A Big Data-Driven Approach to Financial Analysis and Decision Support System Design

Sa Zhang
[1]School of Accountancy, Zhengzhou Vocational College of Finance and Taxation, Zhengzhou 450000, Henan, China
Email: julia6076@163.com

Keywords: big data, financial data analysis, decision support system

Received: September 3, 2024

This study aims to design and implement a financial data analysis and decision support system leveraging big data to enhance financial management and decision-making capabilities in complex market environments. Through a detailed examination of Company A's financial data, key indicators such as sales revenue, cost of sales, net profit, current assets, current liabilities, total assets, and shareholders' equity are selected. Utilizing big data technology, the system achieves efficient data processing, precise financial risk early warning, and scientifically informed investment decision support. Using Hadoop and Spark, the system efficiently processes extensive financial data while monitoring key indicators, such as sales revenue, cost of sales, and profitability. Specific financial models, including net present value (NPV) and internal rate of return (IRR), were tested for their effectiveness in supporting optimal investment decisions, yielding an NPV of 5.12 million and an IRR of 16.8% in the best-performing scenario. Performance metrics indicate a consistent improvement in data processing speed, accuracy reaching 98.9%, and user satisfaction rising from 8.2 to 8.7 over three years. The results indicate that the system performs effectively in terms of data processing speed, accuracy, and user satisfaction, enhancing both the efficiency of financial management and the precision of decision-making processes within enterprises. The financial risk early warning system successfully identifies potential risks in a timely manner, while the investment decision support system aids enterprises in selecting the optimal investment strategy by utilizing indicators such as net present value, internal rate of return, and investment payback period. This study highlights the value of applying big data technology in financial management, offering robust support for enterprises aiming for stable development within a rapidly evolving market environment. Future research will focus on further optimizing system functionality and expanding application scenarios to adapt to the shifting financial management demands of enterprises.

Povzetek: Predstavljen je sistem za finančno analizo in podporo odločanju, ki uporablja tehnologijo velikih podatkov. Integracija Hadoop in Spark omogoča hitro in natančno obdelavo podatkov, napredni modeli (NPV, IRR) pa izboljšujejo napovedi in oceno finančnih tveganj. Sistem omogoča optimizirano finančno upravljanje ter pravočasno odkrivanje tveganj, kar izboljšuje strateško odločanje podjetij v dinamičnem tržnem okolju.

1 Introduction

With the rapid advancement of information technology, global data volumes have grown exponentially, and big data technology has significantly transformed the operations across numerous industries. Financial man agement, as a central aspect of enterprise management, has also been deeply influenced by big data innovations. Big data technology enables the processing and analysis of vast financial datasets and provides real-time, precise decision support, helping enterprises sustain a competitive advantage in an increasingly demanding market environment.

The rapid development of information technology has significantly impacted the global financial industry, with big data emerging as a critical tool in financial management. Financial data is characterized by large volumes, complexity, and real-time requirements, making traditional data analysis methods insufficient. Big data technology offers solutions to these challenges by enabling efficient processing and analysis of massive financial datasets. By leveraging big data, enterprises can extract valuable insights, enhance their financial risk management, and improve decision-making capabilities.

The objective is to develop a comprehensive financial data analysis and decision support system that uses key financial indicators such as sales revenue, cost of sales, net profit, current assets, and liabilities. This system provides early risk warnings and supports investment decisions, helping enterprises remain competitive in rapidly changing markets. The integration of big data in financial management enhances data processing accuracy, enables real-time decision-making, and supports sustained business growth.

To enhance the robustness and accuracy of financial data analysis models, incorporating optimization techniques such as hyperparameter tuning, cross-validation, and feature engineering is essential.

Hyperparameter tuning allows for the adjustment of model parameters to maximize performance, enhancing the predictive power of models in varying financial scenarios. Cross-validation, particularly k-fold cross-validation, ensures that the models generalize well to new data by testing across multiple data subsets, reducing the risk of overfitting. Feature engineering, including techniques like scaling, encoding categorical data, and creating interaction terms, refines the data inputs to optimize model learning. These approaches collectively provide adaptability and precision, crucial for the dynamic financial environment where timely, accurate decision-making is a priority.

Traditional financial data analysis methods face significant challenges due to the large volume of data, the complexity of data types, and high demands for real-time processing. Big data technology effectively addresses these issues by enabling in-depth data mining and analysis, uncovering hidden patterns and trends, and enhancing the accuracy and efficiency of financial analysis. The use of Decision Support Systems (DSS) in enterprise management is increasingly prevalent. Combining DSS with big data technology provides businesses with more intelligent and customized decision support.

In recent years, companies have increasingly recognized the importance of data and are developing and refining their own data analysis and decision support systems. However, in practice, challenges persist, including suboptimal data quality, incomplete analysis models, and insufficient system integration. These issues limit the effective application of big data and decision support systems in financial management to some extent.

The objective of this study is to design a big data-based financial data analysis and decision support system, systematically addressing the limitations of existing systems and enhancing both the accuracy of financial data analysis and the effectiveness of decision support. This research will approach the topic through theoretical analysis, system design, empirical testing, and other methodologies, aiming to provide practical solutions for financial management and to advance the adoption of big data technology within the financial sector.

This study aim to enhance financial data analysis and decision-making capabilities for businesses by using a big data-based system. Specific objectives are as follows:

Improve data processing capabilities: Integrate big data technology to strengthen financial data processing, addressing traditional challenges related to large data volumes, diverse data types, and stringent real-time requirements. Optimize financial analysis models: Develop and refine financial data analysis models using big data technology to extract deeper insights from financial data, thereby improving the accuracy and comprehensiveness of financial analysis and providing valuable insights for business decision-making.

Design a robust decision support system integrated with big data analysis results to deliver real-time, accurate decision support for enterprise management. This system enables businesses to make informed and strategic decisions even in complex and volatile market conditions. Improve system integration by efficiently merging financial data analysis with decision support functions, ensuring smooth coordination across system modules. This enhances overall system efficiency and stability, providing businesses with a unified, comprehensive financial management solution.

By systematically analyzing the application of big data technology in financial data processing and decision support, this research enriches the academic content at the intersection of financial management and information technology. This study investigates specific application methodologies and pathways of big data in financial management, offering a novel perspective for theoretical exploration. Developing a big data-driven financial analysis model demonstrates the potential and strengths of data processing technologies in financial analysis. Based on existing DSS theory and enhanced by big data analysis, this research proposes a new system design and implementation framework, deepening DSS theoretical research and expanding its applicability within financial management. The exploration of innovative financial analysis methods further broadens the theoretical foundation of financial analysis, as the integration of big data processing technologies advances and refines traditional financial analysis models, offering a valuable new reference for related academic research.

At a practical level, this study designs and implements a big data-based financial data analysis and decision support system, offering positive impacts on enterprise financial management practices. The system leverages big data technology to process and analyze substantial volumes of financial data swiftly and accurately, enhancing both the efficiency and precision of corporate financial management. This enables enterprises to make timely and informed decisions in complex market environments. By combining decision support capabilities with big data analytics, the system provides a scientific and accurate decision-making foundation for enterprise management. Through real-time data analysis and forecasting, the system effectively supports financial decision-making processes, reduces decision-making risks, and improves decision quality. With this study, enterprises can better understand and apply big data technology to advance their information infrastructure. This system offers a comprehensive solution for financial data processing and decision support, aiding enterprises in achieving significant progress in digital transformation and information management.

