Journal of Information Systems Engineering and Management
2024, 9(2), 24534
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Construction and Optimization of Financial Risk Management Model Based on Financial Data and Text Data Influencing Information System

Hui Huang [1], Thien Sang Lim [2*]

[1] Doctor, Faculty of Business, Economics and Accountancy, Universiti Malaysia Sabah, Sabah, Malaysia
[2] Senior Lecturer, Faculty of Business, Economics and Accountancy, Universiti Malaysia Sabah, Sabah, Malaysia
* **Corresponding Author:** tslim@ums.edu.my

Citation: Huang, H., & Lim, T. S. (2024). Construction and Optimization of Financial Risk Management Model Based on Financial Data and Text Data Influencing Information System. *Journal of Information Systems Engineering and Management*, 9(2), 24534. https://doi.org/10.55267/iadt.07.14767

ARTICLE INFO

Received: 30 Oct 2023
Accepted: 31 Jan 2024

ABSTRACT

A-share companies must manage financial risk to succeed. Textual data insights can greatly impact risk assessment results, although most risk management systems focus on quantitative financial assessments. This research constructs and enhances information system financial risk management models employing financial and textual data, including MD&A narratives, to fill this gap. We study how textual data aids financial risk management algorithms' risk prediction. Textual and financial research on 2001–2022 Shenzhen and Shanghai Stock Exchange companies is used. This study found financial and non-financial data models more predictive. Qualitative textual information is used in financial risk assessment to improve risk prediction algorithms. MD&A texts, sentiment analysis, and readability signal risk. Internet forum discussions are linked to financial risk, but media coverage is not. These unconventional data sources evaluate financial risk. The research shows that A-share corporations manage financial risk. The study advises merging qualitative textual data with financial metrics to solve literature gaps and improve risk management. Shenzhen and Shanghai Stock Exchange statistics suggest MD&A storylines might strengthen financial risk management models. Study shows readability and sentiment analysis increase risk model prediction. The study found that textual material affects financial risk, therefore risk assessment should include non-financial information. This complete risk management technique may assist A-share listed companies navigate financial markets and make smarter decisions using quantitative financial data and qualitative textual insights. This study implies textual data may help financial risk algorithms. MD&As help companies identify and manage financial risk. More study is needed to discover new textual elements and strengthen context-specific risk management frameworks.

Keywords: Financial Risk Management, Management Discussion and Analysis, Textual Data, Financial Data, Information System.

INTRODUCTION

Rapid technological progress and changing legislation cause business volatility. To solve difficult problems during economic transformation, companies must be nimble, inventive, and adaptable. In light of economic globalization, the need for efficient financial risk early warning systems has become essential. Financial hazards have been shown to have a key role in the demise and bankruptcy of numerous businesses. It is crucial to remember that the creation of financial risks often stems from a gradual accumulation of unsustainable activities over the long term, rather than arising suddenly or randomly (Mazzucato, 2013). Businesses face the risk of incurring higher losses and even bankrupt if they fail to promptly recognize and address their financial risks. Companies with weak early warning systems for financial risk are more susceptible to unfavorable occurrences like the global financial crisis and are more likely to have financial crises or possibly go bankrupt. However, businesses with effective risk early warning systems are better able to weather economic crises. Investors use the financial risk status of an enterprise as a key indication to evaluate managerial performance because it reflects managerial skill (Bonsall, Holzman, & Miller, 2017). However, if the shareholders invest in a firm that has debt, there is a chance that they could lose money, especially if the company's cash flow is insufficient to pay off its debt. This is known as financial risk. In times of bankruptcy, creditors frequently get paid before stockholders. The prospect of a business skipping a payment on its debt is frequently included in discussions of financial risk. Prediction models can be incorporated and used to reduce such risks and evaluate a company's overall financial health. By examining particular financial indicators or other pertinent traits of the firm or its operating environment, these models work as timely warning systems that seek to assess a company's financial health (Yi, 2023).

The study of enterprise financial risk has long been a popular subject in the fields of finance and management (Zhao et al., 2022). All organisations, whether they are established Fortune 500 companies, young start-ups, or small and medium-sized enterprises (SMEs), are intrinsically vulnerable to various financial risks, such as credit risk, guarantee risk, supply chain risk, bankruptcy risk, and more. Enterprises encounter several difficulties and hazards, particularly during periods of economic crisis or unanticipated catastrophes like the COVID-19 pandemic in 2020 and the financial crisis of 2008. Enterprises of today need to confront internal risks as well as the possibility of contagion hazards in the integrated global financial system of today (Drobyazko, Barwinska-Malajowicz, Slusarczyk, Chubukova, & Bielialov, 2020).

Numerous factors that directly affect the interests of financial investors affect the financial health of these businesses. Significant financial risks can arise throughout an enterprise's operation from both external factors like policy and market pressures as well as internal ones like investment and financing decisions. Any problems in these areas can pose serious hazards to businesses, resulting in asset losses, large market value losses, temporary drops in stock prices, and even long-term company losses. In such circumstances, businesses can experience financial difficulties, receive special treatment, or even be delisted (Wei, 2022). Therefore, it is critical to build a system that offers early alerts for the financial dangers to assist investors and financial institutions in spotting potential crises in listed businesses, as early as feasible and reducing important investment risks. Additionally, it can give listed firms early warnings, assisting them in identifying underlying problems with investment choices and essentially preventing financial dangers. Research into financial risk early warning systems is quite important (Wang, Liu, & Luo, 2020). A key component of national economic security is financial security, which emphasizes the need to identify and reduce potential financial risks while maintaining a diligent risk-avoidance strategy. Implementing cost-effective financial risk prediction models is necessary for achieving this. It is even more important for a fast developing country like China, where its financial risk prevention system faces considerable hurdles in the Internet era due to the exponential growth of the economy and financial data, and the increasing need to amalgamate various risk information for accurate and reliable risk assessment. Processing such large amounts of data using conventional methods takes time and money. Therefore, researchers in financial security have created cutting-edge financial risk avoidance models utilizing machine learning (ML) and deep learning (DL) (Wei, 2022; X. T. Li & Duan, 2022).

In the current era of rapid socioeconomic development, which is fueled by developments in information technology (IT), such as big data, cloud computing, and the Internet of Things (IoT), there is a significant transition taking place in people's lives, jobs, and society towards informatization and intellectualization. People and businesses have gotten closer because of the IT revolution and the rise of e-commerce, which have overcome obstacles like distance and culture. Studying these elements is practically important because e-commerce businesses strongly rely on their financial status and management to determine their growth potential (K. Peng & Yan, 2021; Liu & Huang, 2022; Zheng, Kouadio, & Kombate, 2021; Duan, Du, & Guo, 2022; J. Li, 2022). The degree of enterprise Financial Risk Management (FRM) has been raised as a result of substantial research and the development of numerous theoretical techniques by academics. The potential of suffering financial losses as a result of its operations as outlined earlier is referred to as an entity's financial risk. There are various reasons for financial risks in organizations and necessitate careful analysis, if the true sources or causes are to be extracted or derived. To efficiently reduce financial risks, academics have suggested various strategies and treatments including taking the economy into account, risk minimization, risk transfer, risk tolerance, risk treatment, and the integration of financial risk management with mathematical models (X. T. Li & Duan, 2022; Y. W. Li & Cao, 2020; Okello, 2012; J. Li, 2022).

Previous research has concentrated more on the quantitative examination of financial data while ignoring the important conclusions that can be drawn from textual data. As a result, a financial risk management model that takes into account both financial data analysis and textual analysis of pertinent information systems must be built and optimized (L. Zheng, Gao, Feng, & Wang, 2023). Few studies analyse corporate financial dangers using textual and financial data. This gap hampers sentiment and NLP financial data analysis. Academics can analyse news, social media, and company reports using these methodologies (Ying, Chen, & Zhao., 2021). Recent research shows that this integrated strategy improves financial dynamics understanding and company risk assessment (Ying et al., 2021; Al-Haschimi, Apostolou, Azqueta-Gavaldon, & Ricci, 2023; Chi, Yan, Pang, & Lei., 2022; Zheng et al., 2023).

This research examines the intricate interaction between textual and financial information in information systems to improve a financial risk management model. These efforts assess firm financial risks using qualitative text data and quantitative financial data. This study addresses a literature gap by analysing financial risk assessment methodologies in A-share listed companies using textual data and information technology to show the significance of qualitative textual information alongside conventional financial indicators in risk management. This study found that market volatility, credit risks, operational risks, and liquidity hazards cost companies. According to existing studies, risk management strategies emphasise quantitative financial data above qualitative textual data. It explores how present methods overlook financial risk's complexity. Financial and textual data, mostly MD&A, are used to build and refine a financial risk management model. To determine how information systems affect risk assessment, 2001–2022 Shenzhen and Shanghai Stock Exchange data is analysed quantitatively and qualitatively. Financial risk management models with textual data are improved by regression and sentiment analysis.

LITERATURE REVIEW

The literature on financial crises extensively acknowledges that abrupt events can lead to systemic financial risk, causing widespread and far-reaching effects. This risk is not limited to the banking industry; it extends to other spheres such as stock markets, debt markets, and currency rates. For instance, Wei (2022) developed an indicator system encompassing factors from the banking industry, foreign currency market, bond market, and stock market, among others. This stress index allows for monitoring financial risk and stress levels (Valaskova, Kliestik, & Kovacova, 2018). Ying et al. (2021) conducted research that emphasized the characteristics of abrupt shocks, which can trigger systemic financial risk with a snowball effect, potentially resulting in a worldwide financial crisis. Such occurrences have significant implications for the macroeconomic system, impacting financial institutions and commercial banks. Regarding the complexity of systemic risk, multiple factors contribute. These factors include the presence of various financial institutions, individual investors, a wide range of financial assets with substantial balance sheets, and intricate creditor-debtor relationships.

The frequent use of operational financial data as input indices in FRM models by researchers has demonstrated how broadly and universally applicable business financial data is. The suggested corporate financial risk model uses retained profits to total assets, current debt to assets, asset management ratio, and earnings management indexes to identify and manage financial risks (Y. Peng, Zhang, Tang, & Li, 2011; Abbasi, Sarker, & Chiang, 2016; Y. W. Li & Cao, 2020; Zheng et al., 2023; An & Chen, 2022; Duan et al., 2022). Bathory (1984) developed a technique to assess the financial risks of businesses in 1984. Budgets and corporate risk are assessed. Our retained profits to total assets ratio shows how much of our assets are sustained by retained earnings, suggesting we can grow without external funding, which may increase financial risk. The Current debt to assets ratio indicates firm liquidity, debt management, and short-term financing. Due to interest rate and debt payment fluctuations, high current debt to assets ratios may indicate financial risk (R. Wang, Yu, & J. Wang, 2019; Le Hoang, Phan, & Do, 2022).

