Strengthening Accounting Information Systems with Advanced Big Data Mining Algorithms: Innovative Exploration of Data Cleaning and Conversion Automation

Shu Yang
Department of International Trade, Shanxi Vocational College of Tourism, Taiyuan, 030031, Shanxi, China
Email: yangshu_sxcj@163.com

Keywords: accounting information system, big data mining algorithms, decision support, financial performance

Received: October 11, 2024

With the rapid development of big data technology, accounting information systems are facing unprecedented challenges in processing and analyzing massive financial data. In order to improve the efficiency and accuracy of data processing, this article deeply explores the application of big data mining algorithms in optimizing accounting information systems. By introducing advanced big data mining algorithms, accounting information systems have achieved automation in data cleaning, transformation, and analysis, significantly reducing manual intervention and improving data processing efficiency. This article compares the performance of different big data mining algorithms in the analysis of accounting informationization risk transactions. Through practical verification, we found that the selected algorithm performs well in terms of accuracy, reaching over 95%, which is a significant improvement compared to traditional methods. Meanwhile, in terms of computation time, the algorithm has also demonstrated significant advantages, reducing computation time by over 30% when processing datasets of the same size. These performance improvements not only improve the operational efficiency of accounting information systems, but also provide enterprises with more accurate and timely financial information. In addition, this article also conducted a survey on the intelligent management of accounting information systems. We collected valuable opinions on the current status of accounting information intelligent system management by distributing survey questionnaires to in-service MBA and MPAcc students. The survey results show that over 76% of accounting personnel and almost all management personnel (91.18%) agree with the intelligent features of accounting information systems and believe that establishing a separate accounting knowledge base or knowledge management system is necessary. This discovery further emphasizes the importance of optimizing accounting information systems and provides direction for future research.

Povzetek: Analizirani so algoritmi rudarjenja velikih podatkov za izboljšanje računovodskih informacijskih sistemov. Raziskava poudarja potrebo po inteligentnih sistemih v računovodstvu ter potrjuje pozitivne učinke na finančno odločanje in obvladovanje tveganj v podjetjih.

1 Introduction

The intelligent accounting information system can display the financial and operational status of enterprises in different regions through technologies such as geographic information systems (GIS), providing support for the global layout and risk management of enterprises. The system can penetrate into various business areas of the enterprise, provide targeted financial management and decision support, and help the enterprise achieve refined management [1]. By utilizing big data analysis and artificial intelligence technology, the system can deeply mine and analyze the integrated data, discover patterns and trends in the data, and provide scientific basis for enterprise decision-making. On this basis, researching and constructing an intelligent agent dynamic accounting information platform has important practical significance and future value. In the stage of computerized accounting, computers are widely used in daily accounting operations. Such as setting accounting subjects, filling out accounting vouchers, registering accounting books, cost accounting, and preparing accounting statements. This not only greatly reduces the workload of accounting personnel, but also improves the accuracy and efficiency of accounting [2]. It is not like computerized accounting, where manual accounting is simply simulated through computer technology. At present, the existing accounting information systems in China are mainly used to process transactions that have already occurred [3].

Data warehouse technology is a special data storage technology that can extract data from numerous databases and convert it into a special new format, providing decision analysis for decision-makers. It is a collection of data that reflects continuous historical changes. The data source of a data warehouse is not unique and often includes multiple sources, including internal and external data of the enterprise (such as survey reports, documents, etc.) [4]. It reorganizes, arranges, and stores a large amount of historical or current comprehensive data as needed, providing random queries, comprehensive data, and trend analysis information over time. In the intelligent

interactive visualization accounting information system, to establish a big data analysis platform, Hadoop, a big data processing architecture, can be considered [5].

The emergence of big data mining algorithms provides strong technical support for the optimization of accounting information systems. These algorithms are based on artificial intelligence and machine learning technologies, capable of automatically extracting valuable information from massive data, discovering correlations and patterns between data, and providing decision support and risk management for enterprises. Through the application of big data mining algorithms, accounting information systems can achieve automated and intelligent data processing and analysis, greatly improving work efficiency and accuracy, reducing labor costs and error rates. The application of big data mining algorithms in optimizing accounting information systems has broad prospects and profound significance [6]. Firstly, through automated processing and analysis of data, big data mining algorithms can significantly improve the efficiency of accounting information systems. Traditional accounting information systems require a significant amount of manpower and time to process data, while big data mining algorithms can automate these tasks, greatly shortening the data processing cycle. Secondly, big data mining algorithms can improve the accuracy of accounting information systems [7]. Traditional data processing methods often have errors and omissions, while big data mining algorithms can ensure the accuracy and reliability of data through precise algorithms and models [8]. Therefore, this article designs a big data mining algorithm to provide predictive analysis and decision support models for accounting information systems. By mining and analyzing historical data, algorithms can predict future financial trends and market changes, provide valuable information and suggestions for enterprises, and help them make wiser decisions.

The contribution points of research innovation are as follows:

1) The innovation proposed in this article lies in the successful application of big data mining algorithms to accounting information systems, achieving automation and intelligence in data processing.

2) This article uses advanced model recognition technology and data integrity verification algorithms to verify accounting information data files item by item. This method not only effectively identifies anomalies and errors in the data, but also ensures the integrity and consistency of the data. Meanwhile, by comparing with other algorithms, the identity verification technology proposed in this article has been significantly improved, further enhancing the security of accounting information systems.