In recent years, the application of big data technology in financial data analysis and decision support systems has garnered considerable attention. Current research emphasizes enhancing data processing capabilities and decision support through big data technology, adapting to increasingly complex financial management environments and evolving market needs. Casturi and Sunderraman proposed a rule-based, cost-efficient big data analytics aggregation engine for portfolio management, demonstrating big data technology's substantial advantages in increasing data processing efficiency and reducing costs. This study provides a theoretical foundation and practical experience for enterprises utilizing big data technology in financial data analysis [1].

In the educational field, Fahd and Miah designed and evaluated big data analytics methods for predicting student success factors [2]. Although their research primarily focuses on students, the methods and technology they employed are also applicable to financial data analysis. By identifying key factors within large datasets, these techniques enable data-driven decision-making processes. Dutta's research examined the asset-liability management models of life insurance companies, focusing on decision support system outcomes [3]. This research underscores the effectiveness of big data technology in managing complex financial environments and serves as a valuable reference for designing decision support systems.

For credit risk assessment, Lu et al. proposed a decision support method based on dynamic Bayesian networks, showcasing big data technology's potential in risk management by enhancing accuracy and timeliness through dynamic modeling [4]. Talamo et al. addressed organizational and individual decision-making challenges in the design of financial artificial intelligence systems, providing methodological insights that are especially beneficial for managing the complexity of decision processes in this study [5].

Peng and Bao developed an enterprise management analysis framework utilizing big data technology, which demonstrates the extensive applications of big data in enterprise management, particularly in enhancing data analysis and decision support capabilities [6]. Popovic et al. explored the influence of big data analytics on enterprise high-value business performance, concluding that big data analysis can significantly improve business performance, reinforcing its value in financial data analysis [7]. Liu et al. introduced a model for Internet financial risk control based on machine learning algorithms, showcasing the role of machine learning in financial risk management and offering technical guidance and methodological references for this study [8].

The creative contributions of this study are reflected in the application of big data technology to enhance financial data analysis and decision support systems. The use of advanced big data processing technologies, such as Hadoop and Spark, enables the system to handle large volumes of complex financial data with high efficiency and accuracy. Additionally, the design of an integrated financial risk early warning system allows enterprises to identify and mitigate risks through real-time monitoring of key financial indicators, such as the current ratio and asset-liability ratio. Another key innovation is the application of decision-making tools, like net present value (NPV) and internal rate of return (IRR), to assist businesses in making optimal investment choices. The system's design not only improves financial analysis accuracy but also empowers enterprises with data-driven insights to enhance decision-making in competitive markets.

Existing research demonstrates significant advancements in the use of big data technology for financial data analysis and decision support systems [9]. Nonetheless, several persistent challenges remain, such as ensuring high data quality, achieving seamless system integration, and meeting real-time processing demands. Addressing these issues is essential for advancing the implementation and utility of big data technology in financial management, as overcoming these obstacles will facilitate more robust and effective applications within the field.

Table 1: Summary of reviewed research in related work section.

Study Reference	Methods Used	Data Set	Evaluation Metrics	Key Findings
Casturi & Sunderraman [1]	Rule-based analytics, cost aggregation	Portfolio management data	Cost-effectiveness, efficiency	Big data analytics enhance cost-efficiency in portfolio management.
Fahd & Miah [2]	Predictive analytics for success factors	Student academic records	Predictive accuracy	Data analytics effectively predict key success factors for students.
Dutta et al. [3]	Asset-liability management model	Life insurance data	Decision support efficiency	Model improves asset-liability alignment in complex financial environments.
Lu et al. [4]	Dynamic Bayesian networks	Credit risk data	Accuracy, timeliness	Model enhances credit risk assessment accuracy with real-time data.

Talamo et al. [5]	Decision modeling, AI	Financial decision processes	Decision complexity, efficiency	AI-based systems improve decision-making under complex financial conditions.
Peng & Bao [6]	Business management analysis framework	Corporate management data	Performance enhancement	Big data application boosts business analysis and strategic planning.
Liu et al. [8]	Machine learning risk control model	Internet finance risk data	Risk prediction accuracy	Machine learning strengthens risk control in internet finance sectors.

2 Theoretical basis

2.1 Overview of big data technology

The methods outlined focus on systematically addressing financial data analysis and decision support system design. First, data is gathered from multiple sources, followed by preprocessing, which includes cleaning, transformation, and integration to ensure data quality. The system incorporates both relational and NoSQL databases to handle structured and unstructured data efficiently. Big data processing tools, like Hadoop and Spark, are then used to manage large volumes of data swiftly and accurately. The methods also integrate financial analysis models, such as regression analysis and decision trees, to predict and optimize financial decisions. These models are implemented within a user-friendly interface to provide real-time decision-making support. The methodology is coherent and logical, aligning well with the goal of improving financial data processing, risk management, and decision-making efficiency.

2.1.1 Definition and characteristics of big data

Big data refers to datasets too large or complex to be managed by traditional data processing tools. Its primary characteristics include Volume, Variety, Velocity, and Veracity.

Large volume: This is seen in the exponential growth of data amounts, far exceeding traditional database processing capacities. Zheng et al. highlight that modern information systems produce massive quantities of both structured and unstructured data daily, and that conventional database technologies are insufficient for handling this scale [10]. Advanced storage and processing solutions, such as distributed storage and parallel computing, are thus essential.

Variety: This characteristic is represented by the wide range of data sources and types. Xue's research notes that with rapid advances in information technology, data sources have diversified to include sensor data, social media data, and transaction records [11]. These sources contain both structured data and large amounts of unstructured data, such as text, images, and videos, complicating data processing and analysis tasks.

Fast: This characteristic refers to the rapid rate at which data is generated and processed. Modern information systems demand real-time or near real-time data processing and analysis to enable timely decision-making. Xu and Zhou noted that advances in big data technology make real-time data processing feasible, thereby supporting enterprises in making prompt decisions within a rapidly evolving market environment [12]. These real-time requirements place higher demands on data processing technologies, necessitating the adoption of streaming data processing and real-time analytics.

Authenticity: This aspect emphasizes the quality and reliability of data sources. Big data comes from diverse sources, leading to varying degrees of data quality, which makes ensuring data authenticity and accuracy a critical concern.

2.1.2 Big data processing technology

Big data processing technology encompasses the full sequence from data acquisition, storage, processing, and analysis (as shown in Figure 1). Applying big data technology enables the efficient handling and analysis of vast datasets, offering robust technical support for decision support systems.

This study reinforces the critical evaluation of the references, focusing on those sources that directly support the arguments and methods. By selectively referencing the literature, the research ensures consistency with the center's claims and enhances credibility. This study draws on articles published on Informatica, such as theoretical discussions on big data processing and decision support system design. Provides reliable, peer-reviewed insights for research that confirm the application of big data in financial analysis and risk management.

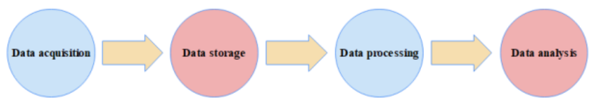

Figure 1: Big data technology processing flow.

The initial step involves data acquisition, gathering information from a wide range of sources. Rodrigues et al. indicate that a significant challenge in data collection arises from the diverse origins and formats of these datasets [13]. Sources such as sensors, social media, and transaction records provide critical data, often in both structured and unstructured formats, necessitating various acquisition and preprocessing techniques.

The second step is data storage, a critical component of big data processing that influences overall efficiency and reliability. Traditional relational databases often cannot handle the scale and complexity of big data, making NoSQL databases and distributed file systems (such as Hadoop HDFS) the primary options for effective storage. Jensen et al. noted that NoSQL databases offer flexible data management across diverse formats, while distributed file systems ensure efficient data storage and access mechanisms [14].

