Hindawi
Mobile Information Systems
Volume 2022, Article ID 4377286, 10 pages
https://doi.org/10.1155/2022/4377286

Research Article

Research on Development Strategy of Ethnic Sports Tourism Resources Based on Stochastic Forest Algorithm

Chunli Nie

College of Education, Guangxi Vocational Normal University, Nanning, Guangxi 530007, China

Correspondence should be addressed to Chunli Nie; ncl1985@126.com

Received 11 May 2022; Revised 15 July 2022; Accepted 28 August 2022; Published 26 September 2022

Academic Editor: Jiafu Su

Copyright © 2022 Chunli Nie. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The interactive development mode of combining sports industry with tourism industry has attracted more and more people's attention and gradually been accepted by people. The organic combination of the two can effectively promote the healthy growth of regional economy. The development and operation mode of minority STRs (sports tourism resources) has become an important topic in the research of sports tourism development. It is not only important for the successful integration of minority traditional sports into tourism industry but also of great significance for promoting the development of sports tourism industry. In this paper, the personalized sports tourism service recommendation system based on multi-objective optimization is studied, and a tourism service combination method based on multi-objective optimization is proposed. This method is based on multi-objective optimization, and the historical data of tourists and their current preferences are considered, respectively. The converted data are used to train the RF (random forest) model offline, and online recommendation only needs to be scored and predicted according to the rules of the RF model. The results show that the online recommendation time of the proposed algorithm is basically below 100 s, which is much lower than that of other recommendation algorithms. The experimental results show that setting the weights of user evaluation information and related tourism information can further improve the matching degree between recommendation results and users' needs.

1. Introduction

China is a multi-ethnic country, and each ethnic group has its own customs and habits. As a part of its excellent culture, ethnic traditional sports embodies the hard work and wisdom of the working people of all ethnic groups in the process of life and production, and it has survived through years of tempering. With the continuous improvement of people's living standard and the change of life concept, people have new requirements for spiritual life on the basis of satisfying material life. In recent years, the domestic tourism demand has been rising year by year. Combining traditional national sports with tourism, promoting the development of ethnic traditional sports, publicizing China's traditional culture, and creating economic benefits on the premise of ensuring sustainable development, how to coordinate the relationship between ethnic traditional sports and tourism has become an urgent problem to be solved.

Compared with foreign countries, the research on the development of sports tourism in China started late. In recent years, the research on sports tourism in China has developed rapidly, but there is still a certain gap compared with foreign countries. Generally speaking, the research on the basic theory of sports tourism in China is scattered and not thorough enough. Pu et al. believe that sports tourism is all forms of tourism activities related to sports activities [1], which are active or passive, accidental or spontaneous, and leaving home or workplace for non-commercial purposes. Sánchez-Martín et al. believe that sports tourism and tourism in general have both similarities and characteristics [2]. Huang et al. think that sports tourism, as a new field produced by the cross-penetration of tourism industry and sports industry, is a new form of tourism based on sports resources, which attracts people to participate in and feel sports activities and natural interests, a special leisure lifestyle combining sports and tourism, and an important part

of sports industry [3]. Szromek et al. put forward the theory of internal and external driving forces of sports tourism marketing. It is believed that the marketing of sports tourism should have horizontal and vertical horizontal and vertical relations, and on the basis of horizontal and vertical relations and professional and correct goals, stable development institutions and positive and effective behaviors should be established [4]. Chen et al. proposed to formulate the intangible ethnic traditional sports cultural heritage protection plan [5], put forward innovative ideas to promote the intangible cultural heritage, that is, the specific measures for the industrialization development of ethnic traditional sports, and then strengthened the protection of intangible cultural heritage and the rational development of ethnic traditional sports.

As an important consumption mode of tourist citizens, it is an important symbol of the living standard of residents in a country or region. Developing tourism and promoting the development of tourism has become a pillar of social economy. The external conditions of sports activities actually refer to STRs (sports tourism resources), especially its natural resources, which is an important factor to attract people to participate in sports tourism activities and also a symbol different from ordinary daily sports activities and sports tourism activities. As one of many algorithms of machine learning, RF (random forest) algorithm can achieve high-precision classification and evaluation of large-scale unknown data by training and learning small-scale known samples. Based on this, this paper uses RF algorithm to study the development of ethnic STR. Through the research on the development and operation of ethnic STR, on the basis of previous studies on STR, we put forward our own opinions and ideas on the development of ethnic STR, so as to make some contributions to the development of China's sports industry and ethnic STR.

2. Related Work

2.1. STR Development Research. As sports tourism has an important influence on the economic development of a region or a country and the protection of local traditional social culture, sports tourism has gradually become the focus of academic and industry attention. To study the development strategy of sports tourism, we must first define the concept of sports tourism scientifically.

Sports tourism is an outbound activity of leaving the circle of life for noncommercial purposes to watch or participate in a sporting event. [6]. Hao et al. believe that sports tourism is essentially sports event tourism, which focuses on the influence and management of events [7]. Khandaker et al. believe that ethnic traditional sports refer to some national folk traditional sports and recreational activities as the predecessor of modern sports. China's ethnic traditional sports should include Han traditional sports and minority traditional sports [8]. Koval et al. started with the natural resources, human resources, and school education of ethnic sports and pointed out that the research on the development of ethnic sports should be combined with the integration of resources, so as to create a broader world for the development of ethnic sports resources in China [9].

