Heliyon 10 (2024) e39770

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Advancing financial resilience: A systematic review of default prediction models and future directions in credit risk management

Jahanzaib Alvi *, Imtiaz Arif, Kehkashan Nizam

Department of Business Administration, IQRA University Karachi, Pakistan

ARTICLE INFO

Keywords:
Default prediction models (DPM)
Credit risk management
Systematic literature review
Financial stability
Hybrid modeling approaches

ABSTRACT

This research presents a systematic review of a substantial body of high-quality research articles on Default Prediction Models published from 2015 to 2024. It is a comprehensive analysis of a DPM wide spectrum approaches including Textual Models, Systematic Review Studies, Hybrid Models, Intelligent Models and Statistical Models. The reason behind this study is rooted in the critical need to mitigate and understand the credit default risk that poses a significant threat to financial stability worldwide. By employing an evidence-based approach and methodological rigorously, this research critically evaluates the gaps, effectiveness and evolution in existing DPM methodologies. It is not only synthesized the current landscape of DPM study but also identified the direction for the future research, by offering novel insights and bridging theoretical gaps for enhancing the strategies of credit risk management. This study stands out by focusing on high citation research from top tier publishers, ensuring the quality and relevance of its analysis. The findings of this study have profound implications for stakeholders across the financial sector, including bankers, investors, regulatory bodies, and researchers. It aims to advance financial stability by providing a comprehensive overview of DPM advancements and pointing towards areas that require further exploration. By doing so, it contributes significantly to the development of more effective and sophisticated DPM strategies, thereby enhancing the robustness of financial institutions against potential defaults.

1. Introduction

Bankruptcy or financial default refers to a condition when an entity has failed to meet its financial obligations (any outstanding debts or regular payments that a party must make) in timely manners thus the entity needs to undergo reorganization of debt to liquidate its assets [1]. Indeed, bankruptcy can cause all the stakeholders in several or mild contexts but it could also be disastrous if it creates domino impact like 2008 financial crisis (which was started due to subprime mortgage hedged to Lehman's Brothers and their default made more firms default ultimately domino effects of financial defaults globally) [2]. In light of worse past experience, estimating defaults before its occurrence gauge the focus of financial analysts [3]. Following this, several statistical and AI techniques were proposed to solve this problem. Further, default prediction through these models still remains with significant room for study and come up with new insights to solve this problem. Such models are quite useful for the bankers, Investors, regulatory bodies and rest of stakeholders [4,5].

The research problem in this domain is deeply rooted in the complex relationship between financial integration and globalization,

* Corresponding author.
E-mail addresses: Janazaib.alvi@iqra.edu.pk (J. Alvi), arif.i@iuk.edu.pk (I. Arif), kehkashan.60003@iqra.edu.pk (K. Nizam).

https://doi.org/10.1016/j.heliyon.2024.e39770
Received 23 April 2024; Received in revised form 21 October 2024; Accepted 23 October 2024
Available online 24 October 2024

which has significantly increased the intricacy of the financial system. Central to this issue is the considerable credit default risk borne by financial institutions, primarily arising from their involvement in credit financing and mortgage lending [6]. This specific type of credit risk exerts a substantial and direct influence on the financial statements of these institutions [4,5]. The primary focus of our research is to understand and analyze how global financial dynamics, particularly in the context of credit default risk, affect the financial stability and reporting accuracy of financial institutions. This aspect of credit default risk represents a critical area for in-depth investigation, underscoring its significance in the broader landscape of financial research. Moreover, assessment of credit risk (by developing default models) is required against the real economy loans financed by the financial institution as per the Basel accord guidelines. The development of loan portfolios assesses diversification effect which also indicates default correlation against each credit asset within the portfolio [7,8]. As per the Basel II and Basel III accord, financial institutions are recommended to establish an Internal-Rating Based system with regular back-testing (Annually). Lagging of IRB system is that the model only covers partial risk (idiosyncratic risk) because of the assumption of "Fine-Grained Loan Portfolio - FGLP" and because of this lagging, the IRB only can measure the unexpected credit losses on loan to cover the capital requirement (EAT and LGD). The model is partially based on the qualitative factors of borrowers which majorly ignores loan portfolio composition.

