EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Development Model of Agriculture + Travel Industry Integration in the Context of Big Data Explore: A case study from Huyi District, Xi'an

Pei Chao Wang[1,*], Xiang Ying Kou[1] and Yuan Zheng[1]

[1] School of Humanities and Management, Xi'an Traffic Engineering Institute, Xi'an 710000, China

Abstract

INTRODUCTION: The new industrial model of agriculture + tourism has been developed for quite some time, however, in the rapid development of information technology, especially the algorithm is further integrated into the agriculture and tourism industry, this fusion industry has ushered in a new round of development opportunities, but with the development of human society, the traditional model of agriculture and tourism will be gradually eliminated.
OBJECTIVES: This paper is aimed at developing the regional needs of agriculture + tourism industry, using advanced big data technology and algorithmic technology to follow the pace of the times, in-depth understanding of the current social needs of agriculture + tourism, so as to better develop their own industries.
METHODS:Through the algorithmic technology to analyze the agro-tourism model that is currently being developed in Xi'an, to analyze the problems that arise in the process of its development, and to use the background of big data and clustering algorithmic technology to put forward the corresponding targeted improvement strategies.
RESULTS: Utilizing Shuangyi District in Xi'an City as a case study to apply the theory and explore new development paths.
CONCLUSION: Shuangyi District, Xi'an City, is rich in soil and water resources, so it has a high level of agricultural development and a favorable geographic location, and also has a huge potential market in tourism. With the support of big data technology, the analysis of the current market demand and the development of local natural and human resources on the basis of maximizing the preservation of the original ecology can promote the development of the local economy.

Keywords: agritourism, big data technology, development model, industrial integration.

Received on 12 February 2023, accepted on 7 August 2023, published on 16 August 2023

doi: 10.4108/eetsis.3749

*Corresponding author. Email: 294580039@qq.com

1. Introduction

With the rapid development of information technology, along with the rise of "Internet +", big data technology has been applied to more and more industries[1]. It is also due to the application of big data technology, contemporary information and data processing and processing methods have been subversive innovation. Therefore, how to make use of big data technology, so that China's agriculture and tourism industries have intersection and can be integrated with each other, in these two different industries through the analysis of big data technology to find the part that can be integrated development, mutual benefit and reciprocity, so that the agriculture and tourism industry can achieve in-depth industrial integration, has become the main focus of the construction of China's characteristic socialism[2]. At present, China is in an important node of economic transformation, in guaranteeing the smooth development

of the primary industry, that is, agriculture, at the same time, we should also pay attention to the proportion of the development of the tertiary industry, therefore, how to connect the agriculture and tourism industry, in order to ensure the development of agriculture at the same time also be able to improve the economic efficiency of the tertiary industry and the proportion of the main issues that need to be explored at present. Therefore, this paper will take the background of the development of big data technology as the premise of industrial integration, so as to explore the development mode of agriculture and travel industry in this context can better synergize the development, improve the economic efficiency and contribute to the coordinated development of the overall economy.

Clustering is a valuable learning method that allows segmenting and aggregating data without labeling the learning model or previous data (e.g., data allocation) [3]. Several clustering algorithms based on different rules have been proposed, which assume a certain data distribution structure. Among them, density and network-based algorithms is an important branch that can effectively classify clusters and boundaries for different types of data distributions, such as in the fields of neural network grouping preprocessing, large-scale data compression, attribute decomposition, image recognition, and image segmentation[4]. The study of the number of physical classes and boundary distribution contributes to the development of clustering theory and practice to some extent. Density-based clustering algorithms assume that the data has density properties, where density is a linear or exponential function of the data distribution. Grid-based methods decompose data into grids that can have different shapes. In this paper, the advantages of density clustering and grid clustering are chosen to transform the similarity measurement principle into boundary-based clustering, which effectively separates and clusters boundary data, improves the feasibility and accuracy of clustering, and applies it to face recognition, rainfall distribution and image analysis[5-7]. The algorithm overcomes some of the shortcomings of existing algorithms: improves data similarity assessment for boundary class data, efficiently identifies rare multi-distributed data regions, filters data distribution attributes, and starts detailed median distributions in non-dense or low-density regions, especially in data point problems. Removing irrelevant data and setting fewer parameters based on the closest data neighbors reduces over-reliance on the actual number of categories and increases the usage of data with high similarity. Advanced boundary classification and recognition modules are developed[8]. In short, data clustering methods not only allow researchers to find hidden distribution models and functions in different data, but also compress large amounts of data, saving time and space. They can also be successfully maintained and applied in many fields, reflecting the specific benefits of data clustering. Therefore, this paper uses the clustering algorithm to analyze the agritourism model that is currently being developed in Xi'an, analyze the problems that arise during its development, and propose corresponding targeted improvement strategies using the background of big data technology.

2.Background

2.1. Identification of rural tourism and agritourism

The two concepts of rural tourism and agritourism do not have a commonly recognized concept in the academic circles at home and abroad, at the same time, due to the close connection between the countryside and agriculture, and taking into account that this paper mainly discusses agritourism may cause some conceptual ambiguities and confusions, therefore, the concept of agritourism explored in this paper is firstly clarified and differentiated from the concept of rural tourism in the public's understanding here[9]. The concept of agritourism discussed in this paper is firstly clarified and distinguished from the popular understanding of rural tourism. First of all, in terms of geographical relationship, rural tourism emphasizes the concept of countryside, that is, the place where the tourism behavior should take place should be in the countryside, so rural tourism has a stricter territoriality, and this name actually excludes tourist attractions located in the inner city and even excludes some tourist attractions located in the outskirts of the city. While agritourism has weakened the locality, because agritourism emphasizes more on the word agriculture, so the place where the tourism behavior occurs is not limited to the countryside, all the tourism behaviors involved in agriculture or located in the production chain of agricultural activities can be called agritourism[10]. Secondly, in terms of tourism behavior, rural tourism focuses more on the experience and scenery of the countryside, focusing on the characteristics of the countryside itself, while agritourism focuses on the agriculture itself, that is, to participate in the activities of agricultural production and life, from which to obtain the pleasure of leisure and entertainment so as to achieve the purpose of tourism. Rural tourism and agritourism have some overlap in some scenes and activities, but they are different according to the region and the purpose of tourism as well as the specific behavior of tourism projects. This paper will take agritourism as the main focus of discussion and research.

2.2. Growth pole theory

The central idea of the growth pole theory is that economic growth is usually achieved by a radial transmission of growth from one or more cores to other regions or industries[11]. In addition, since the concept was developed as a way to study the growth situation of countries and regions, the growth poles selected here are

generally based on the geographical aspect of growth poles.

