Hindawi
Mobile Information Systems
Volume 2022, Article ID 4835540, 9 pages
https://doi.org/10.1155/2022/4835540

Research Article

Adaptive Grid Division and Cluster Analysis of Decision Graph in Real Estate Marketing Using Data Mining

Sheng Lv

School of Economics and Management, Wuhan University, Wuhan, Hubei 430072, China

Correspondence should be addressed to Sheng Lv; 2018101050111@whu.edu.cn

Received 15 March 2022; Revised 11 April 2022; Accepted 19 April 2022; Published 30 April 2022

Academic Editor: Muhammad Babar

China's economy has been in a state of rapid development in recent years. People's lifestyles and quality of life have also undergone earth-shaking changes. More people put forward requirements for the living environment and living environment, which promotes the rapid rise of the real estate industry. The state has promulgated several macrocontrol measures. Real estate companies analyze the needs of various customer groups, affordable housing unit prices, and family populations, and formulate real estate marketing strategies to improve real estate sales. Therefore, this paper practices data mining technology to precisely analyze and locate the target customers. The cluster analyzes the mapping of the real estate marketing bureau. The cluster further divides the adaptive network and selects the data of a real estate company in China for research. It also analyzes the relationship between the family population and the affordable housing unit price and comprehensively improves the accuracy of real estate marketing. The results show that the acceptable housing unit price for families in this area is 9800 yuan. Study its data characteristics to judge the relationship between the population of other families. The analysis and research of data mining technology in the context of real estate can formulate more accurate marketing strategies. Moreover, it can reduce the investment amount of real estate companies in marketing.

1. Introduction

Real estate investment is the main force to improve economic development. Excessive investment or lagging investment restricts economic development. Nowadays, with the rapid development of the market economy, the market subject is the same as the investment subject, and the existing risks and responsibilities are borne by the enterprise. Therefore, while formulating and improving the management mechanism, the enterprise should also reduce the production cost and activate marketing thinking. The most key is the marketing decision, which can better control the enterprise scale, improve efficiency, and optimize the structure. There are many fuzzy factors in the study of real estate, and the decision-making of real estate marketing is no exception. Therefore, this paper selects a data mining algorithm to deal with this problem, to realize the adaptive division of real estate marketing and cluster analysis of decision graphs. At present, data mining technology is widely used in business, such as information management, decision support, and scientific research. Through the organic combination of various industries and data mining technology, it can give full play to its vitality [1].

The real estate industry belongs to a comprehensive market composed of multiple factors and has accumulated a large amount of data [2]. A key parameter used by real estate enterprises is data. Using data mining technology, the real estate industry can find high-value data, convert a large amount of data into high-value knowledge, and make decisions and formulate real estate marketing strategies based on this knowledge. Data mining expertise can be utilized to exactly analyze and locate the target customers. The data clusters can be formed to analyze the mapping of the real estate marketing bureau. The clusters can be utilized to divide the network and chose the dataset of a real estate. The novelties of this paper include

(1) The proposal of a data mining algorithm in real estate marketing to analyze the needs of home buyers' group graph

(2) Establishing a data clustering algorithm for the dynamic grid to divide the real estate marketing grid and calculate the grid density

(3) Cluster analysis of real estate marketing decision-making map and study of the real estate operation data, operation objectives, social resources, internal ability, and external conditions in a scientific way as an important basis for real estate operation and decision-making

(4) Establish a cluster analysis model for data mining, analyze customer demand behavior, select an urban real estate project, analyze the clustering characteristics of home buyers by data mining, and make the decision to analyze the relationship between family population and affordable housing unit price

The rest of the paper is organized as follows. In Section 2, related work is presented. In Section 3, clustering analysis using data mining is presented, which is followed by building a clustering analysis model of real estate marketing using data mining in Section 4. In Section 5, empirical results and discussions are provided which is followed by conclusion of this paper in Section 6.

2. Related Work

China's market economy has risen rapidly in recent years. As an important pillar industry of China's economic development, the real estate industry pays great attention to real estate marketing [2]. Chinese scholars Zheng et al. used correlation analysis to predict China's house prices and tested the accuracy of house price prediction in a community in Guangzhou, and the result was that the difference between the predicted house prices and the actual house prices was less than 5% [3]. Milutinovi et al. use the grey system GM (1, 1) model to study the problem of house price prediction, improves the model according to the actual situation, selects multiple data from the recent samples to build the model, calculates the average house price in China, and tests and analyzes it by comparison [4].

