A Comprehensive Survey on Enterprise Financial Risk Analysis: Problems, Methods, Spotlights and Applications

Yu Zhao, Member, IEEE, Huaming Du

Abstract—Enterprise financial risk analysis aims at predicting the future financial risk of enterprises. Due to its wide and significant application, enterprise financial risk analysis has always been the core research topic in the fields of Finance and Management. Although there are already some valuable and impressive surveys on enterprise risk analysis from the perspective of Finance and Management, these surveys introduce approaches in a relatively isolated way and lack recent advances in enterprise financial risk analysis. Due to the rapid expansion of the enterprise risk research area, especially from the Computer Science and Big Data perspective, it is both necessary and challenging to comprehensively review the relevant studies. This survey attempts to connect and systematize the existing enterprise financial risk studies, i.e. to summarize and interpret the problems, methodologies, spotlights and applications of enterprise financial risk analysis in a comprehensive way, which may help readers to have a better understanding of the current research status and ideas. Unlike previous surveys, this paper attempts to provide a systematic literature survey of enterprise risk analysis approaches from the fields of Finance, Management, as well as Computer Science, which reviews more than 300 representative articles in the past almost 50 years (from 1968 to 2022). In particular, we first introduce the problems of the types, granularity, intelligence and evaluation metrics of enterprise financial risk, and summarize the representative works in terms of them respectively. Then, we compare the analysis methods used to learn enterprise financial risk, and summarize the spotlights of the most representative works. Finally, the applications of enterprise risk analysis are also briefly introduced. Our goal is to clarify current cutting-edge research and its possible future directions to model enterprise risk, aiming to fully understand the mechanisms of enterprise risk generation and contagion and its application on corporate governance, financial institution and government regulation.

Index Terms—Fintech, Enterprise Financial Risk, Small and Medium-sized Enterprises, Machine Learning, Deep Learning, Big Data, Enterprise Governance.

I. INTRODUCTION

Enterprise financial risk analysis has long been a focus of scholarly research in the fields of Finance and Management [1]. Whether young start-ups, small and medium-sized enterprises (SMEs) or famous Fortune 500 enterprises, to some extent they all inevitably face one or multiple financial risks, such as credit risk, guarantee risk, supply chain risk, bankruptcy risk, etc [1]. In particular, enterprises face big risk challenges when an economic crisis or a gray rhino incident happened, such as the financial crisis in 2008 [1] and COVID-19 epidemic in 2020 [2]. In the modern economic system, as the global financial system becomes more deeply interconnected, enterprises face both the intra-risk and contagion risk [3], [4]. Predicting the financial risk of enterprises is of great importance for both government policymakers and financial institutions.

Studies on enterprise financial risk analysis, which originated in the fields of Finance and Management, have gradually attracted increasing numbers of researchers from Computer Science (CS). The earliest research mainly focused on the three most commonly used statistical econometric methods, e.g. multivariate discriminant analysis, linear probability model, and logistic regression, and studied their application in enterprise risk prediction [5]. With technological advancements in natural language processing (NLP) and artificial intelligence (AI), researchers have begun to evaluate enterprise financial risk from a big data perspective. Besides traditional enterprise financial index, more enterprise risk intelligence is taken into consideration, such as non-financial textual information and relational data [],. Specifically, to deal with textual risk information, NLP techniques, including sentiment analysis and event extraction are used to dig enterprise risk signals from non-financial textual data. To model the risk momentum spillover on enterprise relational data, AI techniques, including deep learning and Graph Neural Networks (GNN), are utilized to evaluate the enterprise contagion risk. This series of studies opened up new avenues for understanding the mechanism by which enterprise financial risk generate and contagion from a big data perspective.

Figure [1] shows the number of relevant publications in the past twelve years and their distribution in terms of different research directions. Due to the rapid expansion of the enterprise risk research area, especially from the Computer Science and Big Data perspective, it is both necessary and challenging to comprehensively review the relevant studies. In this work, we systematically review nearly 300 representative articles on enterprise financial risk from the fields of Finance, Management, as well as Computer Science. We endeavor to

1In this work, we have collected nearly 2,000 publications about enterprise financial risk and have selected 320 of the best and most recent publications on enterprise risk.
summarize the problems, methodologies, spotlights and applications of enterprise financial risk analysis in a comprehensive way, which may help readers to have a better understanding of the current research status and ideas. Figure 1 shows the framework of this survey.

The remainder of this article is organized as follows. Section II formalizes the problems of enterprise risk analysis from four different directions, including the risk types, analysis granularity, risk intelligence and evaluation metrics. In Section III we will systematically introduce the approaches used more frequently. Section IV summarizes the spotlights of the most representative studies. Section V introduces the application of enterprise risk analysis. Section VI provides guidance for future research directions. Finally, we conclude the survey in Section VII.

II. PROBLEM FORMALIZATION

A. Financial Risk Types

In this section, we review the previous research on enterprise risk analysis in terms of risk types, i.e., credit risk, bankruptcy risk, guarantee risk, and financial risk, which are summarized in Table I.

1) Credit Risk: The enterprise credit risk refers to the possibility of loss due to the default of one party or the change of credit quality or morality in the transaction. The characteristics of enterprise credit risk are comprehensive, twoway, transitive and diffusive, cumulative, hidden, sudden and uncertain. Credit risks in the work mainly include risks faced by various credit transactions between financial institutions and enterprises, enterprises and enterprises, governments and enterprises [30].

There are two cases in which an enterprise cannot repay the loans of banks and financial institutions on time: one is that the enterprise has no ability to execute the contract and is forced to default; The second is that the enterprise has no intention of executing the contract subjectively and intentionally breaches the contract. According to an assumption of economic individuals in economics, enterprises may obtain additional benefits if they breach the contract. Therefore, enterprises will always make trade-offs between executing the contract and breaching the contract to maximize utility.

2) Bankruptcy Risk: Corporate bankruptcy risk refers to the risk that corporate assets are insufficient to repay their liabilities. Generally, it’s the result of many other risks, and now, bankruptcy risk management is attracting much attention from many managers and researchers owing to its essential role in investment and decision-making. Sheppard et al. [31] suggests that efforts should be put in early to fight against corporate decline. Thus, many works focus on risk prediction. But, why enterprises failed? It’s the base of any prediction work. Hence, we overview previous studies and divide factors into two aspects, external and internal factors.

As for external factors, many researchers find a relationship between the macroeconomic environment and bankruptcy. Early in 1998, Everett et al. [32] investigates the impact of some macroeconomic indexes on corporate failures like interest rate and unemployment rate. They are positively associated with bankruptcy rate. Further, Arcuri et al. [33] demonstrates the relationship between a higher financial development level and a lower probability of bankruptcy. Besides, macro environment policy is considered another inescapable shock. Industry-specific policies affect the risk level of companies in the industry [34]. Additionally, some global shock events also have strong power on corporate bankruptcy risk. Cowling et al. [35] suggests when Black Swan events occur, precautionary saving for SMEs is critical to enhancing resilience. When it comes to internal factors, we have to talk about accounting factors first. As Sun et al. [36] concludes, better financial positions mean the ability to control costs and good solvency. Thus, accounting factors really deserve our attention. Factors like firm size, current liquidity and financial structure have been explored by researchers [37] [38]. Further, researchers find the strong power of corporate governance structure in bankruptcy prediction [39]. Later, more and more studies focus on the relationship between directors and bankruptcy [40]. Besides, many other aspects have been explored by researchers as well. Like leveraged buyout transactions [41], downsizing [42], corporate social responsibility [43], and so on. In short, those comprehensive analyses of the reasons help us better
### TABLE I: Literature comparison in terms of enterprise risk types.

| Enterprise Risk Types | Literature | Aspect | Period | Size | Metric | Category | Methodology |
|-----------------------|------------|--------|--------|------|--------|----------|-------------|
| **Credit Risk**       | Individual Enterprise Risk | financial index | the whole spectrum | China | 1999-2004 | 120 | average misclassification | ML, SVM, LR |
|                       | Individual Enterprise Risk | financial and non-financial index | manufacturing | China | 2015-2017 | 1091 | AUC, KS | Statistical measurement, LR |
|                       | Systemic Risk | financial and non-financial index | manufacturing | China | 2011-2019 | 924 | average accuracy, Type I error, Type II error | DL, the multiview graph-based learning method |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | Italy | 2014 | 2.4M | GCS/CC | Hybrid Model, multinomial logic, classification tree and neural network |
| **Bankruptcy Risk**   | Individual Enterprise Risk | financial index | the whole spectrum | China | 2012-2013 | 48 | average accuracy, Type I error, Type II error, F-Measure | ML, Random Subspace-Real AdaBoost |
|                       | Supply Chain Risk | financial index | automotive | China | 2019 | 192 | classification accuracy | Hybrid Model, the genetic algorithm combined with support vector machine and RBF neural network |
| **Guarantee Risk**    | Individual Enterprise Risk | non-financial and relational index | the whole spectrum | China | 2016-2018 | 330,000 | AUC, IEMP | ML, two-stage multiobjective feature-selection method |
|                       | Individual Enterprise Risk | relational index | the whole spectrum | Belgium, UK | 2011-2014 | 2,800,000 | AUC | ML, Linear kernel training SVM |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | Japan | 2002-2016 | 2062 | average identification rates | DL, Convolutional neural networks |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | Korea | 1995-1998 | 168 | Correct classification, Type I error, Type II error | ML, the BP network and the Kohonen network |
|                       | Individual Enterprise Risk | financial and non-financial index | the whole spectrum | US | 1994-2014 | 118,27 | AUROC, accuracy, cumulative decile-ranking | DL, Word Embedding, CNN, LR, Random Forest, SVM |
| **Financial Risk**    | Systemic Risk | financial index | industrial | US | 1981-1995 | 9999 | statistical regression coefficient | Statistical measurement, z-score |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | Slovakia | 2015 | 625,33 | P-values | ML, Multiple regression analysis |
|                       | Individual Enterprise Risk | financial and non-financial index | the whole spectrum | world-wide | 1992-2009 | 9000 | correlation coefficients | Statistical measurement, panel data regression |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | world-wide | 2008-2013 | 539 | robust standard errors | ML, regression model |
|                       | Individual Enterprise Risk | high tech industries | the whole spectrum | worldwide | 2015-2020 | 660 | predict error | ML, Biometric neural network model |
|                       | Individual Enterprise Risk | financial and non-financial index | E-commerce | China | 2016 | 3617 | prediction accuracy | DL, deep learning algorithm with data fusion technology |
|                       | Individual Enterprise Risk | financial index | the whole spectrum | worldwide | 1996-2013 | 170,57 | standard errors | Statistical measurement, Option pricing model, regression model |
understand the bankruptcy risk and lay the foundation for subsequent predictions.

3) Guarantee Risk: Guarantee risk is the possibility that a credit guarantee agency will suffer losses due to various uncertainties in the course of its guarantee business operations. Many SMEs have difficulties in raising finance due to their small size and relatively weak risk resistance and solvency, and a study of guarantee risk assessment and its system will help to enhance the guarantee capacity of enterprises and improve their financing difficulties. This section will mainly introduce some theoretical achievements of guarantee risk and guarantee methods.

Among the theories about guarantee risk, Cowan et al. in [44] found that partial credit guarantees can severely affect the default rate of insured loans, with insured loans 1.67% more likely to be in default after 24 months compared to similar loans without insurance, but this phenomenon hardly occurs for borrowers with large assets of their own, this result suggests that guarantees affect firms’ incentives to repay loans; GROPP et al. in [13] empirically analyzed the impact of public guarantees on bank risk-taking through a natural experiment, which found that banks that break away from government guarantees adopt a safer credit strategy, cutting off loans to the riskiest borrowers and thus reducing credit risk; Wu et al. in [45] studied the feasibility of letter-of-guarantee securitization according to the theory of asset-backed securities, and proposed two important indexes when constructing the return model of letter-of-guarantee securities: the risk probability of asset-backed securities and the loss rate of asset-backed securities, so as to analyze the cash flow of securities. And simulation experiments were also designed to investigate the factors influencing the model’s return. The results of the experiments show that the security has a higher return in boom periods and that an increase in the probability of cash flow risk or a shorter maturity also enhances the return; Jian et al. in [19] analyses the determinants of Chinese firms’ participation in guarantee circles from both macroeconomic and microeconomic perspectives. The results show that in regions with high economic growth but lacking a well-developed banking system and legal protection, firms are more likely to obtain guarantees from companies related to their controlling shareholders.

Depending on the financing needs of the enterprise, the guarantee methods have diversified, with common guarantee methods including equity swap guarantees [46], [47], related party guarantees [20], government guarantees [21], [48], [49] and mutual guarantees. Equity swap guarantee is a new type of credit agreement between a bank, an insurance company and an SME, whereby the insurance company acts as a guarantor and allows the bank to lend to the enterprise, and if the
enterprise defaults on the loan, the insurance company pays the interest and principal, while the enterprise is required to allocate part of its equity to the insurance company as a guarantee cost. Related-party guarantees refer to guarantees that occur between related or indirect related companies, mostly when the controlling shareholder asks the company to provide guarantees for related-party companies out of their own interests. This type of guarantee seriously infringes on the interests of minority shareholders; Government guarantees generally refer to the loan guarantees provided by the government to banks for SMEs, which increase the risk tolerance of banks; Mutual guarantees are a common form of guarantee for SMEs, in which both enterprises use their creditworthiness to guarantee each other’s financing to banks on the premise of mutual assistance, and form a guarantee chain. The characteristic of this guarantee method is that once a debt crisis breaks out in a company on a certain node of the guarantee chain, it will affect the enterprises in the entire guarantee chain, and even cause financial turmoil.

4) **Financial Risk**: Financial risk is a problem that most enterprises must face in financial management. It is objective and impossible to eliminate. Therefore, for enterprise managers and investors, financial risk warnings and management are very important for them to make the right decision. As an important part of enterprise risk management, financial risk has been studied by many scholars. This section will review the existing literature from the aspects of financial risk prediction and assessment, and influencing factors.

With the continuous development of the world economy, the business environment of enterprises is becoming more and more complex. At the same time, the cases of enterprises facing financial difficulties and even bankruptcy are increasing. The financial crisis of enterprises will not only affect the managers and investors but also bring a series of problems. Therefore, enterprise financial risk warning is particularly important for enterprises. And many scholars have studied financial risk early warning. For example, [50] uses convolutional neural networks and financial data to build a financial risk prediction system to help investors and enterprises find possible financial crises in listed companies. [51] proposed an algorithm based on a hybrid PSO-SVM model to provide early warning for enterprise financial risks. [50] and [52] extract the cash flow statement, income statement, balance sheet and other financial data of the listed company, on the basis of which the financial risk warning model is established, and the company with the collected data is taken as the forecast target to predict the financial risk of the company.

In addition, many scholars have studied the factors that affect the financial risk of enterprises. For example, [53] proves that there is a close relationship between the board management process and enterprise financial risk based on quantitative empirical evidence. [54] uses multiple regression analysis to test the mechanism of influence of scientific and technological innovation investment on corporate financial risk. In addition, many scholars have studied the relationship between social performance and financial risk. For example, [55] studies how corporate social responsibility affects the level of financial distress risk, and the research shows that the two are negatively correlated. [56] uses dual risk measurement to test the impact of corporate social performance which is represented by ESG (Environmental, social and governance) assessment on corporate financial risk.

**B. Risk Analysis Granularity**

In this section, we review the previous research on risk analysis in terms of different granularity, i.e., individual enterprise risk, enterprise chain risk, and system risk, which are summarized in Table I.

1) **Individual Enterprise Risk**: Individual enterprises will encounter many risks in the process of operation, and many management activities at the enterprise level will also be affected by various pressures. In the process of trade, multiple enterprises usually cooperate to achieve the best benefits. In this section, we comprehensively classify according to different industries, and we find that the current research on individual enterprises mainly focuses on industry and manufacturing.

