Research On Fuzzy Clustering Recommendation System For Intelligent Tourism Management

Luxi Chen[1] and Yanna Huang[2*]

[1] School of History and Law, Yulin Normal University, Yulin, Guangxi 537000, China
[2] Student Affairs Office, Yulin Normal University, Yulin, Guangxi 537000, China
[*] Corresponding author. E-mail: hyn8545@126.com

Received: May 09, 2026; Accepted: Jun. 26, 2026

Tourism analytics is essential for understanding traveler behavior and supporting effective destination management. This study proposes a hybrid framework that combines deep learning, evolutionary optimization, and fuzzy clustering to improve tourism data segmentation. A realworld tourism dataset is preprocessed using Min-Max normalization and one-hot encoding. A Variational Autoencoder (VAE) is employed to extract compact latent features from highdimensional data, while Reinforced Evolution Strategy (RES) selects the most informative features by optimizing information gain and reconstruction quality. Fuzzy C-Means (FCM) clustering is then applied to categorize tourism records while handling uncertainty through soft cluster memberships. The clustering performance is assessed using Partition Coefficient (PC), Partition Entropy (PE), and Xie-Beni Index (XB). Results demonstrate effective cluster formation, achieving a PC of 0.65, PE of 0.68 , and clear cluster separation. The proposed framework also demonstrated strong destination retrieval and ranking capability, achieving Precision@10 of 0.89 , Recall@10 of 0.86 , F1-Score of 0.88 , MAP of 0.91 , and NDCG of 0.93 . Compared to previous works, the model showed improved compactness and reduced ambiguity in cluster membership.

Keywords: Tourism Recommendation, Variational Autoencoder, Reinforced Evolution Strategy, Fuzzy C-Means Clustering, Clustering Evaluation.

http://dx.doi.org/10.6180/jase.202611_34.031

1. Introduction

Tourism serves as a foundation for economic development, cultural exchange, and regional identity [1]. With increased globalization and better availability of information, travelers increasingly rely on intelligent recommender systems to help them design a trip according to their special needs [2]. Recommendation systems are able to help tourists navigate destination choices, travelling routes, and accommodation planning with their preferences and historical data to suggest options [3, 4]. Fuzzy logic has demonstrated some advantages in uncertain environments over the traditional recommendation systems based on crisp values [5]. Given the variety of factors geographical, emotional, and financial that influence a visitor's experience, intelligent tourism management systems require sophisticated tools that are adaptive to user needs, can learn, cluster and work with changing content, and recall user interests [6]. Machine learning and soft computing methods are being exposed to and incorporated into tourism analytics [7]. Theoretically they will improve the decision-making performance and ability to cluster tourist attraction choices [8]. When working with the complexity of individual and non-individual datasets, recommendation systems based on clustering have become more desirable, as users are now demanding more intelligent, personalized, and sensitive recommendations [9]. Research in this area will begin with creating frameworks that are accurate, scalable, and explainable, while staying consistent with the way visitors now approach tourism [10]. The novelty of the proposed framework lies in the hybrid integration of Variational Autoencoder (VAE), Reinforced Evolution Strategy (RES), and Fuzzy C-Means (FCM) clustering for intelligent tourism segmentation. This integration improves cluster compactness, reduces membership ambiguity, and better captures overlapping tourism preferences in high-dimensional tourism data. The major contributions of this study are summarized as follows:

- A hybrid intelligent tourism clustering framework combining VAE, RES, and FCM is proposed for effective tourism behavior segmentation.
- A VAE-based latent representation strategy is employed to reduce high-dimensional tourism data while preserving important behavioral patterns.
- The proposed framework improves fuzzy clustering quality by reducing cluster ambiguity and enhancing cluster compactness, as validated using PC, PE, and XB evaluation metrics.

1.1. Research Objectives

- Analyze the capability of VAE in extracting meaningful low-dimensional latent features from high-dimensional tourism data.
- Implement RES to optimize the selection of informative features for improving clustering precision.
- Apply FCM clustering to identify overlapping tourism behavior patterns and classify destinations based on fuzzy membership.

1.2. Research Organization

The paper is divided into sections. The first section introduces. Section 2 reviews the literature, Section 3 deals with the dataset and preprocessing. Section 4 deals with the VAE for feature extraction and RES for feature selection. Section 5 concerns the treatment of fuzzy clustering generating membership, and clustering analysis. Section 6 analyzes the experiments. Finally, the last section ends the work with a conclusion of the study and scope for future work.