In modern society, the growth pole theory is usually used to guide areas or regions with unbalanced or uneven levels of development, but some scholars have put forward different opinions on the theory in the process of applying the theory in practice. Some scholars believe that the growth pole theory emphasizes the so-called "growth pole" which is chosen as the core growth point, so the growth pole should be able to develop rapidly in a period of time and attract different industries to gather so as to realize the radiation function of the growth pole itself; however, some scholars believe that the selection of the growth pole may not be the highest level of development of the region, but may also be achieved through its own development[12]. However, some scholars believe that growth poles can also be selected not as the region with the highest degree of development, but also through its own development potential and other aspects to determine the growth pole. Different points of view still recognize the same fact, that is, we need to select an industry or a center as the main development goal, and on this basis, through the growth pole's own development advantages to attract other industries or radiate other regions, so as to realize the concept of growth pole theory[13].

As for the agriculture+tourism industry, the significance of the growth pole theory lies in the fact that when fostering such emerging industries, especially the mixed industries where different industries are integrated with each other, it is also possible to realize the enhancement of one's own tourism function and expand the market by finding a main development focus, such as agricultural production, natural resources, and other specific directions, and establishing the elements for the development of tourism on this basis, i.e., transportation, food and beverage, and accommodations. market, radiate to the neighboring areas and industries, and eventually produce a clustering effect, making the place a well-known agrotourism scenic spot[14].

2.3. Point-Axis Theory

The point-axis theory is similar to the growth pole theory in some of its points of view, so it can be viewed as a supplement and extension of the growth pole theory[15]. The core idea of the point-axis theory is that in the process of economic development of a region or a country, most of the social resources and various economic factors needed for development will be gathered at certain "points" in the region, and through these "points" and "points", the economic factors will be gathered at certain "points" in the region. Through the continuous development of these "points" and "points", different "points" are connected and become an economic development axis. These economic development axes are not limited to the abstract theory of economic development, but can be figuratively regarded as a variety of transportation routes, communication routes, energy transportation routes, and other economic development required for the delivery of materials connected to the line. With the growth pole theory is different, point - axis theory more clearly and specifically pointed out the direction of the radiation area after the development of different growth poles, can follow the core region of common development is mostly located in the core region of the economic factors of the main road, which is also another embodiment of the effect of industrial agglomeration. Axis theory of mixing matrix classification is shown in Figure 1.

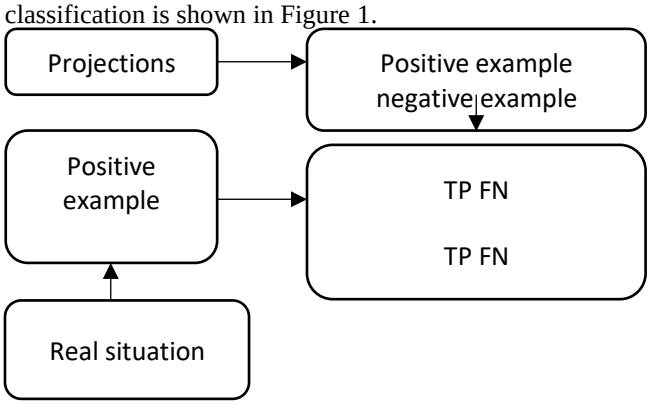

Figure. 1 Mixed matrix classification for axis theory

And in the development process of agriculture + tourism industry, this theory has more important practical significance. Due to the uniqueness of the two industries of agriculture and tourism, agriculture is usually developed as a whole, and most of the agriculture in the same region will be concentrated in the industrial line of several agricultural products, so the area where the agricultural activities take place is naturally the core area of growth. The development of agriculture usually has a very high demand for transportation, energy and other requirements, which is very similar to the requirements of the tourism industry, so agriculture + tourism usually forms the agriculture as the center of growth poles to drive the different areas along the transportation routes to become the axis of economic growth development model. Eventually, these first development centers will become the core point of local industrial development, leading to a variety of traffic arteries and energy transportation lines on the development of scenic spots, and then by the newly developed scenic spots as the core to repeat the previous development pattern, thus driving the development of the entire region.

2.4. Core - Edge Theory

The central idea of the core-periphery theory lies in the fact that there are always core areas in different countries and regions that concentrate all kinds of resources and capital, which are also known as developed areas, while the peripheral areas are relative to these core areas[16].

Core and peripheral areas have different development dynamics due to their different development speeds and different factors of production that they can absorb. However, with the passage of time, the aggregation effect of the core area is not static, but will change according to the changes of various factors of production, thus making the status of the core area and the peripheral area changes, and ultimately realize the coordinated development of different regions under the joint action of internal and external forces. Compared with the previous two theories, the core - edge theory emphasizes more on the development of different regions in the time scale, that is, relying on the core development area and the establishment of other industries may also become the core growth pole of the local area, so as to realize the synergistic development and growth of industries.

As far as the agriculture + tourism industry is concerned, since most of the agricultural industry is located in rural areas, it will naturally rely on the core growth resources of the city in the development process of agritourism, and exist as a supplement to urban tourism and entertainment. For agritourism that cannot rely on location factors such as famous mountains and rivers and historical and cultural monuments as its main selling point, the degree of development of the urban metropolitan area in the core area will directly affect the establishment of an agritourism area. At the same time, with the development of time, an agritourism area will also manage the stage from establishment to maturity, and after maturity, it will also be transformed from a fringe area to a sub-core area, which will promote the formation of the next fringe area, and ultimately lead to the diffusion of the industry.

