SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

Journal homepage: www.dmame-journal.org
ISSN: 2560-6018, eISSN: 2620-0104

AI-Driven Financial Risk Early-Warning Using TD-DS-RF: A Decision Support Model Integrating Multi-Source Data

Tianjiao Zhang[1], Qiangxing Chen[1], Qingqing He[1*], Weidong Zhu[2],

1 Anhui Business and Technology College, Hefei, 231131, Anhui, China.
2 HeFei University of Technology, Hefei, 230009, Anhui, China.

ARTICLE INFO

Article history:
Received 5 October 2024
Received in revised form 19 Jan 2025
Accepted 27 February 2025
Available online 30 March 2025

Keywords:
Two-Dimensional Evidence Theory; Random Forest Algorithm; Financial Risk Early-Warning; Management Discussion and Analysis; Decision-Making Support

ABSTRACT

With the swift advancement of information technology, digitalisation and intelligence have emerged as intrinsic drivers of reform and progress in enterprise risk management. Conventional approaches to financial risk early-warning predominantly rely on structured data, exhibit limited exploration of unstructured information, inadequately address the varying quality of information sources, and generally apply uniform treatment to multi-source data. In response to these shortcomings, the present study harnesses the advantages of evidence theory in managing uncertain information fusion by incorporating two-dimensional evidence theory into the financial risk early-warning framework. By integrating this with a random forest algorithm, a novel financial risk early-warning model—termed TD-DS-RF (Two-Dimensional Dempster-Shafer with Random Forest)—is established. Additionally, a financial risk early-warning lexicon is devised for publicly listed enterprises. Empirical validation utilises data from Chinese manufacturing firms listed between 2012 and 2021. The findings affirm that the TD-DS-RF model exhibits strong performance in real-world contexts, furnishing reliable decision-making support for both stakeholders and regulatory bodies, while contributing a novel conceptual lens to the domain of financial risk early-warning.

1. Introduction

Amidst the rapid evolution of information technology, financial risk early-warning for listed enterprises has taken on new features, notably the emergence of multi-source heterogeneous data and heightened uncertainty in evaluating and integrating risk-related information. The proliferation of enterprise data, combined with the dynamism of external conditions and the increasing complexity of financial risk, underscores the urgent necessity for robust financial risk management. For listed entities, early-warning mechanisms are vital to address potential risks proactively and to support sustained, healthy development. Management should proactively adopt data-driven approaches, integrate artificial intelligence (AI) technologies, construct intelligent early-warning systems, and thereby accurately identify and pre-empt financial risks, ensuring the enterprise's sustainable growth and operational stability.

* *Corresponding author.*
E-mail address: hqq86188@163.com

https://doi.org/10.31181/dmame8120251391

Compared to conventional warning techniques, financial risk early-warning models underpinned by AI algorithms exhibit marked improvements in accuracy and stability [1]. A range of algorithms, including XGBoost [2], random forest (RF), evidence theory [3], and support vector machines [4], have been widely employed in financial risk forecasting. While the big data environment affords an abundance of relevant financial risk data, discrepancies in the relevance and dependability of information from various sources pose significant challenges to model performance. Prevailing models often process all data uniformly, failing to accommodate the differing quality levels inherent in diverse data sources. This shortcoming highlights the pressing need for a novel early-warning model capable of intelligently assimilating multi-source data.

In selecting early-warning indicators, scholars typically formulate index systems across several dimensions: financial [5], corporate governance [6], and managerial cognition [7]. Among these, Management Discussion and Analysis (MD&A)—a key component of listed firms' disclosure practices—blends retrospective and forward-looking insights, encompasses both financial and non-financial information, and presents investors with a managerial perspective on prospects. MD&A has proven valuable in predicting financial distress [8] and constitutes a crucial source of non-financial indicators. Despite the big data environment's ability to deliver diverse financial risk information, scholars often focus on mining emotional sentiment embedded within this data. However, comprehensive extraction and risk profiling from unstructured financial risk content remain a significant research challenge. Thus, effectively employing text mining techniques to extract and interpret unstructured financial risk information constitutes a pivotal issue in the domain of financial early-warning systems.

To overcome the limitations observed in current studies, particularly the inadequate mining of unstructured risk data and the neglect of quality variance in information sources, this study integrates two-dimensional evidence theory, recognising the fundamental uncertainty of financial risks. Expanding upon the authors' previous research [3], the study incorporates this theory within a random forest framework to propose a novel financial risk early-warning model for listed firms, termed TD-DS-RF. This model is capable of fusing risk-related information from multiple data sources. Within the two-dimensional evidence framework, financial indicators from annual financial statements and non-financial indicators extracted from the 'Possible Risks' section of the MD&A are adopted as primary sources of evidence. The initial dimension of evidence is derived by applying random forest to obtain the basic probability assignment (BPA) from early-warning outputs. To account for the quality of the evidence sources, factors such as audit opinion and text readability within the annual reports are introduced as the second dimension, reflecting the reliability of both financial and non-financial indicators. Following the adjustment of the first evidence dimension, the Dempster-Shafer (D-S) combination rule is applied to derive the final early-warning outcome. The model is empirically validated using data from Chinese manufacturing firms listed on the stock exchange. By employing two-dimensional evidence theory to integrate heterogeneous, multi-source financial risk data, the model offers meaningful managerial insights and decision support to stakeholders and enhances regulatory oversight capabilities.

The main contributions of this study are summarised in three key areas:

(1) A financial risk early-warning model for listed companies is developed, based on the TD-DS-RF approach. This model incorporates the quality and reliability characteristics of risk information sources into the early-warning process and optimises the synergy among various AI methodologies. The integration of a second evidential dimension into classical evidence theory allows the model to reflect variations in data quality. Given the objective and uncertain nature of financial risks, inconsistencies in data quality can substantially affect predictive outcomes. The TD-DS-RF model mitigates these effects by exploiting the strengths of complementary AI algorithms, thereby

overcoming the limitations of prior models that overlooked information source quality. This enhancement significantly improves both the performance and practical utility of the model.

(2) Through the application of text mining techniques, this research articulates the types and severity of risks embedded within unstructured financial data disclosed in listed firms' annual reports. Specifically, by extracting and refining the 'Possible Risks' section in MD&A reports, the study identifies keywords and classifies different risk categories. This process results in the construction of a comprehensive keyword dictionary for financial risk indicators, enabling more precise identification and assessment of the nature and magnitude of risks. Empirical findings validate the effectiveness of the proposed non-financial indicators, demonstrating that their integration alongside traditional financial metrics significantly enhances the predictive accuracy and dependability of the financial risk early-warning model.

(3) This study advances the application of two-dimensional evidence theory and broadens its practical relevance. For the first time, this theoretical framework is applied to financial risk early-warning in the context of listed firms, enabling differentiated assessments of data source reliability and improving the precision of early-warning signals. The developed model provides solid decision-making support for both stakeholders and regulatory authorities, thereby contributing to the stability of the capital market. Overall, this research offers both theoretical enrichment and practical insights, reinforcing the foundations of financial risk management.

The remainder of this paper is structured as follows: The next section presents a review of the relevant literature. Section 3 outlines the construction of the financial risk early-warning index system. Section 4 details the development of the proposed model. Section 5 contains the empirical analysis, followed by a concluding section that summarises the principal findings.

