Problem Chosen	2021 Shu Wei Cup Summary Sheet	Team Control Number
C		202111046833

Future development plans for tournism under the Corona-virus disease

Summary

The epidemic of new coronary pneumonia has caused varying degrees of impact on all walks of life. Have a greater negative impact on the tourism industry. This article uses the establishment of genetic algorithm and analytic hierarchy process to optimize the evaluation and analysis of the future development plan of the tourism industry, which has certain research significance.

Regarding question 1, this article obtained the geographic location and reception capacity of the main A-level scenic spots and B-level scenic spots in several provinces in my country by consulting the data. Through multiple regression analysis, each geographic location is fitted and analyzed to determine the difference between each location.

For the question 2, this article selects A-level scenic spots to obtain their corresponding attractions and reception capabilities by consulting data. The reception ability can be replaced by the maximum daily reception ability, and then the clustering algorithm is established for this scenic spot to evaluate it. And classification.

For question 3, this article uses the Forbidden City in Beijing as an example to determine the main factors of current limitation, and then determine the objective function that can occur epidemic infection, using a 0-1 programming model, using MATLAB software to compile programs, and use genetic algorithms for optimization analysis. So as to determine the most reasonable current limiting scheme.

For question 4, this paper obtains the membership matrix of epidemic prevention and control, tourism income, and tourist experience by consulting data, and then uses analytic hierarchy to determine the corresponding main factors, and then establishes the objective function of the epidemic infection rate, and uses the neural network prediction model to proceed. Analyze and simulate the expected effect if controlled.

Regarding question 5, this article draws up a differentiated management plan for different scenic spots during the epidemic based on the results of questions 1 to 4 above.

Finally, the model used in this article is evaluated and analyzed to make it more scientific and reliable and in line with actual needs.

Keywords: new crown epidemic; analytic hierarchy process; genetic algorithm; suggestion

Content

1 Raising and restating the question

In the context of COVID-19, some countries have also taken the initiative to explore new tourism models, but it is believed that a general consensus has been reached. For a country like China with a large population and many holidays, how to weigh the interests of all parties in the context of the epidemic, and then propose a scientific tourism development plan, seems very meaningful. Attachment 1 gives the names and latitude and longitude information of some scenic spots in Beijing. Please solve the following five problems in combination with related theories of mathematical modeling:

Question 1: Please refer to the data in Annex 1. Please collect basic data on scenic spots in all provinces across the country. In particular, it is necessary to collect the current evaluation level and daily maximum reception capacity of scenic spots, and visually display the distribution characteristics of these scenic spots;

Question 2: Can the collected attractions and reception capabilities of these scenic spots be used for quantitative evaluation and reasonable classification?

Question 3: Under the background of the epidemic, current restriction in various scenic spots has become a new model. Can you provide a reasonable current limit model for scenic spots according to different popularity levels, and give specific quantitative analysis results using a scenic spot as an example;

Question 4: Can the mathematical modeling method be used to propose a quantitative model that considers epidemic prevention and control, tourism income, and tourist experience at the same time, and simulate and analyze the expected effects?

Question 5: Provide government management departments with a differentiated management plan for different scenic spots during the epidemic period of no more than 2 pages.

2 Problem analysis

For question one, This article obtains the geographic location and reception capacity of major A-level scenic spots and B-level scenic spots in several provinces in my country by consulting data. The data is classified and summarized to obtain the geographic locations of A-level scenic spots and B-level scenic spots, and then a fitting analysis is performed on each geographic location through multiple regression analysis to determine the difference between each location.

For the second question, This paper selects A-level scenic spots to obtain their corresponding attractions and reception capabilities by consulting data. The reception capabilities can be replaced by the maximum daily reception capabilities, and then clustering algorithms are established for the scenic spots to evaluate and classify them by establishing a clustering algorithm.

For question three, this article uses the Forbidden City in Beijing as an example to determine the main factors of current limit , And then determine the objective function that can occur epidemic infection, adopt 0-1 planning model, use MATLAB software to compile the program, and use genetic algorithm for optimization analysis, so as to determine the most reasonable current limiting plan.

For question four, This paper obtains the membership matrix of epidemic prevention and control, tourism income, and tourist experience by consulting data, and then uses hierarchical analysis to determine the corresponding main factors, and then establishes the objective function of the epidemic infection rate, uses neural network prediction model to analyze, and simulates If you control the expected effect.

Regarding question 5, this article draws up a differentiated management plan for different scenic spots during the epidemic based on the results of questions 1 to 4 above.

3 Basic assumptions

(1) The statistical data is accurate.
(2) In the statistical analysis of the data, the most reasonable method is adopted as much as possible, and the error is the smallest.
(3) In addition to the indicators given in this article, other factors have little influence and can be ignored.

4 Symbol variable description

This article makes the following explanations for some important symbols:

Table 1 Symbols and meanings

Symbol	Meaning
x_1、x_2、…、x_n	Impact index
$B(k)$	A pairwise comparison matrix of the k criterion of the scheme level to the criterion level
B_{ij}^k	The ratio of the influence of options P_i and P_j on the superiority of criterion C_k
ω	Weight vector
λ	Largest characteristic root
CI	Consistency index
CR	Consistency rate

RI	Random consensus index

5 Model establishment and solution

5.1 Problem one solution

For question one, it is required to collect basic data on scenic spots in all provinces across the country. In particular, it is necessary to collect the current evaluation level and daily maximum reception capacity of scenic spots, and visually display the distribution characteristics of these scenic spots. This article obtains the geographic location and reception capacity of the main A-level scenic spots and B-level scenic spots in several provinces in my country by consulting data. The data is classified and summarized to obtain the geographic locations of A-level scenic spots and B-level scenic spots, and then a multiple regression analysis is performed on each geographic location to determine the difference between each location.

The Lagrangian model refers to a multi-element geographic environment system where multiple (more than two) elements also have mutual influence and correlation. Therefore, the multiple geographic regression model is more universal.

Establishment of Lagrangian model

Definition: Suppose the function y = f (x) is defined on the interval [a, b], and the function value $y_i = f(x_i)(i=0,1,2,\ldots,n)$ on a series of points is given, and the n-degree polynomial $L_n(x)$ is

$$L_n(x_i) = y_i\ (i = 0,1,2,\ldots,n)$$

The constructed polynomial of degree n can be expressed as:

$$L_n(x_i) = \sum_{k=0}^{n} l_k(x) y_k = \sum_{k=0}^{n} \left(\prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j} \right) y_k$$

The fitting value of this set of data is $\hat{Y} = X\hat{\beta}$, and fitting error $e = Y - \hat{Y}$ is called residual, which can be used as an estimate of random error ε ,then:

$$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Is the residual sum of squares (or residual sum of squares), namely $Q(\hat{\beta})$.

(3)Statistical Analysis

The following results are given without proof:

$\hat{\beta}$ is the linear unbiased minimum variance estimate of β. It means that $\hat{\beta}$ is a linear function of β; the expectation of $\hat{\beta}$ is equal to β; in the linear unbiased estimate of β, the variance of $\hat{\beta}$ is the smallest.

$\hat{\beta}$ Obey normal distribution

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$$

(iii) For the residual sum of squares Q, and

$$EQ = (n - m - 1)\sigma^2$$

$$\frac{Q}{\sigma^2} \sim \chi^2(n - m - 1)$$

Unbiased σ^2 estimate

$$s^2 = \frac{Q}{n - m - 1} = \hat{\sigma}^2$$

s^2 is the residual variance (variance of the residual), and s is called the residual standard deviation.

(Iv) Decompose the sample variance $S = \sum_{i=1}^{n}(y_i - \bar{y})^2$ of Y, there are

$$S = Q + U, \quad F = \frac{U / m}{Q/(n - m - 1)} \sim F(m, n - m - 1)$$

Among them, Q residual sum of squares reflects the influence of random error on y, and U is called regression sum of squares, reflecting the influence of independent variables on y.