The motivation of this study was began after observing the unprecedented scale and impact of financial defaults, underscored by the 2008 financial crisis, have intensified the imperative for advanced Default Prediction Models (DPMs). This study is driven by the critical need to navigate the increasingly complex interplay between global financial integration and credit default risks, which significantly threatens financial stability and the integrity of financial reporting in institutions engaged in credit financing and mortgage lending. Our research delves into the evolution, effectiveness, and gaps in existing DPM approaches, encompassing statistical, AI-based, and hybrid methodologies. By examining these models through a lens that balances global financial dynamics against the backdrop of regulatory frameworks and the real economy, this study endeavors to fortify financial forecasting and risk management strategies, thereby contributing to the overarching goal of enhancing financial stability and mitigating the ramifications of potential financial crises.

The high default rate significantly impacts not only stakeholders but also the broader economy [9]. Consequently, the need for developing default prediction models (DPM) gain high attention from researchers, economists, investors and so on. Two mega streams of research have emerged in this area: (1) Feature selection and (2) Technique selection [10–12]. Both of these things are important to predict the credit default of a firm. Even comparison metrics can be established to illustrate multiple models and its accuracy (comparison of statistical, intelligent and hybrid models) [13]. Some of the initial research or it may be said as foundation of this topics are given in Table 1.

In 21st century, AI algorithms (machine learning and deep learning) are gaining massive popularity in research society. Artificial Neural Networks (ANN) & Enhanced Neural Network (ENN) are the top ranked tool for DPM models [19]. ANN & ENN are being adopted due is its simplest algorithmic architecture plus in several studies it has further augmented with propagation algorithm of ANN for DPM models among others [20,21]. However, the ANN has several issues with its architecture such as it incorporates undesirable features with extensive computation with falsification error (against of logic and financial theories) in response to various input values or features' data [10,22,23]. Moreover, an ANN with small dataset when ANNs are known to need large samples for optimal performance [24]. Current studies are being conducted in two domains (1) Statistical DPM, (2) Intelligent DPM which is divided in two broad categories as Machine Learning Algorithms and Deep Learning Algorithms [25,26]. Following are the details of recent trends and emerging development; table explicitly illustrate aforementioned.

Table - 2 reveals a dynamic landscape in credit risk modeling witnessing prominent trends and emerging developments. A notable shift is the increasing adoption of hybrid models, which combine the strengths of different techniques, such as integrating machine learning algorithms with traditional statistical methods or optimization algorithms. This approach aims to create more robust and accurate assessments by leveraging the advantages of each individual method. Simultaneously, the use of deep learning models, especially LSTM and CNNs, is gaining momentum. These models are particularly well-suited for capturing complex patterns and non-linear relationships within financial data, potentially leading to more precise predictions. Complementing this trend is the continued reliance on ensemble methods like Bagging and Boosting, which enhance prediction accuracy and model robustness by combining multiple models. Furthermore, there's a growing emphasis on explainable AI (XAI), driven by the need for transparency and interpretability in credit risk models. This push for explain ability reflects the demand for models that not only provide accurate predictions

Table 1
Early studies with Preliminary proposed models.

Study (in chronological order)	Accounting or market based?	Estimation technique
[10]	Accounting	Discriminant analysis
[11]	Accounting	Regression analysis, (logit)
[14]	Accounting	Regression analysis, (probit)
[15]	Mix	Hazard analysis
[16]	Market	Hazard analysis
[17]	Mix	Hazard analysis
[18]	Mostly Accounting	Data Envelopment analysis (additive)

Note: Table-1 was extracted from SLR taken in this research, the idea was to give an insight of DPM and its origins. Above table not only represents Studies Study in Chronological Order but also explains details of taken features in each study (exhibited in the last column). Column 1 illustrates the author of the study with the highest citation around the globe, column 2 the nature of the study, and column 3 showcases estimation techniques used by the author.

Table 2
Emerging trends and recent developments in credit risk assessment.