2. Literature Review

Financial risk early-warning remains a prominent area of interest within academic research [9]. This paper reviews the extant literature from three principal perspectives: studies concerning the determinants influencing financial risk; research focused on the methodologies employed in financial risk early-warning; and investigations related to the development and application of two-dimensional evidence theory.

2.1 Relevant Research on the Influencing Factors of Financial Risk

The financial risks encountered by enterprises are shaped by a wide range of internal and external factors, collectively heightening the uncertainty surrounding their financial status [10]. In financial risk early-warning research, scholars typically construct an indicator system and then evaluate the financial health of enterprises based on the risk information reflected in these indicators [11]. Existing research on the factors influencing the financial risk indicator system mainly includes financial, corporate governance, and management cognition dimensions [9].

Financial dimension indicators are typically derived from the financial statements of listed companies. Altman was one of the first to use seven financial indicators, such as the interest coverage ratio, return on total assets, and current ratio, to predict the likelihood of a company encountering financial risk [12]. Other scholars have selected a set of 12 financial indicators, including ROE, accounts receivable turnover, and operating profit growth rate, to evaluate financial risk [13]. In these studies, scholars often choose indicators that reflect a company's cash flow levels, operational capacity, profitability, and solvency when conducting financial risk early-warning [14]. Corporate governance factors typically encompass equity structure and the board of directors' characteristics. Research has demonstrated that board independence plays a significant role in influencing the likelihood of financial risk within companies [15].

The cognitive dimensions of management largely focus on information disclosed in the MD&A. MD&A reflects management's willingness to communicate with the market [16], provides future-oriented internal data [17], and helps predict future company performance. Most studies employ the Loughran-McDonald Dictionary of Financial Emotion [18] to analyse managerial tone in MD&A, treating it as a primary indicator. It has been shown that an emotional tone can reflect a company's stance on sustainable operations, thus aiding in the accurate prediction of its future performance.

With the advancement of artificial intelligence, scholars are increasingly focusing on extracting valuable insights from text data and applying them to areas such as stock prediction [19] and financial distress forecasting [20]. Previous research has found a connection between the language used in MD&A and corporate bankruptcy [21]. Researchers have suggested that MD&A provides stakeholders with an insight into uncertain factors affecting a company's future, which may lead to complicated texts or instances where management downplays negative information and emphasises positive aspects [22], causing information asymmetry. Much of the current research focuses solely on the tone conveyed in texts, which can distort financial risk warning results. Some studies indicate that integrating financial, managerial, and textual features in risk warning models produces more accurate results than using sentiment indicators alone [23]. As such, further research is needed to focus on the textual content related to risk disclosures in annual reports and extract the information conveyed within the text itself.

2.2 Research on the Financial Risk Early-Warning Model

Initial studies on financial risk warning focused on univariate models, later advancing to multivariate ones. In the big data era, researchers introduced AI-based models for enhanced prediction. Univariate models, such as those using return on total assets, are straightforward but limited, as they fail to reflect the full financial profile of a firm, leading to subjective choices and unstable outcomes [24]. Martin's multivariate logistic model improved bankruptcy prediction over the Z-Score but required high-quality, normally distributed data [25]. Despite relaxed assumptions in logistic models, both approaches rely on static data, causing delays in identifying risks.

To address this, scholars have moved towards dynamic, AI-driven approaches. Neural networks have shown consistent performance in forecasting. Others have applied decision trees [26], RF [3], and SVM [4], improving traditional model accuracy. RF, enhancing decision tree logic, better matches decision-makers' reasoning and surpasses logistic, SVM, and neural network models in accuracy [25]. The authors in [3] applied RF with evidence theory to form the DS-RF model, boosting prediction accuracy for agricultural enterprises. Although AI models improve flexibility and accuracy Zhao et al. [27], challenges remain. These models often neglect the verification of accounting data reliability, which can distort results. Hence, precise financial risk forecasting remains difficult with current methods.

2.3 Two-Dimensional Evidence Theory and Its Application

Dempster originally introduced evidence theory, later expanded by Shafer, as a method for reasoning under uncertainty within AI [28]. Its strength in managing ambiguous data has led to its application in fault diagnosis, pattern recognition, and risk prediction [29]. In financial contexts, scholars have used RF to derive BPA via evidence theory, optimising risk warnings for large firms. The authors in [30] integrated evidence theory with SHAP to build an interpretable model for listed companies. However, traditional evidence theory overlooks variations in source quality, which can skew decisions. To overcome this, the authors in [31] introduced two-dimensional evidence theory, adding a second framework to capture source quality. This model resolves conflicts between data sources and accounts for their reliability, proving useful in areas like investment and fund evaluation [32]. Despite its advantages, this enhanced framework has yet to be applied in financial risk warnings to assess source quality impact.

Current research largely depends on financial statement data, lacking integration of varied risk sources, which limits the development of a holistic risk index system. Though some studies analyse sentiment in MD&A—covering historical, predictive, quantitative, and qualitative content—this information is often biased due to managerial manipulation. As such, deeper analysis of MD&A risk texts is essential. Existing models typically rely on standard financial or industry terms, which limit objectivity and prevent the creation of a standardised financial risk dictionary. Moreover, traditional models often assume data validity and process reports without confirming data quality, leading to flawed warnings. While two-dimensional evidence theory addresses disparities in data source quality and reliability, it has not yet been incorporated into financial risk warning systems.

To address existing gaps, this study focuses on Chinese listed manufacturing firms and introduces a financial risk warning model that integrates RF with two-dimensional evidence theory. The risk index system combines financial data from annual reports and non-financial data extracted from the 'possible risks' section in MD&A. Audit opinion' and 'text readability' are used to assess the credibility of financial and non-financial indicators, respectively, enhancing the classical one-dimensional evidence framework. After adjusting the first evidence dimension, the D-S synthesis rule generates the final risk warning outcome. Through structured integration of diverse data sources, two-dimensional evidence theory offers valuable support for stakeholder decision-making.

2.4 Construction of the Financial Risk Early-Warning Index System

This paper develops a financial risk early-warning index system based on two dimensions. The first dimension consists of financial and non-financial indices that influence the financial risk of enterprises. The second dimension selects indicators that reflect the quality, reliability, and other characteristic information associated with the first dimension's evidence.

2.5 Selection of Two-Dimensional Financial Indicators

(1) Selection of the First-Dimension Financial Indicators

Building on the team's previous research [3] and related findings [4], this study constructs a comprehensive financial risk early-warning evaluation index system based on five key dimensions: profitability, operational capability, solvency, growth potential, and cash flow acquisition capability. A total of 21 financial indicators are selected to ensure scientific rigour and practical relevance in reflecting a company's financial health and potential risk. The initial indicators are presented in Table 1.

Table 1
Financial Risk Early-Warning Initial Financial Index System for Listed Companies

Evaluation Dimension	Indicators
Profitability	Capital Return, ROA, ROE, Operating Profit Ratio, Ratio of Profits to Cost
Operation Capability	Accounts Receivable Turnover, Inventory Turnover, Current Assets Turnover, Fixed Asset Turnover, Total Assets Turnover
Solvency	Current Ratio, Quick Ratio, Interest Cover, Debt to Asset Ratio
Growth Potential	Equity Grow Ratio, Rate of Capital Accumulation, Total Assets Growth Rate, Revenue Growth Rate
Cash Flow Acquisition Capability	Main Income Cash Cover, Company Cash Flow, Free Cash Flow

Due to the large number of initial indicators and the issue of multicollinearity, it is necessary to screen and eliminate redundant indicators to reduce model complexity and enhance the early-warning model's generalisation ability. The Least Absolute Shrinkage and Selection Operator (LASSO) model, known for its fast computation, ease of interpretation, and ability to avoid overfitting, is widely used. By incorporating L1 regularisation in linear regression, the LASSO model compresses insignificant variable coefficients to zero. In this study, LASSO was employed, with an optimal lambda value of 0.0005 derived from LASSO regression. The optimal lambda value was used to obtain the coefficients of the initial financial indicators, as shown in Table 2.