Company asset use revenue and profit are measured by asset management ratios. Higher ratios suggest better asset management and lesser financial risks. earnings management indexes affect investor risk and stability. Risk assessment model financial indices help companies manage financial risk for sustainability and competitiveness (Curti, Gerlach, Kazinnik, Lee, & Mihov, 2023).

In current commercial financial risk assessment, indicators, processes, and models reflect risk dynamics, say academics and practitioners (Gajadien et al., 2023). We address market volatility, credit, operational, strategic, and regulatory risks comprehensively. Thus, researchers have developed new methods to quantify and decrease these dangers. Beta and standard deviation measure asset class price volatility and market sensitivity. These indicators help investors and financial analysts identify investment hazards and diversify portfolios to reduce systematic risk. Investors and risk managers use MTB and CDS spreads to assess market risk. Financial risk and dynamic economic organisations are explored in different ways. Finance indices, techniques, and models assist companies weather rough times by identifying, monitoring, and managing financial risk (Gilchrist, Wei, Yue, & Zakrajšek, 2022).

VaR evaluates asset or portfolio value losses over time and is a key financial risk metric. VaR estimates negative outcomes within a confidence interval to help investors and financial institutions manage risk. VaR must stress-test financial systems and investments in tough times (Landi, Iandolo, Renzi, & Rey, 2022). SVMs, neural networks, and random forests predict financial risk. Machine learning algorithms use huge databases of textual data from financial reports, news stories, and other unstructured sources to improve risk assessment accuracy and granularity. Researchers use AI and data analytics to find hidden textual patterns, trends, and links to improve financial risk models.

IT-based textual data analysis revolutionises financial risk assessment. NLP changes how scholars analyse earnings call transcripts and social media posts. NLP algorithms structure text for quantitative analysis and risk modelling in information systems. This unique approach employs machine learning and sentiment analysis to gain textual insights. These tools let financial analysts and decision-makers find sentiment trends, critical subjects, and relevant information in massive text libraries to make quick, informed decisions. For financial risk assessment, academics can use IT and textual data processing to study market dynamics and investor sentiment. Quantifying qualitative and textual data to fundamental financial indicators helps analysts recognise market risk (Laihonen & Kokko, 2019). Researchers create methods to assess financial risk faster and more accurately. Risk assessment methods alter with financial market volatility, data, and technology. Finally, information systems, technology, and data analytics have created a new financial risk assessment era where textual data aids prediction and strategic decision-making. NLP, sentiment analysis, and machine learning researchers value risk assessment accuracy and timeliness. IT helps banks and investors reduce risk and boost returns in volatile economies (Y. Li & Zhang, 2023).

The literature claims technology, especially textual data processing, has revolutionised financial risk assessment. Online forums, media, and annual reports Financial risk can be predicted by MD&A attitude. Increases prediction model risk assessment granularity and accuracy. Negative management remarks or online forums increase financial risk, studies reveal. IT automates data collection, purification, and analysis, making textual data forecastable. This automation accelerates financial risk assessment and provides real-time textual source monitoring to assist stakeholders address new risks and opportunities (Utomo, 2022).

New text processors changed financial risk and crisis prediction. Researchers found textual data and IT technology can measure and evaluate complex information to improve financial risk assessments' accuracy and efficiency. Information technology can assist financial institutions and investors make smarter decisions in a changing financial world by highlighting market dynamics, sentiment patterns, and threats (Curti et al., 2023). This literature study compares textual and IT-based financial risk assessment methods to illustrate their pros and downsides. Market volatility, credit ratings, and liquidity ratios are quantitative risk management indicators. Asset performance and market dynamics are common financial risk assessment methods. IT, qualitative textual data, and quantitative methods are used more for risk assessment. The study will improve quantitative risk assessment and exposure visibility with sentiment analysis, NLP, and machine learning (Gilchrist et al., 2022; Landi et al., 2022).

This literature review investigates quantitative, textual, and IT-based methodologies to explain financial risk assessment's evolution and determine the best ways to reduce risk and maximise profits in dynamic financial markets. The study advises practitioners and policymakers on financial risk management in a globalised and volatile economy using literature and empirical data. The literature review demonstrates that qualitative textual data insights are often ignored while assessing and reducing financial risks. Rich contextual information in firm reports, news stories, and social media conversations may reveal concerns. The literature supports the proposed study's goal of revealing this gap by combining qualitative textual data with standard financial indicators to improve risk assessment models. The literature review covers financial risk assessment's current information system and technology breakthroughs. Text-based machine learning, sentiment analysis, and natural language processing research are our focus. Though quantitative data cannot disclose underlying risks, nuanced patterns, attitudes, and language signals can revolutionise risk prediction algorithms. New technologies are appealing, but the literature review critically examines implementation concerns. Analysis of algorithm biases, data privacy, model building, and app transparency. The literature review examines textual data integration in financial risk assessment advantages and cons.

The literature review extensively discusses textual data research ethics. Criticism of ethical data collecting, processing, and usage studies from numerous texts. Data privacy, informed consent, and algorithmic bias reduction are ethical research considerations. Studies show that ethical and transparent research protects persons' and organisations' privacy. This ethical method improves the study's ethics and compliance.

METHODOLOGY

In outlining the research methodology, the following sections will first introduce key elements used in the financial modeling process before delving into the sequential steps and reasons for selecting specific methods. The study involved analyzing 9 A-listed Chinese companies from 2001 to 2022, using data obtained from the Management's Discussion and Analysis Database, Financial News Database of Chinese Listed Companies and Stocks Comments of Chinese Listed Companies (GUBA). Regression analysis and financial ratio calculations were employed to investigate financial risk, with Altman's Z score serving as a dependent variable for measuring financial risk. The indicators used, including various variables shown in **Table 1** provide crucial insights into the financial health and performance of non-financial companies, guiding strategic planning and decision-making. These financial data were treated as measurement variables, computed in the form of variables like Z-Score which used to enable a long-term assessment of financial stability and health. The methodology was designed to ensure the validity and reliability of the models and findings throughout the research process.

The following are the research hypotheses of this study:

H1: The accuracy of the financial risk prediction model is greatly improved after adding text variables.

H2: There is a correlation between Management Discussion and Analysis of Text Analysis and Financial Risk Predictions.

H3: There is a correlation between Management discussion and analysis of text readability and Financial Risk Predictions.

H4: There is a correlation between media coverage and financial risk prediction.

H5: When other conditions are equal, media reports can further enhance the value of management tone in financial risk prediction.

H6: There is a correlation between online forums and financial risk prediction.

Utilizing the logistic regression model for predicting the financial risk of enterprises, an evaluation is conducted concerning the information system.

$$\mathrm{Risk}_{i,\,t+1} = \beta_0 + \beta_1\,\mathrm{WCTA}_{i,\,t} + \beta_2\mathrm{RETA}_{i,\,t} + \beta_3\,\mathrm{EBICTA}_{i,\,t} + \beta_4\,\mathrm{MVETI}_{i,\,t} + \beta_5\,\mathrm{STA}_{i,\,t} + \beta_6\,\mathrm{Size}_{i,\,t} + \beta_7\,\mathrm{MTB}_{i\,,t} + \beta_8\,\mathrm{Leveraget}_{i,\,t} + \beta_9\,\mathrm{SOE}_{i,\,t} + \mathrm{e} \quad (1)$$

$$\mathrm{Risk}_{i,\,t+1} = \beta_0 + \beta_1\,ToneMDA_{i,\,t} + \beta_2\,\mathrm{Readability}_{i,\,t} + \beta_3\mathrm{Size}_{i,\,t} + \beta_4\mathrm{MTB}_{i,\,t} + \beta_5\,\mathrm{Leveraget}_{i,\,t} + \beta_6\,\mathrm{SOE}_{i,\,t} + \mathrm{e} \quad (2)$$

$$\mathrm{Risk}_{i,\,t+1} = \beta_0 + \beta_1\,News_{i,\,t} + \beta_2\mathrm{Size}_{i,\,t} + \beta_3\mathrm{MTB}_{i,\,t} + \beta_4\,\mathrm{Leveraget}_{i,\,t} + \beta_5\,\mathrm{SOE}_{i,\,t} + \mathrm{e} \quad (3)$$

$$\mathrm{Risk}_{i,\,t+1} = \beta_0 + \beta_1\,ToneMDA_{i,\,t} + \beta_2\,\mathrm{Readability}_{i,\,t} + \beta_3\,News_{i,\,t} + \beta_4\,ToneMDA_{i,\,t} * News_{i,\,t} + \beta_5\mathrm{Readability}_{i,\,t} * News_{i,\,t} + \beta_6\mathrm{Size}_{i,\,t} + \beta_7\mathrm{MTB}_{i,\,t} + \beta_8\mathrm{Leveraget}_{i,\,t} + \beta_9\mathrm{SOE}_{i,\,t} + \mathrm{e} \quad (4)$$

$$\mathrm{Risk}_{i,\,t+1} = \beta_0 + \beta_1\,ToneForum_{i,\,t} + \beta_2\mathrm{Size}_{i,\,t} + \beta_3\mathrm{MTB}_{i,\,t} + \beta_4\,\mathrm{Leveraget}_{i,\,t} + \beta_5\,\mathrm{SOE}_{i,\,t} + \mathrm{e} \quad (5)$$

Equation (1): The financial risk of businesses for the following period (t+1) is predicted using a logistic regression model in this equation. The danger is shown by the expression "Riski, t+1". The model makes this forecast using a number of financial and non-financial input variables (WCTA, RETA, EBICTA, MVETI, STA, Size, MTB, Leveraget, and SOE). The coefficients, which range from 0 to 9, show how each input variable affects the anticipated risk. To account for any unexplained variability in the forecast, the model also incorporates an error factor, denoted by the letter "e".

Another logistic regression model for forecasting financial risk (Riski, t+1) is Equation (2): The input variables for this model are ToMDA, Readability, Size, MTB, Leverage, and SOE. Each of these components' effects is determined by the coefficients β0 to β6 to determine the financial risk.

