# Competitiveness Evaluation of Tourist Attractions Based on Artificial Neural Network

Bingna Lou[1], Nan Chen[1], Lan Ma[2*]

[1] Management Department, Shijiazhuang University of Applied Technology, Shijiazhuang 050081, China
[2] Center of International Cooperation and Exchange, Hebei Normal University, Shijiazhuang 050024, China

Corresponding Author Email: malan@hebtu.edu.cn

## ABSTRACT

With the rapid development of tourism, tourists have become more aware of tourism level and quality. This triggers fierce competition between tourist attractions. To promote the core competitiveness of tourist attractions, this paper proposes a new evaluation model for the competitiveness of tourist attractions based on artificial neural network. First, a four-layer evaluation index system (EIS) was constructed for the competitiveness of tourist attractions, including detail elements, basic layer, core layer, and characterization layer. Next, all the evaluation indices were optimized through clustering by improved k-modes algorithm. Finally, a backpropagation neural network (BPNN) was established to evaluate the competitiveness of tourist attractions. Experimental results confirm the effectiveness of the proposed method. The research findings provide a reference for the application of artificial neural network (ANN) in other prediction fields.

## 1. INTRODUCTION

The consumption pattern of tourists has been changing constantly with the rapid development of tourism. Tourists have become more aware of tourism level and quality. This triggers fierce competition between tourist attractions. As a result, many scholars shift their attentions to the effective improvement of the core competitiveness of tourist attractions [1-4].

China is implementing supply-side reform of tourism, and upgrading the consumption pattern of tourists. In this context, the competitive advantages of tourist attractions are no longer limited to traditional elements like category, environment, and culture. Many new elements have emerged as the keys to improving the competitiveness of tourist attractions, including but not limited to dynamic management, marketing strategy, and tourist satisfaction. These elements reflect the degree of coordination between a tourist attraction, its managers, and tourists [5-9], which is positively correlated with the competitiveness of the attraction.

On the competitiveness of tourist attractions, the existing studies have covered all levels from country, province, city, attraction, enterprise, to market [10-12]. The main contents fall into three categories: the evaluation indices for competitiveness, the measurement and evaluation of competitiveness, and the improvement of competitiveness [13-15].

In terms of the evaluation indices for competitiveness, Salas-Olmedo et al. [16] constructed an EIS for the soft and hard competitiveness of provincial tourist attractions, which covers three aspects: the attractiveness of tourism resources, the support of the competitive environment, and the influence of the tourism resources market. Maeda et al. [17] conducted a systematic and in-depth study on the competitiveness of tourist attractions from eight aspects (economy, facilities, environment, talents, systems, services, performance, and

openness, and created a reasonable EIS for urban tourism competitiveness.

In terms of the measurement and evaluation of competitiveness, Francalanci and Hussain [18] analyzed the international tourism competitiveness of four countries (China, the United States, Switzerland, and Thailand), modeled the international tourism competitiveness in four dimensions ( market demand, tourism product supply, environmental carrying capacity, and related industry support), and measured the international competitiveness of the tourism industry in these countries, with the aid of data envelopment analysis (DEA). Prayag et al. [19] combined the technique for order of preference by similarity to ideal solution (TOPSIS) and analytic hierarchy process (AHP) to comprehensively evaluate the international tourism competitiveness in 40 countries and regions, selected 30 evaluation indices from the performance, resources, and support of tourist attractions, and explained the evaluation process.

In terms of the improvement of competitiveness, Fahmi et al. [20] explored the tourism competitiveness system of cities in the Yangtze River Delta, and proposed multi-faceted strategies for enhancing the competitiveness of regional tourism cooperation. Under the background of informatization, Ibanez-Ruiz et al. [21] set up an EIS for the competitiveness of tourist attractions from the humanistic approach; Covering sense of belonging, sense of familiarity, sense of dependence, and sense of identity, the established EIS highlights the absolute relationship between tourist perception and national tourism competitiveness.

Despite their sheer number, the existing studies on the competitiveness of tourist attractions have rarely discussed in-depth strategies or models for competitiveness improvement through new methods or from new angles; the evaluation indices are highly correlated, and the evaluation methods and forms are far from diverse. New research on the competitiveness of tourist attractions should focus on the

dynamic changes of the competitiveness, and innovate the evaluation method.

Artificial neural network (ANN) is a mature evaluation tools with excellence in data processing, self-adaptation, and self-learning. This paper puts forward a novel evaluation model for the competitiveness of tourist attractions based on the ANN. Firstly, a scientific EIS was established for the competitiveness of tourist attractions, which involves detail elements, basic layer, core layer, and characterization layer. Next, all the evaluation indices were optimized through clustering by improved k-modes algorithm. After that, an evaluation model for the competitiveness of tourist attractions was designed based on backpropagation neural network (BPNN). The effectiveness of our method was demonstrated through experiments.

## 2. EIS CONSTRUCTION

A scientific EIS is the basis for measuring and evaluating the competitiveness of tourist attractions. According to the provisions in the *Detailed Rules for Rating Service Quality and Environmental Quality* and the *Detailed Rules for Rating Landscape Quality*, the evaluation objective was determined based on the current situation of tourist attractions and innovation, and the traditional evaluation indices for the competitiveness of tourist attractions were merged, drawing on the classification criteria and methods for the quality of tourist attractions.

Referring to the relevant literature, a four-layer comprehensive, integral, scientific, and standard EIS was established for the competitiveness of tourist attractions. Involving detail elements, basic layer, core layer, and characterization layer, the EIS has four driving forces: potential, motivation, support, and attractiveness.