Chen et al. based on the time series analysis method, combined with the law of real estate time series, predicts the trend of house prices, selects Guangzhou as an example to analyze the real estate prices, constructs the house price prediction model, and uses residual analysis to test the error. The results show that there is a small error between the prediction results and the actual results, which can accurately predict the future house prices of real estate [5]. Cena and Gagolewski use OWA operator weighting algorithm to predict the real estate price and use a variety of weighting methods and weighting calculation methods to predict the real estate price. The prediction results show that the weighting method can effectively predict the real estate price and formulate the marketing plan according to the real estate price [6]. Liu et al. when studying the factors affecting China's real estate price selected the multilinear regression method to analyze China's real estate market price [7]. Ojetunde et al. combine data mining technology to explore various factors affecting real estate prices according to local real estate prices and input multiple factors on the MATLAB tool to predict prices [8]. Luo and Zhu use data mining technology to extract various factors related to real estate prices, build a spatial panel model, and analyze its elasticity [9].

The results show that construction cost, per capita disposable income, economic development level, and land price are the key factors of real estate marketing, among which disposable income accounts for the largest proportion. Zheng scholar analyzed the impact of employment rate and housing loan interest rate on house price based on a vector regression model. American housing loan interest rate is an important factor affecting the real estate market [10]. Zhang and Fang scholars use the time series prediction model to predict and study the real estate price based on multiple factors such as the average annual household income, the average real estate price, and the urban population [11]. The results show that the highest correlation with house prices is family income.

3. Clustering Analysis Based on Data Mining

Clustering analysis allows the subdivision of data into subsections (e.g., clusters) to keep similar data points in the clusters. Clustering is to split a set of data into various classes according to similarity. Clustering is a data analysis tool [12]. The cluster analysis model is established based on data mining using the data clustering algorithm based on dynamic grid.

3.1. Data Mining Process. The essence of data mining is to mine the laws of practical application data and high-value patterns of unknown datasets. Its basic steps include data preparation, data mining, evaluation patterns, knowledge application, and knowledge consolidation [13]. Data mining integrates many disciplines to form a new information processing technology. Data mining has many characteristics. The formation of data mining is a data mining technology derived from the need for data processing after accumulating a large amount of data in business and academic fields. At the same time, there are many possibilities of data in data mining, and the data structure is also diversified, such as incompleteness, complexity, large dimension, and noise. Data mining integrates many disciplines, such as computers, statistics, and mathematics. In the application of data mining, we should first determine the goal, then explain the problems to be handled, and clarify the types to be handled. The common types include time-series pattern, association analysis, clustering, classification, and deviation. Figure 1 shows the basic process of data mining.

3.2. Data Mining System. Knowledge discovery is composed of many parts, which are closely related. The knowledge discovery process with data mining as the core is the data mining system, and all algorithms serve the mining system. The purpose of studying the data mining system is to build a scientific system structure, which is convenient for the embedding and reuse of mining algorithms and the better

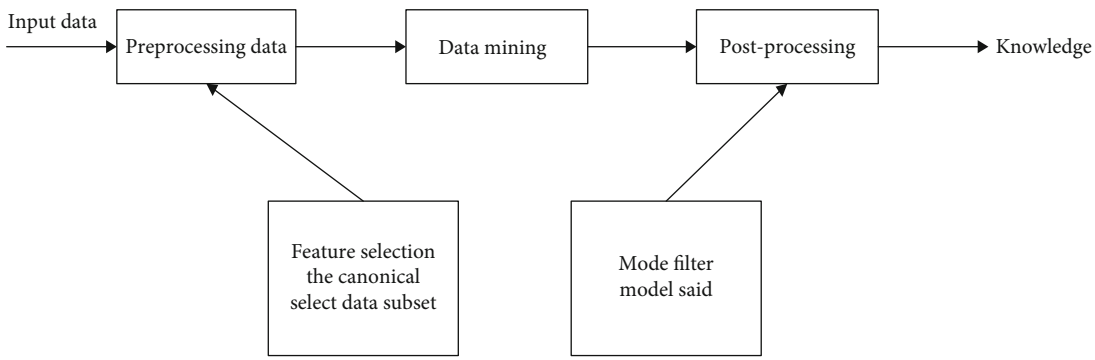

Figure 1: Basic process of data mining.