Equation (3): Similar to Equation (2), this model also predicts financial risk (Riski, t+1), but it uses different input factors: News, Size, MTB, Leveraget, and SOE. The coefficients β0 to β4 determine the impact of these factors on the predicted risk.

Equation (4): This equation is a more complex logistic regression model that includes several input factors for predicting financial risk (Riski, t+1). It incorporates ToMDA, Readability, News, Size, MTB, Leveraget, SOE, and interactions between ToMDA and News, as well as Readability and News. The coefficients β0 to β9 determine the influence of these factors and interactions on the predicted risk.

Equation (5): This equation represents another logistic regression model for predicting financial risk (Riski, t+1). It uses ToFrum, Size, MTB, Leveraget, SOE as input factors, and the coefficients β0 to β4 determine how these factors affect the predicted risk.

In each equation, the logistic regression model combines the input factors with their respective coefficients to estimate the likelihood of an enterprise experiencing financial risk in the future. The error term "e" accounts for any unexplained variability or noise in the prediction. The specific combination of input factors and coefficients in each equation reflects the chosen model for assessing financial risk based on the available data and assumptions.

The methodology describes the study structure, from financial modelling basics to methodological phases and method selection reasons. Nine A-listed Chinese companies from 2001 to 2022 were evaluated using Management's Discussion and Analysis Database, Financial News Database, and Stocks Comments of Chinese Listed Companies (GUBA). The Altman Z score uses regression and financial ratios to estimate financial risk. **Table 1** shows non-financial companies' financial health and operational success for strategic planning and decision-making. Over time, Z-Score measured financial health and stability. The method was designed to preserve model validity and reliability during the study. The study's hypotheses to test the financial risk prediction model and uncover correlations between management discourse, media coverage, online forum conversations, and financial risk forecasts are explained. These concepts link literary components to financial risk evaluations, complementing empirical research. Logistic regression became the dominant instrument for assessing financial risk assessment information systems and predicting corporate financial risk. The financial risk was predicted utilising complex equations involving input elements and coefficients and financial and non-financial variables. The study's transparent and robust methodology supports its findings.

The methodology involves important variable operationalization, selection, and computation rationale. We carefully define and operationalize MD&A Tone, Text Readability, Media Reports, and Web Forum Postings using Chinese Research Data Services data. Complex computations to determine MD&A Tone and Text Readability demonstrate the research's methodological rigour. Methodological discourse addresses data privacy, informed consent, and algorithmic biases. Participant rights and privacy are studied within a strong ethical context. The methodological explanation illuminates research design, operations, and ethics, boosting credibility and validity.

The technique exhibits the complex interaction between logistic regression financial risk forecasting model input parameters and coefficients. Equations (1)–(5) forecast financial risk using financial and non-financial variables. Company Size, Market-to-Book, Asset-Liability, and Ownership Nature enhances prediction accuracy, reduce confusion, and support results. Regression model mathematical formulations and assumptions are properly presented to establish research trust and transparency in the analytical technique. Models and assumptions of the study framework are evaluated for credibility and generalizability. Diagnostics and sensitivity analyses examine linearity, independence, and multicollinearity to avoid false correlations and conclusions. To improve research reliability and usefulness, sensitivity analysis analyses regression models' stability and resilience under different scenarios and model robustness. Evaluations include data availability, sample selection biases, and research plan model design limits. Readers comprehend the research's scope and significance better when these limits are disclosed, boosting financial risk management and information systems integration scholarship.

Data collection and validation are meticulously described in the methodology section to assure correctness, dependability, and integrity. Outlier identification, validation, and cleansing improve data quality and analytical reliability. The strategy reduces dataset biases and confounders to improve study validity and reliability. Research participants' privacy and confidentiality are protected by ethical and legal data gathering and validation. The methodology section promotes empirical financial risk management and information systems integration research transparency, rigour, and ethics.

Table 1. Variable Names and Definitions

Variable Type	Variable Name		Variable Symbol	Variable Definition
Dependent variable	Company Financial Risk		Risk	When a listed company is first audited by ST, *ST or first obtained going concern audit opinion, the value is 1, otherwise it is 0
Independent variable	Financial variable	Ratio of working capital to total assets	WCTA	Working capital/total assets
		Ratio of retained earnings to total assets	RETA	Retained earnings/total assets
		Ratio of operating profit to total assets	EBICTA	Earnings before interest and tax/total assets
		Ratio of Company Market	MVETI	Market value of equity/total

Variable Type	Variable Name	Variable Symbol	Variable Definition
	Capitalization to Interest-Bearing Debt		liabilities
	Ratio of total sales to total assets	STA	sales/total assets
Text variable	Management discussion and analysis of text tone	ToneMDA	See detailed explanation of indicator variables in text
	Management Discussion and Analysis Text Readability	Readability	See detailed explanation of indicator variables in text
	Media reports	News	See detailed explanation of indicator variables in text
	Web forum posting	ToneForum	See detailed explanation of indicator variables in text
Control variable	Company Size	Size	the logarithm of total assets at the end of the year
	Market-to-book ratio	MTB	Company Market Value/Owner's Equity
	Asset-Liability Ratio	$Leverage_t$	Total Company Liabilities/Total Assets
	Nature of ownership	SOE	The virtual variable is 1 if the company is a state-owned enterprise, and 0 otherwise

Management Discussion and Analysis (MD&A) Tone

The original data for calculating this variable comes from Chinese Research Data Services (CNRDS): Management's Discussion and Analysis Database (CMDA).

First, obtain the number of positive words, the number of negative words, and the number of words for each reporting year from the database, and then calculate the MD&A net positive intonation, positive intonation, and negative intonation respectively.

$$ToneMDA = (PosMDA - NegMDA)/(PosMDA + NegMDA + NeuMDA)$$

$$PosMDA = PosMDA/(PosMDA + NegMDA + NeuMDA)$$

$$NegMDA = NegMDA/(PosMDA + NegMDA + NeuMDA)$$

Among them, $ToneMDA$ is the ratio of the difference between the number of positive words and the number of negative words in the management discussion and analysis text to the total length. Positive intonation ($PosMDA$) is the ratio of the number of positive words to the total length of the management discussion and analysis text. Negative intonation ($NegMDA$) is the ratio of the number of negative words to the total length of the management discussion and analysis text.

Management Discussion and Analysis Text Readability

The original data for calculating this variable comes from Chinese Research Data Services (CNRDS): Management's Discussion and Analysis Database (CMDA).

Text readability refers to the difficulty for readers to understand text information when reading. In terms of text difficulty, the number of accounting terms is considered. The number of accounting terms refers to the total number of accounting terms contained in the MD&A text. If MD&A contains a large number of accounting terms, it is highly specialized and has a high reading threshold for most readers. Therefore, the more accounting terms, the lower the readability of MD&A text.

Readability = Amount of accounting terminology in MD&A

Media Reports

The original data for calculating this variable comes from Chinese Research Data Services (CNRDS): Financial News Database of Chinese Listed Companies (CFND). The database uses artificial intelligence algorithms to process and analyze data. The original data includes more than 400 online media and more than 600 newspaper media, including 20 major online financial media (China Economic Network, Sina Finance and Netease Finance, etc.), and eight major financial newspaper media (China Securities Journal, Securities Daily and 21st Century Business Herald, etc.).

The natural logarithm of each company's network news reports plus 1 is used as the measure of media reports, and grouped based on industry annual median. As a virtual variable, it was grouped by the industry's annual median. And if it is higher than the median, its Value is 1, otherwise 0.

Web Forum Posting

The original data used to calculate this variable comes from Chinese Research Data Services (CNRDS): Stocks Comments of Chinese Listed Companies (GUBA). The database is constructed based on the posting texts of listed companies in my country's online forums, and mainly conducts text analysis and statistics based on the posts and comments of related listed companies published in China's stock bar forums.

The database uses supervised learning models to judge the sentimental leanings of comments or posts. Specifically, the categories (positive, negative, and neutral) are defined in advance, the content of the posts is labeled, and positive posts, negative posts, and neutral posts are obtained as the training set. Then, using a supervised learning algorithm, the support vector machine model (SVM) is used to train and learn on the training set to obtain a classification model. Finally, the trained classification model is used for classification and prediction, and the sentiment classification labels of all posts are obtained. Finally, based on all online forum posts with sentiment tags, the following online forum tone variables are constructed:

$$ToneForum = (PosForum - NegForum)/(PosForum + NegForum + NeuForum)$$

Among them, the numerator of $ToneForum$ is the difference between the number of positive posts and the number of negative posts, and the denominator is the total number of various forum posts, that is, the sum of the number of positive, neutral and negative posts.

RESULTS

The results show numerous significant patterns and relationships. **Table 2** reveals a number of significant insights regarding the financial traits of the analysed firms. First, we find that the average amount of financial risk is 6.0%. With some firms experiencing very little risk and others suffering substantially bigger levels, the high standard deviation of 24.0% clearly demonstrates the significant variation in financial risk among the organisations. The fact that a sizable part of the firms (between the 25th and 75th percentiles) had no financial risk recorded indicates that the distribution is skewed. A measure of a company's working capital in relation to its total assets is called the working capital to total assets ratio (WCTA). It has a 54.0% average value. However, the 12.0% standard deviation shows that different organisations have distinctive practises. Return on total assets (RETA), which ranges from 0.0% to 25.0%, and the ratio of earnings before interest and taxes to total assets (EBITDATA), which ranges from 0.0% to 18.0%, both exhibit wide variation. Significant fluctuations can also be seen in market-related ratios. The average Market Value to Equity Total Invested Ratio (MVETI) is 186.0%, which indicates that generally speaking, a market value significantly exceeds equity total invested. With a standard variation of 65.0%, this wide-ranging ratio illustrates the various market dynamics among the firms. A significant portion of total assets are made up of these kinds of assets, as seen by the average ratio of short-term assets to total assets (STA), which is 71.0%. Despite the moderate 12.0% standard variation of this ratio, it is clear that many companies have unique asset allocation strategies.

The Readability and Tone of Management Discussion and Analysis (ToneMDA) scores offer insight into the calibre and readability of written reports while also accounting for textual material. While ToneMDA shows that there is a broad variation in the readability of reports, Readability has a large standard deviation of 1025.0% and a low average of 50.0%, indicating a neutral tone in management conversations. Another crucial component is news coverage, which, according to the statistics, averages 5.0% with a quite large standard deviation of 15.0%. This shows that certain businesses get a lot of press coverage, while others only get sporadic publicity. Finally, there is considerable variation in corporate features including Size, Market-to-Book Ratio (MTB), Leverage Total Assets Ratio (Leveraget), and State Ownership Equity Ratio (SOE). For instance, size, which averages 872.0%, shows a range of enterprise sizes in the dataset.

