data cleaning and handling outliers. Data cleaning includes handling missing values and outlier detection, where tourism data often contains missing values due to incomplete surveys, reporting errors, or seasonal variations. The numerical data imputation using mean/median considers from

$$X_{missing} = \frac{\sum_{i=1}^N X_i}{N} \quad (1a)$$

Table 2. Summary of Tourism Dataset.

Category	Details
Year of Analysis	2017-2021
Visitor Days	27.5M-32.1M days
Visitor Numbers	Maximum of 29M visitors/ year
Average Length of Stay	0.9-1.1 days
Visitors	Domestic Overnight Visitors, Day Visitors, Overseas Inbound Visitors
Economic Impact	Direct Expenditure (£M)
Employment Impact	Direct, Indirect (jobs)
International Visits	North America, Europe, Other Countries and Total world
	Average Nights Stayed, Average Spend Per Visit, Average Spend Per Day
Demographics (Age groups)	(0-15), (16-24), (25-34), (35-44), (45-54), (55-64), >65
Purpose of Visit	Holiday, Part of Inclusive Tour, Business, Visiting Friends & Relatives (VFR), study, Others
Business Tourism Value (£)	(£) 409.5M- (£)577.7M

<https://doi.org/10.1371/journal.pone.0321992.t002>

The mode imputation for visitor categories replaces missing purpose of visit values with the most frequent category. Outliers in tourism datasets can arise due to data entry errors, extreme events like pandemic impact and for detecting these anomalies, apply z-score method. Preventing anomalies form misclassifying visitor groups based on spending behavior and stay duration discussed in Eqns. (1a) and (1b)

$$Z = \frac{X - \mu}{\sigma} \quad (1b)$$

If $Z > 3$, the data point is considered an outlier, and is either removed or replaced using interpolation.

$$IQR = Q_3 - Q_1; Q_1 = 25^{th} \text{ percentile } (l), Q_3 = 75^{th} \text{ percentile } (u)$$

Where

$$l = Q_1 - 1.5 \times IQR; u = Q_3 + 1.5 \times IQR$$

The data points outside this range are flagged as outliers. Thus, sudden seasonal peaks or falls in hotel occupancy and visitor numbers could mislead model predictions. Also, the international visit trends like political factors, exchange rate fluctuations might cause extreme variations.

3.2 Variational encoder process

From the provided dataset the variables of interest given for visitor numbers termed as V can be formulated as

$$V = [v_1, v_2, v_3, \dots, v_t] \quad (2)$$

where v_t in Eq. (2), represents the number of visitors at time period t indicating in terms of years. An average length of stay can be represented as $L = [l_1, l_2, l_3, \dots, l_t]$ where l_t denotes the average length of stay at time t indicating in terms of particular years. Likewise, economic impact, $E_t = [e_{i1}, e_{i2}, \dots, e_{it}]$ varied with respect to specific year with in a given range, and for different age group i . For encoder process the variational layer model's latent variable z_t indicates visitor preferences, spending habits, and economic impact drivers at time t is used to comprehend and forecast tourist economic results, these unseen elements are vital and can be defined in Eq. (3a & 3b).

$$q(z_t | x_t) = \mathcal{N}(z_t | \mu_t, \sigma_t^2) \quad (3a)$$

$$z_t \sim p(z_t | z_{t-1}, h_{t-1}) \quad (3b)$$

The variational distribution $q(z_t | x_t)$ captures the uncertainty in the z_t in Eq. (3) gives the observed data x_t like visitor demographics, visitor days and other attributes that impacts the economic assessment.

3.3 Latent vector space representation

For the tourist data, the latent vector component models include latent variables, that stand for undiscovered trends and unidentified factors. It uses variational methodologies to capture complicated correlations and variances in elements impacting economic impacts and

tourist behaviors, such as visitor preferences and spending patterns. In order to analyze demographic information and purpose of visit data for tourism, deep learning's hidden layers store abstract traits and patterns, which allows for precise demand forecasting and strategic decision-making.

The latent variables are expressed as Latent Variable 1 (Age Group): To create age-specific marketing campaigns, it is necessary to capture visitor age segments (e.g., children, adults, seniors). The second latent variable, "Spending Behavior," shows how people spend their money, differentiating between those on a tight budget and those able to splurge. Third latent variable, "travel purpose," sorts visitors into three categories: vacation, business, and family. This allows for more personalized advertising. Tracks patterns of recurrent visits to facilitate the creation of loyalty programs; this is latent variable 4 (Frequency of Visits). The process of segmentation involves visualizing groups of people in order to conduct targeted marketing, such as high-spending leisure travelers. Campaign optimization: Maximizes engagement and conversions by coordinating marketing campaigns with visitor behaviors.

Fig 2 incorporates the variational autoencoder at each time step of the recurrent layer and the non-determinism presents in the hidden state h_t helps to capture the variability of the input state of visitor sequence and generates robust output distribution of the identified tourism patterns.

3.4 Recurrent layer for analyzing visitor behavior

Using Recurrent Neural Networks (RNNs), the TourVaRNN model's recurrent layer records the interdependencies between visitor behavior and economic effect data over time. This layer's ability to keep a secret state including past data is what gives the model its trend understanding and prediction capabilities. The accuracy of future economic effect and visitor behavior estimates is enhanced by the recurrent layer, which leverages historical trends. In the tourist sector, this competence aids in more informed strategic planning and decision-making. The following mathematical formulation serves as a contextual information that influences the computation of the current hidden state h_t and subsequent predictions or decisions in the model.

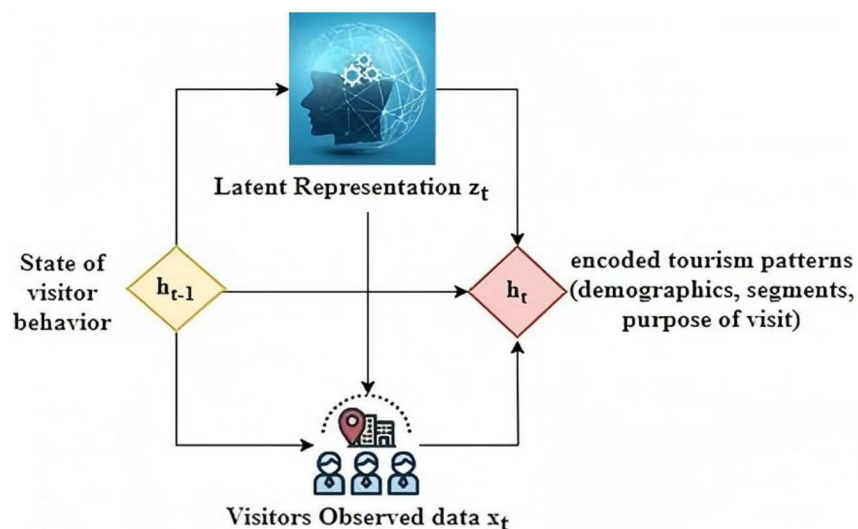


Fig 2. Learning Tourism Patterns from Hidden State Representation.

