Hindawi

Retraction

Retracted: Discussion on Redundant Processing Algorithm of Association Rules Based on Hypergraph in Data Mining

Journal of Robotics

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Zhu, "Discussion on Redundant Processing Algorithm of Association Rules Based on Hypergraph in Data Mining," *Journal of Robotics*, vol. 2022, Article ID 1007464, 9 pages, 2022.

Research Article

Discussion on Redundant Processing Algorithm of Association Rules Based on Hypergraph in Data Mining

Jintan Zhu

School of Electronic Information, Xi'an Railway Vocational and Technical College, Xi'an City, Shaanxi Province 710014, China

Correspondence should be addressed to Jintan Zhu; zjt_19810513@163.com

With the rapid advancement of big data, it is becoming a great problem for people to find objective information in the database. The relevance data processing rule for digging the information can be the way. Relevance data processing rule for digging the information is mainly studied in three aspects: data dimension, data abstraction level, and processing variable type. In the aspect of rules, the research mainly focuses on three aspects: active relationships, passive rules, and uncommon relationship rules. Association rules of digging the data can be the most well-employed investigation goal and aim for the data digging. Along with the advancement of the information scale, the time rate of traditional relationship rules exploration of counting ways is too low. How to increase the time rate for the way of counting is the main research content of relevance data processing rules for digging the information. Current relevance data processing rules for digging the information have two limitations: (1) metrics such as support, confidence, and lift rely too much on expert knowledge or complex adjustment processes in value selection; (2) it is often difficult to explain rare association rules. Based on the existing research, this paper proposes a Markov logic network framework model of association rules to address the above shortcomings. The theory of a hypergraph and system is proposed, and the method of a hypergraph in 3D matrix modeling is studied. Aiming at the new characteristics of big data analysis, a new super edge definition method is introduced according to the definition of the system, which greatly enhances the ability to solve problems. In the cluster analysis and calculation of hypergraph, this paper will use the hypergraph segmentation operator hMETIS to carry out the cluster analysis method in order to achieve higher accuracy in cluster analysis and calculation. As for the test of cycle ones, which is in line with the relevance of the hypergraph with clear directions, the thesis will offer a brand-new way to make an analysis and turn the rule of relevance into the hypergraph with clear directions with a new definition of the near linking matrix, and it will change the dealing way from the test of the cycle and more ones into the linking *f* bricks and circles, which is a new way to explore. This paper uses two datasets of different sizes to conduct rule prediction accuracy experiments on the Markov logic network framework model algorithm of association rules and the traditional association rule algorithm. The results show that compared with the traditional association rule algorithm, the rules obtained by the Markov logic network framework model of association rules have a higher prediction accuracy.

1. Introduction

Since the 21st century, Internet technology and computer hardware technology have developed rapidly and become popular [1]. The data stock of the Internet has increased exponentially, but there are only a handful of data containing valuable information [2]. Therefore, mining valuable information from massive data has changed into the most essential factor or the content in society at present [3]. For the old and ancient dataset digging information skills, the relationship rules which are unseen from the data will be analyzed with an effective rate, what is more, the result of the old-fashioned data digging skills and it will cause the results of massive data because it cannot predict the things happened in the long future [4]. The appearance of the AI will give more skills to the information-digging process. And the big process will be the digging of the data information, which has created the message of big data [5]. At present, data mining has an important application value in many aspects. By describing the existing data, it can effectively predict the

future pattern of the data. Data mining technology has become a promising research direction in today's world. Data mining technology has been applied in scientific research, production process monitoring, quality management, decision support, and design.

The Internet, the Internet of Things, cloud computing, and other information technologies are constantly updated, and they are constantly integrated with the human world in various fields such as economy, politics, military, scientific research, and life [6]. Information visualization is the use of images to express information clearly and effectively. The image representation can be used to find the original data and the information relationship that cannot be observed. Data visualization can enhance users' understanding of multidimensional and large-scale information and plays an important role in the discovery, determination, and understanding of association rules. As a major information discovery and pattern recognition technology, the association processing law of mining data is to find the most meaningful information that can be described. The visualization of association rules is an inseparable subset of association rule theory. Its main task is to display information and help users further grasp the rules of association processing in order to discover the information results [7].

Data mining can discover the hidden laws in the data and effectively exert the value of the data [8]. The relevance data processing rule for digging the information can extract potential and valuable frequent patterns or correlations between attributes from the data [9]. Frequent patterns and correlations can be displayed clearly and intuitively in the form of text, but due to the limited cognitive ability of users, the value of relevant data processing rules for digging the information cannot be fully reflected [10]. Therefore, it is urgent to study the visualization method of association rules in depth, combined with human-computer interaction technology, to help users analyze and process data resources from multiple perspectives, gain insight into valuable information, and support their decision-making and planning [11].

