Analysis of Association Rules of Investment Risk and Accounting Based on Artificial Intelligence Model

Xuanjun Chen

Abstract In recent years, with the rapid development of artificial intelligence technology, data mining technology has been used in various fields to analyze massive data and obtain valuable information. In the financial field, investment risk exists all the time. In order to better analyze the relationship between investment risk and accounting, this paper proposes an association rule analysis method based on artificial intelligence model. Aiming at the fact that the traditional AI model is disturbed by inaccuracy factors and the association rule analysis method cannot deal with financial data well, this paper studies the association rule analysis method in depth, improves the pruning process based on IUA algorithm, and proposes an interactive association rule mining algorithm HIUA based on Hash structure. By comparing and testing the performance of each algorithm, it is proved that the proposed algorithm greatly improves the mining efficiency.

Keywords Artificial intelligence · Data mining · Investment risk · Association rule analysis

1 Introduction

With the advent of the era of big data and artificial intelligence, the rapid development of science and technology has provided more effective, diversified and combined analysis tools for data analysis and risk assessment, and made it possible to provide more accurate and timely investment risk analysis. As one of the current research hotspots, association rule analysis method is widely used in finance, medicine, transportation, education and many other fields [1–3]. From the current application results, the further study of association rules analysis method is helpful to solve various practical problems encountered in the actual enterprise management, which has important theoretical value and practical significance.

X. Chen (✉)
Department of Construction Management and Real Estate, Chongqing Jianzhu College, Chongqing 400072, China
e-mail: 535023045@qq.com

At present, the research of association rule analysis method has received extensive attention in various fields. For example, in [4], the author proposes an association rule analysis method, which can generate rules related to the correlation ratio and limit the number of final rules even at very low support level (0.5%). This method has been used in case-crossover studies to explore the effects of various air pollutants on the risk of childhood asthma. In [5], the author applies association rule analysis to the construction field and uses association rules to mine the relationship between LEED credits. In [6], the author proposes a new association rule analysis method and applies it to the retail industry. The algorithm can help retailers make correct decisions by mining customer behavior, thus greatly improving sales volume. In [7], the author uses three different methods to analyze the landslide sensitivity map and verifies that the association rule analysis method has better performance than other methods through comparative experiments. In [8], the author identifies the causes of accidents in major casualties by using association rule analysis technology. The analysis of association rules reveals the causes of serious casualties in different situations. In [9], the author proposes an association rule analysis method based on FP growth algorithm, which is mainly used to find the relationship between different urban attractions. In [10], the author combines genetic algorithm and association rule analysis method and applies it to construction industry. The results show that this method provides a powerful tool for analyzing useful knowledge in historical defect cases.

In this paper, many association rule analysis algorithms cannot solve the problem of frequent itemsets without changing the confidence threshold. The pruning process of the original IUA algorithm is improved based on the reduction of support threshold, and an association rule analysis method based on artificial intelligence model is proposed.

2 Method

2.1 Relevant Concepts

Investment risk

In the process of investment, all kinds of risks may be faced at any time. Investment risk generally refers to the deviation between expected return and actual return in the process of investment. It has a complex relationship with many influencing factors such as market fluctuation, time change and so on. How to analyze the relationship between investment risk and accounting variables is of great significance to investors and managers.

Accounting information

Accounting information is obtained by the use of specific technology or software by professionals engaged in the field of accounting to sort out the financial information of enterprises. It can often accurately reflect the financial data of an enterprise and through these data can reflect the operation of the enterprise. For enterprise managers, it can provide better reference for enterprise project decision-making. For investors, it enables investors to judge whether to invest through data as much as possible. In order to reduce investment risk better, this paper chooses important accounting variables from accounting information and takes accounting variables as independent variables and investing risk for association rules analysis.

2.2 Association Rules Analysis

Association rule analysis was originally proposed in the retail industry to solve the shopping basket problem. By mining valuable information from a large number of transaction data, it can provide targeted marketing strategies. The relevant definitions of association rule analysis are as follows:

Definition 1: Assuming that $I = \{I1, I2,…, Ik\}$ is the set of all data items, if there are $X \subset I$, $Y \subset I$, then X, Y are the set of data items, and if the number of items is k, it is called the set of K items.

Definition 2: Assume that the whole project set of the target transaction library is $I = \{I1, I2,…, Ik\}$, $D = \{T1, T2,…, Tn\}$ and that $Ti = \{Ii1, Ii2,…, Iik\}$, and that in Ti, any element $Iij(j \in [1, k]) \subseteq I$ is called a transaction in the database.