There are five primary indices in the EIS, including the experience of tour items, the experience of tourism services, the experience of tourism infrastructure, the experience of tourism management, and the experience of landscape atmosphere. The five primary indices are supported by 31 secondary indices.
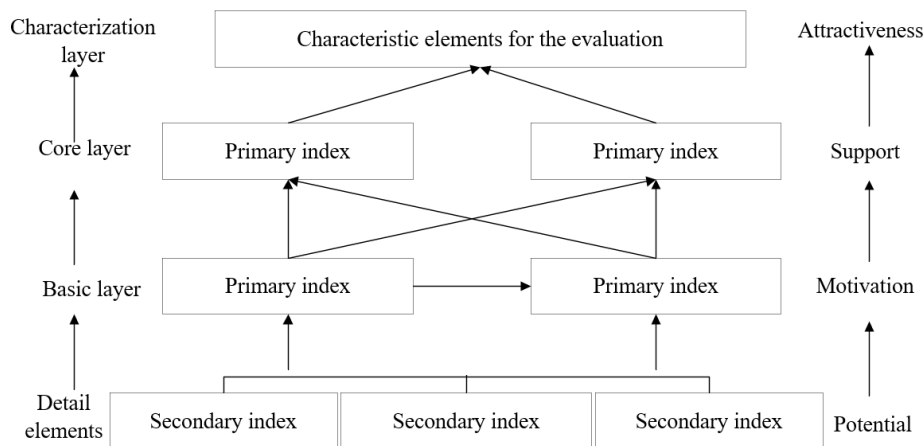


**Figure 1.** The EIS for the competitiveness of tourist attractions

**Table 1.** The correlation levels between the primary indices

| Interval | Type | Level | Interval | Type | Level |
|---|---|---|---|---|---|
| [0, 0.1) | | Extremely uncorrelated | [0.5, 0.6) | | Nearly correlated |
| [0.1, 0.2) | | Strongly uncorrelated | [0.6, 0.7) | | Weakly correlated |
| [0.2, 0.3) | Uncorrelated | Moderately uncorrelated | [0.7, 0.8) | Correlated | Moderately correlated |
| [0.3, 0.4) | | Weakly uncorrelated | [0.8, 0.9) | | Strongly correlated |
| [0.4, 0.5) | | Nearly uncorrelated | [0.9, 1] | | Extremely correlated |

As shown in Figure 1, the entire hierarchical EIS is as follows:

Layer 1 (goal): $C$={competitiveness of tourist attractions};

Layer 2 (primary indices): $C$={$C1$, $C2$, $C3$, $C4$, $C5$}={experience of tour items, experience of tourism services, experience of tourism infrastructure, experience of tourism management, experience of landscape atmosphere};

Layer 3 (secondary indices);

$C1$={$C11$, $C12$, $C13$, $C14$, $C15$, $C16$, $C17$, $C18$, $C19$}={rich entertainment activities, reasonable food and beverage prices, unique food and beverages, good quality goods, unique goods, reasonable goods prices, reasonable accommodation prices, healthy and clean food and beverages, clean and tidy accommodation};

$C2$={$C21$, $C22$, $C23$, $C24$, $C25$, $C26$, $C27$}={patient and friendly attitude of service staff, high professionality of service staff, timely and efficiency handling of complaints, satisfactory handling of complaints, comprehensive and responsible services of tour guide, reasonable design of the travel route};

$C3$={$C31$, $C32$, $C33$, $C34$, $C35$}={clear and definite signages, convenient transportation, comfortable rest areas, sufficient and clean toilets, available facilities for special groups};

$C4$={$C41$, $C42$, $C43$, $C44$, $C45$, $C46$}={diverse and interesting performances, reasonably scheduled performances, coordination between tour items and atmosphere, matching between publicity and actual situation, reasonable ticket prices, proper control of the number of tourists};

$C5$={$C51$, $C52$, $C53$, $C54$}={clear theme, beautiful natural scenery, good security, strong historical and cultural atmosphere}.

This paper mainly analyzes the correlations among the five primary indices by the following correlation function:

$$U = 5\sqrt{\frac{d_1 d_2 d_3 d_4 d_5}{d_1 + d_2 + d_3 + d_4 + d_5}} \quad (1)$$

where, $U \in [0, 1]$ is the value of the correlation function (if U=0, the correlations are the best, and the five primary indices are fully correlated and developing orderly; if U=1, the correlations are the worst, and the five primary indices are not correlated and developing disorderly; the closer the U value is to 0, the poorer the correlations; the closer the U value is to 1, the better the correlations); $d_1$-$d_5$ are the levels of experience of tour items, experience of tourism services, experience of tourism infrastructure, experience of tourism management, experience of landscape atmosphere, respectively:

$$d_1 = \sum_{i=1}^{9} \lambda_{1i} C_{1i}, d_2 = \sum_{i=1}^{7} \lambda_{2i} C_{2i}, d_3 = \sum_{i=1}^{5} \lambda_{3i} C_{3i},$$
$$d_4 = \sum_{i=1}^{6} \lambda_{4i} C_{4i}, d_5 = \sum_{i=1}^{4} \lambda_{5i} C_{5i} \quad (2)$$

where, $\lambda_{1i} \sim \lambda_{5i}$ are the weights of the five primary indices, respectively; $C_{1i} \sim C_{5i}$ are the dimensionless values of the secondary indices under the five primary indices, respectively. Formula (1) must be corrected in case that the primary indices, which are highly correlated, have low correlation values:

$$U^* = \sqrt{U(\alpha_1 d_1 + \alpha_2 d_2 + \alpha_3 d_3 + \alpha_3 d_3 + \alpha_3 d_3)} \quad (3)$$
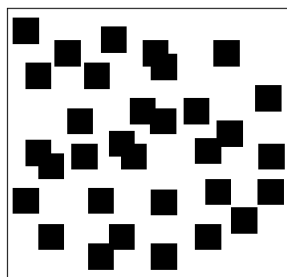
where, $U^*$ is the degree of coordinated development of the contents in five primary indices; $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0.2$ are the coefficients of the contents in five primary indices (the coefficients are equal because the contents in different primary indices are of equal importance). Table 1 divides the correlations between the primary indices.