<https://doi.org/10.1371/journal.pone.0321992.g002>

$$h_t(V) = f(h_{t-1}, z_t, x_t) \quad (4)$$

This Eq. (4), defines the function $h_t(V)$ of all previous inputs and ensures that both the variability of the input sequence and temporal dependence between z_t up to the current time step across the distinct time period is captured. The inner function of x_t acts as a feature extractor and captures the l_t and v_t . The term h_{t-1} represents the state of visitor behavior and trends up to the previous point in time, aiding in understanding and predicting future patterns based on past data.

By combining VAE with RNN, the architecture shown in Fig 3 known as the TourVaRN models the temporal interdependence and uncertainty in tourist data forecasting. After the autoencoder $q(z_t|x_t)$ from Eq. (2) converts the input sequences V, L, E_t to a latent space, stochastic relationships are captured by variational inference. The recurrent decoder uses learnt latent representations to recover future states. This design optimizes uncertainty quantification and temporal dependencies simultaneously, which guarantees strong forecasting. Include assessment feedback in the data preprocessing block, suggesting that learning from model performance could lead to additional preprocessing step refining.

3.5 Variational decoder process

For decoder process the reconstruction layer can be used to predict economic impact e_{it} based on z_t is derived using an Eq. (5). In charge of using the model's learned encoded latent variables and temporal dependencies to reconstruct economic impacts and make predictions about visitor actions. Decisions can be made with more knowledge on the impact of changing visitor demographics and habits on economic outcomes.

$$p(e_{it} | z_t) = \mathcal{N}(e_{it} | \check{e}_{it}, \epsilon) \quad (5)$$

The term $p(e_{it} | z_t)$ predicts economic impacts based on learned latent representations z_t , offering insights into how various other factors like visitor behaviors and spending patterns influence economic outcomes.

3.6 Economic impact assessment

The direct spending and the effect of tourism on jobs are examined in this section. In doing so, it hopes to provide stakeholders with vital criteria for gauging the sector's whole economic health and possible growth areas by quantifying the economic benefits of tourism activities. By

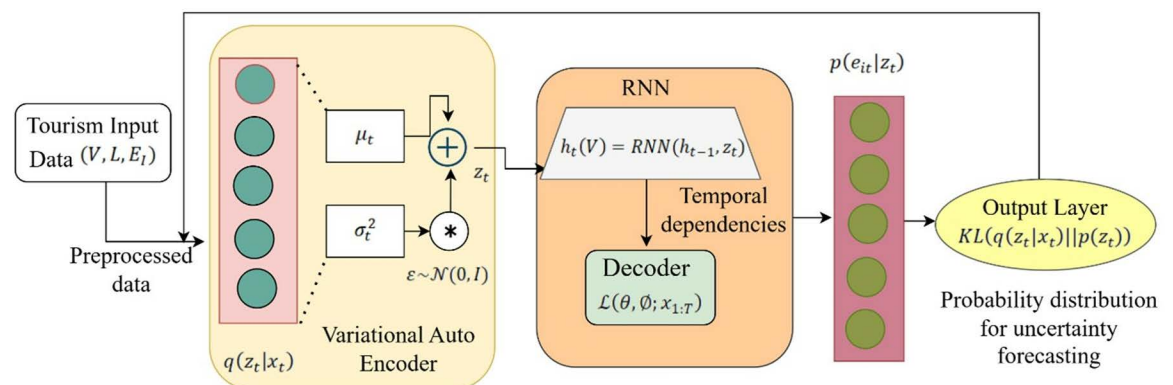


Fig 3. Integration of VAE and RNN Architecture.

<https://doi.org/10.1371/journal.pone.0321992.g003>

using variational lower bound, the elbow variational lower bound $\mathcal{L}$ derived in Eq. (6) optimizes the model parameters θ and $\varnothing$ of the inference network to accurately predict economic impacts while ensuring that the latent variables capture meaningful patterns in the data.

$$\mathcal{L}(\theta, \varnothing; x_{1:T}) = -\mathbb{E}_{q(z_t | \mu_t)} [\log p(e_{it} | z_t)] + KL(q(z_t | x_t) || p(z_t)) \quad (6)$$

This method optimizes the model by balancing reconstruction accuracy with the regularization term KL divergence. This is crucial for robustly modeling tourism economics, ensuring that the model learns z_t aligns with the observed x_t data. The variable μ_t defines the mean vector and σ_t^2 defines the variance-covariance matrix, indicating where z_t is likely to lie in the latent space. The conditional probability distribution $p(e_{it} | z_t)$ defines the e_{it} given z_t assumes Gaussian distribution parameterized by $\tilde{e}_{it}$ as mean prediction and ε as variance.

$$KL(q(z_t | x_t) || p(z_t)) = \frac{1}{2} \sum_{i=1}^N \left(1 + \log((\sigma_t^2)_i) - (\mu_t)_i^2 - (\sigma_t^2)_i \right) \quad (7)$$

Where in Eq. (7), $KL(q(z_t | x_t) || p(z_t))$ defines the approximate posterior $q(z_t | x_t)$ is close to the prior $p(z_t)$ in the latent space and it is computed element-wise across the dimensions of z_t and also the variable N defines the dimensionality of z_t .

$$E_i = \left(\frac{s_i \times P_i \times W_i}{\text{tot}(E_i)} \right) \times 100 \quad (8)$$

Utilizing the economic impact E_i calculated from Eq. (8) helps for each group i is calculated by integrating predicted total spending s_i from analyzed visitor data from encoder process, the proportion of the population P_i represented by that age group, and a weighted factor W_i that signifies the economic influence of each group. By adjusting the weights based on empirical data the assessment reflects varying spending behaviors from direct Expenditure.

3.7 Visitor segmentation

Using the acquired latent variables, it divides visitors into segments according to their demographics and how they engage. In order to increase customer happiness and financial returns, it is helpful to segment visitors based on their preferences. This allows for more targeted marketing and better services. From Eq. (9), latent variables z_t to segment visitors based on spending behavior, demographics, and purpose of visit like identifying the travel motivations.

$$\text{segment}(t) = \text{argmax}_k (p(\text{segment} = k | z_t)) \quad (9)$$

$$p(\text{segment}_k | z_t, a_g) = \frac{\exp(W_k^T z_t + b_k + V_k^T a_g)}{\sum_{j=1}^K \exp(W_j^T z_t + b_j + V_j^T a_g)} \quad (10)$$

In Eq. (10), the mathematical formulation segments visitors into K different categories k based on latent variables z_t and age group features termed as a_g . The other indicator variables like W_k, b_k , and V_k are learnable parameters for each segment, determining the probability of a visitor belonging to segment k for evaluating marketing strategy. Apply these segments to forecast spending patterns $\tilde{s}_{t+1}(k)$ for targeted marketing. The spending data comprises Expenditure per visit, average spend per day to forecast future spending. Incorporate these predictions into spending patterns $\sum_k \tilde{s}_{t+1}(k)$ for budget allocation. For allocating the marketing budget B_t at time t derive the Eq. (11) of the form.

$$B_t = \alpha \check{e}_{t+1} + \beta \sum_k \check{s}_{t+1}(k) \quad (11)$$

The parameters α and β indicates the weights that can be adjusted based on strategic priorities and model predictions. The term α indicates the weight reflecting the importance of predicted economic impact $\check{e}_{t+1}$, similarly β represents the weight reflecting the importance of spending patterns $\check{s}_{t+1}(k)$ of segment k at time $t+1$, would prioritize revenue optimization from specific visitor segments.

