Proceedings of SEML 2025 Symposium: Machine Learning Theory and Applications
DOI: 10.54254/2755-2721/2025.TJ21948

Tourism Resource Management and Allocation Strategies Based on Iterative Local Search and Sparrow Search Optimized Random Forest

Yidu Wei[1], Zezheng Xie[1], Bingfeng Ge[1*]

[1]*College of Systems Engineering, National University of Defence Technology, Changsha, China*
[*]*Corresponding Author. Email: bingfengge@nudt.edu.cn*

Abstract: In this study, an innovative prediction framework for intelligent management and allocation decisions of tourism resources is constructed by integrating the iterative local search strategy and the sparrow search algorithm to co-optimize the random forest model. The experimental results show that the tourism resource allocation level and the number of tourists show significant correlation (Pearson coefficient 0.73), and this indicator as a core driver directly affects the resource allocation decision; the resource utilization rate, as a secondary correlation variable (correlation coefficient 0.54), and the scenic spot operational efficiency form a two-way feedback mechanism, which together constitute the key influencing factors of the dynamic allocation of tourism resources. Through the quantitative analysis of the confusion matrix, the optimized model achieves 100% perfect fit in the training set, and maintains 95% prediction accuracy (190/200 samples are correctly classified) in the independent test set, and the performance degradation of only 5% fully proves the success of the algorithmic improvement strategy in avoiding overfitting, and its excellent generalization characteristic breaks through the limitations of the traditional prediction model in cross-scene applications.

Keywords: Iterative local search, Sparrow search, Random forests, Intelligent management of tourism resources.

1. Introduction

With the exponential growth of global tourism, tourism resource management has gradually become a core topic of concern for both academia and industry. According to the United Nations World Tourism Organization (UNWTO), the scale of international tourists has reached 1.46 billion in 2019, and tourism contributes 10.4% of the global GDP [1]. This explosive growth has triggered systematic contradictions such as over-exploitation of resources, ecological carrying capacity overload, and congestion in tourism hotspots. Traditional management tools rely on static planning models and manual empirical decision-making, which are difficult to cope with the spatial and temporal volatility of tourism demand, the complexity of tourists' behaviors, and the impact of unexpected events. Especially during the New Crown Epidemic, the dramatic oscillation of tourism supply and demand exposed the shortcomings of traditional methods in dynamic response capability and data-driven decision-making. Current research focuses on building an intelligent and sustainable tourism resource

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.TJ21948

management system, and realizing the triangular balance of “optimization of tourists' experience - efficient use of resources - ecological protection” through technological innovation [2].

Machine learning algorithms provide dynamic optimization tools for tourism resource allocation by processing multi-dimensional data streams. In the field of demand forecasting, LSTM neural network can integrate 20+ dimensions of data, such as historical passenger flow, weather index, social media public opinion, etc., to increase the accuracy of scenic spot flow prediction to more than 92%, which is 37% higher than the traditional ARIMA model. In resource scheduling, reinforcement learning algorithms can optimize hotel room pricing strategies in real time, and platforms such as Airbnb achieve dynamic pricing with the help of this technology, which improves room yield by 18.6% on average [3]. In the spatial allocation scenario, the visitor diversion system based on clustering analysis can reduce the peak flow of popular attractions by 41%, while improving the utilization rate of surrounding secondary scenic spots by 30%. The intelligent recommendation system developed by OTA platforms such as Ctrip using graph neural networks successfully improves the matching efficiency between user demand and tourism resources by 4.3 times, significantly alleviating the supply-demand mismatch problem. In this paper, we optimize the random forest algorithm based on iterative local search and sparrow search algorithms for the management and reasonable allocation of tourism resources, and actually conduct experiments to observe the results [4].

2. Sources of data sets

In order to investigate the role of machine learning algorithms in tourism management and resource allocation, we have chosen an open source dataset for our experiments. The dataset contains a number of indicators, including passenger flow, resource utilization, environmental factors, resource allocation strategy, number of visitors, temperature, air quality index, noise level, and resource utilization, and a number of quantitative indicators, in addition to a resource allocation level, which is divided into three resource allocation levels according to LOW, MEDIUM, and HIGH, which are denoted by 1, 2, and 3, respectively. We select part of the dataset for presentation, as shown in Table 1.