2. Literature review

J. Cui et. al., 2025 [11] constructed a smart tourism recommendation system combining big data analytics and contextual ontology modeling. With this technique, the system properly collects user preferences through conversion of discrete and interval attributes into matrix and vector forms and supplies personalized travel recommendations. J. Shen et. al., 2016 [12] introduced a hybrid recommendation algorithm involving fuzzy clustering and attribute-level feature fusion. By applying their approach to preprocess user data, cluster tourists and destinations, as well as fine-tune the fusion of the most essential behaviour-based features improve the view of tour route suggestions.

L. Yang, 2022 [13] were the first studied over travel recommendation to incorporate user interaction and collective intelligence in an attempt to overcome the limitations of traditional collaborative and content-based systems. Later, Huang et al., [14] proposed a smart tourism system by integrating FCM clustering with the Bat Optimization Algorithm (BOA), attaining a clustering accuracy of 98% and showing the usefulness of its IoT integration. Another system was provided by Asaithambi et al., [15] proposed a hybrid model for thematic travel planning through multimodal big data analytics, CNN, NLP, and spark-based learning techniques.

2.1. Problem Statement

Traditional fuzzy clustering approaches frequently exhibit instability and a high sensitivity to noise and variation in datasets, typically undermining their reliability in tourism segmentation studies [16]. A number of tourism papers make the clear mistake of using appropriate clustering methods on mixed-type data irrespective of the clustering algorithm's suitability [17].

3. Proposed fuzzy clustering-based tourism recommendation methodology

The proposed framework begins with the acquisition of a publicly available tourism dataset from Kaggle containing records from countries such as India, Japan, etc. The overall work is shown in Fig. 1.

3.1. Dataset Description

The Kaggle Tourism Dataset [18] contains 5,989 tourism records with both numerical features (visitors, ratings, revenue) and categorical features (country, tourism category, and travel purpose). The dataset contains both categorical and numerical attributes. Min-Max normalization and one-hot encoding were applied during preprocessing, followed by VAE-based feature extraction, RES-based feature selection, and FCM clustering for tourism behavior segmentation.

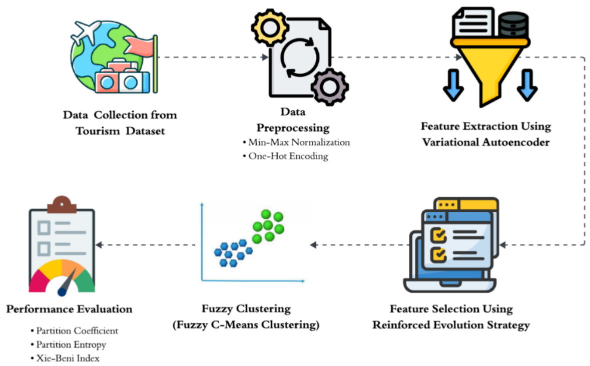

Fig. 1. Proposed Intelligent Fuzzy Clustering Framework for Tourism Data Analysis.

3.2. Data Preprocessing

The Kaggle tourism dataset contains both numerical and categorical attributes that require distinct preprocessing strategies.

3.2.1. Min-Max Normalization

This is a process used to put a range to data scaling, wherein values of a feature variable are brought into a pre-specified range, usually [0, 1]. Thus, the standard transformation is described by Eq. (1):

$$\mathrm{X}' = \frac{\mathrm{x} - \mathrm{x}_{\min}}{\mathrm{x}_{\max} - \mathrm{x}_{\min}} \tag{1}$$

where, X refers to the original value, $X_{\min}$ is the minimum value for the feature, whereas $X_{\max}$ is the maximum value. Min-Max normalization scales numerical tourism attributes such as visitors, ratings, and revenue into a uniform range between 0 and 1. This prevents features with larger numerical values from dominating the clustering process and improves the stability and accuracy of VAE feature extraction and FCM clustering.

3.2.2. One-Hot Encoding

It generates n binary features for any categorical variable having n unique classes. One-Hot Encoding transforms categorical tourism attributes such as country, tourism category, and travel purpose into binary numerical vectors without introducing ordinal relationships.