Traditional ethnic minority sports is a field of traditional sports studies, focusing on the sports activities with ethnic characteristics that have been formed in the process of production, practice, and life of all ethnic groups for thousands of years. This kind of sports activity is presented to the world in a certain form of sports events. Livieris et al. believe that STR of ethnic minorities is a kind of exploitable resource form based on the basic points of ethnic minorities, which has obvious ethnic characteristics and the attributes of sports. It embodies the unique advantages of ethnic resources, natural resources, human resources, location resources, and sports resources [10]. Gao et al. discussed the research background of ethnic minority sports from the perspective of development and put forward their own opinions and feelings on the further research of ethnic minority sports. It is emphasized that ethnic minority sports under the background of the development of modern society is a new field worth studying [11].

2.2. Research Status of RF Algorithm. RF algorithm is an ensemble learning algorithm based on DT (decision tree), which has high prediction accuracy, good tolerance for outliers and noise, and is not prone to overfitting. It is used in a wide range of disciplines.

Mallinis et al. used NCL (neighborhood cleaning rule) technology to process the unbalanced data when using RF algorithm and then applied RF algorithm to improve the classification accuracy [12]. Ma et al. proposed a new oversampling method [13] to solve the problem of poor classification performance of RF algorithm on unbalanced datasets. Liu et al. sorted the DTs in the RF model in descending order according to AUC (area under curve) value, selected the DTs with high AUC value, then calculated the similarity of these DTs, and generated the similarity matrix. After clustering, the highest AUC value in each class was selected to form RF model [14]. Ghorbanian et al. applied RF algorithm to recommendation system and proposed a multi-dimensional context-aware recommendation method based on improved RF algorithm. Experimental results show that this method can reduce the average absolute error and root mean square error [15]. Although the application of this algorithm is more and more extensive and its function is more and more powerful, the random forest algorithm is not perfect, and there is still some room for improving the performance of the algorithm in dealing with different datasets and the classification accuracy of the algorithm. Therefore, this paper proposes a new algorithm to reduce the number of continuous variables, which can provide a concise dataset for C4.5 algorithm, thus improving the execution efficiency of C4.5 algorithm.

Zhang et al. proved the most direct simplified version of the original RF, that is, when selecting split features and split values, it randomly selects one of all features as split feature, and at the same time, it randomly selects one value from the selected feature value as split value [16]. Bao et al. put forward another version with consistency which is very close

to the original RF. The difference lies in the selection of splitting values. They first randomly selected a subset of all the eigenvalues and then searched for the optimal splitting on this subset [17]. Zhang et al. further put forward the deep neural decision forest. They connected RF to the representation learning process of the deep neural network and turned the split function into a random decision function, so that it could update the parameters of the whole network through backpropagation [18]. He et al. also proposed a depth structure based entirely on RF [19]. They regarded RF as a node in the depth neural network. Select the feature with the best Gini value as the segmentation. In addition, they used the sliding window to scan the original features to generate new feature vectors, which were used as the input of the deep forest. It can be considered that the original features were transformed before being built. Sun et al. proposed a bi-objective elite differential evolution algorithm for multi-valued logic networks [20]. Faccioli et al. proposed a multi-objective evolutionary algorithm for protein structure prediction [21].

3. Research Method

3.1. Multi-Objective Optimization of Ethnic Sports Tourism Service Combination. STR development should not only consider the difficulty of resource development and the economic, social, and ecological benefits that resource development can receive but also consider the location and traffic conditions of resource location. According to the location and traffic conditions, STR development is generally divided into resource-based, customer-based, and resource-customer-based. If ethnic minority sports want to get rid of the previous traditional development model, they must integrate ethnic minority sports resources into the market and stand the test and choice of the market. Therefore, only by combining the ethnic minority sports cultural resources with the unique natural scenery and rich ethnic culture in minority areas can the sports resources in minority areas continue to develop and reflect the profound charm of traditional sports resources.

The resources of minority traditional sports tourism include natural resources, and much are humanistic resources. The development of traditional sports of ethnic minorities will certainly promote the development of various industries and realize the common development between tourist areas, local governments, and travel agencies. Therefore, in the process of developing the traditional STR of ethnic minorities, we must respect nature, love nature, and shape nature, so as to realize the organic unity of man and nature and achieve real harmonious coexistence. The strong appeal brought by sports tourism and entertainment activities is people's recognition and appreciation of their own national culture and history. Moreover, the joyful atmosphere, singing and dancing, and the warm atmosphere of unity in the festival enhance the consistency and coordination of the people within the clan.

In fact, the structure of multi-objective optimization problem is similar to that of single-objective optimization problem, which mainly includes objective function, constraints, and decision variables [20]. The difference is that the multi-objective optimization problem contains multiple objective functions and needs to be optimized simultaneously. A multi-objective optimization problem contains two or more objectives that need to be optimized at the same time, and the final result is a set of compromise solutions, the so-called Pareto optimal solution set. When designing algorithms, we should consider parallel computing, which mainly includes parallel computing when optimizing goals, parallel computing of fitness, and so on.

Tourism service composition is different from traditional web service composition, and tourism service composition has certain particularity. The final solution of the multi-objective service composition optimization model is not the only solution but a compromise of time, cost, quality of service, and other optimization objectives in the execution process. In this paper, the multi-objective problem is transformed into a single-objective problem. Firstly, by simplifying some constraints, a set of Pareto optimal solutions that can meet the constraints is obtained. Then, on this basis, these optimal solutions are sorted to get the most satisfactory combination for users.