For industry, [73] estimates the probability of bankruptcy for 5784 industrial firms, showing better calibration and discriminative power than inferred in the standard Black and Scholes framework and the KVM framework. And [74] presents a new framework for modeling corporate bankruptcy and a method for assessing the vulnerability of industrial economic activity. A method is proposed to analyze the vulnerability of industrial economic activities in various countries and industries. [75] analyzes the traditional statistical methods used for distress classification and prediction (i.e. linear discriminant (LDA) or logit analysis with neural networks (NN) artificial intelligence algorithm). This study analyzes more than 1,000 healthy, fragile and unhealthy Italian industrial companies, where the results show a balance of accuracy and other beneficial characteristics between LDA and NN. In the manufacturing and construction industries, considering the real environment of Spanish firms from the construction industry, [76] proposes a new hybrid approach to predicting firm insolvency. In any competitive economy, the risk of bankruptcy is ubiquitous. [77] aims to improve the predictive ability of corporate bankruptcy and insolvency risks by introducing new processing and verification methods. The article examines the broad application of the Z-score model to predict the economic and financial stability of Romanian manufacturing and extractive firms. In the retail industry, [78] investigates the types and management of risks faced in the supply chain of a large US retailer. Among them, the article divides risks into inherent or high-frequency and interrupted or infrequent risks, and investigates mitigation strategies for these risks. [79] uses a back-propagation neural network (BNN) to conduct a multi-industry investigation of Korean corporate bankruptcy, including construction, retail, and manufacturing. This study proposes an industry-specific bankruptcy prediction model by selecting appropriate independent variables. For the financial industry, the article focuses more on the banking industry. Using multiple financial ratios, [80] applies multiple methods (neural network techniques, support vector machines, and multivariate statistics) to the problem of bank failure prediction.
**TABLE II: Literature comparison in terms of risk analysis granularity.**

| Risk Analysis Aspect | Literature Type | Industry | Country | Period | Size | Metric | Category | Methodology |
|----------------------|-----------------|----------|---------|--------|------|--------|----------|-------------|
| **Intelligence** | Credit Risk | financial index | industry | the U.S. | 1971-2001 | 28,000 | statistical | Statistical measurement | maximum likelihood estimation |
| | Bankruptcy Risk | financial index | industry | the U.S. | 1985-2003 | 4000 bonds | correlation | Statistical measurement | The Percentage Zeros and the LOT Model |
| | Supply Chain Risk | financial index | various | Tailemba | 1992-2009 | 500 companies | statistical | Statistical measurement | C-index & LDA |
| | Enterprise Risk | Financial risk | non-financial index | financial markets | the U.S. | 1992-2009 | 500 composite index | statistical | Statistical measurement | LDA, t-test, LR |
| | | Supply Chain Risk | non-financial index | supply chain professionals | - | - | - | - | Statistical measurement | structural equation |
| **Enterprise Chain Risk** | Credit Risk | financial and non-financial index | commerce | Greek | 1986-2005 | over 62,000 observations | non-compensatory character | Statistical measurement | machine learning model | ELECTRE methods |
| | guarantee risk | financial index | bank | Japan | 1986-2002 | 100 banks | statistical | Statistical measurement | machine learning model | machine learning model |
| | Bankruptcy Risk | financial and non-financial index | various | French | 2000-2016 | 14 million firms | accuracy | Statistical measurement | machine learning model | self-organizing neural network |
| | Supply Chain Risk | financial index | various | Taiwan | 2013-2016 | 950 companies | accuracy and interpretability | Statistical measurement | machine learning model | XGBoost, artificial neural network |
| | systemic risk | non-financial index | various | Australia | 1992-2002 | 395 instances | evaluation accuracy | Statistical measurement | machine learning model | Neural Network Model Based on Back Propagation Learning Algorithm |
| **Systemic Risk** | Credit Risk | financial index | commerce | the U.S. | 2000-2005 | 950 firms | statistical | Statistical measurement | TALIS |
| | systemic risk | financial index | commerce | China | 2000-2015 | 86 institutions | - | - | Statistical measurement | differed estimation method |
| | Bankruptcy Risk | financial index | industry | the U.S. | 1992-2001 | 500 firms | coefficient | Statistical measurement | principal component analysis | regression model |
| | Credit Risk | financial index | bank | the U.S. | from 1975 onwards | 500 banks | statistical | Statistical measurement | Multiple regression analysis |
| | Bankruptcy Risk | non-financial index | financial | US, UK, and Canada | 1994-2006 | 1717 firms | statistical | Statistical measurement | CAPMI |
| | Credit Risk | financial index | various | Australia | 2004-2011 | 1770 individual stock prices | coefficient | Statistical measurement | - |
on historical real cases. At the same time, it is also necessary to analyze the risk contagion of the bank and fit the risk propagation process, which has important guiding significance. For example, [81] creates a bilateral banking network model consisting of banks and bank assets, and proposed a cascading failure model to describe the risk propagation process during a crisis and test the financial Systemic risk pressure. This model uses the bank’s balance sheet data to populate the model, which can be used to stress test banks’ systemic risk and other financial systems. In addition, [82] conducts research on all non-financial UK companies, showing that considering different misclassification costs and loan pricing factors, the area under the ROC curve A small difference in the z-core makes an economically large difference in the profitability of users of the credit risk model, with z-core models yielding much higher returns on risk-adjusted revenue, profits, and risk-weighted assets.

2) Enterprise Chain Risk: In a large financial social system, companies tend to collaborate and influence each other. We can call it an enterprise chain if there is a linkage and some kind of transmission relationship between multiple enterprises. Companies at different points play a role in the formation and stability of the enterprise chain (supply, sales, etc.). Financing, investment, production, and operation activities are always happening in the enterprise chain, which are accompanied by profits and risks. This section will provide an introduction to the main types of risk that may be encountered in the above-mentioned financial activities and list the contributions made by some scholars in the field. This section provides an introduction to enterprise risk according to the type of chain, which mainly includes supply chain, security chain, transaction chain and credit chain.

[83] [84] [85] have analysed and studied supply chain risks from different perspectives, [83] finds that while awareness of supply chain risk management has increased, understanding of what exactly supply chain risks mean, what information should be monitored and how to design risk management and mitigation based on these risks is different, and to address this research gap, Heckmann et al. in [83] based on a literature review, identify core characteristics used to define, quantify and model supply chain risks. [84] investigates the impact of disruption risk when a retailer in a supply chain trades with a risky competing supplier that may default during the production cycle and finds that competition between suppliers affects equilibrium wholesale prices. [85] Presents a set of propositions on how firms can manage supply risk in a financial crisis, highlighting how their approach to risk management has shifted and illustrating how they are relevant to enterprise risk management. In the context of guarantee chains, [86] studies the prediction of successive default events for network-guaranteed loans and proposes a temporal default prediction model that integrates SME features from temporal guarantee networks, loan behaviour sequences and circular attention mechanisms. Regarding research on transaction chains, without the need for financial (accounting) data, [12] proposes a SME bankruptcy prediction model using transactional data and payment network-based variables, and both offline and online test results confirmed the predictive power and economic efficiency of the transactional data-based variables. As for credit chains, [87] in which firms are linked through supplier-customer relationships involving trade credit extensions, a simple production network model is proposed.

3) Systemic Risk: Systemic risk refers to the risk that has not been discovered or paid attention to due to various external or internal adverse factors accumulated for a long time. Because systemic risk has a wide impact on the financial system, it is of great significance for financial participants to explain it comprehensively. In this section, we give an overview of systemic risk according to its sources.

Systemic events such as the epidemic of infectious diseases, terrorist attacks, and other emergencies always have a significant impact on the global economy. Overall, the COVID-19 pandemic has increased systemic risk for countries [88]. Yet the outbreak has also affected different industries inconsistently. For example, the pandemic increased activity in the healthcare and technology sectors, which positively impacts revenues [67]. But for tourism-linked industries, movement restrictions have been imposed across the globe because of the highly contagious nature of the coronavirus. This has led to a decrease in the number of people traveling, and the tourism industry, including transportation, hotels, and restaurants, has been severely affected by COVID-19. At the same time, [89] shows that during the COVID-19 pandemic, micro-enterprises may face the direct risk of liquidity crisis if they cannot generate a revenue stream within a few months. In addition to global health and medical events, terrorist events also have an impact on systemic risk. [90] shows that the September 11 attacks significantly impacted a range of airline stocks listed on international stock markets. The occurrence of this incident has led to a considerable increase in systemic risk in the airline industry, which impacted portfolio diversification and the ability of airlines to raise capital.

The implementation of national policies is usually closely related to the economic operation of the whole country. The implementation of interbank lending regulation reduces the risk of commercial banks compared with unregulated non-bank financial institutions [68]. The deregulation reform of the US banking industry has led to an extraordinary increase in entrepreneurship and business failures, with the most significant number of new businesses among the businesses that fail [70]. [91] finds that deregulated banks can better spread credit risk geographically. [92] finds that intrastate bank deregulation increases the local market power of banks and reduces the level of innovation and risk of young private firms. In contrast, interstate banking deregulation reduces the regional market power of banks and increases the level of innovation and risk for young private firms. In addition to regulatory measures, the impact of monetary policy on systemic risk has also received much attention, especially in the real estate sector. A factor-augmented vector automatic regression model is used in [93] to analyze the monetary transmission through private sector balance sheets comprehensively, credit risk spreading and asset markets, showing that expansionary monetary policies and bullish stock markets tend to accelerate the growth of housing prices, while restrictive monetary policies slow down the growth of housing prices. [94] suggests that the impact of
monetary policy on the real estate sector is highly significant and persistent, while the impact on risk spreads in money and mortgage markets is relatively short-term.

As economic entities become increasingly interconnected, shocks in the financial network can trigger major cascading failures throughout the system. [95] introduces DebtRank, a new measure of system impact inspired by feedback centrality, which can be used to identify systemically important nodes in a network. [96] shows that the systemic risk of individual trades is 1,000 times greater than the corresponding credit risk. Financial entities can be linked to each other through a network of different financial contracts, such as credit, derivatives, foreign exchange, and securities. Considering only one aspect would underestimate total systemic risk. [4] proposes an instrumental variable regression approach that can deal with potential endogeneity issues in the analysis of contagion measures as determinants of tail risk. [97] establishes the probability model of contagion among banks by establishing the bow-tie structure of the weighted guide graph describing the structure of interbank loans. The characteristics of the systemic risk related to infectious diseases calculated by this model are consistent with the characteristics of overt stress testing. By establishing a dynamic model that combines credit risk technology with the contagion mechanism of the interbank risk exposure network, [98] reveals the emergence of a robust contagion mechanism, in which a lower default correlation among banks corresponds to a higher loss. [99] reduces systemic risk by increasing the transparency in the network and reducing interbank lending of systemically risky nodes. [100] introduces a new framework that provides an interpretable map of default dependence among institutions, highlighting possible contagion patterns and institutions that may pose a systemic threat. [101] obtains the results of the analysis of the interbank lending network model based on the directly related financial parameters (such as interest rate and leverage ratio) from the specific formula, and also obtains the closed-form formula to calculate the critical value of the number of creditors of each bank for the propagation of a single shock in the network.

C. Enterprise Risk Intelligence

Table [111] presents previous studies on risk analysis in terms of enterprise risk intelligence, i.e., financial index, non-financial textual information, relational data, and intelligence integration. And Figure [4] shows the publication treads of enterprise risk intelligence over the years.

1) Financial Index: Financial indicators refer to relative indicators for enterprises to summarize and evaluate their financial status and operating results. By analyzing the financial indicators of the enterprise, it is possible to identify the potential risk problems of the enterprise while detecting the current status and achievements of the enterprise and provide a reliable basis for the enterprise to carry out risk identification or risk prediction, to achieve the purpose of reducing the risk loss of the enterprise. Due to the accurate forecasting ability of financial indicators, their analysis plays an important role in enterprise risk forecasting. In this section, we start with three classifications of financial indicators, namely, solvency indicators, operating capacity indicators, and profitability indicators, and make a systematic review of the impact of financial indicators on corporate risk.

Solvency refers to the company’s ability to pay current liabilities with current assets, including working capital, quick ratio, asset-liability ratio, etc. It mainly depends on whether the company’s capital structure is reasonable and stable, and the size of the company’s long-term profitability. [113] examines specific bankruptcy determinants using balance sheets. Firm productivity inefficiency, an important ex-ante indicator of firm failure, also adds additional explanatory power to models that include the balance sheet and qualitative variables. At the same time, non-balance sheet items also significantly improved the explanatory power of the predictive bankruptcy model. [116] explains the common risk factors of the cross-section of corporate bond returns; [117] develops a structured bond valuation model using bond price data to capture corporate [118] provides a new method for pricing and hedging derivative securities involving credit risk, which can be applied to corporate debt and OTC derivatives. [119] finds a term structure using risk-neutral implied probabilities of default obtained from the market prices of a group of bonds of the same issuer. For the quick ratio, [120] uses the "traditional" factors such as the current ratio and quick ratio of SMEs to predict the credit risk of SMEs in supply chain finance and proposed an enhanced hybrid integrated ML method to improve The accuracy of predicting the credit risk of SMEs while improving their financing ability. Both [121] and [122] study the relationship between credit default swap (CDS) data and firm risk. Using CDS spread data, [123] finds that systemic risk among U.S. sovereigns is much lower than among euro area sovereigns using a multifactor affine framework that allows for systemic and sovereign credit shocks, and that systemic sovereign Risks are closely related to financial market variables. Operating capacity refers to the operating capacity of an enterprise, that is, the ability of an enterprise to use various assets to earn profits. The financial analysis ratios of enterprise operating ability include: inventory turnover rate, accounts receivable turnover rate, business cycle, current asset turnover rate, and total asset turnover rate, etc. In the study of asset liquidity, [124] develops a dynamic bankruptcy model to reduce the risk of finding an acquirer before bankruptcy in the case of insufficient asset liquidity, but we found that this asset sale policy cannot Fully be offsetting the impact of lower asset liquidity, which increases the probability of default and reduces equity, debt, and firm value, may lead to a more positive stock price response at lower asset liquidity. By studying the typical financial characteristics of Japanese companies when they went bankrupt, [125] finds that the capital flow of one type of failed company had a significant decline. By using the ratio and absolute amount based on the cash-based financial statement data of the three years before the failure as predictive variables, A corporate bankruptcy prediction model with higher prediction accuracy is constructed. Profitability refers to the ability of an enterprise to earn profits. Relevant indicators include gross sales rate, net sales rate, net asset rate, and return on equity. [126]
## TABLE III: Literature comparison in terms of enterprise risk intelligence.

| Enterprise Risk Intelligence | Literature | Type | Aspect | Period | Size | Metric | Methodology | Category | Methodology |
|-----------------------------|------------|------|--------|--------|------|--------|-------------|----------|-------------|
| Financial Index             | Bankrupt Risk | Individual Enterprise Risk | 1995-2001 | 592 | - | ML | logit model based on CART |
|                             | Systematic Risk | Systemic Risk | 2004-2015 | 830000 | Accuracy | DL | GNN |
| Non-financial Textual Information | Credit Risk | Individual Enterprise Risk | 1980-2004 | 2770 | Delta-based standard errors | ML | Supervise the LDA |
|                             | Systemic Risk | Systemic Risk | 2002-2016 | 132060 | The ROC curve | ML | Random forest regression tree |
| Relational Data             | Systemic Risk | Systemic Risk | 2009 | 870 | Finance Metric | ML | Econometric model |
|                             | Financial Risk | Individual Enterprise Risk | 2007 | 250 | AUC | ML | Deep Risk method for neural networks |
| Intelligence Integration    | Credit Risk | Individual Enterprise Risk | 2014-2018 | 83 | AUC | DL | Deep Risk method for neural networks |
|                             | Bankrupt Risk | Individual Enterprise Risk | 2006-2008 | 165 | STRONG ERM | ML | Discrete hazard model |
|                             | Guarantee Risk | Enterprise Chain Risk | 1980-2009 | 169566 | AIC measures | ML | SVM |
|                             | Systemic Risk | Systemic Risk | 2002 | 43 | - | ML | Randomized null-model |
|                             | Financial Risk | Individual Enterprise Risk | 2017-2018 | 2935 | Statistical significance | ML | LR |
|                             | Bankrupt Risk | Individual Enterprise Risk | 2011-2014 | 2.4 million | AUC | ML | Linear kernel training SVM |
|                             | Supply Chain Risk | Enterprise Chain Risk | 2011-2018 | 423764 | TVOL | ML | Ranking |
|                             | Credit Risk | Enterprise Chain Risk | 2014 | 47 million | Accuracy | ML | Classification trees |
|                             | Bankruptcy Risk | Individual Enterprise Risk | 2000-2021 | 11523 | Precision | DL | GNN |
|                             | Financial Risk | Individual Enterprise Risk | - | 1 million | AUC | DL | Spatiotemporal perception map nerve |
|                             | Credit Risk | Individual Enterprise Risk | 2015-2017 | 10 | AUC | ML | LR |
|                             | Bankruptcy Risk | Individual Enterprise Risk | 1989-1997 | 300 | Z-value | ML | The logical model |
|                             | Financial Risk | Individual Enterprise Risk | 2001-2018 | 641667 | Accuracy | ML | BP Neural Network |
|                             | Credit Risk | Enterprise Chain Risk | 2011-2019 | 924 | Average accuracy | ML | Adaptive heterogenous multiview graph learning method |
|                             | Systemic Risk | Systemic Risk | 2020 | 1584 | Statistical significance | ML | DCC-GARCH model |
combines the company’s profit rate and other data to calculate the company’s profit rate and supply and demand stability, so as to establish a logistic regression model to help banks predict credit enterprises and then determine whether the bank will lend, loan amount, interest rate and loan period, etc., to obtain more effective credit strategy. Using interest rate data, [127] proposes a dynamic multi-agent model of the banking system with a central bank, showing that central bank intervention can alleviate financial distress and liquidity shortages in the interbank market, but in the long run, central bank activities have little effect on liquidity supply expectations. The impact will be smaller.

2) Non-financial Information:

Non-financial variables: Non-financial variables is the measurement of business performance using metrics that are not related to a business’s finances. Generally, non-financial variables are divided into three categories: operations, customers, and employees.