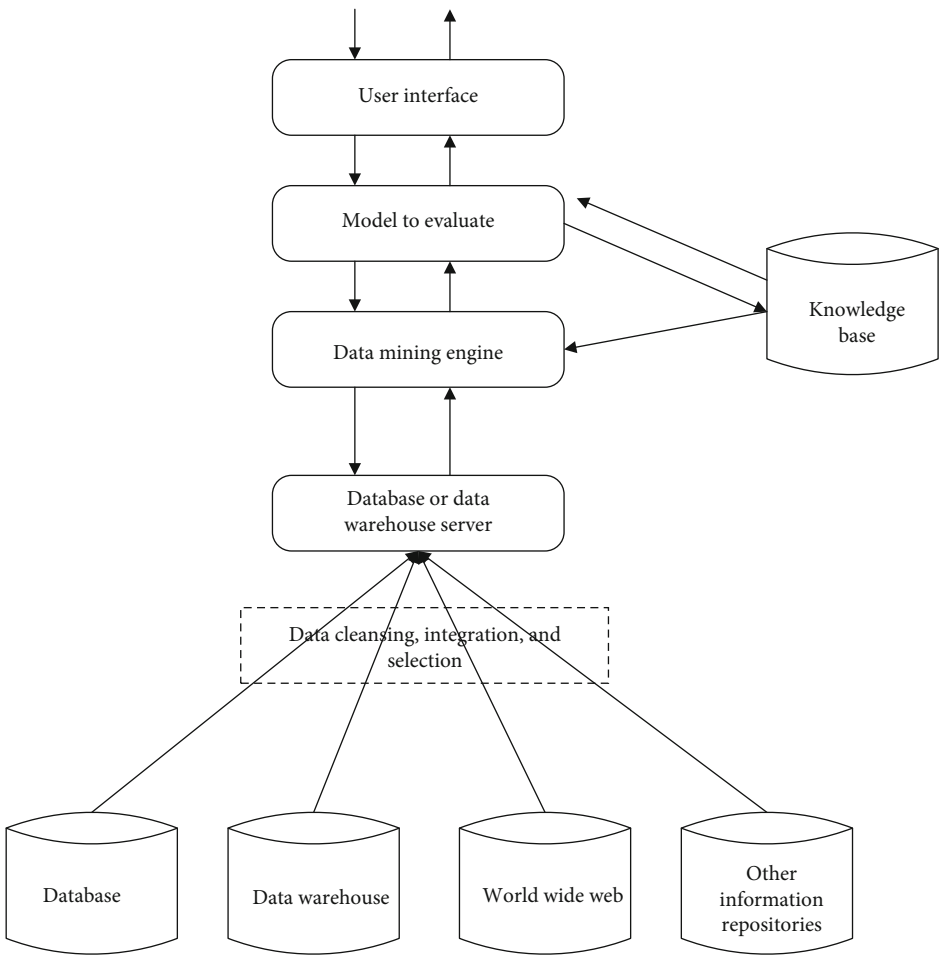

Figure 2: Data mining system.

combination of algorithms with other modules [14]. Figure 2 shows the structure of the data mining system.

3.3. Cluster Analysis. Clustering is to divide a group of individuals into multiple categories according to similarity. As a widely used data analysis tool, multiple physical objects and abstract objects form a set [12]. Clustering refers to the generation of a set of class objects. Therefore, in most applications, the data objects in the class are analyzed. The object is highly like the objects in the same class but different from the objects in other classes. Compared with the classification technology, the category is marked on the classified data object, and the specific class information on the database is very clear. We need to mark the class to which each data belongs. Clustering is to divide those unknown data and unmarked data and calculate and determine them by the clustering algorithm. Therefore, the data attribute category in classification is known, while the data category in clustering is unclear and relatively fuzzy. At this time, clustering judgment should be used.

The data relating to cluster analysis consists of the following types: the first is the data matrix, the second is the data dissimilarity matrix, and there is also the data similarity matrix, which is like the dissimilarity matrix.

3.3.1. Data Matrix. Assuming that the data objects are n $Qi(i = 1, 2, \cdots, N)$ and each data object has k attributes, Qij $(i = 1, 2, .., N; j = 1, 2, \cdots, N)$ represents the object data and attributes. The form of the data structure is a two-dimensional table or n below $\times K$ matrix.

$$\begin{bmatrix} O_{11} & \cdots & O_{1k} \\ \cdots & \cdots & \cdots \\ O_{N1} & \cdots & O_{NK} \end{bmatrix}. \quad (1)$$

3.3.2. Data Dissimilarity Matrix. The dissimilarity relationship between N different objects is saved in the dissimilarity matrix, and N is calculated by the following formula $\times N$ matrix representation

$$\begin{bmatrix} 0 & & & & \\ P(2,1) & & & & \\ P(3,1) & P(3,2) & 0 & & \\ \cdots & \cdots & \cdots & & \\ P(N,1) & P(N,2) & \cdots & \cdots & 0 \end{bmatrix}. \quad (2)$$

$P(ij)$ in the above matrix is the dissimilarity or measurement difference between data objects i and j. Generally, $P(i, j) \geq 0$, and $P(i, j) = P(j, i)$, and $P(i, i) = 0$. If objects i and j are very similar, the value approaches 0. If the value approaches 1, infinitely, there is a large difference between the two objects.