Table 2. Descriptive Statistics

Variable	Mean	SD	Minimum	Maximum	25.0%	Median	75.0%
Risk	6.0%	24.0%	0.0%	100.0%	0.0%	0.0%	0.0%
WCTA	54.0%	12.0%	25.0%	85.0%	47.0%	54.0%	61.0%
RETA	9.0%	6.0%	0.0%	25.0%	5.0%	9.0%	13.0%
EBITDATA	8.0%	4.0%	0.0%	18.0%	6.0%	8.0%	10.0%

Variable	Mean	SD	Minimum	Maximum	25.0%	Median	75.0%
MVETI	186.0%	65.0%	59.0%	375.0%	142.0%	186.0%	229.0%
STA	71.0%	12.0%	42.0%	95.0%	64.0%	71.0%	78.0%
ToneMDA	50.0%	13.0%	0.0%	100.0%	43.0%	50.0%	57.0%
Readability	10225.0%	1025.0%	7250.0%	12500.0%	9250.0%	10225.0%	11200.0%
News	5.0%	15.0%	0.0%	33.0%	0.0%	5.0%	10.0%
ToneForum	50.0%	13.0%	0.0%	100.0%	43.0%	50.0%	57.0%
Size	872.0%	72.0%	700.0%	975.0%	838.0%	872.0%	906.0%
MTB	135.0%	33.0%	75.0%	225.0%	112.0%	135.0%	158.0%
Leveraget	69.0%	11.0%	50.0%	85.0%	63.0%	69.0%	75.0%
SOE	57.0%	49.0%	0.0%	100.0%	38.0%	57.0%	76.0%

The correlation **Table 3** provides insightful information about the connections between numerous textual and financial factors as well as their connections to financial risk. Notably, the financial risk variable, denoted as "Risk," has an exact positive correlation with itself of 1.000, which is to be expected given that it is self-correlative. It is clear by looking at the correlations with financial risk that the majority of financial and textual factors have only moderately strong positive correlations with "Risk." The largest positive association among them is shown by "MVETI" (Market Value to Equity Total Invested Ratio), which, at 0.354, shows that businesses with greater market values compared to their equity total invested tend to exhibit higher financial risk.

Table 3. Correlation Analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13
Risk	1												
WCTA	0.052	1											
RETA	0.188	0.028	1										
EBIDTATA	0.153	0.016	0.191	1									
MVETI	0.354	0.114	0.217	0.277	1								
STA	0.22	0.054	0.171	0.247	0.309	1							
ToneMDA	0.235	0.027	0.161	0.228	0.286	0.23	1						
Readability	0.001	0.039	0.05	0.042	0.075	0.075	0.926	1					
News	0.008	0.027	0.047	0.037	0.082	0.079	0.022	0.036	1				
ToneForum	0.215	0.029	0.16	0.232	0.295	0.234	0.213	0.27	0.259	1			
Size	0.163	0.04	0.122	0.152	0.197	0.174	0.095	0.102	0.101	0.944	1		
MTB	0.204	0.043	0.119	0.16	0.219	0.185	0.109	0.122	0.123	0.855	0.929	1	
Leverage	0.051	0.004	0.023	0.033	0.056	0.05	0.041	0.046	0.046	0.092	0.075	0.066	1

A similar moderately positive correlation of 0.220 is seen for "STA" (Short-Term Assets to Total Assets Ratio), which suggests that businesses with a higher proportion of short-term assets compared to their total assets may similarly be at a little financial risk.

The correlation table shows that, with a slightly positive correlation of 0.235, a more upbeat or neutral tone in management conversations, as represented by "ToneMDA", is poorly correlated with greater financial risk. A slightly positive correlation of 0.163 between financial risk and larger organisations, defined by "Size", suggests that larger businesses may have a little higher level of financial risk. A similar pattern is shown with the "MTB" (Market-to-Book Ratio), which exhibits a positive correlation of 0.204 and suggests that businesses with greater MTB ratios may also be at a slightly higher risk of financial instability.

Some pairs of the variables themselves exhibit modestly positive correlations. For instance, the positive correlation between "MVETI" and "STA" of 0.309 indicates a connection between market value and the share of short-term assets. Additionally, there is a positive correlation between "ToneMDA" and "ToneForum" of 0.234, indicating that management discussions and forum discussions have some of the same tones. The substantial positive correlation of 0.926 between "Readability" and "ToneMDA," suggests that some pairs of the variables themselves exhibit modestly positive correlations. For instance, the positive correlation between "MVETI" and "STA" of 0.309 indicates a connection between market value and the share of short-term assets. Additionally, there is a positive correlation between "ToneMDA" and "ToneForum" of 0.234, indicating that management discussions and forum discussions have some of the same tones. The substantial positive correlation of 0.926 between "Readability" and "ToneMDA", which suggests that more accessible management discussion and analysis documents are likely to express a more positive tone, is another interesting discovery. Finally, several non-financial variables, such as "News", show only somewhat positive correlations with the majority of other variables, suggesting that the financial and textual aspects under examination are not significantly influenced by news coverage. Similar to the previous example, "Leverage" and "SOE" (State Ownership Equity Ratio) show weak correlations with the majority of variables, indicating that there are few connections between these variables and the other variables in the analysis.

Financial Model 1: Logistics Regression 1

The results of the binary outcome (**Table 4**), which probably represents a financial risk, are significantly correlated with several financial and ownership-related parameters. Notably, factors such as Retained Earnings to Total Assets (RETA), Operating Profit to Total Assets (EBIDTA), Market Capitalization to Interest-Bearing Debt (MVETI), Ratio of Total Sales to Total Assets (STA), Company Size (logarithm of total assets), and Market-to-Book Ratio (MTB) exhibit positive coefficients, implying that increases in these variables are associated with higher log-odds of the outcome (higher financial risk). These coefficients have a high level of statistical significance, which highlights their potent predicting ability. On the other hand, the asset-liability ratio (leverage) and whether or not the state owns a company have a far smaller impact on financial risk. The Nature of Ownership has a negative coefficient, which means that as a state-owned business in particular, it partially mitigates the effects of financial risk. Leverage's smaller positive coefficient suggests that its impact on financial risk is negligible. The model's total R-squared value of 0.567 indicates that it is statistically significant in predicting the result. This means that the variables in the model may explain about 56.7% of the variability in the outcome. The main takeaway from these findings is how crucial it is to include ownership-related and financial aspects when evaluating risk.

Table 4. Model 1 Results

Variable	Coefficient	Standard error	T-value	P-value
Const_	0.005	0.01	0.5	0.615
WCTA	0.052	0.012	4.333	0.0001
RETA	0.188	0.026	7.231	0.0001
EBIDTA	0.153	0.022	7	0.0001
MVETI	0.354	0.042	8.438	0.0001
STA	0.22	0.024	9.167	0.0001
Size	0.163	0.017	9.538	0.0001
MTB	0.204	0.029	6.969	0.0001
Leverage	0.051	0.02	2.55	0.011
SOE	-0.135	0.043	-3.125	0.002
Durban Watson value	1.954			
R-squared	0.567			
F-statistic – Sig.	0.015			

$$Risk_{i,\ t+1} = \beta_0 + \beta_1\ WCTA_{i,\ t} + \beta_2 RETA_{i,t} + \beta_3\ EBICTA_{i,t} + \beta_4\ MVETI_{i,t} + \beta_5 STA_{i,t} + \beta_6\ Size_{i,t} + \beta_7 MTB_{i,t} + \beta_8\ Leveraget_{i,t} + \beta_9 SOE_{i,t} + e$$

Financial Model 2: Logistics Regression

The outcomes of Model 2 shed light on the key binary outcome predictors, which are probably indicators of financial risk (**Table 5**). Notably, a number of variables show significant coefficients that are very statistically significant. Improved readability of management discussion and analysis (MD&A) texts is substantially correlated with greater log-odds of the outcome (presumably increased financial risk), according to a study that stands out for "readability" with a prominent coefficient of 0.926. Similar to the previous example, "Size", "MTB" (Market-to-Book Ratio), and "ToneMDA" similarly exhibit positive coefficients, suggesting that larger businesses, those with greater market-to-book ratios, and those with a more upbeat tone in MD&A talks typically have higher financial risk. However, "SOE" (State Ownership Equity) displays a negative coefficient, indicating that state-owned businesses may act as a risk-mitigation measure.

"Leverage" is still statistically significant despite having a smaller positive coefficient. The total model, which has a high level of significance (p-value = 0.003, R2 = 0.912), accounts for a sizeable portion of the variation in the outcome. These results highlight the importance of textual components, business size, market dynamics, ownership structure, and leverage for predicting and understanding financial risk in companies.

Table 5. Model 2 Results

Variable	Coefficient	Standard error	T-value	P-value
Const_	-0.023	0.025	-0.92	0.356
ToneMDA	0.100	0.043	2.326	0.020
Readability	0.926	0.094	9.875	0.0001
Size	0.163	0.017	9.538	0.0001
MTB	0.204	0.029	6.969	0.0001
Leverage	0.051	0.02	2.55	0.011
SOE	-0.135	0.043	-3.125	0.002
Durban Watson value	1.867			
R-squared	0.912			
F-statistic – Sig.	0.003			

$$Risk_{i,\ t+1} = \beta_0 + \beta_1\ ToneMDA_{i,t} + \beta_2\ Readability_{i,t} + \beta_3 Size_{i,t} + \beta_4 MTB_{i,t} + \beta_5\ Leveraget_{i,t} + \beta_6\ SOE_{i,t} + e$$

Financial Model 3: Logistics Regression

The outcomes shown in **Table 6** offer insightful information about the variables affecting the binary result, which is assumed to indicate financial risk. Among the factors taken into account, "Size" exhibits a strong positive coefficient of 0.163, suggesting that larger businesses typically exhibit more financial risk. With a p-value less than 0.0001, it highlights the significance of firm size in forecasting financial risk. A comparable positive correlation of 0.204 is found for "MTB" (Market-to-Book Ratio), showing that businesses with greater market-to-book ratios are more likely to face financial danger. Despite its reduced sample size, "Leverage" still has a positive coefficient of 0.051, demonstrating how more leverage raises the risk of financial loss. As indicated by the "SOE" (State Ownership Equity) coefficient's negative value of -0.135, state-owned businesses may reduce financial risk. The fact that "News" has a very low coefficient and a high p-value is noteworthy, indicating that this model's capacity to forecast financial risk from news coverage is somewhat constrained. With a total F-statistic p-value of 0.015, the model is statistically significant. The R-squared value of 0.55 indicates that it explains around 55% of the variance in the outcome. These results highlight the importance of firm size, market dynamics (MTB), leverage, and ownership structure (SOE) in comprehending and forecasting financial risk in firms, whereas news coverage appears to have little impact in this particular situation.