Table 2
LASSO Regression Results Based on the Best Lambda Values

Indicators	Coefficients
Capital Return	0
ROA	-0.129394
ROE	-0.034815
Operating Profit Ratio	0
Ratio of Profits to Cost	0
Accounts Receivable Turnover	0
Inventory Turnover	0.010938
Current Assets Turnover	0
Fixed Asset Turnover	0.013243
Total Assets Turnover	-0.070747
Current Ratio	0.022675
Quick Ratio	0
Interest Cover	0
Debt to Asset Ratio	0.107551
Debt to Asset Ratio	-0.000995
Rate of Capital Accumulation	0
Total Assets Growth Rate	0
Revenue Growth Rate	-0.032913
Main Income Cash Cover	0.000024
Company Cash Flow	0.021889
Free Cash Flow	-0.000246

Moreover, as shown in Table 2, following LASSO regression, a total of 12 indicators with non-zero coefficients were retained, while those with zero coefficients were excluded. The final financial index system for this study, after the selection process, is presented in Table 3.

Table 3
Financial Index System of Financial Risk Early-Warning of Listed Companies

Evaluation Dimension	Indicators
Profitability	ROA, ROE
Operation Capability	Inventory Turnover, Fixed Asset Turnover, Total Assets Turnover
Solvency	Current Ratio, Debt to Asset Ratio
Growth Potential	Equity Grow Ratio, Revenue Growth Rate
Cash Flow Acquisition Capability	Main Income Cash Cover, Company Cash Flow, Free Cash Flow

(2) Selection of the Second Dimension Financial Indicators

In two-dimensional evidence theory, the second-dimensional evidence reflects the quality and reliability of the first-dimensional evidence. Consequently, when selecting the second dimension for financial data, it is essential to choose relevant factors that reflect the quality of information, particularly its reliability. Accounting information serves as a critical foundation for investors' decisions, and its quality directly influences investment risk. Recent years have seen a rise in financial fraud cases among listed companies, emphasising the essential role of external audit supervision.

Within external audits, the audit opinion issued by certified public accountants acts as a signal transmitting hidden information from the listed companies, thus indicating the reliability of the accounting information. Some scholars have discovered that going-concern audit opinions can prevent management from manipulating the tone in forward-looking texts. Therefore, this study selects third-party audit opinions as the second-dimension indicator for financial data to assess the reliability of the financial information used in early-warning financial risk analysis. The five types of audit opinions on financial statements are: standard unqualified opinion; unqualified opinion with an emphasis of matter paragraph; qualified opinion; adverse opinion; and unable to comment.

2.6 Selection of Two-Dimensional Non-Financial Indicators

(1) Selection of the First-Dimension Non-Financial Indicators

Currently, scholars are focusing on extracting more valuable information from semi-structured or unstructured data. Both domestic and international researchers have confirmed that the MD&A section in the annual reports of listed companies serves as a significant source of information for financial risk warnings. According to the ‘Content and Format Rules for Information Disclosure by Publicly Issued Securities Companies No. 2 - Content and Format of Annual Reports’ issued by the China Securities Regulatory Commission (CSRC), it is stipulated that companies should offer a projection of their future development. This includes an analysis of their strategies, business plans for the upcoming year, and potential risks, with an emphasis on quantitative analysis. Specifically, for ‘possible risks’, the announcement outlines detailed guidelines: ‘Companies should disclose risk factors that could adversely affect the realization of their future development strategies and business objectives based on their own characteristics, adhering to the principles of relevance and materiality (such as policy risks, industry-specific risks, exchange rate risks, interest rate risks, technological risks, raw material price and supply risks, financial risks, impairment risks of goodwill and other assets, as well as significant impacts on core competitiveness due to equipment issues or the resignation of technical personnel, etc.)’

This paper focuses on selecting and mining non-financial indicators from the 'possible risks' section disclosed in the MD&A of annual reports. Text mining will be applied to the 'possible risks' section of listed companies, using the risk types disclosed in the CSRC announcement and high-frequency risk categories identified through manual reading of texts as seed words to construct a financial risk warning dictionary for listed companies. The detailed process flowchart for constructing non-financial indicators is presented in Figure 1.

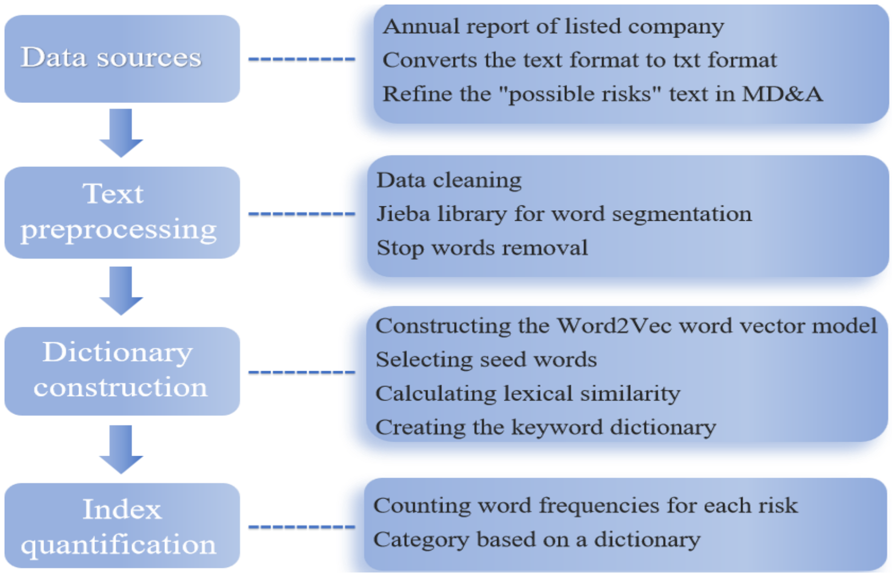

Fig 1. Flowchart of Non-Financial Index System Construction for Financial Risk Early-Warning

The extraction and mining of non-financial indicators in this study are carried out through the following steps:

Step 1: Data Sources

Non-financial indicators are derived from the annual reports of Chinese listed manufacturing companies from 2012 to 2021, sourced from the China Research Data Service Platform (CNRDS) and the Jushao Information Network (http://www.cninfo.com.cn/new/index). The ‘possible risks’ text from the MD&A section is extracted from these reports, and the format is converted from PDF to TXT.

Step 2: Text Pre-Processing

Data Cleaning: Punctuation marks and numbers are removed, leaving only Chinese characters to facilitate word segmentation.

Text Segmentation: The Jieba thesaurus is used for word segmentation, which efficiently classifies words using a dictionary tree. After segmentation, a stop word list is applied, removing irrelevant terms from the text.