The third step is data processing, where distributed computing technology is essential. MapReduce serves as the primary computing model within the Hadoop framework, dividing data into smaller units for parallel processing across multiple nodes, thereby achieving efficient large-scale data handling. Rodrigues highlighted that the MapReduce model is particularly advantageous in massive data processing scenarios that require intensive calculation and analysis [13]. As a big data processing engine, Spark enhances both speed and processing efficiency by utilizing in-memory computing technology, providing faster performance compared to MapReduce.

The fourth step is data analysis, which represents the ultimate objective of big data processing. Through data mining, machine learning, and statistical analysis, valuable insights and knowledge can be extracted from vast datasets. Jensen emphasized that machine learning algorithms in big data analysis significantly improve predictive accuracy and pattern recognition [14]. Techniques such as classification, clustering, and regression are widely applied in financial data analysis, market prediction, and risk assessment.

Data security and privacy are critical in big data processing. It is essential for big data technologies to implement effective strategies to safeguard data confidentiality and integrity while maintaining processing efficiency. Common security measures include encryption techniques, access control, and data anonymization, all of which play a vital role in protecting sensitive information. Proper encryption ensures data is unreadable to unauthorized users, while access control restricts data access based on user roles and permissions. Additionally, data anonymization techniques help preserve privacy by masking personal identifiers within datasets, thus preventing unauthorized disclosure of individual information.

2.2 Theoretical framework of financial data analysis

2.2.1 Financial data analysis methods

Financial data analysis primarily involves data preprocessing, data mining, and data visualization.

Data preprocessing is the initial stage of financial data analysis and includes data cleaning, transformation, and reduction. Li and Chen noted that preprocessing helps remove data noise, fills in missing values, and standardizes data formats to support subsequent analysis and mining processes [15]. When dealing with financial statements, it is essential to ensure all data is thoroughly cleaned and standardized for reliable analytical outcomes.

Data mining is the core phase of financial data analysis. Wang developed an investment recommendation model based on time series data, utilizing data mining techniques from financial time series analysis [16]. Common data mining approaches include classification, regression, clustering, and association rule mining. These methods enable analysts to detect patterns and trends within financial data, supporting informed investment decisions and risk management. Classification aids in credit risk assessment, clustering assists in customer segmentation, and association rule mining reveals correlations among financial indicators. Specifically, fuzzy clustering and the Apriori algorithm are frequently applied, with Guo highlighting their use in the interactive analysis of financial management software to identify frequent item sets and association rules within datasets [17].

Data visualization represents the final stage in financial data analysis, presenting complex results in a form accessible to decision-makers. Data visualization techniques include charts, dashboards, and interactive reports [18]. These tools enhance communication of analytical results and assist users in identifying hidden patterns and anomalies within data. Trends in financial indicators are easily observable through time series graphs, while correlation levels between various financial indicators can be visualized using heat maps. Each step of financial data analysis offers distinct benefits and challenges, necessitating tailored choices of techniques and methods aligned with specific analysis objectives and data characteristics.

2.2.2 Financial data analysis index system

The index system of financial data analysis plays a vital role in comprehensively assessing an enterprise’s financial condition and operational performance by measuring and evaluating various key indicators. A well-constructed index system for financial data analysis includes core financial indicators such as operational efficiency, profitability, debt repayment capacity, and growth potential [19]. These indicators evaluate an enterprise's financial health from multiple perspectives, offering a scientific foundation for informed decision-making. The specific details are presented in Table 2 below.

Table 2: Financial data analysis indicator system.

Indicator Category	Indicator Name	Calculation Formula	Description
Profitability	Gross Profit Margin	(Sales Revenue - Cost of Goods Sold) / Sales Revenue	Measures basic profitability of sales activities
Net Profit Margin	Net Profit / Sales Revenue	Measures overall profitability of the enterprise	
Return on Assets (ROA)	Net Profit / Total Assets	Measures efficiency in using assets	
Return on Equity (ROE)	Net Profit / Shareholder's Equity	Measures return on shareholder's investment	
Solvency	Current Ratio	Current Assets / Current Liabilities	Evaluates short-term debt-paying ability
Quick Ratio	(Current Assets - Inventory) / Current Liabilities	A stricter evaluation of short-term solvency	
Debt to Asset Ratio	Total Liabilities / Total Assets	Evaluates financial structure stability	
Operating Efficiency	Inventory Turnover Ratio	Cost of Goods Sold / Average Inventory	Measures inventory management efficiency
Accounts Receivable Turnover	Sales Revenue / Average Accounts Receivable	Evaluates efficiency of accounts receivable recovery	
Total Asset Turnover	Sales Revenue / Average Total Assets	Measures asset utilization efficiency	
Growth Ability	Sales Growth Rate	(Current Period Sales Revenue - Previous Period Sales Revenue) / Previous Period Sales Revenue	Measures the growth rate of sales revenue
Net Profit Growth Rate	(Current Period Net Profit - Previous Period Net Profit) / Previous Period Net Profit	Measures the growth rate of net profit	
Asset Growth Rate	(Current Period Total Assets - Previous Period Total Assets) / Previous Period Total Assets	Measures the growth rate of asset scale	

2.2.3 Common models in financial data analysis

Commonly used models in financial data analysis include the regression analysis model, time series analysis model, and decision tree model. Among these, the regression analysis model is foundational and widely applicable, enabling companies to forecast financial indicators by examining the relationships between independent and dependent variables. The linear regression model, in particular, is a widely adopted variant of the regression analysis model, with its basic form represented as follows (Formula 1):

$$Y = \beta 0 + \beta 1 X + \epsilon \quad (1)$$

Where, Y is the dependent variable, X is the independent variable, β_0 and $\beta 1$ are the intercept and slope of the model, respectively, and ϵ represents the error term. By estimating parameters β_0 and $\beta 1$ using the least squares method, the linear relationship between the independent and dependent variables can be established, enabling prediction and analysis based on the model.

The time series analysis model is utilized to process sequential data, examine historical data trends, and forecast future financial metrics. Common models within time series analysis include the autoregressive integrated moving average model (ARIMA) and the exponential smoothing model (ETS). The ARIMA model uniquely integrates autoregressive and moving average components to effectively capture temporal dependencies and trend patterns in data, and its basic form is as follows (Formula 2):

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

Y represents the observed value at time t, c is the constant term, ϕ is the autoregressive coefficient, θ_j is the moving average coefficient, and ε_t denotes the error term. Utilizing the ARIMA model enables the capture of trends and cyclical patterns within time series data, facilitating effective financial forecasting.

The decision tree model is a decision analysis tool structured as a tree, widely used for both classification and regression tasks. It works by recursively partitioning the dataset into smaller subsets, selecting the optimal split point at each node until a stopping criterion is met. The construction process follows these basic steps:

1. Select the optimal feature as the basis for node segmentation.
2. Divide the data set into subsets according to different values of feature.
3. Recursively build subtrees for each subset until the stop condition is met.

The decision tree model describes the node segmentation process through (Formula 3):

$$Gini(D) = 1 - \sum_{i=1}^{k} p_i^2 \quad (3)$$

$\mathrm{Gini}(D)$ represents the Gini coefficient of node D, while k denotes the total number of classes, and p_i indicates the proportion of Class E samples. The Gini coefficient serves as a criterion for identifying the optimal segmentation point, ensuring the highest possible purity of the resulting subset.

