Application of Apriori Improved Algorithm in Intelligent Accounting

Wenjie Dong*

dwenjie0428@126.com*

School of Accounting, Dalian University of Finance and Economics, Dalian 116000, Liaoning, China

Abstract. In order to understand the application of Apriori improved algorithm in accounting intelligence, a research on the application of Apriori improved algorithm in accounting intelligence is put forward. First, this paper proposes to know the ways of financial intelligence based on the research context of financial intelligence. Second, by introducing the Apriori smart learning algorithm based on the Boolean-Mapping matrix, smart data can be made only by combining the code in the financial certificate. Correlations between original voucher symbols and accounting vouchers are extracted and produced by an improved character induction learning algorithm. Finally, a good work on intelligent financial systems is created, which is used in the development and application of intelligent financial systems research, and explores new avenues of research to integrate accounting and machine learning algorithms.

Keywords: Apriori improved algorithm; Accounting; intelligentize.

1 Introduction

Traditional accounting methods often rely on manual work by accountants, which is not only difficult but also error-prone. It takes a lot of time and manpower and costs very little money. At present, the level of information technology in the world has reached a certain height, and the development of information technology has not slowed down, but the environment is more scientific and necessary. Information technology is used in all spheres of life and its availability is very high, which is the basis for the further development of some traditional industries. Financial accounting has gradually evolved from basic accounting to financial accounting, and accounting has moved from traditional accounting methods to intelligent accounting methods. At the core of any business, the financial department must incorporate information technology into accounting. How to integrate finance with information technology to know real time and budget is a challenge for business owners. The key to realizing accounting intelligence is to enable computers to learn, make intelligent decisions, and think independently. With the help of financial rules obtained by learning algorithms, business transactions can be intelligently analyzed and resolved and financial reports can be generated. The author believes that it is not wise for human financial experts to pre-set management rules and know how to use automated financial systems because computers do not have autonomous operations. when learning and accounting rules change, it is necessary to manually change the voucher template; Financial robots based on RPA (Robotic Process Automation) technology are not financial intelligence because RPA essentially helps people perform standard tasks. Regular rules and high reproducibility of manual operation of the computer by checking

MSIEID 2023, December 08-10, Guangzhou, People's Republic of China

DOI 10.4108/eai.8-12-2023.2344730

according to regular rules - fixed procedures. established rules and predetermined procedures. In this paper, an Apriori algorithm based on Boolean-Mapping matrix is developed to extract active products from the voucher database, self-learning financial management rules, and extracted from the original voucher database by behavior induction learning algorithm. create rules for money management so that the computer can learn on its own in an unsupervised mode. According to the old vouchers and financial rules, with the help of technology, computers are responsible for counting financial certificates, and finally, the entire financial system is intelligent, which helps Chinese workers move from accounting to management. decision making [1-2]. As shown in Figure 1:

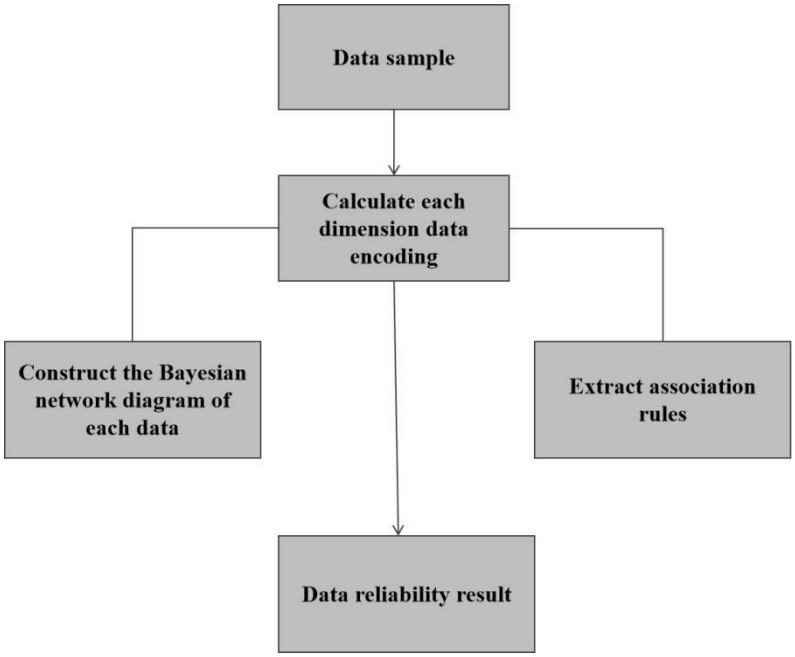

Fig. 1 Application of Apriori Improved Algorithm in Intelligent Accounting.

2 Enterprise accounting intelligence status quo

2.1 The degree of informatization is low

Although China has been committed to promoting enterprise data informationization and intelligent financial accounting in recent years, as far as the present situation is concerned, most enterprises in China still have manual bookkeeping, and this phenomenon is very common. The phenomenon of intelligent accounting of domestic enterprises is that the informatization popularization of most enterprises is low and the informatization process is very slow. Many enterprises' understanding of informatization stays in the simple operation of computer office software, and many employees will only use computers to operate some very basic office software, such as partial use of spreadsheets, which will not only bring substantial help to enterprises, but also prolong the dissemination time of information, and information users and decision makers will not be able to obtain accurate and reliable information in time.

In the end, the decision-makers make wrong decisions: and this way can't generate useful information in time, which leads to the phenomenon of information asymmetry. Enterprises can't correctly judge the market demand and delay the best opportunity for the production and sales of their products, which will make their products lose market competitiveness and bring huge economic losses to enterprises[3-4].

2.2 Insufficient flexibility

In the early stage of accounting intelligence, a certain amount of manpower and material resources need to be invested, and accounting intelligence can not directly bring benefits to enterprises. Some enterprise managers can't directly see the advantages brought by accounting intelligence, and think that the traditional accounting method is enough for the development and operation of enterprises, so they ignore the investment in accounting intelligence. This led to obstacles in the development of accounting intelligence. Although enterprises gradually began to pay attention to accounting intelligence in the later period, due to the lack of a good foundation in the early stage, accounting personnel's understanding of accounting intelligence was very low, and the popularization of software was also very difficult. Some enterprises have gradually discovered this problem and started to train accounting personnel. However, due to the high degree of software professionalism and the lack of professional software operators, accounting personnel can only use software for some very basic accounting, and cannot be flexibly applied to actual cases to play its role[5-6].Traditional accounting is the process by which accountants understand the importance and choices based on accounting policies, financial assumptions, accounting methods, and methods in accordance with the rules and regulations of the financial system. According to the financial statements, after the accounting is recorded and the ledger is written, all costs are incurred. However, accountants should consider the principles of financial analysis, financial guidelines, and financial statements individually, based on primary warrants and other relevant documents, when compiling financial statements and their preliminary connections. This process involves the decision of several experts. Whether computers can replace accountants in professional decision-making and financial reporting is key to determining whether China's financial system will shift from automation to artificial intelligence.The intelligence of the financial intelligence system is based on the premise that the computer should be able to learn autonomously in an unsupervised mode, update knowledge, make decisions and reason, store knowledge, etc. Therefore, on the one hand, an intelligent financial system must have a sufficient machine learning model, and on the other hand, it is necessary to introduce intelligent machine learning algorithms. At present, China's main archives only have financial certificates, financial books and reports, and important documents of intelligent machine learning algorithms (related to old vouchers and business documents) are still missing. With the Ministry of Finance's statement on electronic account audits and the increasing popularity, publication and implementation of electronic account lists, it is not only necessary but necessary to collect and store the original voucher data from the information system. Those. The main basis of financial evaluation is to determine the name of the fund, the direction of the fund and the amount of money that should be in the financial market. Different economies have different financial conditions and the effects of consolidation and financing policies will change over time and at different levels.construction. The intelligent learning algorithm correctly extracts and stores the combination of rules in the voucher database, simultaneously learns from the voucher database, the quality of the original voucher and accounting voucher

(basically, the characteristics of the original voucher and account number, financial direction, accounting amount, price the relationship between values), so that the financial management system has an intelligent character and can make professional decisions instead of using accountants and writing automatically. financial statements[7-8].