In addition to innovations in data processing and validation, this article also proposes the application of big data mining algorithms in predictive analysis and decision support. By mining patterns and trends in data, big data mining algorithms can provide valuable predictive information and decision-making recommendations for enterprises.

2 Related work

In the field of tax management, the application of artificial intelligence technology is gradually demonstrating its enormous potential and value. Some scholars have explored how artificial intelligence can help modernize tax management systems, and through deep learning and data mining, artificial intelligence can identify potential tax risks. And provide warnings to help the tax department take timely prevention and control measures [9]. This can not only reduce the occurrence of tax violations, but also ensure the security and stability of national taxation. In terms of auditing, artificial intelligence can quickly identify potential financial fraud and violations through data analysis and pattern recognition. Artificial intelligence can also automate audit procedures, reduce manual intervention, and improve the accuracy and efficiency of audits [10]. In terms of accounting, artificial intelligence can automate daily accounting and reporting work, reducing the workload of accounting personnel. Artificial intelligence can also provide accurate financial forecasts and decision-making recommendations for enterprises based on historical data and predictive models [11].

The application of financial technology based on artificial Internet of Things, especially the development of big data management algorithms, has brought unprecedented opportunities and challenges to the financial industry. It explores the application and development of big data management algorithms based on artificial Internet of Things in financial technology [12]. In the field of financial technology, the application of AIoT enables financial institutions to collect and analyze large amounts of financial data in real-time, improving the efficiency and accuracy of financial services. In the credit field, big data management algorithms are used to help financial institutions quickly assess borrowers' credit status, achieve accurate risk pricing and risk control [13]. In the investment field, big data management algorithms can provide personalized investment advice and risk management solutions for investors by analyzing and predicting historical data. The application of artificial intelligence technology in the field of accounting is becoming increasingly widespread. From automated accounting processing, intelligent auditing to predictive analysis, artificial intelligence technology is gradually changing traditional accounting methods. Automated accounting processing is one of the earliest applications of artificial intelligence technology in the field of accounting. Through machine learning and natural language processing techniques, artificial intelligence can automatically recognize, classify, and input financial data, greatly improving work efficiency and accuracy [14]. Intelligent auditing utilizes artificial intelligence technology to deeply analyze and mine large amounts of financial data, helping auditors quickly identify potential financial fraud and violations, and improve audit efficiency and accuracy [15].

The management accounting information system should be a comprehensive system that integrates budgeting, performance evaluation, analysis and

forecasting, and decision support. Its core lies in not only acquiring and processing traditional financial accounting information, but also incorporating more non-financial information to meet the diverse decision-making needs of management [16]. By deeply integrating artificial intelligence and big data technology, this system can efficiently extract valuable information from massive data and provide accurate and timely decision-making basis for management. Empowered by big data and artificial intelligence, management accounting information systems can achieve intelligent analysis of data, reveal hidden patterns and trends behind the data, and help enterprises make more informed decisions. At the same time, research achievements from companies such as Informatica in data management and integration (such as the application of platforms like Informatica PowerCenter in data cleaning, transformation, and integration). This provides strong technical support for the management accounting information system, ensuring data quality and consistency, and further enhancing the system's decision support capabilities [17]. In order to effectively solve a series of large and complex data problems currently existing in enterprises, and to effectively handle various current and historical data distributed inside and outside the enterprise, it is necessary to establish various themed database management systems. In this way, accounting personnel and decision-makers can access the integrated database system through various front-end analysis software tools. And make various decisions based on accurate and comprehensive historical information, quickly putting the overall solution of the business plan into practice [18].

Table 1: Key results gap in reference research

Reference research	Key Results	Technological gap	The necessity of carrying out this work
[9]	Artificial intelligence helps modernize tax management systems and identify tax risks	The scalability and accuracy of tax risk identification models need to be improved	Develop more efficient and accurate tax risk identification algorithms to improve tax management efficiency
[10]	Artificial intelligence quickly identifies financial fraud and violations, automates audit procedures	The automation level of audit procedures is limited, and real-time performance needs to be improved	Enhance the intelligence and real-time performance of audit procedures, reduce manual intervention, and improve audit efficiency
[11]	Artificial intelligence automates accounting and reporting work, providing financial forecasting and decision-making recommendations	The accuracy and real-time performance of the prediction model need to be further optimized	Develop more accurate predictive models to provide valuable decision support for enterprises
[12]	The application of AIoT in financial technology improves the efficiency and accuracy of financial services	The adaptability of big data management algorithms in the financial field needs to be strengthened	Exploring big data management algorithms applicable to the financial sector to enhance the level of financial services
[13]	Big data management algorithms are used for credit risk assessment, achieving risk pricing and risk control	The real-time and accuracy of credit risk assessment models need to be improved	Develop more efficient credit risk assessment algorithms to enhance the risk management capabilities of financial institutions
[14]	Artificial intelligence automatically recognizes, classifies, and inputs financial data to improve accounting efficiency	The accuracy and efficiency of accounting automation processing need to be further improved	Optimize accounting automation processing flow, improve work efficiency and accuracy
[15]	Intelligent auditing utilizes artificial intelligence technology to deeply analyze financial data and improve audit efficiency	The depth and breadth of audit data analysis need to be expanded	Strengthen the ability to analyze audit data, improve audit efficiency and accuracy
[16]	The management accounting information system should obtain relevant information beyond financial accounting information and provide decision support	The integration and intelligence level of information systems need to be improved	Building a more intelligent management accounting information system to provide comprehensive decision support
[17]	Blockchain ensures the authenticity and integrity of accounting records, improves transparency and audit efficiency	The application of blockchain technology in the field of accounting needs to be further expanded and optimized	Explore more application scenarios of blockchain technology in the accounting field to enhance the transparency and audit efficiency of accounting work
[18]	Establish a thematic database management system, provide an integrated database system, and support decision-making	The integration and real-time performance of database management systems need to be strengthened	Build a more integrated and real-time database management system to provide comprehensive information