3.8 Pseudocode for visitor segmentation

```

function VisitorSegmentation (  $z_t, a_g$  ):
  Init parameters (  $K, W_k, b_k, V_k$  )
  for  $k = 1$  to  $K$  do # Segment probability calculation
    compute  $\exp(W_k^T z_t + b_k + V_k^T a_g)$ 
    compute  $\sum_{j=1}^K \exp(W_j^T z_t + b_j + V_j^T a_g)$ 
     $segment(t) = \operatorname{argmax}_k p(segment_k | z_t, a_g)$ 
  end for
return  $segment(t)$ 
function ForecastSpendingPatterns (  $segment, t$  ):
  for  $k = 1$  to  $K$  do # Forecast spending for each segment
     $\check{s}_{t+1}(k) = \operatorname{PredictSpending}(segment_k, t+1)$ 
  end for
return  $\sum_k \check{s}_{t+1}(k)$ 
function AllocateMarketingBudget(  $\check{e}_{t+1}, \check{s}_{t+1}(k)$  ):
  init (  $\alpha, \beta$  )
  compute  $B_t$ 
return  $B_t$ 
function OptimizeBudgetAllocation(  $K, total\ budget$  ):
  init  $Rev = 0$ 
  evaluate  $B_k = 0$  for  $k$  in range (  $K$  )
  While  $B_k \leq total\ budget$  do
    for  $k = 1$  to  $K$  do
      if  $B_k \geq MinBudget_k$  then
         $ROI_k = \operatorname{calculate} ROI(k, B_k)$ 
         $Rev += \alpha_k \cdot B_k \cdot ROI_k$ 
      else
         $B_k > total\ budget$ 
      end for
    end for
return  $Rev, B_k$ 
# Main segmentation and budget allocation process
 $z_t = \operatorname{GetLatentVariables}(\operatorname{visitor\_data})$ 
 $a_g = \operatorname{GetAgeGroupFeatures}(\operatorname{visitor\_data})$ 
 $segment = \operatorname{VisitorSegmentation}(z_t, a_g)$ 
 $\check{s}_{t+1} = \operatorname{ForecastSpendingPatterns}(segment, \operatorname{current\_time})$ 
 $\check{e}_{t+1} = \operatorname{ForecastEconomicImpact}(z_t)$ 
 $B_t = \operatorname{AllocateMarketingBudget}(\check{e}_{t+1}, \check{s}_{t+1})$ 
 $total\ budget = B_t$ 
return  $tot(E_t)$ 

```

The provided pseudocode outlines a structure for the tourist industry's visitor segmentation and marketing budget allocation. It separates visitors by latent characteristics and age group, forecasts spending habits, and optimizes expenditures on marketing. This technique targets marketing strategies and optimizes resource allocation, improving tourism marketing's economic impact and efficacy. The VisitorSegmentation function assigns visitors to the most likely segment using latent variables z_t and age group attributes a_g . ForecastSpendingPatterns anticipates segment spending trends and totals them. OptimizeBudgetAllocation distributes the budget across segments to optimize revenue. The core procedure uses visitor data to identify hidden characteristics and demographic features, segment visitors, forecast expenditure and economic impact, distribute the marketing budget, and return the total predicted economic impact $tot(E_t)$.

3.9 Marketing strategy formulation

Marketing strategy formulation optimizes marketing budget allocation and maximizes the revenue across various visitor segments and creates strategic campaigns. Optimizing marketing efforts and improving visitor experiences are the goals of this strategy formulation, that draws on data from demand forecasting and visitor segmentation.

3.10 Budget allocation utilization

The Budget Allocation Utilization (BAU) influenced by market potential, historical performance, and strategic priorities.

$$\text{Maximize } Rev = \sum_{k=1}^K \alpha_k \cdot B_k \cdot ROI_k \quad (12)$$

subject to the constraints:

$$\begin{aligned} \sum_{k=1}^K B_k &\leq \text{total budget} \\ \sum_{k=1}^K B_k &\geq \text{MinBudget}_k \quad \forall k \end{aligned}$$

The term α_k represents weight for segment k based on its strategic importance and B_k indicates budget allocated to segment k followed by ROI_k represents return on investment for segment k . The goal is to maximize total revenue by allocating the marketing spend optimally across different categories. The given constraints make sure there is enough money in the budget for everything and that it doesn't go overboard. More efficient use of marketing resources can increase income, which can be achieved through optimizing budget allocation.

From Table 3 demonstrates the data from time period 2017–2021 provide insights into international visitors, that could inform both economic analysis and marketing strategies for different regions. Tourism officials can maximize revenue through targeted marketing and personalized service offers by analyzing the spending patterns $\check{s}_{t+1}(k)$ and stay durations of visitors across various regions. Developing regional advertising strategies to entice North American tourists with large spending power and international visitors with extended stays l_t . Creating discounted long-term packages and shorter-stay options for tourists from Europe and other continents. Improving the tourist experience according to local tastes and spending habits to increase happiness and encourage return visits.

4 Results and discussion

For comparison study the existing models such as DNN [20], Att-LSTM [21], and BC-RT [23] are contrasted with the proposed TourVaRNN in terms of performance metrics such as

Table 3. International Visitor Analysis for Marketing in Different Regions.

Region	Avg. Nights Stayed I_t	Avg. Spend per visit (£)	Avg. Spend per Day (£)
North America	7	617	83
Europe	4	244	64
Other Countries	13	910	68
Total World (Finland, Romania, Poland, Spain, Germany, Greece)	6	398	69

<https://doi.org/10.1371/journal.pone.0321992.t003>

economic impact assessment, visitor segmentation efficiency, inference time analysis, and budget allocation utilization for effective economic and marketing strategy analysis in the tourism industry.

TourVaRNN surpasses DNN in performance due to its superior uncertainty modeling using VAEs, less need for human feature engineering, and improved predictive accuracy, as shown in Table 4. It provides shorter inference times for real-time forecasting and efficiently captures long-term patterns compared to Att-LSTM. Compared to BC-RT, TourVaRNN is superior in adaptability to visitor behaviors, output reliability, and probabilistic outputs.

For this comparison analysis, two variants are selected; one is visitors' demographics from [29] based on different age groups (0–15), (16–24), (25–34), (35–44), (45–54), (55–64), >65. The second variant is based on tourism data analysis from [30] from 2017 to 2021.

Table 5 presents the key hyperparameters for the TourVaRNN model. It uses a learning rate of 0.0001, 128 hidden neurons, and two layers with a batch size of 64 for training over 100 epochs. The model employs 30-time steps, early stopping patience as 10, and a latent space of 64 for effective sequence learning and economic impact prediction.