Customer-related non-financial indicators include customer turnover rate, personal rating, etc. By using balance sheets and other information, such as location, industry, year of establishment, ownership, [128] constructs an operational model of credit assessment for innovative SMEs, which can be a useful tool for banks to innovate their lending processes. Based on the combination of issuer rating, bank default probability composite index and individual rating, the default risk of state-owned banks is lower than that of private banks, but the operational risk is higher than that of private banks, indicating that the existence of government protection leads to higher risk taking [129]. By analyzing non-financial information such as industry prospect risk, enterprise operation risk, management quality and enterprise’s influence on upstream and downstream, [130] analyzes the relationship between annual loan interest rate and customer churn rate, establishes a multi-objective programming model aiming at minimizing customer churn rate and maximizing bank profitability, and determines the bank’s loan strategy.

The most commonly used employee-related variables in the previous literature is the board structure. The board structure index is a commonly used index to reflect the size of the board of directors, institutional setting and the proportion of independent directors. Using data such as director turnover rate, board size (such as number of shareholders, number of subsidiaries and number of independent directors) and CEO duality, [131] [132] [133] studies the impact of these data characteristics on corporate credit risk, and [134] studies the relationship between these indicators and bankruptcy of private enterprises. [135] [136] also use the gender ratio of the board of directors to study the relationship between the proportion of female directors and corporate risk. Using data such as the number of directors in common, [137] finds that directors overlap is the relevant channel of bankruptcy contagion effect.

Non-financial variables related to company operation mainly include business cycle and macroeconomic variables such as GDP. Business cycle is the business rhythm under the condition of mature market economy. [138] finds that when cash flow is more dependent on current economic conditions, firms can adjust their default and financing policies according to economic conditions at the stage of the business cycle. [64] considers the business cycle and models the specific history of the enterprise using self-organizing neural networks and data spatial segmentation. A method is developed in [139] to incorporate longitudinal models. They use both financial variables (such as SMEs’ ability to receive non-bank financing, degree of financial leverage and return on assets) and qualitative predictors (such as payment history, firm life cycle stage and customer order profile). As qualitative information, they take into account unstructured data from experts’ opinions about payment history, stages of the company’s life cycle (development, stagnation or decline) and the status of customer orders (good or decline). Macroeconomic variables can also provide important information when studying firm risk. Using the balance sheet, S&P 500 index, international government bond index and other information, [140] finds that the measurement and management of operational risk is highly relevant to insurance companies and should be included in the enterprise risk management framework. By studying financial factors such as changes in equity returns, volatility of equity returns, and macroeconomic variables such as GDP to predict default rates, [141] finds that the corporate bond market suffered from clusters of repeat defaults much worse than during the Great Depression. [58] shows that the absolute level of risk can also be considered by incorporating firm-specific characteristics and macroeconomic conditions into the model for predicting default risk. Using non-financial variables such as energy price risk, resource risk, inflation risk, and retroactive feed-in tariff reduction, [142] shows that unexpectedly low inflation increases real debt and default risk.

Textual Information: By employing news coverage and news sentiment to quantify textual information in news articles and quantify qualitative risk disclosures by individual companies in their corporate filings, namely Forms 10-K and 10-Q, [143] finds that more news coverage and negative news sentiment increased credit risk. By analyzing a large amount of text-based financial market data, [144] finds that changes in the emotional content of market narratives are highly correlated among data sources. These data show changes in sentiment before and after the global financial crisis, and the indicators in the paper have predictive ability over other commonly used sentiment and volatility indicators. [145] comprehensively excavate bank risk factors from qualitative textual risk disclosures reported in financial statements through a new semi-supervised text mining method proposed. In addition to the analysis of the text, the company’s legal documents also tell us valuable information. Information relating to legal actions taken by creditors to recover outstanding debts, company filing history, consolidated audit report/opinion data and firm-specific characteristics have contributed significantly to improving the default prediction ability of risk models developed specifically for SMEs [146]. [7] uses information from legal judgments involving firms and their principles, combining it with financial information, to propose a framework to identify legal judgments that effectively predict credit risk, and empirical evaluations show that features extracted from valid legal judgments significantly improve the discrimination performance and grant performance of the model. In addition,
conference calls are an important source of information when analyzing corporate risk. [104] uses a machine learning approach to create a composite measure of credit risk based on qualitative information disclosed in the conference call and 10-K Management Discussion and Analysis section. This criterion accounts for intra-firm changes in future credit events relative to previous risk indicators, and the indicators developed in this paper improve the ability to predict credit events (bankruptcies, spreads, and credit rating downgrades). Using interviews with managers directly involved in global sourcing decisions, supported by documentary evidence, [147] presents a new classification of global sourcing risk and provides characteristics of global sourcing risk mitigation strategies applicable to different industries. In China’s coastal industrial zones, local suppliers face serious labor mobility problems. In the form of a questionnaire survey, [148] combines with quantitative techniques to determine the reasons for workers to leave export factories in China, in an attempt to find out the root cause of job dissatisfaction leading to turnover, and provide management enlightenment that may help managers to cope with labor-related supply chain risks.

3) Relational Data: Relational data is data that defines relationships between entities [13]. Relational data is widely used in the study of the relationship between individual enterprises or enterprises, such as equity relationship, cooperation relationship, supply chain relationship, management relationship, etc. In this section, we will systematically review the application of relational data from both enterprise and supply chain perspectives.

We find that in previous studies, relational data are usually used to study the relationship network among SMEs and enterprise cooperation relationship. For example, in [13] high-dimensional data of company directors and managers are used to establish the network between small and medium-sized enterprises, and the relationship between enterprises is converted into bankruptcy prediction score, and the bankruptcy correlation between two enterprises is studied. In [112], an innovative financial risk analysis framework is proposed and the effectiveness of the recommendations is verified by using the relationship between SEMs and financial data. [111] constructs the payment network of enterprises through the transaction data between enterprises, and empirically studies the interaction and risk distribution between enterprises. Through the payment network between enterprises, it highlights the complex interdependence between enterprises. [149] propose a risk identification and assessment system for sustainable interoperability enterprise collaboration based on fuzzy logic. Enterprise partnership risk is assessed by analyzing historical data sets.

Relational data is also applied to the study of supply chain relationships. For example, [150] uses a questionnaire survey conducted by 800 SEMs as an experimental data set to determine the basic dimensions of supply chain management (SCM) practices and empirically test a framework to determine the relationship among SCM practices, operational performance and SCM related organizational performance. [110] constructs a multi-tier supply network using supplier and customer data from the FactSet Revere business relationship database. It examines the influence relationships between primary and secondary suppliers. The results show the importance of considering the influence of the subdivided supply network structure in the portfolio optimization process.

4) Intelligence Integration: In the assessment of corporate risk, different sources of information have been taken into consideration. Generally, financial data have been combined with non-financial data to obtain comprehensive information. Non-financial information, like firm size, corporate governance indexes, audit opinions, has already been demonstrated to be useful in risk prediction [151] [152] [153]. For example, social responsibility data and stock return data are captured to investigate the relationship between corporate social responsibility and firm risk [154]. What’s more, much credit information like credit history, credit loan records has been widely used especially in credit risk [62] [7]. On some occasions, personal information is combined with financial performance to study the impact of directors’ features on corporate risks [40] [135]. Besides, combination of relational data and financial data is another big tendency in risk management. Cleofas [155] combines accounting statements with business information, bank customer data and credit applicant information to predict the bankruptcy. Such data depicts relationship between company and linked enterprises, or relationship between company and banks [156] [157]. Integration also happens among the three types of data mentioned above. For instance, Zhao et al. [158] utilize transaction data, web news as well as company relations and executives’ information to predict the stock movement. Thus, we believe that, in future research, more and more integration of data will be explored and bring about great improvements in the study of enterprise risks.

D. Enterprise Risk Evaluation

1) Datasets: SMEsD. The SMEsD consists of 4229 SMEs and related persons in China from 2014 to 2021, which constitutes a multiplex enterprise knowledge graph. All enterprises are associated with their basic business information and lawsuit events spanning from 2000 to 2021. The enterprise business information includes registered capital, paid-in capital and established time. Each lawsuit consists of the associated plaintiff, defendant, subjects, court level, result and timestamp. The company’s short and long-term debt. The sample in [57] is for the period 1980-2004, with short-term and long-term debt data from Compustat’s quarterly and annual files. This dataset is used to evaluate models of US industrial companies. Greek Commercial Bank Loan Portfolio. Data in [159] from the loan portfolio of the Commercial Bank of Greece, with 1,411 companies. Included in two datasets. The first, provided by the Bank, includes financial data on 200 companies for the period 1994-1997. Based on the latest available information on these companies (1997), the Bank’s credit officers classified half of them as high-credit-risk companies. The remaining 100 companies in the training sample were assessed as low-credit-risk firms. The second data sample (the holdout sample) consisted of 1211 companies, divided into two groups with the training sample. This sample was used to validate the credit risk assessment model to assess its ability to generalize data on
| Datasets | Tasks | Descriptions | period | size |
|----------|-------|--------------|--------|------|
| Listed SEMs | Credit Risk Prediction | The listed SEMs dataset of Sm 500 companies. | 2014-2015 | 53 |
| SMEsD | Bankruptcy Risk Prediction | This dataset is a benchmark dataset constructed by collecting SEM real data from multiple sources. | 2014-2020 | 4229 |
| The company’s short and long-term debt | Multi-period corporate default forecast | Moody’s Default Risk Service and the CRSP/Compustat database. | 1980-2004 | 2770 |
| Greek Commercial Bank Loan Portfolio | Credit risk assessment | Financial data provided by the Bank and Sample holdout. | 1994-1997 | 1411 |
| Bankruptcy Financial Data | Estimated retail banking portfolio rates | Accounting data of Norwegian Limited liability company. | 1995-1999 | 15401 |
| Interview data | Supply Chain Risk Management | Interview record of purchasing supervisor or purchasing manager. | 2009-2010 | - |
| Daily quotes for Euro-denominated bonds and CDS | Two basic credit risk instruments for accurate pricing | Daily quotes for Euro-denominated bonds and CDS for various corporate borrowers. | 2003-2005 | 730 |
| American Property Insurance Company rating | Enterprise risk management | Standard & Poor’s ERM Quality Ratings available on all U.S. property insurance company data. | 2006-2013 | 76 |
| Standard & Poor’s rating data | Enterprise risk management | Annual observations of banks and insurance companies in S&P’s Direct rating database. | 2006-2008 | 165 |
| Data of listed insurance companies | The influence of EMR method on capital cost of enterprises | Data of listed insurance companies from 1996 to 2012 are mostly included in the combined CRSP/Compustat database | 1996-2012 | 761 |
| Listed SEMs data | compare the effect of ML method in credit risk prediction of small and medium-sized enterprises | Data of 46 smes listed on the SME Board of Shenzhen Stock Exchange and 7 smes listed on the main board of Shanghai Stock Exchange and Shenzhen Stock Exchange between March 31, 2014 and December 31, 2015. | 2014-2015 | 53 |
companies different from those developed for the model and its classification performance. This sample consists of 1093 firms with low credit risk and 118 firms with high credit risk.

**Bankruptcy Financial Data.** The accounting data in [160] includes all Norwegian limited liability companies for the period 1995-1999. Approximately 50 variables are registered from the accounting reports each year, providing a large database for testing how various financial data affect default rates. In addition to the accounting data, there is a “fixed” set of data (such as organization number, name of company, industry code, starting data of firm, geographical location, and number of employees) describing each company that existed in the period 1995-1999, providing the possibility to test the relevance of static variables for insolvency.

**Interview data.** Data in [165] are collected from 2009-2010 through interviews with at least one senior purchasing executive and up to two purchasing managers in charge of SCRM. Detailed notes were taken during the interview. Other materials such as slides and other documents (e.g., supplier self-assessment questionnaire, follow-up supplier development Action plan) were also used for the triangulation. A case study database was developed on the NVivo platform to store and organize relevant information and supplementary materials.

**Daily quotes for Euro-denominated bonds and CDS.** The dataset in [161], which includes daily quotes from 2003 to 2005 on euro-denominated bonds and CDS on senior unsecured plain coupon bonds and senior unsecured debt, is used to analyze the extent to which standard deterministic reduced form models can price both bonds and CDS.

**US Property and Casualty Insurance Company Ratings.** The sample in [162] included all U.S. property insurers for which S&P ERM quality ratings were available between 2006 and 2013. After limiting the sample to groups and non-affiliated companies with positive premiums and ERM quality ratings, the final sample consisted of 76 unique companies (66 groups and 10 non-affiliated companies; 51 publicly traded companies and 25 private companies) and a total of 467 company-year observations. All listed companies were insurance groups.

**Financial services companies.** The sample in [107] consists of a dataset of banks and insurance companies covered in S&P’s Direct rating database. It contains annual observations for 165 companies. After collecting sample data, they used the data to investigate ERM program quality.

**Listed Insurance Companies.** The sample in [163] includes all publicly traded insurers in the CRSP/Compustat consolidated database from 1996 to 2012. This dataset is used to verify whether enterprise risk management will reduce the cost of capital for enterprises.

**Listed SMEs.** The sample in [164] select from 46 SEMs listed on the Small and Medium-sized Board of the Shenzhen Stock Exchange and 7 SMEs listed on the Shanghai Stock Exchange from 2014 to 2015 to compare the performance of the ML method in the credit risk prediction of SMEs in supply chain finance.

**Loan Level Information.** [165] Data by LPS Applied Analytics, Inc. Provides, including loan-level information collected from residential mortgage servicers. As of July 2008, the data includes loans from nine of the top 10 mortgage servicers, representing about two-thirds of the U.S. mortgage market, or more than 39 million active mortgages. The dataset is used to analyze the factors that lenders choose to securitize loans.

**China Bond Yield and Rating Data.** [166] The sample period used was from 2009 to 2015. The final sample comprised 6,528 bonds. Of these, These include 1,560 bonds with floating coupons and 4,968 with fixed coupons.

2) **Evaluation Metrics:** The evaluation index and evaluation system of the model is an important link in the modeling process. For different types of projects and models, different evaluation index and system should be reasonably selected. In this section, we selected eight evaluation indicators for a detailed introduction.

**Kolmogorov-Smirnov statistic (KS):** The K-S test is a goodness-of-fit test for exploring the distribution of continuous random variables, using sample data to infer whether the overall population from which the sample is drawn follows a certain theoretical distribution. Chang et al. in [7] performed 10 independent 10-fold cross-validations against KS in order to predict the recognition performance of the model and obtained 100 performance estimates.

**The area under the receiver operating characteristic curve (AUC):** The AUC curve is the area under the ROC curve and enclosed by the axes. Represents the degree or measure of separability, which is an index to measure the quality of a learner. And as a value, the classifier corresponding to a larger AUC is better. Chang et al. in [7] performs 10 independent 10-fold cross-checks against the AUC and
obtained 100 performance estimators to predict the recognition performance of the model. In order to develop a two-step method to evaluate the classification algorithm of financial risk prediction, three multi-criteria decision making (MCDM) methods are introduced in [167], and AUC is taken as one of the evaluation indexes of the classifier to rank the classifier. In order to test the prediction effect of TSAIB_RS method based on two-stage adaptive integration of multivariate heterogeneous data, Huang et al. in [168] takes China’s microcredit default risk as the test object and takes AUC as one of the evaluation indexes to compare the prediction effect of various test methods. The results show that TSAIB_RS method can significantly improve the prediction effect. In [152], AUC is also taken as one of the evaluation indexes of the model.

**G-measure and F-measure:** The G-measure is the geometric mean of Recall and Precision. It is of high reference value when the data is unbalanced. F-Measure is the weighted harmonic average of Precision and Recall, which is usually used to evaluate the quality of classification models. Sun et al. conducted 100 iterative experiments on the 6 models in [169], take F-measure as the evaluation index and get the most suitable model to deal with the enterprise credit evaluation data with grade imbalance. [170] evaluates the two IEML(RS-boosting and multi-boosting) on 377 data sets by taking F-measure as the evaluation standard, which shows that the IEML method is better than the integrated machine learning method and individual machine learning method in predicting the credit risk of small and medium-sized enterprises. [152] also take F-measure as one of the evaluation indicators to evaluate the model.

**MCDM-based approach for clustering algorithms evaluation:** [171] propose a new evaluation method. The MCDM method is used to evaluate the quality of clustering algorithms in the field of financial risk analysis. In their experiments, six clustering algorithms, 11 performance metrics, three MCDM methods and three financial risk datasets are analyzed, providing a useful tool to evaluate clustering algorithms based on a combination of valid metrics.

**Accuracy:** Accuracy is the ratio of the number of correctly classified samples to the total number of samples, which is often used as an evaluation index of machine learning models. For example, [170] compared the average accuracy of the IEML method (i.e., multi-boosting and RS-boosting) with three other EML methods (i.e., bagging, boosting and RS) and one IML method (i.e., DT). [172] takes accuracy as an evaluation index and verifies that machine learning technology greatly improves the accuracy of bankruptcy prediction. By comparing the average prediction accuracy of 21 different models, [173] find that combining self-organizing mapping (SOMs) with MLP classifier integration works best.

**Type I error, and Type II error:** Type I error, rejecting H0(null hypothesis), actually holds, that is, rejecting the null hypothesis when it is true. The second type of error is to accept a H0 that is actually invalid, that is, to accept the null hypothesis if it is not true. For example, aiming at the problem of small sample size in SME credit risk prediction, [113] proposes an adaptive heterogeneous multi-view learning method. In order to evaluate the effectiveness of the proposed method, type I errors and type II errors are adopted as evaluation indexes than the other two criteria. Finally, it is found that type II error is more important to establish a stable supply chain finance system. By comparing the average prediction accuracy and Type I & II errors of 21 different models, [173] find that combining self-organizing mapping (SOMs) with MLP classifier integration works best.