3 Research on Intelligent Learning Algorithm of Accounting

3.1 Research on the Improvement of Apriori Algorithm Based on The Boolean-Mapping Matrix

Apriori association rule mining algorithm was put forward by Agarwal R et al., which is mainly used to mine association rules between itemsets in customer database, and to mine association rules with Support and Confi-dence not lower than a given threshold from a specified record set. Among many algorithms for mining association rules, Apriori is the most basic and famous one.Let the item set R=(l1,l2,…,lm} be the combination set of a company's first-level account or non-variable end-level account and bookkeeping direction, for example, l1 stands for cash on hand–debit, l2 stands for cash on hand–credit, l3 stands for bank deposit–debit, l4 stands for bank deposit–credit, … Transaction set W={T1,T2,…,Tn} is a set of accounting vouchers for a certain period of a company. Each transaction T1 in W is a subset of R, T ∈ R. The research on the law of account combination in intelligent accounting can be abstracted as mining a frequent transaction T1 from the transaction set W, and the subset of R that constitutes the transaction T1 is the frequent account combination. To sum up, the introduction of Apriori algorithm in this paper fully meets the research goal of intelligent accounting.The basic concept of Apriori Algorithm is that it is a machine learning algorithm based on a large number of ideas, which continuously extracts objects through a layer-by-layer search. The Apriori algorithm will generate a large number of candidate sets during iteration, and assuming that some budget intersections are empty, the iterative process of the Apriori algorithm will generate a large number of illegal competitors; At the same time, the Apriori algorithm needs to scan the data many times to complete the cutoff and frequency statistics, which will greatly reduce the mining efficiency of the algorithm.

3.2 Research on Improvement of Apriori Algorithm

Focusing on the shortcomings of traditional Apriori algorithm, this paper proposes the development of Apriori algorithm based on Boolean-map, and the development ideas and steps are as follows.Create a Boolean combination matrix where the first column represents the number of transaction set W, the second to penultimate column represents each item of item set R, and the last column represents the total number of items in each transaction. Those. The rules of the Boolean network are as follows: if an item appears on the exchange Ti, the elements in the column equal to the sum of the items of row I of boolean numbers are represented as "1", and the column in which there is no item in row I is marked as "O". Those. By extracting the inner vector product to form a matrix of Boolean combinations, the Apriori processing algorithm can achieve the desired performance and efficiency with a single analysis of the data.In order to solve the problem of the traditional Apriori algorithm generating a large number of illegal candidates in the process of iteration, this paper puts forward a development strategy to generate candidates according to accounting rules: First,

after making accounting entries. is broken down in simple accounting, and all transaction changes are of three types: "one loan to one loan", "one loan to many loans" and "one loan to many loans". Those. According to the accounting rules of the debit-credit accounting method, no more than two debit accounts can appear when creating the contest "one debit many credits". All candidates at the same time; Similarly, when creating a "one credit-multiple credit" competitor trading setup, no more than two credit cards can appear in each competitor's setup at the same time. The second is to check the relationship between the libraries and use all cases where the combination number is empty to prevent false candidates. For example, in the configuration of the transaction "one loan-multiple loans", the first item is the combination "document-loan", when creating the candidate configuration "payment - loan", "written payment-loan", "short" -term loan-loan, "long" term loan", "paid-in capital (shared capital)-loan" should be used. Those. Third, according to the registration rules of the debit-credit accounting method, if there is only one thing in the algorithm iteration process, "If there is credit, there must be credit and they are equal." transactions should be removed from the learning model. Those. After improving the Apriori algorithm developed in this paper to optimize the candidates identified by the above three ideas, the candidates not generated by the improved Apriori algorithm will be greatly reduced during iteration, and the pruning cost and statistical frequency of the algorithm will be reduced. will be reduced and the performance of the algorithm will be improved. Table 1 shows the flow and pseudocode of the improved Apriori algorithm based on Boolean maps.