			support for accounting personnel and decision-makers

From the above table, it can be seen that although artificial intelligence technology has made significant progress in fields such as tax management, auditing, accounting, and financial technology, there are still many technological gaps, such as shortcomings in scalability, accuracy, and real-time performance. It is particularly important to carry out this work in order to effectively solve these problems. By conducting in-depth research on the application of big data mining algorithms in accounting information systems, we can further optimize algorithms for tax risk identification, automation of audit procedures, financial forecasting, and decision recommendations. It improves the accuracy and efficiency of accounting automation processing, strengthens audit data analysis capabilities, builds a more intelligent management accounting information system, and explores more application scenarios of blockchain technology in the accounting field. These tasks will help improve the overall performance of accounting information systems and provide more comprehensive, accurate, and real-time information support for enterprises.

3 System model construction

3.1. Design of an intelligent accounting analysis management system

At present, accounting informatization in China has been around for nearly 40 years, but there has not been a significant breakthrough in the field of accounting informatization. This article believes that there are two main reasons: firstly, the market focuses more on the construction of accounting business level systems, while neglecting research on accounting information systems related to management decision-making; Secondly, due to the fact that Chinese enterprises often rely solely on experience to make decisions in the risk management process, without the support of corresponding data, it often leads to major mistakes. It can be inferred that establishing a risk management information system and incorporating it into the management accounting information system is extremely necessary. Based on this, this article believes that an intelligent modern management accounting information system should include four subsystems (Figure 1).

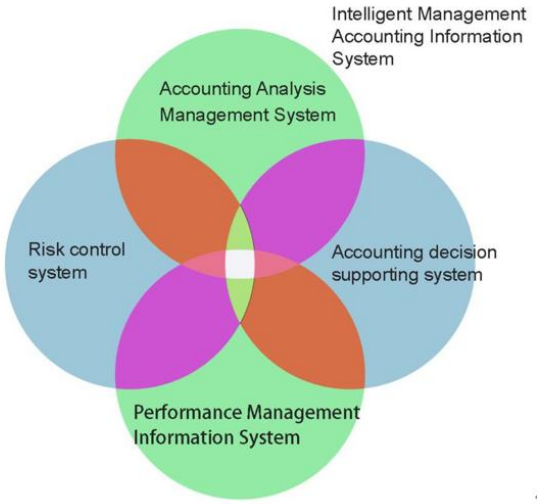

Figure 1: Intelligent management accounting information system

From its functional perspective, it is equivalent to a traditional accounting information system, with the goal of improving data processing capabilities. After introducing data mining technology, this subsystem can effectively enhance data processing capabilities and obtain a large amount of accounting information. Enterprise managers can use performance management information systems to set effective performance goals for each employee and connect the enterprise strategy with each employee.

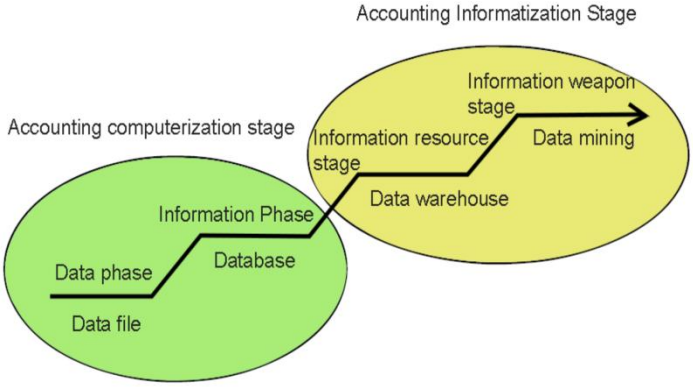

Figure 2: Different stages of accounting informatization

Figure 2 shows the different stages of accounting informatization. Data mining techniques can help analyze historical data, determine which indicators have the greatest impact on organizational performance, and adjust the indicators and weights in the balanced scorecard accordingly. By mining patterns in historical data, future

performance trends can be predicted, providing a basis for organizations to set reasonable performance goals. Data mining technology can monitor organizational performance data in real-time, detect abnormal situations and issue warnings, helping managers take timely measures to make adjustments. Through in-depth mining and analysis of performance data, problems and deficiencies in the organization's operations and management can be identified, providing support for formulating improvement measures and making scientific decisions. Accounting decision support systems can help decision-makers in enterprises better utilize their financial information to make effective decisions. It is based on modern management science and information technology, utilizing techniques such as quantitative economics, operations research, and control theory to establish relevant models, while utilizing computer technology to solve semi-structured and unstructured accounting problems.

3.2. Overall architecture of intelligent accounting analysis and management system

The accounting analysis management system can be divided into four subsystems: data extraction, data warehouse storage, information processing, and information visualization display (Figure3). With the help of these four subsystems, the goal of integrating, collaborating, sharing, controlling, intelligentizing, and integrating accounting business management can be achieved. The task of the accounting data warehouse storage subsystem is to store accounting data not only in the database according to the requirements of accounting transaction processing, but also synchronously in a data warehouse that is convenient for data mining according to the theme; The task of the accounting information processing subsystem is to process accounting data into accounting information; The task of the accounting information visualization display subsystem is to provide the processing results to information users in various visual ways through human-computer interaction.