(a) Least square estimation of parameter vector β

Definition 1: In the Lagrangian model,the least square estimator β of parameter $b = (b_0, b_1, b_2, \ldots, b_m)$ should minimize the sum of squares of error $Q(b)$.

$$Q(b) = \min_{ALL\beta} Q(B) \tag{1}$$

Among them:

$$Q(\beta) = \sum_{t=1}^{n} \varepsilon_t^2 = \sum_{t=1}^{n} \left[yt - (\beta_0 + \beta_1 x_{t1} + \ldots + \beta_m x_{tm})\right]^2 = (Y - C\beta)'(Y - C\beta) \tag{2}$$

Theorem 1: Let $rank(c) = m + 1 \le n$,Then, $\hat{\beta} = b = (C'C)^{-1} C'Y$ is the least square estimation of the parameters.

$$C'C\beta = C'Y \tag{3}$$

Then,We often call the above equations a normal equation.
Using the principle of maximum likelihood ratio, the maximum likelihood estimates of parameter σ^2 is easily obtained as:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[y_i - (b_0 + b_1 x_{i1} + \ldots + b_m x_{im})\right]^2 \\ &= \frac{1}{n}(Y - Cb)'(Y - Cb) \\ &= \frac{1}{n}Q(b)\end{aligned} \tag{4}$$

But because $\hat{\sigma}^2$ is not an unbiased estimator of σ^2, we usually take

$$s^2 = \frac{1}{n-m-1}Q(b) \tag{5}$$

As an estimator of σ^2, it is an unbiased estimator of σ^2.

(2) Estimation of parameter function $\alpha'\beta$

In Lagrangian model analysis, finding the least square estimator b of parameter β is not our ultimate goal. Our purpose is to estimate Y, and Y is a linear function of $x_1, x_2, \ldots, x_m$,then:

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m = (1, x_1, \cdots x_m)\beta = \alpha'\beta \tag{6}$$

That is, the linear function β of the parameter vector $\alpha'\beta$ is estimated.
Significance test of Lagrangian model.Like the univariate linear regression model, when the Lagrangian model is established, it also needs to be tested for significance. The fluctuation or difference between the observed value $y_1, y_2, \ldots, y_n$ of the dependent variable y is caused by two factors, one is due to the difference of the independent variable $x_1, x_2, \ldots, x_k$, and the other is caused by the influence of other random factors.In order to distinguish them from the sum of squared deviations of y, it is necessary to perform analysis of variance on the Lagrangian model, that is, to decompose the sum of squared deviations of y in S_T into two parts, namely the regression sum of squares U and the remainder The sum of squares Q, so we replace

s^2 with σ^2, when $H_0^{(J)}$ is true, we have

$$t_j = \frac{\hat{\beta}_j / \sqrt{c_{ij}}}{\sqrt{Q/(n-m-1)}} \tag{7}$$

For the given α, if $|t_j| < t_{\frac{\alpha}{2}}(n-m-1)$, accept $H_0^{(J)}$, otherwise reject.

Formula (6) can also be used to estimate the interval of β_j. At the confidence interval level of $1-\alpha$, the confidence interval of β_j is:

$$\left[\hat{\beta}_j - t_{\frac{\alpha}{2}}(n-m-1)s\sqrt{c_{ij}}, \hat{\beta}_j + t_{\frac{\alpha}{2}}(n-m-1)s\sqrt{c_{ij}}\right] \tag{8}$$

Among them：

$$s = \sqrt{\frac{Q}{n-m-1}}.$$

Then this paper obtains the specific geographic location and daily maximum passenger flow of the A-level and B-level scenic spots in my country's main provinces by consulting the data, and then obtains the geographic locations of the A scenic spot and the B scenic spot through the data competition, and then uses the multiple regression model to fit the latitude and longitude. Analysis, you can get the result graph as shown below:

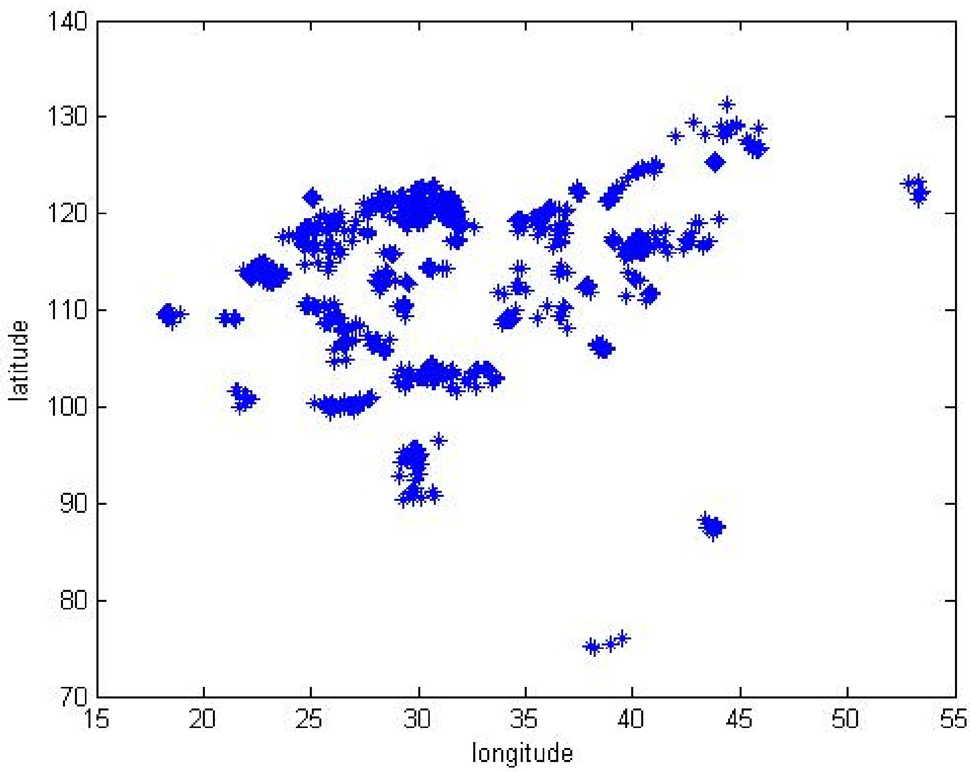


Figure 1 Distribution characteristics of scenic area A

From the above figure, we can see that Scenic Area A is mainly distributed in a relatively concentrated longitude range of 20 to 45, and a latitude range of 90 to 130.

In the same way, the result map of the main distribution characteristics of B-level scenic spots can be obtained, as shown below:

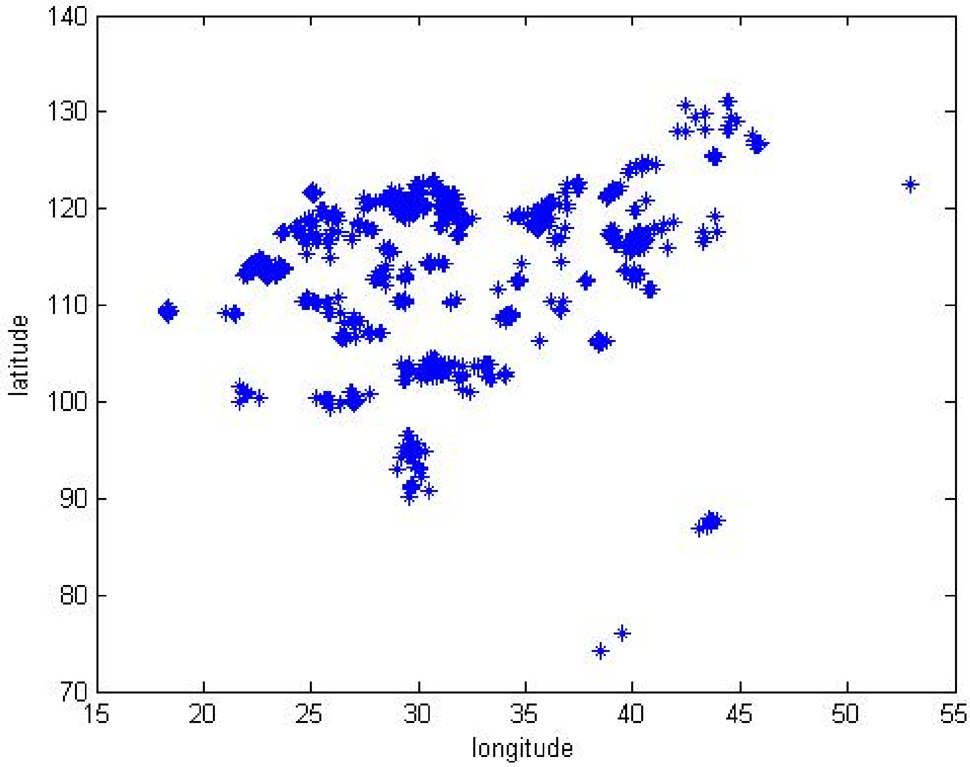


Figure 2 Distribution characteristics of B-level scenic spots

From the above figure, it can be found that the category B scenic spots are mainly distributed in the longitude range of 20 to 45, and the latitude range is in the range of 90 to 130. Through comparison, it can be found that the relative distribution of category A and category B. Basically, category A has category B. The two and regional differences have a greater impact.