3.3. Feature Extraction

In this framework, feature extraction is performed using a Variational Autoencoder (VAE) to transform the high-dimensional preprocessed tourism data into a compact latent representation.

3.3.1. Variational Autoencoder

The VAE architecture consists of a symmetric encoder-decoder structure. The encoder comprises two fully connected dense layers with 128 and 64 neurons respectively, each followed by ReLU activation, which introduces non-linearity while avoiding vanishing gradient issues. The encoder outputs two vectors mean (μ) and log-variance $(\log \sigma^2)$ each of dimension 16, representing the latent space parameters. The reparameterization trick samples latent vector z as $z = \mu + \sigma \odot \varepsilon$, where $\varepsilon \sim \mathrm{N}(0, \mathrm{I})$,

A reconstruction convergence analysis was conducted to verify the stability of latent feature learning during VAE training. The reconstruction loss decreased steadily from 0.184 at epoch 10 to 0.132 at epoch 20, 0.098 at epoch 30 , and 0.081 at epoch 40 , indicating rapid learning of the underlying tourism data patterns during the initial training phase. Further reductions were observed at epochs 50 and 60 with reconstruction losses of 0.072 and 0.068 , respectively. After epoch 60, the loss values stabilized gradually, reaching 0.065, 0.063, 0.062, and finally 0.061 at epochs 70, 80, 90, and 100.

VAE Loss Function

The VAE model consisted of encoder and decoder dense layers with ReLU activation. The latent dimension was set to 16 after empirically evaluating dimensions $\{8, 16, 32, 64\}$. Smaller dimensions (e.g., 8) caused high reconstruction error, losing discriminative tourism features, while larger dimensions (e.g., 32, 64) introduced redundancy, reducing FCM clustering stability. The model was trained using the Adam optimizer with a learning rate of 0.001 , batch size of 32, and 100 epochs.

The VAE latent embeddings, Euclidean distances between mean latent representations (μ) of the three tourism categories were computed in the 16 -dimensional latent space. The distances Beach vs. Cultural (2.84), Beach vs. Historical (2.61), and Historical vs. Cultural (1.93) confirm that semantically distinct categories are positioned farther apart, while thematically similar categories remain closer.

KL Divergence

Kullback−Leibler (KL) Divergence quantifies the difference between two probability distributions, typically a learned approximation and a prior, as in Eq. (2).

$$\mathrm{D}_{\mathrm{KL}}(\mathrm{P} \| \mathrm{Q}) - \sum_{\mathrm{i}} \mathrm{P}(\mathrm{i}) \cdot \log \frac{\mathrm{P}(\mathrm{i})}{\mathrm{Q}(\mathrm{i})} \tag{2}$$

Were, $P(i)$: True distribution (e.g., latent encoding from encoder), $Q(i)$: Prior distribution (usually standard normal $N(0, 1)$, Σ_i : Summation over all latent dimensions. The expected loglikelihood term in the VAE loss function minimizes the reconstruction error between the input tourism records and the reconstructed output, enabling the model to preserve important tourism characteristics such as visitor preferences, ratings, and travel patterns.

3.4. Feature Selection

Following VAE-based feature extraction, a feature selection step is introduced to identify the most informative latent dimensions for FCM clustering.

3.4.1. Reinforced Evolution Strategy

RES was preferred over GA, PSO, and ACO since GA has high computational overhead, PSO converges prematurely in high-dimensional latent spaces, and ACO involves complex pheromone mechanisms unsuitable for continuous feature optimization. The perturbation operator used in RES is isotropic Gaussian mutation RES was implemented with a population size of 50 and executed for 100 iterations using Gaussian mutation for feature exploration.

Fitness Function (Information Gain + Error)

The fitness function evaluates the quality of selected features by combining their information relevance and reconstruction error. Eq. (3) defines the function as:

$$F(x) = \alpha \cdot \mathrm{IG}(x) + \beta \cdot \mathrm{RE}(x), \text{ subject to } \alpha + \beta - 1 \tag{3}$$

Here, α and β are tunable weights, Information Gain(x) quantifies feature usefulness, and Reconstruction Error(x) measures how accurately VAE reconstructs the input. RE(x)Reconstruction Error measures VAE reconstruction quality. α, β-Weighting coefficients; $\alpha = 0.6,\ \beta = 0.4$ in this study. The RES framework, Information Gain evaluates the relevance of features by measuring their ability to reduce uncertainty and improve discrimination among tourism patterns.