3.4. Data Clustering Algorithm Based on Dynamic Grid. A specific clustering algorithm is proposed based on dynamic grid.

3.4.1. Meshing. By dividing each dimension of d-dimensional data space into disjoint intervals with the same length, multiple grid cells are formed. The grid cells on each dimension are in the form of left closed and right open intervals. This paper assumes that each dimension is a uniformly divided k different intervals, then k^d represents the grid in the data space, and the side length of the one-dimensional grid is calculated by the following formulaδj:

$$\delta_j = \frac{h_j - l_j}{k}. \quad (3)$$

The data is not evenly distributed in the space. Theoretically, there will be a phenomenon that k^d is higher than the total data object. However, setting a reasonable k value in the application can make k^d lower than the total data. Empty grids without data objects will appear on the divided grid cells. At this time, the nonempty grid cells should be reserved to save some memory space.

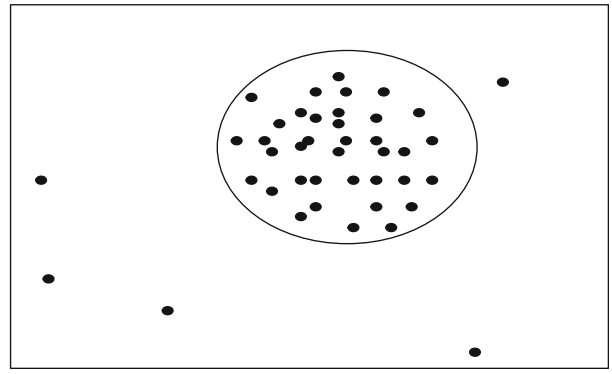

Figure 3: Clustering algorithm original data class.

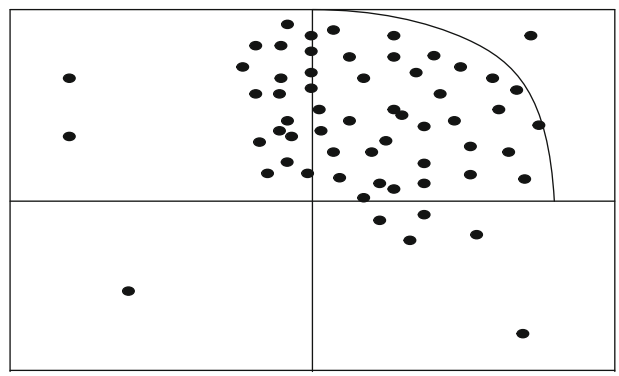

Figure 4: Clustering algorithm loses boundary class.

3.4.2. Calculate Grid Density. Most grid-based clustering algorithms convert the number of data points on grid cells into grid cell density values in the process of solving grid density. This method has simple operation and strong feasibility [15]. However, when used in clustering, the influence information in the space near the data points will be lost, which will cause the points of the same kind to be divided into different classes. This problem occurs more on the boundary of fruit classes. Figure 3 below shows the influence of the density calculation method on clustering results. The results of Figure 4 can be obtained after running the grid algorithm for oval subclasses shown in multiple figures on the original dataset. As a result, most of the data such as subclasses at boundary points are mistaken for noise data and discarded. Due to a large number of data points in the upper-right cell calculated by the traditional algorithm, it can be regarded as dense cells. The remaining three cells are regarded as sparse cells because they contain fewer data points. When clustering, only all data points in the dense cell grid are taken into account, resulting in the whole tribe of clustering boundary points in three adjacent grids and eventually being lost.