2.3 Basic theory of decision support system

2.3.1 Definition and composition of decision support system

A Decision Support System (DSS) is a computer-based information system developed to assist decision-makers in handling complex decision-making tasks. By integrating data, analytical models, and user interfaces, DSS enables users to analyze issues, formulate decision strategies, and evaluate outcomes effectively [20]. Widely applied across fields such as enterprise management, healthcare, and financial investment, DSS provides a scientific foundation for decision-making that enhances both accuracy and efficiency. Through data analysis, model computation, and expert systems, DSS equips decision-makers with structured tools and methodologies for each phase of the decision-making process, including problem identification, data collection, model development, solution evaluation, and feedback.

The composition of a DSS typically includes several key subsystems:

The Data Management subsystem is responsible for the collection, storage, and management of data essential for decision-making. Data sources include both internal data (e.g., enterprise financial and production data) and external data (e.g., market and economic information) [21]. This subsystem houses the Database Management System (DBMS), which facilitates efficient data access and management. A primary function of the Data Management subsystem is to ensure data accuracy, consistency, and timeliness, providing a reliable foundation for subsequent analysis and decision support.

The Model Management subsystem oversees the storage and administration of analytical models required for decision support. These models may include mathematical, statistical, optimization, or simulation types [22]. This subsystem allows for model definition, creation, modification, and execution, supporting flexible application and updating of models. By selecting appropriate models, complex scenarios can be simulated and analyzed, enabling decision-makers to evaluate the potential impacts of various options.

The User Interface subsystem serves as the intermediary between the DSS and the user. It offers an accessible interface that allows decision-makers to input data, select models, execute analyses, and view results. Key components of this subsystem include a graphical user interface, reporting tools, and visualization options, which present data and analysis outcomes clearly to facilitate users' understanding and application of system information.

The Knowledge Management subsystem plays a crucial role by gathering, storing, and managing decision-related knowledge, such as expert insights, business rules, and historical decision examples. This subsystem employs a Knowledge Base to convert expert knowledge into system rules, thereby assisting decision-makers in making more informed and scientifically grounded choices.

The Communication subsystem enables information exchange between the DSS and the external environment. Through this subsystem, the DSS can acquire up-to-date market and economic data from external sources, while also transmitting analysis results and decision recommendations to relevant decision-makers or implementing departments.

2.3.2 Application of decision support system in financial management

The Decision Support System (DSS) offers scientifically grounded decision-making support for enterprises by integrating data analysis, model computation, and user interfaces. DSS is particularly effective in budgeting and control, as it leverages historical financial data and market forecasts to help enterprises develop rational budgets and make dynamic adjustments during implementation, ensuring budget accuracy and execution effectiveness. Kwan et al. demonstrates that computerized DSS can enhance decision-making accuracy, improve efficiency, and minimize human errors [23].

Through automated data processing and analytical functionalities, DSS swiftly generates financial statements and analytical reports, enabling enterprise management to understand financial health and operational outcomes in a timely manner. This automation enhances reporting efficiency and improves data accuracy and transparency.

DSS also identifies, assesses, and manages various financial risks faced by enterprises using an integrated risk assessment model. Lutz et al. noted that DSS performs strongly in complex analysis and forecasting tasks, aiding enterprises in comprehensive risk management [24]. By analyzing market data and financial indicators, DSS allows for accurate financial risk predictions and provides actionable strategies to help enterprises mitigate potential exposure.

In addition, DSS plays a vital role in investment decisions by integrating diverse investment analysis models. This capability assists companies in evaluating the benefits and risks of different investment scenarios, thereby supporting informed, data-driven investment decisions. DSS effectively processes vast amounts of historical and real-time data, and integrates macroeconomic factors and market trends to provide comprehensive investment analysis and guidance.

3 Research design

3.1 System design principles

System design should prioritize user needs, ensuring usability and operability to allow decision-makers streamlined access to and analysis of financial data. The system requires high-capacity data processing, able to efficiently manage large-scale and diverse financial datasets, thus maintaining data timeliness and accuracy. Advanced data storage and processing technologies, such as distributed databases and parallel computing, should be employed to maximize processing efficiency. The system must exhibit robust scalability, enabling flexible expansion and upgrades to meet evolving enterprise requirements. Multi-level security protocols should be implemented to guarantee the confidentiality and integrity of financial data. Additionally, intelligent analysis and decision-support functionalities should be integrated through data mining, machine learning, and artificial intelligence technologies, facilitating in-depth financial data analysis and informed decision-making.

To ensure clarity and replicability, the methods need precise parameters at each stage of data processing, modeling, and evaluation. For data processing, specific configurations used in the ETL process should be included, such as extraction intervals, transformation rules, and the cleaning criteria. The exact ETL software solution, such as Apache Nifi or Talend, and any SQL or Python scripts used should be specified, detailing the functions or libraries employed for each transformation. In the modeling stage, the algorithms, model parameters, and training settings, such as learning rates or epoch counts in machine learning, should be explicitly listed to allow exact replication. Similarly, evaluation metrics, such as accuracy, recall, or F1 score, and their respective thresholds for decision-making should be clarified. These details ensure that each methodological step is replicable and the results are verifiable, enhancing the study's reliability.

3.2 System architecture design

The architecture design of a financial data analysis and decision support system utilizing big data must thoroughly address data collection, storage, processing, and analysis to ensure effective data management and intelligent decision support. This system adopts a hierarchical structure, comprising the data layer, processing layer, analysis layer, and application layer. The data layer manages the acquisition and storage of financial data from diverse internal and external sources, using distributed databases to facilitate efficient access and management. The processing layer employs big data technologies, such as Hadoop and Spark, to conduct data cleaning, transformation, and pre-processing, establishing a reliable data foundation for subsequent analysis. In the analytics layer, various data mining and machine learning models are integrated to enable comprehensive analysis of financial data through real-time and batch processing, providing predictive and trend analysis capabilities. Finally, the application layer features a user-friendly interface supporting multiple interaction methods, such as chart presentation, report generation, and customized queries, aiding decision-makers in interpreting and applying the analytical outcomes. The architectural design also emphasizes high availability and scalability, incorporating fault-tolerant mechanisms and dynamic scaling techniques to maintain system stability under high loads and in dynamic environments.

3.3 Data source and data processing

3.3.1 Data acquisition and preprocessing

This study primarily collects a range of financial data from external sources, including market data, economic indicators, industry reports, and social media data. These data are accessed through APIs, web crawlers, and third-party data providers. To maintain data quality and consistency, pre-processing steps are undertaken, involving data cleaning, transformation, and integration. Data cleaning addresses missing values, duplicate entries, and outliers by employing methods such as interpolation, mean replacement, and statistical detection. Data transformation standardizes and normalizes data to remove unit and scale disparities between different sources, while data integration consolidates data into a consistent format to ensure logical and structural uniformity.

In this study, ETL (Extraction, Transformation, Loading) tools are used to automate data acquisition and pre-processing, enhancing processing efficiency. These tools extract data from source systems according to predefined workflows and scripts, clean and transform it, and load it into a data warehouse, establishing a high-quality data foundation for subsequent analysis. This process not only improves data accuracy and processing efficiency but also lays the groundwork for the system's real-time data analysis capabilities.

3.3.2 Data storage and management

The data storage scheme selection in this study carefully considers data characteristics, including scale, structure, access frequency, and security, to address diverse storage requirements effectively. A hybrid storage approach is implemented, integrating both relational databases and

NoSQL databases. This combined storage solution accommodates various data types while ensuring flexibility and robustness to support the study's analytical objectives, as shown in Table 3 below.

Table 3: Data storage solutions comparison.