The hypergraphs are widely used in many fields of information science. In the past, information visualization technology and visual data analysis technology mainly focused on analyzing simple binary signals inside data objects [12]. However, studies have shown that multiple associations can more naturally represent the sum pattern of internal connections implicit in signals [13]. A Hypergraph is a generalization of ordinary topological relations and can be easily expressed as multiple relations [14]. This also provides strong conditions and theoretical support for the visualization of association rules. The hypergraph model combines the characteristics of hypergraphs and multidirectional graphs and can visually describe association rules. In the graph, nodes represent data items and edges represent association relationships. The support and reliability of rules can also be described in different ways and with different values. Therefore, the intuitive display of multirelationships by the hypergraph provides strong theoretical support for further in-depth research on visualization methods of frequent item sets and association rules.

2. State of the Art

2.1. Data Mining Overview and Research Status. In 1990, the first KDD International Seminar was held in Detroit, USA. Since then, people have shown interest in data mining terminology [15]. Information is used by people to describe specific events in actual society. This is an abstract description of the information. With the development and progress of the times, there are more and more aspects of human exploration, and digital has become a tool to support human exploration. Indispensable means that due to the development of science and technology, human beings are exploring the physical universe more and more extensively, which makes the range of numbers more and more extensive, and the complexity of data is rapidly increasing [16]. In this case, people can no longer find hidden laws through simple logical reasoning. Therefore, people began to pay attention to the importance of data and eagerly hoped to find the value and meaning hidden behind the data. It is precisely to meet this demand that data mining technology was born. Mining, also known as information mining, is the process of discovering potential, meaningful, and interesting things from a large amount of incomplete and noisy information stored in large transaction databases or data warehouses. The information mined can often help us to conduct in-depth exploration. Of course, it is not a knowledge exploration method that can retrieve information at any time. Therefore, through the search engine to find the web pages and search databases that you are interested in, records cannot become knowledge discovery. These methods are only looking for information that meets specific conditions, but they do not explore the things behind big data. Data mining is not a panacea. That is, what is found in large transaction databases is not always correct or valuable [17]. This needs to meet special commercial conditions. People study business and do statistical analysis. Under these premises, the use of information mining is more likely to mine valuable and instructive messages [18].

In the procedure of digging the data will be decided by requests from businesses and data features [19]. But it will be classified as the changing data, preprocessing of data, data mining, and knowledge assessment steps. The pre-processing of data will deal with the data and transform it. It costs a lot of time to deal with the data [20]. The image of the data preprocessing process is shown in Figure 1.

Data mining technology originally came from abroad, but its research and development direction are varied. At present, the most common way to deal with analysis problems is decision tree induction. The corresponding calculation methods are C4. 5 calculation, ID3 algorithm, ID4 algorithm, IDS algorithm, and quest calculation. Complex structured data learning, the slio algorithm, the sprint algorithm, and "rainforest" calculations are used for building a decision tree. Both emphasize the establishment of a decision tree with scalability. The decision tree pruning algorithm includes cost complexity pruning, error reduction pruning, and pessimistic evaluation pruning. Some methods, such as Bayesian classification, the back propagation algorithm, the neural network
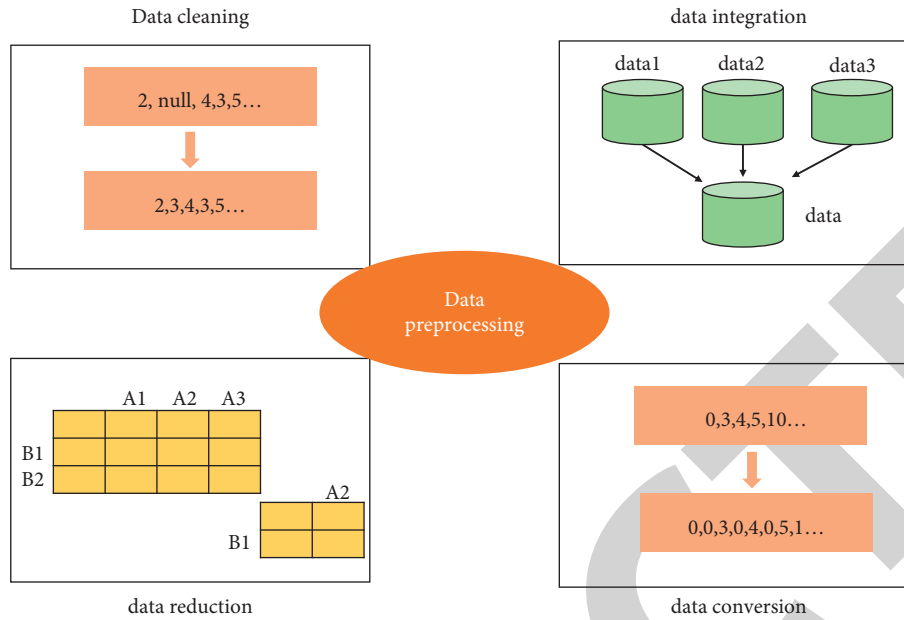
Figure 1: Data preprocessing process.

method, the machine learning method, the CAEP classification method, and the rough set method, are all applied to analysis and data mining. There are also many data mining techniques and methods to deal with data clustering. Common systems divide clustering methods into the k-means method and the basic condensed hierarchical clustering method. DBSCAN is a clustering algorithm based on density, while optics is a clustering algorithm based on density.