However, the dynamic change of the object in time is not fully considered. To solve the problem, the original model was improved by adding the adjustment coefficient of time change:
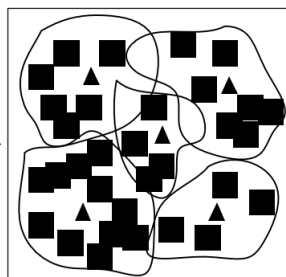
$$U_{kj}^* = U_{kj} \times (1 - \beta_k)^{j_{max} - j} \quad (4)$$

$$\Delta d_{iy} = \frac{(AV_{kj} - AV_{k(j-1)})}{AV_{k(j-1)}}, \beta_{k1} = \frac{\sum_{j=j_{min}}^{j_{max}} \sum_{i=1}^{N_k} \lambda_{ki} C_{ki}}{j_{max} - j_{min}} \quad (5)$$

where, $AV_{kj}$ is the mean of the k-th primary index of all tourist attractions in the 31 provincial administrative regions in China in year j; $\Delta d_{kj}$ is the growth rate of the k-th primary index in year j relative to the previous year; $\beta_k$ is the adjustment coefficient of the k-th primary index; $U_{kj}$ is the correlation value of the k-th primary index in year j; $U_{kj}^*$ is the new correlation value of the k-th primary index in year j after being adjusted by the adjustment coefficient.

Figure 2 below displays the new correlation values of C1 experience of tour items given by ten experts.
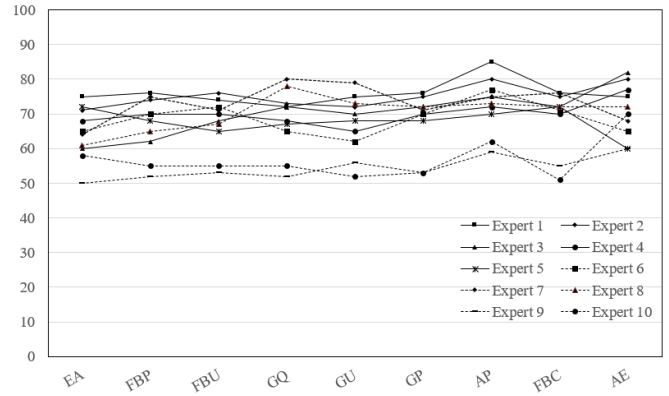


**Figure 2.** The new correlation values of C1 experience of tour items

Note: EA is Entertainment Activities, FBP is Food and Beverage Prices, FBU is Food and Beverages Uniqueness, GQ is Goods Quality, GU is Goods Uniqueness, GP is Goods Prices, AP is Accommodation Prices, FBC is Food and Beverages Cleanliness, AE is Accommodation Environment.

## 3. OPTIMIZATION THROUGH K-MODES CLUSTERING

Besides the accurate competitiveness evaluation of tourist attractions, this paper needs to fully consider the robustness and interpretability of the evaluation model. For this purpose, the traditional k-modes algorithm (Figure 3) was used to preliminarily cluster the data on the evaluation indices. The k-modes algorithm is suitable for local clustering and optimization of continuous numerical data. The objective function is the Euclidean distance from the data point to the prototype.

The traditional k-modes algorithm is easy to fall into local extreme points. Besides, the clustering classes often ignore the impact of the indices with low-frequency attributes.



**Figure 3.** The clustering process of traditional k-modes algorithm

This paper improves the k-modes clustering method to overcome the two defects. Firstly, the list of classification attributes of the evaluation indices was defined as $CP=(U, A, V, h)$, where $U=\{u_1, u_2, \ldots, u_n\}$ is the set of index samples; $A=\{a_1, a_2, \ldots, a_m\}$ is the set of index attributes; $V=\cup_{a\in A}V_a$ is the union of attribute intervals; $h$ is a function of $C\times A\to V$ (for any $a\in A$, and $u\in U$, there exists $h(u, \alpha)\in V_a$). In addition, it is assumed that $V_{aj}=\{a_j^{(1)}, a_j^{(2)}, \ldots, a_j^{(nj)}\}$, $a_j^{(l)}$ is the l-th optional attribute value of attribute $a_j$, and $n_j$ is the number of optional attributes.

Assuming that $W=[\omega_1, \omega_2, \ldots, \omega_k]^T$, the p-th clustering of an index can be expressed as:

$$\omega_p = [\omega_{p11}, \omega_{p12}, \ldots, \omega_{p1n_1}, \omega_{p21}, \omega_{p22}, \ldots, \\ \omega_{p2n_2}, \ldots, \omega_{pm1}, \omega_{pm2}, \ldots, \omega_{pmn_m}]^T \tag{6}$$

where, $k$ is the number of clusters; $1\le p\le k$; $1\le j\le m$.