Evolutionary Update Rule

The reward signal adapts dynamically based on each candidate subset's fitness relative to the population mean. Subsets performing above the mean receive a positive reward, reinforcing that search direction, while below-average subsets receive a negative reward, suppressing it. The strategy adds weighted noise-based exploration to improve the solution iteratively. It is given by Eq. (4):

$$\theta_{t+1} = \theta_t + \eta \cdot \frac{1}{n} \sum_{i=1}^{n} r_i \cdot \epsilon_i \tag{4}$$

Where, θ_t denotes Current parameter vector at iteration t, θ_{t + 1} - Updated parameter vector, η - Learning rate controlling update step size, n-Population size ($\mathrm{n} = 50$ in this study), r_i -Reward signal for the i-th candidate solution, ε_i-Gaussian noise vector for exploration. The reward signal $\mathrm{r_i}$ is computed as $\mathrm{r_i} = \mathrm{F}(\mathrm{x_i}) - \overline{\mathrm{F}}$, where $\overline{\mathrm{F}}$ is the population mean fitness. Subsets with above-average fitness receive positive rewards, reinforcing beneficial mutations, while below-average subsets receive negative feedback, suppressing poor feature choices.

3.5. Fuzzy Clustering

Fuzzy C-Means (FCM) is adopted in this framework to assign soft membership values to tourism records, enabling partial cluster membership that reflects overlapping tourist preferences.

3.5.1. Fuzzy C-Means Algorithm

The fuzzifier parameter m was set to 2 , the most widely adopted value in FCM literature, as it provides a balanced trade-off between crisp and overly diffuse membership assignments. When $m \to 1$, memberships approach hard clustering (0 or 1), losing the ability to capture overlapping preferences; when m is large, all memberships converge toward equal values ($1/C$), producing meaningless soft boundaries. Fuzzifier values of $\{1.5, 2.0, 2.5, 3.0\}$ were tested. At $\mathrm{m} = 1.5$, hard boundaries the membership function remains sensitive to inter-centroid distances, enabling the model to reflect genuine uncertainty in tourist preferences where a visitor may simultaneously belong to Beach and Cultural clusters without collapsing into ambiguity. Centroid initialization sensitivity in high-dimensional latent spaces, FCM was executed with multiple random. FCM centroid initialization was performed by randomly generating a normalized fuzzy membership matrix, from which initial centroids were derived.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \tag{5}$$

This membership function (Eq. (5)) calculates the degree of membership uij of data point xi in cluster j. The term $\|xi - cj\|$ represents the Euclidean distance between the data point and the cluster centroid cj. The FCM algorithm, each tourism record is assigned a membership value between 0 and 1 for all clusters based on the Euclidean distance from the cluster centroids. Higher membership values are assigned to nearer clusters. The centroids are iteratively updated using weighted membership values, and the objective function is minimized through repeated updates until convergence is achieved.

4. Result and discussion

All experiments were conducted using identical preprocessing and clustering settings to ensure consistency and reproducibility. Numerical features were normalized using Min-Max normalization, while categorical attributes were transformed using one-hot encoding. The VAE architecture employed latent dimension size 16, batch size 32, learning rate 0.001, and 100 training epochs.

4.1. Performance metrics

Performance Metrics-are the primary indicators that go to evaluate the fuzzy clustering outcome. The Partition Coefficient (PC) quantifies the degree of crispness in a fuzzy clustering solution sum of squared membership degrees of data points in the clusters using Eq. (6).

$$PC = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{2} \tag{6}$$

Partition Entropy tests the amount of fuzziness in a clustering solution. This can be analysed by Eq. (7),

$$PE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}\log\left(u_{ij}\right) \tag{7}$$

XB maintains the balance between cluster compactness and separation. This can be analysed by Eq. (8),

$$XB = \frac{\sum_{i-1}^{N}\sum_{j-1}^{C} u_{ij}^{m}\left\|x_i - v_j\right\|^2}{N\cdot \min_{j+k}\left\|v_j - v_k\right\|^2} \tag{8}$$

4.2. Model Validation

The bar chart in Fig. 2 shows the proportions of tourism data points found in the three fuzzy clusters: Historical, Cultural, and Beach. The bar chart also provides a visual representation of the popularity or density of each tourism type based on fuzzy membership assignments.