Definition 3: Call $A \Rightarrow B$ an association rule, in which A, B must satisfy $\{A, B|A \subset I, B \subset I, A \cap B = \phi\}$ at the same time, A is the precondition of the rule, B is the latter.

Definition 4: If the s% transactions in the target transaction library contain items A and B, the s% is called support of $A \Rightarrow B$.The formula is as follows:

$$\text{Support}(A \Rightarrow B) = P(A \cup B) = s\% \tag{1}$$

Definition 5: Call c% $A \Rightarrow B$ confidence, if c% of all transactions in the target transaction library containing itemset A also include B. The formula is as follows:

$$\text{Confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} = c\% \tag{2}$$

Definition 6: $(X \Rightarrow Y)$ is called strong association rule. If there are $(X \Rightarrow Y)$'s support degree and confidence degree c, they are greater than or equal to the preset minimum support degree and minimum confidence degree, respectively.

2.3 Association Rule Analysis Based on Artificial Intelligence Model

In all kinds of association rule analysis methods, there are many algorithms that cannot solve the problem of changing confidence threshold without changing frequent itemsets. In this paper, a new association rule analysis method based on IUA algorithm is proposed, considering only the reduction of support threshold. Aiming at the possible omission of some frequent patterns in the original IUA algorithm, this paper improves the pruning process of the IUA algorithm and proposes an interactive association rule mining algorithm HIUA based on Hash structure. By counting the support degree of the pattern obtained during the execution of the Hash structure preservation algorithm, the mining efficiency is improved. The specific flow of the algorithm is as follows:

(1) For the first time, the algorithm will degenerate into Apriori algorithm to compute frequent itemsets and preserve frequent itemsets of hierarchical data under corresponding thresholds.
(2) Secondly, the results of step (1) are filtered to get new frequent itemsets of hierarchical data while the support threshold rises.
(3) Then, when the support threshold decreases, the original frequent itemset is LOrg1, and the new frequent itemset L1 under the new threshold is calculated, and the new frequent itemset LNewl that does not belong to the original frequent itemset is further obtained.
(4) Finally, IUA GenLk functions are self-connected to get their respective Ck. The new support threshold Lk is obtained by screening the Ck, until the Lk is the end of the space-time algorithm.

3 Experiment

In the course of the experiment, because the financial data of enterprises are generally not easy to obtain, and there are common problems in the accounting investment risk of domestic listed companies, the ST shares of domestic listed companies originating from the financial terminal software of Oriental Wealth CHOICE are selected as the research data.

4 Result

Result 1: Performance test comparison of each algorithm.

In order to verify the effectiveness of the proposed method, the traditional association rule analysis algorithms Apriori and IUA are selected for comparative experiments. Based on the fin ratio of the enterprise's actual financial data set, 720 records
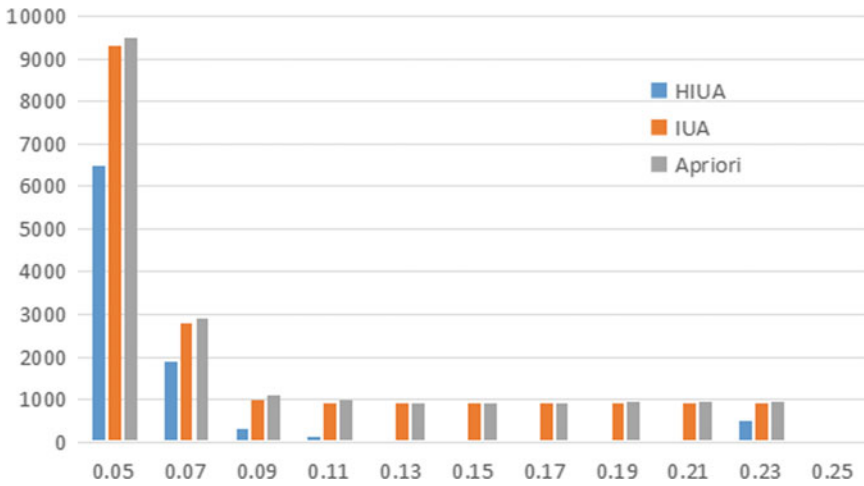
Fig. 1 Performance comparison results of each algorithm

were selected for various financial indicators, such as profitability, operating capacity, cash flow and so on. The abscissa support threshold is set to 0.02 step size, from 0.25 to 0.05, and the ordinate is to calculate the running time of frequent pattern sets. The results of performance test are shown in Fig. 1 by extracting rules from the pattern set for correlation analysis of investment risk and accounting.