The above formula represents the weight vector considering the attribute values of all indices. Under the above assumptions, the distance between index sample $u_i$ and cluster head c can be calculated by:

$$DIS(c_p, u_i) = \sum_{a\in A}\eta_a(c_p, u_i) \tag{7}$$

where,

$$\eta_a(c_p, u_i) = \begin{cases} 1, & h(c_p, a)\ne h(u_i, a) \\ 0, & h(c_p, a)=h(u_i, a) \end{cases}, 1\le p\le k, 1\le i\le n \tag{8}$$

The distance obtained by formula (7) characterizes the correlation difference between evaluation indices. Suppose $V^* = V_{a1}\times V_{a2}\times\ldots\times V_{a|A|}$. The density of index sample $u_i$ belonging to $V^*$ can be expressed as:

$$DEN(u_i) = \left(\frac{\left|\{v_i\in U\mid h(u_i, a)=h(v_i, a)\}\right|}{|U|} - 1\right) \tag{9}$$

The density obtained by formula (9) is the mean difference between an index sample and all the other index samples. The mode c of U that reflects the common feature of all indices can be calculated by:

$$DEN(c) = \max_{u_i\in V^*} Dens(u_i) \tag{10}$$

The distance from the cluster head c directly bears on the possibility of index sample $u_i$ being a boundary point. Thus, the k samples with the farthest distance from c and the highest density were taken as the class samples for iterative selection. The first class samples can be obtained by:

$$CS_1(u_i) = DEN(u_i) + DIS(u_i, c) \tag{11}$$

The other class samples can be obtained through iteration:

$$CS_{p+1}(u_i) = DEN(u_i) + \min_j^p DIS(v_i, c_j) \tag{12}$$

Let $u_i$ be an index sample of class $p$, corresponding to candidate cluster head $CC_{ui}=\{cc_1, cc_2, \ldots cc_{|A|}\}$, where $cc_i$ is a mode in the set (12):

$$M_i = \{v_i\in U\mid DIS(v_i, u_i)\le i\}, 1\le i\le |A| \tag{13}$$

The final cluster head was selected from $CC_{ui}$. The selection principle is to choose the index sample with the highest inter-class distance, smallest intra-class distance, and highest density from $v_i$ in $CC_{ui}$. The cluster head of the first class can be obtained by:

$$Center_1(v_i) = DEN(v_i) + DIS(v_i, c) - DIS(v_i, u_i) \tag{14}$$

The cluster heads of the other classes can be obtained through iteration:

$$center_{p+1}(v_i) = DEN(v_i) + \min_{j=1}^p DIS(v_i, c_j) \\ - DIS(v_i, u_i) \tag{15}$$

To consider the attribute values of all indices, the obtained cluster heads were transformed into the input weight cluster mode by the definition of mode W. If cluster head $u_{ij}=a_j^{(s)}$, the weight $\omega_{pjs}$ equals 1; otherwise, $\omega_{pjs}$ equals 0. Note that $1\le p\le k$, $1\le j\le m$, and $1\le s\le n_j$. Thereby, the characteristic clustering class $W_1$ was obtained for k classes. The objective function of the clustering can be defined as:

$$O_N(H, W) = \sum_{p=1}^k\sum_{i=1}^N\theta_{pi}DIS_N(w_p, u_i) \tag{16}$$

where, $H=[\theta_{pi}]$ is an $k\times N$-dimensional 0-1 binary matrix. If index sample $u_i$ belongs to class p, $\theta_{pi}$ equals 1; otherwise, $\theta_{pi}$ equals 0. The distance between index samples can be calculated by:

$$DIS_N(w_p, u_i) = \sum_{j=1}^m\rho_{aj}(w_p, u_i) \\ = \sum_{j=1}^m(1-\omega_{pjt})^2 + \sum_{s=1, s\ne t}^{n_j}\omega_{pjs}^2 \tag{17}$$

If $u_{pj}=a_j^{(t)}$, then $1\le t\le n_j$. To minimize the objective function, H and W were iteratively updated by:

$$\hat{\omega}_{pjt} = \frac{|n_{pjt}|}{|n_p|} \tag{18}$$

$$\hat{\theta}_{pi} = \begin{cases} 1, & DIS_N(\hat{\omega}_p, u_i)\le DIS_N(\hat{\omega}_g, u_i), 1\le g\le k \\ 0, & Otherwise \end{cases} \tag{19}$$

where, $|n_{pjt}|$ is the number of index samples in which the j-th attribute is of the value t in class p; $|n_p|$ is the total number of index samples in class p. Since the updates of H and W satisfy the optimal Karush–Kuhn–Tucker (KKT) conditions, the objective function $O_N$ can converge. Moreover, the cluster heads obtained through iteration fully represent the features of the k clustering classes for the evaluation indices, which effectively prevents the falling into local extreme points.

Hence, the optimal clustering results can be directly imported to the competitiveness evaluation model, making the evaluation more accurate. The clustering process of the improved algorithm is illustrated in Figure 4.
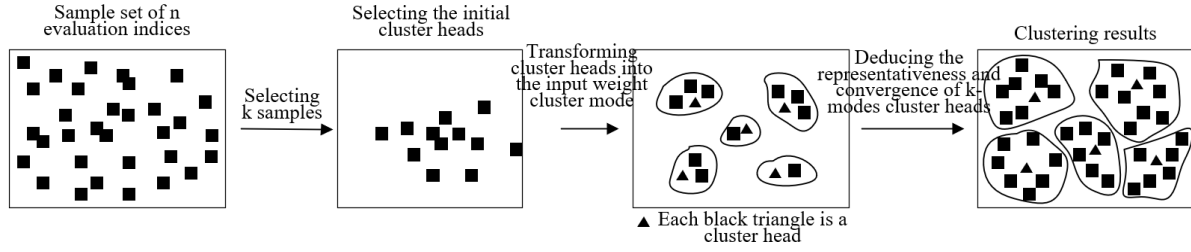


**Figure 4.** The clustering process of improved k-modes algorithm

## 4. MODEL CONSTRUCTION

The BPNN learns through fastest gradient descent, and adjusts the thresholds and weights of the hidden layer through error backpropagation, thereby minimizing the sum of squares for error (SSE) between actual output and the desired output. To evaluate the competitiveness of tourist attractions, this paper constructs a BPNN in the following steps:

Step 1. Assign values $\in$ [-1, 1] to the connection weights of each layer, configure the network structure, and determine the maximum number of learning and calculation accuracy.