The best results that users want are the shortest time, the lowest cost, and the best service quality. For convenience, this paper defines time d, cost p, and quality of service s as a function, as follows:

$$\begin{aligned} d &= f(x_1, x_2, \cdots, x_n), \\ p &= g(x_1, x_2, \cdots, x_n), \\ s &= h(x_1, x_2, \cdots, x_n), \end{aligned} \tag{1}$$

where $x_i\,(i = 1, 2, \cdots, n)$ is the parameter attribute of d, p, s.

0-1 variable x_{ij} is introduced. When $x_{ij} = 1$, it means that the jth alternative to realize the ith tourism service element is selected in the product design; otherwise, it is 0. Thus, the total cost function of product design can be obtained:

$$C = \sum_{i=1}^{6} \sum_{j=1}^{m} c_{ij} x_{ij} + c^*, \tag{2}$$

where c^* means the fixed cost of tourism product design, including staff salaries, research expenses, and so on. It is used to indicate the maximum cost that travel agencies can accept when designing travel products.

Tourism products are different from other service products. Optimizing the design of tourism products will inevitably involve the development of tourist attractions, which will inevitably have a certain impact on the ecological environment of tourist attractions. Let x be the eigenvector, and the final optimization problem can be expressed as

$$\begin{cases} \text{Min}\left(\sum_{i=1}^{n} w_i x_i + c \sum_{i=1}^{n} w_i^2\right), \\ s.t. \sum_{i=1}^{n} w_i = 1. \end{cases} \tag{3}$$

Among them, the first term of the optimization objective is the above optimization equation, the second term is the

regularization term of the parameters, and c is the balance parameter between them. Because we simplify the objective function to the linear combination of features, the constraint equation requires the sum of weights to be 1.

In the RF algorithm, C4.5 algorithm adopts the binary separation and dispersion method, which makes the efficiency of the algorithm have a great relationship with the value of continuous variables. Therefore, if we can effectively reduce the dataset and reduce the value of continuous variables, we can improve the execution efficiency of RF algorithm. In this paper, a new CVD (continuous variable disclusion) algorithm based on x^2correction is designed, and its calculation process is as follows:

(1) Calculate the number k of decision attributes in two adjacent intervals of a certain attribute value.

(2) Calculate the theoretical number E_{ij}, and the formula is

$$E_{ij} = \frac{R_i \cdot C_j}{N}, \tag{4}$$

where R_i is the number of samples in i interval; C_j is the number of class $j$samples; and N is the total number of samples in two adjacent intervals.

(3) The sequence of merging intervals is determined by the value of D, and its formula is as follows:

$$D = \frac{x_a^2 - x^2}{\sqrt{2v}}, \tag{5}$$

where v is the degree of freedom.

(4) Select the interval with the smallest D value to merge.

After all the continuous attribute variables in the dataset are subjected to the above steps, the reduction process of the dataset is completed, that is, the discretization of continuous variables is realized. The flowchart of CVD algorithm is as follows (Figure 1).

From the perspective of travel agencies or enterprises, we hope to design and optimize tourism products at the lowest cost and bring the greatest satisfaction to tourists. However, in real life, it is often difficult to achieve high-quality tourists' satisfaction and low cost at the same time, which requires enterprises or travel agencies to give weight to both according to the actual situation and make a scheme choice according to the final actual results.

3.2. STR Recommendation Based on RF. Every nation has its own unique traditional festivals. At present, these distinctive ethnic traditional festivals have become important tourism resources. Ethnic minorities will hold all kinds of traditional ethnic arts and sports activities when their traditional festivals come. However, the emergence of a new tourism model is bound to require the support of a large number of related manpower. Vigorously developing the research of ethnic traditional STR can enable more local people to return to their good jobs and show their local ethnic characteristic culture. As the market propagandist of traditional sports tourism of ethnic

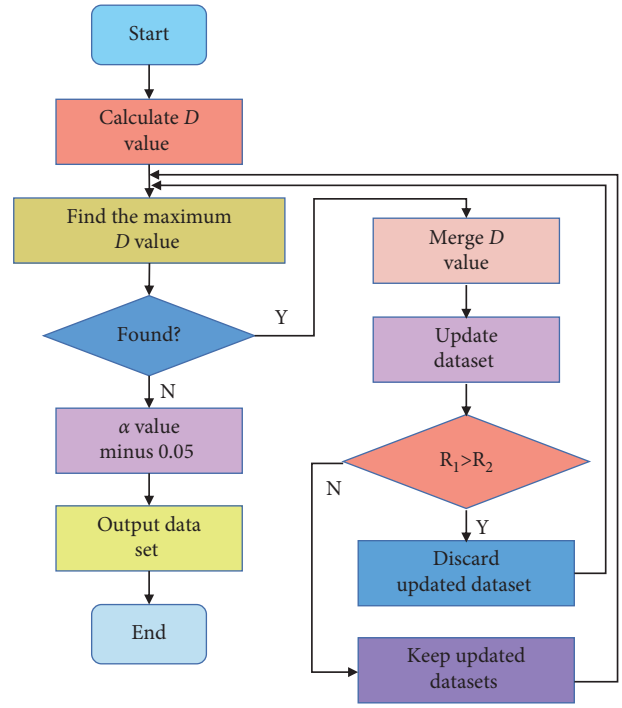

Figure 1: Flowchart of CVD algorithm program.

minorities, it is more important to promote themselves to the greatest extent. It is necessary to publicize the traditional sports tourism of ethnic minorities in various aspects and in all directions. Take various measures to encourage and mobilize citizens to participate in traditional sports of ethnic minorities, cultivate their habit of participating in traditional sports of ethnic minorities, and then cultivate the tourism market of traditional sports of ethnic minorities.