Step 3: Dictionary Construction

The processed texts are compiled into a corpus, and a Word2Vec word vector model is established to transform words into word vectors [33]. Based on risk types disclosed by the CSRC and manual reading, the study identifies key risk categories such as 'policy risk', 'industry risk', 'operating risk', 'environmental risk', 'exchange rate risk', 'interest rate risk', 'technical risk', 'management risk', 'investment risk', 'credit risk', and 'default risk'. Using the cosine similarity algorithm, the word vectors of these risk categories are ranked, and the top 50 most similar words for each category are selected. These are then used to form an initial risk keyword dictionary. The final dictionary categorises risks into five categories: core competitiveness risk, operating risk, debt and liquidity risk, industry risk, and macro environment risk. Some of these dictionaries are presented in Table 4.

Table 4
Listed Companies’ Financial Risk Early-Warning Part of the Dictionary

Risk Category	Keyword Dictionary
Core Competitiveness Risk	Technical Risk, Innovation, New Products, New Technologies, Process Development, Upgrading, Renewal and Replacement, Core Technology, Talent Loss, Design, Confidentiality, R&D, Achievements, Performance, Production, Biotechnology, High-tech, Precious Metals, Technological Advancement, Professional Knowledge, Advanced Technology, R&D Personnel, etc.
Operating Risk	Operational Risk, Slowdown, Economic Downturn, Stagnation, Business Pressure, Long Cycle, Dependence, Management, Revenue, Business Model, Fixed Costs, Challenges, Asset Losses, Liquidity, Loss of Control, Price Competitiveness, Material Consumption, Gross Margin, Asset Quality, Price Differential, Safety Stock, Economic Regulation, Cyclical Fluctuations, etc.
Debt and Liquidity Risk	Credit Risk, Default, Fulfilment of Obligations, Credit Limit, Press for Money, Limit, Pricing Method, Delayed Payment, Contract Disputes, Negotiation, Periodic, Hidden Hazards, Commitment, Fixed Costs, Termination of Contract, Payments, Credit Record, Audit, Receivable Bills, Signing, Recall, Overdue, Liability, Credit Rating, Repayment, Lease Term, Illegal, etc.
Industry Risk	Industry Risk, Cyclical Fluctuations, Cyclical Recession, Industry Regulation, Impact, Decline, Economic Shock, Economic Volatility, Supply Exceeding Demand, Insufficient Demand, Consumer Demand, Severe, Negative Reporting, Significant Impact, Capacity Shortfall, Overcapacity, Downwards Volatility, Decrease, Downstream Market, Instability, Slowing Growth, Economic Stagnation, Price Decline, Excess Capacity, etc.
Macro Environment Risk	Legal Risk, Reputation, Claims, Legal Liability, Litigation, Contracts, Disputes, Legal Advisor, Violation, Non-compliance, Compensation, Corporate Law, Plaintiff, Force Majeure, Legal Rights, Legal Proceedings, Settlement Currency, Infringement, Domestic and Foreign, Freezing, Mortgage, Limits, Cases, Operational Efficiency, Asset Loss, Penalties, etc.

Finally, based on the 17 subdivided risk types, this study compiles and structures the non-financial indicator system, which encompasses five risk categories. The detailed indicator system is presented in Table 5.

Table 5
Non-Financial Index System of Financial Risk Early-Warning of Listed Companies

Evaluation Dimension	Indicators
Core Competitiveness Risk	Technical Risk, Management Risk
Operating Risk	Business Risk, Cost Control Risk, Operational Risk, Investment Risk
Debt and Liquidity Risk	Credit Risk, Cash Flow Risk, Receivables and Payables Risk
Industry Risk	Industry Cyclical Risk, Industry Market Risk
Macro Environment risk	Legal Risk, Exchange Rate Risk, Interest Rate Risk, Policy Risk, Macroeconomic Risk, Environmental Protection Risk

Step 4: Index Quantification

In accordance with the risk types of the non-financial indicator system, keyword frequency statistics are performed using the financial risk early-warning dictionary. The frequency count of each risk type is then used as the indicator value to quantify non-financial risk information.

(2) Selection of the Second-Dimension Non-Financial Indicators

The second dimension of non-financial indicators focuses on the quality and reliability of the ‘possible risks’ text in the MD&A section. Management may manipulate readability to obscure negative information, misleading investors. To assess readability, this study uses the ratio of specialized vocabulary to total words, as excessive financial jargon can reduce comprehension and mislead investors. This measure reflects the reliability of non-financial indicators.

2.7 Financial Risk Early-Warning Index System for Listed Companies Based on Two-Dimensional Evidence Theory

In summary, the researchers identified a total of 29 indicators pertinent to financial risk early-warning, comprising 12 financial indicators and 17 non-financial indicators. Drawing upon two-dimensional evidence theory, secondary indicators were selected within both the financial and non-financial domains to assess the reliability of the corresponding information sources. A comprehensive outline of the indicator system is presented in Table 6.

Table 6
Financial Risk Early-Warning Index System of Listed Companies Based on the Two-Dimensional Evidence Theory

Information Dimension	The First Dimension		The Second Dimension
	Evaluation Dimension	Indicators	
Financial Indicators	Profitability	ROA	Audit Opinion
		ROE	
	Operation Capability	Inventory Turnover	
		Fixed Asset Turnover	
		Total Assets Turnover	
	Solvency	Current Ratio	
		Current Ratio	
	Growth Potential	Equity Grow Ratio	
		Revenue Growth Rate	
	Cash Flow Acquisition Capability	Main Income Cash Cover	
		Company Cash Flow	
		Free Cash Flow	
Information	The First Dimension		The Second Dimension

Dimension	Evaluation Dimension	Indicators	
Non-Financial Indicators	Core Competitiveness Risk	Technical Risk,	Text Readability
		Management Risk	
	Operating Risk	Business Risk	
		Cost Control Risk	
		Operational Risk	
		Investment Risk	
	Debt and Liquidity Risk	Credit Risk	
		Cash Flow Risk	
		Receivables and Payables Risk	
	Industry Risk	Industry Cyclical Risk	
		Industry Market Risk	
	Macro Environment Risk	Legal Risk	
		Exchange Rate Risk	
		Interest Rate Risk	
		Policy Risk	
		Macroeconomic Risk	
		Environmental Protection Risk	

3. Model Construction

3.1 Two-Dimensional Evidence Theory

Two-dimensional evidence theory, developed by Professor Zhu Weidong as an extension of classical evidence theory, incorporates a secondary dimension that encompasses characteristic information such as the quality and reliability of evidence sources. This additional framework enables the expression of objective events through a more comprehensive evidential structure [31]. In contrast to classical evidence theory, the two-dimensional approach effectively addresses conflicts between evidence and differentiates the credibility of evidence sources. These enhancements contribute to improved accuracy and reliability in decision-making, thereby increasing its applicability in practical problem-solving scenarios.

Definition 1. Two-dimensional frame of discernment: Beyond the traditional frame of discernment Θ employed in classical evidence theory, an auxiliary dimension Ψ is introduced. This expanded two-dimensional framework enables the formal representation of attributes pertaining to the quality and reliability of evidence sources, or the characteristic information encountered during the evidence-gathering process. The structure of this two-dimensional frame of discernment is expressed as follows:

$$\Gamma:\{\Psi_1,\Psi_2,\cdots\Psi_n\}\rightarrow\{\Theta_1,\Theta_2,\cdots\Theta_m\} \tag{1}$$

In this context, $\Psi=\{\Psi_1,\Psi_2,\cdots\Psi_n\}$ represents two dimensions. Different elements within this frame can represent n features of the evidence sources, where the i-th feature of the evidence source is denoted as Ψ_i. Ψ_i can be formally expressed as $\Psi_i=(Feature\ \mathrm{i};\ Feature\ value\ V_i)$. $\Theta=\{\Theta_1,\Theta_2,\cdots\Theta_m\}$ denotes the evidence decision framework, which is constructed using methods that are consistent with those used for the frame of discernment in classical evidence theory, and the elements within the frame of discernment Θ must satisfy the conditions of being exhaustive and mutually exclusive.