2.2. Research Status of Association Rules.

Up to now, large-scale exploration has been carried out in the field of data mining, but the research in the field of data mining is still popular. That is, although the current mining method is quite perfect, its data mining efficiency is very low in the face of large-scale information. The result is not ideal. Therefore, the information accumulated by the network e-commerce industry is extremely rich and complicated, resulting in a large number of useless bits of information and garbage materials. For such a large amount of information, the preprocessing steps of data mining will be very difficult. Once the pretreatment is poor, the whole mining process may even fail. At the same time, although the preprocessing process is relatively smooth, conventional mining methods may not obtain valuable and meaningful data in a large number of databases, even if the mined data are useless and meaningless. All these show that there are still many problems in data mining. Considering the differences in the main research directions and mining methods of data mining, there are still many challenging research topics in the field of data mining applications. These topics are closely linked, mainly involving information fusion technology of information discovery and data warehouses, visual data mining, super large-scale data mining of complex types, network data mining, and

network security technology. The visualization technology of a two-dimensional matrix is usually used to represent the features on the bar graph. The items in the front part and the rear part are arranged on two axes in turn. The width and color of the bar in turn represent the support and confidence, as shown in Figure 2.

2.3. Hypergraph Overview.

A hypergraph is a subset system of a finite set, which is a generalization of graph theory and plays a very important role in discrete mathematics. The term "hypergraph" was first proposed by Berge in the monograph "Hypergraphs" in 1966. The original purpose was to promote some classical results in graph theory. Later, people gradually realized some theorems in graph theory. It can be generalized in the unified form of a hypergraph, thus opening the prelude to the study of hypergraph theory, making it a huge new branch of graph theory. Compared with the study of general graphs, the study of hypergraphs is more complicated. Some important structures and properties in general graphs no longer exist in hypergraphs, which complicates the discussion of many similar problems in graph theory. At present, hypergraphs have been widely used in circuit division, knowledge representation and organization methods, cellular communication systems, and the representation of molecular structures of atypical compounds and polycyclic conjugated molecules. There are two types of hypergraphs: directed hypergraphs and undirected hypergraphs. Since the 1960s, after decades of unremitting efforts, the development of hypergraph theory has made great progress.

A hypergraph is a binary pair $H = (V, E)$, where $V = \{v_1, v_2, v_3, \ldots v_n\}$, denotes the n vertices of the hypergraph, and $E = \{e_1, e_2, e_3, \ldots e_n\}$, denotes the m hyperedges of the hypergraph. A hyperedge set E is a subset defined on a vertex set V, that is $\forall e_j \subseteq V, \quad j = 1, 2, \ldots m$, and satisfies
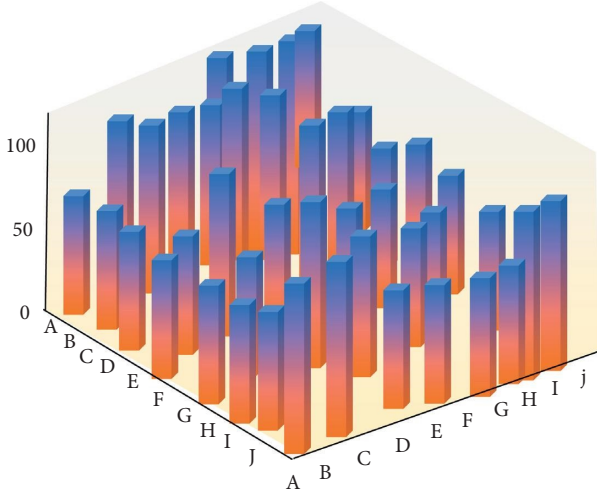
Figure 2: Visualization is based on a two-dimensional matrix.