At present, the level of social technology and the changing trend have relatively little direct impact on the ethnic traditional sports tourism, but the indirect impact covers all aspects of the ethnic traditional sports tourism. Technological changes and breakthroughs have made the supporting industries of the ethnic traditional sports tourism more and more prosperous. In recent years, China's economy has been developing continuously, and residents' income level has been improved to a great extent. However, compared with developed countries, it is still at a lower level, and the proportion of income that can be used for traditional ethnic sports tourism is not very large, and it is still low compared with western developed countries. By developing and operating the ethnic traditional STR and forming the tourism industry with the ethnic traditional sports culture as the core, it can help more local minority people to engage in sports culture with ethnic characteristics and solve their employment difficulties.

With the popularization and development of network technology, the infinite potential of ethnic traditional sports tourism can be tapped to the maximum extent by the east wind of the network. The perfection of the construction of tourism platform is of great significance to the development of tourism. The recommendation system is used to recommend items and information that users may be interested in, which improves the information pushing efficiency of

information providers and the advertising efficiency of commodity providers, so that specific information can be delivered to specific people, thus achieving a win-win situation for businesses and users.

The DT in the recommendation system can be used in the model-based method. The user can construct a DT after evaluating only two items, and the characteristics of each item are used to build a model to explain the user's rating. The DT can only be used to simulate a special part of the system.

The traditional CF algorithm calculates the similarity between this user and other users by using the paired evaluation scores of common scoring items [19] but ignores the scores that only one user evaluates. In this chapter, DCF (distributed collaborative filtering) algorithm is proposed. In addition, in order to significantly improve the efficiency, we provide the map reduce-based distributed version of this algorithm running on Hadoop. The framework of the algorithm Map_DCF is shown in Figure 2.

In the process of recommendation, the problem of personalized matching with users can be dealt with by means of knowledge reasoning. Therefore, it is necessary to establish rules that conform to reasoning according to data sources. Therefore, a general user data ontology needs to be established here.

The feedback factor N of ethnic sports tourism users is introduced, and the weight μ of the feedback factor is considered at the same time, and then the final total matching degree value M is calculated by

$$M = D(1-\mu) + N\mu. \tag{6}$$

CF recommendation algorithm based on proximity relation can be divided into UCF (user-based CF) algorithm and ICF (item-based CF) algorithm. The basic idea of the former is that similar users have common preferences, while the latter assumes that users will like projects similar to those they liked in the past, so ICF recommends projects similar to those they liked in the past.

In the UCF algorithm, the target user a scores the item i as follows:

$$\hat{r}_{ai} = \overline{r}_a + \frac{\sum_{u\in N(a)}(r_{ui} - \overline{r}_u)sim(a,u)}{\sum_{u\in N(a)}|sim(a,u)|}, \tag{7}$$

where $N(a)$ represents the k users who are most similar to the target users among the users who have scored item i; r_{ui}indicates the rating of user u on item i; $\overline{r}_a, \overline{r}_u$ represents the average value of user's a, u history score; and $sim(a,u)$ indicates the similarity between two users.

For the problem of dividing a sample into a certain category, that is, discrete variables, the DT in RF generally uses Gini value as the evaluation standard. It is defined as

$$Gini = 1 - \sum P(i)^2. \tag{8}$$

Among them, $P(i)$ is obtained according to the sample classification ratio, that is, the ratio of i classification to the total sample.

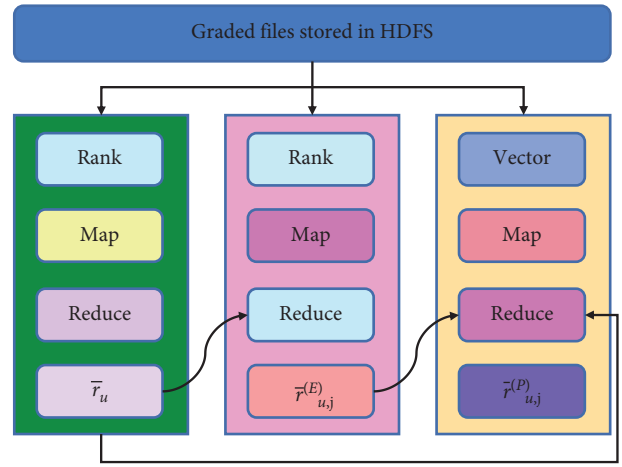

Figure 2: Map_DCF algorithm principle block diagram.

RF is a classifier that contains multiple DTs as the basis, and the final prediction result depends on the votes cast by the prediction results of the DTs, with the class with the most votes as the final prediction result. The advantage of the classifier is that its results are more accurate than those of the individual classifiers. The accuracy of RF is equal to AdaBoost, but it is more robust to errors and outliers.