Definition 2. Basic probability assignment (BPA): Assume that Θ is the frame of discernment in evidence theory and that the mass function $m:2^{\theta}\rightarrow[0,1]$ is the BPA of Θ. A represents the focal elements for any $A\subset\Theta$:

$$Bel(A)=\sum_{B\subset A}m(B) \tag{2}$$

Let the function $\mathrm{Bel(A)}$ be the mass function on the frame of discernment Θ. The belief measure assigned to a given proposition should be equivalent to the aggregate of the support degrees attributed to all its subpropositions.

Definition 3. Comprehensive Optimization Coefficient Method: This method uses a correction coefficient α to adjust the BPA of evidence and passes the untreated $1-\alpha$ portion to the unknown part of the frame of discernment Θ. The magnitude of α determines the amount of adjusted BPA [31]. Let Θ be the frame of discernment; the BPA is denoted asm_1, and the focal elements are also$\mathrm{A}_\mathrm{i}.\alpha_\mathrm{i}$ represents the correction coefficients for the evidence. The corrected BPA m_1 can be expressed as:

$$\mathrm{m}_1'(A_\mathrm{i}) = \begin{cases} \alpha_\mathrm{i}\mathrm{m}_1(A_\mathrm{i}) & A_\mathrm{i} \neq \Theta \\ 1-\sum\limits_{A_\mathrm{i}\subset\Theta}\mathrm{m}_1'(A_\mathrm{i}) & A_\mathrm{i} = \Theta \end{cases} \tag{3}$$

Definition 4. Two-Dimensional Evidence Theory Synthesis Rules: Upon modifying the BPA of the first dimension using the quality characteristics of the second-dimensional information source, the BPAs from different evidence sources can be combined through the application of synthesis rules. The specific formula for this synthesis is as follows:

$$\mathrm{m}'_{1,2}(A) = \begin{cases} 0 & A = \varnothing \\ \dfrac{\sum\limits_{A_\mathrm{i}\cap B_\mathrm{j}=A}\mathrm{m}_1'(A_\mathrm{i})\mathrm{m}_2'(B_\mathrm{j})}{1-K} & A \neq \varnothing \end{cases} \tag{4}$$

Where, $K = \sum\limits_{A_\mathrm{i}\cap B_\mathrm{j}\neq\varnothing} m'_1(A_\mathrm{i})m'_2(B_\mathrm{j})$, the value of K is proportional to the degree of conflict between the pieces of evidence. $(1-K)^{-1}$ is a normalization factor that compensates for the confidence of the empty set proportionally to the other nonempty sets to ensure that the sum of the confidence of all the synthesized elements is 1.

3.2 Random Forest Algorithm

The Random Forest (RF) algorithm is an ensemble learning technique that utilizes multiple decision trees as base classifiers, which are combined using the bagging method. The final classification prediction is determined through majority voting among these classifiers. During the training of the various decision tree models, the RF approach incorporates random feature selection. As the number of base learners increases, this process results in a reduction in generalisation error, ultimately improving model performance. From the perspective of the model construction process of the RF, after training iterations, a series of different decision tree models, denoted as $\{h_1(x), h_2(x), \cdots h_i(x)\}$, are obtained. These distinct decision tree models are then aggregated into an ensemble, and the final classification is determined through majority voting. The formula for this process can be expressed as follows:

$$H(x) = \underset{\mathrm{t}\in T}{\mathrm{agrmax}}\sum_{\mathrm{i}=1}^{I} I(\mathrm{h}_\mathrm{i}(x) = \mathrm{t}) \tag{5}$$

Where, $H(x)$ represents the ensemble classification prediction model, $\mathrm{h}_\mathrm{i}(x)$ represents a single decision tree model, t represents the target variable belonging to class t, I represents the number of decision trees, and $I(\ )$ represents the indicator function.

3.3 TD-DS-RF Model

In our previous study [3], we introduced the DS-RF model, using RF to obtain BPA for evidence theory, enhancing financial risk early-warning. Building on this, the current study integrates two-dimensional evidence theory and text mining to develop the TD-DS-RF model. This model addresses the DS-RF's limitation of solely relying on quantitative data by incorporating the quality of information sources, offering a more comprehensive and objective risk warning framework. The model flow chart is shown in Figure 2.

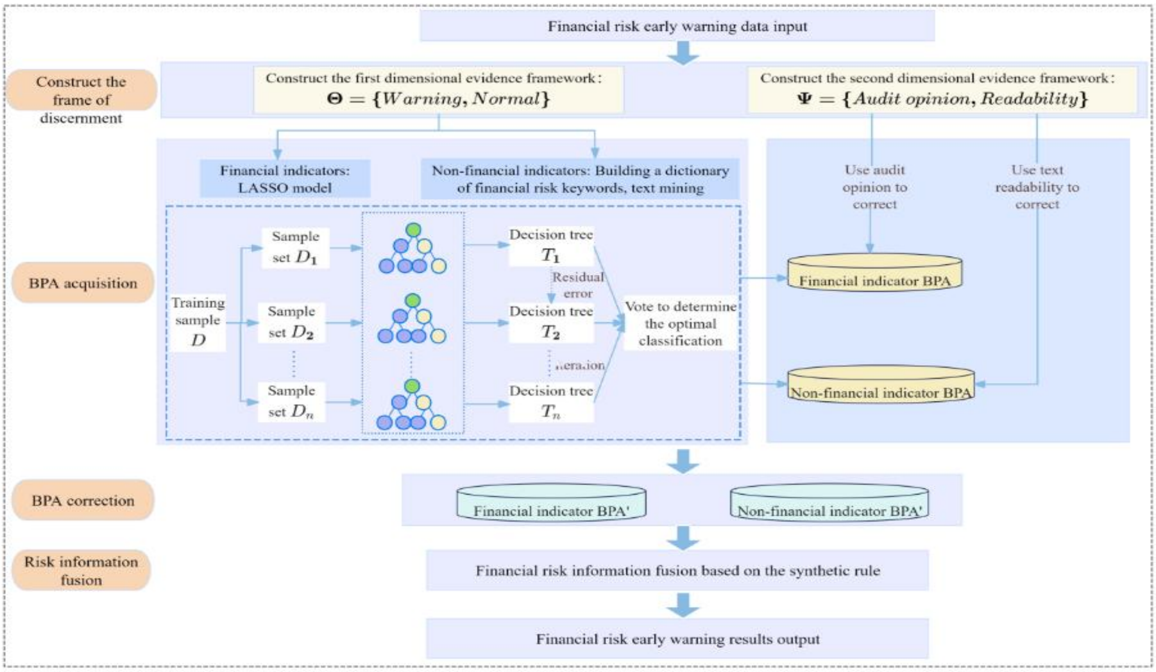

Fig. 2 Flowchart of the Financial Risk Early-Warning Model of Listed Companies Based on the TD-DS-RF Model

The specific steps of the TD-DS-RF model are as follows:

Step 1: Input financial data and non-financial data related to financial risk early-warning.