$$e_j \neq \varnothing, \quad j = 1, 2, \ldots m,$$

$$\sum_{i=1}^{m} e_j = V. \tag{1}$$

The size of a graph is generally uniquely determined by the number of vertices N and the number of super edges M. In hypergraph, its size also depends on the cardinality of each super edge. We can define the size of a hypergraph as the cardinality of each hyperedge, and it is given using the following equation:

$$\text{size}(H) = \sum_{e_i \in E} |e_i|. \tag{2}$$

As an important branch of hypergraph theory, directed hypergraph theory has not been studied deeply for a long time, and the results are relatively small. The foreign paper that has a far-reaching impact on the development of directed hypergraph theory is directed hypergraph and its application by Giorgio Gallo. Here, the author systematically summarizes the previous achievements in the field of directed hypergraphs. In China, based on the demand for electrical engineering and automation research, Professor Huang has provided another new way to describe directed hypergraphs and has done relevant research work on this basis. Similar to the theoretical study of undirected graphs and directed graphs, directed hypergraphs add directions to the super edges in undirected hypergraphs, and then describe the arrangement of the vertices of the super edges. On this basis, we introduce some properties of undirected hypergraph into directed hypergraph and find the special properties of directed hyper-tree. Of course, the main purpose of studying directed hypergraphs is to solve problems in practical applications. In this paper, people try to describe the relevant laws through the directed hypergraph and solve the redundancy and circulation problems in the relevant data processing rules by using the attributes in the directed hypergraph, and mining information.. Figure 3 is a representation of a directed hypergraph H.
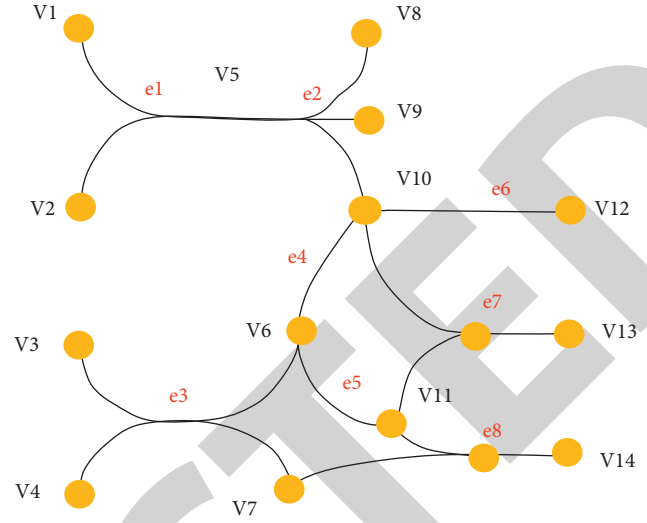

Figure 3: Representation of a directed hypergraph H.

A directed hypergraph is $\overline{H} = (V, E)$, V is the vertex set, and E is the directed hyperedge set. Its adjacency matrix is given as follows:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

$$a[i, j] \in \{0, 1, 2, \ldots\ldots\}, \tag{3}$$

aij represents the number of directed hyperedges from the start point 1 to the end point 1, where v_i and v_i are the vertices of $\overline{H} = (V, E)$.

Like an undirected hypergraph, the size of a directed hypergraph is defined as the sum of the cardinality of each super edge, and the rank (low rank) is also defined as the largest (small) multiple of the cardinality of each super edge. However, it is difficult to form a one-to-one correspondence between the adjacency matrix and other matrices in a directed hypergraph due to the characteristics of clusters in the hypergraph itself. People can also understand that the directed hypergraph corresponds to the adjacency matrix, but people cannot uniquely recover the original directed hypergraph from the adjacency matrix. Therefore, how to reduce these factors is also an important topic in the study of directed hypergraph theory.

3. Methodology

3.1. Relevance Data Processing Rule for Digging the Information Process. The purpose of a relevance data processing rule for digging the information is to find some credible rules from massive data. These rules usually have potential value and significance and can help enterprise managers analyze the current market situation and make correct decisions. The relevance data processing rule for digging the information system searches for association rules based on two minimum thresholds, which are the minimum support threshold min sup and the minimum confidence threshold min conf, which

can usually be specified by the user. The relevance data processing rule for digging the information work is mainly divided into two stages: (1) Find all item sets not less than the minimum support threshold min_sup, that is, frequent sets; (2) For each frequent set, find not less than the minimum confidence threshold. Association rules for min_conf. The process of mining association rules in stage 2 is as follows: there is a frequent set L. For any proper subset $L' \subset L, L' \neq \emptyset$ in the itemset L, if Support (L)/Support (L') 2 min_conf, the association rule $L' \Rightarrow L - L'$ is a credible rule.

The following uses an example to illustrate the specific process of relevance data processing rules for digging the information. Suppose all items in database D constitute an itemset, and all records in transaction database D are shown in Table 1:

Looking at Table 1, it is easy to calculate all 1-, 2-, 3-, and 4-item sets. The four item sets and their corresponding support information are shown in Table 2.