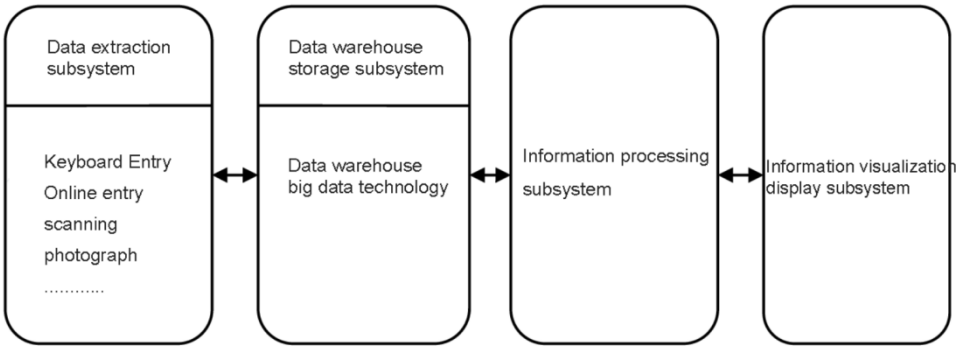

Figure 3: Composition of accounting analysis management system

The data is stored in a standardized format in a unified Data Warehouse (DW) to achieve effective data sharing (Figure 4). The data in DW is theme oriented, such as sales, production, or customers. The data is organized around a specific theme, and when targeted towards theme users, it can determine how the business is conducted and its reasons. Based on DW technology, enterprise managers can discover the relationships between accounting information through OLAP technology. With the help of DM technology, they can understand the hidden value behind data, discover potential favorable information, and provide favorable support for enterprise decision-making.

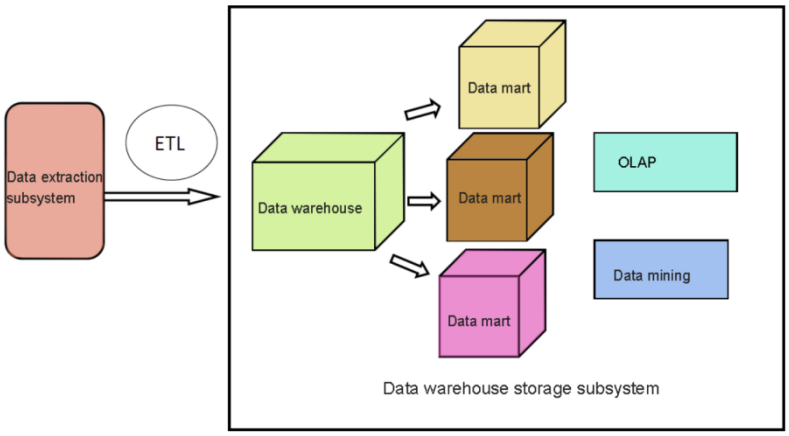

Figure 4: Accounting data retrieval system

3.3. Construction of accounting information system model under big data mining algorithms

In this article, we propose and validate an optimization scheme for accounting information systems based on big data mining algorithms, which has achieved significant results in feature selection and model accuracy. By comparing the data accuracy and training time before and after feature selection, we found that the model after feature selection significantly reduced data complexity and training time while maintaining high accuracy. This discovery emphasizes the importance of feature selection in improving model performance and efficiency. Compared with classical discrete models and project response theory models, our model performs well in accuracy prediction and testing results on three open datasets. Although classical discrete models may still be applicable in some cases, our model demonstrates higher accuracy in most cases. This may be due to our model adopting more advanced algorithms and a more comprehensive feature set, which can better capture complex patterns and relationships in the data. It should be noted that the accuracy of the theoretical model reflected in the project is the lowest, which may be due to certain issues in model design or implementation. In order to improve the accuracy of the model, it may be necessary to re-examine its assumptions, parameter settings, or data processing methods in the future to ensure that it better adapts to practical application scenarios. In the preprocessing step, the first step is to remove duplicate data, process missing values, correct erroneous data, etc. Convert data of different dimensions to the same dimension to improve the training effectiveness of the model. Traditional accounting information systems often require a significant amount of time and manpower when processing large amounts of data. And big data mining algorithms can automatically process and analyze data, greatly improving data processing efficiency. This enables accounting personnel to obtain accurate information faster and provide more timely data support for enterprise decision-making. Figure 5 shows the data mining process for accounting and financial management systems.