Step 2: A two-dimensional frame of discernment is established, with firms marked 'ST' or '*ST' serving as indicators of financial risk early-warning, following common domestic research practice. Therefore, the first-dimensional evidence framework constructed is represented as $\Theta = \{\theta_1, \theta_2\} = \{\text{Warning, Normal}\}$. For financial indicators and non-financial indicators, audit opinion' and 'text readability' are used, respectively, to reflect the reliability of first one-dimensional information. Therefore, the second-dimensional evidence framework is expressed as $\Psi = \{\Psi_1, \Psi_2\} = \{\text{Audit opinion, Readability}\}$.

Step 3: The financial risk early-warning index system is constructed by utilising LASSO regression for the attribute reduction of financial indicators, thereby selecting the most pertinent variables while eliminating those with minimal influence. Additionally, a text mining algorithm is employed to extract key terms and risk-related information from non-financial indicators, enhancing the comprehensiveness of the early-warning framework.

Step 4: By integrating the RF algorithm, the financial risk outcomes are predicted using both financial and non-financial indicators. This process yields the BPA for both financial and non-financial indicators.

Step 5: Use the second two-dimensional evidence framework $\Psi = \{\text{Audit opinion, Readability}\}$ to adjust the BPA derived in Step 4, and 'Financial indicators BPA' and 'non-financial indicators BPA'' are obtained to consider the impact of differences in the quality of evidence sources on risk warning.

Step 6: The modified BPA is synthesised using the two-dimensional evidence theory fusion rule (Formula (4)) to integrate the risk information effectively.

Step 7: The financial risk early-warning results are generated based on the synthesised evidence.

4. Empirical Analysis

4.1 Data

This study uses listed manufacturing companies in China from 2012 to 2021 as research samples. Companies marked with 'ST' or '*ST', indicating financial risk, make up about 20% of the sample. To avoid sample imbalance and overfitting, the propensity score matching (PSM) method is employed to match non-risk companies with similar characteristics, such as industry and asset scale. A total of 392 companies marked as 'ST' or '*ST' (warning) and 784 matched non-risk companies (normal) are selected, resulting in 1176 total companies. Data is sourced from the CNRDS, CSMAR databases, and Jushao Information Network.

4.2 BPA Acquisition Via the RF Algorithm

In this study, the Ensemble module from Python's Scikit-learn machine learning library was utilised in combination with the RF algorithm to predict financial risk early-warning results and calculate the BPA based on both financial and non-financial indicators. The methodology involves establishing two RF models, with financial and non-financial indicators as input data. If the financial risk warning result for a listed company is normal, the label is '0'; if it is flagged with 'ST' or '*ST', the label is '1'. Through cross-validation, the training-to-test set ratio for the RF models constructed is 8:2, with the training set consisting of 940 listed companies and the test set consisting of 236. To train the optimal model, the GridSearchCV algorithm is used to optimize the parameters 'n_estimators', 'min_samples_leaf' and 'max_depth' of the RF and set the range and step size changes of the three parameters. When the model iteration is complete, the optimal parameters are output. The optimal parameters of the two RF models in this study are shown in Table 7.

Table 7
Optimal Parameter Combinations of the RF Models

Random Forest Model	n_estimators	min_sample_leaf	max_depth
Based on Financial Indicators	100	5	15
Based on Non-Financial Indicators	400	15	15

After model training is complete, the probabilities of the '0' and '1' prediction outcomes for the test set data are obtained using the 'predict_proba()' function, which is regarded as the BPA of early warning. In this study, a test set comprising 236 samples was utilised. To illustrate the steps of the model, five listed companies were selected as examples. The BPA outputs and risk warnings for these five companies, based on both financial and non-financial indicators, are presented in Table 8.

Table 8
Financial Risk Early-Warning Results and the BPA Based on the Random Forest Model

Listed Company	Financial Indicators Early-Warning BPA			Non-Financial Indicators Early-Warning BPA		
	Normal	Warning	Results	Normal	Warning	Results
A	0.969	0.031	Normal	0.834	0.166	Normal
B	0.126	0.874	Warning	0.342	0.658	Warning
C	0.840	0.160	Normal	0.386	0.614	Warning
D	0.176	0.824	Warning	0.706	0.294	Normal
E	0.442	0.558	Warning	0.659	0.341	Normal

As shown in Table 8, for Company D, the early-warning results based on financial and non-financial indicators are inconsistent. The financial indicators predict a 'warning' with a BPA of 0.824, while the non-financial indicators predict 'normal' with a BPA of 0.706. In such cases, decision-makers cannot directly determine the final prediction. Additionally, variations in the quality of information in annual reports can affect the reliability of evidence sources and influence the results. Therefore, after obtaining the BPA, it is corrected using the second dimension of evidence, reflecting the quality of information sources.

4.3 Evidence Correction Based on Two-Dimensional Evidence Theory

In the second-dimensional evidence framework constructed in this study, the correction coefficients corresponding to audit opinions and text readability are α_1 and α_2, respectively. The higher the value of the correction coefficient, the more reliable the information from the evidence source, indicating better data quality. Regarding audit opinions, the correction coefficient α_1 for the five audit opinions is set as follows: a standard unqualified opinion, indicating the highest reliability of the financial report, is assigned a coefficient of 1; an unqualified opinion with an emphasis on a matter paragraph, confirming the report's reliability but noting some uncertainty, receives a coefficient of 0.8; a qualified opinion, signifying errors in the report, is assigned a coefficient of 0.6; an adverse opinion, indicating serious reliability issues, receives a coefficient of 0.4; and an inability to comment, where the auditor cannot assess the report's reliability, is assigned a coefficient of 0.5. For text readability, the ratio of professional words to total words in the text is used as an index, with professional terms taken from the 'Sogou Finance and Economics Vocabulary Collection'. This ratio is calculated using Python, and the higher the use of specialized terms, the less readable the text is. Therefore, this study reversibly normalized the acquired readability index and mapped it to the interval [0.5, 1] to calculate the correction coefficient α_2 of the non-financial index. The specific formula is as follows:

$$\alpha_2 = 0.5 + \frac{0.5}{X_{\max} - X_{\min}} \times \left(X_{\max} - X\right) \tag{6}$$

Where, $X_{\max}$ is the maximum value in this set of data and where $X_{\min}$ is the minimum value in this set of data. Thus, the correction coefficients for the reliability and quality features of evidence sources, based on both financial and non-financial indicators, are presented in Table 9.

Table 9

Correction Coefficients of the Financial Index and Non-Financial Index

Listed Company	Financial Indicators		Non-Financial Indicators	
	Audit Opinions	Correction Coefficient α_1	Text Readability	Correction Coefficient α_2
A	Standard Unqualified Opinion	1.00	0.86	0.61
B	Standard Unqualified Opinion	1.00	0.78	0.68
C	Qualified Opinion	0.60	0.86	0.61
D	Unqualified Opinion with An Emphasis of Matter Paragraph	0.80	0.84	0.63
E	Qualified Opinion	0.60	0.85	0.62

After obtaining the correction coefficient, it must be applied to revise the evidence. For non-financial indicators, this paper adopts the comprehensive optimization coefficient method (Formula 3) for evidence correction. Regarding financial indicators, the selected audit opinion reflects the reliability of the annual report. If the comprehensive optimization coefficient method were directly applied, both the 'warning' and 'normal' BPA for the financial risk of listed companies would be reduced. However, in the case of four audit opinion categories (except for the "Standard unqualified opinion"), the financial risk early-warning of listed companies would increase. Therefore, this study allocated the $1 - \alpha_1$ part to the revised ‘warning’ BPA to obtain the revised BPA' and set the correction coefficient as α_1. The specific correction formula is as follows:

$$\mathrm{m}_1{}'(A_i)=\begin{cases}\alpha_1\times \mathrm{m}_1(A_i)+(1\text{-}\alpha_1) & A_i\neq\Theta,\ \mathrm{i}=1\\ \alpha_1\times \mathrm{m}_1(A_i) & A_i\neq\Theta,\ \mathrm{i}=2\end{cases} \tag{7}$$

Where i=1 indicates ‘Warning’, and i=2 indicates ‘Normal’. After calculation by the above formula, the revised BPA' can be obtained, with the results shown in Table 10.