**Recall Rate:** The recall rate is the percentage of a sample in which positive cases are correctly predicted. Usually used to evaluate the detector to cover all the detected targets. For example, [168] takes the recall rate as one of the evaluation indexes of the model and tests the reliability of the TSAIB_RS method for the adaptive integration of multi-source heterogeneous data in the two-stage adaptive integration. Cheng et al. in [174] proposes a risk assurance relationship predictive attention neural network based on dynamic graphs, and takes recall rate as an evaluation index to verify that the model can improve the accuracy of risk prediction.

**Kappa measure:** The Kappa coefficient is an index to measure classification accuracy based on confusion matrix calculation. It is a good evaluation index to measure the classification effect of data with unbalanced samples. The higher the Kappa coefficient is, the better the prediction effect will be. In [152], the Kappa coefficient is used as the main model evaluation index to evaluate the model.

### III. Methodologies

Table [VI] presents the previous studies on risk analysis in terms of analysis models, i.e., statistical model, machine learning model, and deep learning model. In [5] we plot the distribution of representative works of some major models on the timeline. And Figure [6] shows the publication trends of enterprise risk intelligence over the years.

#### A. Statistical Econometric Methods

The study of enterprise risk originated from the traditional economic disciplines. Before computer science technology was widely used to quantify enterprise risk, scholars used a large number of traditional economic and statistical models to study this problem. After sorting out and screening the literature, we found that the traditional models used by scholars are rich in variety and quantity. In this section, we will briefly introduce Regression Model, Z-score Model, CAPM Model, Merton Model, KMV Model, Complex Network, Structural Equation Model and Factor Model. Table [VII] shows the core formulas of the main statistical models in this section, and Table [VIII] explains the meaning of the parameters of these models.

1) **Regression Model:** The regression model is very significant in economic research, and it can be used to verify the relationship between variables and predictors. [39] uses logistic regression to test the proposed hypothesis. The key variables are board composition, CEO-board chairperson structure, and composition-structure interaction. Control variables include financial indicators, the number of shares held by each entity, and the quality of the board of directors. Finally, the author uses the “likelihood ratio approach” to test, verifying that
**TABLE V: Evaluation index calculation formula**

| Metric     | Formula                                                                 | Parameter                                                                 | Explanation                                                                 | Reference |
|------------|-------------------------------------------------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------------|-----------|
| KS         | $D_n = \sup_{x} |F_n(x) - F(x)|$                                                                           | $F_n(x)$                                                                   | Cumulative distribution function |            |
|            |                                                                         | $F(x)$                                                                   | Theoretical distribution                                                  | [7]       |
|            |                                                                         | $\sup_{x}$                                                                | The upper bound of distance                                               |           |
| AUC        | $AUC = \sum_{i \in \text{positiveClass}} rank_i - \frac{M+M_1}{M \times N}$ | $M$                                                                       | The number of positive samples T                                           | [7] 167   |
|            |                                                                         | $M_1$                                                                     | The number of negative samples                                            | 168 152   |
|            |                                                                         | $N$                                                                       | Just add up the numbers of the positive samples                           | 13        |
| F-measure  | $F_{\beta} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta \times \text{Precision} + \text{Recall}}$ | $rank_i$                                                                  | Article i Serial number of the sample                                      |           |
|            |                                                                         |                                                                          | Based on a lot of work experience doing salient object detection          | [169] 170 152 |
| Accuracy   | $A = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$ | $\text{TP}$                                                               | True Positive                                                              |           |
|            |                                                                         | $\text{FP}$                                                              | False Positive                                                            | [170] 172 173 |
|            |                                                                         | $\text{FN}$                                                              | False Negative                                                            |           |
| Recall Rate| $rec = \frac{\text{TP}}{\text{TP} + \text{FN}}$                         |                                                                          |                                                                           |           |
| Kappa measure | $k = \frac{p_0 - p_e}{\sqrt{p_0 \times p_e}}$                     | $p_0$                                                                     | Overall classification accuracy                                           | [152]     |
|            |                                                                         | $P_e$                                                                     | $P_e = \frac{\sum_{i=1}^{n} a_i \times b_i}{n^2}$                          |           |
|            |                                                                         | $a_i$                                                                     | True sample number                                                        |           |
|            |                                                                         | $b_i$                                                                     | The number of predicted samples in each category                          |           |

![Fig. 5: Chronological overview of methods to studying enterprise risk. Methods in red, green and blue are statistical econometric methods, machine learning methods and deep learning methods, respectively.](image)

Financial indicators, board quality, and corporate governance structure are all important factors that can predict corporate bankruptcy. [180] uses the ordinal logistic regression model to verify that the execution of enterprise risk management requires the leadership of senior managers. The author focuses on the impact of the existence of a chief risk officer, independence of the board of directors, management’s attitude towards enterprise risk management, type of auditor, size of the company, industry, and country to which it belongs on the implementation of enterprise risk management, and conducts sensitivity tests. [7] uses the company’s financial data and legal judgment information to solve the problem of information asymmetry in the credit application of the company. The author uses the logistic regression model to verify that when the legal judgment amount exceeds 12.5% of the company’s annual income, the default probability of the company will be greatly improved. [116] uses panel regression to show that bond excess returns have significant statistical significance with key explanatory variables such as market downside risk, credit risk, and bond-level illiquidity. The authors repeated the Fama-MacBeth regression using orthogonalized risk characteristics for robustness. The results show that the new risk factor is better than other models in explaining future corporate bond returns. [198] uses the probit model to calculate the possibility...
| Category         | Literature Type | Aspect                          | Focus                          | Experiment | Methodology |
|------------------|----------------|---------------------------------|-------------------------------|------------|-------------|
|                  |                |                                 | Intelligence | Industry | Industry | Period | Size | Metric |
| Statistical Econometric Methods | Guarantee Risk | Individual Enterprise Risk | financial statements | Germany | bank | 1995-2006 | 1329 | significance | Z-Score |
|                  | Credit Risk    | Individual Enterprise Risk | stock market/financial statements | China | / | 2004-2006 | 80 | expected default frequency | KMV |
|                  | Credit Risk    | Systemic Risk | CDS Spreads | U.S | / | 2006-2012 | 100 | Risk probability | copula |
|                  | Credit Risk    | Systemic Risk | Financial ratio, etc | 64 countries | bank | during the COVID-19 | 1584 | CoVaR | CoVaR |
|                  | Bankruptcy Risk | Individual Enterprise Risk | Balance sheet/questionnaire | Italian | / | 1989-1997 | / | significance | logit model |
|                  | Stock Movement Risk | Individual Enterprise Risk | financial data, etc | China | / | from 2007 listed companies | / | significance | regression |
| Machine Learning Methods | Bankruptcy Risk | Enterprise Chain Risk | Financial ratios/Executive information, etc | Belgian and UK | / | 2011-2014 | 400000 | AUC | SVM |
|                  | Bankruptcy Risk | Individual Enterprise Risk | financial statements | / | / | 1975-1982 | 129 | Accuracy | neural network |
|                  | Financial Risk | Systemic Risk | stock market data | U.S | 51 U.S. industries | 1998-2010 | / | Absorption Rate | principal component analysis |
|                  | Bankruptcy Risk | Systemic Risk | Financial ratios | / | 21 industries | / | 66,000 | accuracy | random forest |
|                  | Bankruptcy Risk | Individual Enterprise Risk | Financial ratios | Turkish | / | 2005-2011 | 27.22 | Accuracy | Decision tree |
|                  | Bankruptcy Risk | Individual Enterprise Risk | Financial ratios | / | / | 2040 | accuracy/AUC | PTVPSO-FKNN |
| Deep Learning Methods | Bankruptcy Risk | Individual Enterprise Risk | financial statements | Japanese | / | 2002-2016 | 2164 | AUC,etc | CNN |
|                  | Guarantee Risk | Systemic Risk | Loan data | China | / | / | / | AUC | graph neural network |
|                  | Bankruptcy Risk | Individual Enterprise Risk | Financial data, etc | North America | / | 2001-2018 | 644,667 | accuracy | Deep neural network |
|                  | Credit Risk    | Enterprise Chain Risk | Financial ratios | Polish | / | 2002-2013 | / | AUC | BSM-SEAS |
|                  | Credit Risk    | Enterprise Chain Risk | Corporate relations, etc | / | / | / | / | F1/AUC/Cetc | Deep neural network |
| Hybrid approach  | Credit Risk    | Individual Enterprise Risk | Securities market data/audit rating | China(Taiwan Province) | / | 2003-2008 | 2470 | significant | KMVRF/RST |
|                  | Bankruptcy Risk | Individual Enterprise Risk | financial data/Stock data | China(Taiwan Province) | IT industry | 1999-2006 | 321 | Accuracy | CART-CBB/KST-GRAC-CBR/RST-CBB-GRAC-GAN |
|                  | Credit Risk    | Enterprise Chain Risk | Market share/profit, etc | / | / | / | / | multiple performance measures | BNN/PTA |
of enterprises adopting ERM, and then calculates the Inverse Mill ratio as an explanatory variable. The author uses three regression models to test three hypotheses, including that the implementation of ERM by enterprises can reduce enterprise risk costs and increase changes in corporate profits and lags risk change. The authors then use a subsample for robustness testing, which shows that firms using ERM are more profitable, less risky, and less volatile in stock returns.

2) Z-score Model: Z-score is a very classic bankruptcy prediction model. This method through the financial statements calculates a set of financial ratios that reflect the extent of financial crisis of the company, and then measures the impact of the financial crisis according to these ratios to obtain a comprehensive calculation named risk score, aka Z-score. The independent variables in the Z-score model represent the firm’s asset size, liquidity, profitability, financial structure, debt servicing capacity and asset utilization. When the Z-value is less than 1.8, there is a high probability that the company will go bankrupt.

Since Altman created the Z-score model, it has been widely applied to bankruptcy forecasting. [199] combines the z-score model with other forecasting models to predict the bankruptcy risk of Japanese listed companies. [200] uses the Z-score model to measure the size of the bank’s risk-taking, and finally verifies whether the German government guarantee has a causal relationship with the bank’s risk-taking. [17] uses
the Z-score model and other models to measure a company’s bankruptcy risk, with the goal of studying the relationship between bankruptcy risk and return on investment. [82] compares the Z-score model with the contingent-claims valuation approach in bankruptcy forecasting, and the conclusion proves that the two have little difference in predictive ability, in addition, the Z-score model has more advantages under certain conditions.

3) **CAPM Model:** The Capital Asset Pricing Pricing Model is a very classic model proposed by Sharpe in 1964 to study the relationship between expected returns and risks in the securities market. Although the assumptions of this model are relatively harsh, it is established in a highly idealized securities market. However, its high-risk, high-reward conclusion dominates modern financial theory. [154] uses the CAPM model to measure the risk-return of the target company, and then judge the relationship between the company’s systemic risk in the securities market and corporate social responsibility. In studying how green policies affect the systemic risk of the industry, the authors refer to the CAPM model to estimate the abnormal returns of the underlying stocks. [72].

4) **Merton Model:** Merton model is one of the famous credit risk measurement models proposed by Merton in 1974. It mainly judges the risk of corporate default from the aspect of option pricing. [199] uses the Merton model to detect signs of corporate bankruptcy from option price information. [48] uses Merton proportional dividend option pricing model to value the company’s financial claims, and finally can use this value to judge the quality of corporate loan guarantees. [201] researches methods for assessing the systemic failure of global banks, in which the authors use the Merton model to measure defaults in the securities market, and then estimate the default status of banks. [179] uses the Merton model to gauge the default risk of individual companies and to analyze the effect of default risk on stock returns. [202] uses the Merton model to simulate the default state of the company, and the author derives a formula for failure beta from the model. The results show that the joint default probability of the company is related to the failure beta.

5) **KMV Model:** The KMV model is a method of estimating the default probability of a borrowing company. The model is based on Merton’s option pricing theory. It mainly uses the data of the stock market to consider the company’s loan repayment problem, so that can obtain the company’s actual default probability. However, the model still has harsh assumptions. [75] uses the KMV model to estimate the bankruptcy probability of an enterprise, but because the KMV model has many restrictive assumptions, the author proposes a better performance prediction model on this basis, and uses the results of the KMV model as a comparison item. [181] adjusts the parameters based on the KMV model to improve the validity of the model, and the results show that the adjusted KMV model is very well suited to predict credit defaults of listed SMEs in China. In order to study the relationship between technological competition and enterprise bankruptcy, [203] used three bankruptcy risk measurement models including KMV model as control variables.

6) **Complex Network:** Complex network is a network structure consisting of a large number of nodes and complex relationships between nodes. Many scholars use complex networks to model financial connections in the real world, and then analyze risk transmission or systemic risks in financial networks. [95] developed an application called DebitRank based on a complex network approach to analyze changes in systemic risk caused by the Federal Reserve’s emergency lending from 2008 to 2010. [204] used the complex network method to study the linkage relationship between the stock index returns of the United States and other countries, with the purpose of analyzing the contagion of international financial risks. [205] proposes a SICM (susceptible agents, infected agents, contagious agents, and immune agents.) model based on complex networks to talk about the contagion mechanism of financial risk, which sheds new light on the nonlinear dynamics of risk contagion. [2] adopted scale-free networks in complex networks as the basis for the entire study, and focused on analyzing the impact of the spread of COVID-19 on corporate credit risk on supply chain networks.

7) **Structural Equation Model:** Structural equation model is a statistical method for evaluating the relationship between variables, and it is a universal general linear model, which is often used to analyze some questionnaires and variables that are difficult to measure accurately. [206] studies the impact of liquidity risk and credit risk on the probability of bank default, in which the author uses structural equation model to analyze whether there is a correlation between liquidity and credit risk. [61] in order to study whether a firm’s supply chain disruption orientation can improve the firm’s supply chain resilience, the author designed a special firm’s resilience scale and used structural equation modeling to analyze the results of the survey. [207] uses structural equation modeling to study whether the supply chain risk of manufacturing firms is related to firm agility performance. In order to study the relationship between supply chain integration and supply chain risk management, [208] uses structural equation model to analyze the questionnaire data of manufacturing companies, and the results show that there is a significant positive correlation between the two. Against the backdrop of supply chain interruptions due to the COVID-19 outbreak, [209] uses structural equation model to investigate whether supply chain risk management increases supply chain resilience and robustness number.

8) **Factor Model:** The factor model is similar to the regression model, both of which study the influencing factors of the explained variables, but the factors of the factor model cannot be directly observed. The more commonly used models are single-factor model, three-factor model and multi-factor model. For example, [210] uses factor model in studying the risk of quarterly heterogeneity in supply chains. [211] uses a single-factor model to study the reasons related to supply chain elasticity. [212] studies the probability of default of SMEs in France and Germany using a one-factor credit risk model. [123] uses CDS spread data and a multi-factor model to study financial market systemic risk in sovereign countries. [213] designs an application containing a multi-factor model to test the credit risk stress of a portfolio.

In the traditional economic or management field, there still
have many common statistical models to study enterprise risk. For example, [57] uses the method of maximum likelihood estimation to study the term structure of the conditional probability of enterprise bankruptcy. Based on the multi-criteria model, [182] establishes a multi-criteria credit risk model using soft information to evaluate innovative small and medium-sized enterprises. [78] uses the Nash equilibrium in the game model to study the competition and cooperation between suppliers and retailers, which reflects one of the sources of risk in the supply chain. In order to analyze the impact of oil price changes on the systemic risk among banks in the Gulf countries, [183] uses the Delta CoVaR method to measure the systemic risk of each bank. Of course, there are many statistical models that we have not shown, but these methods have laid a solid foundation for studying various risks of enterprises on a mathematical basis.

B. Machine Learning Methods

1) Logistic Regression: Logistic regression is a classification algorithm. Logistic regression is supported by linear regression theory, but it introduces nonlinear factors through Sigmoid function, so it can deal with classification problems. [214] proposed a global model of bankruptcy forecasting to predict the risk of bankruptcy in various regions of the world. [42] analyzes the problem of corporate bankruptcy from the perspective of layoffs. When a company falls into a trough, it often try to Isingle-fac-torees. But the article’s research finds that layoffs can disrupt organizations, and increases the probability of bankruptcy. By utilizing a logistic regression model, the article finds that downsizing firms are twice as likely to get bankruptcy as non-downsizing firms. [32] makes a special study of the relationship between macroeconomics and small-firm bankruptcy, and uses logistic regression to study the relationship between corporate bankruptcy and systemic and non-systematic risk factors. In particular, it explores the macroeconomic impact on small business bankruptcy, and has some unexpected findings, such as a better economic environment may also make some businesses go bankrupt, which stems from their idea of exit when profit. [184] further explores the logistic regression model, and approximates the logistic regression through Taylor series expansion, and verifies the validity of the Taylor model in bankruptcy prediction.