With the increase of the number of trees in the forest, the generalization error of the forest gradually converges. Therefore, RF can avoid the problem of overfitting. For the conversations without purchasing behavior, further bootstrap sampling is carried out, and then the class is weighted by using the idea of weighted RF. The corresponding weight item is designed for one class, and a few classes are given more weight, so that it is valued.

Specific algorithm flow is as follows:

(1) Randomly extract the same amount of data to replace it, so that the samples in the majority class and the minority class form the sample data subset D_i^*, and ensure that the data of each class are included in the subset.

(2) Design a corresponding weight item for one class and assign a larger weight to a few classes, so that they are valued and assigned the corresponding weight value $w_1, \cdots, w_c$.

(3) The division of each node in DT is determined by the weighted majority vote, and the weighted vote of a class at the end node is multiplied by the number of data contained in the class. Get another prediction function sequence $h_1^*, \cdots, w_n^*$.

(4) Repeat the above steps k times, integrate the results of two random trees, and use the prediction function to vote on the classification problem to make the final prediction.

4. Result Analysis

For the sake of generality, Iris, Credit, and Blood datasets of UCI machine learning database are selected as experimental data. In the process of discretization, the algorithm has to obtain the chi-square quantiles. When it is implemented, the

chi-square quantiles are all obtained by software R3.0.2, and the RF algorithm is implemented by using the language package random Forest4.6-6 of software R3.0.2.

The continuous variables in Iris are discretized by three algorithms, respectively, and the data reduction effects are shown in the following (Tables 1–3).

In the dataset, different reduction algorithms have good performance and achieve the purpose of discretization of continuous variables. From the reduction effect, the reduction rate of CVD algorithm is obviously higher than that of Ref [15] algorithm. Therefore, from the degree of information reduction, CVD algorithm is outstanding, which shows that CVD algorithm can reduce redundant information in datasets on a large scale.

Using the above datasets, the initial dataset and the datasets reduced by three different algorithms are used, respectively, and 200 experiments are carried out under the condition of keeping the parameters of RF algorithm consistent. The mean values of OOB (out-of-bag) estimation of RF on different datasets are compared, and the discretization performance of CVD algorithm is analyzed. The experimental data are as follows (Figure 3).

It can be seen that in the RF algorithm, the OOB estimation value of the algorithm is reduced after using CVD algorithm to reduce the dataset, which shows that the performance of the reduced dataset not only does not decrease but also improves the accuracy of RF. From the measurement standard of continuous variable discretization, CVD algorithm not only improves the dataset reduction rate but also keeps the information of the original dataset.

The construction of RF is the core of model building. The number of DT directly affects the operation speed and classification effect of RF classification algorithm, so the number of DT is very important for modeling [13]. As shown in Figure 4, with the increase of DT trees in RF, the accuracy of the model increases, but the recall rate does not change significantly.

It can be seen that, for this dataset, when the number of DT trees reaches 10, with the increase of DT trees, the classification accuracy rate of RF does not continuously increase, but the time cost will increase because of the increase of DT trees. It is ideal to select 10 DTs in the RF classification algorithm, which makes the recommendation system get accurate results quickly and meet the time limit of the recommendation system.

In this experiment, the Map_DCF algorithm proposed in this paper is compared with UCF, ICF, and Ref [13]. Figure 5 shows the change of online recommendation time of each algorithm with the number of clusters.

UCF and ICF algorithms have no clustering process, so the online recommendation duration is not affected by the change of cluster number. The online recommendation time of Map_DCF is basically below 100 s, which is much lower than that of other recommendation algorithms. As the number of clusters increases, the online recommendation time of Ref [13] decreases because the search range of the nearest users is narrowed. The online recommendation time of Map_DCF hardly changes with the number of clusters because there is no process of searching for the nearest user (or project) in online recommendation of Map_DCF, and it only needs to be recommended according to the rules obtained in advance by the RF model, so the online recommendation time is shorter than that of other models in the long run.

All features in an interval have a high correlation, and all features in the remaining intervals have a low correlation. When selecting features, on the one hand, the important features are selected according to Gini index; on the other hand, the features in two interval features are selected according to a certain proportion, which is partially random, so as to balance the strength and correlation of features. Repeat this process to build the whole forest. It can be found that, as shown in Figure 6, the effect of the model is improved.

The time of RF modeling and prediction is fast, and the time cost of feature evaluation by chi-square test is very small. Therefore, this method can ensure the running efficiency of the model. In addition, through the correlation analysis of features, the features are divided into intervals to improve the process of feature selection and get the effect of RF model optimization.

After resampling data, a lot of information will be lost. Therefore, on the basis of using the idea of balanced RF, we sample the conversations without purchasing behavior, and then use the idea of weighted RF to weight the classes, design the corresponding weight items for each class, and assign more weight to a few classes, so that they are valued. By weighting, the accuracy of the model can be guaranteed, and at the same time, the difference between DTs can be guaranteed. Combining the DT of the two to obtain the balanced weighted RF, the result is shown in Figure 7, and the result $F1$ is improved.

From the experimental results, it can be seen that after using balanced weighted RF, $F1$ index has been greatly improved, which reflects that threshold adjustment and class weighting method have obviously improved the unbalanced classification problem. Combined with the balance and weighting method of RF, the improvement of experimental results verifies the effectiveness of the improved method for unbalanced classification and also shows the importance of taking into account the retention and non-redundancy of data information and ensuring that a few classes are valued at the same time.