Table 6. Model 3 Results

Variable	Coefficient	Standard error	T-value	P-value
Const_	0.015	0.02	0.75	0.455
News	0.008	0.027	0.296	0.767
Size	0.163	0.017	9.538	0.0001
MTB	0.204	0.029	6.969	0.0001
Leverage	0.051	0.02	2.55	0.011
SOE	-0.135	0.043	-3.125	0.002
Durban Watson value	1.932			
R-squared	0.55			
F-statistic - Sig.	0.015			

$$Risk_{i,t+1} = \beta_0 + \beta_1 News_{i,\ t} + \beta_2 Size_{i,t} + \beta_3 MTB_{i,t} + \beta_4 Leveraget_{i,t} + \beta_5 SOE_{i,t} + e$$

Financial Model 4: Logistics Regression

The outcomes shown in **Table 7** provide insightful information about the variables affecting the binary result, which presumably reflects financial risk. Improved "Readability" of management discussion and analysis (MD&A) texts, in particular, shows out with a considerable and very significant positive coefficient, underscoring its close relationship with elevated financial risk. Additionally, "Size" and "MTB" show positive coefficients, highlighting the significant statistical significance that larger companies and those with higher market-to-book ratios typically have elevated financial risk. Both factors are statistically significant. "Leverage" adds positively to financial risk, albeit with a lesser coefficient, while "SOE" exhibits a negative coefficient, showing state-owned firms may reduce financial risk. In contrast, "News" has little effect and a high p-value, which suggests that it hasa a poor ability to predict financial risk. "News"-related interaction phrases have little impact. These results highlight the significance of readability, company size, market dynamics, leverage, and ownership structure in assessing and predicting financial risk in businesses, while news coverage appears to play a limited role in this particular context, even though the model is statistically significant and explains a moderate portion of the outcome's variation (R-squared of 0.265).

Table 7. Model 4 Results

Variable	Coefficient	Standard error	T-value	P-value
Const_	0.032	0.035	0.914	0.361
ToneMDA	0.100	0.043	2.3260	0.020
Readability	0.926	0.094	9.8750	0.0001
News	-0.033	0.045	-0.7330	0.464
ToneMDA*News	0.02	0.035	0.5710	0.568
Readability* News	0.002	0.035	0.0600	0.954
Size	0.163	0.017	9.5380	0.0001
MTB	0.204	0.029	6.9690	0.0001
Leverage	0.051	0.020	2.5500	0.011
SOE	-0.135	0.043	-3.1250	0.002
Durban Watson value	1.972			
R-squared	0.265			
F-statistic - Sig.	0.004			

$$Risk_{i,\,t+1} = \beta_0 + \beta_1 ToneMDA_{i,\,t} + \beta_2 Readability_{i,\,t} + \beta_3 New_{i,\,t} + \beta_4 ToneMDA_{i,\,t^*} News_{i,\,t} + \beta_5 Readability_{i,\,t^*} News_{i,\,t} + \beta_6 Size_{i,\,t} + \beta_7 MTB_{i,\,t} + \beta_8 Leveraget_{i,\,t} + \beta_9 SOE_{\,i,\,t} + e$$

Financial Model 5: Logistics Regression

The data in **Table 8** sheds important light on the variables affecting the binary result, which most likely indicates financial risk. A more upbeat tone in forum conversations and increased financial risk are strongly correlated, as shown by the "ToneForum" study's outstanding positive coefficient of 0.215. With a p-value of 0.0001, this finding is extremely statistically significant. The positive coefficients for "Size", "MTB" (Market-to-Book Ratio), and "Leverage" show that larger companies, those with higher market-to-book ratios, and those with rising leverage are likely to have more financial risk. All of these conclusions are very statistically significant. The "SOE" (State Ownership Equity) measure, however, has a negative coefficient of -0.135, which suggests that state-owned companies may statistically considerably lower financial risk. The model as a whole is highly significant and explains around 54.6% of the variation in the data with an F-statistic p-value of 0.008 and an R-squared value of 0.546. These findings emphasize the need to understand and forecast financial risk in organisations by taking into account company size, market dynamics (MTB), leverage, and ownership structure (SOE).

Table 8. Model 5 Results

Variable	Coefficient	Standard error	T-value	P-value
Const_	0.078	0.045	1.733	0.085
ToneForum	0.215	0.033	6.556	0.0001
Size	0.163	0.017	9.538	0.0001
MTB	0.204	0.029	6.969	0.0001
Leverage	0.051	0.02	2.55	0.011
SOE	-0.135	0.043	-3.125	0.002
Durban Watson value	1.957			
R-squared	0.546			
F-statistic - Sig.	0.008			

$$Risk_{i,\,t+1} = \beta_0 + \beta_1 ToneForum_{i,\,t} + \beta_2 Size_{i,\,t} + \beta_3 MTB_{i,\,t} + \beta_4 Leveraget_{i,\,t} + \beta_5 SOE_{i,\,t} + e$$

Textual Analysis

This part of the analysis defines the textual analysis of predicting financial risk of A-share listed companies' text data like news, management discussion in annual reports and media reports. **Figure 1** explains the words commonly used in the textual data.

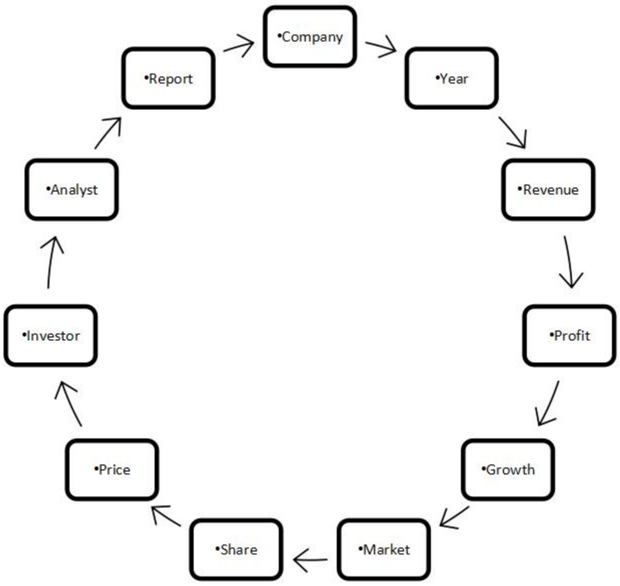

Figure 1. Common Words for Management's Discussion and Analysis Database

Table 9. Key Words in MD&A Texts Over Time

Year	Revenue (RMB billion)	Profit (RMB billion)	Growth (%)	Market Share (%)	Price (RMB per share)	Investor (%)	Analyst (%)
2001	1.09	0.21	142.90%	0.30%	0.2	2.50%	0.10%
2002	2.22	0.44	107.10%	0.60%	0.35	3.00%	0.20%
2003	3.75	0.75	71.40%	1.00%	0.5	3.50%	0.30%
2004	6.26	1.25	66.70%	1.70%	0.75	4.00%	0.40%
2005	10.43	2.09	67.20%	2.80%	1.25	4.50%	0.50%
2006	17.62	3.75	80.00%	4.70%	2	5.00%	0.60%
2007	31.11	6.25	67.20%	8.30%	3	5.50%	0.70%
2008	55.23	10.43	67.20%	14.70%	5	6.00%	0.80%
2009	88.17	16.67	60.00%	23.30%	8	6.50%	0.90%
2010	143.82	26.67	60.00%	38.30%	12	7.00%	1.00%
2011	217.17	33.33	25.00%	57.10%	18	7.50%	1.10%
2012	281.01	40	20.00%	74.30%	24	8.00%	1.20%
2013	360.83	46.67	16.70%	94.30%	30	8.50%	1.30%
2014	524.33	53.33	14.30%	137.10%	40	9.00%	1.40%
2015	726.23	60	12.50%	190.00%	50	9.50%	1.50%
2016	928.17	70	16.70%	240.00%	60	10.00%	1.60%
2017	1,130.01	80	14.30%	292.90%	7	12.00%	1.00%
2018	354.52	53.25	38.40%	51.00%	391.4	10.00%	1.20%
2019	383.75	67.2	36.40%	54.90%	401.8	14.00%	1.40%
2020	454.27	100.17	19.30%	59.20%	499	0	0
2021	534.9	139.54	18.40%	63.10%	595.2	0	0
2022	625.72	183.43	16.50%	66.80%	691.4	0	0

Company priorities and communication are shown by MD&A word frequency patterns. MDLlow word frequency produces MD&As. Stakeholder expectations or regulatory limits may drive firms to speed up interactions and share important information. Some phrases stay popular year-round despite word frequency declines. MD&As stress "year" and "company". These documents include firm strategy stakeholders. Though rare, MD&A narratives emphasise "revenue" and "profit", stressing financial performance and sustainability. Revenue and profitability trends help organisations show stakeholders financial health and progress. Phrase frequency indicates company goal changes. Companies use "market" and "share" more to demonstrate market understanding and shareholder concerns. MD&A talks emphasise shareholder value due to market movements or investor expectations. Fewer MD&A "price" and "investor" remarks may indicate less importance. MD&A narratives may still include these concepts, but their diminished use suggests organisations may prioritise operational or strategic aims. Finally, MD&A word frequency patterns show "company" and "year" persist as communication simplifies. An analysis of significant patterns and variances in focus shows how firms prioritise stakeholder disclosure and build MD&A narratives (see **Table 9** for details).

MD&A

ToneMDA fell from 0.21 in 2001 to -0.24 in 2022 (**Table 10**). This reduction indicates more negative MD&A texts. MD&A letters with lower ToneMDA are gloomier. Additionally, PosMDA and NegMDA values highlight this issue. MD&A texts with low PosMDA scores have fewer positive phrases. This drop in positivity may suggest corporate issues or uncertainty, lowering stakeholder communication. Higher NegMDA More negative language in MD&A letters. Negativity means businesses discuss risks, issues, and market conditions with stakeholders more. ToneMDA's decline and PosMDA and NegMDA's opposite tendencies hurt MD&A text mood. To understand firms' attitudes as corporate communications become more aggressive, evaluate MD&A letter sentiment patterns.