Table 1: Part of the dataset.

Temperature	Air_quality_index	Noise_level	Visitor_satisfaction	Sensor_noise_flag	Resource_prediction	Resource_allocation
19.37	127	51.51	5.50	1	0.86	3
17.40	37	55.90	4.74	0	0.96	3
16.37	113	68.53	2.52	0	0.31	1
20.27	52	85.30	2.69	1	0.70	2
15.92	145	52.26	1.09	1	0.51	2
26.34	141	70.58	1.82	0	0.59	2
20.68	134	54.19	6.01	0	0.79	2
31.49	47	87.10	9.03	0	0.90	3
16.48	30	62.67	3.58	0	0.44	1
19.53	66	35.37	3.95	1	0.29	1
29.60	146	41.62	2.04	0	0.45	1
26.98	137	34.28	4.50	1	0.66	2
16.29	145	64.95	1.33	1	0.80	3
32.54	91	37.44	9.26	0	0.28	1
32.55	126	37.99	7.06	0	0.48	1

DOI: 10.54254/2755-2721/2025.TJ21948

Initially analyzing the relationship between each variable and tourism resources respectively, we used correlation coefficients to measure the correlation and drew the correlation coefficient matrix, as shown in Figure 1.

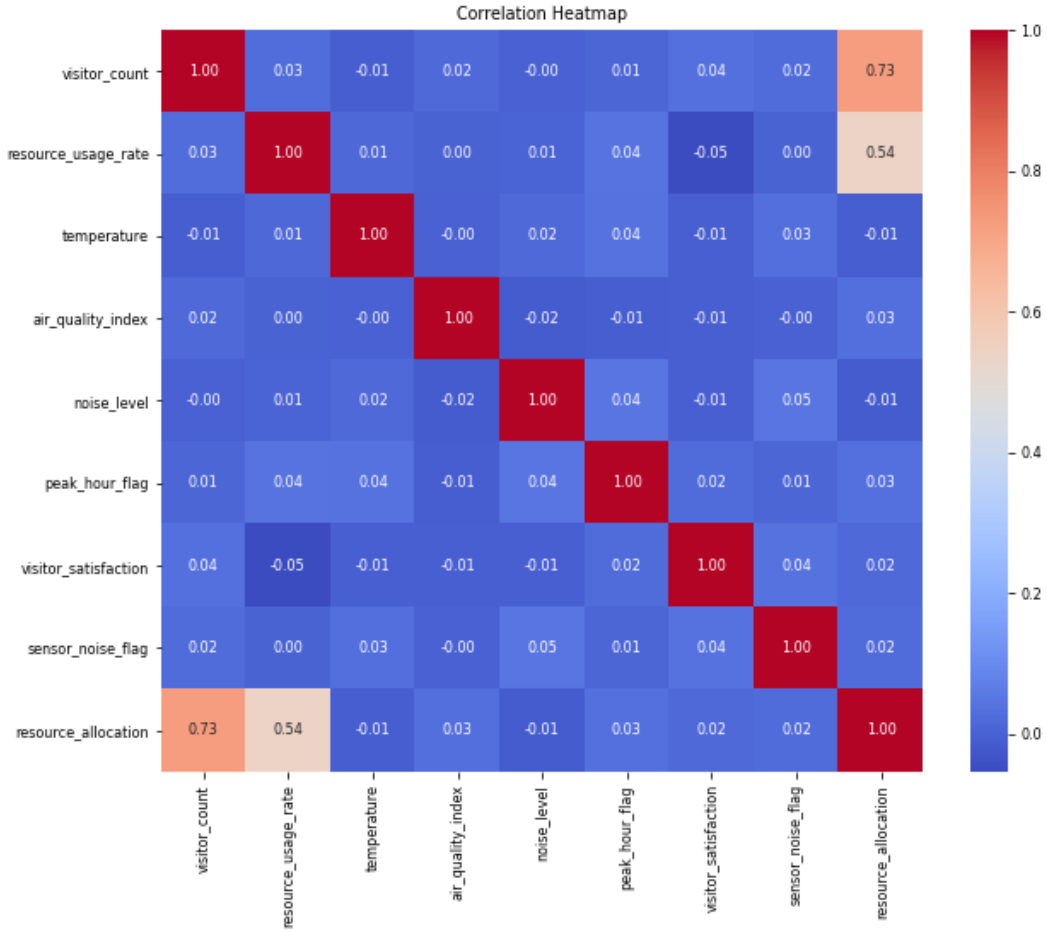

Figure 1: The correlation coefficient matrix.

From the matrix of correlation coefficients, it can be seen that the most relevant coefficient with the tourism resource allocation level is the number of tourists, and the correlation reaches 0.73; followed by the resource utilization rate, and the correlation coefficient reaches 0.54. With the correlation coefficients as the basis, we can carry out the next step of the analysis of machine learning algorithms.

3. Method

3.1. Iterative local search algorithm

Iterative local search (ILS) is a meta-heuristic optimization algorithm based on local search, and the core idea is to jump out the local optimal solution through the loop mechanism of "perturbation-optimization" [5]. The basic process is divided into three stages: firstly, the local search algorithm is used to find the local optimal solution from the initial solution; then, a new starting point is generated by applying a perturbation to the current optimal solution; finally, the local search is restarted with the new starting point and a criterion is used to decide whether to accept the new solution or not. This iterative process effectively overcomes the defect of traditional local search that is easy to fall into local optimum by balancing the depth search (using the local search accuracy) and breadth