In addition to being applied to financial data, macroeconomics and other data, textual information is also introduced into logistic regression models for bankruptcy prediction. [215] uses a logistic regression model to provide incremental information for bankruptcy prediction by analyzing textual disclosure information with a logistic regression model. Based on traditional financial data, [216] introduces text features. By extracting sentiment from textual reports as textual features, combing with financial data, the performance of logistic regression model for bankruptcy prediction is enhanced.

2) Neural Networks: Neural network is a machine learning algorithm that imitates the behavioral characteristics of animal neural network, which can handle end-to-end data processing. The neural network method is applied to small business loan decision-making and has achieved good performance. The input data includes corporate financial data and credit situation, etc. Neural network is widely used in enterprise risk prediction. [217] uses a neural network to analyze a company’s future financial position. The neural network identifies patterns in financial data, and distinguish between companies that are operating normally and companies that are about to go bankrupt. The research of the paper shows that neural network has superior performance than multiple discriminant analysis. [15] analyzes neural networks for corporate bankruptcy prediction, namely supervised neural network and unsupervised neural network. The result of classification for supervised learning models is usually better, but for real-time business environments, training of supervised learning may not be realistic, and this situation may be more suitable for unsupervised learning. [218] combines multilayer perceptrons and self-organizing maps (SOM) for the financial domain. Not only can the probability of bankruptcy be predicted, but it can also visualize the probability of bankruptcy three years before the bankruptcy occurs. The method can assess the strengths and weaknesses of banks in the short, medium and long term by combining the outputs of three multilayer perceptron models in a 2D map using SOM.

Improvements on neural performance of network can be achieved through better training methods, better architecture, or better input data. In [219], the author analyzes the neural network model in three aspects: data span, neural network architecture and iteration times. It is found that rich data can improve the performance of the model, and the latter two need to be properly valued to avoid overfitting. [220] introduces new input data, namely stock price information, in company bankruptcy prediction. The stock price can reflect the company’s future performance expectations to a certain extent. Bankruptcy prediction models often rely on static data, however bankruptcy is a process that changes continuously over time. [64] considers the effect of time and finds that differences in historical data through time may be a better way to predict bankruptcy. Ensemble learning is a good choice for improving the performance of neural networks. In [186], Bagging and Boosting are applied to bankruptcy prediction to improve the classification performance of neural networks, and obtain better performance than traditional neural networks. The bankruptcy factors of companies in different industries are often different. [79] considers the industry in which the company is located, and proposes a neural network bankruptcy prediction model for a specific industry. By t-test and correlation analysis, the variables that are more relevant to the bankruptcy of companies in this industry are selected. The results show an improvement in performance after considering the industry.

3) Support Vector Machine: Support vector machine (SVM) is a kind of generalized linear classifier that performs binary classification on data according to supervised learning, and its decision boundary is the maximum margin hyperplane for the sample data. SVM can be used to analyze the credit status of customers of financial institutions. [221] uses the least squares support vector machine (LS-SVM) classifier to analyze the creditworthiness of corporate customers with financial data as input. In order to improve the efficiency of
government funds, a default prediction model based on SVM is proposed \cite{222}. The paper chooses factors such as financial ratios, economic indicators and technical assessments as input, and achieves good forecasting results. \cite{223} applies support vector machine to bank bankruptcy analysis, and the input data is financial ratio. The study finds that Gaussian kernel support vector machine can extract useful information from financial data. Many researchers have done a lot of work on feature selection so that SVM can obtain better prediction performance. \cite{224} describes the advantages of support vector machines for bankruptcy prediction. Compared with the neural network, it can obtain the optimal solution with a smaller training set and does not require too much parameter tuning.

Some methods are used to improve the performance of SVM in enterprise risk analysis. In \cite{187}, Partial Least Squares (PLS) is used for financial data feature selection. And the processed data is then fed into a Support Vector Machine for bankruptcy prediction. \cite{225} use direct search and features ranking technology to optimize features selection before using SVM. Some scholars are devoted to the optimization of SVM model parameters. \cite{226} is devoted to finding the optimal support vector machine model. Firstly, feature selection is performed to screen out important financial ratio features. The preprocessed data is then used as the input of the support vector machine. Then the optimal parameter of the SVM kernel function is found by grid search technique and 5-fold cross-validation.

4) Random Forest: Random forest (RF) is a classifier of ensemble model that uses multiple trees to train and predict samples. \cite{189} applies the random forest method for credit default prediction and find that non-traditional variables have a significant impact on classification accuracy. \cite{74} builds a separate random forest model for each industry to build an industry bankruptcy prediction model. It can well analyze the financial health of the industry. Random forests can also be used to output intermediate results. \cite{227} combines evidence theory and random forest model. Firstly, the classification result of random forest is used as the basic probability distribution for financial risk characteristic evaluation. Then, the evidence synthesis rule of evidence theory is used as a technical tool for multi-source information fusion, and information fusion is carried out to obtain the financial risk level of the enterprise and its corresponding probability distribution. The model can better reveal the characteristics and causes of financial risks.

5) Epidemic Spreading Model: It is a a model derived from epidemics. The epidemic spreading model regards the company as a member of a system. When a company falls into a crisis, it may spread to other companies, and even cause systemic risks. According to our findings, there is not much research on enterprise risk based on this model. A few articles mainly focus on the risk of the financial system. For example, \cite{228} uses the epidemic model to analyze the impact of the 2008 US subprime crisis on Korean listed companies, and discusses the effectiveness of using the epidemic model.

6) Genetic Algorithm: The genetic algorithm is designed and proposed according to the evolution law of organisms in nature, and it can usually obtain better optimization results faster. In \cite{185}, bankruptcy rules are extracted through genetic algorithms (GAs) and applied to bankruptcy prediction. \cite{153} applies genetic programming to bankruptcy prediction and provides insight into the tangled interaction of different bankruptcy related factors. In \cite{229}, a robust bankruptcy prediction model is constructed by combining genetic programming algorithm with rough set theory.

7) Decision Tree: Decision Tree is a classical classification and regression algorithm in machine learning. Some researcher employed decision tree to study enterprise risk. For instance, \cite{193} applied a two-step analysis methodology to study firm performance: first, using exploratory factor analysis (EFA) to validate underlying dimensions of the financial ratios, second, four popular decision tree algorithms are used to explore the potential relationships between the firm performance and financial ratios. \cite{230} adopts decision tree as the bankruptcy prediction model because of its great interpretability. The paper considers that the bankruptcy risk of enterprises is different under the two macroeconomic conditions of normal economic and economic crisis. Therefore, predicting models are established for the two situations separately to improve the accuracy of forecasting. When there are too many variables to deal with, decision trees can also be used for variable selection \cite{231}.

Many scholars have proposed the optimization scheme of decision tree in order to obtain better performance. \cite{169} proposes a new decision tree ensemble model DTE-SBD which combine the synthetic minority over-sampling technique (SMOTE) and the Bagging ensemble learning algorithm to deal with the imbalanced enterprise credit evaluation problem. Risk exposures between the various business units of an enterprise are shared and relied upon in the corporate structure. \cite{188} develops an integrated optimization framework through a copula-based decision tree interface to help companies to achieve specific goals in enterprise risk management decisions. \cite{232} predicts the financial distress of the restaurant industry through decision tree and AdaBoosted decision tree. AdaBoosting is used to overcome the sensitivity problem of decision trees, and the results show that the results of AdaBoosted decision trees are better.

8) Others Methods: Some papers try to combine theory from other fields with machine learning models or different machine learning models to propose a better model. \cite{233} combines DEA, rough sets, and support vector machines to improve the ability of support vector machines to predict company bankruptcy. \cite{234} propose the application of genetic programming-based neural logic networks, which can explain the network structure through a set of expert rules. Most models attempt to evaluate all bankruptcy scenarios, either using the same set of variables or the same set of samples. In many cases, these two assumptions do not hold. \cite{235} uses biclustering and neural network-based ensembles to predict bankruptcy. Biclustering methods can estimate subgroups of data while considering possible relationships between samples and variables. \cite{236} studies the integration of genetic algorithm and support vector machine. It proposes a method to improve the performance of support vector machine from two aspects of feature subset selection and parameter optimization. Genetic algorithms are used to simultaneously optimize
### TABLE IX: Main Formular

| Model                  | Main Formular                                                                 |
|------------------------|------------------------------------------------------------------------------|
| Logistic Regression Model | \( p = \frac{1}{1 + e^{-\left(\alpha x + \beta \right)}} \)             |
| Neural Networks         | \( a^l_j = \sigma \left( \sum_k w^l_{jk} a^{l-1}_k + b^l_j \right) \)     |
| Support Vector Machine  | \( \min_{w,b} \frac{1}{2} ||w||^2 + \lambda \sum_{i=1}^m \left[ \sum_{y=1}^{K_i} \log \frac{\left( \sum_{k=1}^m p_k \exp(w'x_i + b) \right)}{p_{y_i}} \right] \) |
| Genetic Algorithm       | \( P_i = \frac{\exp(f_i)}{\sum_{i=1}^m \exp(f_i)} \)                      |
| Decision Tree           | \( Ent(D) = -\sum_{k=1}^m p_k \log_2 p_k \)                                |

| Logistic Regression Model | \( p \) Probability of \( y = 1 \)                                      |
|--------------------------|-------------------------------------------------------------------------|
|                          | \( \alpha \) The intercept term of the regression function             |
|                          | \( \beta \) Row vectors of regression coefficients                      |
|                          | \( \lambda \) Column vectors of explanatory variables                   |
|                          | \( a^l_j \) The \( j \)th activated neuron in layer \( l \)            |
|                          | \( \sigma \) Activation function                                        |
|                          | \( w^l_{jk} \) Weight from the \( k \)th neuron in layer \( l-1 \) to the \( j \)th neuron in layer \( l \) |
|                          | \( a^{l-1}_k \) The \( k \)th activated neuron in layer \( l-1 \)      |
|                          | \( b^l_j \) The offset of the \( j \)th neuron in layer \( l \)         |
|                          | \( w \) Parameter vector                                                 |
|                          | \( b \) Offset term                                                      |
|                          | \( y_i \) Label                                                          |
|                          | \( x_i \) The \( i \)th sample                                          |
|                          | \( P_i \) Probability of individual \( i \) being selected               |
|                          | \( M \) Population Size                                                   |
|                          | \( f_i \) Fitness of individual \( i \)                                  |
|                          | \( D \) Sample set                                                       |
|                          | \( m \) Number of sample categories                                       |
|                          | \( p_k \) Proportion of the class \( k \) sample                         |

feature subsets and parameters of support vector machines for bankruptcy prediction. [194] proposes a bankruptcy prediction model based on the adaptive fuzzy k-nearest neighbor (FKNN) method, and the continuous particle swarm optimization (PSO) method is used to determine the neighborhood size and fuzzy parameters of the model.

### C. Deep Learning Methods

Deep learning is an important branch of machine learning, based on an artificial neural network architecture, where multiple processing layers are built inside a deep learning model and then the model parameters are adjusted in a back-propagation manner to enable autonomous exploration of data and feature information generation. This design pattern reduces the incompleteness caused by human-designed features. Thanks to this learning model, deep convolutional neural networks and recurrent neural networks have been the first to make breakthroughs in image processing and speech recognition respectively.

As artificial intelligence continues to develop and advance, deep learning is beginning to appear frequently in financial research work. Compared to classical machine learning models, deep learning models have better performance, generalization and fitting capabilities and are suitable for processing complex data information such as unstructured data, with applications in finance including algorithmic trading, risk management, portfolio management and fraud detection. Study on early warning of E-commerce enterprise financial risk based on deep learning algorithm. This section will introduce some deep learning models applied in the field of financial risk.

**LSTM+CNN model:** [237] extracted structured data and unstructured text from corporate annual reports and constructed a financial risk prediction system using CNN+LSTM model, which can help investors and the companies themselves to detect possible financial crises of listed companies as soon as possible and help companies to deal with their financial risks in a timely manner.

Corporate failure prediction: An evaluation of deep learning vs discrete hazard models.

**BSM-SAES:** [238] proposed a novel approach based on deep learning, called BSM-SAES. This approach combines a boundary synthesis minority oversampling technique (BSM) and a stacked autoencoder (SAE) based on the Softmax classifier. The aim is to develop an accurate and reliable bankruptcy prediction model.

**CNN model:** [175] proposed a method to apply CNN to bankruptcy prediction, and in order to facilitate the training and optimization of CNN networks, the financial scale data from financial statements were converted into grayscale images, and 102 bankrupt companies and 2062 listed companies in Japan were selected as samples for experiments, and the results showed that the model outperformed traditional machine learning models.

**SAE model:** [176] combined stacked autoencoders (SAE) and softmax classifiers in a bankruptcy prediction model, using SAE to extract the best features from the training dataset and train the softmax classification layer to predict the category labels. The experimental results demonstrate the effectiveness of the method.

**Deep Grassmannian Networks:** [14] proposed an optimal deep learning model for predicting corporate failures using panel data structures. Using Deep Grassmannian Network for bankruptcy prediction of US companies, Deep Grassmannian Network (GrNet) borrows ideas from standard deep convolutional neural networks (CNN), and it is shown that
fully connected convolutional layers with normalization have considerable predictive power in the context of Grassmannian networks.

DeepRisk: [8] proposed a Deep Risk method that fuses firm statistics and financing behavior data to predict credit risk of supply chain SMEs, and the results showed that the method significantly outperformed the benchmark model.

D. Hybrid Models

In this section, we will introduce some papers using hybrid models. In the field of enterprise risk research, most scholars use a single model to study this problem. However, a single model may have low accuracy of prediction or evaluation, as well as various limitations and problems existing in the single model. Therefore, some scholars combine several models to learn from each other’s strengths, they expect to avoid some limitations and improve the accuracy of model evaluation. Next, We will introduce some hybrid models in terms of hybrid types, and Table X shows the distribution of the individual models used in the hybrid model. Since scholars have used so many different approaches in their hybrid models, we only show here some of the more used models, where CBR is case-based reasoning method, SVM is support vector machines, RST is rough set theory, the ‘Other’ represent such as process based reasoning method, SVM is support vector machines, so many different approaches in their hybrid models, we only.

Next, We will introduce some hybrid models in terms of hybrid types. Such hybrid models combine some traditional statistical and metrological methods, or combine various recently popular methods in machine learning and deep learning. Generally speaking, such hybrid models appear earlier than cross-hybrid models.

Both [239] and [240] use contagion models with Bernoulli mixture models, but they focus on different points. In [239], the contagion effect model and Bernoulli mixed model are used to study the importance of overall credit loss risk to large financial institutions. [240] applied the credit contagion model and the Bernoulli mixed model to study the distribution of losses in large portfolios caused by credit crises. Some scholars use Bayesian Belief Networks, principal component analysis, and Merton’s structural model combined with other models to analyze risk transmission problems. [197] proposed a supply chain risk network management (SCRNM) process based on Bayesian Belief Networks (BBNs) and Expected Utility Theory (EUT), and the authors also proposed a ‘weighted net evaluation of risk mitigation’ method, which technically employs safety and reliability engineering, decision making under uncertainty, and multi-criteria decision analysis. Systematic risks such as financial crisis are difficult to predict, [1] uses principal component analysis and Granger causality test to simulate systematic risks. To analyze the risk spread of interbank networks, [241] uses the Maximum Influence Diagram (MID) and Merton’s structural model to investigate this problem. In addition, [77] uses the Z-score model and Stepwise Least Squares Estimation through the Forward method to study bankruptcy risk and tried to extend the time frame of the model’s predictions.

Next is the hybrid model between machine learning algorithms, and due to the variety of features of these algorithms, many predictive studies have improved performance when using hybrid models of multiple algorithms.

Some scholars have combined two machine learning algorithms together in an attempt to improve prediction performance. [242] develops a hybrid model combining artificial neural networks and data mining techniques to predict financial distress. [243] combined genetic algorithms with the case-based reasoning method to improve bankruptcy prediction performance. To better predict the credit risk, [164] proposes an enhanced hybrid integrated machine learning method called RS-MultiBoosting, which combines two classic integrated machine learning methods of random subspace (RS) and MultiBoosting. [244]’s evaluation method combines genetic algorithms and support vector machines, which is used to study supply chain credit risk.

Other more radical researchers have combined three or even more machine learning algorithms in an attempt to further improve prediction performance. [245] proposes a data mining method combining attribute-oriented induction, information gain, and decision tree to predict financial distress problems. Similarly, [260] also proposes a financial distress prediction method, but the method is based on the serial combination of multiple classifiers, which combines support vector machine, multiple discriminant analysis and case-based reasoning. [246] proposes a selective integration model, which combines a decision tree, back-propagation neural network and support vector machine, this model can better forecast bankruptcy. [233] pays special attention to the data of the company management efficiency and proposes a new hybrid model to predict business failure, this model combines the rough set theory, support vector machine method and data envelopment analysis method. To improve the accuracy of bankruptcy forecasts, [196] develops hybrid models also based on case-based reasoning(CBR): RST–CBR (combining Rough Set Theory with CBR), RST–GRA–CBR (integrating RST, Grey Relational Analysis, and CBR).