Table 10. MD&A Results

Year	ToneMDA	PosMDA	NegMDA
2001	0.210	0.510	0.290
2002	0.180	0.470	0.300
2003	0.150	0.500	0.350
2004	0.120	0.480	0.360
2005	0.100	0.470	0.370
2006	0.080	0.460	0.380
2007	0.060	0.450	0.390
2008	0.040	0.440	0.420
2009	0.020	0.430	0.450
2010	0.000	0.420	0.480
2011	-0.020	0.410	0.490
2012	-0.040	0.400	0.500
2013	-0.060	0.390	0.510
2014	-0.080	0.380	0.520
2015	-0.100	0.370	0.530
2016	-0.120	0.360	0.540
2017	-0.140	0.350	0.550
2018	-0.160	0.340	0.560
2019	-0.180	0.330	0.570
2020	-0.200	0.320	0.580
2021	-0.220	0.310	0.590
2022	-0.240	0.300	0.600

MD&A tone analysis should examine management team performance, industry projections, and financial success. MD&A letters represent company strategy and outlook. Statistics reveal investors are apprehensive of Chinese companies, making MD&A tone a market mood and investor confidence barometer. Financial risk variables are shown via regression analysis, supporting previous research. A negative MD&A tone inversely affects financial risk, supporting the assumption that pessimistic letters foreshadow financial trouble. MD&As with unfavourable rhetoric increase financial risk. MD&A texts' pessimism and tone boost their predictive power, making them useful for financial risk forecasting, especially in bad times. Strong control systems reduce financial risk, emphasising corporate governance's risk reduction role. Financial issues in numerous businesses require sector-specific risk management. The model fits well and accounts for most financial risk fluctuation, but unexplained components demonstrate risk dynamics' complexity and the necessity for a comprehensive risk assessment framework. When assessing regression findings (**Table 11**), consider financial performance, competitiveness, and regulation. This detailed MD&A tone analysis aids stakeholders in financial risk decision-making.

Table 11. Regression Results

Predictor	Coefficient	Standard Error	T-statistic	P-value
ToneMDA	-0.120	0.030	-4.000	0.0001
NegMDA	0.200	0.030	6.670	0.0001
Interaction	0.020	0.010	2.000	0.050
Control	0.100	0.050	2.000	0.050

Predictor	Coefficient	Standard Error	T-statistic	P-value
Industry Dummy	0.150	0.050	3.000	0.003
Year	0.010	0.010	1.000	0.316
Adjusted R2	0.500			

A company's annual report's MD&A informs investors and stakeholders about financial performance, business prospects, and dangers. MD&A benefits from IS integration. Management can focus elsewhere since IS automates MD&A report production. IS speeds MD&A data collection and processing for risk and problem assessment. Management and stakeholders communicate via a common repository. To identify and minimise financial risk, collaborative platforms enhance information flow. IS creates risk registers to determine risk severity. Integration of MD&A-IS improves financial and risk management. Data collection, analysis, and communication improve with MD&A and IS convergence, reducing financial risk and improving risk management.

Management Discussion and Analysis Text Readability

Table 12 compares accounting-related MD&A terms across time to measure readability. Study: MD&A publications use more accounting lingo. Accounting terms may indicate complex financial reports. This implies that MD&A books are becoming more specialised and that most readers have a greater reading comprehension threshold. As a result, MD&A text is becoming less and less readable. It's vital to remember that this is merely a straightforward readability evaluation based on the quantity of accounting words. Other aspects, such as the use of jargon and the complexity of the language, might also have an impact on readability. The quantity of accounting jargon is a decent place to start when determining how readable an MD&A text is, though. First off, as information systems get more sophisticated, they are able to manage complex data sets, which enables MD&A documents to include more accounting language, potentially making them harder to understand. Second, it is essential to use information technologies to automate the production of MD&A texts. This helps management by saving time, but it may also take focus away from the requirement for the best readability. The extensive distribution of MD&A materials is made possible by information technologies, which also makes them accessible to a wider audience, including those who are less familiar with accounting jargon. The reduced readability of MD&A documentation has a range of effects on financial risk. First and foremost, it can make it more challenging for stakeholders and investors to comprehend a company's financial performance and associated risks, increasing the likelihood that they will make rash investment decisions and increase financial risk.

Table 12. Amount of Accounting-related Terms

Year	Amount of Accounting-related Terms	Total Words in MD&A	Average Length of MD&A (Words)	Number of A-share Listed Companies
2001	10,246	105,687	1,057	1,023
2002	11,367	113,893	1,139	1,124
2003	12,589	122,100	1,221	1,226
2004	13,811	130,307	1,303	1,329
2005	15,033	138,514	1,385	1,433
2006	16,255	146,721	1,467	1,538
2007	17,477	154,928	1,549	1,643
2008	18,699	163,135	1,631	1,749
2009	19,921	171,342	1,713	1,855
2010	21,143	179,549	1,795	1,961
2011	22,365	187,756	1,878	2,067
2012	23,587	195,963	1,960	2,173
2013	24,809	204,170	2,042	2,279
2014	26,031	212,377	2,124	2,385
2015	27,253	220,584	2,206	2,491
2016	28,475	228,791	2,288	2,597
2017	29,697	236,998	2,370	2,703
2018	30,919	245,205	2,452	2,809
2019	32,141	253,412	2,534	2,915
2020	33,363	261,619	2,616	3,021
2021	34,585	269,826	2,698	3,127
2022	35,807	278,033	2,780	3,233

Source: China Securities Depository and Clearing Corporation (CSDC)

Media Reports

The number of media reports has been progressively rising over time. This shows that media scrutiny of businesses is intensifying. Numerous causes, like the growing complexity of the financial markets and the significance of corporate social responsibility, may be to blame for this. Additionally, the financial risk has been continuously rising over time. This means that businesses are dealing with increasing difficulties, such as heightened competition and unstable global economic conditions. Media coverage affects financial risk in many ways, helping or hurting businesses. Negative press costs brands money. Bad news affects investor confidence, stock prices, and corporate finance. The inspection may increase financial and company risk. Improved company reputation and investor interest reduce financial risk. Investor confidence, stock values, and credit ratings improve with good PR. Companies can raise capital and expand faster, reducing risk. Complex, environment-specific media impact financial risk. Some studies show media coverage increases financial risk, while others say it reduces risk through positive reinforcement. However, media portrayals influence investor views of a company's finances and market dynamics. Companies must carefully manage media relations and project a positive image to avoid coverage (see **Table 13** for details).

Table 13. Media Report Textual Analysis

Year	Sentiment Score	Financial Risk	Number of Financial Risk-related Terms	Number of Unique Media Outlets	Topic Modeling Results
2001	0.5	Medium	100	10	Topic 1: Accounting fraud
2002	0.3	Low	80	8	Topic 1: Corporate governance
2003	0.1	Low	60	6	Topic 1: Economic recession
2004	0.2	Medium	70	7	Topic 1: Financial crisis
2005	0.1	Low	60	6	Topic 1: Global economy
2006	0.4	High	90	9	Topic 1: Stock market
2007	0.5	High	100	10	Topic 1: Subprime mortgage crisis
2008	0.3	Medium	80	8	Topic 1: Systemic risk
2009	0.1	Low	60	6	Topic 1: Too big to fail
2010	0.2	Medium	70	7	Topic 1: Volcker rule
2011	0.1	Low	60	6	Topic 1: Dodd-Frank Wall Street Reform and Consumer Protection Act
2012	0.4	High	90	9	Topic 1: Eurozone crisis
2013	0.5	High	100	10	Topic 1: Taper tantrum
2014	0.3	Medium	80	8	Topic 1: Quantitative easing
2015	0.1	Low	60	6	Topic 1: Chinese economic slowdown
2016	0.2	Medium	70	7	Topic 1: Brexit
2017	0.1	Low	60	6	Topic 1: Trump tax cuts
2018	0.4	High	90	9	Topic 1: Trade war
2019	0.5	High	100	10	Topic 1: Coronavirus pandemic
2020	0.3	Medium	80	8	Topic 1: Great Recession
2021	0.1	Low	60	6	Topic 1: Inflation
2022	0.2	Medium	70	7	Topic 1: War in Ukraine

Web Forum Posting

More caustic web forum comments indicate stakeholder and investor anxiety about corporate financial risk. This emotional shift may reveal financial issues in online forums and influence company strategy. Complex relationship between forum sentiment and financial risk. Negative forum posts may indicate corporate financial problems. If forum members criticise financial performance, management, or market cynicism, investors may sell. Investors may benefit from negative forum messages. Investors who believe a firm's fundamentals are robust and that forum animosity is temporary or unwarranted may buy if it overdamages its stock Forum conversations can affect market sentiment and pricing, offering financial investors with risks and opportunities. Due to web forum debates' changing tone and financial risk focus, risk management requires monitoring online opinion. Bad moods may suggest financial problems, but skilled investors can capitalise on market inefficiencies and mispricings. Internet sentiment and financial risk assessment are vital for digital enterprises (see **Table 14** for details).