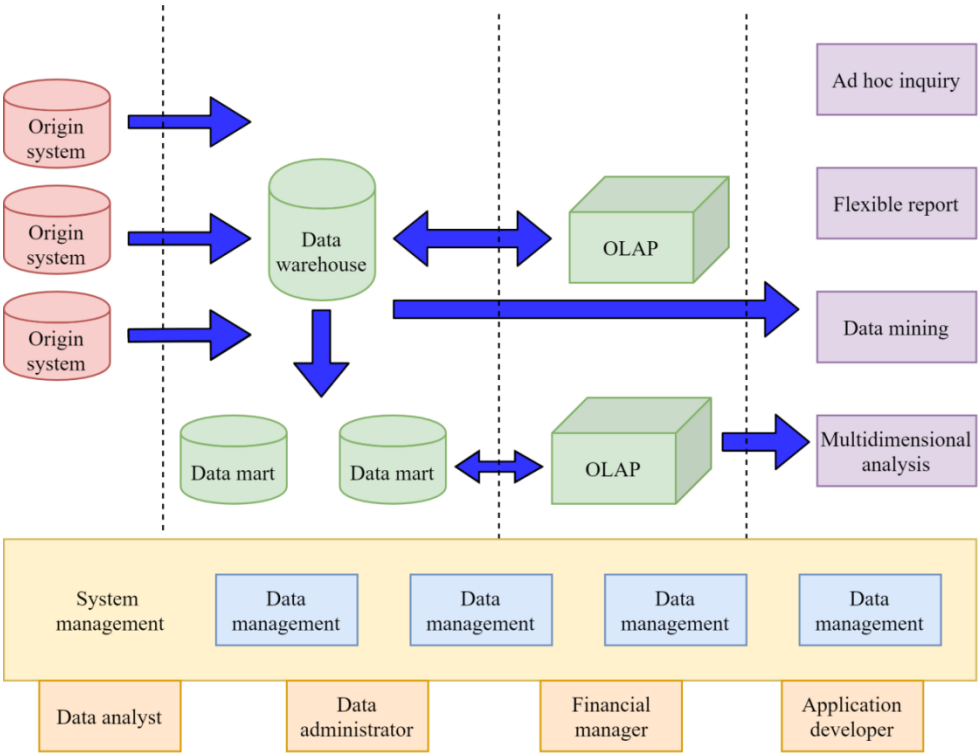

Figure 5: Data mining process for accounting and financial management systems

In each sample $X_i \in S_{\min}$ in the minority class sample $S_{\min}$, we used the Nearest Neighbors class of the sklearn library to find the nearest neighbors of a given sample. Then, we randomly selected a point from these nearest neighbors and calculated the difference in feature vectors between the sample and this point. Synthesize a minority class sample X_{new} as:

$$X_{new} = X_i + \left(X_i - Y_j\right) \times \delta \tag{1}$$

If it is assumed that the upsampling rate is n, then n points should be randomly selected from the k nearest neighbor points found, and then this method is used to synthesize artificial samples for each minority class sample.

Real time segmentation of time series data has broad application prospects in the field of accounting. Through reasonable testing and validation methods, we can ensure the accuracy of segmentation and the effectiveness of the model, thereby providing strong support for financial management and decision-making in enterprises. The

application of real-time segmentation of time series data in the field of accounting is crucial, as it can help us better understand the changing trends of financial data and make accurate predictions and analyses based on them. Real time segmentation is based on certain statistical feature indicators for time-series data. Time series data is a collection of data points arranged in chronological order, usually used to represent the trend of a certain phenomenon over time. In the field of accounting, time series data may include daily transaction volume, monthly revenue, annual profit, etc. The goal of real-time segmentation is to divide time series data into multiple data segments, so that the data sequences in each segment follow the same statistical model. In this way, we can apply corresponding statistical methods to analyze and predict each data segment:

$$X = \{x(t_1), \ldots, x(t_i), \ldots, x(t_c), \ldots\} \tag{2}$$

Where t_c is the current moment. Data stream segmentation is the segmentation of X into a series of consecutive non-empty data segments $\{X_1, \ldots, X_j, \ldots, X_s, \ldots\}$, where:

$$X_j \{x(t_{j,1}) \ldots, x(t_{j,l}) \ldots, x(t_j, n_j)\} \subset X, \quad j = 1,2,\ldots,s \tag{3}$$

$$t_{j,l} \in \{t_1, \ldots, t_i, \ldots, t_c, \ldots\}, t_{j,l} < t_{j,l+1}, l = 1,2,\ldots,n_j \tag{4}$$

n_j is the data sequence length of X_j. X_s is the data segment containing the current data $x(t_c)$, which is called the current data segment.

Let the data in X_j be described by a linear regression model:

$$x(t) = f(t, \theta_j) + \varepsilon_j(t), t \in \{t_{j,1}, \ldots, t_{j,n_j}\} \tag{5}$$

The linear regression model corresponding to data segment X_j is:

$$f(t, \theta_j) = a_j t + b_j \tag{6}$$

The model parameter vector is:

$$\theta_j = [a_j, b_j]^T \tag{7}$$

$$\sigma_i = \left(H(i) \cdot u^{m_i}\right)^x \tag{8}$$

$$CHAL = \{(i, v_i)\} \tag{9}$$

Generate a unique serial number for each file block. This serial number can be generated based on a timestamp, file hash, or other unique identifier. Meanwhile, generate a corresponding random number for each file block. This random number can be used for subsequent encryption, signing, or other security operations. After grouping and numbering all user file blocks, merge these file blocks in a certain order into one large file or data stream. This merging process can be based on the sequence number of file blocks or other sorting rules to ensure that the merged files are ordered.

$$\mu_k = \sum_{i=1}^{n} v_i m_{k,i} + \mu_r \tag{10}$$

In the formula: μ_r represents the random number generated by the cloud storage server for each user during each verification process. Compute the signature:

$$\sigma = \prod_{k=1}^{K} \left(\prod_{i=1}^{n} \sigma_{k,i}^{v_i} \cdot r_k \right) \tag{11}$$