Step 2. Randomly extract h samples from the n-dimensional input vector of evaluation indices, and express the desired output and characteristic input by:

$$\begin{cases} e(h) = \left(e_1(h), e_2(h), ..., e_m(h)\right) \\ x(h) = \left(x_1(h), x_2(h), ..., x_n(h)\right) \end{cases} \quad (20)$$

Step 3. Calculate the input $HI_p(H)$ and output $HO_p(H)$ of each hidden layer node backward layer by layer:

$$HI_p(h) = \sum_{i=1}^{n} \omega_{ip} x_i(h) - \varepsilon_p, \, p = 1, 2, ..., m \quad (21)$$

$$HO_p(h) = g\left(HI_p(h)\right), \, p = 1, 2, ..., m \quad (22)$$

Step 4. Based on the desired output and actual output of index samples, calculate the partial derivation of the error function relative to each output layer node:

$$\frac{\partial e}{\partial o_i} = \frac{\partial\left(\frac{1}{2}\sum_{i=1}^{m}\left(e_i(h) - o_i(h)\right)^2\right)}{\partial o_i} \quad (23)$$
$$= -\left(e_i(h) - o_i(h)\right)o_i(h)$$

By formula (23), compute the partial deviation of the error function relative to the connection weights between output layer and hidden layer:

$$\frac{\partial e}{\partial \omega_{op}} = \frac{\partial e}{\partial O_i} \cdot \frac{\partial O_i}{\partial \omega_{op}} = \frac{\partial e}{\partial O_i} \cdot \frac{\partial\left(\sum_{p=1}^{q}\omega_{op}HO_p(h) - \varepsilon_p\right)}{\partial \omega_{op}} \quad (24)$$
$$= \frac{\partial e}{\partial O_i} \cdot HO_p(h)$$

Step 5. Based on the partial derivatives obtained above and the output of each hidden layer node, compute the partial derivative of the error function relative to each hidden layer node:

$$\frac{\partial e}{\partial HI_p(h)} = \frac{\partial\left(\frac{1}{2}\sum_{i=1}^{m}\left(e_i(h) - o_i(h)\right)^2\right)}{\partial HO_p(h)} \cdot \frac{\partial HO_p(h)}{\partial HI_p(h)} \quad (25)$$
$$= -\left(\sum_{p=1}^{m}\frac{\partial e}{\partial o_i} \cdot \omega_p\right) \cdot \frac{\partial HO_p(h)}{\partial HI_p(h)}$$

By formula (25), compute the partial deviation of the error function relative to the connection weights between hidden layer and input layer:

$$\frac{\partial e}{\partial \omega_{ip}} = \frac{\partial e}{\partial HI_p(h)} \cdot \frac{\partial HI_p(h)}{\partial \omega_{ip}} \quad (26)$$
$$= \frac{\partial e}{\partial HI_p(h)} \cdot \frac{\partial\left(\sum_{i=1}^{n}\omega_{ip}x_i(h) - \varepsilon_p\right)}{\partial \omega_{ip}} = \frac{\partial e}{\partial HI_p(h)} \cdot x_i(h)$$

Step 6. Based on the connection weights between hidden layer and output layer and the partial deviation of the error function relative to each output layer node, adjust and update the weights through backpropagation of gradient, using the fastest gradient descent method:

$$\begin{cases} \Delta\omega_{op}(h) = -\eta\frac{\partial e}{\partial \omega_{op}} = -\eta\frac{\partial e}{\partial o_i} \cdot HO_p(h) \\ \omega_{op}^{t+1} = \omega_{op}^{t} + \tau\frac{\partial e}{\partial o_i} \cdot HO_p(h) \end{cases} \quad (27)$$

Step 7. Based on the connection weights between hidden layer and input layer and the partial deviation of the error function relative to these connection weights, adjust and update the weights through backpropagation of gradient, using the fastest gradient descent method:

$$\begin{cases} \Delta\omega_{ip}(h) = -\eta\frac{\partial e}{\partial \omega_{ip}} = -\eta\frac{\partial e}{\partial HI_p(h)} \cdot x_i(k) \\ \omega_{ip}^{t+1} = \omega_{ip}^{t} + \tau\frac{\partial e}{\partial HI_p(h)} \cdot x_i(k) \end{cases} \quad (28)$$

Step 8. Calculate the global error:

$$E = \frac{1}{2m} \sum_{k=1}^{m} \sum_{i=1}^{q} \left(d_i(k) - y_{oi}(k)\right)^2 \qquad (29)$$

If the number of learning times reaches the preset maximum number or the global error meets the preset evaluation accuracy, terminate the iterative process; otherwise, return to Step 3 and execute the following steps until the network converges.

## 5. EXPERIMENTS AND RESULT ANALYSIS

The contrastive experiments were conducted on a computer with Intel® Xeon® Processor E5620 (2.40 GHz) and 32G memory. The proposed algorithm was compared with k-nearest neighbors (KNN), traditional k-modes, k-means clustering (KMC), and fuzzy c-means (FCM)-KNN, in the Java language environment.