Given two thresholds, support threshold min_sup = 30% and confidence threshold min_conf = 70%, then, according to Table 2, all frequent sets have {a}, {b}, {c}, {d}, {a, b}, {a, c}, {a, d}, {c, d}. For each frequent set whose number of item sets is greater than 1, association rules can be mined from it. For example, two rules {a} = {b} and {b} > {a} can be mined from the frequent set {a, b}. Examining each such frequent item set, we can get all the association rules as shown in Table 3:

Since the confidence threshold min_conf = 70%, according to Table 3, all the strong association rules have {a} = {b} and {b} = {a}. At this point, the mining of association rules is over.

3.2. Apriori Algorithm.

The Apriori algorithm is the most classic algorithm for relevance data processing rules for digging the information. The algorithm is easy to understand in the principle of the algorithm, and concise and convenient in the realization of the algorithm. The principle of the Apriori algorithm is that the more times the two items appear in pairs in the transaction data, the greater the correlation between the two items, and the two items are items with a strong correlation. The implementation process of the Apriori algorithm needs to scan the database multiple times. The first scan counts the number of occurrences of each item in the data and deletes some items that do not meet the minimum support requirements. Before the second scan, the items obtained from the first scan should be combined in pairs. Then, the second scan of the data is performed to count the occurrences of the combination, and some combinations that do not meet the minimum support are deleted. After that, the combination and scan are repeated until no new combination is generated, and the implementation process of the Apriori algorithm ends. Figure 4 below shows the implementation process of the Apriori algorithm when the minimum support is 0.2.

3.3. The Redundant Processing Method of Association Rules Based on Hypergraph.

The rule dug up from the relationship rule will include some projects that users do not need. Or it is in line with the information that users are familiar with, or it

Table 1: Transaction database D.

TID	Phase set
T011	a, b, c
T021	a, c, d
T031	a, b
T041	c, d
T051	a, b, d

Table 2: All item sets and support.

Item sets	Support (%)
{a}	88
{b}	65
{c}	65
{d}	65
{a, b}	65
{a, c}	48
{a, d}	45
{b, c}	21
{b, d}	20
{c, d}	45
{a, b, c}	25
{a, b, d}	25
{b, c, d}	25
{a, b, c, d}	0

Table 3: All association rules.

Association rules	Confidence (%)
{a} {b}	77
{b} {a}	90
{a} {c}	55
{c} {a}	65
{a} {d}	51
{d} {a}	65
{c} {d}	65
{d} {c}	66

gets the same rule table that expresses the same meaning. Such rules will not bring more messages to users or offer any effective help. In most cases, the number of the redundant rule is larger than the meaning rule numbers.

Redundancy rules can generally include two types: one is subordinate tone rules. For example, the conclusion of rule Xi is consistent with that of XJ, and the premise of Xi meets the sufficient condition for the existence of XJ, that is, XJ is redundant. Therefore, repetition rules are also regarded as subordinate rules. Special circumstances. The second is the repeated path principle. If there are selectors Xi and XJ in the rule base at the same time, and there must be two paths between Xi and XJ, the principle of redundancy can be judged.

The subordination principle can be expressed by the following formula (4):

$$\begin{cases} X_2 & \longrightarrow & X_4, \\ X_2 X_3 & \longrightarrow & X_4. \end{cases} \tag{4}$$

Tid	Items
10	A,C,D
20	B,C,E
30	A,B,C,E
40	B,E

(a)

1stscan →

Itemset	sup
{A}	0.4
{B}	0.6
{C}	0.6
{D}	0.2
{E}	0.6

(b)

L_1 →

Itemset	sup
{A}	0.4
{B}	0.6
{C}	0.6
{E}	0.6

(c)

Itemset	Sup
{A,C}	0.4
{B,C}	0.4
{B,E}	0.6
{C,E}	0.4

(f)

← L_2

Itemset	Sup
{A,B}	0.2
{A,C}	0.4
{A,E}	0.2
{B,C}	0.4
{B,E}	0.6
{C,E}	0.4

(e)

← 2ndscan

Itemset
{A,B}
{A,C}
{A,E}
{B,C}
{B,E}
{C,E}

(d)

Itemset
{A,B,C}
{A,C,E}
{B,C,E}

(g)