5.2 Solving Problem 2

For the second question, it is required to use the collected attractions and reception capabilities of these scenic spots to quantitatively evaluate and reasonably classify them. This paper selects A-level scenic spots to obtain their corresponding attractions

and reception capabilities by consulting data. The reception capabilities can be replaced by the maximum daily reception capabilities, and then clustering algorithms are established to evaluate and classify the scenic spots by establishing a clustering algorithm.The K-means clustering algorithm proposed by MacQueen is also known as the K-means algorithm. It is mainly used for the classification of data objects. Clustering is to divide data with similar characteristics into the same set according to certain rules. In the collection. This algorithm is an iterative algorithm. It first calculates the distance between the sample object and the center point of each cluster, and then divides each sample into the cluster domain closest to itself.
The definition of the clustering problem is: a data matrix is composed of p-element observation data of n samples:

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{9}$$

Among them, each row represents a sample, each column represents an attribute, and x_{ij} represents the observed value of the j-th attribute of the i-th sample.
Set up the P element observation data of n samples

$$x_i=(x_{i1},x_{i2},\cdots,x_{ip})^T \qquad (i=1,2,\cdots,n)$$

At this time, each sample can be considered as a point in the P-element space, that is, a P-dimensional vector. The distance between the two vectors is $d(x_i,x_j)$. The most commonly used in cluster analysis is the Euclidean distance, that is

$$d(x_i,x_j)=\left[\sum_{k=1}^{p}(x_{ik}-x_{jk})2\right]^{1/2}$$

Before using K-means clustering, the classification number k should be determined according to the actual case. The selective samples in each class are called cluster points, which are generally selected according to the principle of minimum and maximum.
For two data objects in a two-dimensional sample space, the similarity measurement methods between any two data objects X_i and X_j generally include: Euclidean distance, Minkowski distance and Mahalanobis distance.
The flow of the clustering algorithm is shown below.
(1) Let the set of the k-th initial gathering point be $L^{(0)}=\left\{x_1^{(0)},x_2^{(0)},\cdots,x_k^{(0)}\right\}$.

Referred to as $G_i^{(0)}=\left\{x:d(x,x_i^{(0)}),j=1,2,\cdots,k,j\neq i\right\}\ (i=1,2,\cdots,k)$. Then, divide the sample into k disjoint categories to get an initial classification:

$$G^{(0)}=\left\{G_1^{(0)},G_2^{(0)},\cdots,G_k^{(0)}\right\}$$

(2) Calculate the new cluster set $G^{(0)}$ from the initial class $L^{(1)}$, and calculate

$$x_i^{(1)} = \frac{1}{n_i} \sum_{x_l \subseteq G_i^{(0)}} x_l \quad (i = 1,2,\cdots,k)$$

Where n_i is the number of samples in the class $G^{(0)}$, to get a new set

$$L^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \cdots, x_k^{(1)}\}$$

Start from $L^{(1)}$ and then classify it as

$$G_i^{(1)} = \{x : d(x, x_i^{(1)}) \le d(x, x_j^{(1)}), j = 1,2,\cdots,k, j \ne i\} \quad (i = 1,2,\cdots,k)$$

Get a new category

$$G^{(1)} = \{G_1^{(1)}, G_2^{(1)}, \ldots, G_k^{(1)}\}$$

Repeat the above steps m times to get

$$G^{(m)} = \{G_1^{(m)}, G_2^{(m)}, \ldots, G_k^{(m)}\}$$

Among them, x_i^m is the center of gravity of class $G_i^{(m-1)}$, x_i^m is not necessarily a sample. When m gradually increases, the classification also tends to be stable. At the same time, x_i^m can be regarded as the center of gravity of $G_i^{(m)}$ approximately, that is,

$$x_i^{(m+1)} \approx x_i^{(m)}, G_i^{(m+1)} \approx G_i^{(m)}$$

calculation can stop at this time.

The above has completed the introduction of the clustering algorithm model, and then checked the information to determine the relevant information about the political position of the Republican candidate Donald Trump and the Democratic opponent Joe Biden on the general election, as shown below.

(1) New coronavirus

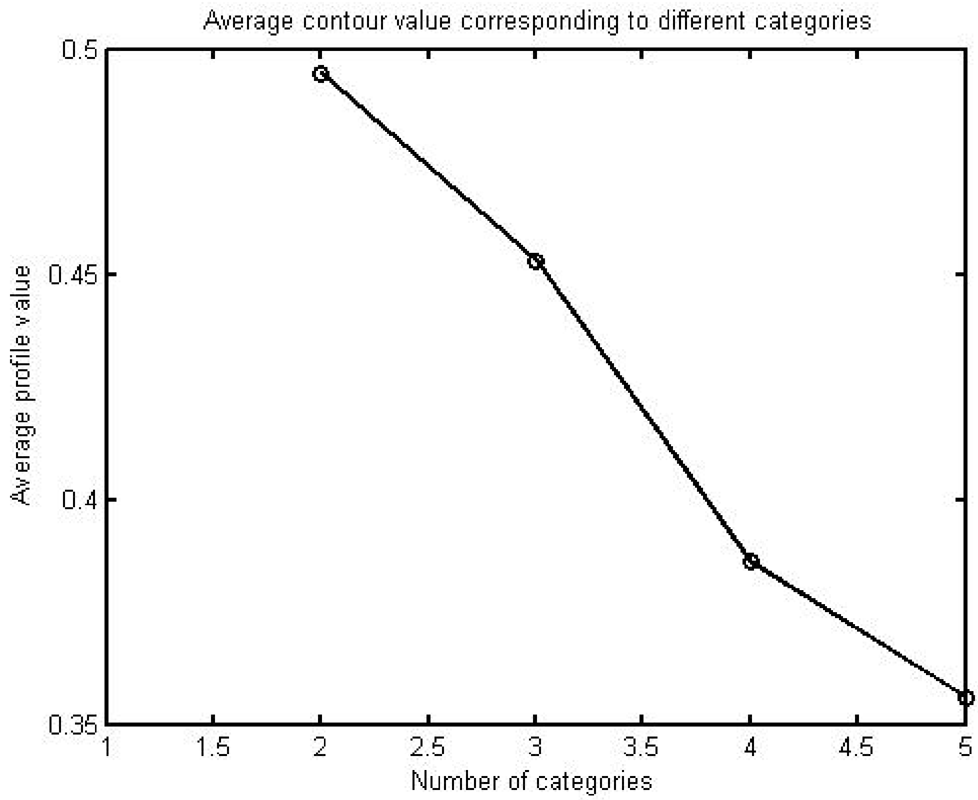
Trump established a new crown epidemic working group at the end of January this year. He said that the current focus of the working group has shifted to "safety and open our country." Trump also prioritized speeding up the progress of new crown treatments and vaccines, and he has allocated $10 billion in funds for such projects.

Biden hopes to establish a national contact tracking program, set up at least 10 testing centers in each state, and provide free new crown testing for everyone. He supports the introduction of mandatory masks across the country. If so, everyone in all federal industries will be required to wear facial masks.

