Application of Data Mining Technology in Evaluating Real Estate Investment Plan Based on GRA-AHP

ZHUANG MEIER, WEN-TSAO PAN*, SHI ZHUOHONG, ZHOU YINGYING and ZHONG ZUCHANG
Guangdong University of Foreign Studies, Guangzhou, Guangdong, China

*E-mail: teacherp0162@126.com

Abstract. Data mining and information technology is increasingly important in the era of big data. Many researches have been discussing the application of data mining and information technology in all walks of life. To further explore the practical value of data mining technology in economic and business field, this paper studies its application in evaluating real estate investment plans. R software is applied as the statistical computing console and relative program code is created to implement grey relational analysis-analytical hierarchy process (GRA-AHP), which establishes an improved evaluation model of decision-making. Firstly, the related data of the research is divided into several layers, namely the target layer, criteria layer and plan layer. In the criteria layer, the processing method is different from previous research because in this paper the indicator value of the plan layer to the criteria layer is calculated based on grey relational analysis rather than the pairwise comparison judgment matrix proposed by previous scholars, which makes this analysis more scientific and accurate. Then the plan layer single sort weight and the total sort weight of each plan can be calculated to provide reference for the final investment decision. This paper aims to propose a better evaluation method in the real estate investment and promote the application of data mining technology in economic fields.

1. Introduction

As an up-to-date analytical method, data mining technology has been frequently used in various fields in the modern research. Among them, the combination of data mining technology and real estate industry is an increasingly important research hotspot [1-2].

So far, plenty of researchers have proposed all kind of related theories in the field. Jovanovic conducted sensitivity analysis for investment projects under risk conditions [3]; Borgonovo and Peccati discussed the overall situation of investment decisions under uncertain conditions [4]; Kollmann and Kuckertz analysed venture capitalists' risk investment decision-making processes [5]; Ke Xiaoling established a different evaluation system from several aspects [6].

Different from previous studies, in this paper data mining technology will be applied to analyse the real estate investment. The specific calculating methods used in the study are grey relational analysis (GRA) and analytic hierarchy process (AHP). According to the GRA theory, the real estate investment decision-making problem is a grey system where both known and unknown information are included. The indicators of real estate investment decisions are not independent of each other. [7-8]. As one of the analytical methods of data mining technology, AHP is a useful tool for decision-making and evaluation of complex problems proposed by American operations researcher T. L. Saaty [9]. It is
suitable for analysing decision-making problems of complex systems composed of many interrelated factors. [10]. AHP has the advantages of system and flexibility and is widely applied in all sorts of fields such as economical management and public administration [11].

Therefore, to further explore the practical value of data mining technology in the real estate investment decision-making field, this paper combines GRA and AHP to construct the evaluation model of the real estate investment development plan. The association hierarchy analysis and evaluation system will be proposed in the paper, as well as the real estate investment evaluation model method based on GRA-AHP. In brief, what we focus on is the application of data mining technology in a real estate industry and how it spurs the development of this industry.

2. Research Methods

2.1. Analytic Hierarchy Process(AHP)
AHP is to decompose related elements into goals, criteria, programs and other levels, and conduct qualitative and quantitative analysis on this basis. AHP first proposes a general goal based on the nature and requirements of the problem. Then the problem is decomposed into layers, and the factors in the same level are determined by the pairwise comparison method to determine the respective weight coefficients relative to the upper layer target. This layer-by-layer analysis continues until the last level, giving a ranking of all factors relative to the overall goal.

2.2. Grey Relational Analysis(GRA)
GRA is a direct and comprehensive evaluation method when conducting data mining. The degree of association is essentially the degree of difference in geometry between curves. Therefore, the difference between the curves can be used as a measure of the degree of association. The research object of this method is the non-dimensionalized sequence, which is divided into a parent sequence and a subsequence. In most cases, the parent sequence, also called reference sequence, is a standard sequence in the grey relational analysis, reflecting the behaviour characteristics of the system. On the other hand, the subsequence, also named comparison sequence, is a sequence of objects in the gray relational analysis, which is a data sequence composed of factors affecting system behaviour.

2.3. Specific Steps of AHP based on GRA
First, establish a hierarchical structure model. Next, construct all judgment matrices in each level. Finally, conduct the hierarchical single sorting, hierarchical total sorting and consistency checking.