3. Research methodology

3.1. Segmentation of the clustering algorithm

The basic idea of distribution-based clustering methods is to partition the data object into parts and minimize the error squared assigned to certain parts of the corresponding cluster centers. A typical partition-based clustering algorithm is K-Means[18]. The advantages of this algorithm are the ease of understanding and implementation of the design concept, the wide range of applications, the linear distribution of time complexity, and the efficient and scalable analysis of big data. The K-Means method decomposes the elements of the dataset n into K-clusters i.e., by selecting the centers of the K-clusters, each cluster is a cluster. After selecting the set the following method can be used:

$$\begin{cases} C_i \bigcap C_j = \varnothing, C_i \neq \varnothing, i = 1,2,...,k, \\ C = \sum_{i=k}^{k} C_i \end{cases} \quad (1)$$

Equation (1) utilizes the form of a set in order to better figure out the intersection of extensions;

$$P_{ij}^{k} = \begin{cases} \dfrac{(\tau_{il}^{\mu} / d_{ij}^{v})}{\left(\sum_{l=1}^{n} l \notin j, \tau_{il}^{\mu} / d_{ij}^{v}\right)}, j \notin l \\ 0, j \notin l \end{cases} \quad (2)$$

Clustering is used at this point for better path optimization; where the

$$\Delta \tau_{ij}^{k} = Q / L_k \quad (3)$$

The separation principle Km is as follows: firstly, the distance between each data point and the original clustering center chosen by K is calculated and the data points are assigned to the class closest to the clustering center until all data points are available. Second, the horizontal mean of all data points in each cluster is recomputed as the center of the new cluster. Finally, the method is used to iterate successively through the clustering centers until the specified object function (sum of squares of errors) reaches its minimum value. The purpose of this method is to compile the data points corresponding to the class without updating all other points in the class, maximizing the distance between classes and minimizing the distance within classes.

The disadvantages of the KM algorithm can be summarized in two words: firstly, if the data distributions are inhomogeneous and different from each other, the number of cluster centers can only be determined based on the generalized bending algorithm, bias statistics, and canopy algorithms. k. In addition, a number of methods have been developed to adjust the number of cluster centers. However, defining "k" is a problem for many researchers if the database structure is uniform or circular. The proposed principle only considers the effect of the distance factor and does not take into account the effect of other situations on the algorithm, making the algorithm highly sensitive to extreme data.

In order to reduce the influence of K and initial mean selection on the K-mean, foreign researchers proposed the K model, which uses the Euclidean K-mean measure instead of the Hamming distance and is chosen as the median value for measuring data similarity. However, the algorithm still has some shortcomings in terms of the number of clustering centers and the heterogeneity of data separation. The DH-K-Maans algorithm is proposed, which dynamically defines the initial clustering centers according to the clustering density and distance parameters and replaces the fixed data resources to avoid the local optimization problem. Meanwhile, a hierarchical clustering method is used to combine similar small clusters to accurately handle multiple datasets. However, determining the number of cluster centers for different datasets remains a challenge. To avoid defaulting the number of classes, a scalable K-means algorithm is proposed. The algorithm develops a coherence function to evaluate the similarity of different separation structures and harmonic forms of the data, breaking the data separation limit set by the Euclidean index. However, the

method is not efficiently recognized if there is a lot of overlapping and bubbly data in the data, which significantly degrades the clustering performance. The flow of the clustering algorithm is shown in Figure 2.

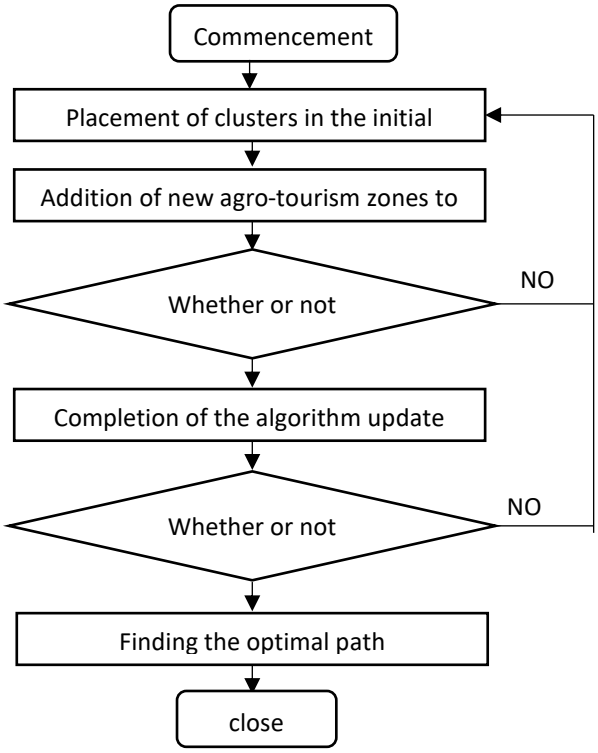

Figure. 4 Flow of clustering algorithm

3.2.Cluster density analysis

Density clustering algorithms are based on the distribution of spatial data points that represent the clustering properties and define the data density in different formats. There are several forms of Gaussian distribution, Cauchy distribution or density expressions to choose from[19]. Low density regions, in particular, can be noise points or boundaries. A schematic diagram of clustering ARI regions with different step sizes is shown in Figure 3.

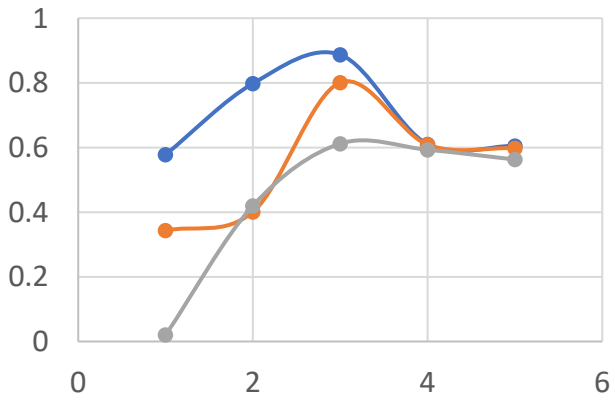

Figure. 3 Schematic diagram of clustering ARI regions with different step sizes

Unlike the traditional K-means algorithm, DBSCAN can only be used to group convex sets of samples.The DBSCAN algorithm is suitable for irregular data distributions in different formats. For noisy data, the DBSCAN algorithm can only be used to group convex samples and the grouping is also effective. The algorithm proposes the concept of density availability based on the distribution of iron wheels. Whenever two nuclear vortices encounter a part of the "biological chain", these nuclear vortices are connected along the path of the biological chain to form a category that recognizes sound or static data during the collection process. the EPS and MinPTS settings are difficult to configure and the clustering performance is low for large datasets.

Another popular density-based clustering algorithm is the advanced density clustering algorithm proposed by the scientific community. This algorithm provides a similarity measure that combines density and distance and visualizes the cluster centers by constructing a decision diagram. In addition, the design principles of DPC are easy to understand and implement, and it improves the clustering performance compared to other methods. Based on the context in which the DPC algorithm was created, the drawbacks of the DPC algorithm are summarized, including the need to manually select the percentage of crossover densities, the actual number of classes, and the similarity of distance duplicates computed due to spatio-temporal complexity. Therefore, grouping large datasets into large dimensions is not ideal. A schematic diagram of clustering ACC regions with different step sizes is shown in Figure. 4.