3.5. Decision Graph Clustering Analysis. Real estate marketing decision-making is used to express the decision-making made in the process of operating real estate. The essence of business decision-making is the process of market investigation, scientific research and analysis, and making a scientific decision. This process integrates people's creative thinking

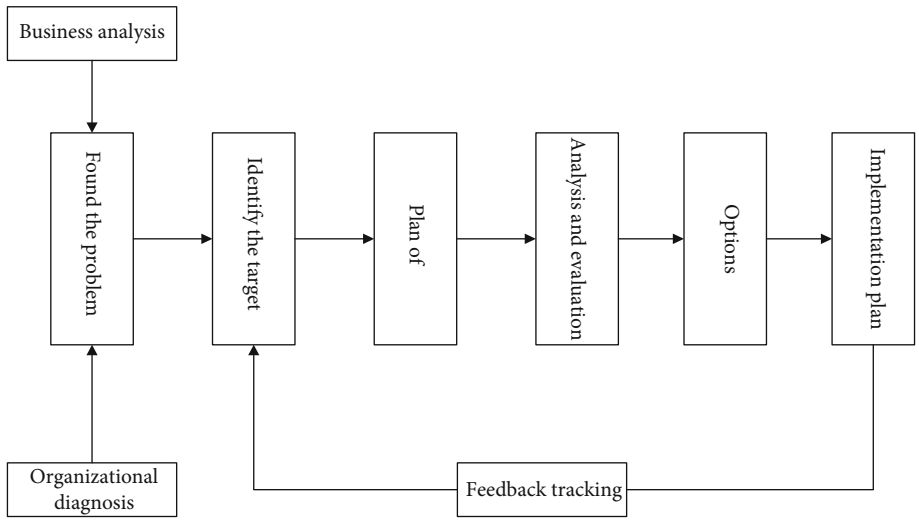

Figure 5: Real estate marketing decision-making procedure chart.

ability and logical thinking such as early analysis, reasoning, and judgment [16]. The marketing decision-making contents of real estate companies mainly include business analysis and decision-making. Business analysis is from the perspective of real estate company management and adopts a scientific way to analyze the real estate company's business data, business objectives, external conditions, social resources, and internal capabilities [17]. Decision-making is a feasible scheme formulated on the premise of business analysis, and the best marketing decision is selected by combining judgment with other nonquantitative conditions and human factors [18]. In Figure 5, the procedural chart of real estate marketing decision-making is shown.

4. Building a Clustering Analysis Model of Real Estate Marketing Using Data Mining

A key index in real estate marketing decision-making is sales forecast, so we should choose the best way to analyze the problem of real estate sales forecast. The traditional prediction of real estate sales has certain limitations. We should use a scientific way to predict the sales capacity and level of each real estate project, provide a decision-making basis for real estate, analyze the fuzziness and particularity affecting sales based on the analysis of the factors affecting each real estate project, and study the real estate marketing based on the cluster analysis model.

4.1. Establish a Cluster Analysis Model Based on Data Mining. Real estate sales are affected by many factors, such as whether the location is prosperous, traffic convenience, sales price, real estate environmental quality, performance quality, promotion mode, supporting facilities, house type, consumption level, and economic policy. Assuming that n factors affect real estate production and sales, and an evaluation factor set of $U = \{u_1, u_2, \cdots, u_n\}$ is formed, each factor in this set affects the real estate sales volume to different degrees, so it should be given corresponding weight. $A = \{a_1, a_2, \cdots, a_n\}$ is the weight set, and a_1 represents the corresponding weight of the i factor u_i, which must meet the following conditions:

$$1 \geq a_i \geq 0, \tag{4}$$

$$\sum_{i=1}^{n} a_i = 1. \tag{5}$$

In this paper, five grades of poor, poor, average, good, and good are selected to represent the real estate sales situation. If m represents different grades of sales situation, its constituent sales levels are collectively expressed as $Y = \{y_1, y_2, \cdots, y_m\}$. All factors on U have corresponding membership degrees in Y, that is, the grade degree of each factor on Y, for the i factor, the membership degrees of different grades are R_i fuzzy subsets on Y:

$$Ri = \{r_{i1}, r_{i2}, \cdots, r_{im}\}. \tag{6}$$

All R_i forms the following R fuzzy comprehensive evaluation matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}. \tag{7}$$

The fuzzy subset B in the following Y is obtained based on the fuzzy matrix synthesis operation, as shown below:

$$B = (b_1, b_2, \cdots, b_n) = (a_1, a_2, \cdots, a_n) \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}. \tag{8}$$

According to the above conclusion, the core of the fuzzy mathematical model is to clarify the membership degree and weight.

4.2. Customer Demand Behavior Data Analysis. Cluster analysis is divided into several different types according to distance and similarity. Generally, distance is used in sample cluster analysis, and similarity coefficient is used for cluster analysis between variable classes [19]. Customer demand behavior data analysis is calculated in the using the Euclidean distance. Suppose the number of samples is n, expressed as $X_1, X_2, \cdots, X_n$, the number of indicators in each sample is P, which can be expressed as $X_i = (X_{i1}, X_{i1}, X_{i1})$, $X_i = (X_{i1}, X_{i2}, \cdots, X_{ip})$, $i = 1, 2, \cdots, n$. Assuming that d_{ij} is the distance between X_i and X_j, the common methods for calculating the distance are Minkowski distance, Euclidean distance, and absolute distance.