Table 1. Pseudocode of Improved Apriori Algorithm Based on Boolean Mapping Matrix.

1; Input:Amxn,Minsup

2:for (k=2,Lk=1≠ Φ ; k++)// Generate frequent k itemsets from frequent k-1 itemsets.
3:begin
4:C_k=apriori_gen(Lx7); //Call the function apriori_gen to generate Ck from Lk=1.
5: for all transactions TD do $\subset$ D do
6: begin
7:C1=subset(Ck,t); //Not the candidate set contained in t.
8: for all candidates c $\subset$ C
9:m=Ml,i) Ml,i)// Calculate the vector inner product of the item corresponding to the candidate set, where i is the number of columns of the matrix A.
10: qi=cOUNT(mi)// Count the sum of the items in column i.
11: if q < Mlinsup Then Delete Cuij//calculate and generate Am,n-x row intersection matrix Cmxm.
12:Then delete c from Ckj// delete c
13:endif

After mining all frequent itemsets from database D, the corresponding association rules can be obtained relatively easily. As shown in Equation (1):

$$\text{conf}\,(X \Rightarrow Y) = P(Y \mid X) = \frac{\text{sup\_count}\,(X \cup Y)}{\text{sup\_count}\,(X)} \tag{1}$$

Where, $\text{sup\_count}\,(X \cup Y)$ is the number of transaction records containing item set $X \cup Y$; $\text{sup\_count}\,(X)$ is the number of transaction records containing itemset X.

This algorithm uses a new data coding technique, and the experiment proves that this new algorithm is very effective for mining association rules in relational database.The core idea and steps of Coding-Apriori algorithm are as follows: firstly, the relational database is mapped

to attribute names and attribute values, and then the mapped database is encoded, and finally the frequent itemsets are obtained by replacing the traditional Apriori algorithm with the encoding "and" operation.

Attribute mapping in relational database includes attribute name mapping and attribute value discretization. The transformation of attribute name mapping is mainly for the consideration of simplified pattern matching, and the purpose of the transformation of attribute value discretization is to transform relational database into transaction database, so that frequent itemsets can be obtained by using the idea of Apriori algorithm. The general approach to discretization of attribute values is to discretize the attributes related to the mining task hierarchically into a common character set according to predefined concepts, such as {1,2... ,n}, then the relational database is converted to a transactional database.

However, the items represented as characters in this common character set do not reflect the relationship between items; some items belong to the same property, while others do not. This also leads to the fact that dimension constraints cannot be taken into account when "pruning" the set of candidates. When analyzing the characteristics of association rules in relational database, we find that it is impossible to form association between items with the same attribute discretized in relational database. In other words, if the items after attribute discretization can reflect the relationship between them, the subsequent mining algorithm can delete the candidate sets generated by the items belonging to the same attribute discretization, thereby reducing the number of candidate sets and improving the mining efficiency.

The mapping method is as follows: for a Boolean attribute, map it to two sets of strings, such as $\{A1, A2\}$; For the classification attribute, it is mapped to n string sets according to its number of classification values n, such as $\{B1, B2, \ldots, Bn\}$. When the classification attribute has many different values, it should be generalized at a certain conceptual level, and then the generalized classification values are mapped to string sets. For quantization attributes, they are discretized using predefined conceptual layers. The discretized numerical attributes have interval values and can be mapped to m string sets like classification attributes, such as $\{C1, C2, \ldots, Cm\}$. In the mapping process, the same attribute value is mapped to a string with the same first character, which can facilitate the judgment of intra-dimensional association and delete items from the same attribute in the candidate itemset, thus reducing the size of the candidate itemset to a large extent.