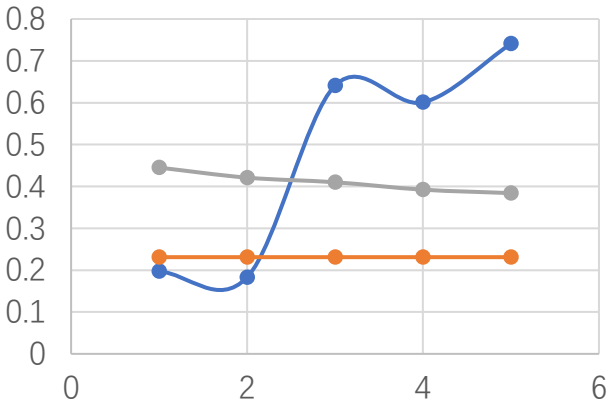

Figure. 4 Schematic diagram of clustering ACC regions with different step sizes

3.3.Clustering algorithm lattices and modeling methods

Grid-based clustering methods divide the data room into several individual grid cells without forgetting all the internal data attributes[20]. Like density-based methods,

these methods perform clustering by determining the properties of these grids (e.g., number of points in the grid, grid size). An obvious advantage of data networks is their speed and clustering efficiency. At the same time, there are some shortcomings of the network technology. Secondly, grid grouping has some drawbacks in arrogant datasets because the number of grid cells increases exponentially with the data size. Therefore, grid clustering algorithms can be combined with powerful computational algorithms and density clustering algorithms can be combined with any clustering algorithm. Most density-based clustering algorithms require the determination of distances between data to determine the density, which can complicate the algorithm time. Using network segmentation to obtain representative data can significantly reduce the amount of data, reduce unnecessary time, and improve algorithm performance. Based on analyzing these principles, many researchers have proposed an improved clustered network density algorithm. Clustering algorithm is a clustering algorithm proposed by foreign researchers that combines grid density and attributes to obtain basic data. On this basis, the subspace data will improve the computational power.GDILC algorithm is proposed by national researchers who use grid criterion to describe the distribution of data and group the data according to their distribution characteristics to improve the clustering efficiency. In recent years, some researchers have proposed an SGC algorithm that divides the data space into multiple grid cell structures and uses the concept of sliding windows to maintain the grid density and reduce the complexity of the algorithm time. Some have been working on network-based density clustering algorithms. First the density of each grid cell is calculated and then the separation density of each data point and neighboring grid cells is discussed. The magnitude of the difference determines whether the final port density is connected to the output port or not. Based on this, a DPC cluster is created. The combination of the density clustering method and the grid clustering method reduces the spatio-temporal complexity of the peak density clustering method to a certain extent, avoids the curse of dimensionality, which restricts the performance of the grid-based clustering algorithm, and improves the efficiency of the clustering algorithm.

The basic idea of model-based clustering algorithms is to classify each subcategory into a corresponding model, and then compare all the data with a hypothetical model until the model is satisfied to ensure that the data in the same category have the same probability distribution. The classical Gaussian mixture model (GMM) can be effectively used in data clustering and density estimation algorithms. The model assumes that each data type follows a Gaussian distribution and that all data types are described as a mixture of several finite Gaussian models. However, the actual data does not contain data that exactly conforms to a normal distribution, so the generalized torque algorithm is an approximate grouping method whose results may vary significantly. To address this limitation, some studies have introduced adaptive methods for GMM clustering methods that can effectively improve the performance of measurement clustering. Some studies have proposed a common EM adaptive algorithm that prevents overlapping and local convergence.

Probabilistic models can contain not only observed variables, but also hidden variables that are assumed to be absent or invisible. The EM algorithm proposed by foreign researchers is an iterative method that utilizes density estimation and data grouping to estimate the highest probability or highest inverse probability of the hidden parameters in a variable model. However, this iterative method leads to poor performance of the algorithm, especially for complex distributions and relatively small sample sizes. In some studies, the EM algorithm is used to collect locally sufficient statistical information, and locally sufficient aggregated centralized statistical information is used to generate complete global statistical information until the algorithm converges. This algorithm not only decomposes the work of the EM algorithm and reduces the number of iterations, but also improves the convergence speed of the EM algorithm.

3.4.Hierarchy and groupwise optimization

Level groups collect data at the same level in different ways. First, each data point is connected to the boundary from the bottom up according to the definition of the similarity metric; second, downlink similarity is enforced throughout the data segment until certain conditions are met. Hierarchical clustering methods are based on similarity matrices between data points and support the ability to analyze adjacent data structures to create tree diagrams and combine or separate data pairs to create hierarchical clustering methods. The performance of the algorithm is significantly degraded if the data contains different classes or a few constants and is sensitive to noise distributions and limit values. The processing algorithm is one of the representative algorithms in hierarchical clustering. The method solves the problem of significant changes in the size of clusters and subcategories of aspherical data. The method uses a strategy based on the center of mass and representative elements of representative categories to select multiple representative points from the data space instead of selecting the center of mass or elements in representative categories. On this basis, stratified groupings are made based on various shrinkage factors until the number of natural classes is reached. A national study proposed a birch tree algorithm that uses a tree structure to create master classes and then calculates the similarity of these subcategories and hierarchical clusters. The chameleon algorithm proposed in a foreign study generates the underlying distribution of nearest neighbor data and groups the data based on the similarity of the regions.The LCCV algorithm determines the density of the data points by releasing the nearest neighbors and selects the point

with the highest density of data as the representative point.