2) Cross hybrid model: Cross hybrid models combine different types of models, such as statistical models and machine learning models. Such hybrid models better reflect the combination of economics and computer science. Generally, such models may achieve a good balance between prediction accuracy and interpretability. When machine learning methods emerged, there were attempts to combine statistical methods with machine learning, and this has remained a worthwhile way to explore until now. This paragraph presents literature that includes only one statistical method combined with one machine learning method. On the issue of enterprise risk prediction and classification, [75] first compares the performance of traditional statistical methods (including linear discriminant analysis and logit analysis) with neural networks, and the author recognizes the potential of neural networks in predictive problems, so the author proposes to combine traditional discriminant analysis with neural networks. [247] combines neural network learning and logit analysis to build a hybrid model named RBFN (radial basis function network) to predict financial distress. [6] proposes an integrated binary discriminant rule (IBDR) predict
TABLE X: Hybrid Model Distribution

| Literature | Neural Networks | Data Mining | Genetic Algorithms | CBR | SVM | Decision Tree | RST | Logit | Random Forests | Other statistical model | Other ML Model | Other | Total |
|------------|-----------------|-------------|--------------------|-----|-----|---------------|-----|-------|-----------------|------------------------|----------------|-------|-------|
| 239        | 2               | 2           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 240        | 2               | 2           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 197        | 1               | 1           | 2                  |     |     |               |     |       |                 |                        |                |       | 3     |
| 195        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 194        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 243        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 164        | 2               | 2           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 241        | 1               | 1           | 2                  |     |     |               |     |       |                 |                        |                |       | 3     |
| 196        | 1               | 1           | 1                  |     |     |               |     |       |                 |                        |                |       | 3     |
| 195        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 240        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 247        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 195        | 1               | 1           | 1                  |     |     |               |     |       |                 |                        |                |       | 3     |
| 196        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 249        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 120        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 252        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 253        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 254        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 255        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 3     |
| 257        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |
| 258        | 1               | 1           | 1                  |     |     |               |     |       |                 |                        |                |       | 2     |
| 259        | 1               | 1           |                    |     |     |               |     |       |                 |                        |                |       | 2     |

Some scholars are not satisfied with the performance improvement brought by hybrid two models, and they try to cross-hybrid using three different methods, and this segment has cross-hybrid models based on statistical methods, and cross-hybrid models based on machine learning or deep learning methods. [250] proposes a mixed model using decision tree classification method and logical regression method, the author uses the principal component analysis method to extract variables. [251] develops a business intelligence model to predict corporate bankruptcy, which combines the financial ontological model, the data mining algorithm and the Altman Z-score method. [120] proposes an integrated model called NIM to predict bankruptcy, which combines convolutional neural network oriented deep learning, support vector machine method and soft set theory. [252] develops a risk decision support system that automatically selects supply chain risk management strategies, this system combines methods such as Systematic review, Correspondence analysis and Fuzzy inference system. [253] uses back-propagation neural networks, analytic hierarchy processes, and data mining methods to predict the financial risk of listed companies. On bankruptcy forecasting, [254] uses three taxonomic methods to deal with data, including the least-squares approach to anomaly detection, an isolation forest, and one-class support vector machines. In order to improve the accuracy of financial indicators to measure systemic financial risks, [255] uses logistic regression methods, artificial neural networks and time-varying parameter vector autoregressive model (TVP-VAR) to process multiple data.

There are also some scholars who intend to combine the strengths of each school, they hybrid four and more statistical and machine learning models, and they have achieved
Some studies may fall into multiple categories. Here, we only list the representative articles in single categories according to their foci, i.e., existence and methods. This categorization of relevant work is not mutually exclusive. Some studies may fall into multiple categories. Here, we only include the main topic of the article. Table XIII summarizes the findings of these representative works for easy review.

The first category is the existence category, which focuses on demonstrating the comovement between enterprise financial risk and various types of data including financial index, textual information, relational data and intelligence integration. [39] conducted a study that revealed there was a direct connection between the corporate governance structure and enterprise bankruptcy. [206] examine the relationship between liquidity risk and credit risk in the banking industry from different perspectives. [271] analyze supply chain tweets and gain insight into the potential role of Twitter in supply chain practice and research.

The second category is the methods category that is devoted to utilizing and enhancing various types of techniques, especially the latest advances in the fields of natural language processing and machine learning, to capture valuable information from textual media and relation data and bridge the connections between enterprise multi-source heterogeneous data and enterprise financial risk. [172] use machine learning models to predict the probability of enterprise bankruptcy and find a higher average performance than usual models. [276] use a reduced-form econometric model to predict short- and long-term corporate bankruptcies and failures, while showing that stocks with a high risk of failure tend to deliver lower average returns.

### IV. Spotlights of Representative Works

In this section, we first make a brief introduction of the representative works and then present their unique contributions from several different perspectives.

We select the representative works based on citations and methodological progress. The citation for each reference is based on the citation statistics from Google Scholar as of November 1, 2022. Table XIV presents the most representative studies during the period from 1968 to 2019 in terms of citations. In addition, we rank the journals in terms of the citations of their articles and present the statistical results in Table XII. These journals cover a wide range of topics, from finance and management science to computer science. Journal of Finance, Journal of Banking & Finance and Expert Systems with Applications is ranked first with 4 publications. These representative studies with high recognition from peer researchers can be roughly classified into two categories according to their foci, i.e., existence and methods. This categorization of relevant work is not mutually exclusive. Some studies may fall into multiple categories. Here, we only mention the most prominent contributions.
### TABLE XI: The Representative Works in Terms of Citations.

| Reference | Title                                                                 | Journal                                                                 | Year | Citations |
|-----------|------------------------------------------------------------------------|------------------------------------------------------------------------|------|-----------|
| 26.1      | Financial ratios, discriminant analysis and the prediction of corporate bankruptcy | J FINANC                                                               | 1968 | 22519     |
| 26.2      | Systemic risk and stability in financial networks                      | American Economic Review                                               | 2015 | 1835      |
| 26.3      | A neural network model for bankruptcy prediction                        | IJCNN                                                                  | 1990 | 1401      |
| 26.4      | Benchmarking state-of-the-art classification algorithms for credit scoring | Journal of the Operational Research Society                           | 2003 | 1138      |
| 26.5      | An application of support vector machines in bankruptcy prediction model | Expert systems with applications                                       | 2005 | 1041      |
| 26.6      | Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters | Expert systems with applications                                       | 2005 | 1126      |
| 26.7      | Ambidexterity and performance in small-to-medium-sized firms            | Journal of management                                                  | 2006 | 2506      |
| 26.8      | The pivotal role of top management team behavioral integration          | Journal of Business Venturing                                           | 2011 | 2184      |
| 26.9      | Is innovation always beneficial? A meta-analysis                        |                                                                        |      |           |
| 26.10     | of the relationship between innovation and performance in SMEs          |                                                                        |      |           |
| 26.11     | Bankruptcy and corporate governance: The impact of board composition and structure | Academy of Management journal                                          | 1994 | 1305      |
| 26.12     | Competitive and Cooperative Inventory Policies in a Two-Stage Supply Chain | Management science                                                    | 1999 | 874       |
| 26.13     | Enterprise risk management: An empirical analysis of factors associated with the extent of implementation | Journal of Accounting and Public Policy                              | 2005 | 1072      |
| 26.14     | The relationship between liquidity risk and credit risk in banks        | Journal of Banking & Finance                                           | 2014 | 418       |
| 26.15     | Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors | Journal of financial economics                                        | 2012 | 2312      |
| 26.16     | Is the risk of bankruptcy a systematic risk?                            | J FINANC                                                               | 1998 | 1362      |
| 26.17     | Modelling credit risk for SMEs: Evidence from the U.S. market          | ABACUS                                                                 | 2007 | 1018      |
| 26.18     | The Determinants of Credit Default Swap Premia                          | Journal of Financial and Quantitative Analysis                       | 2009 | 870       |
| 26.19     | Machine learning models and bankruptcy prediction                       | Expert Systems with Applications                                       | 2017 | 555       |
| 26.20     | Corporate Social Responsibility in the Supply Chain: An Application in the Food Industry | Journal of Business Ethics                                           | 2006 | 1337      |
| 26.21     | Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence | Institute for Operations Research and the Management Sciences           | 2019 | 858       |
| 26.22     | Evaluation of clustering algorithms for financial risk analysis using MCDM methods | Information Sciences                                              | 2014 | 206       |
| 26.23     | Rollover Risk and Credit Risk                                           | THE JOURNAL OF FINANCE                                                | 2012 | 625       |
| 26.24     | Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research | Production Economics                                                 | 2015 | 467       |
| 26.25     | Principal Components as a Measure of Systemic Risk                      | The Journal of Portfolio Management Summer                             | 2011 | 416       |
| 26.26     | Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns | Journal of Banking & Finance                                           | 2011 | 572       |
| 26.27     | Cross-border interbank networks, banking risk and contagion             | Journal of Financial Stability                                         | 2015 | 134       |
| 26.28     | Pricing derivatives on financial securities subject to credit risk      | Journal of Financial                                                | 1995 | 3034      |
| 26.29     | The power of flexibility for mitigating supply chain risks              | International Journal of Production Economics                         | 2008 | 1115      |
| 26.30     | The severity of supply chain disruptions: design characteristics and mitigation capabilities | Decision sciences                                                      | 2007 | 1734      |
| 26.31     | In search of distress risk                                              | The Journal of Finance                                                | 2008 | 2641      |
| 26.32     | A markov model for the term structure of credit risk spreads            | The review of financial studies                                       | 1997 | 2336      |
| 26.33     | The effect of credit guarantees on credit availability and delinquency rates | Journal of Banking & Finance                                         | 2015 | 83        |
| 26.34     | The Impact of Public Guarantees on Bank Risk-Taking: Evidence from a Natural Experiment | Review of Finance                                                   | 2014 | 329       |
| 26.35     | Deep Learning Models for Bankruptcy Prediction using Textual Disclosures | European journal of operational research                              | 2019 | 203       |
| 26.36     | Bankruptcy prediction using imaged financial ratios and convolutional neural networks | Expert systems with applications                                       | 2019 | 192       |
| 26.37     | Expropriation through loan guarantees to related parties: Evidence from China | Journal of Banking & Finance                                         | 2009 | 450       |
TABLE XII: Journal Influence in Terms of Citations.

| Journal                                      | Total | Number | Papers |
|----------------------------------------------|-------|--------|--------|
| J FINANC                                    | 23881 | 4      | 261    |
| Academy of Management journal                | 1305  | 1      | 39     |
| Management science                           | 1732  | 2      | 178    |
| Journal of Accounting and Public Policy      | 1072  | 1      | 180    |
| Journal of Banking & Finance                 | 1523  | 4      | 296    |
| Journal of financial economics               | 2312  | 1      | 1      |
| Journal of Financial and Quantitative Analysis | 870  | 1      | 268    |
| Expert Systems with Applications             | 2914  | 4      | 172    |
| ABACUS                                       | 1018  | 1      | 1      |
| THE JOURNAL OF FINANCE                       | 625   | 1      | 158    |
| Int. J. Production Economics                 | 467   | 1      | 171    |
| Information Sciences                         | 700   | 1      | 152    |
| Journal of Business Ethics                   | 1337  | 1      | 159    |
| Review of Finance                            | 329   | 1      | 15     |
| European journal of operational research     | 203   | 1      | 16     |
| IJCNN                                        | 1399  | 1      | 263    |
| Journal of the Operational Research Society  | 1138  | 1      | 264    |

recognize the crisis it is facing and thus react in advance. By introducing new technologies, it can also assist companies in better management. Also, timely attention to rapidly changing financial markets may be beneficial to the management of corporate risk.

1) Enterprise risk analysis: It is not easy for enterprises to operate. In the face of rapidly changing markets and constant competition, enterprises need to pay close attention to their own performance at all times. The enterprise bankruptcy model can help companies to better understand their own performance, and to adjust their own strategies in time, in order to avoid bankruptcy.

The data mining methods and decision tree models for predicting financial distress constructed by [245] are important for the sustainability of the company. [258]'s proposed financial distress prediction model based on a weighted majority voting combination of multiple classifiers can help companies effectively anticipate financial distress, which is important for their survival and growth. [131] use a large number of databases to develop discrete-time risk models that assess data publicly available at the point of incorporation of new firms and early in the lifecycle and their predictive power in determining the likelihood of bankruptcy. The model achieves acceptable in-sample and out-of-sample classification accuracy for the rating system and demonstrates the predictive contribution of director and board characteristics. [124] establishes a dynamic bankruptcy model with asset illiquidity, and analyzed the company’s exit policy in the face of bankruptcy, its time, default probability, and equity, debt and company value. By discriminatingly analyzing the financial characteristics of failed Japanese companies, [125] provides a new idea for improving the accuracy of predicting bankruptcy. The article shows that the use of data from two or more years prior to bankruptcy together with other indices can lead to more accurate predictions of corporate bankruptcy, while emphasizing that the basis for compiling financial statement data for each sample company should not be overlooked. Methods like convolutional neural networks [14] can also be applied in enterprise bankruptcy prediction. [278] present a novel parameter reduction method to select financial ratios for business failure prediction. Based on deep learning algorithms, [28] conducts research from the perspective of establishing a financial early warning model based on deep learning and building a financial risk early warning mechanism for e-commerce companies, and analyzes and predicts the financial risks of listed companies. [50] aims to improve the ability of enterprises to deal with financial crises, find some countermeasures to prevent potential financial risks, and establish an early warning system model for financial risks. In [219], people can evaluate and predict the firm’s future financial status by using data from the firm’s financial reports. [227] combines the evidence theory and random forest algorithm to establish an enterprise financial risk warning model, and it also can trace the causes of financial risks. [279] uses a logit scoring model to predict the probability of default, which is helpful for SMEs to understand their default risk, and also can be used to estimate their risk-adequate debt cost. [120] proposes a new integrated model (NIM) for predicting corporate bankruptcy in China’s energy industry and applies it to real-world data of Chinese energy-listed companies. [242] takes the operating rules violated by Taiwan stock exchange companies as the scope, extracts adaptive variables, uses artificial neural network and data mining technology to build a financial distress prediction model, and verifies that the proposed method can predict the financial distress of listed companies. [139] discusses models used to predict SME defaults, and proposes a new approach to credit risk assessment that fuses data from longitudinal and survival duration models, providing a novel approach for more accurate risk assessments. [146] supplemented limited accounting data using non-financial, regulatory compliance, and “event” data from private companies and ultimately found that data related to legal actions taken by creditors to recover unpaid debts, company filing history, consolidated audit report/opinion data, and company-specific characteristics, It has made a significant contribution to improve the default prediction ability of risk models specially built for small and medium-sized
| Group | Reference | Contribution |
|-------|-----------|--------------|
| Existence | 39 | The authors used logistic regression to analyze the relationship between bankruptcy and corporate governance structure, which gave rise to the trend of studying the relationship between corporate governance structure and financial position. The authors use a game model to study the relationship between inventory strategy and competitive cooperation in supply chains, and the conclusion that competition may reduce inventory brings a new insight to the study of supply chain risk. The authors examine what are the reasons for companies to adopt ERM and the conclusions provide a preliminary basis for future research on ERM deployment. This is the first paper to examine the relationship between liquidity risk and credit risk in the banking industry from different perspectives, and the findings have significant implications for subsequent research on the relationship between liquidity risk and credit risk. This paper fills a gap in the literature by formalizing and testing the channels through which CSR policies affect corporate systemic risk and value. Using Z-score and O-score model for bankruptcy prediction, it demonstrates that higher bankruptcy risk doesn’t bring about higher returns which is different from previous opinions. This paper contributes to the supply chain community by proposing a novel analytical framework to analyze supply chain tweets and gain insight into the potential role of Twitter in supply chain practice and research. This study fills a key gap in the literature on whether structural shocks in the oil market affect systemic bank risk. It extends the literature examining the interlinkages between oil shocks and bank risk in GCC countries by simultaneously examining two major crises, namely the global financial crisis and the ongoing COVID-19 pandemic. The study finds that modeling credit risk for SMEs lower the capital requirements for banks under certain circumstance. The results also suggests that credit risk management of small and medium size enterprises are different from large corporates which may be insightful for banks. It analyzes the linear relationship of theoretical determinants of default risk and default swap spreads and identifies explanatory power of volatility and leverage in regressions. This paper contributes to the supply chain community by proposing a novel analytical framework to analyze supply chain tweets and gain insight into the potential role of Twitter in supply chain practice and research. This paper provides a measure of the vulnerability of the Italian interbank market to financial contagion, thus enriching the evidence available to date in most industrialized countries. |
| Methods | 1 | The authors use principal component analysis and Granger causality networks to study systemic risk among banks, insurance, funds and dealers, the model has been influential in studying systemic risk. The authors use principal component analysis and Granger causality networks to study systemic risk among banks, insurance, funds and dealers, the model has been influential in studying systemic risk. This study tests the performance of machine learning models in bankruptcy prediction and it finds a higher average performance than usual models. Thus, it’s a substantial improvement in accuracy using machine learning methods. The framework of CSR in food supply chain proposed in this paper lays a foundation for the food industry to further study the elements of CSR in supply chain. This paper provides a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. This paper provides a new pricing and hedging theory for derivative securities involving credit risk, and introduces a new metric DebtRank which can avoid infinite reverberation by excluding the repeated walking of the edge. This study tests the performance of machine learning models in bankruptcy prediction and it finds a higher average performance than usual models. Thus, it’s a substantial improvement in accuracy using machine learning methods. The framework of CSR in food supply chain proposed in this paper lays a foundation for the food industry to further study the elements of CSR in supply chain. This paper provides a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. The framework of CSR in food supply chain proposed in this paper lays a foundation for the food industry to further study the elements of CSR in supply chain. This paper provides a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. This paper provides a new pricing and hedging theory for derivative securities involving credit risk, and introduces a new metric DebtRank which can avoid infinite reverberation by excluding the repeated walking of the edge. This study tests the performance of machine learning models in bankruptcy prediction and it finds a higher average performance than usual models. Thus, it’s a substantial improvement in accuracy using machine learning methods. The framework of CSR in food supply chain proposed in this paper lays a foundation for the food industry to further study the elements of CSR in supply chain. This paper provides a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. This paper provides a new pricing and hedging theory for derivative securities involving credit risk, and introduces a new metric DebtRank which can avoid infinite reverberation by excluding the repeated walking of the edge. This study tests the performance of machine learning models in bankruptcy prediction and it finds a higher average performance than usual models. Thus, it’s a substantial improvement in accuracy using machine learning methods. The framework of CSR in food supply chain proposed in this paper lays a foundation for the food industry to further study the elements of CSR in supply chain. This paper provides a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. This paper provides a new pricing and hedging theory for derivative securities involving credit risk, and introduces a new metric DebtRank which can avoid infinite reverberation by excluding the repeated walking of the edge. |