(2) Climate

Trump is a climate change skeptic, and he hopes to expand the use of non-renewable energy. His goal is to increase oil and gas drilling and reduce further environmental protection activities. He is committed to withdrawing from the "Paris Climate

Agreement", an agreement reached by the international community to address climate change. The United States will officially withdraw from this framework later this year.This article takes the scenic spot A in question 1 as an example, obtains the corresponding attraction ability by consulting the data, the specific data is shown in the attachment, and then the clustering algorithm is used for analysis, and the MATLAB software is used to compile the program,and the following results can be obtained:



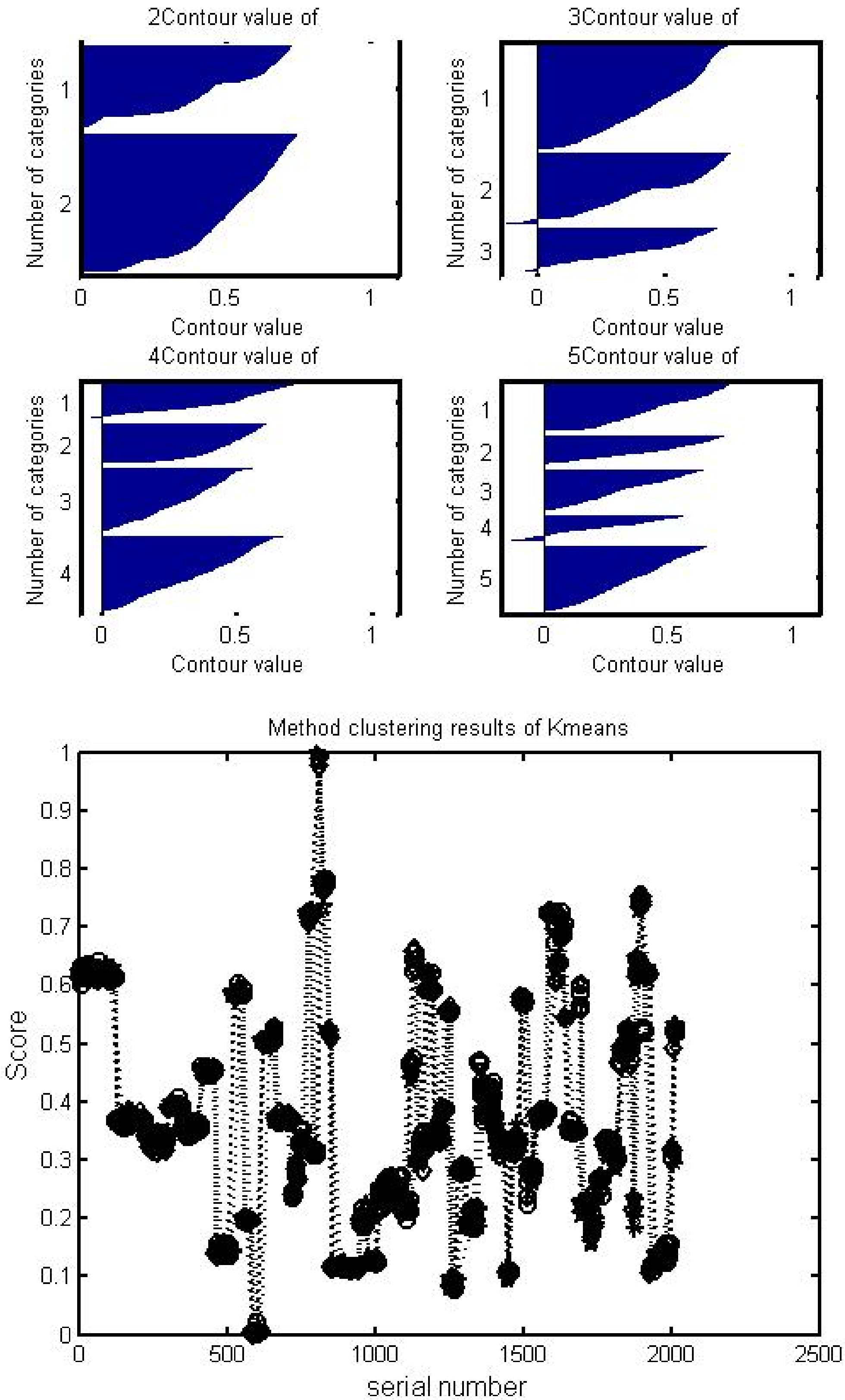


Figure 3 Clustering result graph

According to the analysis, the corresponding specific classification results can be determined as shown in the attachment.

5.3 Solving Problem Three

For question three, according to different prevalence, provide a reasonable current limit model for scenic spots, and take a certain scenic spot as an example, give specific quantitative analysis results; this article takes Beijing Forbidden City as an example to determine the main factors of current limit , And then determine the objective function that can occur epidemic infection, adopt 0-1 planning model, use MATLAB software to compile the program, and use genetic algorithm for optimization analysis, so as to determine the most reasonable current limiting plan.
Genetic Algorithms is a parallel random search optimization method proposed by Professor Holland of Michigan University in 1962 to simulate the genetic mechanism of nature and the theory of biological evolution.

(1) Select operation

Selection operation refers to selecting individuals from the old group to the new group with a certain probability. The probability of an individual being selected is related to the fitness value. The better the quality of the individual's usage, the greater the probability of being selected.

(2) Cross operation

The crossover operation refers to the selection of two individuals from individuals and the exchange and combination of the two dyes to produce a new outstanding individual. The crossover process is to choose two chromosomes from the population and randomly select one or more chromosomal positions for exchange.

A:1100 0101 1111 $\longrightarrow$ A:1100 0101 0000
B:1111 0101 0000 $\quad\quad\quad$ B:1111 0101 1111

(3) Mutation operation

The mutation operation refers to selecting an individual from the group and selecting a point in the chromosome to mutate to produce a better individual.

A:1100 0101 1111 $\longrightarrow$ A:1100 0101 1101

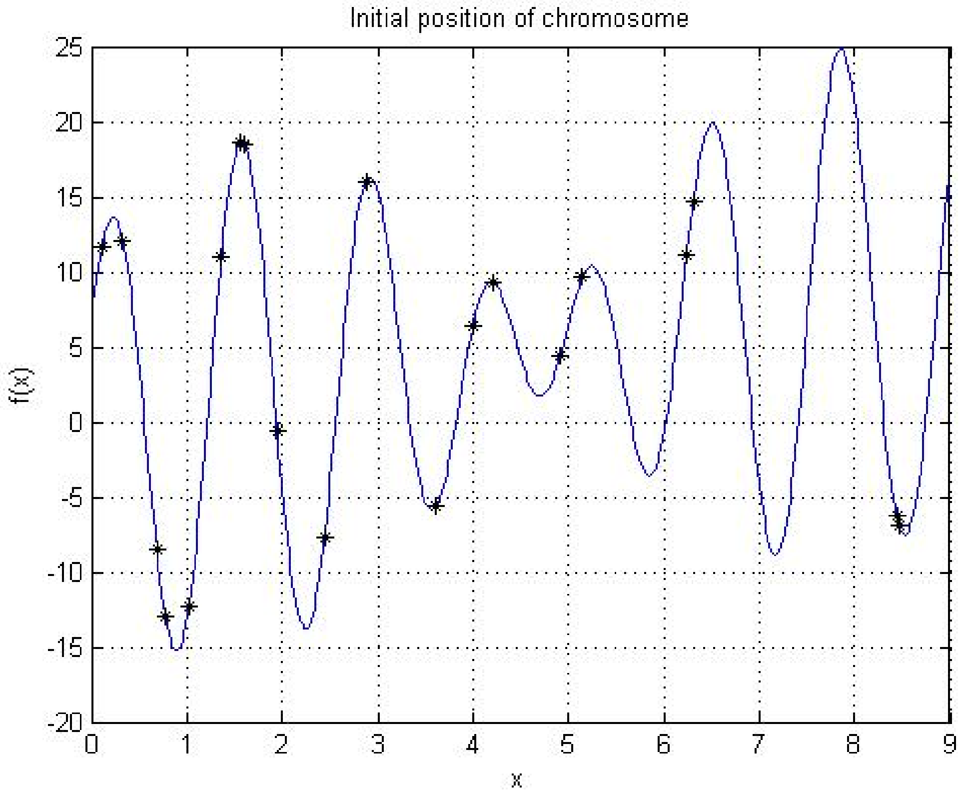
Genetic algorithm has the characteristics of efficient heuristic search and parallel computing. It has been applied in function optimization, combinatorial optimization and production scheduling.
The basic elements of genetic algorithm include fitness function of chromosome encoding method, genetic operation and operating parameters.
The chromosome encoding method refers to the individual encoding method, which currently includes binary method and real number method. The binary method refers to encoding an individual into a binary string, and the real number method refers to encoding an individual into a real number string.
The fitness function refers to a function that is written according to the target to calculate the fitness value of an individual. The fitness value of each individual is calculated through the fitness value function and provided to the selection operator for selection.
The operating parameters are the parameters determined during the initialization of the genetic algorithm, which mainly include the population size M, the genetic algebra G, the crossover probability Pc and the mutation probability Pm.