Classical clustering optimization algorithms include flocking optimization, particle clustering optimization, saloon algorithm, fruit fly algorithm and so on. In response to the limitations of cluster intelligence algorithms in terms of local search capability, search results and convergence speed, many effective improvement schemes have been proposed to balance the global and local search capabilities. An SVD particle cluster optimization algorithm for text processing is proposed. The algorithm optimizes the lens function from both global and local perspectives, generates the output randomly, and then performs the SVD.Some foreign research institutes have proposed a method to determine the data similarity using higher correlation instead of linear measure. This method introduces three local search methods in the artificial parval algorithm, studies the data state, and promotes the further development of the algorithm. In recent years, a finite artificial bee colony algorithm has also been proposed abroad. The algorithm improves the search ability by introducing three new search equations. In view of the effective optimization of the clustering intelligent algorithm, many researchers have applied it to the clustering algorithm to reduce the influence of input parameters on the clustering results, improve the self-regulation ability of the algorithm, and improve the accuracy and generality of the clustering algorithm. The algorithm adopts the Chinese K-DBSCAN algorithm, which combines the harmony search algorithm with the DBSCAN method, and uses the harmony search algorithm to optimize the data points in different environments. The efficiency function of the algorithm consists of two parts, one part aims to minimize the number of predicted and actual classes, and the other part aims to maximize the minimum number of clusters. Both methods are effective in determining the number of target classes and improving the clustering accuracy. To address the shortcomings of the DBSCAN algorithm, a new clustering classification algorithm based on particle clustering density optimization was proposed, and the optimal parameters for clustering classification were determined. In recent years, local researchers have proposed a two-stage clustering algorithm that uses an improved DBSCAN algorithm based on the BAT optimization algorithm to create initial clusters. The improved DBSCAN algorithm uses profiles as adaptive functions to control the parameter optimization process. Cluster centers are automatically defined on the basis of the initial categories and then the second stage of clustering is performed using the peak density clustering method. To address the shortcomings of the particle day optimization algorithm, a multidimensional gravitational learning factor is introduced which uses a Gaussian kernel function to determine the number of clustering centers. This study combines the particle pair algorithm with clustering to improve the coordination and development of the algorithm.

4. Results and discussion

4.1. Analysis of the main tourism industry model of agriculture in Xi'an City

Located in the Guanzhong Plain area, Xi'an City is rich in natural resources and deep in history and culture, with abundant tourism resources, and at the same time, it has laid a good foundation for the development of the agro-tourism industry. At present, relying on Xi'an's good economic development as well as its profound agricultural industry foundation, Xi'an has basically formed an agricultural industry belt with its own characteristics. In addition, in the long-term development process, the agritourism industry in Xi'an also takes farmhouse, leisure farm and leisure agriculture park as the main industry mode, while field complex and leisure agriculture town have emerged as new forms of agritourism in recent years.

4.1.1.Hierarchy and groupwise optimization

The following table shows the main types and percentage of leisure agritourism industry models in Xi'an at present. As shown in Figure 5.

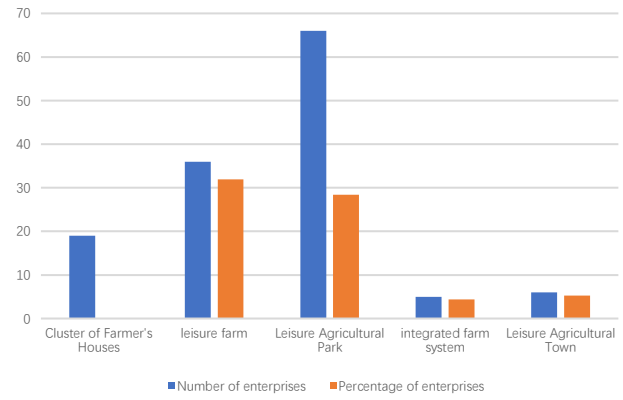

Figure 7 Distribution of agritourism industry model types in Xi'an City

As the current mode of agritourism is more, this paper will mainly divide agritourism into the above categories and analyze the status quo according to the specific different industrial development modes.

(1) Agri-house model

As the mode of Nongjiale is more special, it is not convenient to count by enterprises, so the number of Nongjiale is counted by Nongjiale clusters, up to now Nongjiale around Xi'an city has formed 19 clusters, which include the participation of 1102 households. And the cluster phenomenon of these Nongjiale is also very obvious, mainly concentrated in Chang'an District, Shuangyi District and Lantian County of Xi'an. At present, the agro-tourism industry development mode of Nongjiale has formed a more mature industrial chain due

to its early and long development time. Compared with other types of industrial modes, Nongjiale has the most diverse functions and can provide the most types of services, and thus has become the preferred industrial mode of the agro-tourism industry. Nongjiale, with its unique way of scattering clusters, has also formed different characteristic services in different gathering areas. For example, the Nongjiale cluster in Shangwang Village offers rural private kitchens as a selling point, while the Nongjiale in Shuangyi District mainly focuses on various cultural heritages, as well as various conference venues and so on. Thus, it can be seen that different Nongjiale clusters are still formed mainly on the basis of local characteristic production resources, which has become the most basic condition for the development and growth of Nongjiale clusters.

(2) Leisure farm model

Compared with Nongjiale mode, leisure farm mode is closer to the current sense of agriculture + tourism integration industry development mode, these farms usually rely on the development of the core characteristics of the resources and advantages of the agricultural industry, according to the different areas of the local conditions to carry out the corresponding development of the farm industry. Most of the leisure farms in Xi'an are located in Chang'an, Shuangyi, Lintong, etc., with strawberries and grapes as the specialty crops in Chang'an, common in Shuangyi, and pomegranates and all kinds of vegetables in Lintong. These areas are able to develop the agro-tourism mode of leisure farms, on the one hand, because their own agricultural industry is more mature and has its own characteristics of natural resources, and on the other hand, another important factor is the proximity to the urban areas, which has a broader potential market and development potential. In addition, leisure farms have a more advanced business philosophy compared to the farmhouse, so the development of a series of such as personal picking and harvesting, farming experience, tasting special food and other local agricultural industry has a strong relevance to the tourism and entertainment projects, which is the leisure farm model to attract a lot of tourists to come to experience the play.

(3) Leisure Agricultural Park Model

The leisure agriculture park model is more large-scale and industrialized, and offers more ways of experiencing crops than simply participating in the process of harvesting and tasting them. Since the development of agricultural parks requires more production factors, most of the leisure agricultural parks are located in Chang'an, Gaoling and Lantian, which have their own industrial lines. The leisure agriculture parks in Chang'an District have built a number of leisure agriculture parks integrating leisure tourism visits and agricultural products production demonstrations by demonstrating the relationship between high technology and agricultural production; while Gaoling District has established a leisure park along the Weihe River relying on the local characteristics of the natural landscape, which not only provides natural scenery sightseeing, but also attracts many people with such agricultural tourism and sightseeing through the development of water surface agriculture and underwater agriculture. This leisure agriculture park not only provides natural scenery sightseeing function, but also attracts many tourists with this kind of agritourism sightseeing experience, and even develops the new agritourism mode of adopting happy farms with individual tourists as the main unit on this basis. It can be seen through the development direction of different leisure agriculture parks that leisure agriculture parks emphasize more on the contribution of local economic and technological capacity in the development of agro-tourism projects, and most of these leisure agriculture parks need to rely on mature agricultural production lines as well as mature tourism business models in order to be able to develop and grow.