The difference between ethnic sports tourism and other tourism modes is that tourists can not only visit mountains and rivers, enjoy beautiful scenery, and experience traditional ethnic culture but also participate in them and feel the charm of sports. Scientifically set up ethnic sports activities that can enable tourists to participate in them, highlight the experience-oriented tourism mode, and form a cultural tourism scenic spot with "ethnic culture as its connotation and leisure experience as its feature." Only by balancing viewing and participation in this way can tourists feel happy physically and mentally when they play in the scenic spot and achieve the sublimation of their bodies and souls. This makes ethnic sports tourism a unique cultural and tourism approach with market competitiveness among many tourism resources.

Table 1: The algorithm reduces the effect table of Iris dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	33	40	23	—
Ref [15]	30	40	23	4.86
Ref [17]	20	35	18	17.62
Algorithm in this paper	16	32	14	33.69

Table 2: The algorithm simplifies the effect table of Credit dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	30	35	34	—
Ref [15]	28	32	34	3.36
Ref [17]	28	32	31	3.41
Algorithm in this paper	24	25	26	11.67

Table 3: The algorithm simplifies the effect table of Blood dataset.

Algorithm	Variable 1	Variable 2	Variable 3	Reduction rate (%)
Initial dataset	68	91	26	—
Ref [15]	17	13	7	91.33
Ref [17]	17	13	7	90.24
Algorithm in this paper	15	5	3	96.28

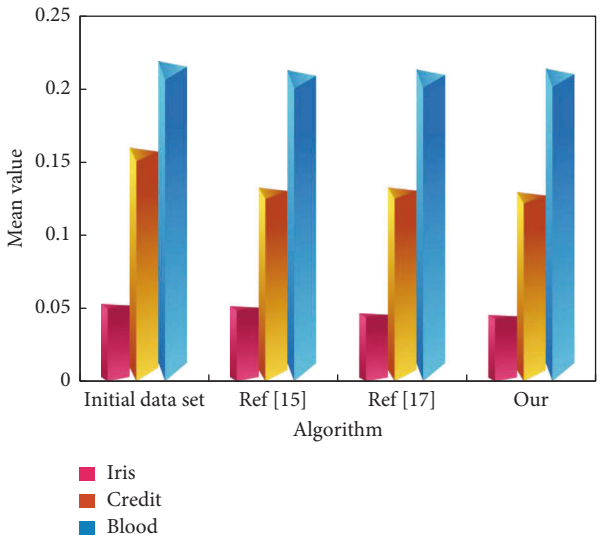

Figure 3: Use of the average value of OOB estimation after processing various datasets by RF.

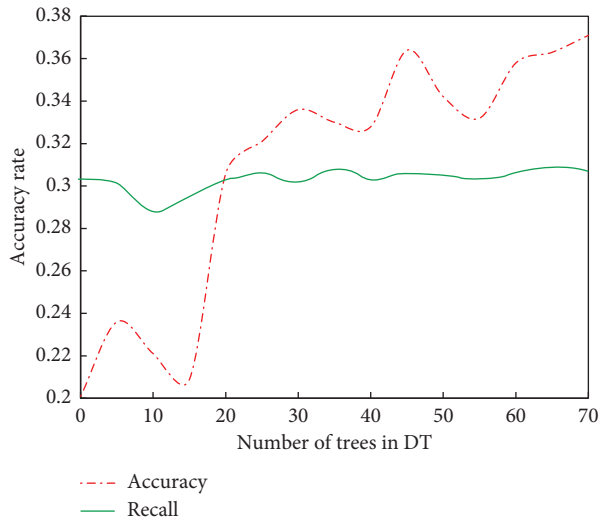

Figure 4: Influence of DT tree number on accuracy and recall rate.

Nowadays, the mainstream direction of the development of the world tourism market is from single sightseeing to diversified and thematic tourism. When developing ethnic traditional sports tourism projects, the relevant functional departments should keep the original ethnic style, dig deep into the connotation of ethnic traditional sports, gradually transform into modern people's tourism cultural concepts, aesthetic characteristics, and value orientation, develop towards the trend of international sports tourism integration, and speed up its socialization and industrialization development process in combination with the cultural characteristics of ethnic areas.

To meet the needs of tourism development and ethnic minority sports work, we should focus on building three teams: first, professional teams who know sports and are familiar with ethnic customs and culture, so as to prepare for the deep excavation of STR development of ethnic minorities. Second, it will manage a pioneering management team, promote the combination of ethnic minority sports resources and tourism, and plan excellent projects that can enter the market and manage related industries. Third, the all-round team of waiters, in the service place, reflects the quality level of service and leaves a good impression on tourists.