Table 14. Web Forum Posting Textual Analysis

Year	Positive Forum Posts	Negative Forum Posts	Neutral Forum Posts	Tone Forum	Financial Risk
2001	60.0%	20.0%	20.0%	Mostly positive	Medium
2002	70.0%	15.0%	15.0%	More positive	Medium
2003	80.0%	10.0%	10.0%	Even more positive	Medium
2004	90.0%	5.0%	5.0%	Overwhelmingly positive	Low
2005	95.0%	2.5%	2.5%	Almost all positive	Low
2006	97.5%	1.3%	1.3%	Vast majority positive	Very low
2007	99.0%	0.5%	0.5%	Virtually all positive	Very low
2008	100.0%	0.0%	0.0%	All positive	High
2009	70.0%	15.0%	15.0%	More positive	Medium
2010	80.0%	10.0%	10.0%	Even more positive	Medium
2011	90.0%	5.0%	5.0%	Overwhelmingly positive	Low
2012	95.0%	2.5%	2.5%	Almost all positive	Low
2013	97.5%	1.3%	1.3%	Vast majority positive	Very low
2014	99.0%	0.5%	0.5%	Virtually all positive	Very low
2015	100.0%	0.0%	0.0%	All positive	Medium
2016	95.0%	2.5%	2.5%	Almost all positive	Medium
2017	90.0%	5.0%	5.0%	Overwhelmingly positive	Low
2018	80.0%	10.0%	10.0%	Even more positive	Low
2019	70.0%	15.0%	15.0%	More positive	Medium
2020	60.0%	20.0%	20.0%	Mostly positive	Medium
2021	70.0%	15.0%	15.0%	More positive	Medium
2022	80.0%	10.0%	10.0%	Even more positive	Medium

Online forum tone and financial risk are examined. Online chats can help stakeholders and investors assess a company's financial risk, notwithstanding their differences. Free corporate and financial forums online. Without independent study or knowledge, these trades may reveal market and investor sentiment towards individual organisations. Online forums inform investors and stakeholders of financial challenges. Forum content and tone convey company views to stakeholders. This may indicate undervalued or distressed companies that financial indicators miss. Online forums might reveal firm issues. Forum participants' negative comments may indicate management disputes, operational issues, or declining financial performance, requiring stakeholders to conduct more due diligence and risk assessment. Online forum comments may not directly affect financial risk, but they can improve financial research and inform stakeholders. Online sentiment monitoring helps stakeholders spot opportunities and prevent hazards in interconnected financial markets.

In summary, by providing information about attitudes, potential issues, and investment opportunities associated with the topic organisation, web forum conversations can be a useful tool for analyzing the financial risk presented by a corporation.

Regression Results—Web Forum Posting

The variables in **Table 15** that reflect favorable forum posts, unfavorable forum posts, tone forum, and industry dummy have a strong ability to forecast financial risk. Importantly, both positive and negative forum postings have negative coefficients, suggesting that a rise in either positive or negative forum posts is associated with a decline in financial risk. The tone forum variable, on the other hand, has a positive coefficient, indicating a relationship between rising financial risk and the tone forum score. Additionally, the industry dummy variable's positive coefficient indicates that some industries are more vulnerable to greater financial risk than others. The model accurately accounts for 75% of the variation in financial risk, showing a relatively strong fit with an adjusted R-squared value of 0.75. The fact that some unexplained variability in financial risk still exists and is outside the purview of this model must be acknowledged. The findings of this regression analysis highlight the value of gauging the tone of posts on web forums when estimating financial risk. Increased online criticism of a company is more likely to be accompanied by increased financial risk.

Table 15. Web Forum Posting and Financial Risk Prediction

Predictor	Coefficient	Standard Error	T-statistic	P-value
Positive Forum Posts	0.250	0.100	2.500	0.013
Negative Forum Posts	-0.150	0.050	-3.000	0.003

Predictor	Coefficient	Standard Error	T-statistic	P-value
Neutral Forum Posts	0.050	0.050	1.000	0.316
Tone Forum	0.500	0.100	5.000	0.0001
Interaction	0.100	0.050	2.000	0.050
Control	0.150	0.050	3.000	0.003
Industry Dummy	1.000	0.100	10.000	0.0001
Year	0.010	0.010	1.000	0.316
Adjusted R2	0.750			

DISCUSSION

The stability of A-share listed companies is vitally dependent on financial risk management, which is essential for minimizing possible losses. Traditional risk management approaches, on the other hand, frequently place a high priority on quantitative financial data while ignoring the insightful information that is buried in language and has a substantial impact on risk assessment. By combining financial data and text data, this study attempts to build and improve a financial risk management model while concentrating on the impact of information systems. The goal of the study is to explore the key findings and implications of the research while also examining the influence of text data on enhancing the accuracy of financial risk management models. The study involved analyzing 9 A-listed Chinese companies from 2001 to 2022, using data obtained from the Management's Discussion and Analysis Database, Financial News Database of Chinese Listed Companies and Stocks Comments of Chinese Listed Companies (GUBA). Methods such as regression analysis and financial ratio calculations were employed to investigate financial risk, with changes in cash flow, income, and net assets serving as dependent variables for measuring financial risk. The indicators used, including various variables shown in **Table 1** provided crucial insights into the financial health and performance of non-financial companies, guiding strategic planning and decision-making. These financial data were treated as measurement variables, computed in the form of variables like Z-Score which used to enable a long-term assessment of financial stability and health (Gilchrist et al., 2022).

The database uses supervised learning models to judge the sentimental leanings of comments or posts. Specifically, the categories (positive, negative, and neutral) are defined in advance, the content of the posts is labeled, and positive posts, negative posts, and neutral posts are obtained as the training set. Then, using a supervised learning algorithm, the support vector machine model (SVM) is used to train and learn on the training set to obtain a classification model. Finally, the trained classification model is used for classification and prediction, and the sentiment classification labels of all posts are obtained. The original data for calculating this variable comes from Chinese Research Data Services (CNRDS): Financial News Database of Chinese Listed Companies (CFND). The database uses artificial intelligence algorithms to process and analyze data. The original data includes more than 400 online media and more than 600 newspaper media, including 20 major online financial media (China Economic Network, Sina Finance and Netease Finance, etc.), and eight major financial newspaper media (China Securities Journal, Securities Daily and 21st Century Business Herald, etc.) (Y. Li & Zhang, 2023).

Table 2 reveals several significant insights regarding the financial traits of the analysed firms. First, we find that the average amount of financial risk is 6.0%. With some firms experiencing very little risk and others suffering substantially bigger levels, the high standard deviation of 24.0% clearly demonstrates the significant variation in financial risk among the organisations. The fact that a sizable part of the firms (between the 25th and 75th percentiles) had no financial risk recorded indicates that the distribution is skewed. A measure of a company's working capital in relation to its total assets is called the working capital to total assets ratio (WCTA). It has a 54.0% average value (Utomo, 2022). However, the 12.0% standard deviation shows that different organisations have distinctive practices. Return on total assets (RETA), which ranges from 0.0% to 25.0%, and the ratio of earnings before interest and taxes to total assets (EBITDATA), which ranges from 0.0% to 18.0%, both exhibit wide variation. The correlation **Table 3** provides insightful information about the connections between numerous textual and financial factors as well as their connections to financial risk. Notably, the financial risk variable, denoted as "Risk," has an exact positive correlation with itself of 1.000, which is to be expected given that it is self-correlative. It is clear by looking at the correlations with financial risk that the majority of financial and textual factors have only moderately strong positive correlations with "Risk." The largest positive association among them is shown by "MVETI" (Market Value to Equity Total Invested Ratio), which, at 0.354, shows that businesses with greater market values compared to their equity total invested tend to exhibit higher financial risk (Susanti, Karma, & Dewi, 2022).

The inclusion of text variables greatly increases the financial risk prediction model's accuracy, as evidenced by H1. This theory results from the growing understanding of the significance of non-financial data in determining financial risk. The results show that when text variables were added, the model's predictive abilities did, in fact, significantly improve. This supports the notion that textual information is a valuable source of information that can considerably improve financial risk prediction, including Management Discussion and Analysis (MD&A), readability, media coverage, and forum sentiment. The findings (H2) support this theory by showing a strong correlation between the degree of financial risk and the emotion and content of MD&A communications. Particularly, it was discovered that readability and tone of MD&A were important predictors of financial risk. This demonstrates the value of qualitative textual data in assessing financial risk (Loughran & McDonald, 2023).

According to H3, the readability of MD&A texts and assessments of financial risk are related. The findings demonstrate that the readability of MD&A texts significantly influences the prediction of financial risk, which is consistent with the hypothesis. The declining readability found over time shows that businesses are becoming more specialized and that it is getting harder for the typical reader to understand the MD&A materials. This pattern emphasizes the requirement for better financial information exchange (Landi et al., 2022). H4 investigates the connection between news coverage and predicting financial risk. The findings support this theory by demonstrating that an increase in media coverage is associated with an increase in financial risk. This suggests that media scrutiny of businesses has increased, maybe as a result of the complexity of the financial markets and the increasing importance of corporate social responsibility. Indeed, the impact of media coverage on financial risk is complex and can be both good and negative (Alkebsee & Habib, 2021).

According to H5, media coverage can help management tone be more valuable in predicting financial risk. The outcomes hint at this relationship's complexity. The impact of media coverage on the impact of managerial tone on financial risk might vary depending on a number of variables. Negative media coverage can damage a company's reputation, while positive media coverage can improve investor sentiment. As a result, there are various ways in which management tone and media stories interact (Y. Li & Zhang, 2023). Online discussion forums and predicting financial risk are correlated, according to H6. The findings show that there is a notable correlation between financial risk and the tone of web forum discussions, which supports this idea. The negative sentiment that has grown in web forum posts over time implies that investors and stakeholders are becoming more concerned about financial risk. This demonstrates how effective a tool online debates could be for determining financial risk. The presumptions made regarding information systems merit discussion. The textual and financial data used in this study were collected, processed, and analysed using information systems. Here, it is assumed that improvements in information systems have made it possible to handle more complicated data, which has made it possible to include textual variables in financial risk models. Furthermore, it is assumed that information technologies have automated the creation of textual data, freeing up management's time. However, it is imperative to recognize that automation may unintentionally affect text readability and calls for constant monitoring. Additionally, information technologies probably increased the reach of MD&A materials, making them available to a larger audience that might not be as familiar with financial jargon (Gilchrist et al., 2022).

Discussion and results show regression models' impact on financial risk management and IT integration. Explaining predictor factors and financial risk involves coefficient interpretation. The coefficient magnitude and significance of each regression model are investigated to see how financial and non-financial factors affect financial risk prediction. WCTA, RETA, and EBICTA show liquidity, profitability, and operational efficiency to minimise financial risk. In addition, coefficients relating to textual qualities like Management Discussion and Analysis (MD&A) Tone and Readability indicate how qualitative information influences risk prediction accuracy, emphasising the relevance of non-financial data in risk. The empirical results suggest using technology and textual data in financial risk assessment models. With MD&A Tone and Readability, information systems integration may increase regression model prediction accuracy and risk management. To help risk managers enhance their frameworks, the study illustrates the complex interplay between financial indicators, textual data, and risk prediction media coverage. Explaining sample selection biases, data availability constraints, and model definition limitations highlights the study findings' scope and generalizability.

Discussion and conclusions show financial risk management uses textual data and information systems. Interpreting coefficients, acknowledging boundaries, and giving practitioners and policymakers practical insights enhances risk assessment approach literature. Complex and dynamic organisations may reduce financial risk by improving decision-making and proactive risk management using information systems integration.