(4) Rural Complex Model

Xi'an has the most concentrated and mature development of idyllic complexes in Lantian, Gaoling and Zhouzhi. Among these three regions, Zhouzhi has the most mature development of idyllic complexes, and part of the industry has even gradually radiated to other regions, promoting the development of neighboring regions. Compared with the industrial model of farmhouse, leisure farm and leisure agricultural park, the proportion of agricultural elements in the idyllic complex is relatively small, and this development model emphasizes more on agriculture as the basis, and vigorously develops the mode of one-stop tourism, which has different main bodies in different areas, but all of them use their respective agricultural industries as a component of tourism, rather than the main form of participation. Among them, the complexes in Gaoling District have been combined with a variety of functions such as farm sharing, business incubation, cultural and creative industries, tourist resort hotels, crop production landscapes, and commercial cuisine. As seen in the development of GaoLing District and other idyllic complexes, this industrial model as agriculture + tourism weakens the dominant position of agriculture and emphasizes more on the dominant position of tourism.

(5) Leisure Agricultural Town Model

The agriculture+tourism industry model of leisure agriculture town is the industry model with the highest degree of maturity and completion among the various agricultural tourism industry models at present. The formation and development of this industry requires a strong economic foundation as well as a potential market as an important industrial support, so this industrial model is mainly distributed in Chang'an, Lantian and other regions. The development idea of these areas is to take the local agricultural industry as the theme of building a leisure agriculture town, and according to this theme, to carry out the construction of other kinds of industries, and introduce the elements of tourism, so as to realize the integrated development of the two industries. For example, the flower industry characteristic town in Chang'an District is mainly constructed with flowers as

the main landscape and theme, and other commercial complex functions are developed on this basis to finally realize the construction of the whole characteristic town. These leisure and agricultural towns usually require high level planning and careful design, as well as appropriate matching of tourism experience programs, in order to ensure the unity of the entire leisure and agricultural town. Therefore, this type of industrial model is also the highest threshold of development and the highest investment in industrial development mode.

4.2. Problems currently facing agriculture + travel

4.2.1.Simple production and business model

At present, the main development mode of agrotourism in Xi'an is still the primary stage development mode such as farmhouse, so most of the agrotourism industry has the situation that the small volume of agriculture leads to the small volume of tourism. In addition, due to the geographical limitation of the market, the main development areas of agritourism are still gathered in the suburbs of Xi'an, and most of these places lack of arable land and technology, and the agricultural industry itself has not formed a complete industrial chain, mainly based on the local villagers' scattered planting. Most of the agricultural products in the agrotourism areas have almost no deep processing process, which leads to the low added value of agricultural products and the overall competitiveness is also weak.

There are no outstanding features in agriculture, which makes it impossible to develop the local tertiary industry. Most villages do not have a specific positioning of agricultural tourism, which leads to a high rate of duplication with other regions, no unique attraction, unable to meet the needs of tourists of various types and levels of tourism, and thus a lack of competitiveness of tourism, which ultimately leads to a shorter duration of stay of tourists, and is unable to promote the development of other industries, and is unable to form a larger-scale industrial agglomeration.

4.2.2.Simple production and business model

Through the survey of areas developing agritourism, it was found that the lack of facilities in Xi'an and the imbalance in the development of various types of supporting facilities are mainly reflected in three aspects. The first aspect is the different degree of development of different types of supporting facilities. Relying on the location of Xi'an, the level of infrastructure in most of the areas developing agritourism is basically higher than the average, but the level of construction of other public service facilities is generally on the low side, and most of the areas are lacking in supporting facilities related to agriculture+tourism, which leads to poor experience of all kinds of basic services needed by tourists in tourism, and ultimately leads to the slow pace of development of the local agritourism industry. The second aspect lies in the low level of basic public services in the industrial areas where agritourism is developed, especially in education, healthcare and pension, which lag far behind the urban areas in terms of development. The most direct impact of such public services is the loss of local personnel, so there is a lack of human resources needed in the development of agritourism. In addition, the lack of human resources will also inhibit the level of development of agritourism, resulting in the majority of agritourism can only be based on a lower level of agritourism industry model. The third aspect lies in the fact that apart from the basic public service facilities, other public service facilities about agriculture and tourism are even smaller in scale, which not only do not dominate in quantity, but also fail to gain the recognition of tourists in terms of quality. In addition, since most of the public facilities for tourism are constructed in a piecemeal manner, they are usually not integrated with the basic public service facilities, and there is no unity or connection with each other.

The lack of most of these facilities, or the different levels and degrees of development of different infrastructures, are the most important constraints to attracting more tourists to the area. However, in addition to these obvious constraints, there are some other aspects that have an impact on the overall development progress but need to be solved in the long run, which have also been neglected. For example, most of the agrotourism zones are not well developed in terms of internet, so tourists either don't know the local situation or don't come back or don't recommend it to others after traveling in the area due to the internet or signal problems, which will lead to the decline of the attractiveness of the area. In addition, most of the agrotourism areas do not have enough funds and manpower to support the local hygiene and long-term maintenance of equipment, which affects the perception of a significant portion of tourists who come to visit the area.

4.2.3.Agro-tourism characteristics are not prominent

Most of the current agrotourism adopts the same industrial model, as mentioned before, there are five main types of agrotourism in Xi'an, but the content of tourism in most cases of these five types of industry is almost the same, and the homogenization is serious. Among them, the cluster industry of agro-tourism is the worst-hit area. Since most of the villagers who develop agro-tourism don't know the industrial knowledge of agriculture + tourism, they mostly borrow the contents of other agro-tourism in the process of developing their own agro-tourism. And due to the influence of infrastructure level,

most of the contents that can be provided by agro-parlors are at the primary level, with less contents and lower sense of participation of tourists, which affects the actual experience of playing and fails to form their own special competitiveness.

For agrotourism developed in part on the basis of the cultural heritage of the respective regions, most of the regions lacked the same planning and acted in a fragmented manner, thus providing similar tourism contents and seriously homogenizing the tourist activities. In addition, the process of modernization will affect these primary agrotourism areas, and the way of life and production in different areas, especially in the rural areas, will be greatly affected, so that practitioners of agrotourism in some areas may choose to give up the industry and turn to other industries if they cannot attract more tourists, thus creating a vicious circle above the tourism industry.