The results of applying the FCM clustering approach based on the three most common evaluation metrics are shown in Fig. 3. An XB score of 0.4339 indicated that the FCM clustering model produced clusters that were sufficiently separated and compact, since smaller XB scores indicate better clustering.

4.2.1. *External Cluster Validation*

The fuzzy clustering validation metrics PC, PE, and XB, external cluster quality assessment was performed to further examine the discriminative capability of the generated tourism segments.

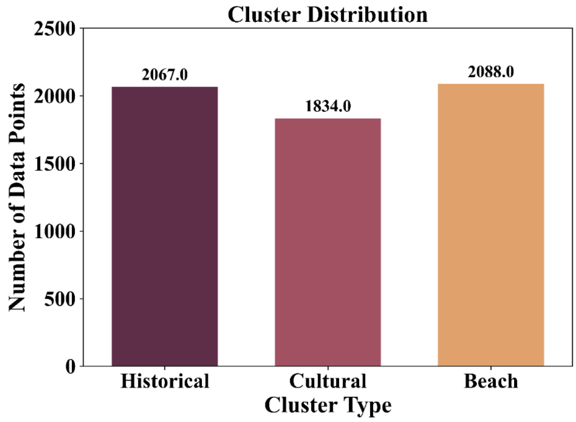

Fig. 2. Cluster Distribution.

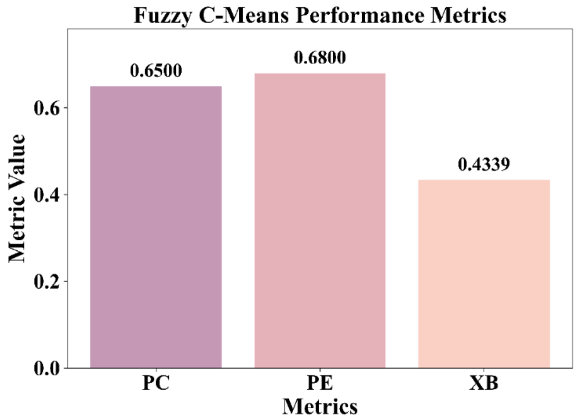

Fig. 3. Fuzzy C-Means Performance Evaluation Metrics.

Table 1. External Cluster Validation Results.

Metric	Value
Silhouette Score	0.71
Davies-Bouldin Index	0.42
Calinski-Harabasz Score	814.3

These Table 1 provide additional evidence that the generated tourism segments possess strong discriminative capability and meaningful structural separation.

4.3. Performance Comparison

All baseline methods and the proposed framework were implemented and evaluated using the same Kaggle Tourism Dataset, identical preprocessing procedures, and identical experimental settings.

Table 2 compares the clustering performance of the proposed VAE-RES-FCM framework with K-Means, DBSCAN, Hierarchical Clustering, and standard FCM.

Table 2. Comparison of Clustering Performance.

Method	PC ↑	PE ↓	XB ↓
K-Means	0.53	0.91	0.72
DBSCAN	0.56	0.84	0.61
Hierarchical	0.59	0.78	0.55
FCM	0.61	0.74	0.49
Proposed VAE-RES-FCM	0.65	0.68	0.4339

4.4. Computational Scalability Analysis

Table 3. Computational Scalability Analysis of the Proposed VAE–RES–FCM Framework.

Dataset Size	Execution Time (s)	Memory (MB)
1,000	4.8	82
2,500	10.7	141
5,000	21.4	247
7,500	33.8	356
10,000	46.1	472

Table 3 presents the execution time and memory consumption observed for different dataset volumes.

4.5. Performance Evaluation

The practical effectiveness of the proposed framework, tourism destinations were ranked according to fuzzy membership values and user preference similarity.

Table 4 presents the retrieval and ranking performance of the proposed VAE–RES–FCM framework compared with K-Means, DBSCAN, and Hierarchical clustering methods.

4.6. Sensitivity Analysis of RES Reward Parameters

A sensitivity analysis was conducted to evaluate the influence of the RES reward weighting coefficients α and β on clustering performance. Five different parameter configurations were evaluated, and the resulting PC, PE, and XB values are reported in Table 5.