Table 10
Results of the Revision of Evidence

Listed Company	Financial Indicators Early-Warning BPA		Non-Financial Indicators Early-Warning BPA	
	Normal	Warning	Normal	Warning
A	0.969	0.031	0.511	0.102
B	0.126	0.874	0.233	0.448
C	0.504	0.496	0.235	0.375
D	0.141	0.859	0.445	0.185
E	0.265	0.735	0.408	0.211

4.4 Financial Risk Early-Warning Information Fusion Based on Two-Dimensional Evidence Theory

The corrected BPA was synthesized using the synthesis rule of two-dimensional evidence theory. After synthesis, it was converted into the final decision result using the probability transformation rule. The formula for probability transformation is as follows:

$$\mathrm{m}^*(A_i)=\mathrm{m}^*(A_i)\times\left(\frac{\mathrm{m}(\Theta)}{\sum \mathrm{m}(A_i)}+1\right) \tag{8}$$

In the final decision stage, based on the Max (bel) criterion, the proposition in the first one-dimensional recognition framework corresponding to the maximum probability after conversion is selected as the final risk warning result, as shown in Table 11. A comparison between Table 8 and Table 11 shows that the financial and non-financial indicator-based risk early-warning results and BPA are generated using the RF method. When the prediction outcomes diverge, users of the information cannot immediately draw conclusions. To address this, the model developed in this study incorporates the quality characteristics of the information sources.

Table 11
Financial Risk Early-Warning Fusion Results and Probability Transformation Results

Listed Company	Fusion Result		Probability Conversion Result		Early-Warning Result	Actual Label
	Normal	Warning	Normal	Warning		
A	0.559	0.004	0.994	0.006	Normal	0
B	0.040	0.529	0.070	0.930	Warning	1
C	0.171	0.268	0.389	0.611	Warning	1
D	0.106	0.269	0.283	0.717	Warning	1
E	0.168	0.241	0.411	0.589	Warning	1

The Dempster-Shafer (D-S) synthesis rule is employed to merge multi-source information, effectively integrating early-warning signals. The results indicate that the prediction outcomes from the proposed early-warning model align entirely with the actual early-warning situations of listed companies. To more accurately assess the performance of the proposed early-warning model, a statistical analysis was conducted on the prediction results for 236 listed companies in the test set, alongside their corresponding actual labels. Four performance indicators—precision, accuracy, F1-score, and recall—were chosen to evaluate the model. The confusion matrix for the model is shown in Table 12. The precision, accuracy, F1-score, and recall of the proposed two-dimensional evidence theory-based financial risk early-warning model for listed companies were 0.81, 0.84, 0.75, and 0.71, respectively.

Table 12
Confusion Matrix of Financial Risk Early-Warning of Listed Companies

Actual Result	Forecast Result	
	Warning	Warning
Warning	TP=58	FN=24
Normal	FP=14	TN=140

4.5 Comparative Analysis of Model Performance

To validate the model's effectiveness, four comparative models were tested. Logistic regression (Model 1), a classical early-warning method, was used to benchmark against traditional models. To assess the role of non-financial indicators, RF with only financial data (Model 2) and RF with both financial and non-financial data (Model 3) were compared. To test the contribution of two-dimensional evidence theory, Model 4 employed classical evidence theory, while Model 5 used the proposed TD-DS-RF model. The performance of all five models is detailed in Table 13.

Table 13
Model Performance Comparison of the Five Models

Model	Acc	Pre	F1-Score	Recall	Concrete Model
Model 1	0.70	0.65	0.67	0.70	Logistic Regression
Model 2	0.73	0.72	0.70	0.68	Random Forest (Financial Indicators)
Model 3	0.78	0.75	0.72	0.69	Random Forest (Financial and Non-Financial Indicator)
Model 4	0.83	0.81	0.73	0.67	Random Forest & Classical Evidence Theory
Model 5	0.84	0.81	0.75	0.71	Random Forest & Two-Dimensional Evidence Theory（TD-DS-RF）

A comparative analysis of Pre, Acc, Recall, and F1-score shows that the TD-DS-RF model outperforms the other four models. The findings are as follows: (1) The TD-DS-RF model enhances predictive accuracy over traditional models and assigns probability estimates to outcomes, allowing for more precise risk assessment. (2) Incorporating non-financial indicators significantly boosts model accuracy compared to using financial data alone. (3) The integration of evidence theory outperforms the use of RF alone, confirming its strength in managing uncertain data. (4) The two-dimensional evidence theory, by accounting for the quality and reliability of evidence sources, yields superior results to the classical framework in the same decision context. These results confirm the model's effectiveness and practical value.

5. Conclusion

This study constructs a financial risk early-warning model for listed firms using TD-DS-RF, combining financial and non-financial indicators within a one-dimensional evidence framework. Non-financial indicators are extracted through text mining from the "possible risks" section of MD&A disclosures. A tailored keyword dictionary is developed, covering five risk categories to assess risk type and severity. The model enhances classical evidence theory by introducing a second dimension—'audit opinion' and 'text readability'—to correct evidence sources. The D-S rule then integrates predictions from both dimensions. Using data from 392 risk-listed and 784 normal manufacturing firms (2012–2021), the model improves information reliability and enhances early-warning performance. By integrating AI techniques, TD-DS-RF boosts source quality and supports stakeholder decisions while aiding regulators in real-time monitoring, contributing to capital market stability. In terms of adaptability, TD-DS-RF shows promise across sectors and regions. For banks, merging financial and qualitative data can refine credit risk forecasts. In insurance, sector-specific risks can improve strategy design. Globally, parameter adjustments can tailor the model to regulatory, economic, and market conditions, enabling cross-border risk prediction and investor support. Nonetheless, the synthesis method for two-dimensional evidence needs refinement. This study uses sequential fusion; future work should explore simultaneous integration of both dimensions to minimise information loss and improve results. Optimising this synthesis algorithm is a key direction for future research.

Author Contributions

Conceptualization, Q.H.; methodology, T.Z.; software, T.Z.; validation, Q.H.; formal analysis, T.Z.; investigation, Q.C. and T.Z.; resources, W.Z.; data curation, T.Z. and Q.C.; writing-original draft preparation, T.Z.; writing-review and editing, T.Z.; visualization, Q.H.; supervision, W.Z.; project administration, Q.C.; funding acquisition, W.Z. All the authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the National Natural Science Foundation of China [grant numbers 72471074 and 71774047].