← 3rdscan

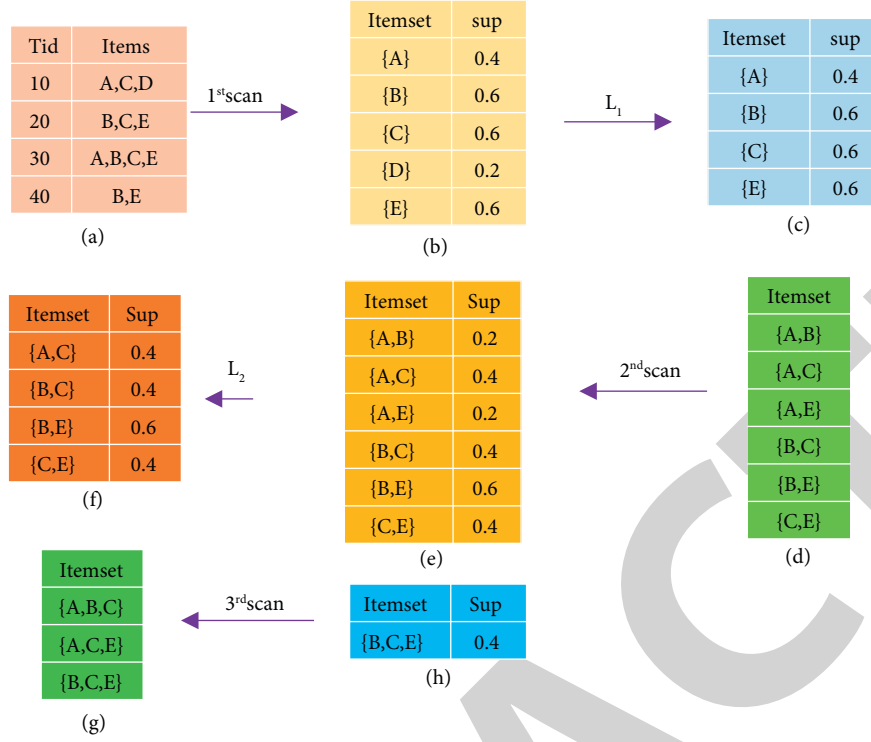| Itemset	Sup
{B,C,E}	0.4

(h)

FIGURE 4: Example diagram of the mining process of the Apriori algorithm.

The repeating path rule can be expressed by the following given formula:

$$\begin{cases} X_1 & \longrightarrow & X_2 X_3 & \longrightarrow & X_4, \\ X_1 & \longrightarrow & X_5 & \longrightarrow & X_4. \end{cases} \quad (5)$$

The adjacency matrix in a directed hypergraph completely illustrates the adjacency problem between nodes of a graph. In the directed hypergraph based on association rules, the association between the checked items and the association rules can be expressed by the adjacency matrix. Based on the concept of redundancy rules in circuit science and its related characteristics, the redundancy checking method can be realized on the basis of the directed hypergraph.

$$\text{design } G = (V(G), E(G)). \quad (6)$$

Its path is a finite, nonempty sequence, and it is given as following:

$$W = v_0 e_1 v_1 e_2 \ldots e_k v_k. \quad (7)$$

That is, the staggered sequence of vertices and edges, where $e_i \in E(G)$, $v_j \in V(G)$, e are associated with V_{i-1}, v_i, respectively, $1 \le i \le k$, $0 \le j \le k$ is denoted as (v_0, v_k) path, and the vertex v_0, v_k is called the starting point and end point of the path W, respectively. It is called a path $v_1, v_2, \ldots, v_{k-1}$ is the inner vertex of W, k will be named the long W.

From the graph theory, the procedure of dealing with the redundant ones with the hypergraphs into the linking bricks and also it will change it into formative tress, because every line in the hypergraph represents a linking rule. When the linking picture will become the formative trees, we need remove it, and this line is the redundant rule.

Given a hypergraph is as follows:

$$\begin{aligned} H &= (X), \\ X &= \{X_1, X_2, \ldots \ldots, X_n\}, \\ E &= \{E_1, E_2, \ldots \ldots, E_m\}. \end{aligned} \quad (8)$$

If H is connected and does not contain any hyperloops, then H is called a hyper-tree.

4. Result Analysis and Discussion

4.1. Model Establishment. In the Markov logic network, this paper regards the items in the transaction dataset as nodes in the Markov logic network. In this way, the weight between the two nodes can be regarded as the degree of association between the two items. In the Markov logic network, adding weights to the rules is adopted so that the knowledge base of the first-order predicate logic is not so rigid. The higher the weight value attached to the rule, the greater the restriction on each group in the Markov logic network. When the weights of all rules in the knowledge base are infinite, the Markov logic network is the same as the standard one. The logical reasoning framework for order predicates is the same.