4.3. Targeted optimization strategies using big data technologies

Through the analysis of the problems arising in the process of developing the agriculture + tourism integration industry, the following three aspects will be put forward to help optimize the targeted strategies. In this process, a large number of optimization strategies are based on an in-depth understanding of the local and its own potential market for detailed research, so the technology of big data is particularly important. Through the use of modern information technology, it is conducive to understanding as soon as possible and formulate appropriate enhancement strategies.

4.3.1. Industrial Economic Development Strategy

(1) Development ideas

The industrial integration development mode of agriculture + tourism is ultimately to realize the villagers' income generation, so as to promote the development of industrial economy. First of all, sort out the local special tourism resources, and other areas to form a differentiated competition, and then the local basic public service resources and tourism service resources to make up for the shortcomings and enhance the development of the agro-tourism industry as a driving force to drive the development of other industries, and ultimately to achieve the integrated development of the industry.

(2) Development planning

Local special tourism or agricultural resources need to be identified to create a special local tourist spot. After identifying the local special tourism resources, the industry needs to be positioned according to the specific content. That is to say, natural resources-based areas need to maximize the restoration of local natural features while promoting the development of supporting peripheral industries, while human resources can be integrated into other local industries as cultural and creative products of the local tourism industry.

The development mode of the industry is determined through local agritourism resources, and is generally categorized into the village collective-driven mode, the capital prying mode and the combination of the two. These need to be decided according to the specific local conditions, and the generally recommended choice is the third one, which has become the preferred choice in most areas because most agro-tourism areas do not have enough financial resources for development, but need to maintain their own style on the basis of commercialization.

4.3.2. Public Facilities Layout Strategy

Due to the good overall development of Xi'an, the infrastructure of its neighboring areas in terms of agro-tourism is relatively well-developed, but the overall problem lies in the unbalanced and unsupported development of infrastructure and other public service facilities, so it is necessary to focus on the construction and layout of public service facilities.

(1) Configuration ideas

Improvement of public service facilities need to take into account the regional location, topography and other natural factors, but also need to take into account the region's development potential, population composition and the status quo of public service facilities, etc., the use of big data technology to help form a set of more complete and comprehensive and targeted public facilities configuration strategy, so as to realize the fair allocation of public service facilities and other basic service facilities. And through the configuration of network and other information technology service equipment, to achieve the modernization of various types of public service facilities.

(2) Construction content

The most important thing is the improvement of various infrastructures, such as transportation, water, electricity, communication, location change, irrigation channels, etc., as well as the advance planning of future development, so as to avoid the problem of insufficient supply of various resources. According to the big data technology to analyze the demographic composition of tourists and the potential market and all kinds of demand preferences, and use this to adjust the number of different types of public service facilities, content, etc., so as to more targeted configuration of local public service facilities. Especially in areas with a high number of special populations, such as the elderly and children, more attention should be paid to the needs of these two groups. In order to control development costs and subsequent maintenance costs, public service facilities should be efficiently utilized, and all kinds of public service facilities and tourism facilities should be located in places where the demand of the population is high or where

transportation is convenient. Some of the facilities for activities can be decentralized to cover more service space.

4.3.3. Strategies for rural landscape construction

(1) Construction Ideas

The construction of rural landscape generally needs to consider both natural and humanistic landscapes, and also consider the overall architectural style of the local community with the natural and humanistic elements. As most of the agrotourism areas are also used as local people's living areas, it is necessary to integrate the villagers' daily life with the special humanistic and natural elements of the local area through the overall integrated planning, so as to achieve the final harmonization effect.

(2) Construction content

Maintaining, repairing and remodeling existing buildings, especially those with a history, especially traditional houses and ancestral halls of various kinds with rich humanistic value, and choosing the corresponding means of refurbishment according to actual needs. For some of the old and worn-out buildings, they should be restored as much as possible according to the results of field investigation and the impression of local people, so as to maintain the original appearance and continue the role of history. The original local production and living buildings should be renovated, and on the basis of not changing their original functions, through the renovation of the appearance and structure, the local regional characteristics should be added, such as the color of the façade, the overall shape and structure of the house, etc., all of which reflect the local characteristics. Some buildings that have been abandoned for a long time and have no humanistic value are demolished or transformed to improve the utilization rate of local limited space and land, and at the same time reduce the overall development cost, and finally realize the rational use of functions through the replacement of different functions.

4. Specific case study of Shuangyi District

4.4.1. Analysis of the current situation in Shaanxi District

Shuangyi District of Xi'an is located in the southwest of Xi'an, with the Qinling Mountains in the south and the Wei River in the north, with superior natural landscape conditions, among which Taipingyu Area is located in the mountainous area at the northern foot of the Qinling Mountains, 35 kilometers away from the center of Xi'an in a straight line, with great potential for tourists in the tourism industry. Meanwhile, the surrounding transportation is well developed, with new and old mountain roads and highways interspersed, the basic conditions are superior. At present, most of the planned land for agro-tourism in Shuangyi District is still dominated by agricultural production land, most of which is farmland, with villages banded along the old mountain-ring road. In addition, the establishment of some science and technology industrial bases in the area has led to an accelerated pace of infrastructure development in the surrounding area. There is a small amount of green space in the area of the district of Shaanxi, and there is a lack of major landscape plants at present, but most of the major green plant landscapes are economic landscapes such as economic forests, orchards, and farmlands, which is conducive to the development of agro-tourism projects. In addition, there are important various cultural and historical sites around the farmland such as Daewonsa Temple, Song Village Ruins and Taepyeong Palace Ruins. Natural resources and tourism resources are relatively rich. The city of Xi'an has also issued a corresponding "Planning and Protection" program for Shanyi District, which clearly points out that there are 55 projects that need to be protected, so there are abundant tourism resources available for development.

4.4.2. Industrial integration of agriculture + travel under big data technology

Utilizing the ability of big data technology to integrate and process various types of information, as well as through research and studies, this paper has compiled the following table to show the different levels of factors that are currently affecting the development of the agriculture + travel fusion industry. As shown in Figure 6.