Author (s) & Year	Methodology/Model	Data Used	Key Findings/Contributions	Limitations
[27]	Hybrid model combining LSTM and Attention Mechanism	Financial data of listed companies in China	Improved prediction accuracy compared to traditional models; Effectively captures long-term dependencies in financial data	Limited to the Chinese market; Generalizability to other markets needs further investigation
[28]	Fuzzy logic-based credit risk assessment model	Data from a Pakistani commercial bank	Provides a robust framework for handling uncertainty in credit risk assessment; Offers practical implications for banking institutions	Reliance on expert knowledge for fuzzy rule definition; Scalability to large datasets may be challenging
[29]	Ensemble learning approach using Random Forest and Gradient Boosting	Data from Nigerian banks	Demonstrated superior performance compared to individual models in predicting credit defaults	Focus on a specific geographical region; Further validation required for different economic contexts
[30]	Deep learning model based on Convolutional Neural Networks (CNN)	Financial and textual data of loan applicants	Highlights the potential of CNNs for extracting complex patterns from diverse data sources	Computational complexity and interpretability challenges associated with deep learning models
[26]	Hybrid model incorporating machine learning and expert judgment	Data from a microfinance institution in India	Emphasizes the value of combining quantitative and qualitative information for more accurate credit risk assessment	Subjectivity inherent in expert judgment; Potential bias in data collection
[31]	Machine Learning Techniques (Logistic Regression, SVM, Decision Tree)	Data from Indian banks	Compared the performance of different ML models; Identified key factors influencing credit risk	Limited to traditional ML models; Did not explore more advanced techniques like deep learning
[32]	Hybrid model using Support Vector Machines and Genetic Algorithms	Financial data of publicly traded companies	Improved feature selection and model optimization through the use of Genetic Algorithms	Computational cost associated with Genetic Algorithms; Requires careful parameter tuning
[33]	Ensemble learning using Bagging and Boosting	Credit bureau data and macroeconomic indicators	Enhanced prediction accuracy and robustness through ensemble techniques	Potential overfitting; Requires a large and diverse dataset for effective training
[34]	Artificial Neural Networks (ANN)	Financial data of SMEs in Nigeria	Showed the effectiveness of ANN in capturing non-linear relationships in credit risk data	Black-box nature of ANN; Difficulty in interpreting the model's decision-making process
[35]	Machine Learning for Credit Scoring	German Credit Data	Comprehensive comparison of various ML models for credit scoring; Provides insights into model selection	Limited to a specific dataset; Generalizability needs further validation
[36]	Deep Learning and Feature Engineering for Credit Risk Prediction	Data from Ghanaian Banks	Demonstrates the importance of feature engineering in improving the performance of deep learning models	Limited geographic scope; Focus on a specific banking sector
[37]	Hybrid Model using ANN and Genetic Algorithm	Financial data of Indian companies	Showed the effectiveness of hybrid approaches in optimizing model parameters and improving prediction accuracy	Computational complexity; Requires expertise in both ANN and Genetic Algorithms
[38]	Explainable AI for Credit Risk Assessment	Simulated and Real-world Credit Data	Emphasizes the need for transparency and interpretability in credit risk models; Explores techniques for explaining model predictions	Relatively new area of research; Further development of explainable AI methods is needed

but also offer insights into their decision-making processes, fostering greater trust and understanding. Finally, researchers are increasingly leveraging a diverse range of data sources, going beyond traditional financial statements to include credit bureau data, macroeconomic indicators, and even textual information. This broader perspective allows models to capture a more holistic view of borrowers' creditworthiness, potentially leading to more accurate risk assessments. These trends collectively underscore the ongoing evolution of credit risk modeling and the pursuit of more sophisticated and reliable prediction methods.

Many research were conducted to provide a robust model for DPM. The aim of this research is to make analyses on this research conducted in the domain of DPM and contribute by giving an assessment on the strength and weakness of proposed models. This study will analyze recent literature that will also exhibit the trends and future aspects for research more in this domain. One study conducted a systematic review based on thirty research articles focused on Neural Networks (NN). The review examined the strengths and weaknesses of NN algorithms, particularly in classification problems. It highlighted issues such as feature selection, sample size, and the impact of input variables on the final results [39]. A detailed systematic review covering articles published from 1968 to 2005 on classification problems for both statistical and AI techniques [26]. He commented on the method used (classification methods) in the articles with the sequence of authors, technique, and results. He further compared DPM with accuracy rate of every model in accordance with the dataset and used features. A proposed systematic review focused on soft computing algorithms for DPM [40]. Another study examined statistical techniques used to predict corporate defaults [41].

Our study is different from said examples as above, such as we employed methodology of systematic review. Secondly, we focused on top-tier publishers (measured by the volume of citation on each article), high quality researches and recent thesis based on default prediction models, we further covered 5 mega domains following such as, (1) Statistical Models, (2) Intelligent Models, (3) Hybrid Models, (4) Systematic Review Studies, and (5) Textual Models for DPM (segregated on the basis of used research methodology in the each article) [42,43]. We conducted our analysis using a substantial collection of articles published in high-quality journals between 2015 and 2024.