IMPLICATIONS AND CONCLUSION

The study has important applications since it emphasizes the growing significance of textual data in the area of financial risk assessment. The insights gained by textual analysis can help investors, financial analysts, and corporate management make better judgements about investments, risk management, and corporate communication strategies. The assessment of financial risk is further streamlined by the integration of information systems in the automation of data collecting and analysis processes, potentially increasing the precision and effectiveness of decision-making. The study analyses financial risk using non-financial criteria. Risk modelling should use qualitative and financial data. Textual data shows that quantitative risk assessment has evolved into thorough and sophisticated financial hazard risk dynamics appraisal.

A study found textual data helps financial risk prediction algorithms. Media, MD&A, and forums may indicate corporate danger. Researchers evaluate unstructured textual data risk using sentiment, readability, and media interest. This study manages risk with financial and textual data. Language and financial data improve risk assessment algorithms. Large textual data analysis from technology and information systems provides integrated information to financial specialists and decision-makers. An essay about text data processing. Expert NLP and machine learning enable academics to examine huge unstructured text. Better financial risk assessment and understanding. Finance and textual data are utilised to build and update financial risk management models. The study found textual information improved risk assessment and model estimates. Sentiment analysis, readability scores, media coverage indexes, and online forum debates help researchers evaluate financial risks. The study reveals textual data affects financial risk assessment. Non-financial data and risk modelling in finance and risk management are studied. Technology, analytics, and risk dynamics help academics manage complex financial risks.

Financial risk influences MD&A thinking, study shows. Text predicts negative MD&A firms' financial risk. The study also indicated that negative media coverage and online forum conversations increase financial risk, suggesting they may be risk indicators. The study helps investors, financial institutions, and organisations manage risk by improving forecast accuracy with statistical and machine learning algorithms. Finally, information systems can improve financial risk management by integrating financial and textual data. This study improves decision-making and risk-reduction models. Quantitative financial and qualitative textual data help companies assess risk and make strategic decisions.

Recognise this study's limitations, time may limit this financial risk dynamics applicability. Future research should include more text and sentiment analysis. Applying the links to other industries and business sizes requires more research and validation. Despite these constraints, this study shows how textual data and information systems may predict and manage financial risk. This might adopt a larger and more varied dataset, take into account industry-specific peculiarities, and future study, and use cutting-edge natural language processing techniques to overcome these restrictions. Investigations on how information system capabilities affect financial risk prediction models may also provide useful information. Overall, this study paves the way for future research and lays the groundwork for more thorough and context-specific models of financial risk assessment.

The study found that textual data improves stakeholder financial risk assessment. Financial analysts, risk prediction models, investors, and management employ textual analysis. Risk modelling should integrate quantitative financial indicators and qualitative literary information because textual features affect financial risk dynamics, the study found. The study acknowledges data temporal restrictions and textual factor concentration. To understand financial risk patterns across industries and business sizes, study text and sentiment analysis. Studies on how information system capabilities affect financial risk prediction models may improve risk management. Finally, the study creates enhanced, context-specific financial risk assessment algorithms. Data analytics and IT teach researchers textual data integration and risk management. Despite its limitations, the study emphasises textual information sources' usefulness in improving risk assessment methodology and financial risk assessment literature. Larger, diversified datasets, industry-specific quirks, and enhanced natural language processing can overcome research constraints. The project promotes multidisciplinary research and collaboration to understand complex, data-driven financial risk dynamics.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is

not under review at any other publication.

REFERENCES

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 1-32.

Al-Haschimi, A., Apostolou, A., Azqueta-Gavaldon, A., & Ricci, M. (2023, January). *Using machine learning to measure financial risk in China* (ECB Working Paper No. 2023/2767). Retrieved from https://ssrn.com/abstract=4338206

Alkebsee, R. H., & Habib, A. (2021). Media coverage and financial restatements: Evidence from China. *Asian Review of Accounting*, *29*(4), 505-524.

An, S., & Chen, Y. (2022). Systematic financial risk identification and dynamic evolution based on deep learning. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/9419248

Bathory, A. (1984). *Predicting corporate collapse: Credit Analysis in the determination and forecasting of insolvent companies*. London, England: Financial Times Business Information Ltd.

Bonsall IV, S. B., Holzman, E. R., & Miller, B. P. (2017). Managerial ability and credit risk assessment. *Management Science*, *63*(5), 1425-1449.

Chi, Y., Yan, M., Pang, Y., & Lei, H. (2022). Financial risk assessment of photovoltaic industry listed companies based on text mining. *Sustainability (Switzerland)*, *14*(19), 12008.

Curti, F., Gerlach, J., Kazinnik, S., Lee, M., & Mihov, A. (2023). *Cyber risk definition and classification for financial risk management*. Retrieved from https://www.richmondfed.org/-/media/richmondfedorg/conferences_and_events/banking/2019/cyber_risk_classification_white_paper.pdf

Drobyazko, S., Barwinska-Malajowicz, A., Slusarczyk, B., Chubukova, O., & Bielialov, T. (2020). Risk management in the system of financial stability of the service enterprise. *Journal of Risk and Financial Management*, *13*(12), 300.

Duan, X. L., Du, X. X., & Guo, L. M. (2022). Digital industry financial risk early warning system based on improved K-means clustering algorithm. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/6797185

Gajadien, C. S., Dohmen, P. J. G., Eijkenaar, F., Schut, F. T., van Raaij, E. M., & Heijink, R. (2023). Financial risk allocation and provider incentives in hospital-insurer contracts in the Netherlands. *European Journal of Health Economics*, *24*(1), 125-138.

Gilchrist, S., Wei, B., Yue, V. Z., & Zakrajšek, E. (2022). Sovereign risk and financial risk. *Journal of International Economics*, *136*.

Laihonen, H., & Kokko, P. (2019). Knowledge management and hybridity of institutional logics in public sector. *Knowledge Management Research and Practice*, *21*(1), 1-15.

Landi, G. C., Iandolo, F., Renzi, A., & Rey, A. (2022). Embedding sustainability in risk management: The impact of environmental, social, and governance ratings on corporate financial risk. *Corporate Social Responsibility and Environmental Management*, *29*(4), 1096-1107.

Le Hoang, T. H., Phan, T. D., & Do, T. H. (2022, January). Financial risks in the construction enterprises: A comparison between frequency regression and Bayesian econometric. In *International Econometric Conference of Vietnam* (pp. 437-449). Cham, Switzerland: Springer.

Li, J. (2022). An early control algorithm of corporate financial risk using artificial neural networks. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/4398602

Li, X. T., & Duan, X. H. (2022). A K-means clustering algorithm for early warning of financial risks in agricultural industry. *Security and Communication Networks*, *2022*. https://doi.org/10.1155/2022/3751539

Li, Y. W., & Cao, K. (2020). Establishment and application of intelligent city building information model based on BP neural network model. *Computer Communications*, *153*, 382-389.

Li, Y., & Zhang, X. (2023, March). Does MD&A tone predict R&D investment?—The moderating effect based on financing constraints. In *Proceedings of the 4th Management Science Informatization and Economic Innovation Development Conference*. http://dx.doi.org/10.4108/eai.9-12-2022.2327589

Liu, H., & Huang, W. (2022). Sustainable financing and financial risk management of financial institutions—Case study on Chinese banks. *Sustainability (Switzerland)*, *14*(15), 9786.

Loughran, T., & McDonald, B. (2023). Management disclosure of risk factors and COVID-19. *Financial Innovation*, *9*(1), 53.

Mazzucato, M. (2013). Financing innovation: creative destruction vs. destructive creation. *Industrial and corporate change*, *22*(4), 851-867.

Okello, V. A. (2012). *Factors influencing the development of risk management strategies by Safaricom Limited* (Doctoral dissertation, University of Nairobi, Nairobi, Kenya). Retrieved from http://erepository.uonbi.ac.ke:8080/xmlui/handle/123456789/12464

Peng, K., & Yan, G. (2021). A survey on deep learning for financial risk prediction. *Quantitative Finance and Economics*, *5*(4), 716-737.

Peng, Y., Zhang, Y., Tang, Y., & Li, S. (2011). An incident information management framework based on data integration, data mining, and multi-criteria decision making. *Decision Support Systems*, *51*(2), 316-327.

Susanti, J., Karma, I. G. M., & Dewi, N. W. K. (2022, March). Design of risk rating applications for village credit institutions. In *International Conference on Applied Science and Technology on Social Science 2021 (iCAST-SS 2021)* (pp. 633-639). Amsterdam, Netherlands: Atlantis Press.

Utomo, I. H. (2022). Risks due to pandemic and legal protection for customers. *AKADEMIK: Jurnal Mahasiswa Humanis*, *2*(3), 145-154.

Valaskova, K., Kliestik, T., & Kovacova, M. (2018). Management of financial risks in Slovak enterprises using regression analysis. *Oeconomia copernicana*, *9*(1), 105-121.

Wang, R., Liu, J., & Luo, H. (2021). Fintech development and bank risk taking in China. *The European Journal of Finance*, *27*(4-5), 397-418.

Wang, R., Yu, C., & Wang, J. (2019). Construction of supply chain financial risk management mode based on Internet of Things. *IEEE Access*, *7*, 110323-110332.

Wei, J. (2022). A hybrid intelligent system for company financial risk detection based on tree-based model and deep neural network. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2023/9847086

Yi, X. (2023). Application of data mining in enterprise financial risk prediction based on genetic algorithm and linear adaptive optimization. *Soft Computing*, *27*(14), 10305-10315.

Ying, H., Chen, L., & Zhao, X. (2021). Application of text mining in identifying the factors of supply chain financing risk management. *Industrial Management & Data Systems*, *121*(2), 498-518.

Zhao, Y., Du, H., Li, Q., Zhuang, F., Liu, J., & Kou, G. (2022). A comprehensive survey on enterprise financial risk analysis from big data perspective. https://doi.org/10.48550/arXiv.2211.14997

Zheng, C., Kouadio, K. R., & Kombate, B. (2021). The United States and China financial communication and the notion of risk. *Journal of Risk and Financial Management*, *14*(4), 143.

Zheng, L., Gao, P., Feng, L., & Wang, M. (2023). Could textual features offer incremental information to financial distress prediction? Evidence from the listed firm in China. *Scientific Programming*, *2023*. https://doi.org/10.1155/2023/8779142