4.1 Tourism predictive insights

Direct Expenditure (£M) is the accumulation of all the money tourists spend in a particular region. Everything directly linked to the tourists' spending falls under this category, including Accommodation: The amount of money spent on various forms of housing, including hotels, hostels, bed & breakfasts, and more. Expenditures on food and drink at multiple establishments such as restaurants, cafes, and bars. Transportation: Expenses related to getting around town include taking the bus, a cab, renting a car, and gas—expenditures on amusements, museums, tours, performances, and other entertainment. Merchandise tourists buy may include apparel, trinkets, and other mementoes. Miscellaneous Expenses: Tips, medical

Table 4. Comparison of Models: Limitations Vs. TourVaRNN Advantages.

Model	Limitations	TourVaRNN Advantages
Deep Neural Network (DNN)	• Overfitting risk with limited data.	• VAEs quantify uncertainty for reliable forecasting.
	• Requires extensive feature engineering.	• Requires less manual feature engineering and enhance predictive accuracy.
	• Lacks uncertainty quantification.	
Attention-Based Long Short-Term Memory (Att-LSTM)	• Computationally expensive due to attention mechanisms.	• Efficiently captures long-term temporal patterns with faster inference times for real-time prediction
	• Struggles with long-term dependencies without proper tuning.	• Optimized inference speed for real-time predictions.
	• Requires large, high-quality datasets.	
Bagged Classification and Regression Trees (BC-RT)	• Susceptible to bias from training data and lacks accuracy	• More flexible and adaptive in learning complex visitor behaviors and improves forecast accuracy
	• Rigid structure limits adaptation to nonlinear relationships.	• Provides probabilistic outputs for better decision-making.
	• Lacks uncertainty modeling.	

<https://doi.org/10.1371/journal.pone.0321992.t004>

Table 5. Hyperparameter Parameter Setting.

Parameter	Range with its Description
Learning Rate	0.0001 - Rate of update for model weights.
Hidden Layer Size	128 - Number of neurons in each hidden layer.
Number of Layers	2 (VAE+RNN) - Total recurrent and variational layers in the model.
Batch Size	64 - Number of samples per gradient update.
Training Iterations	100 - Total epochs for training the model.
Loss Function	Mean Squared Error (MSE) - Metric for evaluating model performance.
Sequence Length	30 - Number of time steps taken as input to the RNN.
Early Stopping Patience	10 - Number of epochs without improvement before stopping training.
Latent Space Dimension	64 - Dimensionality of latent variables in the variational autoencoder.
$W_k, b_k, \text{ and } V_k$	Learnable parameters

<https://doi.org/10.1371/journal.pone.0321992.t005>

services, and personal care are examples of visitor expenditures that cannot be categorized into the above categories.

As tourists' spending grows, so does the overall economic contribution because of the positive link between direct Expenditure and economic impact. The correlation analyzed below suggests that fluctuations in direct Expenditure can serve as a predictive factor for economic impact, reinforcing the importance of tourism as a key economic driver.

$$E_t = \gamma_0 + \gamma_1 \check{s}_{t+1}(k) + \varepsilon \quad (13)$$

Where γ_0 in Eq. (13) represents the marginal economic impact per unit increase in tourist spending, A positive γ_1 confirms a direct proportionality, meaning increased expenditure leads to a higher economic impact. where $\gamma_1 > 0$ confirms that higher visitor spending leads to a directly proportional increase in economic impact. The statistical significance is typically validated using regression analysis, where a high R^2 value and a p-value < 0.05 indicate a strong and meaningful relationship.

The graphical illustration provided in Fig 4 shows that the direct spending by visitors correlates with the overall economic impact over the years. As direct expenditure increases, there will be a corresponding rise in the economic impact, suggesting a strong positive relationship between these two variables. The left y-axis indicates the direct Expenditure (£M) evaluated for millions of pounds sterling, and the right y-axis defines the economic impact in terms of (£M). Demand forecasting helps future visitor numbers V_t and visitor days can be given as $\check{v}_{t+1}$ can be evaluated using the proposed TourVaRNN to predict future demand.

i) Economic impact assessment. For evaluating economic impact assessment among different algorithms, the proposed TourVaRNN outperforms DNN, Att-LSTM, and BC-RT in economic impact assessments for most age groups. This shows that TourVaRNN's capacity to capture complex patterns using variational inference blocks for analyzing spending patterns s_i Concerning consequences across many age groups, they lead to more precise forecasts.

$$E_t = \left(\frac{s_i \times P_i \times W_i}{\text{tot}(E_t)} \right) \times 100 \quad (14)$$

On the other hand, DNN, Att-LSTM, and BC-RT all show different levels of economic impact E_t derived in Eq. (14) depends on age group, which shows how well they model and predict economic dynamics within specific demographics varied on different W_i . Policymakers and tourist managers rely heavily on these findings for long-term planning. To maximize

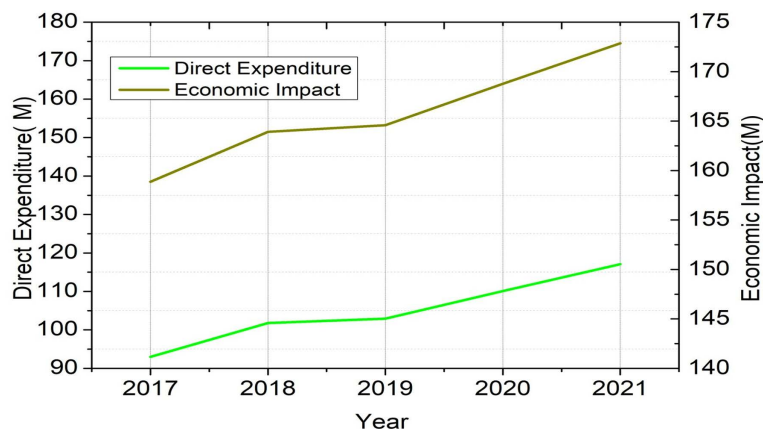


Fig 4. Direct Expenditure Vs. Economic Impact (2017-2021).

<https://doi.org/10.1371/journal.pone.0321992.g004>

the economic benefits of tourism, officials need to know which age groups have the most significant influence on the industry's bottom line. Only then can they modify marketing campaigns, distribute budget funds wisely, and plan for infrastructure upgrades accordingly based on these $s_i \times P_i \times W_i$ evaluation factors for economic analysis. The results also show the importance of using advanced DL techniques with the combined effort of the variational auto encoder and recurrent neural network as TourVaRNN to improve economic effect evaluations and back decisions in tourism and economic planning with data.

Based on the computed percentage improvement provided in Table 6, which comprises roughly 7.63%, TourVaRNN has a higher economic impact assessment across the age group of 0–15 than DNN. This method gives a holistic picture of how various age groups impact tourism's bottom line. Stakeholders can allocate budgets effectively and enhance the economic advantages received from tourism operations with the help of this methodology, which evaluates the economic impact of different age groups and guides strategic decisions in the management of tourism and marketing.