(4) Implementation of genetic algorithm

Genetic algorithm optimization neural network is to use genetic algorithm to optimize the initial weight and threshold of BP neural network, so that the optimized BP neural network can better predict output.

In the formula, amax is the upper bound of gene aij; amin is the lower bound of gene aij; f(g)=r2(1-g/Gmax); r2 is a random number; g is the current iteration number; Gmax is the maximum evolution number; r is a random number between [0,1].

According to the requirements of the subject, using the data objective function and using MATLAB software to analyze, the results shown below can be obtained.



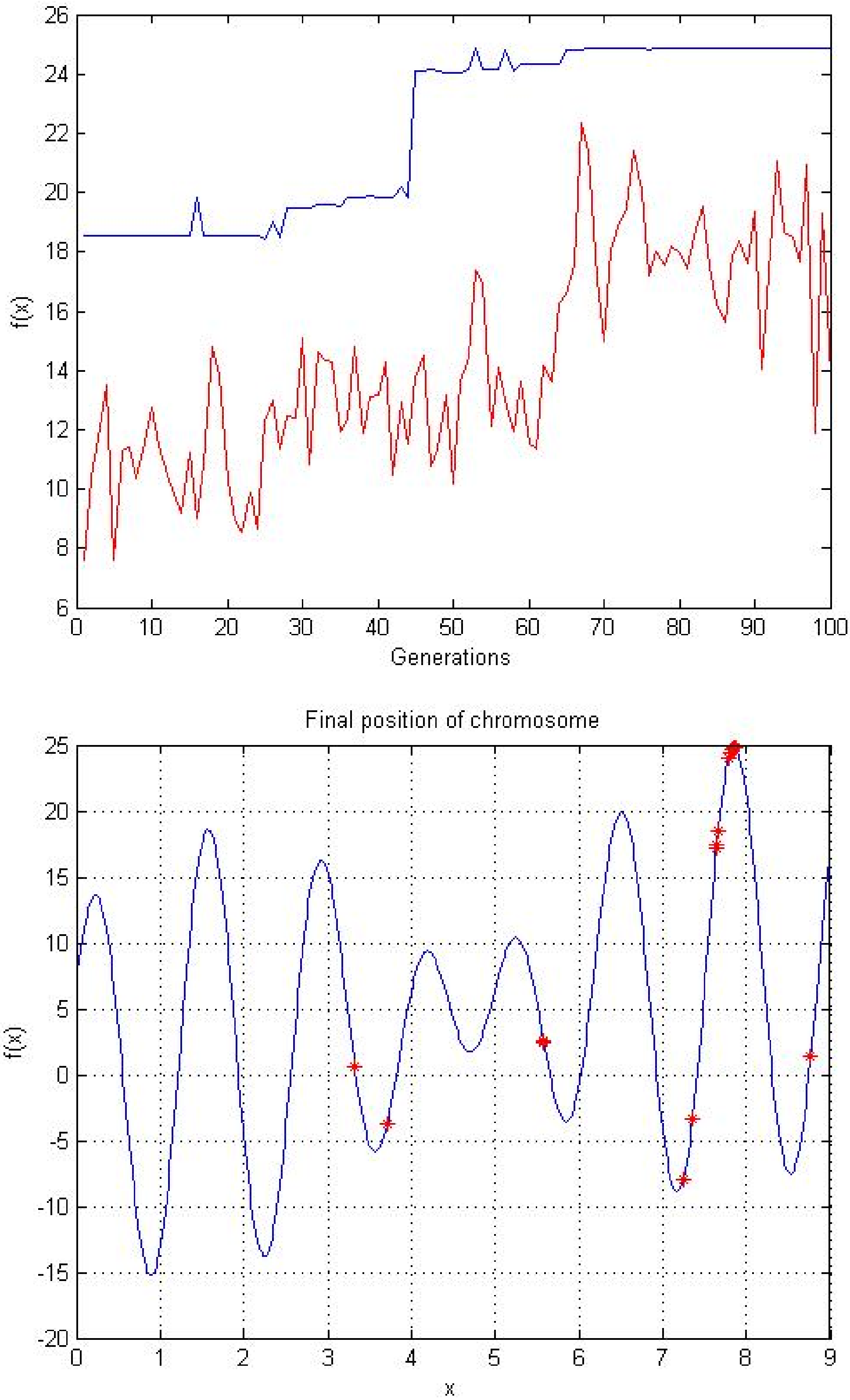


Figure 4 Results of genetic algorithm

After analysis, it can be determined that a reasonable flow restriction policy is a reasonable control of the normal number of people in the scenic area. When the number of people in the scenic area reaches a certain index, the number of people in the scenic area should be limited. At the same time, personnel activity areas can be set to avoid excessive population concentration.

5.4 Solving the fourth question

For question four, In this paper, the membership matrix of epidemic prevention and control, tourism income, and tourist experience is obtained by consulting data, and then the corresponding main factors are determined by analytic hierarchy process, and then the objective function of the epidemic infection rate is established, and the neural network prediction model is used to analyze and simulate If you control the expected effect.

Analytical Hierarchical Process (AHP) is a decision-making method that decomposes the elements related to the decision-making problem into goals, standards, plans and other levels, and performs qualitative and quantitative analysis on this basis. The steps in the process of analyzing the hierarchy are as follows:

(1) Establish a hierarchical structure model

On the basis of detailed analysis of actual problems, the relevant factors are divided into several levels from top to bottom according to various attributes. Factors of the same level are subordinate to or affect factors of higher levels, while leading the next level.

(2) Constructed into a pair of comparison matrix

Take two factors x_i and x_j each time, and use a positive number a_{ij} to express the ratio of importance of x_i and x_j. Get the matrix from all the results:

$$A = (a_{ij}) \tag{10}$$

It is called a pairwise comparison matrix. Apparently:

$$a_{ij} = \frac{1}{a_{ji}}, \quad a_{ij} > 0, \ 1 \le i, \ j \le n \tag{11}$$

(3) Calculate the weight vector and do the consistency test

If the comparison of n decision factors is logically consistent, there should be a relationship between elements a_{ij} in the paired comparison matrix:

$$a_{ij} \cdot a_{jk} = a_{ik}, 1 \le i, \ j, \ k \le n \tag{12}$$

In fact, the importance of each factor has an importance index. Suppose the importance index of factor x_i is ω_i, then according to the ratio of the importance of x_i to x_j, that is

$$a_{ij} = \frac{\omega_i}{\omega_j} \tag{13}$$

$$a_{ij} \cdot a_{jk} = \frac{\omega_i}{\omega_k} \cdot \frac{\omega_j}{\omega_k} = a_{ik}, 1 \le i, \ j, \ k \le n \quad (14)$$

Let A be a consistent matrix. Expressed by the corresponding ω_i ,then

$$A = \begin{pmatrix} \frac{\omega_1}{\omega_1} \frac{\omega_1}{\omega_2} \cdots \frac{\omega_1}{\omega_n} \\ \frac{\omega_2}{\omega_1} \frac{\omega_2}{\omega_2} \cdots \frac{\omega_2}{\omega_n} \\ \frac{\omega_3}{\omega_1} \frac{\omega_3}{\omega_2} \cdots \frac{\omega_3}{\omega_n} \\ \frac{\omega_n}{\omega_1} \frac{\omega_n}{\omega_2} \cdots \frac{\omega_n}{\omega_n} \end{pmatrix} \quad (15)$$

If A is not consistent, $\lambda_{\max}(A) > n$ can be proved. Moreover, the greater the degree, the greater the difference. At this time, the feature vector Y corresponding to $\lambda_{\max}(A)$ cannot accurately reflect the ratio of $X = \{x_1, x_2, \ldots, x_n\}$ to the target Z.