Euclidean distance formula: calculate the Euclidean distance between the k sample and the i sample, which is expressed as follows:

$$D_{ik} = \sqrt{\sum_{j=1}^{p} \left(X_{ij} - X_{ki}\right)^2 \max \lim_{x \longrightarrow \infty}}, \tag{9}$$

D represents the distance, X is the variable, and K and I represent the samples.

The distance between two samples is the square sum and square root of the difference between the values of two different variables. The Euclidean distance square is the Euclidean distance square, and the sum of the squares of the value differences of each variable is the distance between [20]. Minkowski distance is extended based on Euclidean distance, that is, the 4th power root of the sum of the absolute value of the q-power value of the value difference of each variable.

$$D_{ik} = \left[\sum_{j=1}^{p} \left|X_{ij} - X_{kj}\right|^q\right]^{1/q}. \tag{10}$$

Absolute value distance: the sum of the absolute values of the difference values of all variables.

$$Dik = \sum_{j=1}^{p} \left|X_{ij} - X_{kj}\right|. \tag{11}$$

Table 1: China's real estate industry and the speed of economic development.

The speed of national economic development	State of the real estate industry
<4%	Shrinking
4%-5%	To stop or go backward
5%-8%	Stable development
>8%	Rapid development
10%-15%	Speed development

Table 2: Weight table of clustering influencing factors of house purchase plan.

Name	Importance	Measurement	Type	Label
Q52	0	Binary	Num	Q52
Q53	0	Ordinal	Num	Q53
Q54	0.3907	Ordinal	Num	Q54
Q55	0	Ordinal	Num	Q55
Q61	0.669	Ordinal	Num	Q61
Q62	1	Ordinal	Num	Q62
Q7	0	Ordinal	Num	Q7
Q8	0	Ordinal	Num	Q8

Table 3: Description of the relationship between family population and affordable housing unit price.

	N	Minimum	Maximum	Mean	Std. deviation
Q8	685	6800	12800	9800	3000
Valid N (listwise)	685	—	—	—	—

Table 4: Relationship between family population and affordable housing unit price.

	N	Minimum	Maximum	Mean	Std. deviation
Q8	68	6800	12800	9732.46	2932.46
Valid N (listwise)	68	—	—	—	—

The similarity of different classes is defined by the distance between classes, and the representative points of classes are described by the following methods:

(1) Select the longest distance and the shortest distance between two classes as the representative points of the class

(2) Calculate the average distance between stores of S and t as the square of the distance between the two classes:

$$d^2(S, T) = \frac{1}{n_s n_t} \sum_{i \in S, j \in T}^{n} d^2{}_{ij}. \tag{12}$$

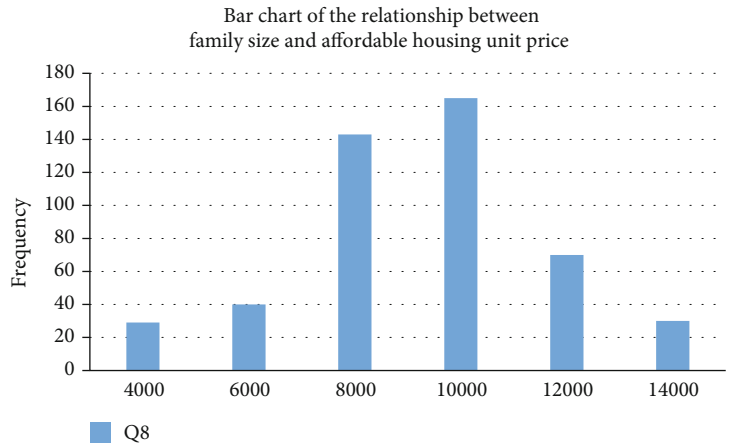

Figure 6: Histogram of the relationship between family population and affordable housing unit price.

n_s is the number of points on class s, n_t is the number of points on class T, the distance between i and j is represented by d_{ij}, and the distance between class T and class s is represented by $d(S, T)$. When calculating the distance between classes, it is less dependent on some points in the class.

5. Empirical Analysis

The rationale of the empirical analysis is the investigation of the real estate data of the region in 2020.