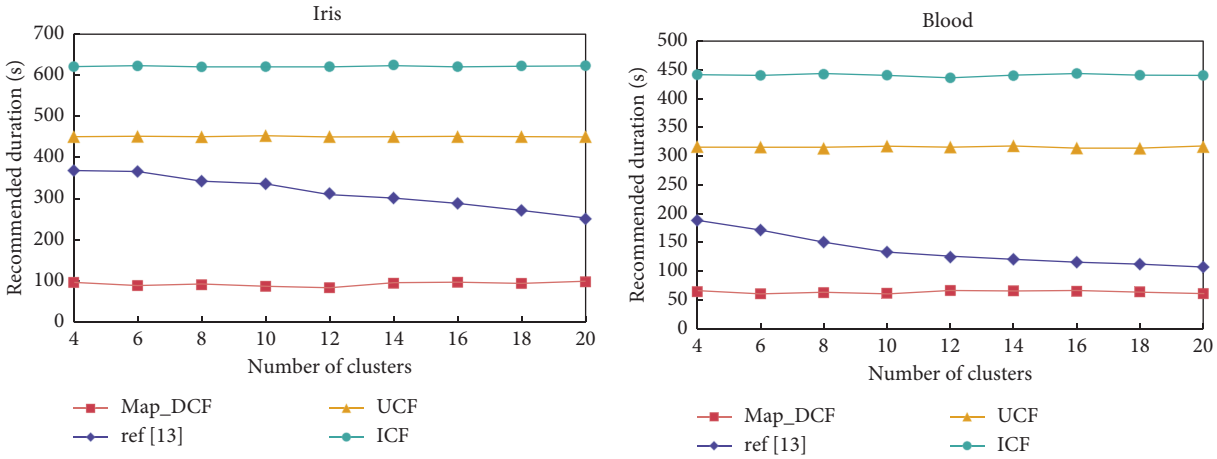

Figure 5: Comparison of recommended online time.

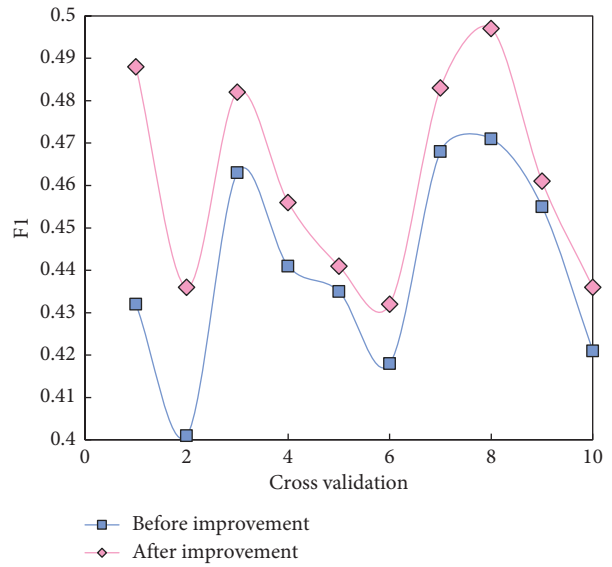

Figure 6: Comparison of results before and after feature selection and improvement.

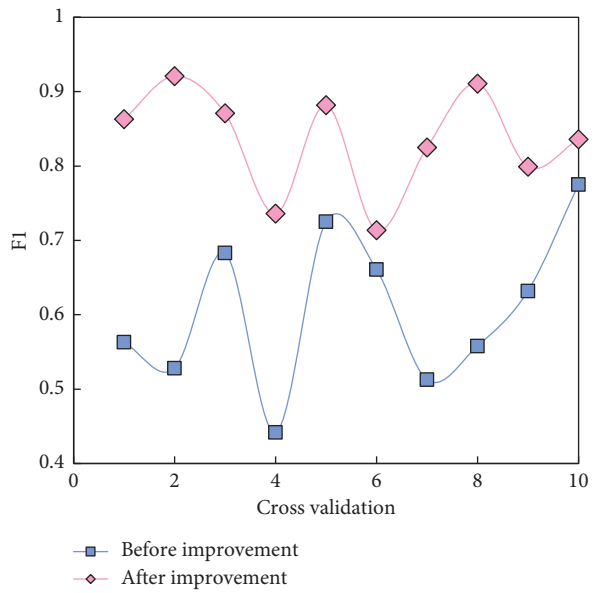

Figure 7: Comparison of effect before and after balanced RF.

5. Conclusion

The key to the development of ethnic traditional sports lies in the formation of a healthy cultural atmosphere, the establishment of sustainable development strategy, and the construction of internal development motivation. The integrated development of ethnic sports tourism not only enriches the tourism product system and promotes the transformation and upgrading of tourism but also revitalizes various resources and promotes the quality and efficiency of new industries [22]. Combining the characteristics of tourism services, this paper puts forward a multi-objective tourism service composition method. Finally, experiments show that our proposed method has high accuracy and efficiency in the multi-objective optimization of tourism service composition. Aiming at the defects of today's tourism recommendation system, an ethnic sports tourism information feedback recommendation algorithm based on RF is proposed. The algorithm reduces the dimensions of user and item vectors through clustering process, and the system only needs to make score prediction according to the RF model constructed when offline, without looking for the nearest user or item, thus greatly improving the recommendation efficiency. Through the set of balanced RF and weighted RF, the data information defect and redundancy caused by resampling can be solved, and all rare data can be selected. At the same time, the accuracy and difference of the model can be ensured by weighting.

Under the background of big data, it is a hot topic in the field of RF algorithm optimization in the future. How to realize distributed and concurrent computing of RF algorithm will be an important research direction.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This research was supported by Ministry of Education Humanities and Social Sciences Research Youth Fund Project (21XJC890002) and Guangxi Philosophical and Social Science Planning Research Topics (18CTY001).

References

[1] J. Pu and Y. Meng, "Application of visual design in the development of marine tourist resources," *Journal of Coastal Research*, vol. 98, no. sp1, p. 125, 2019.