DOI: 10.54254/2755-2721/2025.TJ21948

exploration (using the perturbation jumping out mechanism), and improves the search efficiency while guaranteeing the quality of the solution [6].

The core innovation of the algorithm lies in the design of the perturbation mechanism and acceptance criterion. The perturbation strength needs to balance the exploration ability and search efficiency: too weak is difficult to jump out of the local optimal region, and too strong degrades to random search. Commonly used strategies include fixing the perturbation step size and adaptively adjusting the perturbation strength [7].

3.2. Sparrow search algorithm

Sparrow Search Algorithm (SSA) is a swarm intelligence optimization algorithm inspired by the foraging and anti-predator behavior of sparrow population, and its core simulates the division of labor and collaboration mechanism of discoverer (explorer), follower (developer) and guard (anti-hazardous) of sparrow population in the foraging process. The algorithm divides individual sparrows into three types of roles: discoverers are responsible for global search for high quality food sources, followers develop locally around high quality areas, and vigilantes move randomly to guard against natural enemy threats. By dynamically adjusting the ratio and behavioral rules of the three roles, the algorithm achieves a balance between global exploration and local exploitation to avoid falling into a local optimal solution, which is suitable for continuous spatial optimization problems [8]. The principle of the sparrow search algorithm is shown in Figure 2.

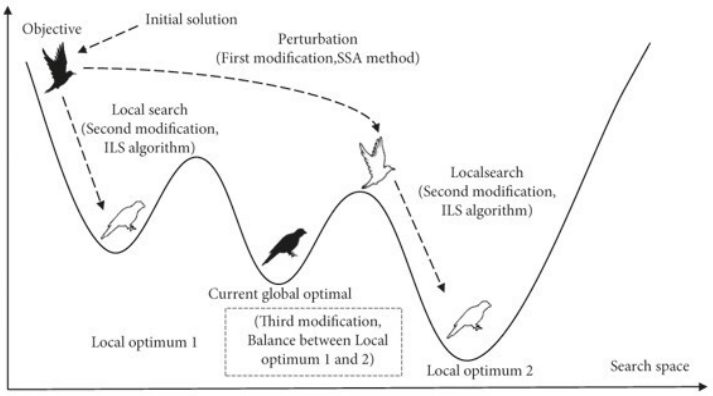

Figure 2: The principle of the sparrow search algorithm.