Table 2 compares the performance of KNN, traditional k-modes, and our algorithm in competitiveness evaluation of tourist attractions. It can be seen that our algorithm greatly outshined the other two methods in accuracy and recall.

The clustering results of all indices for the competitiveness of tourist attractions were summarized. Figure 5 presents the clustering results of primary index C2. Obviously, C21 belongs to class 1, C27 belongs to class 11, and C22-26 belong to class 5.

The proposed clustering algorithm was adopted to preliminarily cluster the evaluation indices. Figure 6 shows the relationship between minimum SSE and the number of classes K. It can be seen that the inflection point of the curve appeared at K=2 and 3; the number of index samples was correlated with the K value. The K value changes with the scale and type of the evaluation problem, and needs to be configured careful in actual application.

**Table 2.** The comparison between three clustering algorithms in competitiveness evaluation of tourist attractions

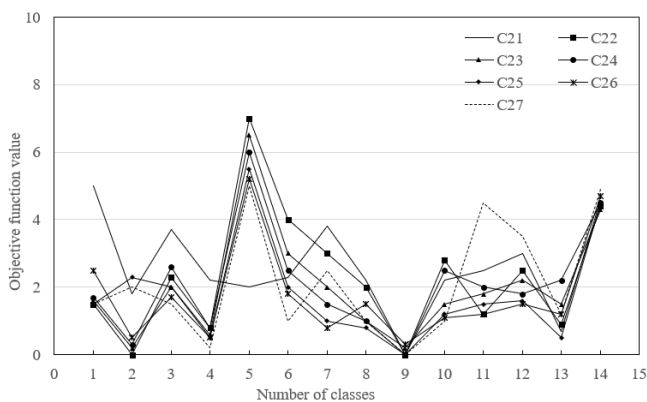| Algorithm | Cluster heads | Number of samples | Number of result samples | Minimum SSE | Accuracy | Recall |
|---|---|---|---|---|---|---|
| KNN | (3,4,5,6) | 1,546 | 775 | 1,564 | 0.9201 | 0.9213 |
|  | (7,6,4,3) |  | 656 |  |  |  |
|  | (4,6,5,2) |  | 583 |  |  |  |
| Traditional k-modes | (6,4,7,5) | 1,546 | 526 | 947 | 0.9379 | 0.9476 |
|  | (3,5,2,5) |  | 673 |  |  |  |
|  | (4,5,7,3) |  | 711 |  |  |  |
| Our algorithm | (4,5,6,3) | 1,546 | 803 | 703 | 0.9728 | 0.9841 |
|  | (4,5,3,2) |  | 523 |  |  |  |
|  | (7,3,6,4) |  | 476 |  |  |  |



**Figure 5.** The clustering results of primary index $C_2$
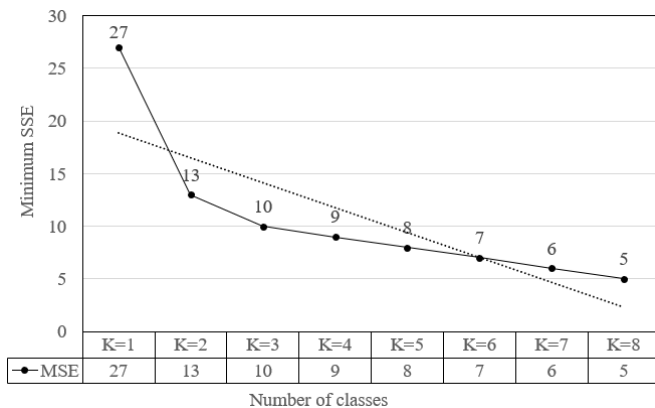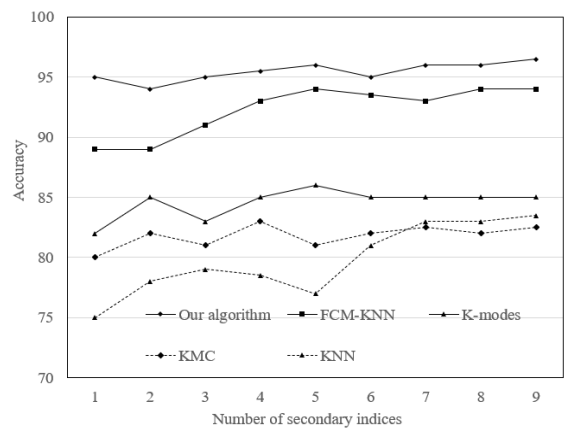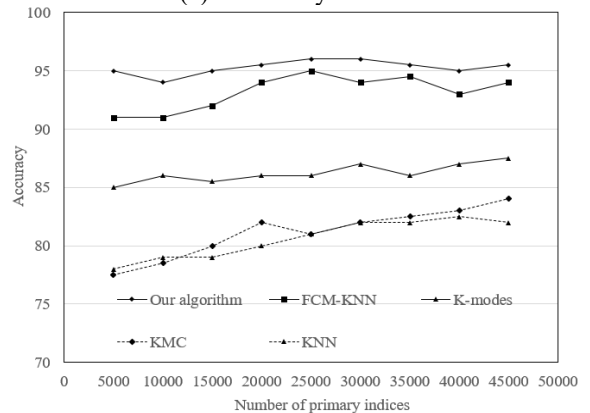


**Figure 6.** The influence of number of classes on evaluation error



(a) Secondary indices



(b) Primary indices

**Figure 7.** The influence of algorithm and number of index samples on accuracy

To verify the effect of different datasets with equal size on clustering accuracy, 31,224 test index samples were evenly divided into six groups for experimental analysis. Figure 7(a) presents the experimental results. It can be seen that KNN and traditional k-modes were low in accuracy, while our algorithm had a clear edge over the other four algorithms in evaluation accuracy. The main reasons are as follows:

The traditional K-modes computes the similarity between sample attributes based on distance to the boundary. The evaluation accuracy is greatly affected by the uncertain number of sample attributes. The KNN often mistakenly allocates lots of samples to the class of the largest group of index samples. The FCM-KNN generates the initial clusters randomly, which reduces the accuracy of the final classes. By contrast, the proposed algorithm can mitigate the negative effect of random initial clusters. In our algorithm, the initial cluster heads are selected based on both distance and density, preventing the falling into local extreme points. Through thorough consideration of the attribute values of all samples, the distance measure is effectively optimized. In this way, our method can adapt well to index data with class attributes, and realize high classification accuracy.