This research applies analysis of variance (ANOVA) to determine if there are statistically significant variations in the mean E_i evaluations according to a_g .

Key Steps in ANOVA Implementation. Formulation of Hypothesis:

Null Hypothesis (H_0): E_i means across all a_g are equal, indicating no significant influence of age on economic contribution.

Alternate Hypothesis (H_1): At least one a_g has a significantly different mean economic impact, suggesting that age group affects tourism spending behavior s_i .

Table 6. Economic Impact Assessment by Age Group Comparison.

Age Group	DNN	Att-LSTM	BC-RT	TourVaRNN
(0-15)	11.8	10.5	12.0	12.7
(16-24)	14.5	13.2	15.1	15.6
(25-34)	17.9	18.3	17.5	19.4
(35-44)	19.6	20.1	20.0	21.9
(45-54)	18.9	19.5	18.7	20.7
(55-64)	16.7	17.0	16.8	18.5
>65	15.8	15.2	16.0	17.3

<https://doi.org/10.1371/journal.pone.0321992.t006>

The F-statistic is calculated by comparing variance between and within age groups, and if a high F-value with a p-value < 0.05 indicates the difference in E_i across age segments.

There are noticeable disparities in the economic impact shown in Table 7, with the 25–44 age group making the most significant contribution, suggesting that they spend more on tourism. The preferences of active travel groups (16–64) exhibit moderate to high diversity.

ii) Visitor segmentation efficiency analysis. Through behavioural and preference-based tourist segmentation, TourVaRNN facilitates the development of targeted marketing campaigns. Engagement and conversion rates can be enhanced with this tailored strategy. Proactive marketing that meets expected demand, making the most efficient use of resources while increasing impact, is made possible by the model's predictions of future visitor behaviors and spending habits.

An essential measure for assessing the efficacy of segmentation tactics in marketing and tourism is visitor segmentation efficiency, depicted in Figs 5(a) and 5(b). Companies can improve their tactics to increase consumer engagement and economic effect by measuring the accuracy with which segments represent visitor diversity and forecast behaviors. The proposed TourVaRNN patterns and places to improve your segmentation techniques by plotting

Table 7. ANOVA Validation for Economic Impact Across Age Groups.

Age Group	Mean Economic Impact	Variance (Within Group)	Variance (Between Groups)	Possible Reasons for Economic Impact Differences
0-15	12.7	Low	High	Lower spending power due to financial dependence on guardians.
16-24	15.6	Moderate	High	Young adults travel for education, adventure, and social activities.
25-34	19.4	High	High	Peak earning phase, higher disposable income, frequent travelers.
35-44	21.9	High	High	Family-oriented travel, business trips, and higher spending on accommodations.
45-54	20.7	Moderate	Moderate	Work-life balance, moderate disposable income, and steady tourism participation.
55-64	18.5	Moderate	Moderate	Retirement planning and leisure trips but budget-conscious spending.
>65	17.3	Low	Moderate	Fixed incomes lead to lower travel spending and a preference for relaxed tourism.

<https://doi.org/10.1371/journal.pone.0321992.t007>

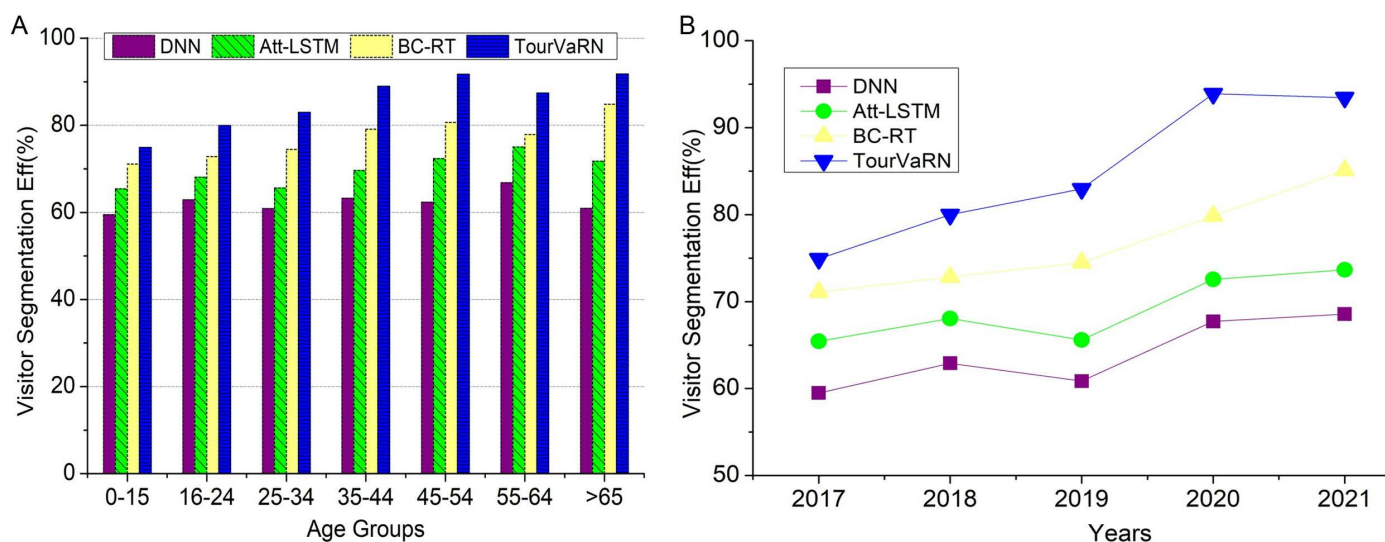


Fig 5. (a). The Efficiency of Visitor Segmentation Over a Different Age Population. (b). Efficiency of Visitor Segmentation Over a Year.

<https://doi.org/10.1371/journal.pone.0321992.g005>

the efficacy of visitor segmentation over time or across many different categories, such as age groups and across different years). If the segmentation approach can capture all the variations and nuances throughout visitors, it allows for more targeted marketing activities and resource allocations, resulting in higher efficiency.

The TourVaRNN model provides visitor spending forecasts across segments, exhibiting varying uncertainty levels in [Table 8](#). Leisure and family vacation segments demonstrate good reliability with narrower confidence ranges. There is a moderate level of dependability in business travel and group tours because of the higher variance. With less variation and a narrower confidence range, couples' getaways are the least reliable.

[Table 9](#) presents four types of visitors, each with its traits, spending habits, and validation data from real-world tourism marketing campaigns. The marketing campaign outcomes match each segment's profile, including average stay, spending habits, and preferences. The segmentation model is validated and found to be practically relevant when segment behaviors align with marketing outcomes like return on investment (ROI) from promotions or program conversions.

iii) Inference time calculation. TourVaRNN fared better than similar models such as DNN, Att-LSTM, and BC-RT in conducting economic effect assessments across various age groups. It proved more efficient in distinguishing between different types of visitors, which made targeted marketing and resource allocation possible. Additionally, TourVaRNN demonstrated the quickest inference times, which facilitated the implementation of real-time decision-making in the tourism management sector. The model's full research informs strategic marketing campaigns and policy decisions regarding tourist trends, employment effects, and expenditure patterns. This analysis can strengthen local economies while also providing valuable insights.