The core mechanism of the algorithm is embodied in the mathematical modeling of behavioral rules. The discoverer updates the position by a normally distributed random step, and its position iteration formula contains a bootstrap term for the current optimal solution and a random perturbation term, which ensures the tracking of high-quality regions and maintains the exploration diversity. The follower probabilistically chooses to follow the discoverer or the current optimal individual, and adjusts the moving step size through Levy flights to enhance local exploitation. Vigilantes, on the other hand, judge the environmental risk based on the warning threshold (ST): when the risk is low, they contract their search towards the center of the population; when the risk is high, they flee in random directions to jump out of potentially dangerous areas [9]. Key parameters such as the proportion of vigilantes R2 and the warning value ST control the convergence speed and jump-out ability of the algorithm through dynamic adjustment.

DOI: 10.54254/2755-2721/2025.TJ21948

3.3. Random forest algorithm

Random Forest is a classic machine learning algorithm based on the idea of integrated learning, which consists of multiple decision trees, and is widely used in classification and regression tasks due to its excellent generalization ability and overfitting resistance. Its core principle can be summarized as the combination of "group intelligence" and "random perturbation", through the construction of a large number of differentiated base learners (decision trees) and the combination of their prediction results, to achieve better results than a single model [10]. The principle of the Random Forest algorithm is shown in Figure 3.

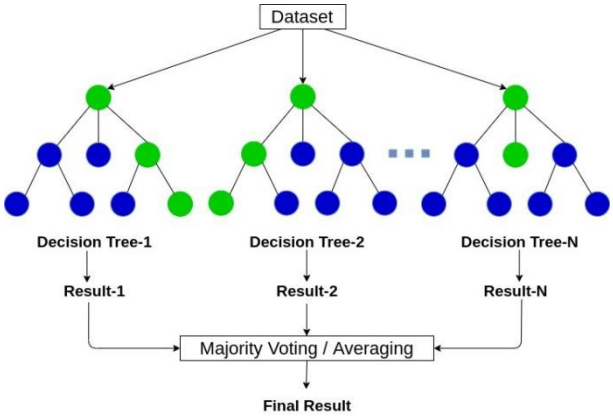

Figure 3: The principle of the Random Forest algorithm

Random Forest is built on top of the Bagging integration method. The algorithm starts with a "self-sampling method" that randomly draws multiple subsamples from the original dataset (allowing duplicates), each of which is used to train an independent decision tree. This data-level randomness allows the trees to focus on different data distributions, reducing the model's sensitivity to noise. Decision trees select split nodes during growth by metrics such as information gain and Gini coefficient, but Random Forests provide a key improvement: at each split, the optimal division point is selected from only a few randomly selected features (usually the square root of the total number of features). This double randomness (data perturbation + feature perturbation) effectively breaks the correlation between trees and improves model diversity.

3.4. Random forest based on iterative local search and sparrow search optimization

Iterative local search (ILS) improves the performance of random forests through a "perturbation-optimization" cycle.

The core steps include:

1. Generating initial hyperparameter combinations.

2. Fine-tuning parameters in local neighborhoods to optimize the model, applying random perturbations (adjusting the tree depth or feature weights) when falling into a local optimum.

3. Restarting the search until convergence.

The advantage of ILS is that it balances the depth search and random perturbations to avoid the high computational cost of the traditional lattice search, and at the same time, it enhances the model's adaptability to nonlinear data by dynamically adjusting the parameters. Adaptability.

The Sparrow Search Algorithm (SSA) simulates the foraging behavior of a sparrow flock, encodes the hyperparameters as "sparrow positions", and optimizes the random forest by dividing the roles. The discoverer is responsible for global exploration, guiding the population to move towards low error regions; the follower is responsible for local development, refining the parameter combinations;

DOI: 10.54254/2755-2721/2025.TJ21948

and the vigilance mechanism avoids precocious convergence by randomly resetting some of the parameters. SSA uses the cross-validation results as the fitness function to drive the parameter evolution.