Then:

$$CI = \frac{\lambda_{max}(A) - n}{n-1} \quad (16)$$

CI is used as a measure of the degree of inconsistency in the comparison matrix A, and CI is called a consistency index.

Sati proposed to use the average random consistency index RI to test whether the comparison matrix A of the pair is sufficiently consistent. we can get:

$$RI = \frac{\lambda'_{max} - n}{n-1} \quad (17)$$

Then

$$CR = \frac{CI}{RI} \quad (18)$$

In practice, the following method can also be used to calculate the approximate value of $\lambda_{\max}$ and the corresponding feature vector.

Pair $A = (a_{ij})$, let

$$u_k = \frac{\sum_{j=1}^{n} a_{kj}}{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}} \quad (19)$$

In this paper, the membership matrix of epidemic prevention and control, tourism income, and tourist experience obtained by consulting data is as follows:

$$\begin{bmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 2 \\ 1/3 & 1/2 & 1 \end{bmatrix}$$

We import its weight matrix into Matlab software, and use analytic hierarchy process and calculation to get:

$$\lambda_{\max} = 1.8532 \quad CI=0.0046 \quad CR=0.0088<0.1$$

Can pass the consistency check. The corresponding weight vector is:

$$\omega_1 = (0.5396, 0.2970, 0.1634)^T$$

From this, it can be determined that the ratio of epidemic prevention and control, tourism income, and tourist experience can be approximated as: 5:3: 2. This article selects the first two factors for analysis, and obtains the corresponding target and establishes the target of the epidemic infection rate. function:

$$\min Z = \sum_{i=1}^{M} \sum_{j=1}^{L_i} \left[\sum_{k=1}^{n_{ij}} d_{r_{ij(k-1)} r_{ijk}} + d_{r_{ij0} r_{ijn_{ij}}} f(n_{ij}) \right]$$

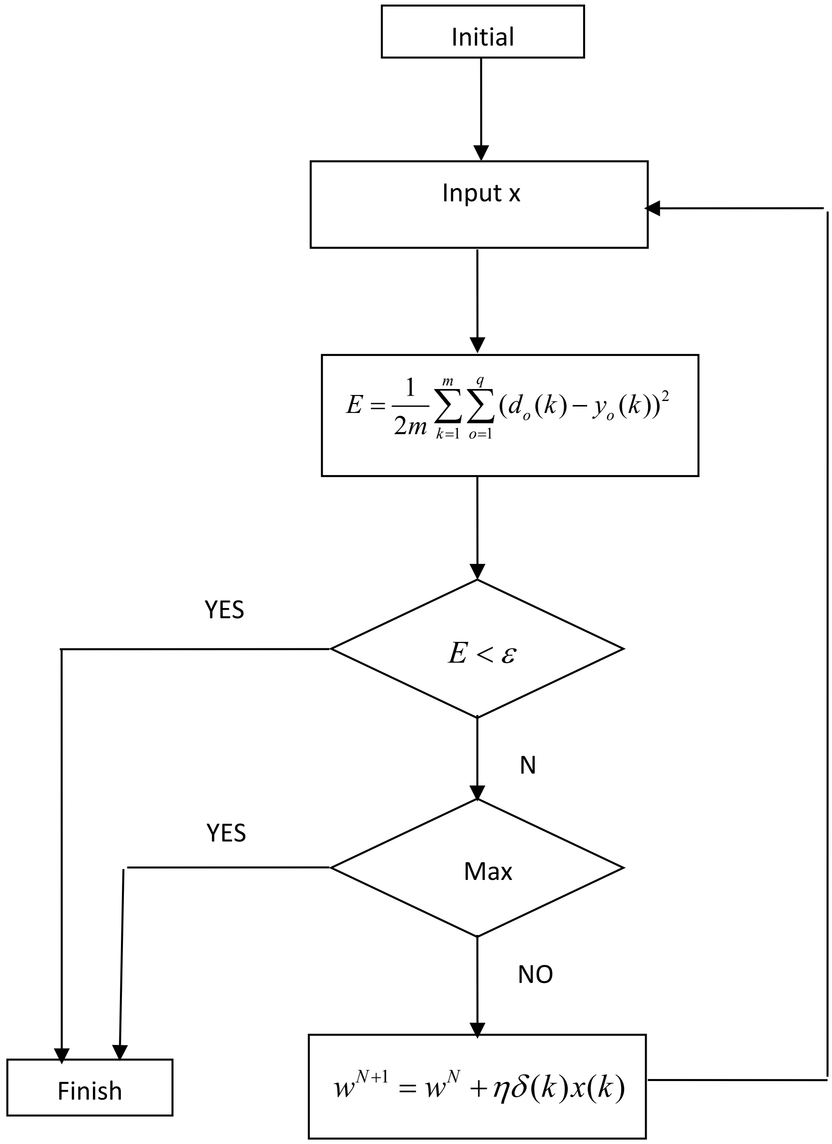
$$\sum_{k=1}^{n_{ij}} q_{r_{ijk}} \le Q_{ij}$$

$$0 \le n_{ij} \le N \qquad 20)$$

$$\sum_{i=1}^{M} \sum_{j=1}^{L_i} n_{ij} = N \qquad (20)$$

Then, the neural network algorithm is used for predictive analysis. Input signal x_i Acting on the output node through the intermediate node , through the non-linear transformation, generating the output signal y_k , Each sample of network training includes an input vector x and expected output d , the Relative error between the network output value y and expected output d . Adjust the connection strength of the input node and the hidden layer node w_{ij} and the strength of the connection between the hidden layer node and the output node T_{jk} and the threshold value. Causing the error to fall in the gradient direction, and after repeated learning training, the network parameters corresponding to the minimum error are determined, and the training stops. At this time, the trained neural network can process the non-linear conversion information with the smallest output error for the input information of similar samples.

According to the algorithm flow of the neural network, the corresponding algorithm flow chart can be determined as follows:



Figue5 Algorithm flowchart

Based on the above objective function, the neural network algorithm is used for predictive analysis, and the corresponding forecast result of the target function of the infection rate in the scenic spot can be determined as follows:

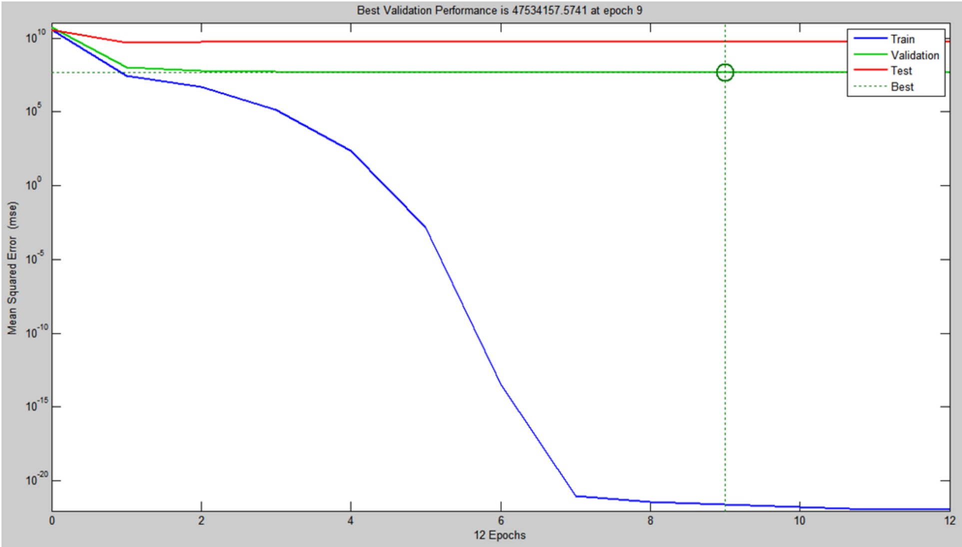


Figure 6 Process diagram of predictive analysis

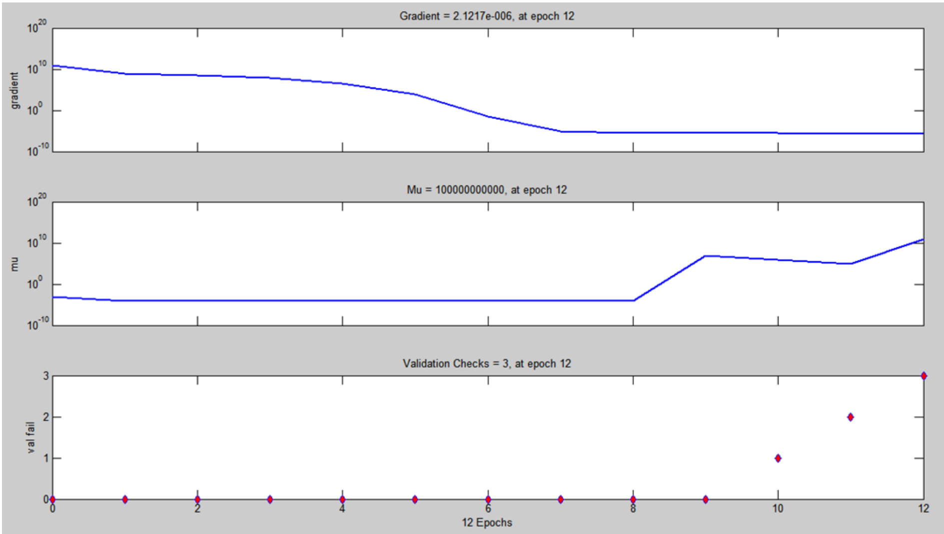


Figure7 Training process diagram

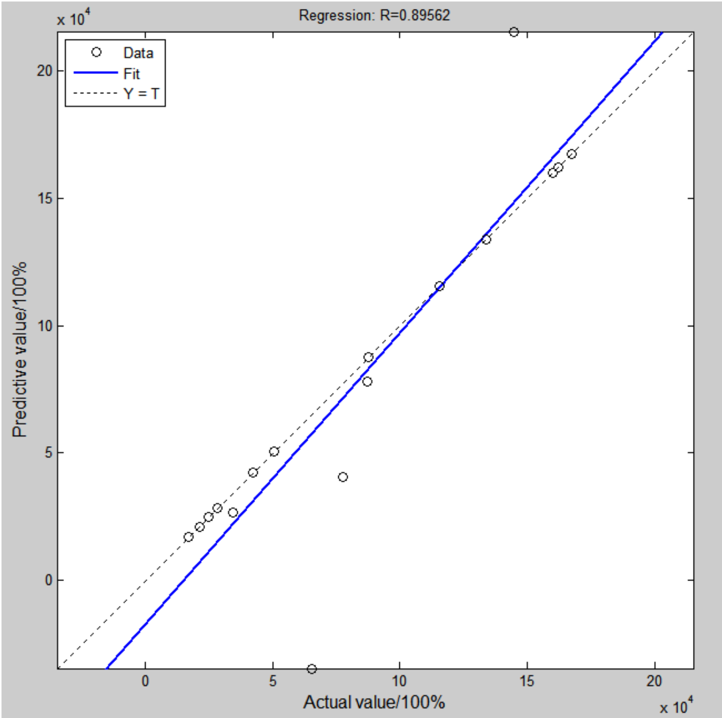


Figure 8 Error analysis diagram

According to the results, it is found that if reasonably controlled, the incidence of epidemic infections can be effectively reduced. Therefore, the results of the analytic hierarchy process should be used to reasonably control the spread of the epidemic.

5.5 Solving Problem Five

Regarding question 5, this article draws up a differentiated management plan for different scenic spots during the epidemic based on the results of questions 1 to 4 above.

The outbreak of the epidemic has eliminated speculators in the industry and promoted the acceleration of the transformation and upgrading of the tourism industry. The old saying goes: "Those who do not seek the overall situation are not enough to seek a domain; those who do not seek the world are not enough to seek a moment." In the process of transformation and innovation of the tourism industry, Internet thinking, business model innovation, and industrial chain thinking are very important. Thought is the forerunner of action. The research team believes that tourism companies should use the following thinking and strategies when planning their future development:

One is the Internet thinking of scenic area management. The epidemic has brought great negative effects to people in life and production, but it has unexpectedly promoted the rapid spread of various mobile Internet applications across the country. After more than two months of repeated training, most people have learned and adapted to the way of life and work on the Internet, whether they like it or not. For the tourism industry, how to achieve the organic integration of offline tourism and online operations will become the key to the success or failure of the restoration of production after the epidemic.

The second is the online strategy of innovative business models. First, the scenic spot cloud viewing. Affected by this epidemic, new forms of tourism such as cloud live broadcasting and cloud viewing have become increasingly popular. Using cloud viewing methods to carry out promotional activities and festival promotion can well stimulate tourists' desire to consume. Second, online tourism marketing. Use self-built platform and large-scale OTA platform for network distribution. When making full use of WeChat, Weibo, official websites, and official APP clients to carry out online marketing, there should be trade-offs-"Baidu" is the first choice for information search; "Sina Weibo" is the focus of hot topic discussion and secondary diffusion; Instant It is recommended to choose “WeChat” for social and topic placement; “Tik Tok” is recommended for interactive scenes of creative videos. Third, live travel and shopping. It is understood that during the epidemic period, traditional small and medium-sized travel agencies gave full play to their advantages in mastering consumer end customers, and carried out online travel product purchasing and sales activities, and achieved good results; Unlike traditional travel agencies, Liang Jianzhang, chairman of the board of directors of Ctrip, a large domestic OTA company exist

At the time of the first live broadcast, the transaction volume of Sanya hotel packages reached within one hour exceeded 10.25 million yuan. The sound waves in the live broadcast room also floated in the top 20 on the entire platform for a long time. The organic combination of live broadcasting and tourism shopping provides an opportunity for the integration of the entire industry chain of tourism products, and it is worth learning from the majority of scenic spots.

The third is the convergent thinking of the entire industry chain. First, give full play to the integrated functions of industry associations. Integrate the resources of local large, medium and small tourism enterprises, realize the complementary advantages of resources, and the development of mutually beneficial and mutually beneficial cooperation between large and small tourism enterprises, and create a community with a shared future for regional tourism. Second, we must have industrial chain thinking. The tourism industry is composed of all kinds of enterprises that meet the six aspects of people’s travel activities: “food, housing, transportation, travel, shopping, and entertainment”. A win-win situation can only be achieved by adopting the development model of competition and cooperation. Therefore, each tourism enterprise must focus on itself, consider the vertical and horizontal issues of integration between upstream and downstream industry chain enterprises and horizontal and vertical similar or complementary tourism enterprises, and plan cooperation, mutual benefit, and mutual benefit with the collective thinking of the entire industry chain. Win-win is a development model with overall, strategic and long-term nature.

The fourth is to improve the quality of tourism. Improving quality and efficiency is the internal driving force and requirement for the development of tourism. Tourism has become a "strategic pillar industry of the national economy", and improving the "quality" of the tourism industry is an important means to solve the "insufficient supply of effective tourism" in my country. From the perspective of planning and operation, the scenic area is a "stage" that attracts landscapes, assists facilities, and serves as a means to meet tourists' travel experience needs. The A-level evaluation standard for scenic spots is a bottom-line standard for tourism quality, not the highest standard. In layman's terms, the improvement of tourism quality can only be better, not the best. For the vast number of scenic spots, the epidemic is both a crisis and a turning point. Scenic spots that have been fully prepared during the epidemic will resume production after the epidemic, and the market can be quickly restored through product updates and quality improvements. After the epidemic, all scenic spots should introduce quality management while increasing product innovation. The quality of tourist experience should be the center and the full participation of all employees in the scenic spot should be based on the goal of satisfying tourists and achieving sustainable and healthy development of the scenic spot.