To obtain the parameter value in the data, adjust the value of the data. The most common way to deal with modulus problems, such as infinite numbers, is the step reduction method. In computer mathematics, the ladder lowering methods generally include batch ladder lowering, small batch ladder lowering, random gradient lowering, and batch ladder lowering. The batch ladder lowering method is to calculate the lowest point (or the highest point along the
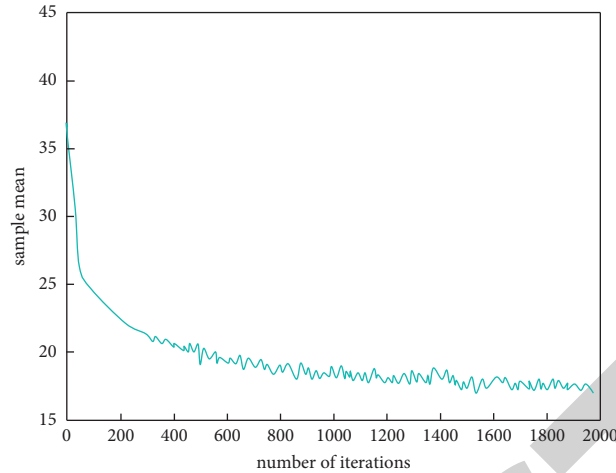
Figure 5: Convergence diagram of the Markov logic network model algorithm.

Table 4: Section BSAKETS data sheet.

Fruit veg	Fresh meat	Dairy	Canned veg	Canned meat	Frozen meal	Beer	Wine	Softer ink	Fish	Confectionery
1	1	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	0	0	0	0
0	1	0	0	0	0	0	1	0	0	0
1	0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0	1	1	0
0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	0
0	0	0	0	0	0	1	0	1	0	0
0	1	1	0	1	0	1	1	1	0	0

gradient rising trend) along the ladder falling trend during the calculation process, and the gradient direction change can be obtained by functional derivation. A training set is selected for improvement in each iteration. Different from batch ladder reduction, small batch ladder reduction refers to selecting local training samples for improvement in each iteration. Random gradient reduction refers to randomly selecting a training sample to improve in each iteration process.

4.2. Experimental Results and Analysis. To prove correctness, parameters in the Markov logic network framework with relationship rules are learned through real data to see if the model converges. This paper uses a dataset from a grocery store to learn the parameters of the Markov logic network framework model of association rules, and the data set contains 75 variables. There are a total of 3956 pieces of information. By using the stochastic gradient descent method to learn the parameters of the Markov logic network framework model of association rules, the convergence graph of the Markov logic network framework model of association rules is obtained. Figure 5 indicates that when the number reaches 1000, the algorithm begins to converge, indicating that the one logic net sample of the relationship rules is converged.
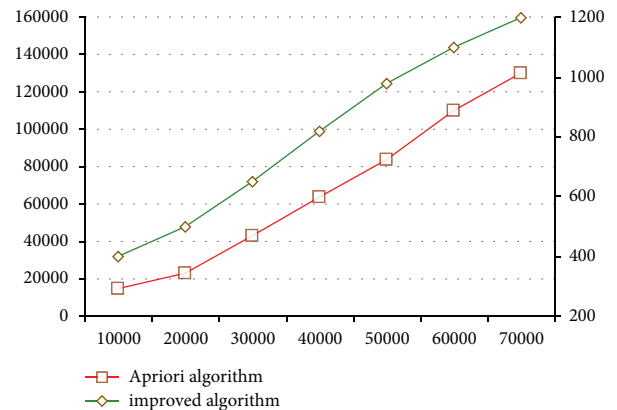

Figure 6: Comparison of the execution time of two algorithms with different numbers of records.

This chapter conducts an example analysis of a small dataset. The data is the data that comes with SPSS Clementine11.1, named BSAKETS1n. The data contains 18 fields and 1000 records, mainly including customer information, total purchase amount, and purchased items. The sighs include vegetables, fruit, meat, fish, and soft drinks. We use part of this data, the purchase item information, for association rule analysis. By processing the data, we get the data in Table 4 (including only part of the processed data).

This experiment continues to compare the time efficiency of the Apriori algorithm and the improved algorithm. The dataset used in the experiment is the basket data retail.dat file of a retail store in Belgium from the CSDN blog. The dataset has a total of 88,162 records and 16,470 binary attributes. The experimental test results are shown in Figure 6. As can be seen from the figure, the execution time of the improved algorithm has been greatly reduced compared with the Apriori algorithm.

5. Conclusion

In today's big data era, relationship rule mining can be a popular investigation direction. More and more people start to study association law and they are employed in many areas. In the research about association rules, people have obtained a lot of results and gained good consequences. But there are still some shortcomings. To address these problems, this paper proposes a Markov net model with relationship rules. Most of the current relevance data processing rules for digging up the information algorithms are constructed under a unified framework model. The main contents of this paper include the following aspects: (1) Combining Markov logic network and association rules, a new model, the Markov one framework of relationship rules, will be proposed. (2) Employ the random counting method of the Markov logic network framework. By comparing the prediction accuracy of the Markov logical network framework model and the traditional relevance data processing rule for digging the information, the Apriori algorithm, through different datasets, achieved higher accuracy than the traditional association rule algorithm (Apriori algorithm). (3) In line with the test about the relationship with the rule redundancy for the hypergraph, the author will change the relation rule into the hypergraph and give a new concept to the term about the matrix to offer a new way to the test of the circle and linking bricks in the hypergraph.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] S. Doganer and H. E. Akolu, "The effect of sports science students' social media adddictions on redundant purchasing behavior," *Asian Journal of Education and Training*, vol. 6, no. 4, pp. 24–30, 2020.