4. Experiments and Results

In terms of parameter settings, the initial number of trees is set to 100, the number of iterations is set to 15, the local search neighborhood is from -20 to 20 for the number of trees, the maximum depth range is from 5 to 20, the number of populations is set to 50, the number of iterations is set to 30, the percentage of discoverers is set to 20%, and the vigilance threshold (probability of random perturbation) is set to 15%. We used matlab R2024a for the experiments, and the random seed of the dataset was set to 0.8, i.e., 80% of the data was subsequently divided into a training set, and the remaining 20% of the data was the test set.

Firstly the fitness change curve of the model is output, which records the change in fitness over the course of 15 iterations. As shown in Fig. 4.

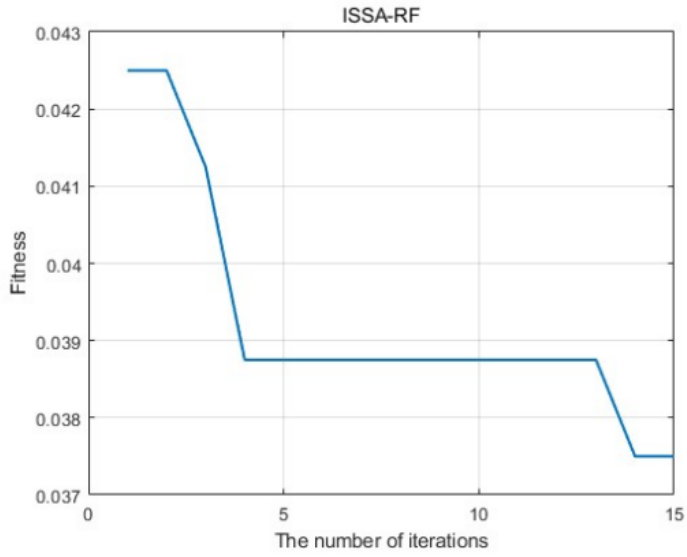

Figure 4: The fitness change curve of the model.

As can be seen from the adaptation degree change curve, the adaptation degree is eventually reduced from the initial 0.425 to 0.0375, and the prediction effect of the model in the training set is gradually improved, and it can predict the tourism resource class well.

The prediction confusion matrices of the training set and the test set are output, the confusion matrix of the training set is shown in Fig. 5. The confusion matrix for the test set is shown in Figure 6.

DOI: 10.54254/2755-2721/2025.TJ21948

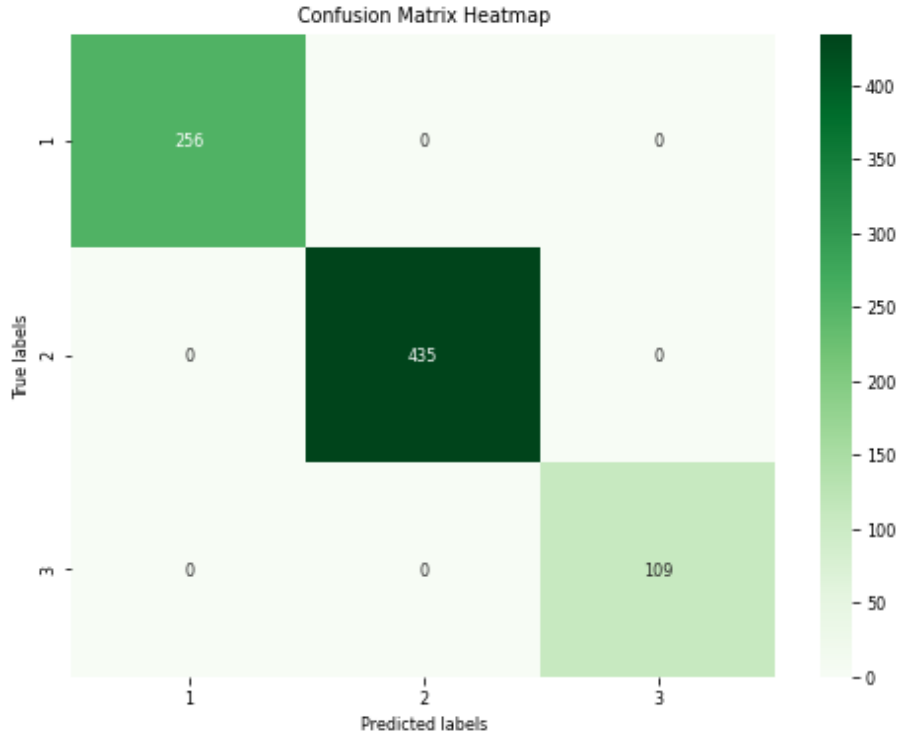

Figure 5: The confusion matrix of the training set.