A reduced-form econometric model is implemented to predict short- and long-term corporate bankruptcies and failures, which gives rise to the trend of studying the relationship between corporate governance structure and financial position. The authors use a game model to study the relationship between inventory strategy and competitive cooperation in supply chains, and the conclusion that competition may reduce inventory brings a new insight to the study of supply chain risk. The authors examine what are the reasons for companies to adopt ERM and the conclusions provide a preliminary basis for future research on ERM deployment. This is the first paper to examine the relationship between liquidity risk and credit risk in the banking industry from different perspectives, and the findings have significant implications for subsequent research on the relationship between liquidity risk and credit risk. This paper fills a gap in the literature by formalizing and testing the channels through which CSR policies affect corporate systemic risk and value. Using Z-score and O-score model for bankruptcy prediction, it demonstrates that higher bankruptcy risk doesn’t bring about higher returns which is different from previous opinions. This paper contributes to the supply chain community by proposing a novel analytical framework to analyze supply chain tweets and gain insight into the potential role of Twitter in supply chain practice and research. This study fills a key gap in the literature on whether structural shocks in the oil market affect systemic bank risk. It extends the literature examining the interlinkages between oil shocks and bank risk in GCC countries by simultaneously examining two major crises, namely the global financial crisis and the ongoing COVID-19 pandemic. The study finds that modeling credit risk for SMEs lower the capital requirements for banks under certain circumstance. The results also suggests that credit risk management of small and medium size enterprises are different from large corporates which may be insightful for banks. It analyzes the linear relationship of theoretical determinants of default risk and default swap spreads and identifies explanatory power of volatility and leverage in regressions. This paper contributes to the supply chain community by proposing a novel analytical framework to analyze supply chain tweets and gain insight into the potential role of Twitter in supply chain practice and research. This study fills a key gap in the literature on whether structural shocks in the oil market affect systemic bank risk. It extends the literature examining the interlinkages between oil shocks and bank risk in GCC countries by simultaneously examining two major crises, namely the global financial crisis and the ongoing COVID-19 pandemic. The paper proposes a unified framework and 5 stylized models. The findings highlight the power of flexibility while clarifying the associated benefits of different levels of flexibility, providing insights into deploying flexibility to reduce supply chain risk. The article examines how and why supply chain disruption is a bigger problem than the other. Not only does it add to existing knowledge related to supply chain risk, vulnerability, resilience and business continuity planning, but it also calls into question the wisdom of pursuing practices such as supply base reduction, global sourcing and sourcing from supply clusters.
enterprises. [280] introduces a hybrid default prediction model considering market information. Experimental results show that the proposed hybrid model can well replace the existing accounting based default prediction standard method for smes, and provides a new model for smes default prediction.

2) Supply Chain Management: Few enterprises can control the whole industrial chain. The vast majority of enterprises are part of the supply chain. Therefore, the research on the supply chain can help enterprises reduce the supply chain risk. [209]'s findings from a survey of French companies in the context of the COVID-19 outbreak can provide guidance on the specific conditions and cooperation for companies to adopt supply chain risk management practices. The approach to supply chain integration proposed after [207] empirical study can enable firms facing supply chain risks to enhance their agility performance. [211]'s research on supply chain resilience can help motivate supply chain managers to improve their integration capabilities and increase supply chain resilience. The results of [208]'s research on supply chain integration and supply chain risk management can help supply chain managers develop integration practices to manage risk and improve operational performance. [197]'s supply chain risk network management process can help decision-makers rank risks and strategies. [210] studies the propagation of financial risk in supply chain networks finding that the risk to the focal firm is higher when secondary suppliers are highly shared, concluding that firms should actively identify and monitor these secondary suppliers, which can effectively mitigate supply chain risk. [78] studies risk handling and mitigation strategies in supply chains, and identifies general strategies for supply risk, demand risk, and other risks. [40] empirically testify the factors that affects stock market when the supply chain disrupts. This study has essential implications on firms' strategies of their supply chain management especially during the period of disruptions. Specially, great attention have been paid in small and medium-sized enterprise (SMEs). [10] propose an integrated ensemble machine learning method for predicting the credit risk of China’s small and medium-sized enterprise in supply chain finance. The method possesses an outstanding prediction performance and it is especially suitable for forecasting the credit risk of China’s SME in SCF. In order to help supply chain risk managers choose appropriate risk management strategies, [252] focuses on the risk mitigation steps in the supply chain risk management (SCRM) process and develops a decision-making framework for recommending effective risk mitigation strategies.

3) Assisting Enterprise Management: Through data analysis, association analysis, model building, more useful information for enterprises can be provided and help companies make better management decisions. [251] has improved a new business intelligence approach that can facilitate business operations through the analysis of big data. [256] proposes two hybrid models to extract valuable information from big data so that companies can better define their own strategic plans, such as risk control, crisis management, or growth management. [25] carries out a study on the relationship between corporate social performance and financial risk, and found that corporate social responsibility was negatively correlated with the company’s systemic risk, while corporate social irresponsibility was positively correlated with financial risk. This finding has implications for future corporate management. It has important research significance. [26] explores the relationship between corporate social irresponsibility (CSI) and financial risk, and its findings complement existing theories on the risk mitigation effects of corporate social responsibility. At the same time, these insights provide complementary strategies for enterprise risk management. [281] analyze the relationship between maturity of enterprise risk management and firm's value. [188] develops an integrated optimization framework through copula based decision tree interface, so that enterprise risk management decisions can meet the specified enterprise goals in a multi-stage setting. [163] based on a sample of listed insurance companies in the United States, this paper simulates the impact of Enterprise risk management(ERM) and ERM on the cost of capital, and finds that ERM significantly reduces the cost of capital and provides a new direction for enterprise risk management. [282] finds that companies benefit from higher ratings and lower interest spreads if their relative ESG performance matches that of the corresponding country. [274] examines four different types of risk mitigation strategies and finds that businesses can reap most of the benefits in low levels of flexibility. This fact can make it easier for managers to justify investments in flexibility even when the impact and likelihood of different types of risks cannot be accurately estimated.

4) Utilizing Financial Market Information: The timeliness of the financial market helps enterprises to understand the situation they face more quickly and make better investment and financing decisions. [122] examines whether transactions in credit default swaps (CDS) increase the credit risk of the reference entity by quantifying the impact of CDS transactions on a company’s credit risk, finding companies with relatively large numbers of CDS contracts outstanding, as well as those with “no-restructuring” contracts companies are more likely to be adversely affected by CDS transactions. [117] establishes a model to illustrate the impact of liquidity risk on the yield spread of corporate bonds, and the results showed that the liquidity spread is positively correlated with the probability of default. [49] establishes a real option model and discusses the investment and financing strategies of small and medium-sized enterprises under the conditions of the above-mentioned partial guarantee contract. Explicit formulas are derived for the pricing and timing of cash flow investment options with diffusion and jump risk. Larger funding gaps or higher levels of guarantees are found to lead to later investments and lower option values. [48] uses option pricing techniques to obtain estimates of the pure money cost of loan guarantees, the company’s interest savings on senior and subordinated debt, and the project’s implied present value profitability index. [141] examines the nature of corporate bond default risk using a large new dataset and finds that changes in stock returns, stock return volatility, and GDP are strong predictors of default rates, providing indicators for corporate bond default risk prediction. [142] provides new evidence for corporate bond investors to price the time-varying risk of debt deflation using international corporate information gap data. The results of this
Paper provide a rich and effective way for further research, and have implications for the optimal capital structure of enterprises. [270] provides a model to analyze the impact of debt market liquidity on corporate credit risk through debt rollover, providing a new direction for corporate credit risk assessment.

B. Financial Institution

This section describes the application of scholars' findings to financial institutions. In the collected papers, we found that scholars' research on enterprise risk can help financial institutions regulate market risk, identify and evaluate various risks, improve loan efficiency and profitability.

1) Regulation and Control: The approach proposed by some scholars can capture the overall risk profile at a macro level, which facilitates regulation and control by central banks or regulators. [283] presents a dynamic multi-agent model for bank systems that have a central bank, which proves that the central bank's decision to intervene can actually reduce the financial distress and liquidity shortage in the interbank market. [284] presents a new systematic risk measure in the context of financial networks, which leads to a composite indicator capable of assessing the overall stress state of the financial system. [223] shows that the gaussian kernel based support vector machine can extract useful information from financial data and can be used as a part of the bank bankruptcy early warning system.

2) Risk Identification and Control: A large part of the role of studying enterprise risk is to help financial institutions identify risks to themselves and their customers. The findings of some scholars can help financial institutions such as banks identify their own risks as well as enhance risk management. [218] develops a neural network model to study the bankruptcy of banks of America, which is helpful to bank identify early financial risks. [145] proposes a new semi-supervised text mining method to analyze the qualitative textual risk disclosure of a large amount of bank risk information in financial statements. This method can identify the bank's own risk factors more accurately. [243] combines genetic algorithms and case-based reasoning methods to build bankruptcy prediction models that can be important for risk management in financial institutions. The results obtained from [239]'s study of total credit losses on large financial positions provide important insights into the measurement of risk in financial institutions. [206]'s analysis of banks' liquidity risk and credit risk does have an impact on bank defaults, and the findings motivate banks to strengthen their management of liquidity and credit risk. [62] proposes an outranking multicriteria modeling approach that enables analysts to learn the characteristics of the model, and make some changes to meet the requirements of banks' risk management. [11] evaluate the credit risk of supply chain finance from the perspective of banks which introduces theoretical support for banks' risk management.

More approaches to studying enterprise risk are used to identify the risk of financial institutions' customers. For example, banks assess the risk status of their customers before lending, and some companies issuing bonds have their risks rated by professional rating agencies. [102] proposes a novel predictor selection procedure based on non-parametric regression and classification tree approach (CART). It has high applicability among bank practitioners. This method is a viable risk assessment tool that is applicable to the real-world environment in which banks operate. In [189], the RF model is used to model the data of micro-enterprises in China, and the advantages of the model are verified through experiments. The algorithm can be used by financial institutions to systematically measure the probability of credit default for both individual and corporate users. [285] develop an intelligent-agent-based fuzzy group decision making model which may be an effective multicriteria tool for decision analysis that can be used by credit granting institutions in the area of credit risk assessment. [280] proposes a method that can extract relevant features from legal judgment information to predict credit risk. It can help financial institutions to assess the credit risk of SMEs. [162] use a sample of P&C insurance companies with S&P ERM quality ratings from 2006-2013, this article examines the independent and aggregate effects of ERM and diversity on their performance, helping large financial institutions institutional build rating standards and optimize management decisions. The bankruptcy prediction model proposed in [185] can provide a basis for the credit rating system, which is very important for financial institutions. [166] studies China's bond market and credit rating, their findings are beneficial and help rating agencies to unify standards and classify bonds according to default risk, increase the information content of bond ratings, and build a healthy and stable bond market.

When a bank lends to a customer there is still a risk of default and some scholars’ research methods can help identify post-loan risks. [221] explains and evaluates the added value of the Bayesian LS-SVM classifier, and uses the bankruptcy data set of the Benelux mid-cap stock market to carry out the experiment, which shows that the proposed classifier based on Bayesian nonlinear kernel has a better performance. It provides a good discriminator for financial institutions to predict enterprise bankruptcy. Banks and financial companies use the bankruptcy prediction model proposed by [235] to estimate the risk of customers’ inability to repay their debts.

3) Loan and Profit: Most financial institutions are for-profit, and most of their profits come from lending to businesses. Scholars' research methods many be able to help these financial institutions make better credit decisions to increase their profits. [76] proposes a new bankruptcy prediction model using Self Organized Maps (SOM) and Multivariate Adaptive Regression Splines (MARS), which can help banks to make effective investment and lending decisions. [217] makes a reliable early prediction of the future health of enterprises, which can help bank loan officers make appropriate loan decisions. [126] builds a regression model combined with the firm's creditworthiness and whether it defaults. This model can be used to help banks predict credit risk and obtain credit strategies. The bankruptcy prediction model proposed in [187] can help financial institutions better carry out credit approval, loan portfolio, security management and so on. The hybrid model obtained by [246]'s selective integration of the
three classifiers can reduce the number of wrong decisions made by financial institutions. [283] presents a dynamic multi-agent model for bank systems that have a central bank, which proves that the central bank’s decision to intervene can actually reduce the financial distress and liquidity shortage in the inter-bank market. Because accurate financial distress forecasts can have a significant impact on a financial institution’s lending decisions and profitability. [250]’s financial distress models using a combination of methods can provide investors with the necessary information. [11] combines genetic algorithm, support vector machine and BP neural network to measure credit risk in supply chain finance, and the method can reduce the probability of bank profit impairment.

C. Government Regulation

Advancements in quantifying the influence of multi-source heterogeneous data on enterprise financial risk in the era of big data can be critical for solving several challenging issues from the perspectives of risk prediction, risk warning and government investment.

1) Systemic Risk Prediction: Based on multi-source heterogeneous data, government can predict industry risks and systemic risks. The DebtRank proposed by [95] could provide a quantitative assessment of systemically important financial institutions, and the method is applicable to regulators’ supervision of systemic risk. The methodology developed by [201] to assess the risk of systemic failure of global banks is a reasonable estimate of systemic risk, which is one of the issues of great concern to regulators and politicians. TALIS3, a novel systemic risk monitoring approach, is proposed to assess the systemic impact of the COVID-19 pandemic on financial markets in North America and Europe. This monitoring method can be used by regulators to measure and more precisely assess the economy-wide costs associated with systemic events [89]. The method proposed in [287] to predict the systemic impact of financial institutions in interconnected systems can be applied to monitor systemic risk in banks and insurance companies in a timely manner. [103] proposes a new approach to help regulators quantify systemic risk and provide clues to containment. In order to measure the effectiveness and accuracy of financial indicators, [255] shows that it is necessary to pay attention to the changes of interest rate, real estate price and stock price when monitoring China’s financial systemic risks, and the influence of exchange rate changes and monetary policies can not be ignored.

2) Systemic Risk Warning: [69] recommends that regulators and investors create indices of implied systemic risk, defined as absorption ratios, for various markets that can act as early warning signs of potential asset depreciation and financial turbulence. With warning of potential problems, policymakers and investors can take steps to prevent them from becoming a reality. A new framework is proposed in [100] to study default dependence and systemic risk in financial networks evolving over time. This methodological framework can help to come up with systemic bank rankings.

3) Guide and Manage Government Investments: Business failure prediction is essential to investors and government. Thus, the RST–SVM model proposed by [233] which has a good performance in failure prediction can provide a guide of investment for them. [222] proposes a support vector machine (SVM) model to predict the default of subsidized SMEs and it can help to effectively manage these government funds.

VI. Directions For Future Work

A. Enterprise Intelligence

Traditional financial information is widely employed in past studies and they still have considerable significance in the future study of corporate risk. But with the development of big data analysis and data mining technologies, data tends to be diversified. Such diversity not only comes from the bigger size of the data but also comes from the various information it delivers to us. When faced with mass data, more technology like big data analytics is available to deal with it. Besides, information from different perspectives may be extracted from the mass data which could be a supplement to traditional data. Like data from customers, industry, linked enterprises, and even the macro environment. For example, can we get any valuable information from the company’s utility bill? The answer might be yes. Such data seems to be inconspicuous, but can’t be ignored. Additionally, information from the government is also non-negligible. It may provide guides for the company’s activities. Sometimes, collecting feedback from customers can also reveal the company’s performance to some extent.

But data is not available all the time. We notice that banks and databases like Center for Research of Security Prices, Compustat, and New York Stock Exchange Trading and Quotation high-frequency database are the main sources of intelligence. And data usually is only available to researchers due to privacy or other needs. This makes it hard for other people to repeat the trial and identify the significance of the findings. Thus, the construction of open databases which could be a source as well as a sharing platform of data should be put on the agenda. The more intelligence we can collect, the more value we can create in the area of risk management. In a word, proper use and sharing of multimodal data will benefit us a lot.