Regardless of age in [Fig 6\(a\)](#) and year in [Fig 6\(b\)](#), the TourVaRN model consistently displays the quickest inference times. The fact that it outperforms simpler models despite its complexity is a result of its streamlined architecture and efficient use of variational inference.

Table 8. Confidence Intervals and Variance Analysis for TourVaRNN Predictions.

Segment	Prediction	Mean Prediction Value	Confidence Interval (95%)	Variance	Reliability Indicator
Segment 1 (Leisure Travel)	Visitor Spending	\$800	750–850	100	High
Segment 2 (Business Travel)	Visitor Spending	\$1200	1150–1250	150	Moderate
Segment 3 (Family Vacations)	Visitor Spending	\$950	900–1000	200	High
Segment 4 (Couples' Getaways)	Visitor Spending	\$700	650–750	80	Low
Segment 5 (Group Tours)	Visitor Spending	\$600	550–650	250	Moderate

<https://doi.org/10.1371/journal.pone.0321992.t008>

Table 9. Visitor Segmentation Validation Accuracy.

Segment	Characteristics	Spending Behaviors	Validation (Marketing Outcome Comparison)
Segment 1: Luxury Travelers	High-income individuals, average, stay 10 nights	Average spend £950 per visit, prefer premium experiences	Matches luxury campaign results, showing increased ROI from high-end offers
Segment 2: Budget Travelers	Cost-conscious tourists, average stay 5 nights	Average spend £250 per visit, prefer economical packages	Aligns with budget campaign results, showing strong conversions from promotions
Segment 3: Business Travelers	Frequent short-stay visitors, typically for work trips	Average spend £300 per visit, high daily spend on accommodations and transport	Correlates with business incentive program outcomes
Segment 4: Family Travelers	Group travellers, average stay 7 nights	Average spend £600 per visit, seek family-friendly services	Consistent with outcomes of family-oriented package promotions

<https://doi.org/10.1371/journal.pone.0321992.t009>

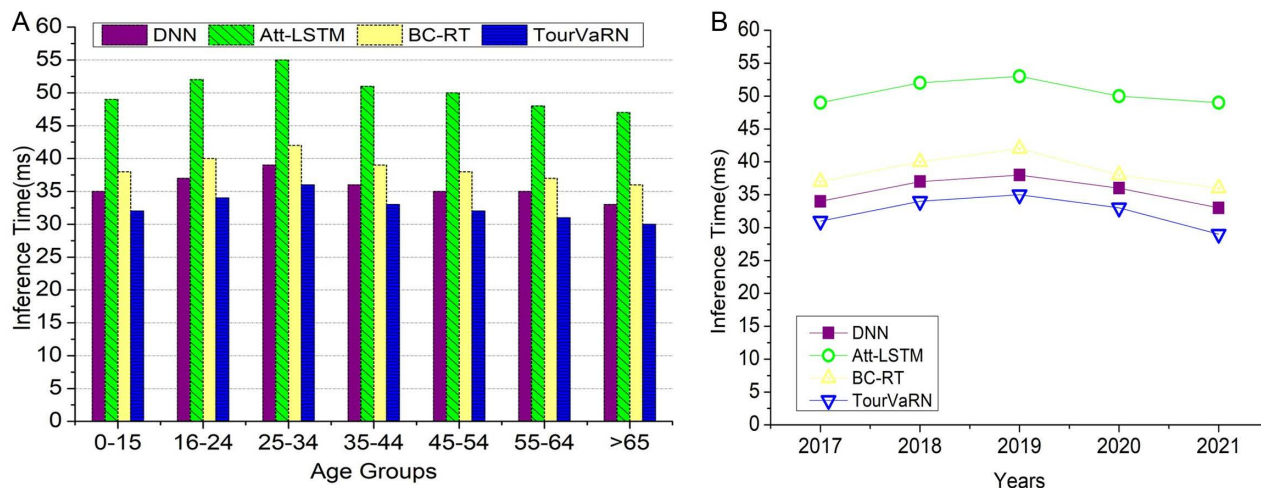


Fig 6. (a). Inference Time Analysis on Distinct Age Groups. (b). Inference Time Analysis Across Years.

<https://doi.org/10.1371/journal.pone.0321992.g006>

Except for TourVaRNN, which keeps its maximum time of 40 ms, this is the highest limit of the inference time, which is set to 60 ms for all models. TourVaRNN's reduced inference times indicate it can handle tourism data faster, allowing for quicker decision-making in marketing and management strategies and real-time applications. The proposed model's capacity to simulate temporal changes and integration of age group data makes it ideal for studying tourism patterns from 2017 to 2021. It may show changing trends in visitor segments based on age. From 2017 to 2021, inference times generally decreased, which could be due to hardware advancements or model optimizations.

iv) Budget allocation utilization. Using this *BAU* calculation illustrated in Figs 7(a) and 7(b), the efficiency of each algorithm distributes funds among different demographic subsets and time intervals like years. For each age group and year, analyze each algorithm, including proposed TourVaRNN, compared to existing DNN, Att-LSTM, BC-RT, and HTAM models to forecast the number of visitors or economic effect. Consider the unique traits with the

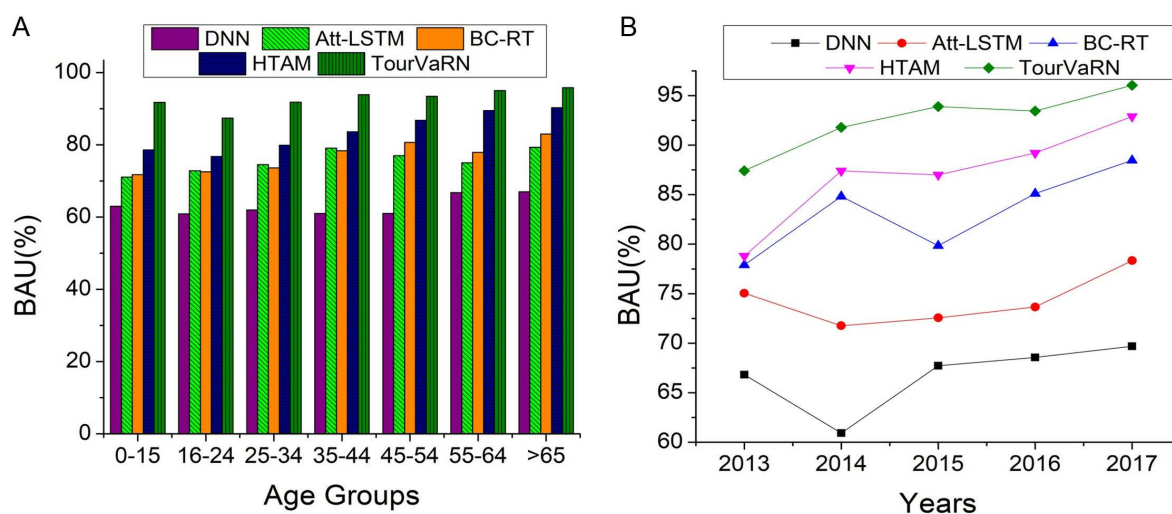


Fig 7. (a). BAU Comparison on Different Age Groups. (b). BAU Comparison on Varying Years.