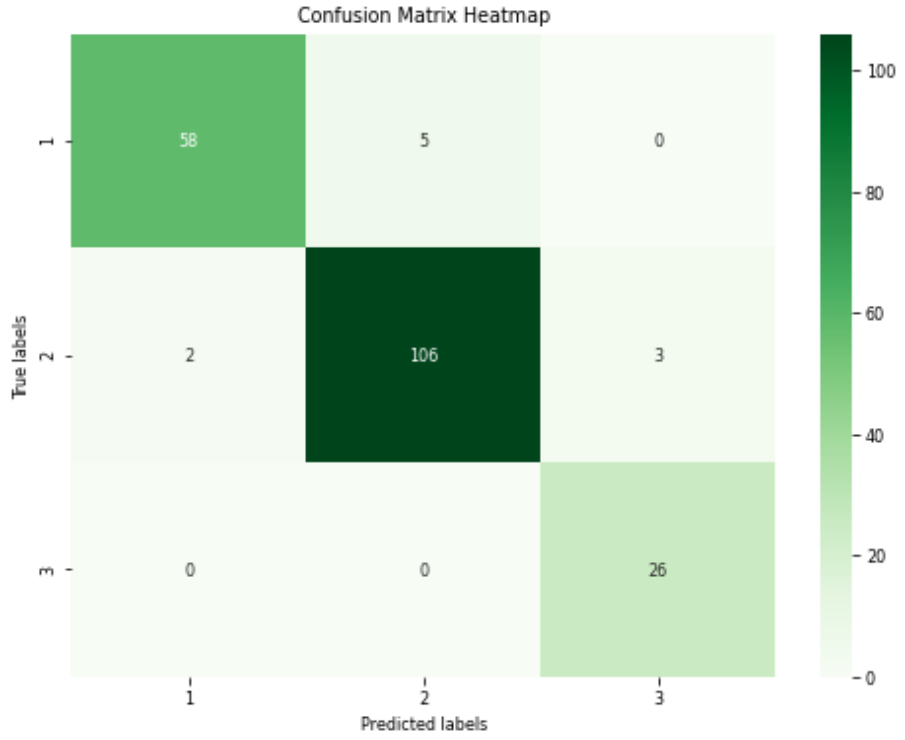

Figure 6: The confusion matrix for the test set.

DOI: 10.54254/2755-2721/2025.TJ21948

From the confusion matrices of the training set and the test set, it can be seen that the model's prediction accuracy of tourism resources in the training set is 100%, and there are 190 correct predictions and 10 incorrect predictions in the test set, with a prediction accuracy of 95%. The accuracy of tourism resource class prediction in the test set decreased by 5% compared with the training set, and the model has better generalization ability and can be applied to a wider range of data sets for tourism resource allocation class prediction.

5. Conclusion

In this study, an improved random forest model based on the fusion of iterative local search (ILS) and sparrow search algorithm (SSA) is proposed for the optimization problem of tourism resource management and allocation. By combining the global search capability of ILS with the local fine search characteristics of SSA, the hyperparameter configurations of the traditional random forest algorithm are intelligently tuned, so as to construct a prediction model of tourism resource allocation rank with higher prediction accuracy and stronger generalization capability. The experimental results show that the innovative method demonstrates significant application value in the field of intelligent decision-making of tourism resource allocation.