Storage Solution	Advantages	Disadvantages	Suitable Scenarios
Relational Database	Strong transaction processing, high data consistency, supports complex SQL queries	Poor scalability, low efficiency in handling large-scale data	Core financial data, structured data
NoSQL Database (MongoDB)	Flexible data model, supports high concurrency and large-scale data storage, strong scalability	Does not support complex transactions, weaker data consistency	Dynamic data, unstructured and semi-structured data
Distributed File System (HDFS)	Efficient handling of large-scale distributed data, supports parallel computing and batch processing	Poor real-time processing, high data access latency	Large-scale data analysis, batch processing tasks

3.3.3 Data cleaning and conversion

In this study, data cleaning includes identifying and handling missing values, duplicate entries, and outliers, while data conversion covers normalization, scaling, and data format transformations.

The initial step addresses missing values. For datasets with relatively few missing entries, records containing missing data are removed. However, when the extent of missing data is significant, and deletion would compromise analysis, interpolation or mean imputation is applied to estimate missing values. The interpolation method predicts missing values based on trends from adjacent data points, as follows (Formula 4):

$$X_i = \frac{X_{i-1} + X_{i+1}}{2} \quad (4)$$

X_i is the missing value, while X_{i-1} and X_{i+1} represent the previous and subsequent non-missing values, respectively.

The second step involves handling duplicate entries, as duplicates can skew analytical results. Identifying and removing duplicate records ensures data accuracy and uniqueness.

The third step addresses outliers, which are data points that deviate significantly from expected ranges. These outliers are typically detected and managed using the Boxplot method or the standard deviation approach, as shown in (Formula 5):

$$Z = \frac{X - \mu}{\sigma} \quad (5)$$

Z represents the standard fraction, X denotes the data point, μ is the mean, and σ (σ) indicates the standard deviation. When Z exceeds a predefined threshold, these data points are classified as outliers and are either processed or excluded as appropriate.

Data transformation involves both standardization and normalization. Normalization adjusts the data into a standard normal distribution with a mean of 0 and a standard deviation of 1, as demonstrated below (Formula 6):

$$X' = \frac{X - \mu}{\sigma} \quad (6)$$

Normalization adjusts data to fit within the range [0, 1], following this formula (Formula 7):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (7)$$

X represents the original data, while X'' denotes the transformed data. Here, X_{min} and X_{max} represent the minimum and maximum values within the data set, respectively. Adjustments are applied to standardize data values between C and D, ensuring consistent scale and reducing variability for more accurate comparative analysis across the data set. This transformation process supports enhanced data normalization and minimizes distortion, especially when dealing with a wide range of values across different categories.

To ensure consistency in data preprocessing, the standardization and normalization techniques are distinctly applied based on the characteristics of specific financial indicators. Standardization is applied to indicators like revenue and net profit, which vary widely across different periods and can benefit from transformation to a standard normal distribution (mean 0, standard deviation 1), allowing consistent scale for predictive modeling. Normalization, on the other hand, is applied to indicators such as liquidity ratios and turnover ratios, converting their values into a range between 0 and 1 to facilitate comparison across different metrics without scale imbalance. This detailed approach in data preprocessing enhances the model's interpretability, maintains consistency, and improves the model's performance by ensuring that each transformation aligns with the unique properties of the respective financial indicator.

3.4 Module function design

3.4.1 Data analysis module

The data analysis module forms the core of the financial data analysis and decision support system, with its primary function to perform in-depth analysis of collected and stored data, ultimately providing valuable decision-making insights. Python is utilized as the main tool for data analysis, leveraging both data analysis libraries and machine learning libraries to enable efficient processing and analysis. Table 4 below presents the primary functions and advantages of the Python data analysis tools applied in this study.

Table 4: Data analysis tools comparison.

Tool	Main Functions	Advantages
Pandas	Data cleaning, transformation, manipulation; supports DataFrame and Series structures	Easy to use, powerful functions, suitable for structured data
NumPy	Numerical computation, supports large-scale matrix operations and mathematical functions	Fast computation speed, ideal for numerical and matrix operations
Scikit-learn	Machine learning library, includes algorithms for classification, regression, clustering, and dimensionality reduction	Comprehensive functions, easy to use, suitable for various machine learning tasks

3.4.2 Decision support module

The decision support module serves as the core of the financial data analysis and decision support system, primarily designed to provide robust decision support for enterprise management through insights derived from data analysis. This study employs linear regression and decision tree models as key decision support tools, catering to a variety of decision-making scenarios. The functions and advantages of the decision support methods applied in this study are detailed in Table 4 below.

Table 4: Decision support methods comparison.

Method	Main Functions	Advantages
Linear Regression	Predicts and explains the linear relationship between dependent and independent variables	Simple to use, strong interpretability, suitable for linear relationship analysis
Decision Tree	Handles complex decision problems, presents decision paths and results through a tree structure	Easy to understand and interpret, handles nonlinear relationships, high efficiency

3.5 System implementation and deployment

In this study, the system is implemented using a microservice architecture to decouple functional modules, enhancing both flexibility and maintainability. The microservice design deploys and manages each module independently, facilitating interactions through lightweight communication mechanisms. This structure not only improves system scalability but also minimizes the overall impact of potential system failures.

The decision support module incorporates linear regression and decision tree models to provide predictive insights for financial data analysis. However, to enhance accuracy, especially for financial data with non-linear characteristics and temporal dependencies, additional advanced models are considered. Random forest and gradient boosting methods offer greater capacity for handling non-linear data patterns, improving the system's ability to capture complex financial trends. Furthermore, recurrent neural networks (RNN) are introduced to address temporal dynamics, enabling the system to better account for sequence dependencies in financial data. Integrating these models improves the robustness and versatility of the decision support module, ensuring more reliable outcomes in a variety of financial decision-making scenarios.

The experimental setup for implementing and deploying the financial data analysis system includes key specifications on data set size, hardware, software environment, and parameter configurations. The data set comprises financial records totaling 500 GB, processed using Hadoop and Spark for distributed data handling. Hardware configurations include a 64-core CPU with 256 GB RAM and 2 TB SSD storage. Software setup includes Hadoop 3.2.1, Spark 3.1.2, and Docker 20.10 for containerization, running on a Linux-based environment with Ubuntu 20.04. Parameters for Hadoop's HDFS include a block size of 128 MB and a replication factor of three to ensure data reliability. Spark configurations use an executor memory of 8 GB per node and a parallelism setting of 64, optimizing for large-scale data processing. These technical details ensure reproducibility and allow other researchers to accurately compare and benchmark system performance.

For deployment, containerization technology is employed, with each module packaged into an isolated operating environment using Docker containers. Docker provides a lightweight, consistent, and rapid deployment solution, which enhances the system's stability and deployment efficiency. Additionally, Kubernetes is utilized for container orchestration and management. Kubernetes offers robust features for automated deployment, scaling, and management, dynamically adjusting resource allocation based on system load, ensuring stable operation under high concurrency and heavy traffic.

4 Case study

4.1 Case background

Company A, a medium-sized enterprise established in 2010, specializes in the sale of mobile phones and related electronic products. Headquartered in the United States, the company's primary market comprises major cities in China. Company A's product lineup includes smartphones, tablets, smartwatches, and various smart home devices, with a focus on the mid-to-high-end market. The company is dedicated to building customer trust through technological innovation and high-quality service.

Amid the rapid expansion of the smart device market and rising competition, Company A faces significant market pressures and challenges. To address these, the company has increased investments in research and development, as well as marketing, introducing new products that are competitive within the market. The company actively works to optimize supply chain management and control costs, which bolsters operational efficiency and enhances its competitiveness and profitability.

Company A's sales revenue and net profit have shown steady growth; however, cost pressures are also gradually rising. The company employs stringent budget control and financial analysis mechanisms, regularly reviewing financial statements and monitoring financial indicators to promptly identify and address operational challenges. The management of current assets and liabilities remains stable, reflected by a high current ratio and quick ratio, which ensures the company's capability to meet short-term obligations.