We employed a methodology of systematic review based on a five-step SLR process [44]. The core purpose of SLR studies is to extract out the evidence based on synthesized assessment on studied topics. SLR has many advantages from various perspective such as it gives detailed insights about the topic by mitigating biases and more generalizability, it provides a detailed comparison not only study to study but also time by time, it further gives healthy room for study in the future by identifying a research gap [45,46]. Considering the importance of SLR studies, researchers have applied this methodology across various domains to identify grey areas, gaps, and generalizability. For instance, SLR has been applied in software engineering research to consolidate findings and highlight emerging trends [47]. In the health domain, SLR has been used to explore obesity in children, offering insights into its prevalence and contributing factors [48]. Similarly, in psychology, SLR has been crucial in examining psychological health, synthesizing data from numerous studies to better understand mental health issues [49].

In the finance domain, SLR has been employed to study serious games in finance, demonstrating their impact on financial education and decision-making [50]. Furthermore, SLR has been instrumental in the management of technical debts within financial systems, providing a structured analysis of how these debts influence financial stability [44]. Lastly, SLR has been used to investigate models of financial defaults, focusing on the methodological issues associated with default prediction models (DPM), thus offering a clearer understanding of the limitations and challenges in predicting financial failures [51]. This broad application of SLR across diverse fields underscores its importance as a methodological tool for synthesizing research findings and identifying areas for future investigation.

This research marks a significant advancement in the field of Default Prediction Models (DPM) through a comprehensive systematic review of over 250 high-quality articles from 2015 to 2024. It critically evaluates a diverse range of DPM approaches Statistical, Intelligent, Hybrid, Systematic Review Studies, and Textual Models using a methodologically rigorous, evidence-based analysis. This study not only synthesizes the current state of DPM research but also highlights potential future directions, addressing unexplored areas and theoretical gaps. Its focus on quality studies (Measure by citation score - from top-tier publishers ensures the relevance and excellence of its findings [42,43]. By offering new insights for effective credit risk management, this study serves as a crucial resource for stakeholders across banking, investment, regulatory, and research domains, aiming to enhance financial stability and develop more sophisticated DPM strategies.

This article will be beneficial for investors, bankers and researchers by providing a detailed comparative analysis and grey area to study in default prediction models (DPM), further this article will highlight all studies conducted in between 2015 and 2024 with all five domains as discussed above in prediction modeling. Remained paper is designed as chapter-2 'Methodology' which describes all the five step process of SLR methodology, then chapter-3 'Results and Discussion' where we will explain findings driven by the SLR, and chapter-4 'Conclusion' will be illustrating summary of the research, contribution, and guideline for future research.

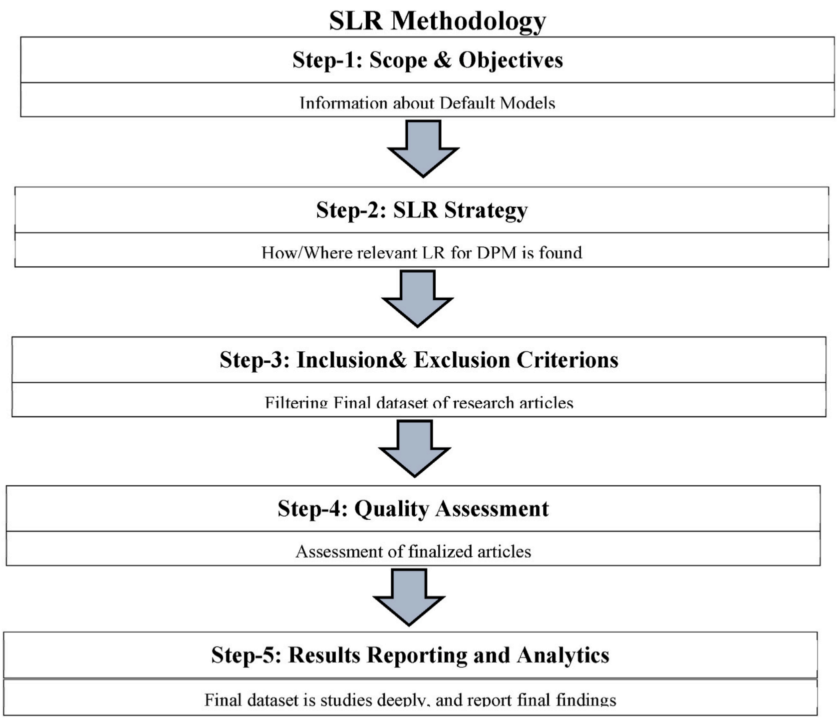

Figure-1. Slr methodology.