B. Analysis Model

Recently, some new methods have emerged in the research on enterprise risk. Here is a discussion on the possible research directions.

1) Graph Neural Network Model: Graph data contains very rich relational information, so graph can represent higher dimensional information. Graph neural network is a new neural network paradigm emerging in recent years. It combines the advantages of graph and deep learning to achieve deeper information correlation. For example, In [288] the authors use high-order graph attention representations to predict loan risk, which provides a new dimension for loan risk assessment and is more trustworthy. At present, the research on graph neural network of enterprise risk is still in its infancy. In this field, combining graph neural network, more meaningful relationship information can be extracted from the dimensions
of inter-enterprise, enterprise data and senior management relationship.

2) Advanced Natural Language Processing Model: In recent years, encoder and decoder mechanisms and have made a splash in the field of natural language processing when combine with attentional mechanisms, a technique that allows models to focus on important information and fully learn to absorb it. Natural language contains a wealth of information about the enterprise, but because its unstructured penguins are difficult for machines to understand, little previous research has analyzed the text at a fine-grained level. By incorporating a priori information about the enterprise through natural language pre-training models, textual information can be better mined and used for enterprise risk analysis. On the other hand, autoencoder is a good choice for feature extraction. The advantage of autoencoder is that it can determine the complex relations on the dataset. There have been some researches in this field. [175] uses stacked autoencoder to extract features, and then the softmax classifier is applied to classify the companies. [174] use a dynamic graph-based attentional neural network to help predicting risk guarantee relationship.

3) Time Series Network Model: The development of enterprises is dynamic, and most of the current research methods use static data, which can hardly reflect the dynamic change information. By introducing time-series information, the model can better determine the risk of enterprises from the time dimension. Most of the current time series are based on traditional measurement methods [289] [152]. With the development of deep learning, the time-series network such as recurrent neural network and long short-term memory network can better determine the risk of enterprises from the time dimension. So the time-series method combines with deep learning methods will be a promising research direction.

4) Fusion Model: For enterprise risk analysis, heterogeneous convergence is a promising direction. A single model often has advantages in one aspect, such as the processing of financial information, textual information, relational data, and time-series data, all of which can reflect enterprise risk information to a certain extent, but all of which focus on one point. If we can combine the advantages of each model and model the multidimensionality of enterprise risk, we will have a more comprehensive presentation of enterprise risk and realize the leap from point to surface, which will improve the analytical performance of the model even more.

C. Contagion Mechanism

In the context of the financial network, many studies have put forward methods and measures to control risks under sub-networks. For example, PD model is introduced in [98], which is a dynamic model that combines credit risk technology with the contagion mechanism of risk network between banks to quantify systemic risk. [272] using a unique data set containing actual bilateral exposures, it analyzes how contagion spread within the Italian interbank market. Through the network of connections between institutions, it is of great significance to study risk transmission among institutions.

Considering the limitations of a single network layer, [96] contributes to the study of inter-bank risk transmission mechanism by analyzing the systemic risk contribution of the four exposure layers of the inter-bank network (derivatives, securities cross-holdings, foreign exchange and interbank deposit and loan markets). In future studies, other potentially important sources of infection, such as overlapping portfolio networks and capital liquidity risks, can be considered. Secondly, how to introduce more network layers is also an important topic for future research.

[99] estimates the risk of nodes in the financial network, such as banks, by using network measures of the inter-bank liability network (e.g., DebtRank). The research proves that the risk transmission of banks can be reduced by increasing the transparency between relevant nodes. However, the real risk network is a multiple network, and the same group of financial agents is connected by various networks, including asset-liability network, derivative network and collateral network. Therefore, how to make comprehensive use of all the information of risk network and derive a practical and effective risk network is a technical challenge in the future. At the same time, it is mentioned in the paper that by increasing transparency, risk nodes can reduce their credit risks by being prevented from lending, while risk-free banks can increase their risks by lending. Finally, it can realize the effect of evenly distributing risks through the network. Therefore, a more market-driven mechanism to obtain the same self-organizing critical regulatory dynamics is subject to further investigation.

[262] provides a framework to study the relationship between financial network architecture and the potential for systemic failure due to counterparty risk contagion. The findings highlight the fact that even if banks fully internalize the impact of their lending (and risk-taking) decisions on direct creditors, they fail to take into account the fact that their lending decisions may also expose many other banks (such as their creditors’ creditors) to greater risk of default, thus proving the existence of financial network externalities. The externalities of financial networks and their attendant inefficiencies also mean that there is scope for government intervention to improve welfare. The analysis of optimal strategy in this kind of model is also of high research value.

D. Risk Interpretability

Despite the success in enterprise analysis, most existing deep learning methods are treated as black-boxes and cannot explain “Why the model to make certain predictions?”, which limit their application in critical areas such as finance and security. Several techniques are proposed to explain the deep models for image and text data. For example, by studying the gradients or weights [290], occluding different input features [291] and exploring the meaning of hidden neurons [292], we can get input-dependent and input-independent explanations. Graph Neural Networks (GNNs) have become increasingly popular since many real-world data are represented as graphs, such as social networks, chemical molecules, and financial data. Several approaches are proposed to explain the predictions of GNNs, such as SubgraphX [293], RC-Explainer [294], SE-GNN [295] and ProtGNN [296], etc. These methods are developed from different angles and provide different levels of explanations.
Although various self-explanatory machine learning, deep learning model have been proposed recently, they will face performance and efficiency limitations when deployed on real-world tasks like bankruptcy prediction and credit evaluation. Therefore, developing a risk detector that simultaneously gives high-quality predictions and explanations is imperative to the deployed enterprise analysis systems. However, such has not yet been explored and demands further investigation.

VII. Conclusion

Enterprise risk is the threat of loss suffered by every link of the enterprise’s reproduction experience activities. No matter in different experience processes such as purchasing, production and sales, or in different functional areas such as planning, organization and decision-making, enterprises may encounter risks, which have an important impact on enterprises. Therefore, enterprise risk research has always been the focus of academic research. In this study, we systematically reviewed 319 articles on enterprise risk published between 1968 and 2022, covering various fields such as finance, management science and computer science. Our goal is to provide a systematic and comprehensive summary. We divided enterprise risk into six risk types: bankruptcy risk, credit risk, guarantee risk, financial risk, supply chain risk and systemic risk. We systematically studied the concept definition, research model, data type and model evaluation index of the first four risk types. We divide the research models into statistical and econometric models, machine learning models, deep learning models and mixed models. We find that most scholars adopt statistical and econometric models and machine learning models when studying enterprise risks. At the same time, we divide the data used in the research into financial data, non-financial data and relational data. The research finds that financial data is the most commonly used index in the study of enterprise risk.

Different from previous studies, we systematically introduced the four types of models and the evaluation indicators of the models. In addition, we also introduced the application of enterprise risk management in three aspects: corporate governance, financial institutions and government supervision. In the process of combing through the relevant literature, we have also obtained some valuable information:

- We have compiled most of the research literature related to enterprise risk from 2010 to 2021, and we can see that there is a significant increase in the literature from 2013-2015, and 2019-2021, and that the tough economic situation is likely to be an important factor contributing to this phenomenon.
- Before 2019, statistical models and machine learning models dominate in the study of enterprise risk, while with the popularity and development of deep learning techniques in the financial field, more and more scholars try to use deep learning methods to study enterprise risk after 2019.
- Enterprise risk research mainly involves financial datasets, but also contains non-financial datasets, relational data, and hybrid data, which are characterized by diversity.
- Approximately 58.67% of scholars engaged in enterprise risk research are from the field of finance, 28% from the field of management, and 13.33% from the field of computer science, which indicates that enterprise risk issues are of interest to scholars from multiple fields, and more scholars from more fields may join this research in the future.

With the in-depth study of enterprise risk, it will help companies to develop strategies to resist risks and build a healthy and stable economic environment.

ACKNOWLEDGMENTS

REFERENCES

[1] M. Billio, M. Getmansky, A. W. Lo, and L. Pelizzon, “Econometric measures of connectedness and systemic risk in the finance and insurance sectors,” Journal of financial economics, vol. 104, no. 3, pp. 535–559, 2012.
[2] W. Chen, Z. Li, and Z. Xiao, “On credit risk contagion of supply chain finance under covid-19,” Journal of Mathematics, 2021.
[3] Y. Zhao, S. Wei, Y. Guo, Q. Yang, Q. Li, F. Zhuang, J. Liu, and G. Kou, “Combining intra-risk and contagion risk for enterprise bankruptcy prediction using graph neural networks,” arXiv preprint arXiv:2202.03874, 2022.
[4] A. Abduraimova, “Contagion and tail risk in complex financial networks,” Journal of Banking & Finance, p. 106560, 2022.
[5] R. A. Collins and R. D. Green, “Statistical methods for bankruptcy forecasting,” Journal of Economics and Business, vol. 34, no. 4, pp. 349–354, 1982.
[6] Z. Hua, Y. Wang, X. Xu, B. Zhang, and L. Liang, “Predicting corporate financial distress based on integration of support vector machine and logistic regression,” Expert Systems with Applications, vol. 33, no. 2, pp. 434–440, 8 2007.
[7] C. Yin, C. Jiang, H. K. Jain, and Z. Wang, “Evaluating the credit risk of smes using legal judgments,” Decision Support Systems, vol. 136, p. 113364, 2020.
[8] C. Wang, F. Yu, Z. Zhang, and J. Zhang, “Multiview graph learning for small-and medium-sized enterprises’ credit risk assessment in supply chain finance,” Complexity, vol. 2021, 2021.
[9] E. Letizia and F. Lillo, “Corporate payments networks and credit risk rating,” EPJ Data Science, vol. 8, no. 1, p. 21, 2019.
[10] Z. You, X. Chi, W. Gang-Jin, and Y. Xin-Guo, “Predicting china’s sme credit risk in supply chain finance based on machine learning methods,” Entropy, vol. 18, no. 5, pp. 195, 2016.
[11] B. Sang, “Application of genetic algorithm and bp neural network in supply chain finance under information sharing,” Journal of Computational and Applied Mathematics, vol. 384, pp. 113170, 3 2021.
[12] G. Kou, Y. Xu, Y. Peng, F. Shen, Y. Chen, K. Chang, and S. Kou, “Bankruptcy prediction for smes using transactional data and two-stage multiobjective feature selection,” Decision Support Systems, vol. 140, p. 113429, 2021.
[13] E. Tobbac, T. Bellotti, J. Moeyersoms, M. Stankova, and D. Martens, “Bankruptcy prediction for smes using relational data,” Decision Support Systems, vol. 102, pp. 69–81, 2017.
[14] T. Hosaka, “Bankruptcy prediction using imaged financial ratios and convolutional neural networks,” Expert systems with applications, vol. 117, pp. 287–299, 2019.
[15] K. Lee, D. Booth, and P. Alam, “A comparison of supervised and unsupervised neural networks in predicting bankruptcy of korean firms,” Expert Systems with Applications, vol. 29, no. 1, pp. 1–16, 2005.
[16] F. Mai, S. Tian, C. Lee, and L. Ma., “Deep learning models for bankruptcy prediction using textual disclosures,” European journal of operational research, vol. 274, no. 2, pp. 743–758, 2019.
[17] I. D. Dichev, “Is the risk of bankruptcy a systematic risk?” the Journal of Finance, vol. 53, no. 3, pp. 1131–1147, 1998.
[18] R. Gropp, C. Gruendl, and A. Guettler, “The impact of public guarantees on bank risk-taking: Evidence from a natural experiment,” Review of Finance, vol. 18, no. 2, pp. 457–488, 2014.
[19] J. Ming and X. Ming, “Determinants of the guarantee circles: The case of chinese listed firms,” Pacific-Basin Finance Journal, vol. 20, no. 1, pp. 78–100, 2012.
A. Maghyereh and H. Abdoh, “The effect of structural oil shocks on S. Angilella and S. Mazzu, “The financing of innovative smes: A M. S. Beasley, R. Clune, and D. R. Hermanson, “Enterprise risk M. Vassalou and X. Yuhang, “Default risk in equity returns,” M. S. Munir, M. S. S. Jajja, K. A. Chatha, and S. Farooq, “Supply X. Chao and M. Xiaojing, “Bankruptcy prediction using deep learning approach based on borderline smote,” Artificial Intelligence, vol. 33, pp. 406–4075, 2021.

J. Hilscher and M. Wilson, “Credit ratings and credit risk: Is one M. Xu and C. Zhang, “Bankruptcy prediction: the case of japanese J. El Baz and S. Ruel, “Can supply chain risk management practices S. Ai, P. L. Brockett, and W. Tianyang, “Optimal enterprise risk management and decision making with shared and dependent risks,” Journal of Risk and Insurance, vol. 84, no. 4, pp. 1127–1169, 5 2017.

M. S. Uddin, G. Chi, M. A. M. A. Janabi, and T. Habib, “Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability,” International Journal of Finance & Economics, vol. 25, no. 1, pp. 126–137, 2020.

C. Xiaohong, W. Xiaoding, and D. D. Wu, “Credit risk measurement and early warning of smes: An empirical study of listed smes in china,” Decision Support Systems, vol. 49, no. 3, pp. 301–310, 6 2010.

Y. Xu, X. Yuan, and G. Zheng, “Internet searching and stock price crash risk: Evidence from a quasi-natural experiment,” Journal of Industrial Economics, vol. 141, no. 1, pp. 255–275, 2021.

B. L. Wilson and R. Shanda, “Bankruptcy prediction using neural networks,” Decision support systems, vol. 11, no. 5, pp. 545–557, 6 1994.

D. Delen, C. Kuzey, and A. Uyar, “Measuring firm performance using financial ratios: A decision tree approach,” Expert systems with applications, vol. 40, no. 10, pp. 3970–3983, 2013.

H.-L. Chen, B. Yang, G. Wang, J. Liu, X. Xu, S.-H. Wang, and D.-Y. Liu, “A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method,” Knowledge-Based Systems, vol. 24, no. 8, pp. 1348–1359, 2011.

Y. Ching-Chiang, L. Fengyi, and H. Chih-Yu, “A hybrid kmv model, random forests and rough set theory approach for credit rating,” Journal of Knowledge-Based Systems, vol. 33, pp. 166–172, 9 2012.

C. Chun-Ling, “Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction,” Information Sciences, vol. 236, pp. 174–185, 7 2013.

A. Qazi, A. Dickson, J. Quigley, and B. Gaudenzi, “Supply chain risk network management: A bayesian belief network and expected utility based approach for managing supply chain risks,” International Journal of Production Economics, vol. 196, pp. 24–42, 2 2018.

D. L. Eckles, R. E. Hoyt, and S. Müller, “The impact of enterprise risk management on the marginal cost of reducing risk: Evidence from the insurance industry,” Corporate Governance: Disclosure, 2014.

M. Xu and C. Zhang, “Bankruptcy prediction: the case of japanese listed companies,” Review of accounting studies, vol. 14, no. 4, pp. 334–358, 2009.

R. E. Crott, A. Guettler, and C. Gründl, “The impact of public guarantees on bank risk taking: Evidence from a natural experiment,” Banking & Financial Institutions eJournal, 2010.

S. M. Bartram, G. W. Brown, and J. E. Hund, “Estimating systemic risk in the international financial system,” Journal of Financial Economics, vol. 86, no. 3, pp. 835–869, 2007.

J. Hilscher and M. Wilson, “Credit ratings and credit risk: Is one measure enough?” Management science, vol. 63, no. 10, pp. 3414–3437, 10 2017.

A. Eisdorfer and P.-H. Hsu, “Innovate to survive: The effect of technology competition on corporate bankruptcy,” Financial Management, vol. 40, no. 4, pp. 1087–1117, 2011.

Y. Zhu, F. Yang, and W. Ye, “Financial contagion behavior analysis based on complex network approach,” Annals of Operations Research, vol. 268, pp. 93–111, 2018.

B. Wu and T. Duan, “Nonlinear dynamics characteristic of risk contagion in financial market based on agent modeling and complex network,” Complex., vol. 2019, pp. 2946018:1–2946018:12, 2019.

B. Imbierowicz and C. Rauch, “The relationship between liquidity risk and credit risk in banks,” Journal of Banking & Finance, vol. 40, pp. 242–256, 3 2014.

M. S. S. Jajja, K. A. Chatha, and S. Farooq, “Impact of supply chain risk on agility performance: Mediating role of supply chain integration,” International Journal of Production Economics, vol. 205, pp. 118–138, 11 2018.

M. S. Munir, M. S. S. Jajja, K. A. Chatha, and S. Farooq, “Supply chain risk management and operational performance: The enabling role of supply chain integration,” International Journal of Production Economics, vol. 227, p. 107667, 2020.

J. El Baz and S. Ruel, “Can supply chain risk management practices mitigate the disruption impacts on supply chains’ resilience and robustness? evidence from an empirical survey in a covid-19 outbreak era,” International Journal of Production Economics, vol. 233, p. 107972, 2021.

Y. I. Wang, J. Li, D. A. Wu, and R. Anupind, “When ignorance is not bliss: An empirical analysis of sub-tier supply network structure on firm risk,” Ross: Technology & Operations (Topic), 2020.