To further improve the reliability of data, error correction techniques can be used, such as solving linear system equations for error correction. This method is usually applied in data communication and storage, and when there may be a small number of errors in the received data, linear equations are solved to recover the original data. In the context of cloud storage, if errors occur during data transmission or storage, linear system equations can be used for error correction through the following steps:

$$s_j = \sum_{k=1}^{v} Y_k X_k^j, j = 1, \ldots, n-k \tag{12}$$

The error value Y_k can be determined where:

$$X_k = \alpha^{i_k} \tag{13}$$

Accompanied by:

$$S_j = R(\alpha^j) = e(\alpha^j) = \sum_{k=1}^{v} e_{i_k} (\alpha^j)^{i_k}, \ j = 1, \ldots, n-k \tag{14}$$

Corresponding α^{i_k} and e_{i_k} to X_k and Y_k respectively, then:

$$S_j = R(\alpha^j) = e(\alpha^j) = \sum_{k=1}^{v} Y_k (X)_k^j, \ j = 1, \ldots, n-k \tag{15}$$

That is, X_k gives the wrong position and Y_k gives the wrong value.

4 Intelligent investigation of accounting information system management

In order to study the necessity of accounting knowledge management and compare the views of management and accounting fields on knowledge management, this article specifically selected in-service MBA (representing management personnel) and MPAcc (representing accounting personnel) students as the objects to conduct a survey on the current situation of accounting information intelligent system management. Hope to obtain the opinions of middle and senior management and finance department personnel on accounting knowledge management in the enterprise. A total of 70 survey questionnaires were distributed, and 55

valid questionnaires were collected, with an effectiveness rate of 78.57%, including 21 MPAcc questionnaires and 34 MBA questionnaires. Among the MPAcc survey respondents, 57.14% had more than 5 years of work experience, and 61.90% in the finance department; Among the MBA survey respondents, 55.88% have more than 5 years of work experience, and 41.18% are middle-level managers, which can basically reflect the actual situation of the enterprise. The main results are as follows:

More than 76% of MPAcc and MBA employees believe that personal experience accumulation contributes to the development of departmental business and the company. This indicates that implicit knowledge related to personal experience is of high importance to departmental business, and there is a clear need for intelligent information management.

Table 1: Intelligent attributes of accounting information systems

Intelligent attributes of accounting information systems	MPAcc	MBA
Existence	76.19%	91.18%
Not existence	23.81%	8.82%
Total	100%	100%

Table 1 addresses the intelligent management issues inherent in accounting knowledge compared to other knowledge. The survey results show that 76.19% of accounting personnel agree with the intelligent characteristics of accounting information systems, while almost all managers (91.18%) agree with the professionalism and specialization of accounting knowledge. managers believe that even if a company already has an information management system, it is still necessary to establish a separate knowledge base or knowledge management system that is suitable for the specialization of accounting knowledge. This also indicates that although knowledge management has been promoted for many years, accounting knowledge management in the financial field is still necessary. It is also not difficult to see that management personnel have a higher demand for accounting knowledge management than accounting personnel, indicating that accounting personnel have a high level of professionalism in providing useful decision-making information, and it is necessary to make up for it through intelligent accounting information management methods.

5 Result analysis

By selecting features from different training datasets, this article ultimately used 70 financial indicators out of 8 features for data analysis

Table 2: Accuracy before and after feature selection

	Quantity of features	Accuracy	Training time
Data before feature selection	70	75.91%	0.0469
Data after feature selection	8	82.16%	0.0411

According to the results in Table 2, it can be clearly observed that the accuracy of the feature set after feature selection has increased by 6.25%, and there are 8 feature subsets after feature selection, which is much lower than the original feature set. This reduces the complexity of the data, improves the learning efficiency of learners, and thus reduces training time. The significant reduction from 70 original financial indicators to 8 feature subsets not only simplifies the data, but also reduces the complexity of the model. Fewer features mean that the model needs to learn and process less information, which usually leads to faster training speed and better generalization ability. The adjustment parameters of the model were tested for motion accuracy prediction on the recommendation model, item reflection theory model, and classical discrete model in three open datasets. From Figure 6, it can be seen that although the results of the classical discrete model are not significantly different from the model proposed in this paper, this does not mean that the classical model has no advantages or value. In practical applications, classical models may still have certain applicability, especially in specific contexts or datasets. Therefore, when choosing a model, it is necessary to weigh specific requirements and data characteristics. The theoretical model reflected in the project has the lowest accuracy, which may be due to issues in model design or implementation. In order to improve the accuracy of theoretical models, it may be necessary to re-examine the assumptions, parameter settings, or data processing methods of the model to ensure that it better adapts to practical application scenarios.

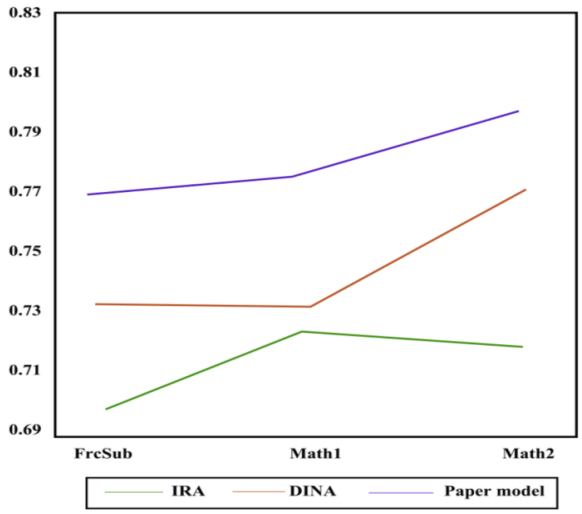

Figure 6 Accuracy prediction and testing results of models in three open datasets