2. Methodology

We adopted a five-step SLR process [44,51]. Steps are defined as follows, (1) Scope & Objectives, (2) SLR Strategy, (3) Inclusion & Exclusion Criterions, (4) Quality Assessment and (5) Results Reporting and Analytics. Further, we have presented the steps in figure-1 as below. Further, this articulated synthesis work will explore prior literature on default prediction models.

2.1. Scope and objectives of SLR

Core object of this synthesis is to evaluate the role of statistical and AI (Machine Learning & Deep Learning) models to predict financial distress/bankruptcy/credit default. This research focused on various methods used to predict default which as classified in five broad categories such as given below.

Table 3 present the organization of default prediction literature into five domains for comprehensive analysis. Hybrid Models blend statistical methods with intelligent techniques, offering a comparative look at their accuracy. Intelligent Models are subdivided into Deep Learning (DL) and Machine Learning (ML) Algorithms, utilizing classification algorithm Deep Learning Models, Decision Trees, Support Vector Machines (SVM), and Neural Network (NN), which employed advanced tools of regression such as Long Short-Term Memory (LSTM) and Convolution Neural Network (CNN). Statistical Models serve as basic tools employing techniques such as Logit-Probit Models and Multivariate Discriminant Analysis (MDA). SLR examines literature reviews in default modeling to integrate previous finding with the present research. Lastly, Textual Models represent an innovative approach, predicting defaults through non-financial data sources like auditor comments and market news. This framework enables the understanding of the default prediction by using diverse methodologies.

2.2. Search strategy

We used Google Scholar as our main search engine with keywords entitled as "Credit Default: Probability of Default: Financial Distress: Predicting Default AI: ML: DL: Financial Ratios: Feature Selections". We further used Scopus database, Research Gate and SSRN to extract relevant articles to use in our SLR study. We set the range in between 2015 and 2024 for only articles that contain keywords and fall in the domain of this SLR study. As a result, we extracted a significant number of research papers published over the last 7–9 years, particularly in the said domain. Indeed, if we change the keywords some of more and different articles could be sought and taken into account but we believe that this study will be very beneficial to produce quality assessment of DPM in recent world.

2.3. Inclusion and exclusion criteria

The inclusion criteria for this synthesis work encompassed a wide range of documents within the domain of DPM, such as research articles, theses, and working papers even those without citations but with significant relevance to DPM. Systematic review studies were also considered important to establish facts based on these studies in conjunction with all four categories. However, articles considered malicious, those focusing on ML or DL for performance evaluation purposes, topics like Credit Default Swap (CDS), or non-English publications were excluded. The entire process was meticulously organized using an Excel sheet, leading to the development of a Spider-Web database that houses over 250 articles. Each article was thoroughly examined and collectively analyzed, as demonstrated in step 5, to ensure a comprehensive and meaningful synthesis.

2.4. quality assessment

The quality of systematic literature research can be assessed through two ways. (1) Source of published material and (2) Relevancy with the topic which is being studied. Therefore, we have only taken articles from reliable & authentic sources as mentioned above further almost 75 % of our articles belongs to top class publishers such as Elsevier, Emerald, Springer, Taylor & Francis, IEEE, Wiley, Sage, InderScience and so on (Details are given in the step-5). As mentioned above in inclusion exclusion criteria, we have selected only those articles which fall in our study domain, so technically we have preliminarily eliminated all those articles which do not cover DPM in any case. We have exhaustively acquired comprehensive datasets of articles and dissected them in excel sheets and then we sketch columns for each head (Details could be given upon request), and at the end we analyzed Abstract, Keywords, Data Collection, Data

Table 3
Classification of studies.

META	Total Articles	Total Citations	Per Article Citation
Hybrid Model (Mixed of Pure Statistical Models and ML Algorithms)	39	989	25.36
Intelligent Model (Only ML and DL Algorithms)	104	3682	35.40
Statistical Model (Only Statistical Models)	49	956	19.51
Textual Model (Based on Textual Dataset e.g. Accounting Disclosures)	21	755	35.95
Systematic Review (Studies on Default Prediction)	37	1697	45.86
Total	250	8079	162.09

Note: Table – 3 is author's illustration.

Source, Research Technique, Evaluation Technique and last results and findings. In case any article which did not qualify for any reason we omitted it at once. In accordance with statements covering the second feature of quality assessment of SLR we believe that collected articles are satisfactory and qualify for SLR.