[2] J. M. Sánchez-Martín, M. Sánchez-Rivero, and J. I. Rengifo-Gallego, "Water as a tourist resource in extremadura: assessment of its attraction capacity and approximation to the tourist profile," *Sustainability*, vol. 12, no. 4, p. 1659, 2020.

[3] W. Huang and C. S. Huang, "Discussion on engineering geological conditions and geological tourism advantage resources of coastal landscape highway in guangdong province," *Ground Water*, vol. 8, no. 10, p. 06, 2019.

[4] A. R. Szromek and K. Wybrańczyk, "Proposal of value for customer of spas: expectations of spa patients and tourist in polish spas," *Sustainability*, vol. 11, no. 13, p. 3598, 2019.

[5] H. Chen, S. Wang, H. Guo et al., "Study of marine debris around a tourist city in East China: implication for waste management," *Science of the Total Environment*, vol. 676, no. 1, pp. 278–289, 2019.

[6] Y. C. Lee and P. T. Liao, "The effect of tourism on teleconnected ecosystem services and urban sustainability: an emergy approach," *Ecological Modelling*, vol. 439, no. 1, Article ID 109343, 2021.

[7] J. Hao, H. Liu, J. Chen, and F. Sun, "Study on city-level optimization of tourism industry spatial organization nodes and organization mode for tourist destinations," *Complexity*, vol. 2020, no. 12, pp. 1–13, 2020.

[8] M. U. Khandaker, K. Asaduzzaman, A. F. B. Sulaiman, D. Bradley, and M. O. Isinkaye, "Elevated concentrations of naturally occurring radionuclides in heavy mineral-rich beach sands of Langkawi Island, Malaysia," *Marine Pollution Bulletin*, vol. 127, no. 7, pp. 654–663, 2018.

[9] G. M. Danilova-Volkovskaya, L. N. Koval, A. P. Babchenko, and N. Volkovskiy, "The tourist directions resources and sustainable development competitive advantage marketing model in the north caucasus," *Materials Science Forum*, vol. 931, no. 2, pp. 1182–1186, 2018.

[10] I. E. Livieris, E. G. Pintelas, T. Kotsilieris, S. Stavroyiannis, and P. Pintelas, "Weight-constrained neural networks in forecasting tourist volumes: a case study," *Electronics Times*, vol. 8, no. 9, p. 1005, 2019.

[11] W. Gao, Q. Zhang, Z. Lu, D. Wu, and X. Du, "Modelling and application of fuzzy adaptive minimum spanning tree in tourism agglomeration area division," *Knowledge-Based Systems*, vol. 143, no. 1, pp. 317–326, 2018.

[12] G. Mallinis, I. Chrysafis, G. Korakis, E. Pana, and A. P. Kyriazopoulos, "A random forest modelling procedure for a multi-sensor assessment of tree species diversity," *Remote Sensing*, vol. 12, no. 7, p. 1210, 2020.

[13] R. Ma, J. Ban, Q. Wang et al., "Random forest model based fine scale spatiotemporal O3 trends in the beijing-tianjin-hebei region in China, 2010 to 2017," *Environmental Pollution (Amsterdam, Netherlands)*, vol. 276, no. 10159, Article ID 116635.

[14] S. Liu, D. Mao, T. Xue et al., "A data-driven approach for online inter-area oscillatory stability assessment of power systems based on random bits forest considering feature redundancy," *Energies*, vol. 14, no. 6, p. 1641, 2021.

[15] A. Ghorbanian, S. Zaghian, R. M. Asiyabi, M. Amani, A. Mohammadzadeh, and S. Jamali, "Mangrove ecosystem mapping using sentinel-1 and sentinel-2 satellite images and random forest algorithm in google earth engine," *Remote Sensing*, vol. 13, no. 13, p. 2565, 2021.

[16] X. Zhang, J. Xu, Y. Chen, K. Xu, and D. Wang, "Coastal wetland classification with GF-3 polarimetric SAR imagery by using object-oriented random forest algorithm," *Sensors*, vol. 21, no. 10, p. 3395, 2021.

[17] S. Bao, H. Y. Pan, W. Zheng et al., "Multicenter analysis and a rapid screening model to predict early novel coronavirus pneumonia using a random forest algorithm," *Medicine*, vol. 100, no. 24, Article ID 26279, 2021.

[18] C. Zhang, J. Ren, F. Liu, X. Li, and S. Liu, "Three-way selection random forest algorithm based on decision boundary entropy," *Applied Intelligence*, no. 2, pp. 1–14, 2022.

[19] S. He, W. Chen, H. Liu et al., "Gene pathogenicity prediction of Mendelian diseases via the random forest algorithm," *Human Genetics*, vol. 138, no. 6, pp. 673–679, 2019.

[20] J. Sun, S. Gao, H. Dai, J. Cheng, M. Zhou, and J. Wang, "Bi-Objective elite differential evolution algorithm for multi-valued logic networks," *IEEE Transactions on Cybernetics*, vol. 50, no. 1, pp. 233–246, 2020.

[21] R. Faccioli, I. Silva, and L. O. Bortot, "A mono-objective evolutionary algorithm for Protein Structure Prediction in structural and energetic contexts," *IEEE Congress on Evolutionary Computation. IEEE*, vol. 13, no. 61, pp. 3–9, 2012.

[22] C. Nie, "Study on the strategy of developing ethnic sports tourism in Guangxi under the perspective of "one belt and one road"," *Journal of Guangxi College of Education*, no. 06, pp. 88–94, 2020.