As can be seen from Fig. 1, the running time of HIUA at the first execution is similar to that of IUA, but the efficiency of HIUA is obviously improved compared with that of IUA as the saved support counts are continuously used in the running process. Compared with Apriori algorithm, HIUA has higher efficiency. When the support threshold decreases, the local frequent pattern set may not change much, and the candidate frequent pattern set Ck of IUA and HIUA will be smaller, and some patterns can be found in the HIUA hash table. Therefore, the association rule analysis method proposed in this paper has good performance and certain practical significance.

Result 2: Analysis of correlation between beta coefficient of investment risk and accounting variables.

By using statistical analysis software, SPSS to analyze some of the results obtained by the previous association rule analysis method, the dependent variable is selected to measure the important parameter beta coefficient of investment risk, and the independent variable is each accounting variable. The statistical results are shown in Table 1.

From Table 1, we can see that operating leverage ($X1$), dividend payment rate ($X4$) and growth rate of main business income ($X5$) are not completely consistent with the assumptions during the sample period. They are not particularly stable. They are positively correlated in some years and negatively correlated in some years. Among them, operating leverage ($X1$) is negatively correlated in some years, but it is not

Table 1 Relevance analysis of β coefficient and accounting variables

Index code	Accounting variables	2015	2016	2017	2018
X1	Operating leverage	0.035	−0.091	0.072	0.025
X2	Financial leverage	0.166	0.096	0.102	0.211
X3	Company size	−0.332	−0.349	−0.681	−0.467
X4	Dividend payment rate	0.029	−0.178	−0.244	−0.165
X5	Increase rate of main business income	0.049	0.075	−0.023	0.014

significant. It may be caused by some investors' behavior. Dividend payment rate ($X4$) is positively correlated in some years, but in some years, it is opposite, and its performance is not stable. It may be because the dividend distribution policies of listed companies are not consistent, and the dividend number of specific investors is not certain, and the volatility is relatively large, so the relationship between the two will be unstable, and there will be fluctuations in some years.

5   Conclusion

With the development of science and technology, artificial intelligence and data mining play an increasingly important role in the financial field. Among them, the analysis of association rules between investment risk and accounting provides great help to business managers, investors and the government. After introducing the concepts of investment risk and accounting, this paper introduces the association rule analysis method, improves the pruning process based on the common IUA algorithm, and proposes the association rule analysis method based on artificial intelligence model. The experimental results show that the proposed method can effectively solve practical problems and has certain feasibility and rationality.

References

1. Qin, W., Z. Kangkang, and C. Huiting. 2018. Association rule mining for the talents introduction strategy: a case study of Zhejiang University of Finance & Economics. *American Journal of Applied Mathematics* 6 (2): 55–61.
2. Khader, N., A. Lashier, and S.W. Yoon. 2016. Pharmacy robotic dispensing and planogram analysis using association rule mining with prescription data. *Expert Systems with Applications* 57: 296–310.

3. Ahmed, M., M.T. Imtiaz, R. Khan, and R.M. Rahman. 2018. Clustering and association rule mining-based traffic analysis and prediction of Dhaka. *International Journal of Knowledge Engineering and Data Mining* 5 (4): 241–276.

4. Toti, G., R. Vilalta, P. Lindner, B. Lefer, C. Macias, and D. Price. 2016. Analysis of correlation between pediatric asthma exacerbation and exposure to pollutant mixtures with association rule mining. *Artificial Intelligence in Medicine* 74: 44–52.

5. Ma, J., and J.C. Cheng. 2016. Data-driven study on the achievement of LEED credits using percentage of average score and association rule analysis. *Building and Environment* 98: 121–132.

6. Kaur, M., and S. Kang. 2016. Market basket analysis: identify the changing trends of market data using association rule mining. *Procedia computer science* 85: 78–85.

7. Erener, A., A. Mutlu, and H.S. Düzgün. 2016. A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). *Engineering Geology* 203: 45–55.

8. Xu, C., J. Bao, C. Wang, and P. Liu. 2018. Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. *Journal of Safety Research* 67: 65–75.

9. Fang, H., C. Chen, J. Lin, X. Liu, & D. Fang. 2017. Association rule analysis for tour route recommendation and application to Wctsnop. ISPRS-international archives of the photogrammetry. *Remote Sensing and Spatial Information Sciences*, 1121–1126.

10. Cheng, Y., and Q. Li. 2015. GA-based multi-level association rule mining approach for defect analysis in the construction industry. *Automation in Construction* 51: 78–91.