To verify the influence of sample number on clustering algorithms, eight groups of different number of primary index samples were adopted to compare the accuracies of different algorithms. As shown in Figure 7(b), our algorithm outperformed the other algorithms, regardless of the change in the number of index samples.

The next task is to verify the feasibility and effectiveness of the constructed BPNN. Under the initial training cycles of 3,000 and step length of 600, the output did not change significantly after 12,000 cycles. Hence, 12,000 was selected as the training cycles of the network.

The training period has a great impact on the evaluation accuracy and runtime of the neural network. If the period is too short, the network will have a poor data fitting effect; if the period is too long, the network could suffer from overfitting, i.e. the accuracy on test set is poorer than that on training set. Hence, this paper trains the network in 1,000 cycles, and observes the variation of evaluation accuracy with training cycles. As shown in Figure 8, the optimal number of training cycles for the BPNN is 200.
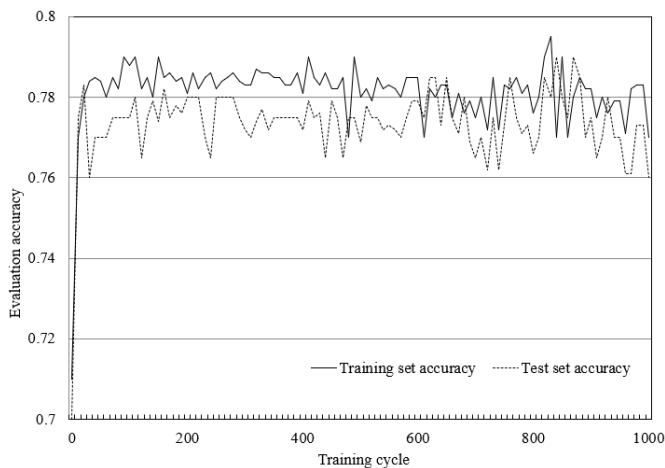


**Figure 8.** The relationship between training cycle and evaluation accuracy

Figure 9 shows the distributions of the absolute error between the actual value and the corresponding output. It can

be seen that the variance of the actual value was greater than that of the output, indicating that the fluctuation between the actual value and the mean value was larger than that between the output and the mean value. The main reason is the gap between the number of different primary indices, adding to the difficulty of network training and learning.
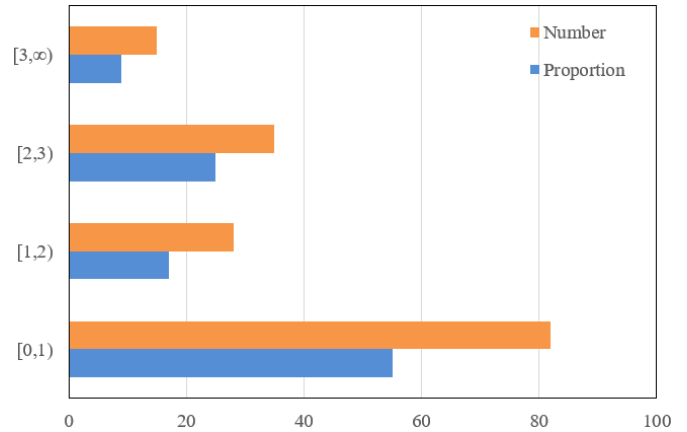


**Figure 9.** The distributions of the absolute error between the actual value and the corresponding output

Figure 10 shows the relationship between the number of classes and the actual results of competitiveness evaluation. It can be seen that the actual values basically coincided with the desired outputs. This fully demonstrates the effectiveness of our BPNN.
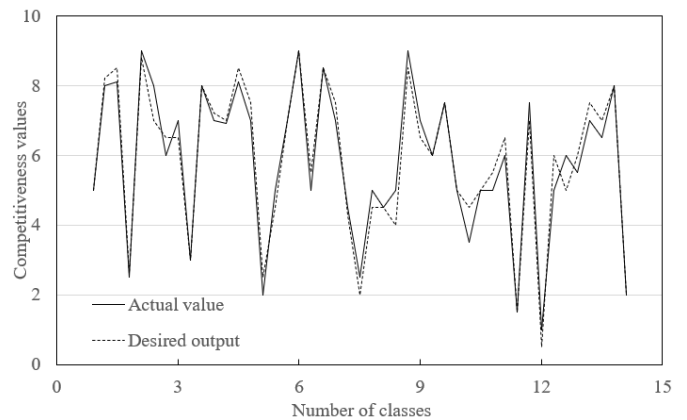


**Figure 10.** The number of classes and the evaluation values

## 6. CONCLUSIONS

This paper introduces the BPNN to the competitiveness evaluation of tourist attractions, and puts forward a competitiveness evaluation model for tourist attractions. Firstly, a scientific EIS was designed for competitiveness evaluation of tourist attractions. The proposed EIS covers four layers, namely, detail elements, basic layer, core layer, and characterization layer, and has four driving forces: potential, motivation, support, and attractiveness. Then, all evaluation indices were optimized through the clustering by improved k-modes algorithm. Experimental results show that the proposed algorithm selects initial cluster heads based on both distance and density, preventing the falling into local extreme points; through thorough consideration of the attribute values of all samples, the distance measure is effectively optimized. Finally,

the authors designed a BPNN for the competitiveness evaluation of tourist attractions. The experimental results show that the actual values basically coincided with the desired outputs, which fully demonstrates the effectiveness of our BPNN.