For each upper element, consider the lower elements that have a logical relationship with them, and make a pairwise comparison between them. The computing results of the judgment are given in quantitative numbers and are represented in a certain matrix, which is called a "judgment matrix".

GRA-AHP is to use the grey relational degree calculation to determine the weight of the relative importance ranking of each element in each level, list the mutual factors of the criterion layer mutual factor weight table and the program layer to the criterion layer. Through the comprehensive evaluation of each level, the total ranking weight of each scheme for the target layer is given.

3. Application of Investment Plan Evaluation Model Based on GRA-AHP

3.1. Annual Measurement Index
Here is a real estate development enterprise that is to carry out investment development planning. There are four investment options to choose from, and the results of the annual indicators of each program are shown in the following table summarized from a previous study [12]. (see table 1)

To improve the accuracy of analysis and reduce the subjectivity of decision-making, certain qualitative indicators such as expected sales situation and customer satisfaction have been graded according to uniform standards, and then processed into quantitative indicators through fuzzy analysis.
Table 1. Annual Indicators of Four Plans.

| Index                              | Plan T1 | Plan T2 | Plan T3 | Plan T4 |
|------------------------------------|---------|---------|---------|---------|
| Economic benefit                   |         |         |         |         |
| Financial internal rate of return (%) | 16.3    | 15.97   | 15.16   | 15.65   |
| Financial net present value (Ten thousand yuan) | 9579    | 10300   | 7931    | 9682    |
| Economic internal rate of return (%) | 9.03    | 9.21    | 8.38    | 8.93    |
| Economic net present value (Ten thousand yuan) | 2925.6  | 4120.5  | 2225.1  | 3461.2  |
| Degree of risk                     |         |         |         |         |
| Investment loss (Ten thousand yuan) | 96.8    | 110     | 97.9    | 110     |
| Project risk (%)                   | 50      | 60      | 50      | 60      |
| Investment returns                 |         |         |         |         |
| Total investment (Ten thousand yuan) | 3100.2  | 3563.4  | 3369.8  | 3647.3  |
| Payback period (Year)              | 4.56    | 7.64    | 7.88    | 7.71    |
| Market effect                      |         |         |         |         |
| Expected sales (%)                 | 60      | 70      | 50      | 60      |
| Customer satisfaction (%)           | 50      | 60      | 60      | 50      |

3.2. Establish an Evaluation Index System

According to the fundamental theory of analytic hierarchy process, the evaluation index system is established in such a way as follows. The corporate real estate development investment plan is set as the target level A, which is the problem needed to be solved. The target layer A includes four sub-targets B1, B2, B3, and B4, each of which is a preferred criterion for a real estate development investment plan, so this layer is called a criterion layer and is represented by B. Each criterion layer contains several indicators, and the values of these indicators are applied to the total target A through their subordination principle B. The four options for development investment become the layer of the program to be evaluated, denoted by T.

The following figure illustrates how the evaluation system works. (see figure 1)

3.3. GRA-AHP Empirical Analysis Steps

This paper is based on GRA-AHP, which is implemented in the R software using the GRA program.
Since each of the overall indicators in the criterion layer is composed of several subordinate indicators, the GRA of the subordinate indicators in the criterion layer is performed to obtain the overall indicators of the matching criterion layers of each project which needs selecting.

First, according to each subordinate index value below the criterion layer, the standard sequence and the comparison sequence of the subordinate index are calculated, and the gray correlation program is used to obtain the overall index value $B$ of each scheme $T$ relative to each criterion.

Specifically, the GRA is implemented on the four subordinate indicators of financial internal rate of return, financial net present value, economic internal rate of return and economic net present value, with the aim to obtain the corresponding economic benefit index values of each program; Similarly, the GRA is carried out on the investment loss amount and the project risk degree, and then the index value of the risk degree is obtained as well. In the same way, we use GRA to implement the data mining process on the total investment amount and the investment recovery period in order to obtain the investment income index value; It is nearly the same way to calculate the data of the sales situation and the customer satisfaction, and therefore, the market effect indicator value is obtained.