Looking ahead, Company A aims to further strengthen its market competitiveness and achieve sustainable development by increasing its investment in research and development over the next five years. These investments will drive product innovation and enhance marketing strategies to expand brand influence. Additionally, Company A seeks to elevate its financial management and decision-making processes by leveraging big data analytics and decision support systems, thereby advancing its strategic goals through informed data analysis and decision-making

4.2 Data selection and processing

To ensure the accuracy and depth of the analysis, data covering key financial indicators of Company A, including sales revenue, cost of sales, net profit, current assets, current liabilities, total assets, and shareholders' equity over the past three years, has been selected. The details are presented in Table 5 below.

Table 5: A company's financial data details.

Year	Sales Revenue (million)	Cost of Sales (million)	Net Profit (million)	Current Assets (million)	Current Liabilities (million)	Total Assets (million)	Shareholder's Equity (million)
2021	58.73	34.21	10.45	25.67	14.32	72.45	45.12
2022	63.29	36.89	11.24	27.89	15.67	78.34	48.67
2023	67.54	39.76	12.13	30.21	16.89	83.56	52.34

The selected data undergoes rigorous cleaning and transformation to ensure consistency and reliability. During data cleaning, missing values, outliers, and duplicate entries are addressed: the interpolation method is applied to estimate missing values, while outliers are identified and managed using the boxplot method. In the data transformation phase, all financial indicators are standardized to remove unit and scale discrepancies, thus enhancing comparability across variables and ensuring uniformity.

4.3 Financial risk early warning analysis

This study uses current ratio, quick ratio, asset-liability ratio and other indicators to evaluate the financial risk of Company A in recent three years. As shown in Figure 2 below.

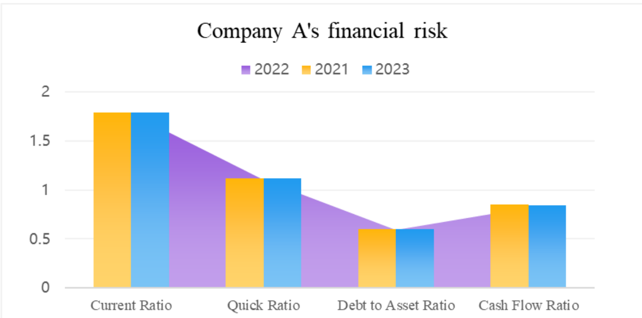

Figure 2: Financial risk of Company A.

In terms of the current ratio, Company A has maintained a level between 1.78 and 1.79 over the past three years, indicating stable short-term debt repayment ability. A current ratio exceeding 1.5 typically suggests that the company possesses sufficient current assets to meet its short-term obligations.

The quick ratio has shown slight fluctuation over the same period, decreasing marginally from 1.12 in 2021 to 1.10 in 2022, before returning to 1.12 in 2023. With a ratio above 1.0, the company demonstrates solid liquidity. The minor dip in 2022 may reflect short-term variations in inventory or accounts receivable management.

Company A's asset-liability ratio has ranged between 0.59 and 0.60 in recent years, reflecting a sound financial structure with a moderate level of debt. A lower asset-liability ratio suggests that the company benefits from leveraging its own capital for expansion and investment, with minimal debt repayment risk.

The financial risk analysis in this section will incorporate a discussion on the rationale for selecting specific ratios, such as the current ratio and quick ratio, by evaluating their direct impact on assessing liquidity and short-term solvency. The choice of these ratios provides a clear snapshot of the company's ability to meet its short-term obligations. To broaden the depth of analysis, a brief comparison with other common financial risk assessment tools, such as Altman's Z-score, is included. Altman's Z-score, a well-regarded predictor of bankruptcy risk, enables a more comprehensive financial risk evaluation by combining profitability, leverage, liquidity, solvency, and activity ratios. Integrating this perspective will complement the traditional liquidity ratios and provide a more robust framework for assessing Company A's financial stability, supporting a holistic approach to financial risk assessment.

Company A's cash flow ratio has remained between 0.82 and 0.85 over the last three years, suggesting that cash flow from operations sufficiently covers short-term liabilities. Although slightly below the ideal benchmark of 1.0, the company would benefit from enhanced cash flow management to ensure sustained financial stability.

The stability of Company A's key financial indicators over the past three years reveals no significant financial risks. However, it remains essential to monitor fluctuations in the quick ratio and manage the cash flow ratio effectively to maintain long-term financial health and mitigate potential risks.

4.4 Analysis of investment decision support

In this study, the net present value (NPV), internal rate of return (IRR), and payback period (PBP) are applied to assess the company's three investment options, guiding the company in selecting the optimal project. These metrics, shown in Figure 3 below, enable a comprehensive comparison of each scheme's profitability, risk, and time to recoup investment, thus providing a robust basis for informed decision-making on capital allocation.

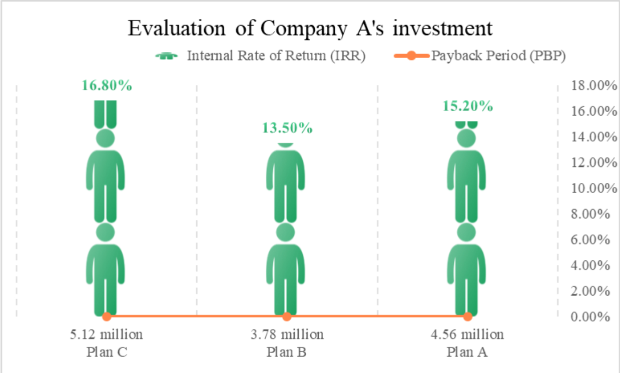

Figure 3: Evaluation of company A's investment plan.

The net present value (NPV) of Plan A is $4.56 million, with an internal rate of return (IRR) of 15.2% and a payback period of 3.5 years. This option presents a favorable combination of NPV and IRR alongside a moderate payback duration, indicating a strong return on investment within manageable risk parameters. Plan B, on the other hand, has an NPV of $3.78 million, an IRR of 13.5%, and a payback period of 4.0 years. Compared to Plan A, Plan B offers a slightly lower NPV and IRR, a longer payback period, and a reduced return on investment, carrying a relatively higher degree of risk. Plan C achieves the highest performance across all indicators, with an NPV of $5.12 million, an IRR of 16.8%, and the shortest payback period at 3.0 years. These results underscore Plan C as the most advantageous option, boasting the highest NPV and IRR and the quickest fund recovery rate.

Therefore, based on the metrics of NPV, IRR, and payback period, Plan C stands out as the optimal investment choice. Plan A ranks as a viable alternative due to its comparatively high return on investment, although it offers a slightly lower IRR than Plan C. Plan B, with its lower NPV, IRR, and longer payback period, is not recommended as the preferred option given its higher associated risks.

4.5 System performance evaluation

To strengthen the performance evaluation of the developed system, a benchmark comparison with state-of-the-art (SOTA) systems is essential. This includes comparing metrics like data processing speed, accuracy, system stability, and user satisfaction against those of current leading systems. Integrating this benchmarking will involve using tables or figures to highlight the developed system's performance in areas such as faster processing speeds, reduced computational costs, or higher system reliability. These metrics, when set against established systems, will clearly demonstrate the advantages and potential improvements achieved by the new system. Such comparative data provides quantitative support for claims of superior performance, offering clearer insights into the system's competitive strengths in financial data processing and decision support.

Comprehensively assess the performance of Company A's financial data analysis and decision support system by evaluating key metrics such as processing speed, data accuracy, system stability, and user satisfaction. This analysis provides insight into the system's operational status and serves as a foundation for future optimization. Details are presented in Table 6 and Figure 4.