Data Availability Statement

The datasets generated or analysed during this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Sarlin, P. (2014). On biologically inspired predictions of the global financial crisis. *Neural Computing and Applications*, *24*, 663-673. https://doi.org/10.1007/s00521-012-1281-y

[2] Liu, W., Fan, H., Xia, M., & Pang, C. (2022). Predicting and interpreting financial distress using a weighted boosted tree-based tree. *Engineering Applications of Artificial Intelligence*, *116*, 105466. https://doi.org/10.1016/j.engappai.2022.105466

[3] Zhu, W., Zhang, T., Wu, Y., Li, S., & Li, Z. (2022). Research on optimization of an enterprise financial risk early warning method based on the DS-RF model. *International review of financial analysis*, *81*, 102140. https://doi.org/10.1016/j.irfa.2022.102140

[4] Huang, Y.-P., & Yen, M.-F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing*, *83*, 105663. https://doi.org/10.1016/j.asoc.2019.105663

[5] Mousavi, M. M., & Lin, J. (2020). The application of PROMETHEE multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. *Expert Systems with Applications*, *159*, 113438. https://doi.org/10.1016/j.eswa.2020.113438

[6] Li, Z., Crook, J., Andreeva, G., & Tang, Y. (2021). Predicting the risk of financial distress using corporate governance measures. *Pacific-Basin Finance Journal*, *68*, 101334. https://doi.org/10.1016/j.pacfin.2020.101334

[7] Muñoz-Izquierdo, N., Camacho-Miñano, M.-d.-M., Segovia-Vargas, M.-J., & Pascual-Ezama, D. (2019). Is the external audit report useful for bankruptcy prediction? Evidence using artificial

intelligence. *International Journal of Financial Studies*, *7*(2), 20. https://doi.org/10.3390/ijfs7020020
[8] Donovan, J., Jennings, J., Koharki, K., & Lee, J. (2021). Measuring credit risk using qualitative disclosure. *Review of Accounting Studies*, *26*, 815-863. https://doi.org/10.1007/s11142-020-09575-4
[9] Mselmi, N., Lahiani, A., & Hamza, T. (2017). Financial distress prediction: The case of French small and medium-sized firms. *International review of financial analysis*, *50*, 67-80. https://doi.org/10.1016/j.irfa.2017.02.004
[10] Li, X., Wang, J., & Yang, C. (2023). Risk prediction in financial management of listed companies based on optimized BP neural network under digital economy. *Neural Computing and Applications*, *35*(3), 2045-2058. https://doi.org/10.1007/s00521-022-07377-0
[11] Li, C., Jin, K., Zhong, Z., Zhou, P., & Tang, K. (2021). Financial risk early warning model of listed companies under rough set theory using BPNN. *Journal of Global Information Management (JGIM)*, *30*(7), 1-18. https://doi.org/10.4018/JGIM.300742
[12] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), 589-609. https://doi.org/10.2307/2978933
[13] Chen, P. (2023). RETRACTED ARTICLE: Research on Financial Risk Evaluation and Control of Tourism Enterprises Based on Improved GA Algorithm. *International Journal of Computational Intelligence Systems*, *16*(1), 146. https://doi.org/10.1007/s44196-023-00317-3
[14] Lin, G., & Chen, W. (2022). Early warning method for enterprise financial informatisation caused by tax difference. *International Journal of Information Technology and Management*, *21*(2-3), 248-263. https://doi.org/10.1504/IJITM.2022.123404
[15] Manzaneque, M., & Priego, A. M. (2016). Corporate governance effect on financial distress likelihood: Evidence from Spain: Efecto del gobierno corporativo en la probabilidad de fracaso empresarial: evidencia española. *Revista de Contabilidad-Spanish Accounting Review*, *19*(1), 111-121. https://doi.org/10.1016/j.rcsar.2015.04.001
[16] Lee, A., & Chae, S.-J. (2018). The effect of management disclosure and analysis on the stock crash risk: Evidence from Korea. *The Journal of Asian Finance, Economics and Business*, *5*(4), 67-72. https://doi.org/10.13106/jafeb.2018.vol5.no4.67
[17] Park, H., & Lee, A. (2016). MD&A disclosure quality and the Accuracy of management forecasts. *Review of Accounting & Policy Studies*, *21*(21), 185-210. https://doi.org/10.7469/JKSQM.2022.50.4.735
[18] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The journal of finance*, *66*(1), 35-65. https://doi.org/10.1111/j.1540-6261.2010.01625.x
[19] Brown, S. V., & Tucker, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, *49*(2), 309-346. https://doi.org/10.1111/j.1475-679X.2010.00396.x
[20] Wang, G., Chen, G., & Chu, Y. (2018). A new random subspace method incorporating sentiment and textual information for financial distress prediction. *Electronic Commerce Research and Applications*, *29*, 30-49. https://doi.org/10.1016/j.elerap.2018.03.004
[21] Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Making words work: Using financial text as a predictor of financial events. *Decision support systems*, *50*(1), 164-175. https://doi.org/10.1016/j.dss.2010.07.012
[22] Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and accrual-based earnings management in the pre-and post-Sarbanes-Oxley periods. *The accounting review*, *83*(3), 757-787. https://doi.org/10.2308/accr.2008.83.3.757

[23] Tang, X., Li, S., Tan, M., & Shi, W. (2020). Incorporating textual and management factors into financial distress prediction: A comparative study of machine learning methods. *Journal of Forecasting*, *39*(5), 769-787. https://doi.org/10.1002/for.2661

[24] Fitzpatrick, P. J. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies*. https://openlibrary.org/books/OL6298050M

[25] Sun, X., & Lei, Y. (2021). Research on financial early warning of mining listed companies based on BP neural network model. *Resources Policy*, *73*, 102223. https://doi.org/10.1016/j.resourpol.2021.102223

[26] Qian, H., Wang, B., Yuan, M., Gao, S., & Song, Y. (2022). Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. *Expert Systems with Applications*, *190*, 116202. https://doi.org/10.1016/j.eswa.2021.116202

[27] Zhao, Z., Li, D., & Dai, W. (2023). Machine-learning-enabled intelligence computing for crisis management in small and medium-sized enterprises (SMEs). *Technological Forecasting and Social Change*, *191*, 122492. https://doi.org/10.1016/j.techfore.2023.122492

[28] Shafer, G. (1976). A Mathematical Theory of Evidence. https://philpapers.org/rec/SHAAMT

[29] Tong, Y., Zhao, R., Ye, W., & Li, D. (2016). Research on energy efficiency evaluation for overhead crane. *Kybernetes*, *45*(5), 788-797. https://doi.org/10.1108/K-09-2015-0225

[30] Zhang, T., Zhu, W., Wu, Y., Wu, Z., Zhang, C., & Hu, X. (2023). An explainable financial risk early warning model based on the DS-XGBoost model. *Finance Research Letters*, *56*, 104045. https://doi.org/10.1016/j.frl.2023.104045

[31] Zhu, W., Ku, Q., Wu, Y., Zhang, H., Sun, Y., & Zhang, C. (2018). A research into the evidence reasoning theory of two-dimensional framework and its application. *Kybernetes*, *47*(5), 873-887. https://doi.org/10.1108/K-05-2017-0156

[32] Zhu, W., Li, S., Zhang, H., Zhang, T., & Li, Z. (2022). Evaluation of scientific research projects on the basis of evidential reasoning approach under the perspective of expert reliability. *Scientometrics*, 1-24. https://doi.org/10.1007/s11192-021-04201-9

[33] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. https://doi.org/10.48550/arXiv.1301.3781