4 Intelligent accounting system design research

According to the accounting practice process and the needs of intelligent accounting, the intelligent accounting information system architecture designed in this paper is shown in Figure 2.

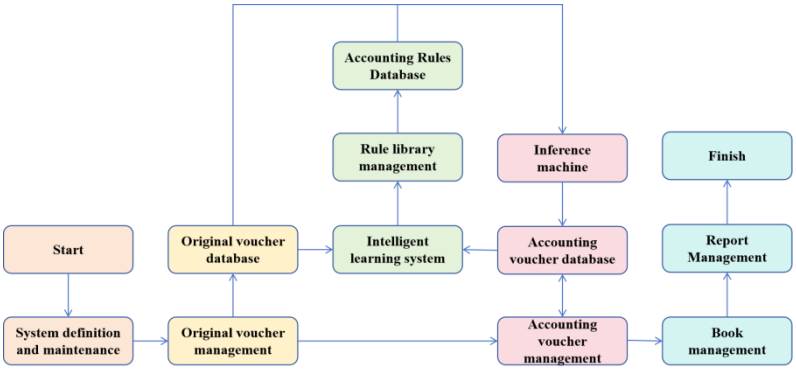

Fig.2 Architecture diagram of the intelligent accounting system

According to the requirements of the financial system and the intelligent financial system, the intelligent information system architecture developed in this paper is shown in Table 2, and its functions are also shown as follows.This job usually involves basic tasks such as budgeting, scheduling, and personnel management. At the same time, the system must collect and store a lot of important information, such as personnel information, organization information, product information, etc., which is the basis for intelligent computer decision-making.This link provides a large number of learning samples for Apriori's improved algorithm. Before the intelligent learning system generates accounting rules, it can manually compile the data entry of vouchers for accountants, and at the same time, it can also confirm or correct the intelligent accounting vouchers automatically compiled by computers.

Table 2. Architecture diagram of intelligent accounting system.

begin	System definition and maintenance	Original Intelligent Learning System Initial Voucher Management	Accounting voucher management	Account book management
	Original voucher database	Intelligent learning system	Accounting voucher database	Report management

Intelligent learning system is the most important module in intelligent accounting system, which mainly learns and extracts relevant rules from the database of original vouchers and accounting vouchers:Pre-processing of learning sample data. First, setting control parameters for system intelligent learning, mainly including Minsup, activation parameters of learning system, attribute learning threshold, etc. The second is to create a Boolean mapping matrix according to the accounting voucher database, which provides learning samples for the improved Apriori algorithm.Mining association rules: by calling the improved Apriori algorithm, learning and training the account combination rules in the accounting voucher database, and outputting frequent accounting account combination itemsets, providing basic data for learning the original voucher attributes[9-10].

The experimental development language is the C # language. Using the Lenovo computer, the memory is 2 GB, the CPU is using Intel i5, and the main frequency size is 1. 6 GHz. The experimental data set is the asset transaction data set purchased in the operation process of five

subways in a certain city. The data volume is 313 MB, including 862,000 asset records, which belongs to the massive data set. After determining the database structure, first regulate the data properties and convert them into the same format according to the quantitative rules to preprocess the data in the same format; then reference the SQL statement statistics function to complete the data statistics; finally use the improved Apriori algorithm to complete the calculation of the candidate item set, so as to construct the frequent item set of strong association rules. Rail transit engineering construction units have a large amount of asset data, and the amount of asset attributes is also large, and the number of related rules mined is more. To verify the mining efficiency and accuracy of the algorithm, the mining time of the traditional Apriori algorithm with different support is compared with the improved algorithm, and the results are shown in Figure 3.