[2] H. Ge, H. Qin, S. Xue et al., "Research on denoising algorithm of thoracic impedance signal for respiratory monitoring during running exercise," *Biomedical Signal Processing and Control*, vol. 70, no. 5, Article ID 102941, 2021.

[3] H. Li, M. Qiao, and S. Peng, "Research on the recommendation algorithm of rural tourism routes based on the fusion model of multiple data sources," *Discrete Dynamics in Nature and Society*, vol. 2022, no. 11, Article ID 2262148, 125 pages, 2022.

[4] S. Maleki, U. Agarwal, and M. Burtscher, "BiPart: a parallel and deterministic multilevel hypergraph partitioner," *International Journal of Approximate Reasoning*, vol. 10, no. 8, pp. 88–95, 2020.

[5] S. Bashkari, A. Sami, and M. Rastegar, "Outage cause detection in power distribution systems based on data mining," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 99, p. 1, 2020.

[6] V. S. Klaus, S. C. Schriever, J. M. Monroy Kuhn et al., "Correlation guided Network Integration (CoNI) reveals novel genes affecting hepatic metabolism," *Molecular Metabolism*, vol. 53, no. 7747, 2021.

[7] N. Lakshmi and M. Krishnamurthy, "Frequent itemset generation using relevance data processing rule for digging the information based on hybrid neural network based billiard inspired optimization," *Journal of Circuits, Systems, and Computers*, vol. 31, no. 8, pp. 113–121, 2022.

[8] V. Soboleva and O. Shipitko, "Raindrops on windshield: dataset and lightweight gradient-based detection algorithm," *Springer International Publishing*, vol. 6, no. 5, pp. 30–45, 2021.

[9] N. Veldt, A. Wirth, and D. F. Gleich, "Parameterized objectives and algorithms for clustering bipartite graphs and hypergraphs," *Pattern Analysis & Applications*, vol. 6, no. 21, pp. 77–85, 2020.

[10] Z. An, G. A. Fei, and Y. A. Mi, "A new rule reduction and training method for extended belief rule base based on DBSCAN algorithm," *International Journal of Approximate Reasoning*, vol. 119, no. 20, pp. 20–39, 2020.

[11] N. Parashar, A. Tiwari, and R. K. Gupta, "Mining minimal non-redundant image association rules using statistical texture feature based frequent itemsets," *Evolutionary Intelligence*, vol. 15, no. 2, pp. 33–41, 2022.

[12] S. Mudgal and M. Parmar, "Mining of the correlations for fatal road accident using graph-based fuzzified FP-growth algorithm," *Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication-BEIESP*, vol. 10, no. 5, pp. 119–128, 2020.

[13] J. Chen, C. Zhao, and Uliji, "Collaborative filtering recommendation algorithm based on user correlation and evolutionary clustering," *Springer International Publishing*, vol. 10, no. 1, pp. 55–62, 2020.

[14] F. Chen, J. Liao, and Z. Lu, "A real-time two-stage and dual-check template matching algorithm based on normalized cross-correlation for industrial vision positioning," *Pattern Analysis & Applications*, vol. 8, no. 8, pp. 64–70, 2021.

[15] Y. S. Wu, Y. Shen, and H. Sun, "Intelligent algorithm-based analysis on ultrasound image characteristics of patients with lower extremity arteriosclerosis occlusion and its correlation with diabetic mellitus foot," *Journal of Healthcare Engineering*, vol. 2021, no. 9, 99 pages, Article ID 7758206, 2021.

[16] S. Govindarajan, J. A. Ardila-Rey, and K. Krithivasan, "Development of hypergraph based improved random forest algorithm for partial discharge pattern classification," *IEEE Access*, vol. 15, no. 99, p. 1, 2021.

[17] C. . Robinson, "Goal agnostic planning using maximum likelihood paths in hypergraph world models," *Springer International Publishing*, vol. 9, no. 2, pp. 121–130, 2021.

[18] Z. Yang, L. Xu, and L. Zhao, "Multimodal feature fusion based hypergraph learning model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9073652, 13 pages, 2022.

[19] B. Zhang, L. Wang, and Y. Li, "Precision marketing method of E-commerce platform based on clustering algorithm," *Complexity*, vol. 2021, no. 10, Article ID 5538677, 10 pages, 2021.

[20] J. Fang, "Research on automatic cleaning algorithm of multi-dimensional network redundant data based on big data," *Evolutionary Intelligence*, vol. 9, no. 25, pp. 1–9, 2021.