<https://doi.org/10.1371/journal.pone.0321992.g007>

constraints and check $B_k \leq \text{total budget}$ and tendencies detected by each algorithm when allocating funds according to these forecasts. The continuous analysis of the marketing budget varies on $B_k \geq \text{MinBudget}_k \forall k$ and how much revenue can be generated through this marketing strategy formulation $\sum_{k=1}^K \alpha_k \cdot B_k \cdot \text{ROI}_k$ and analyze actual B_k spent on each age group and year. Evaluate the algorithms' BAU values across various demographic groupings and periods to determine which generates higher utilization rates. Determine the efficacy of each algorithm by examining patterns and trends in BAU. Identifies when and where algorithms show a higher BAU, aiding marketing strategy optimization. converts forecasts into efficient distributions of funds.

Case Study: 1 High-Spending Behavior in Certain Age Groups

Travellers aged 30–45 were shown to exhibit high-spending tendencies in a regional tourism dataset analyzed with TourVaRNN. The emergence of premium tourist products might be attributed to the fact that this age group spends 35% more on leisure activities than younger travelers (Table 10).

Figs 8(a) and 8(b) visualize spending behavior among age groups and marketing budget allocation before and after optimization. The targeted campaign promotes premium experiences such as exclusive tours and luxury accommodations for the 30–45 age group in Table 11.

Case Study 2: Dynamic Budget Allocation

Table 10. Segment-wise Spending Analysis.

Age Group	Average Expenditure (\$)	Percentage Difference
18-29	100	–
30-45	135	35%
46-60	120	20%
60+	110	10%

<https://doi.org/10.1371/journal.pone.0321992.t010>

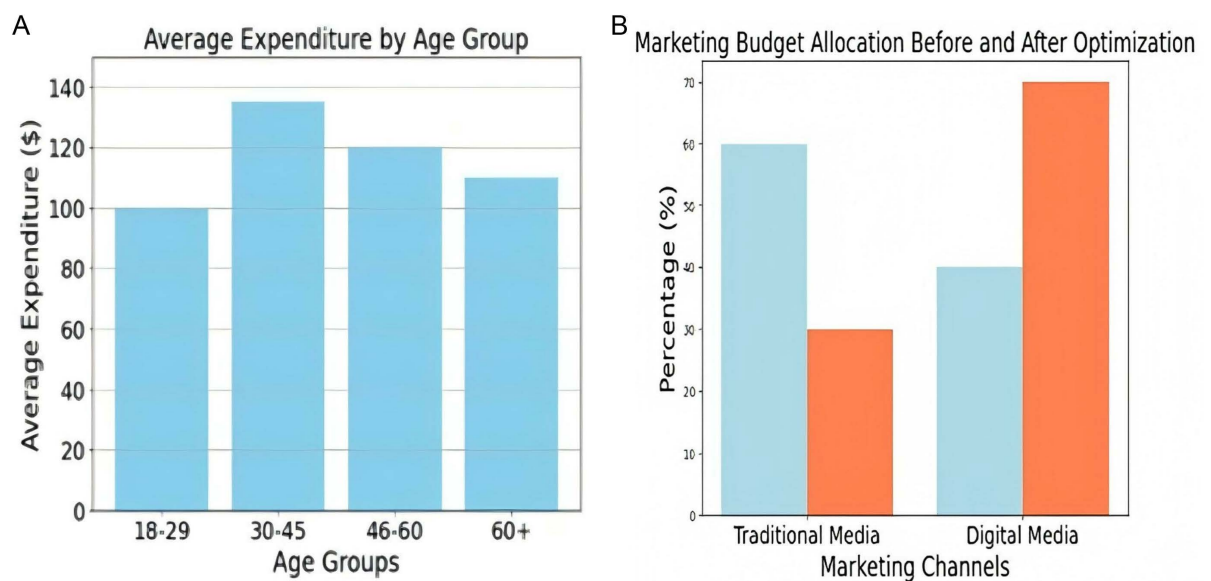


Fig 8. (a) Avg Expenditure Vs. Age Groups (b) Fig. Dynamic Budget Allocation.

<https://doi.org/10.1371/journal.pone.0321992.g008>

Table 11. Marketing Campaign Analysis.

Metric	Before Campaign	After Campaign	Increase
Marketing Investment (\$)	0	1,50,000	–
Revenue (\$)	0	6,00,000	400% ROI

<https://doi.org/10.1371/journal.pone.0321992.t011>

Table 12. Dynamic Budget Allocation Analysis (Pre-Post Optimization).

Marketing Channel	Pre-Optimization (%)	Post-Optimization (%)
Traditional Media	60%	30%
Digital Media	40%	70%

<https://doi.org/10.1371/journal.pone.0321992.t012>

TourVaRNN redirected resources to better-performing platforms after comparing the effectiveness of digital and conventional marketing channels.

[Table 12](#) provides shifted marketing spend to social media platforms, demonstrating higher engagement among younger demographics.

5 Conclusion

Ultimately, the study's findings have led to the creation and launch of the Tourism Variational Recurrent Neural Network (TourVaRNN). This tool improves economic effect forecasts and gives marketers helpful campaign information. Using variational neural networks with recurrent connections and variational autoencoders makes TourVaRNN a significant step forward in tourism research through its application. It delivers vital information for successful marketing strategies and tourist management in a global industry that is continually shifting. These insights are provided by precisely forecasting economic effects and visitor behaviors. The robust performance measures that characterize the model highlight its practical relevance in improving budget allocation and improving procedures for making choices within the industry. TourVaRNN's improved segmentation efficiency of 15.7 percent allows for more targeted marketing, increasing engagement with visitors and income. Decisions may be made in real-time, improving operational efficiency in tourism management, thanks to a 17.5% reduction in inference time (to 40 ms). The most efficient use of funds is guaranteed by a 13.4% rise in budget allocation utilization, leading to maximum economic benefits. These innovations support sustainable economic growth, better visitor experiences, and data-driven tourist policy.

5.1 Limitations

Although TourVaRNN has many advantages, it also has a few constraints, such as the possibility of being sensitive to data quality since it varies in different tourism regions. The interpretability of the model, its scalability to more enormous datasets, and integration with real-time data sources could be the focus of future developments. These enhancements would strive to improve prediction capabilities and responsiveness in tourism contexts that are constantly changing.

5.2 Applications

The research directly improves marketing strategies in the tourism industry by providing actionable insights through precise visitor behavior forecasts, enabling targeted campaigns and efficient budget allocation.