Table 6: Performance evaluation of Company A's decision support system.

Evaluation Metric	2021	2022	2023
Data Processing Speed (s)	2.45	2.30	2.15
User Satisfaction (score)	8.2	8.5	8.7
Data Accuracy (%)	98.5	98.7	98.9
System Stability (%)	99.2	99.4	99.5

In terms of data processing speed, Company A's system has shown consistent improvement over the past three years, decreasing from 2.45 seconds in 2021 to 2.15 seconds in 2023. These results indicate the system's robust capability in handling large-scale financial data efficiently, enabling it to complete complex analytical tasks within a reduced timeframe, thereby enhancing overall operational efficiency. Additionally, user satisfaction scores have increased from 8.2 in 2021 to 8.7 in 2023, which reflects the system's continuous optimization in functionality, usability, and performance, contributing to a strengthened user experience and building greater user trust.

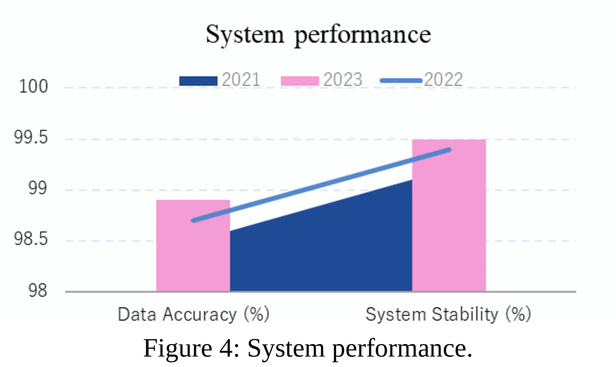

Figure 4: System performance.

The system's accuracy has remained consistently high, increasing from 98.5% in 2021 to 98.9% in 2023. This stability in accuracy ensures the reliability of analysis results, minimizes decision-making errors caused by data discrepancies, and provides a robust data foundation for the company. Operational stability has also improved annually, rising from 99.2% in 2021 to 99.5% in 2023. This trend indicates that the system can maintain steady performance under high-load and extended operation, reducing the likelihood of system failures and downtime, and thus enhancing overall system availability.

In summary, Company A's financial data analysis and decision support system demonstrates strong performance across key metrics, including data processing speed, accuracy, stability, and user satisfaction, fully meeting anticipated performance standards. Continuous monitoring and optimization further enhance the system's capacity to support enterprise financial management needs, offering scientifically reliable decision-making support in a dynamic business environment.

The system performance analysis now incorporates specific criteria to improve clarity and replicability. User satisfaction is evaluated using a structured questionnaire that includes Likert-scale items to measure aspects like system usability, functionality, and efficiency. The questionnaire was distributed to 50 end-users, with results analyzed to generate an average satisfaction score. Processing speed is quantified through system logs that capture average response times during peak and non-peak hours, ensuring consistency in data handling performance. Data accuracy is assessed by tracking error rates in data processing tasks, providing a measure of reliability. Finally, system stability is monitored through uptime statistics and error frequency logs, ensuring sustained performance under variable loads. Including the questionnaire in the appendix enhances transparency, while quantitative metrics solidify the reliability of the performance evaluation.

5 Discussion

5.1 Result discussion

The results show that the financial data analysis and decision support system designed with big data technology has significantly improved the accuracy and efficiency of financial management. By processing large-scale data swiftly, the system enhances data accuracy, reaching up to 98.9% over a three-year evaluation. The system's financial risk early warning function effectively identifies potential risks by monitoring indicators such as the current ratio, quick ratio, and asset-liability ratio. Moreover, the decision support component has helped evaluate investment options through key metrics such as net present value (NPV), internal rate of return (IRR), and payback period (PBP). The system recommended the most favorable investment with an NPV of 5.12 million, IRR of 16.8%, and a payback period of 3.0 years, providing reliable decision-making support for enterprise investment strategies and financial risk control.

The results are compared with the state-of-the-art (SOTA) benchmarks to contextualize the system's performance within current methodologies. Specifically, the system's data processing speed, accuracy, and stability are evaluated against leading benchmarks, highlighting areas of alignment and discrepancy. This analysis reveals that while the current system achieves high processing speed and accuracy, its performance is influenced by factors such as data quality, algorithm efficiency, and scalability. For instance, improvements in data quality directly enhance accuracy, while algorithm efficiency impacts processing speed. The scalability of the architecture, enabled by big data tools like Hadoop and Spark, allows the system to maintain performance with increasing data volume. These results underscore the system's strong positioning relative to SOTA methods, with insights into areas for further refinement.

The detailed analysis of Company A's financial data and decision support system demonstrates the tangible impact of big data technology on financial management. The findings indicate that this financial data analysis and decision support system provides marked advantages in increasing data processing efficiency, enhancing decision accuracy, and optimizing financial management processes.

In terms of data processing speed and accuracy, the system exhibits robust performance with large-scale financial data, completing complex analytical tasks swiftly. High data accuracy contributes to reliable analysis outcomes, reducing errors in decision-making due to data inaccuracies. Implementing big data technology thus improves the efficiency and accuracy of financial data handling, supplying enterprises with a dependable data foundation.

The system's effectiveness in financial risk early warning is also notable. By monitoring and analyzing Company A's key financial indicators, the system identifies potential financial risks promptly, issuing early warnings that enable timely preventive measures. This improves the company's risk management capabilities, supporting a stable financial position amid intense market competition.

Investment decision support analysis results further show that the system can objectively assess various investment options using financial indicators such as net present value (NPV), internal rate of return (IRR), and payback period (PBP), assisting companies in selecting the most advantageous investment projects. This robust analytical support promotes efficient fund utilization, maximizing return on investment. Additionally, performance evaluations verify the system's stability and user satisfaction. The system maintains reliable performance under high loads and continuous operation, reducing system failures and downtime, thereby enhancing availability. Increased user satisfaction reflects ongoing system optimization in functionality, usability, and performance, fostering user trust and engagement.

5.2 Enlightenment and suggestions for enterprise financial management

5.2.1 Improve data processing and analysis capabilities

Enterprises are increasingly adopting big data technology to strengthen financial data processing and analysis capabilities. By employing advanced data storage, processing, and analysis tools, they achieve efficient management and in-depth examination of large volumes of financial data, thereby providing more accurate and timely support for decision-making. This approach improves data processing efficiency, enhances data accuracy and reliability, and supports enterprises in sustaining a competitive edge within a highly competitive market environment.

5.2.2 Strengthen risk management

Big data technology is highly effective in providing early warnings for financial risks, enabling enterprises to establish robust financial risk early warning systems. Such systems allow companies to detect and mitigate potential risks in a timely manner through real-time monitoring and analysis of essential financial indicators. This capability strengthens enterprises' risk management frameworks, fortifying their financial stability and responsiveness to unexpected events. Regular risk assessments enable enterprises to adjust financial strategies proactively, thereby maintaining financial security and resilience.

5.2.3 Optimize the investment decision-making process

Enterprises can leverage big data technology to conduct a systematic evaluation of investment projects, allowing for the selection of optimal investment options. By using comprehensive indicators such as net present value (NPV), internal rate of return (IRR), and payback period (PBP), investment returns and risks can be more accurately forecasted, leading to more informed decision-making. This approach enhances investment efficiency, maximizes return on investment, and facilitates optimal resource allocation, thereby supporting strategic financial management and risk mitigation.