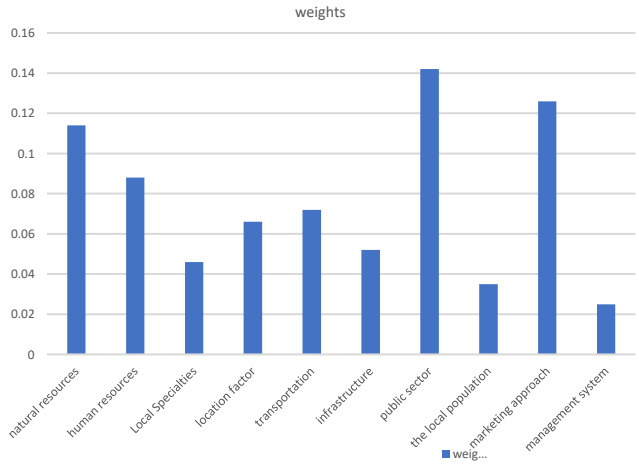

Figure. 8 Impact weight values of agriculture + travel industry factors

As can be seen from the data in the table, the most important factors affecting the development of agriculture + travel integration industry in Shuangyi District are mainly centered on natural resources, public sector participation and marketing methods. It can be seen that these three factors mainly represent the local conditions, development capacity and public awareness, which are

also the three most important factors affecting the development of agriculture industry and tourism industry. Therefore, before realizing the integration of agriculture + travel industry, the primary character is still to strive to improve the basic capacity of agriculture and tourism infrastructure construction, on the basis of a certain acceptance of the capacity of agriculture and tourism through a variety of marketing methods to expand the visibility, and at the same time, the use of big data technology for the accurate placement of advertising marketing, to maximize the benefits, so as to promote the integration of local agriculture + travel industry development, and ultimately promote the common development of related industries and drive the development of the local agriculture + travel industry, which is the most important factor. Ultimately, it will promote the common development of related industries and drive radiation to the surrounding areas.

5. Conclusion

The agriculture + travel integration industry has been developed for a certain number of years, and some regions have also taken shape and accumulated certain experience, so by studying and summarizing the experience of mature regions, and analyzing and summarizing the factors affecting the development of the industry using big data technology, it is conducive to deepening the understanding of the industry, which can assist in making the strategy more suitable for the economic development of different regions. Shuangyi District, Xi'an City, has rich soil and water resources, so the level of agricultural development is high, and the geographical location is favorable, and there is a huge potential market in tourism. With the support of big data technology, through the analysis of the current market demand and the development of local natural and human resources on the basis of preserving the original ecology to the greatest extent possible, the local economic development has been promoted, and the income level of the residents has been raised, thus promoting the modernization process of the whole region.

6. Acknowledgements

The work was supported by 2022 Mid-Young Youth Fund Project (2022KY-74) of Xi'an Traffic Engineering Institute.

References

[1] Liao X , Liao G , Cao J .Modeling of green agricultural environment and rape hyperspectral analysis based on machine learning algorithm.Optik, 2023, 273:170395-.

[2] Zuo L , Bingfang W U , You L ,et al.Big Earth Data Supports Sustainable Food Production: Practices and Prospects.Bulletin of Chinese Academy of Sciences, 2022, 36(8):885-895.DOI:10.16418/j.issn.1000-3045.20210706001.

[3] Duncan E , Rotz S ,André Magnan,et al.Disciplining land through data: The role of agricultural technologies in farmland assetisation.Sociologia Ruralis, 2022(2):62.DOI:10.1111/soru.12369.

[4] Burgers T A , Vanderwerff K J .Vision and Radar Steering Reduces Agricultural Sprayer Operator Stress without Compromising Steering Performance.Journal of Agricultural Safety and Health, 2022(3):28.

[5] Yadav V S , Singh A R , Raut R D ,et al.Exploring the application of Industry 4.0 technologies in the agricultural food supply chain: A systematic literature review.Computers & Industrial Engineering, 2022(Pt.2):169.

[6] Mavrogiorgos K , Kiourtis A , Mavrogiorgou A ,et al.Self-Adaptable Infrastructure Management for Analyzing the Efficiency of Big Data Stores.Journal of Advances in Information Technology, 2022(5):13.

[7] Begel A , Easterly T .Supporters of Beginning Agriculture Teachers Digital Conference: An Origin Story.The Agricultural Education Magazine, 2022(5):94.

[8] Stafford K C , Molaei G , Williams S C ,et al.Rhipicephalus capensis (Acari: Ixodidae), A geographically restricted South African tick, returning with a human traveler to the United States.Ticks and tick-borne diseases, 2022, 13(3):101912.DOI:10.1016/j.ttbdis.2022.101912.

[9] Li W , Mao E , Chen S ,et al.Design and verification of slip rate control system for straight line travel of high clearance self-propelled sprayer:.Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2022, 236(6):1319-1335.DOI:10.1177/09544070211032033.

[10] Nguyen T V , Kumar R , Musolff A ,et al.Disparate Seasonal Nitrate Export From Nested Heterogeneous Subcatchments Revealed With StorAge Selection Functions.Water Resources Research, 2022(3 Pt.2):58.DOI:10.1029/2021WR030797.

[11] Saayman A , Seetanah B , Sandt A V ,et al.Navigating the corn maze: Customizing travel cost models to value market segments in heterogeneous industries:.Tourism Economics, 2022, 28(4):899-919.DOI:10.1177/1354816620976968.

[12] Luo Y .SWOT Analysis of New Trend of Rural Tourism Development in Xinjiang.Asian Agricultural Research, 2022, 14(6):51-54.

[13] Ghimire C P , Bradley S , Ritchie W ,et al.Towards quantifying plot-scale overland flow connectivity using acoustic proximal remote sensing.Agricultural Water Management, 2022, 262:107418-.

[14] Emanuelson K , Covino T , Ward A S ,et al.Conservative solute transport processes and associated transient storage mechanisms: Comparing streams with contrasting channel morphologies, land use and land cover.Hydrological processes, 2022(4):36.

[15] Tomar G , Nagpure A S , Kumar V ,et al.High resolution vehicular exhaust and non-exhaust emission analysis of urban-rural district of India.Science of The Total Environment, 2022, 805:150255-.DOI:10.1016/j.scitotenv.2021.150255.

[16] Brown A , Williams R .Equity Implications of Ride-Hail Travel during COVID-19 in California:.Transportation

Research Record, 2023, 2677(4):1-14.DOI:10.1177/03611981211037246.
[17] Hodgson C , Envelope K C P .Using drone technology to collect school transportation data.Travel Behaviour and Society, 2023, 31:1-9.
[18] May C .The Professional Poet in the Romantic Period: Unpublished letters from Samuel Rogers to William Wordsworth.Romanticism, 2023, 29(1):68-79.DOI:10.3366/rom.2023.0581.
[19] Kumar D .European travel and leisure sector and uncertainties: A risk spillover analysis:.Tourism Economics, 2023, 29(1):48-67.DOI:10.1177/13548166211035954.
[20] Ao Y , Naess P .Travel behavior in smaller urban and rural areas.Transportation Research Part D: Transport and Environment, 2023, 115:103608-.