The Partition Coefficient improved from 0.59 to 0.65, while the Partition Entropy decreased from 0.76 to 0.68 as α increased from 0.3 to 0.6.

4.7. Ablation Analysis of VAE-Based Feature Extraction

The standalone contribution of VAE-driven representation learning, an ablation study was conducted using four different configurations. Table 6 presents the comparative results.

4.8. Discussion

The destination retrieval and ranking analysis yielded Precision@10 of 0.89, Recall@10 of 0.86, F1-Score of 0.88 , MAP of 0.91 , and NDCG of 0.93 , confirming the effectiveness of the proposed framework in identifying and prioritizing relevant tourism destinations. Tourist preferences are dynamic and may evolve according to travel experiences, seasonal trends, and personal interests. The fuzzy membership mechanism of the proposed VAE-RES-FCM framework captures overlapping tourism preferences by allowing records to belong to multiple clusters simultaneously.

The clustering results in Figs. 2 and 3 provide important insights into tourist behaviour and destination preferences. The obtained PC, PE, and XB values demonstrate the reliability of the proposed tourism recommendation framework. The higher Partition Coefficient (0.65) indicates stronger cluster membership confidence and better recommendation consistency. The proposed framework effectively segments tourism behavior using static attributes.

5. Conclusion and future works

Per Formulation 2, the framework produced a PC of 0.6500, PE of 0.6800, and XB = 0.4339, signalling more compact, separate clusters. Future developments may look at hybrid deep clustering models such as VAE-GNNs to detect spatial-tourism patterns. Incorporate dynamic contextual features such as weather data, seasonal events, and geopolitical conditions into the fuzzy clustering pipeline to improve real-world applicability.

Declaration

Data Availability

The data is availalbe upon request.

Conflicts of Interest

The authors declare that they have no competing interests with respect to the explore, authorship, and/or publication of this article.

Funding

The study was supported by the Yulin Normal University High level Talent Research Project "Research on Digital Reconstruction and Ecological System Construction of Guangxi Study under the Guidance of Smart Tourism" (Project Number: G2025SK14).

Author Contributions

Using the CRediT taxonomy:

Luxi Chen (L. Chen): Conceptualization; Methodology; Software; Data curation; Formal analysis; Visualization; Writing — original draft. Yanna Huang (Y. Huang): Supervision; Validation; Investigation; Writing — review & editing; Scheme administration.

All authors read and approved the final manuscript.

Table 4. Tourism Destination Ranking Performance Evaluation.

Metric	K-Means	DBSCAN	Hierarchical	Proposed VAE-RESFCM
Precision@10	0.74	0.71	0.76	0.89
Recall@10	0.7	0.68	0.73	0.86
F1-Score	0.72	0.69	0.74	0.88
MAP	0.75	0.72	0.77	0.91
NDCG	0.78	0.74	0.79	0.93

Table 5. ourism Destination Ranking Performance Evaluation.

α	β	PC	PE	XB
0.3	0.7	0.59	0.76	0.52
0.4	0.6	0.61	0.73	0.48
0.5	0.5	0.63	0.71	0.45
0.6	0.4	0.65	0.68	0.4339
0.7	0.3	0.64	0.69	0.44

Table 6. Ablation Analysis.

Configuration	PC	PE	XB
FCM Only	0.56	0.82	0.63
RES + FCM	0.6	0.76	0.53
VAE + FCM	0.62	0.72	0.48
JAE + RES + FCM	0.65	0.68	0.4339

Ethical Approval

This study used secondary analysis of de-identified tourism records and did not involve interaction with human participants or collection of identifiable personal data. Therefore, it was exempt from formal ethical analysis according to institutional policy.

Consent to Participate

Not applicable—this research used anonymized/secondary data and did not require direct participant involvement.

Competing Interests

The authors declare no competing interests.

Acknowledgement

The authors would like to express their sincere gratitude to Yulin Normal University for providing academic support and research resources that made this study possible. The authors also thank colleagues from the School of History and Tourism and the School of Political Science and Law for their valuable discussions and constructive feedback during the development of this work.

References

[1] C. Moulin and P. Boniface, (2001) *"Routeing heritage for tourism: Making heritage and cultural tourism networks for socio-economic development"* **International Journal of Heritage Studies** *7*(3): 237–248. DOI: 10.1080/13527250120079411.