5.2.4 Improving system stability and user satisfaction

The stability and user satisfaction of a financial management system are critical to its success, as enterprises must continually enhance system functionality to ensure consistent performance under high load and extended operation periods. Regular improvements to system capabilities not only enhance the user experience but also build user trust and reliance on the system, thus increasing its effectiveness and adoption. Enterprises can benefit from systematically conducting system performance evaluations and user feedback surveys to inform continuous improvements and refine system functionality over time, ensuring that the system aligns with user expectations and operational demands.

5.3 Research limitations and future prospects

5.3.1 Research limitations

This study achieved notable progress in applying big data technology to financial data analysis and decision support; however, limitations remain. Despite implementing various data cleaning and preprocessing techniques, data accuracy and consistency are constrained by the limitations of data acquisition channels and the inherent quality of the data itself. Additionally, this study employs linear regression and decision tree models, which may not fully capture the intricacies of certain financial data types and thus may not address all financial management scenarios. Furthermore, the system's performance evaluation relies primarily on simulated data within an experimental setting, meaning that real-world complexity and variability could influence system performance and stability, necessitating further adjustments in practical applications.

5.3.2 Future outlook

Future research should enhance data collection and cleaning techniques to ensure higher data accuracy and consistency. Expanding data sources and improving data processing algorithms will be essential in raising data quality. In terms of model optimization and innovation, exploring additional analysis models and algorithms—such as deep learning and reinforcement learning—will be beneficial for enhancing the system's analytical and predictive capabilities in various financial management contexts. Furthermore, potential applications of big data technology in financial management, including real-time financial monitoring, intelligent report generation, and automated financial decision-making, should be further explored. Integration with emerging technologies like artificial intelligence and blockchain can advance the intelligence and security of financial management systems.

Focusing on improving user experience is also crucial; enhancing the user interface and interaction features will increase the system's usability and user satisfaction. Future research should emphasize interdisciplinary collaboration, combining insights and techniques from fields such as financial management, information technology, and data science to drive continuous innovation in financial data analysis and decision support systems. Ongoing technological advancements and expanding applications will offer more scientifically robust, intelligent, and efficient solutions for enterprise financial management.

References

[1] Casturi R, Sunderraman R (2022). Cost effective, rule based, big data analytical aggregation engine for investment portfolios. Wireless Networks. 28(3):1203-1209. https://doi.org/10.1007/s11276-018-01904-5.

[2] Fahd K, Miah SJ (2023). Designing and evaluating a big data analytics approach for predicting students' success factors. Journal of Big Data. 10(1):159. https://doi.org/10.1186/s40537-023-00835-z.

[3] Dutta G, Rao HV, Basu S, Tiwari MK (2019). Asset liability management model with decision support system for life insurance companies: Computational results. Computers & Industrial Engineering. 128:985-998. https://doi.org/10.1016/j.cie.2018.06.033.

[4] Lu J, Wu D, Dong J, Dolgui A(2023). A decision support method for credit risk based on the dynamic Bayesian network. Industrial Management & Data Systems. 123(12):3053-3079. https://doi.org/10.1108/IMDS-04-2023-0250.

[5] Talamo A, Marocco S, Tricol C (2021). "The Flow in the Funnel": Modeling Organizational and Individual Decision-Making for Designing Financial AI-Based Systems. Frontiers in Psychology. 12:697101. https://doi.org/10.3389/fpsyg.2021.697101.

[6] Peng J, Bao L (2023). Construction of enterprise business management analysis framework based on big data technology. Heliyon. 9(6):e17144. https://doi.org/10.1016/j.heliyon.2023.e17144.

[7] Popovic A, Hackney R, Tassabehji R, Castelli M (2018). The impact of big data analytics on firms' high value business performance. Information Systems Frontiers. 20(2):209-222. https://doi.org/10.1007/s10796-016-9720-4.

[8] Liu M, Gao R, Fu W (2021). Analysis of Internet Financial Risk Control Model Based on Machine Learning Algorithms. Journal of Mathematics. 2021:8541929. https://doi.org/10.1155/2021/8541929.

[9] Yu H, Guo Z (2022). Design of Underground Space Intelligent Disaster Prevention System Based on Multisource Data Deep Learning. Wireless Communications & Mobile Computing. 2022:3706392. https://doi.org/10.1155/2022/3706392.

[10] Zheng T, Chen G, Wang X, Chen C, Wang X, Luo S (2019). Real-time intelligent big data processing: technology, platform, and applications. Science China-Information Sciences. 62(8):082101. https://doi.org/10.1007/s11432-018-9834-8.

[11] Xue CJ (2019). Special Issue on Emerging Technologies for Big Data Processing. Journal of Electronic Science and Technology. 17(1):1-2. https://doi.org/10.3969/j.issn.1674-862X.2019.01.001.

[12] Xu Z, Zhou W (2021). A Data Technology Oriented to Information Fusion to Build an Intelligent Accounting Computerized Model. Scientific Programming. 2021:6031324. https://doi.org/10.1155/2021/6031324.

[13] Rodrigues M, Santos MY, Bernardino J (2019). Big data processing tools: An experimental performance evaluation. Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery. 9(2):e1297. https://doi.org/10.1002/widm.1297.

[14] Jensen MH, Nielsen PA, Persson JS (2023). From Big Data Technologies to Big Data Benefits. Computer. 56(6):52-61. https://doi.org/10.1109/MC.2022.3206032.

[15] Li K, Chen Y (2021). Fuzzy Clustering-Based Financial Data Mining System Analysis and Design. International Journal of Foundations of Computer Science. 33(06N07):603-624. https://doi.org/10.1142/S0129054122420060.

[16] Wang S (2020). Research on Data Mining and Investment Recommendation of Individual Users Based on Financial Time Series Analysis. International Journal of Data Warehousing and Mining. 16(2):64-80. https://doi.org/10.4018/IJDWM.2020040105.

[17] Guo Y (2021). CNS: Interactive Intelligent Analysis of Financial Management Software Based on Apriori Data Mining Algorithm. International Journal of Cooperative Information Systems. 30(1-4). https://doi.org/10.1142/S0218843021500088.
[18] Liang M (2021). Optimization of Quantitative Financial Data Analysis System Based on Deep Learning. Complexity. 2021:5527615. https://doi.org/10.1155/2021/5527615.
[19] Lin Y, Yue H, Liao H, Li D, Chen L (2022). Financial Risk Assessment of Enterprise Management Accounting Based on Association Rule Algorithm under the Background of Big Data. Journal of Sensors. 2022:8041623. https://doi.org/10.1155/2022/8041623.
[20] Zhai Z, Martinez JF, Beltran V, Martinez NL (2020). Decision support systems for agriculture 4.0: Survey and challenges. Computers and Electronics in Agriculture. 170:105256. https://doi.org/10.1016/j.compag.2020.105256.
[21] Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digital Medicine. 3(1):17. https://doi.org/10.1038/s41746-020-0221-y.
[22] Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, Rashidi P, Upchurch GR Jr, Bihorac A (2020). Artificial Intelligence and Surgical Decision-making. JAMA Surgery. 155(2):148-158. https://doi.org/10.1001/jamasurg.2019.4917.
[23] Kwan JL, Lo L, Ferguson J, Goldberg H, Diaz-Martinez JP, Tomlinson G, Grimshaw JM, Shojania KG (2020). Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. BMJ. 370:m3216. https://doi.org/10.1136/bmj.m3216.
[24] Lutz W, Deisenhofer AK, Rubel J, Bennemann B, Giesemann J, Poster K, Schwartz B (2022). Prospective evaluation of a clinical decision support system in psychological therapy. Journal of Consulting and Clinical Psychology. 90(1):90-106. https://doi.org/10.1037/ccp0000642.