5.1. Data Mining Analysis of Real Estate Marketing Consumption Prediction. This paper uses data mining to study the clustering analysis of adaptive division and decision graph of real estate marketing and takes a real estate project in China as the research object for empirical analysis. By investigating the real estate data of the region in 2020, the real estate information is obtained according to the information published online and the actual survey; the consumption characteristics of consumers' real estate are obtained by questionnaire, so the amount of data collected during the survey is large. Therefore, this paper uses data mining technology to analyze project marketing as an important basis for project decision-making. The relationship between the development status of China's real estate industry and the speed of economic development is listed in Table 1.

The following is the basic process based on data mining:

(1) Determine the marketing goal: use data mining to find some significant characteristics of home buyers, and cluster analysis based on these characteristics

(2) Prepare data: extract the required data based on data mining marketing objectives, and then, further check and correct the data, including data attributes, format, processing missing values, loading data, etc., and then, import the collected data to Enterprise Miner in SAS for analysis. By analyzing the data, select the attribute data with significant personal characteristics of consumers, mainly including Q1 (housing status), Q2 (purchase intention), Q7 (purchase purpose), Q8 (current housing area), Q11 (purchase plan), Q52 (gender), Q53 (age), Q54 (family structure), Q55 (education level), Q61 (personal monthly income), Q62 (family monthly income) and select this field and load the data in SAS Enterprise Miner 4.1 software

(3) Research on customer data based on the cluster analysis method of data mining

(4) In this paper, the data is divided into three parts, of which 40% is the training set, 30% is the verification set, and 30% is the test set. Q11 (house purchase plan) is selected as the target variable. SAS enterprise miner software is used to cluster and analyze the collected data. There are two clusters in total. The clustering results are listed in Table 2.

According to the data in the above table, among the statistical consumer variables, the clustering factors affecting the house purchase plan model mainly include q54 (family income), q61 (personal monthly income), and Q62 (family monthly income). The factor with the highest weight is Q62 (family monthly income), followed by q61 (personal monthly income), and the factor with the lowest weight is q54 (family structure).

Through analysis, it is concluded that the factors affecting customers' house purchase plans are family monthly income, personal monthly income, and family structure, that is, Q62 = 7 and q61 = 7. The probability of such customers buying houses reaches 100%.

5.2. Analysis of Real Estate Marketing Decision Chart. This paper selects affordable housing unit prices (Q8) and (Q15) family populations for cluster analysis to master the relationship between family population and affordable housing unit prices. Here, you should first understand the data status of "affordable housing unit price (Q8)", including standard deviation, all attributes, etc. click the analyze = = > descriptive, statistics = = > descriptions menu, and the following dialog box will pop up automatically:

5.2.1. Static Descriptive Analysis. In Table 3, the description of the relationship between family population and affordable housing is highlighted.

According to the data analysis in the above table, the number, minimum value, maximum value, and statistics of Q8 samples are obtained; after cluster analysis, the results show that the average number of 685 data is 9800, and the difference between the mean and the marked difference is 3000.

5.2.2. Conditional Descriptive Analysis. The above is only a simple data description, and data mining should be carried out in other options. By analyzing the "family population" factor and selecting organize output by family population (Q15) for an explanation, multiple descriptions describing the situation of the family population can be obtained. In Table 4, the relationship between family population and affordable housing unit price is shown.

The standard deviation and mean in the data in the above table show that compared with other families, the consumption level of a family of three and a family of four is higher, so the housing price of 9800 yuan is more in line with the market demand. The data characteristics of affordable housing unit prices (Q8) are shown in the following histogram, which is shown in Figure 6.

6. Conclusions

Real estate is not only the pillar industry of China's economic development but also the core part of China's fixed-asset investment. At present, the local real estate industry continues to rise, and the house price also increasing rapidly. It is difficult for customers to make decisions when purchasing real estate. To solve this problem, this paper uses data mining technology to analyze the real estate data, divide the real estate market sales adaptive grid, and judge the best selling price of the real estate market in various regions. This Ayong cluster analysis method analyzes the decision chart to provide customers with accurate data for decision-making and judgment. Through scientific analysis and monitoring of real estate data, provide investors with a decision-making basis, so that they can better avoid risks and find ideal real estate investment projects. This paper selects two indicators of affordable housing unit price (Q8) and (Q15) family population for cluster analysis. The results show that the house price of 9800 yuan is more in line with the market requirements. Draw a histogram to represent the data characteristics of affordable housing unit price (Q8), and formulate the real estate marketing strategy according to the data characteristics.

Data Availability

All the data along with the experimental results are included as part of this work.

Conflicts of Interest

I declare that there is no conflict of interest for the publication of this paper.