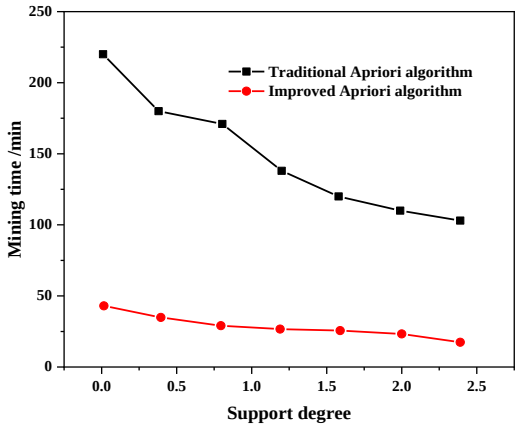

Fig.3. Mining time required for the old and new algorithms with different support levels.

5 Conclusion

With the in-depth application of artificial intelligence technology in accounting field, accounting intelligence has become one of the hottest topics in the development of accounting industry at this stage. Under the background of informationization, the important position of accounting intelligence will become more and more prominent. Intelligent accounting can not only provide comprehensive information for enterprises, but also bring higher economic benefits to enterprises. However, at present, the accounting software system still cannot automatically convert the original vouchers into bookkeeping vouchers, and enterprises need to take other measures to make up for this deficiency. Focusing on the intelligentization of accounting, this paper comprehensively applies the improved Apriori algorithm based on Boolean mapping matrix to maximize the intelligentization of accounting and lay the foundation for the good development of enterprises.

Acknowledgment: This work was supported by basic scientific research project (youth project - supporting project) of Liaoning Provincial Education Department (No. LJBKYF2021001) and basic scientific research project (general project) of Liaoning

Provincial Education Department (No. LJKMR20221957) and The title of the project is research on the strategy of private universities serving rural revitalization, project number LMJX2023265, Liaoning Provincial Association of Private Education Science "14th Five-Year Plan."

References

[1] Yi, Z. , & Zhengquan, Y. . (2021). Application of an improved physarum polycephalum algorithm on qos routing problem. International journal of wireless and mobile computing: IJWMC,24(4), 21.

[2] Zhao, H. , Zhou, H. , & Yang, G. . (2021). Research on global path planning of artificial intelligence robot based on improved ant colony algorithm. Journal of Physics Conference Series, 1744(2), 022032.

[3] Chen, L. , Xiao, Y. , & Yang, T. . (2021). Application of the improved fast iterative shrinkage-thresholding algorithms in sound source localization. Applied Acoustics, 180, 1(0)8101.

[4] Eliguzel, I. M. , & Ozceylan, E. . (2021). Application of an improved discrete crow search algorithm with local search and elitism on a humanitarian relief case. Artificial Intelligence Review: An International Science and Engineering Journal41(6), 54.

[5] Gao, J. . (2021). Research on application of improved association rules mining algorithm in personalized recommendation. Journal of Physics: Conference Series, 1744(3), 032111-.

[6] Ma, J. . (2021). Intelligent decision system of higher educational resource data under artificial intelligence technology. Pediatric obesity., 16(5)74.

[7] Mohamed, W. , & Abdel-Fattah, M. A. . (2022). A proposed hybrid algorithm for mining frequent patterns on spark. International journal of business intelligence and data mining74(2), 20.

[8] Liu, X. , Zhao, D. , Peng, Y. , & Li, J. . (2023). A wavelet basis ann and 5-class decision factor ai algorithm. International Journal of Pattern Recognition and Artificial Intelligence, 37(08)74.

[9] Ma, J. . (2021). Intelligent decision system of higher educational resource data under artificial intelligence technology. International Journal of Emerging Technologies in Learning (iJET), 16(5), 130.

[10] Xiong, Q. . (2021). Smart community security monitoring based on artificial intelligence and improved machine learning algorithm (retraction of vol 38, pg 7351, 2020). Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology45(5), 41.