## REFERENCES

[1] Zhang, X., Luo, H., Chen, B., Guo, G. (2020). Multi-view visual Bayesian personalized ranking for restaurant recommendation. Applied intelligence, 50: 2901-2915. https://doi.org/10.1007/s10489-020-01703-6

[2] Liu, Y., Pan, S.L. (2015). Personalized travel recommendation technology based on friendship of LBSN. Computer Engineering and Applications, 51(8): 117-122.

[3] Park, D.H. (2017). The development of travel demand nowcasting model based on travelers' attention: focusing on web search traffic information. The Journal of Information Systems, 26(3): 171-185. https://doi.org/10.5859/KAIS.2017.26.3.171

[4] Maghrebi, M., Abbasi, A., Rashidi, T.H., Waller, S.T. (2015). Complementing travel diary surveys with twitter data: application of text mining techniques on activity location, type and time. 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas, pp. 208-213. https://doi.org/10.1109/ITSC.2015.43

[5] Abbasi, A., Rashidi, T.H., Maghrebi, M., Waller, S.T. (2015). Utilising location based social media in travel survey methods: Bringing Twitter data into the play. LBSN'15: Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks, NY, US, pp. 1-9. https://doi.org/10.1145/2830657.2830660

[6] Pesonen, J., Lampi, M. (2016). Utilizing open data in tourism. ENTER 2016 Conference on Information and Communication Technologies in Tourism, Bilbao, Spain.

[7] Lim, K.H., Chan, J., Leckie, C., Karunasekera, S. (2018). Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. Knowledge and Information Systems, 54(2): 375-406. https://doi.org/10.1007/s10115-017-1056-y

[8] Chong, W.H., Dai, B.T., Lim, E.P. (2015). Not all trips are equal: Analyzing foursquare check-ins of trips and city visitors. In Proceedings of the 2015 ACM on Conference on Online Social Networks, NY, US, pp. 173-184. https://doi.org/10.1145/2817946.2817958

[9] Padilla, J.J., Kavak, H., Lynch, C.J., Gore, R.J., Diallo, S.Y. (2018). Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter. PloS One, 13(6): e0198857. https://doi.org/10.1371/journal.pone.0198857

[10] Mariani, M., Baggio, R., Fuchs, M., Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: A systematic literature review. International Journal of Contemporary Hospitality Management,

30(12): 3514-3554. https://doi.org/10.1108/IJCHM-07-2017-0461

[11] e Silva, F.B., Herrera, M.A.M., Rosina, K., Barranco, R. R., Freire, S., Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. Tourism Management, 68: 101-115. https://doi.org/10.1016/j.tourman.2018.02.020

[12] Dwyer, L., Tomljenović, R., Čorak, S. (2017). Evolution of Destination Planning and Strategy. The Rise of Tourism in Croatia, Springer.

[13] Calheiros, A.C., Moro, S., Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. Journal of Hospitality Marketing & Management, 26(7): 675-693. https://doi.org/10.1080/19368623.2017.1310075

[14] Bunghez, C.L. (2016). The importance of tourism to a destination's economy. Journal of Eastern Europe Research in Business & Economics, 143495. https://doi.org/10.5171/2016.143495

[15] Sabou, M., Onder, I., Brasoveanu, A.M., & Scharl, A. (2016). Towards cross-domain data analytics in tourism: a linked data based approach. Information Technology & Tourism, 16(1): 71-101. https://doi.org/10.1007/s40558-015-0049-5

[16] Salas-Olmedo, M.H., Moya-Gómez, B., García-Palomares, J.C., Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing Big Data sources. Tourism Management, 66: 13-25. https://doi.org/10.1016/j.tourman.2017.11.001

[17] Maeda, T.N., Yoshida, M., Toriumi, F., Ohashi, H. (2018). Extraction of tourist destinations and comparative analysis of preferences between foreign tourists and domestic tourists on the basis of geotagged social media data. ISPRS International Journal of Geo-Information, 7(3): 99. https://doi.org/10.3390/ijgi7030099

[18] Francalanci, C., Hussain, A. (2016). Discovering social influencers with network visualization: evidence from the tourism domain. Information Technology & Tourism, 16(1): 103-125. https://doi.org/10.1007/s40558-015-0030-3

[19] Prayag, G., Hosany, S., Muskat, B., Del Chiappa, G. (2017). Understanding the relationships between tourists' emotional experiences, perceived overall image, satisfaction, and intention to recommend. Journal of Travel Research, 56(1): 41-54. https://doi.org/10.1177/0047287515620567

[20] Fahmi, U., Ginting, N., Sitorus, R. (2018). Local communities and tourists' perception towards to PLTD Apung sites as tsunami disaster tourism in Banda Aceh City. IOP Conference Series: Earth and Environmental Science, 126(1): 012177. https://doi.org/10.1088/1755-1315/126/1/012177

[21] Ibanez-Ruiz, J., Bustamante-Martinez, A., Sebastia, L., Onaindia, E. (2018). LinkedDBTour: A tool to retrieve linked open data about tourism attractions. 31st IBIMA Conference, Milan, Italy.