References

[1] F. G. Wang, "Application analysis of data mining technology in 4G communication network optimization," *China New Telecommunications*, vol. 20, no. 21, p. 87, 2018.

[2] V. Zahirovic-Herbert, B. D. Waller, and G. K. Turnbull, "Properties that sell at or above listing price: strategic pricing, better broker or just dumb luck?," *The Journal of Real Estate Finance and Economics*, vol. 60, no. 1-2, pp. 53–76, 2020.

[3] W. Zheng, X. Li, N. Guan, and K. Zhang, "Correlation analysis of fiscal revenue and housing sales price," *Mathematical Computation*, vol. 9, no. 1, p. 10, 2020.

[4] A. Milutinović, Z. Gligorić, Z. Gojković, and A. Ganić, "Forecasting the building deformation using the non-equidistance Grey Model (1,1)," *Podzemni Radovi*, vol. 2020, no. 36, pp. 49–59, 2020.

[5] X. Chen, J. Qiu, and Y. Wu, "Research on forecast of Beijing housing price based on spatial gray Markov model," *Journal of Physics: Conference Series*, vol. 1437, no. 1, article 012082, 2020.

[6] C. Anna and G. Marek, "Genie+OWA: robustifying hierarchical clustering with OWA-based linkages," *Information Sciences*, vol. 520, no. 2, pp. 324–336, 2020.

[7] J. Liu, L. Qiao, and N. Zhang, "A research on the regional differences in the influential factors of Inner Mongolia real estate prices based on panel data model," *Credit reference*, vol. 35, no. 6, pp. 65–71, 2017.

[8] I. Ojetunde, A. I. Sule, O. A. Kemiki, and I. A. Olatunji, "Factors affecting the academic performance of real estate students in a specialized Federal University of Technology in Nigeria," *Property Management*, vol. 38, no. 2, pp. 177–198, 2020.

[9] J. Luo and J. F. Zhu, "Empirical analysis of real estate price based on data mining and spatial measurement," *Research on Development*, vol. 5, no. 12, pp. 128–131, 2013.

[10] S. Zheng, "Research on dynamic relationship between interest rate & exchange rate—based on the analysis of marketization of interest rate and exchange rate," *Journal of Economic Management: Chinese and English*, vol. 10, no. 1, p. 16, 2021.

[11] T. Zhang and B. Fang, "Prediction of the real estate industry economics based on LSTM model," *Journal of Physics: Conference Series*, vol. 1802, no. 4, article 042095, 2021.

[12] S. Jha, G. P. Joshi, L. Nkenyereya, D. W. Kim, and F. Smarandache, "A direct data-cluster analysis method based on neutrosophic set implication," *Computers, materials and continuum*, vol. 2020, no. 11, p. 11, 2020.

[13] S. H. Guo, X. Lu, and College of Systems Engineering, National University of Defense Technology, Changsha 410073, China, "Live streaming: data mining and behavior analysis," *Acta Physica Sinica*, vol. 69, no. 8, article 088908, 2020.

[14] S. Jiang, J. J. Ren, and Z. Y. Gao, "Design of large network database mining system based on clustering optimization," *Modern Electronics Technique*, vol. 43, no. 6, pp. 175–177, 2020.

[15] X. Xu, S. F. Ding, and T. F. Sun, "Large-scale density peaks clustering algorithm based on grid screening," *Journal of Computer Research and Development*, vol. 55, no. 11, pp. 2419–2429, 2018.

[16] S. Debika, "Home sweet virtual home: The use of virtual and augmented reality technologies in high involvement purchase decisions," *Journal of Research in Interactive Marketing*, vol. 12, no. 4, pp. 398–417, 2018.

[17] E. Santos and F. O. Tavares, "The profile of Portuguese people when choosing an apartment," *Real Estate Management and Valuation*, vol. 29, no. 2, pp. 38–51, 2021.

[18] T. W. Long and Y. F. Shi, "Investigating the financing decision-making in China's real estate industry based on capital structure theory," *Journal of Engineering Management*, vol. 35, no. 3, pp. 141–146, 2021.

[19] L. M. Du Ruizhong and J. Tian, "Multi-keyword ranked ciphertext retrieval scheme based on clustering index," *Journal of Computer Research and Development*, vol. 56, no. 3, pp. 555–565, 2019.

[20] Y. W. Li, F. X. Fu, and Y. F. Guo, "Construction and application for a class of one-parameter perturbation of copula functions," *Journal of Northwest Normal University (Natural Science)*, vol. 57, no. 3, pp. 32–38, 2021.