5.3 Future research scope

Furthermore, there is the possibility of expanding the use of TourVaRNN beyond economic impact prediction and visitor behavior prediction to encompass more significant areas of tourism sustainability, environmental impact evaluation, and the optimization of individualized tourist experiences. The research scope is extended to dynamically monitor the environment using real-time data from sensors and social media, create multi-task learning models to forecast economic and environmental repercussions, engage with environmental agencies to enhance datasets, implement sustainable tourist policies, and more. These updates will make TourVaRNN more flexible and encourage ecotourism.

TourVaRNN has the potential not just to contribute to more knowledgeable decisions in the tourist management industry but also to play a crucial part in creating tourism practices that are both environmentally responsible and technology-driven on a global scale.

Author contributions

Data curation: Ming Gao.

Formal analysis: Jing Zhang.

Funding acquisition: Ming Gao.

Software: Jing Zhang.

Visualization: Jing Zhang.

Writing – original draft: Jing Zhang.

Writing – review & editing: Ming Gao.

References

1. Puri V, Mondal S, Das S, Vrana VG. Blockchain propels tourism industry—an attempt to explore topics and information in smart tourism management through text mining and machine learning. *Informatics*. 2023;10(1):9.
2. Xu A, Siddik AB, Sobhani FA, Rahman MM. Driving economic success: Fintech, tourism, FDI, and digitalization in the top 10 tourist destinations. *Humanities and Social Sciences Communications*. 2024;11(1):1549.
3. Li P, Zhou Y, Huang S. Role of information technology in the development of e-tourism marketing: A contextual suggestion. *Economic Analysis and Policy*. 2023;78:307–18.
4. Viverit L, Heo CY, Pereira LN, Tiana G. Application of machine learning to cluster hotel booking curves for hotel demand forecasting. *International Journal of Hospitality Management*. 2023;111:103455.
5. Ismunandar I, Andriani NY, Hanis R, Hamzah R, Yusuf M. Grand preanger bandung effective marketing communication strategy in the staycation program. *Jurnal Ekonomi*. 2023;12(01):48–53.
6. Sharma GD, Thomas A, Paul J. Reviving tourism industry post-COVID-19: A resilience-based framework. *Tour Manag Perspect*. 2021;37:100786. <https://doi.org/10.1016/j.tmp.2020.100786> PMID: 33391988
7. Alkan T, Dokuz Y, Ececiş A, Bozdağ A, Durduran SS. Using machine learning algorithms for predicting real estate values in tourism centers. *Soft Computing*. 2023;27(5):2601–13.
8. Jardim S, Mora C. Customer reviews sentiment-based analysis and clustering for market-oriented tourism services and products development or positioning. *Procedia Computer Science*. 2022;196:199–206.
9. Imam T, Ananda J. Machine learning for characterizing growth in tourism employment in developing economies: an assessment of tourism employment in Sri Lanka. *Current Issues in Tourism*. 2022;25(16):2695–716.
10. Bulchand-Gidumal J, Secin E, O'Connor P, Buhalis D. Artificial intelligence's impact on hospitality and tourism marketing: exploring key themes and addressing challenges. *Current Issues in Tourism*. 2024;27(14):2345–62.

11. Ma H. Development of a smart tourism service system based on the Internet of Things and machine learning. *The Journal of Supercomputing*. 2024;80(5):6725–45.
12. Hu M, Li M, Chen Y, Liu H. Tourism forecasting by mixed-frequency machine learning. *Tourism Management*. 2025;106:105004.
13. Dalir S. Innovative strategies to tackle seasonality issue in hospitality and tourism industry. *International Journal of Contemporary Hospitality Management*. 2024;36(5):1690–709.
14. Zhang Y, Li G, Muskat B, Law R. Tourism demand forecasting: A decomposed deep learning approach. *Journal of Travel Research*. 2021;60(5):981–97.
15. Marigliano P. Analyzing tourism reviews using deep learning and AI to predict sentiments. *Quality & Quantity*. n.d.:1–17.
16. Li M, Zhang C, Sun S, Wang S. A novel deep learning approach for tourism volume forecasting with tourist search data. *Journal of Tourism Research*. 2022;25(2):183–97. <https://doi.org/10.1002/jtr.2558>
17. Ugli AAI. Quality and competitiveness of services in the tourism market. *Journal of Marketing, Business and Management*. 2023;1(12):145–8.
18. Dhote S, Vichoray C, Pais R, Baskar S, Mohamed Shakeel P. Hybrid geometric sampling and AdaBoost based deep learning approach for data imbalance in E-commerce. *Electronic Commerce Research*. 2020;20(2):259–74.
19. Polyakova O, Karagiorgos T, Konstantinidis K, Ntoli A, Alexandris K. Can event service quality predict memorable tourism experience in the context of a participatory recurring sport event? The case of oceanman international. *Event Management*. 2025.
20. Neshat N, Moayedfar S, Rezaee K, Amrollahi Biuki N. Sustainable planning of developing tourism destinations after COVID-19 outbreak: A deep learning approach. *Journal of Policy Research in Tourism, Leisure and Events*. 2024;16(1):1–21.
21. Pan H, Wei T, Li L. A strategy study based on artificial intelligence data analysis of rural governance into tourism industry development research. *Applied Mathematics and Nonlinear Sciences*. 2024.
22. Bi JW, Han TY, Yao Y. Collaborative forecasting of tourism demand for multiple tourist attractions with spatial dependence: A combined deep learning model. *Tourism Economics*. 2024;30(2):361–88.
23. Sulong Z, Abdullah M, Chowdhury MAF. Halal tourism demand and firm performance forecasting: new evidence from machine learning. *Current Issues in Tourism*. 2023;26(23):3765–81.
24. Elgarhy SD, Mohamed LM. The influences of services marketing mix (7ps) on loyalty, intentions, and profitability in the Egyptian travel agencies: The mediating role of customer satisfaction. *Journal of Quality Assurance in Hospitality & Tourism*. 2023;24(6):782–805.
25. Kyrlov Y, Hranovska V, Boiko V, Kwilinski A, Boiko L. International tourism development in the context of increasing globalization risks: On the example of Ukraine's integration into the global tourism industry. *Journal of Risk and Financial Management*. 2020;13(12):303.
26. Wei C. Tourist attraction image recognition and intelligent recommendation based on deep learning. *Journal of Computational Methods in Sciences and Engineering*. 2025;14727978251318805. <https://doi.org/10.3233/JCM-220001>
27. Siddik AB, Forid MdS, Yong L, Du AM, Goodell JW. Artificial intelligence as a catalyst for sustainable tourism growth and economic cycles. *Technological Forecasting and Social Change*. 2025;210:123875. <https://doi.org/10.1016/j.techfore.2024.123875>
28. Srinivasan J, Niranjane M, Nandhana C, Azhagiri M. A Text Classification in Tourism Reviews with the Hybrid Transformer-Attention Model in the Information Management of Smart Tourism. In *International Conference on Cognitive Computing and Cyber Physical Systems*. Springer, Singapore. 2025. pp. 479–88.
29. <https://ckan.publishing.service.gov.uk/dataset/visitor-economy-data/resource/ef4eeca4-ec98-4037-8211-cf8ed34f2734>
30. <https://data.europa.eu/en/publications/datastories/high-value-datasets-tourism-eu>