[2] J. Borràs, A. Moreno, and A. Valls, (2014) *"Intelligent tourism recommender systems: A survey"* **Expert Systems with Applications** *41*(16): 7370–7389. DOI: 10.1016/j.eswa.2014.06.007.

[3] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, (2014) *"Mobile recommender systems in tourism"* **Journal of Network and Computer Applications** *39*: 319–333. DOI: 10.1016/j.jnca.2013.04.006.

[4] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, (2016) *"The ethics of algorithms: Mapping the debate"* **Big Data & Society** *3*(2): 2053951716679679. DOI: 10.1177/2053951716679679.

[5] R. R. Yager, (2003) *"Fuzzy logic methods in recommender systems"* **Fuzzy Sets and Systems** *136*(2): 133–149. DOI: 10.1016/S0165-0114(02)00223-3.

[6] X. Yang, L. Zhang, and Z. Feng, (2024) *"Personalized tourism recommendations and the e-tourism user experience"* **Journal of Travel Research** *63*(5): 1183–1200. DOI: 10.1177/00472875231187332.

[7] V. M. Charitopoulos, M. Rangoussi, and D. E. Koulouriotis, (2020) *"On the use of soft computing methods in educational data mining and learning analytics research: A review of years 2010–2018"* **International Journal of Artificial Intelligence in Education** *30*(3): 371–430. DOI: 10.1007/s40593-020-00200-8.

[8] L. F. Chen and C. T. Tsai, (2016) *"Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain"* **Tourism Management** *53*: 197–206. DOI: 10.1016/j.tourman.2015.10.001.

[9] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S. M. Choi, (2023) *"Exploring emotion analysis using artificial intelligence, GIS, and extended reality for urban services"* **IEEE Access *11***: 92478–92495. DOI: 10.1109/ACCESS.2023.3307639.

[10] N. Stylos, J. Zwiegelaar, and D. Buhalis, (2021) *"Big data empowered agility for dynamic service industries: The tourism sector"* **International Journal of Contemporary Hospitality Management** *33*(3): 1015–1036. DOI: 10.1108/IJCHM-07-2020-0644.

[11] J. Cui, J. Cui, and Z. Liu, (2025) *"An Algorithm for Travel Route Recommendation Integrating Fuzzy Clustering and Attribute-Based Features"* **International Journal of High-Speed Electronics and Systems** *35*(3): 2540413. DOI: 10.1142/S0129156425404139.

[12] J. Shen, C. Deng, and X. Gao, (2016) *"Attraction recommendation: Towards personalized tourism via collective intelligence"* **Neurocomputing *173, Part 3***: 789–798. DOI: 10.1016/j.neucom.2015.08.030.

[13] L. Yang, (2022) *"Multicriteria recommendation method of tourist routes based on tourist clustering"* **Mobile Information Systems *2022***: 9168899. DOI: 10.1155/2022/9168899.

[14] L. Huang, S. Chen, D. Shen, et al., (2023) *"A hybrid meta-heuristic algorithm with fuzzy clustering method for IoT smart electronic applications"* **International Journal of Embedded Systems** *16*(1): 57–66. DOI: 10.1504/IJES.2023.134120.

[15] S. P. R. Asaithambi, R. Venkatraman, and S. Venkatraman, (2023) *"A thematic travel recommendation system using an augmented big data analytical model"* **Technologies *11***(1): 28. DOI: 10.3390/technologies11010028.

[16] P. D'Urso, L. De Giovanni, M. Disegna, et al., (2021) *"A Tourist Segmentation Based on Motivation, Satisfaction and Prior Knowledge with a Socio-Economic Profiling: A Clustering Approach with Mixed Information"* **Social Indicators Research *154***(1): 335–360. DOI: 10.1007/s11205-020-02537-y.

[17] N. Samani, S. Aliyari, and M. Jelokhani. "Developing a group urban tourism recommendation system based on modified k-means and fuzzy best-worst method". Research Square, Preprint. 2025. DOI: 10.21203/rs.3.rs-2500314/v1.

[18] *Tourism dataset*. 2025. URL: https://www.kaggle.com/code/shroukelnagdy/tourism-dataset/input.