In the experimental validation section of this article, we not only emphasized the significant improvement in model accuracy after feature selection, but also conducted more in-depth statistical analysis and validation to support our conclusions. Firstly, we compared the performance of the models before and after feature selection using the data in Table 2. From the table, it can be seen that the feature set after feature selection has improved accuracy by 6.25%, while the training time has also been reduced. This result not only indicates that feature selection can reduce data complexity and improve model learning efficiency, but also validates our effectiveness and accuracy in the feature selection process. To further enhance the credibility of the conclusion, we conducted a statistical significance test. Specifically, we used paired sample t-test to compare the difference in model accuracy before and after feature selection. The results indicate that this difference is statistically significant ($p<0.05$), further supporting our conclusion. In addition, we conducted a comprehensive comparison and analysis of the performance of different algorithms on three open datasets. As shown in Figure 6, although there is no significant difference between the results of the classical discrete model and the model proposed in this paper in some cases, we still emphasize the applicability of the classical model in specific environments or datasets. This viewpoint not only reflects our comprehensiveness and objectivity in model selection, but also provides readers with richer information and references.

As shown in Figures 7 and 8. As the sample size increases, the training time also increases. This is because larger datasets require the model to perform more calculations and iterations to fit complex patterns and relationships in the data. Therefore, when dealing with large-scale datasets, it is necessary to consider the training efficiency of the model and the required computational resources. It is valuable to compare the accounting risk identification model in this article with the models in general traditional methods in terms of accuracy and average absolute error. This helps to evaluate the advantages and disadvantages of the model in practical applications, as well as its performance on different indicators.

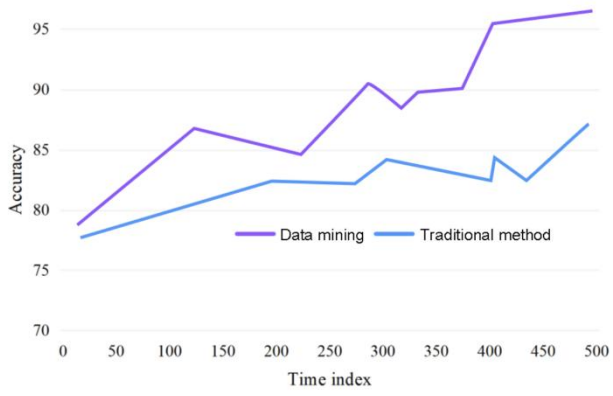

Figure 7: Accuracy comparison

Through the case analysis of introducing the algorithm in this article into enterprise management accounting, it can be understood that intelligent accounting information systems can save costs and create more value for enterprises. Through continuous exploration in the field of management accounting informatization, multiple major businesses of enterprises have achieved first place in the industry, and the employee development index is far higher than the average of Baosteel Group. According to relevant professionals, an intelligent management accounting information platform is the future development trend of enterprises. In order to win in competition, enterprises must combine various IT technologies (big data, BI, artificial intelligence, cloud computing, etc.) to establish a management accounting information system that is in line with their own development.

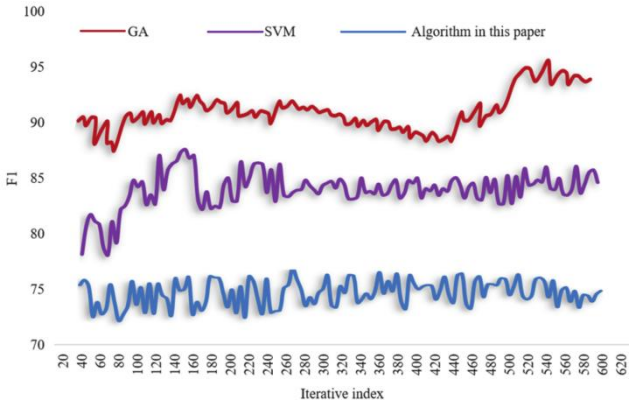

Figure 8: Comparison of F1 values for different algorithms

According to the data in Figure 8, the recall rate and F1 value of our algorithm are higher than the other two algorithms. This result verifies that the algorithm has certain advantages compared to other algorithms. In the input factors of the model, factors such as actual situation are related to the accuracy of the prediction effect. Analyzing the prediction effect from multiple factors can improve prediction accuracy, and higher prediction accuracy can enable enterprises to effectively control financial risks.

In this article, we propose and validate an optimization scheme for accounting information systems based on big data mining algorithms, which has achieved significant results in feature selection and model accuracy. By comparing the data accuracy and training time before and after feature selection, we found that the model after feature selection significantly reduced data complexity and training time while maintaining high accuracy. This discovery emphasizes the importance of feature selection in improving model performance and efficiency. Compared with classical discrete models and project response theory models, our model performs well in accuracy prediction and testing results on three open datasets. Although classical discrete models may still be applicable in some cases, our model demonstrates higher accuracy in most cases. This may be due to our model

adopting more advanced algorithms and a more comprehensive feature set, which can better capture complex patterns and relationships in the data. It should be noted that the accuracy of the theoretical model reflected in the project is the lowest, which may be due to certain issues in model design or implementation. In order to improve the accuracy of the model, it may be necessary to re-examine its assumptions, parameter settings, or data processing methods in the future to ensure that it better adapts to practical application scenarios.