After the gray relational analysis processing, the overall index values of the corresponding criterion layers of each scheme are shown in the following table. (see table 2)

| Criteria layer | Weights |
|---------------|---------|
| Economic benefit $B_1$ | 5/7 |
| Degree of risk $B_2$ | 4/7 |
| Investment returns $B_3$ | 1/2 |
| Market effects $B_4$ | 1 |

Second, calculate the weight of the relative importance order of each element in the criterion layer. Specifically, after determining the index values of the corresponding criterion layers in the solution layer, the grey relational analysis is again used, and the index values obtained in the previous step are brought into the GRA program, and each scheme corresponds to the overall of the criteria. The index value is used as a comparison sequence, and the corresponding maximum value in each comparison sequence is used as a reference sequence, thereby calculating the weight of the relative importance ranking of each element in the criterion layer, and the results are shown in table 3.

| Criteria layer | Weights |
|---------------|---------|
| Economic benefit $B_1$ | 1 |
| Degree of risk $B_2$ | 4/5 |
| Investment returns $B_3$ | 3/2 |
| Market effects $B_4$ | 5/7 |

Then, according to the formulation principle of the judgment matrix in the analytic hierarchy process and the above-mentioned data, the weight values of the overall indicators in the criterion layer with respect to the target layer A are calculated, and the criterion layer B mutual factor weight table is listed. As it is shown in the table below. (see table 4)

| A | Economic benefit $B_1$ | Degree of risk $B_2$ | Investment returns $B_3$ | Market effects $B_4$ |
|---|----------------------|----------------------|--------------------------|----------------------|
| Economic benefit $B_1$ | 1 | 4/5 | 2/3 | 7/5 |
| Degree of risk $B_2$ | 5/4 | 1 | 7/8 | 7/4 |
| Investment returns $B_3$ | 3/2 | 8/7 | 1 | 2 |
| Market effects $B_4$ | 5/7 | 4/7 | 1/2 | 1 |

From the above calculation, it is convenient for us to work out the weights of the scheme layer to each criterion layer with the help of data mining technology, according to the all those known data. Hence, the scheme layer to the criterion layer mutual factor can be listed smoothly, which is illustrated in table 5.
After constructing the judgment matrices, the analysis process is started using the procedure of the AHP. The numerical values in the judgment matrix obtained above are substituted into the AHP, and the results of the analytic hierarchy process data are illustrated in table 6.

### Table 5. Scheme Layer to the Criterion Layer Mutual Factor.

| Economic benefit B₁ | Plan T₁ | Plan T₂ | Plan T₃ | Plan T₄ |
|---------------------|---------|---------|---------|---------|
| Plan T₁             | 1       | 4/3     | 8/9     | 8/7     |
| Plan T₂             | 3/4     | 1       | 2/3     | 6/7     |
| Plan T₃             | 9/8     | 3/2     | 1       | 9/7     |
| Plan T₄             | 7/8     | 7/6     | 7/9     | 1       |

| Degree of risk B₂   | Plan T₁ | Plan T₂ | Plan T₃ | Plan T₄ |
|---------------------|---------|---------|---------|---------|
| Plan T₁             | 1       | 3/2     | 1       | 3/2     |
| Plan T₂             | 2/3     | 1       | 2/3     | 1       |
| Plan T₃             | 1       | 3/2     | 1       | 3/2     |
| Plan T₄             | 2/3     | 1       | 2/3     | 1       |

| Investment returns B₃| Plan T₁ | Plan T₂ | Plan T₃ | Plan T₄ |
|----------------------|---------|---------|---------|---------|
| Plan T₁              | 1       | 3/2     | 5/4     | 3/2     |
| Plan T₂              | 2/3     | 1       | 5/6     | 1       |
| Plan T₃              | 4/5     | 6/5     | 1       | 6/5     |
| Plan T₄              | 2/3     | 1       | 5/6     | 1       |

| Market effects B₄    | Plan T₁ | Plan T₂ | Plan T₃ | Plan T₄ |
|----------------------|---------|---------|---------|---------|
| Plan T₁              | 1       | 1       | 2/3     | 1       |
| Plan T₂              | 1       | 1       | 2/3     | 1       |
| Plan T₃              | 3/2     | 3/2     | 1       | 3/2     |
| Plan T₄              | 1       | 1       | 2/3     | 1       |