6 Analysis of the advantages and disadvantages of the model

6.1 Model advantages

① The principle of the classic multiple linear regression analysis model is clear, the model is simple, easy to understand, and easy to apply. Therefore, people often use multiple linear regression analysis in many research fields in daily life.

②The classic multiple linear regression analysis model can be directly analyzed and calculated with the ready-made EViews software, so its calculation is simple, fast and accurate.

6.2 Model shortcomings

①Multivariate linear regression analysis is very sensitive to a small amount of ill-conditioned data in a large number of samples, and the appearance of a small amount of ill-conditioned data will often affect the fitting effect, thereby making the prediction results inaccurate. In practical applications, some abnormal data will inevitably appear, but the classic multiple linear regression analysis is helpless for these problems.

②Multivariate linear regression analysis models have very high requirements on the amount of data. Generally, the amount of data required for regression analysis is very large, but in real life, collecting data is often very difficult.

7 References

[1] Dai Jingwen, Jiang Xingyi. Research on the development trend and countermeasures of tourist attractions in the post-new crown pneumonia era—— Taking the questionnaire survey and data analysis of Shaanxi Province as an example [J]. Special Economic Zone, 2021(10): 128-132.
[2] Sheng Qiaoling, Wang Yanhua. An Empirical Analysis of the Satisfaction of Scenic Facilities Layout from the Perspective of Tourists' Convenience in the Post-epidemic Period: Taking Chongqing Chiyou Jiuli City Scenic Spot as an Example [J]. Public Standardization, 2021(19): 183-185 .
[3]Sun Yudan. Experts from the Municipal Health Commission interpreted the "three enhancements" [N]. Wenzhou Daily, 2021-10-01 (002).
[4]Guo Yanping. Countermeasures for the high-quality development of Shanxi's tourist attractions under the new crown epidemic[J]. Journal of Jincheng Vocational and Technical College, 2021, 14(05): 41-44.
[5] Guo Qi, Yan Fei, Wu Jing. The integration of cultural and tourism adds a bright spot, hard work to accumulate kinetic energy [N]. Hainan Agricultural Reclamation News, 2021-08-20 (003).
[6]Cheng Li, Zhao Chenyue. Research on the impact of tourists' psychological carrying capacity in outdoor scenic spots under the background of the new crown pneumonia epidemic: Empirical evidence based on visual behavior experiments[J]. Tourism Tribune, 2021, 36(08): 27-40.
[7]Wang Jiale, Wang Mengdi, Chu Can, Kou Daitong, Liu Xiaofan, Jiang Xiyu. Recovery and development of scenic spots in the post-epidemic period: Taking Penglai Pavilion as an example [J]. Industry and Technology Forum, 2021, 20(15): 15- 16.
[8] Ren He, Wang Yahui. Research on New Media Marketing Strategy in Yuntai Mountain Scenic Spot under the Background of New Coronary Pneumonia Epidemic[J]. Inner Mongolia Science Technology and Economy, 2021(12): 43-45+48.
[9] Li Ce. Development countermeasures for 5A-level tourist attractions in Hunan Province under the background of normalization of epidemic prevention and control[J]. Journal of Huaihua University, 2021, 40(02): 44-47.
[10] Zhou Chan. Social responsibility and financial management strategies of tourist attractions under the normalization of epidemic prevention and control[J]. Enterprise Reform and Management, 2021(06): 219-220.
[11] Guide to epidemic prevention and control measures for the reopening of tourist attractions [N]. China Tourism News, 2021-03-19 (002).

Appendix

```
clc
clear;
load ssgs-A.txt;
x=ssgs_A(:,1);
y=ssgs_A(:,2);
plot(x,y,'*');
xlabel('longitude')
ylabel('latitude')
%%
clc, clear all, close all
load ssgs1.txt
x=ssgs1;
[n,m]=size(x); %
[X,ps]=mapminmax(x');
ps.ymin=0.002; %
ps.ymax=0.996; %
ps.yrange=ps.ymax-ps.ymin;
X=mapminmax(x',ps);
A=X';
X=A;
numC=5;
for i=1:numC
kidx = kmeans(X,i);
silh = silhouette(X,kidx);
silh_m(i) = mean(silh);
end
figure
plot(1:numC,silh_m,'ko-', 'linewidth',2)
set(gca,'linewidth',2);
xlabel('Number of categories')
ylabel('Average profile value')
title(' Average contour value corresponding to different categories')
%
figure
set(gca,'linewidth',2);
for i=2:5
kidx = kmeans(X,i);
subplot(2,2,i-1);
[~,h] = silhouette(X,kidx);
set(gca,'linewidth',2);
title([num2str(i), 'Contour value of ' ])
```

```
snapnow
xlabel('Contour value');
ylabel('Number of categories');
end
[idx,ctr]=kmeans(A,4);
c1=find(idx==1); c2=find(idx==2);
c3=find(idx==3); c4=find(idx==4);
figure
F1 = plot(find(idx==1), A(idx==1),'k:*', ...
find(idx==2), A(idx==2),'k:o', ...
find(idx==3), A(idx==3),'k:p', ...
find(idx==4), A(idx==4),'k:d');
set(gca,'linewidth',2);
set(F1,'linewidth',2, 'MarkerSize',8);
xlabel('serial number','fontsize',12);
ylabel('Score','fontsize',12);
title('Method clustering results of Kmeans')
disp('Clustering result：');
disp(['Type 1:' ,'Center point：',num2str(ctr(1)),'    ','This kind of sample number：', num2str(c1')]);

disp(['Type 2:' ,'Center point：',num2str(ctr(2)),'    ','This kind of sample number：', num2str(c2')]);

disp(['Type 3:' ,'Center point：',num2str(ctr(3)),'    ','This kind of sample number：', num2str(c3')]);

disp(['Type 4:' ,'Center point：',num2str(ctr(4)),'    ','This kind of sample number：', num2str(c4')]);

%%
clear all;
close all;
clc;
n=20;ger=100;pc=0.65;pm=0.05;
v=init_population(n,22);
[N,L]=size(v);
disp(sprintf('Number of generations:%d',ger));
disp(sprintf('Population size:%d',N));
disp(sprintf('Crossover probability:%.3f',pc));
disp(sprintf('Mutation probability:%.3f',pm));
xmin=0;xmax=9;
f='x+10*sin(x.*5)+7*cos(x.*4)';
x=decode(v(:,1:22),xmin,xmax);
fit=eval(f);
figure(1);
fplot(f,[xmin,xmax]);
grid on;hold on;
```

```
plot(x,fit,'k*');
title('Initial position of chromosome');
xlabel('x');ylabel('f(x)');
vmfit=[];
vx=[];
it=1;
while it<=ger
    %Reproduction(Bi-classist Selection)
    vtemp=roulette(v,fit);
    %Crossover
    v=crossover(vtemp,pc);
    %Mutation
    M=rand(N,L)<=pm;
    %M(1,:)=zeros(1,L);
    v=v-2.*(v.*M)+M;
    %Results
    x=decode(v(:,1:22),xmin,xmax);
    fit=eval(f);
    [sol,indb]=max(fit);
    v(1,:)=v(indb,:);
    fit_mean=mean(fit);
    vx=[vx sol];
    vmfit=[vmfit fit_mean];
    it=it+1;
end
%%
disp(sprintf('\n'));
disp(sprintf('Maximum found[x,f(x)]:[%.4f,%.4f]',x(indb),sol));
figure(2);
fplot(f,[xmin,xmax]);
grid on;hold on;
plot(x,fit,'r*');
title('Final position of chromosome');
xlabel('x');ylabel('f(x)');
figure(3);
plot(vx);
xlabel('Generations');ylabel('f(x)');hold on;
plot(vmfit,'r');hold off;
runtime=toc
```