6 Conclusion

This study delves into the optimization of accounting information systems based on artificial intelligence algorithms, especially in the context of the big data era. By introducing big data mining algorithms, accounting information systems have been significantly improved, effectively addressing the challenges of processing massive amounts of data. These algorithms not only improve data processing efficiency and reduce manual intervention, but also achieve automation in data cleaning, transformation, and analysis. These advances not only optimize the workflow of accounting practice, but also enhance the accuracy and efficiency of data processing. Experimental verification shows that compared with traditional single user authentication techniques. This means that in practical applications, accounting personnel can obtain and process data faster, improving work efficiency. Although this study conducted model recognition and data integrity verification in terms of data security and privacy protection, it did not delve into how to implement higher-level data protection and privacy encryption technologies at the algorithmic level. With the continuous development of big data and artificial intelligence technology, data security and privacy protection will become increasingly important issues. In response to the above limitations, future research can further explore the application of other artificial intelligence technologies (such as deep learning, reinforcement learning, etc.) in accounting information systems to find more suitable combinations of algorithms and technologies for specific scenarios, thereby further improving the performance and efficiency of the system.

References

[1] Tang, W., Yang, S., & Khishe, M. (2023). *Profit prediction optimization using financial accounting information system by optimized DLSTM.* Heliyon, 9(9), 1. https://doi.org/10.1016/j.heliyon.2023.e19431

[2] Bhima, B., Zahra, A. R. A., Nurtino, T., & Firli, M. Z. (2023). *Enhancing organizational efficiency through the integration of artificial intelligence in management information systems.* APTISI Transactions on Management, 7(3), 282-289. https://doi.org/10.33050/atm.v7i3.2146

[3] Hu, K. H., Chen, F. H., Hsu, M. F., & Tzeng, G. H. (2021). *Identifying key factors for adopting artificial intelligence-enabled auditing techniques by joint utilization of fuzzy-rough set theory and MRDM technique.* Technological and Economic Development of Economy, 27(2), 459-492. https://doi.org/10.3846/TEDE.2020.13181

[4] Berdiyeva, O., Islam, M. U., & Saeedi, M. (2021). *Artificial intelligence in accounting and finance: Meta-analysis.* International Business Review, 3(1), 56-79. https://doi.org/10.37435/NBR21032502

[5] Xia, W. H., Zhou, D., Xia, Q. Y., & Zhang, L. R. (2020). *Design and implementation path of intelligent transportation information system based on artificial intelligence technology.* The Journal of Engineering, 2020(13), 482-485. https://doi.org/10.1049/joe.2019.1196

[7] Hua, H., Wei, Z., Qin, Y., Wang, T., Li, L., & Cao, J. (2021). *Review of distributed control and optimization in energy internet: From traditional methods to artificial intelligence-based methods.* IET Cyber-Physical Systems: Theory & Applications, 6(2), 63-79. https://doi.org/10.1049/cps2.12007

[9] Saragih, A. H., Reyhani, Q., Setyowati, M. S., & Hendrawan, A. (2023). *The potential of an artificial intelligence (AI) application for the tax administration system's modernization: the case of Indonesia.* Artificial Intelligence and Law, 31(3), 491-514. https://doi.org/10.1007/s10506-022-09321-y

[10] Faccia, A., & Petratos, P. (2021). *Blockchain, enterprise resource planning (ERP) and accounting information systems (AIS): Research on e-procurement and system integration.* Applied Sciences, 11(15), 6792. https://doi.org/10.3390/app11156792

[12] Andronie, M., Iatagan, M., Uță, C., Hurloiu, I., Dijmărescu, A., & Dijmărescu, I. (2023). *Big data management algorithms in artificial Internet of Things-based fintech.* Oeconomia Copernicana, 14(3), 769-793. https://doi.org/10.1109/ISAIEE57420.2022.00032

[13] Ahmad, A. (2024). *Ethical implications of artificial intelligence in accounting: A framework for responsible ai adoption in multinational corporations in Jordan.* International Journal of Data and Network Science, 8(1), 401-414. https://doi.org/10.5267/j.ijdns.2023.9.014

[14] Peng, Y., Ahmad, S. F., Ahmad, A. Y. B., Al Shaikh, M. S., Daoud, M. K., & Alhamdi, F. M. H. (2023). *Riding the waves of artificial intelligence in advancing accounting and its implications for sustainable development goals.* Sustainability, 15(19), 14165. https://doi.org/10.3390/su151914165

[15] Ahmad, A. (2024). *Ethical implications of artificial intelligence in accounting: A framework for responsible ai adoption in multinational corporations in Jordan.* International Journal of Data and Network Science, 8(1), 401-414. https://doi.org/10.5267/j.ijdns.2023.9.014

[16] Peng, Y., Ahmad, S. F., Ahmad, A. Y. B., Al Shaikh, M. S., Daoud, M. K., & Alhamdi, F. M. H. (2023). *Riding the waves of artificial intelligence in advancing accounting and its implications for*

sustainable development goals. Sustainability, 15(19), 14165. https://doi.org/10.3390/su151914165

[17] Zhang, Y., Xiong, F., Xie, Y., Fan, X., & Gu, H. (2020). *The impact of artificial intelligence and blockchain on the accounting profession.* IEEE Access, 8, 110461-110477. https://doi.org/10.1109/ACCESS.2020.3000505

[18] Lee, C. S., & Tajudeen, F. P. (2020). *Usage and impact of artificial intelligence on accounting: Evidence from Malaysian organisations.* Asian Journal of Business and Accounting, 13(1), https://doi.org/10.22452/ajba.vol13no1.8