### Table 6. Data Mining Result of GRA-AHP.

| Criteria layer B     | Economic benefit B₁ | Degree of risk B₂ | Investment returns B₃ | Market effects B₄ | Total sort weight |
|----------------------|---------------------|-------------------|-----------------------|------------------|------------------|
| Criteria layer weight value | 0.2569610          | 0.2055688         | 0.1777249             | 0.3597453        |                  |

| Scheme layer single sort weight | Plan T₁ | Plan T₂ | Plan T₃ | Plan T₄ |
|-------------------------------|---------|---------|---------|---------|
| Plan T₁                       | 0.2290909   | 0.2     | 0.1904762 | 0.2727273 | 0.2319459       |
| Plan T₂                       | 0.3054545   | 0.3     | 0.2857143 | 0.2727273 | 0.2890514       |
| Plan T₃                       | 0.2036364   | 0.2     | 0.2380952 | 0.1818182 | 0.2011641       |
| Plan T₄                       | 0.2618182   | 0.3     | 0.2857143 | 0.2727273 | 0.2778386       |

Among the total ranking weights, the scheme T2>T4>T1>T3, which means T2 is the best investment plan and should be adopted or at least taken more consideration than others; T4 is the second appropriate selection, which can be regarded as an alternative for decision makers' reference; while T3 and T4 are relatively poor, they should be excluded. (To better present the data results and highlight the differences in the total ranking weights of the various schemes, the form of the decimals is restored here. In addition, after conducting GRA-AHP, the obtained value of CR is less than 0.1, which means it passed the consistency test and the results are reliable)
4. Conclusions
Data mining and information technology can greatly improve the effectiveness and efficiency of decision-making in the real estate investment plans. As a specific method of data mining technology, GRA-AHP exerts positive effects on the development administration of real estate industry, which is worthy of wide application.

Applying GRA-AHP helps us to comprehensively consider various factors affecting real estate investment decision-making, and organically combine qualitative and quantitative analysis, which can fully reflect the ambiguity of evaluation factors and evaluation process while minimizing the drawbacks caused by individual subjective judgments. Since the data mining technology is conducive to making full use of the data, the GRA-AHP can reflect the real estate development investment situation more realistically, facilitate the horizontal and vertical comparison of comprehensive evaluation, and provide a technical mathematical model for real estate development management. The analysis results based on data mining technology are more objective and practical than those ordinary evaluation methods such as scoring. Therefore, compared with previous evaluation methods, the data mining technology is much more scientific, thus the evaluation model established in this research are far more reliable and valuable.

Acknowledgments
The research is supported by Guangdong University of Foreign Studies Featured Innovation Program (project number: 299-GK18GS53) and National Natural Science Foundation of China (71673064), Innovation team project (Humanities and Social Science) of universities in Guangdong, China(2017WCXTD003).

References
[1] Luo Jing, Zhu Jianfeng. Empirical analysis of real estate price based on data mining and spatial measurement. Research of Development. 2013, (5), p 128-131
[2] Sun Shanshan. Real estate price prediction based on data mining. Modern Electronic Technique. 2017,40(5), p 126-129
[3] Jovanovic P. Application of sensitivity analysis in investment project evaluation under uncertainty and risk. International Journal of project management. 1999, 17(4), p 217 -222
[4] Borgonovo E, Peccait L. Uncertainty and global sensitivity analysis in the evaluation of investment projects. Journal of Production Economics. 2006, 104(1), p 62-73
[5] Kollmann T, Kuckertz A. Evaluation uncertainty of venture capitalists' investment criteria. Journal of Business Research. 2010, 63(7), p 741-747
[6] Ke Xiaoling, Zhu Kejun, Diao Fengqin. Improved model for fuzzy comprehensive evaluation of real estate investment risk. Reform and Strategy. 2010, 26(4), p 43-45
[7] Zhang Ping, Ma Li, Zhai Bowen. The grey fuzzy comprehensive evaluation of real estate investment risk. Journal of Technical Economics and Management, 2018, (4), p 22-25
[8] Zhou Xiang, Jiang Genmou. Risk analysis of construction project based on combination weighting and improved grey interrelated. Construction Management and Modernization. 2018,103(6), p 78-81
[9] Saaty T L. How to make a decision: the analytic hierarchy process. European Journal of Operational Research,1990,48(1), p 9-26
[10] Deng Xue, Li Jiaming, Zeng Haojian, Chen Junyang, Zhao Junfeng. Research on computing method of AHP weight vector and its applications. Mathematics in Practice and Theory. 2012,42(7), p 93-100
[11] Liu Zhaoming, Han Bin. Analysis on regional venture capital environment based on AHP fuzzy comprehensive evaluation. Journal of Southwest Jiaotong University. 2005, 40(3), p 379-384
[12] Sun Haitao. Fuzzy optimization model of real estate investment project based on AHP and its Application. Enterprise Economy. 2011,369(5), p 167-170