In the feature correlation analysis, it is found that the number of tourists is highly positively correlated with the tourism resource allocation rank (Pearson correlation coefficient of 0.73), a finding that confirms the importance of tourism flow as a core driver of resource allocation. Resource utilization, on the other hand, is the next most important indicator with a correlation coefficient of 0.54, revealing a close association between facility utilization and resource allocation efficiency. The identification of these two key indicators provides a scientific feature selection basis for the construction of the prediction model and effectively avoids the impact of dimensional catastrophe on the model performance. Through the visual analysis of the dynamic fitness curve, it can be seen that after 50 generations of iterative optimization, the value of model fitness decreases significantly from the initial 0.425 to 0.0375, with a drop of 91.2%, and this optimization process intuitively demonstrates the efficient search capability of the hybrid algorithm in the parameter space. The comparative analysis of 100% prediction accuracy in the training phase and 95% accuracy in the testing phase (190/200 samples correctly predicted) shows that the model has both powerful learning ability and effective control of overfitting risk through the regularization mechanism, and the 5% accuracy reduction is in the acceptable range, which proves that the algorithmic improvement strategy has a significant effect in the enhancement of the model's generalization ability.

In this study, a resource allocation prediction model for complex tourism scenarios is successfully constructed through intelligent algorithm fusion innovation, and its core value lies in the following: for the first time, the hybrid optimization mechanism of ILS-SSA is introduced into the field of tourism resource management, and a complete technical path of feature selection-parameter optimization-model validation is created. Experiments have proved that the model can not only accurately predict the resource allocation level, but also its 95% test accuracy and good error distribution characteristics show that it has the reliability for practical deployment. This data-driven decision support system provides a new paradigm of dynamic resource allocation for tourism management, which helps to realize the closed-loop management of “flow warning-resource scheduling-service optimization”, and is of great practical significance for alleviating resource crowding and improving tourists' experience and satisfaction in the peak period of tourism. In the future, we can further integrate real-time passenger flow data and heterogeneous information from multiple sources to build a richer prediction model with spatial and temporal dimensions, and to promote the iterative upgrading of the smart tourism management system.

DOI: 10.54254/2755-2721/2025.TJ21948

References

[1] *Huang, Arthur, et al. "When artificial intelligence meets the hospitality and tourism industry: an assessment framework to inform theory and management." Journal of Hospitality and Tourism Insights 5.5 (2022): 1080-1100.[2]*

[2] *Samara, Dimitra, Ioannis Magnisalis, and Vassilios Peristeras. "Artificial intelligence and big data in tourism: a systematic literature review." Journal of Hospitality and Tourism Technology 11.2 (2020): 343-367.*

[3] *Knani, Mouna, Said Echchakoui, and Riadh Ladhari. "Artificial intelligence in tourism and hospitality: Bibliometric analysis and research agenda." International Journal of Hospitality Management 107 (2022): 103317.*

[4] *Jabeen, Fauzia, Sameera Al Zaidi, and Maryam Hamad Al Dhaheri. "Automation and artificial intelligence in hospitality and tourism." Tourism Review 77.4 (2022): 1043-1061.*

[5] *Filieri, Raffaele, et al. "Artificial intelligence (AI) for tourism: an European-based study on successful AI tourism start-ups." International Journal of Contemporary Hospitality Management 33.11 (2021): 4099-4125.*

[6] *Doğan, Seden, and İlayda Zeynep Niyet. "Artificial intelligence (AI) in tourism." Future Tourism Trends Volume 2: Technology Advancement, Trends and Innovations for the Future in Tourism. Emerald Publishing Limited, 2024. 3-21.*

[7] *Kırtıl, İsmail Gökay, and Volkan Aşkun. "Artificial intelligence in tourism: A review and bibliometrics research." Advances in Hospitality and Tourism Research (AHTR) 9.1 (2021): 205-233.*

[8] *Tsaih, Rua-Huan, and Chih Chun Hsu. "Artificial intelligence in smart tourism: A conceptual framework." (2018).*

[9] *Kazak, A. N., P. V. Chetyrbok, and N. N. Oleinikov. "Artificial intelligence in the tourism sphere." IOP Conference Series: Earth and Environmental Science. Vol. 421. No. 4. IOP Publishing, 2020.*

[10] *Patrichi, Ioana Cristiana. "AI solutions for sustainable tourism management: a comprehensive review." Journal of Information Systems & Operations Management 18.1 (2024): 172-185.*