2.5. Data reporting and analysis

We have established our findings on the basis of collected articles, we emphasis on the citation of the articles with respect to domain of the study, indeed [10,52,53], [54], and [55]are the most cited studies in the past, but our this SLR is based on the recent studies in the discipline of default modeling. Moreover, the citations are collected from Google Scholar manually and are placed in the final spider-web, below is the detailed analysis of this study. We have reported results in.

3. Results and discussions

Building on the categorization provided in Table 1, which organizes articles into five domains based on their scope, we will now delve into a detailed analysis of each category, starting with the "Hybrid Model." This section presents the findings from the articles scoped within the Hybrid Model domain.

3.1. Hybrid model analysis

Table 4 illustrates details of Hybrid Models. hybrid models are the mixed of "Core Statistical Techniques" and "AI Techniques (ML and DL)", the analysis reveals a diverse application of hybrid models in financial research, with each technique's estimated accuracy varying widely. Techniques range from Big Data Analytics with accuracies between 87 % and 89 %, to Neural Networks and Self-Organizing Maps achieving as high as 96.15 %. Notably, hybrid models combining machine learning with traditional methods, such as Artificial Neural Networks (ANN) and Logistic Regression, often report superior accuracy to traditional models alone. The scope of accuracy spans from specific (e.g., 90.30 %–82.79 % for Feedforward Neural Network with Cuckoo Search Algorithm) to qualitative assessments like "high" or "superior to traditional models" for many approaches.

Table 5 explains the most common model validation methods when use hybrid models. Validation methods include Grid Search Optimization and various interpretability tools like Shapley Additive Explanations (SHAP), indicating high accuracy or variability depending on the technique. Specific metrics include Precision Recall, ROC and AUC reveal accuracies ranging from 80 % to 91 %. The performance of comparison methods and advance statistical testing use underscores the commitment to evaluate the effectiveness of model.

Table 6 emphasis on the most common feature that selected for the prediction of default risk through the mixed model approach. This study emphasizes the wide range importance of variable in enhancing the predictive power of hybrid models. The financial data and ratios dominate the variables used, featuring the seven papers. The indicators related to macroeconomics and economic as well as specific metrics such as growth metrics, liquidity and profitability also play a vital role. This indicated that a comprehensive approach to building model, incorporate a broad spectrum of indications of financial health. The hybrid model's analysis showcases a high variety of variables, techniques and approaches used to predict the financial results, demonstrating both the potential and complexity of these models in financial research advancing. The emphasis on both validation and accuracy methods highlights the field move

Table 4
Hybrid models.

Technique Used	Count of Research Papers	Estimated Accuracy
Big Data Analytics	1	0.87 to 0.89
Hybrid Models (Machine Learning and Traditional Methods)	1	High
Feedforward Neural Network with Cuckoo Search Algorithm	1	90.30 %–82.79 %
Generalized Additive Models	1	Best for 3–12 months prediction
Machine Learning Techniques (short to long-term prediction)	1	High
Hybrid Artificial Neural Network	1	Superior to traditional models
Machine Learning Techniques (Various Datasets)	1	Varies
Hybrid Model (Logistic Regression and Neural Networks)	1	High
Machine Learning Techniques (ANN, SVM, k-NN)	1	ANN and k-NN most accurate
Logistic Regression and Machine Learning Non-Parametric Algorithms	1	High
Machine Learning Techniques (Credit Scoring and Bankruptcy Prediction)	1	High
Hybrid Model (Traditional and AI Methods)	1	Superior to initial models
Neural Networks and Self-Organizing Maps	1	96.15 %
Hybrid Model (Combining Statistical and Machine Learning)	1	High
Machine Learning (Credit Risk Prediction)	1	High
Machine Learning Techniques (Various)	1	Varies
Stacked Machine Learning Models	1	High
Machine Learning (Bankruptcy Prediction)	1	95 %
Hybrid Model (ANN and Logistic Regression)	1	More accurate than individual models
Hybrid Model (Machine Learning and Statistical)	1	High
Others Mixed Methods	19	Varies

Table 5
Model evaluation.

Evaluation/Validation Technique	Estimated Accuracy
Grid Search Optimization	High
SHapley Additive exPlanations (SHAP) Value	Varies
ROC AUC	0.87 to 0.89
Logarithmic Loss (logLoss)	High
Precision-Recall and ROC Area Under Curve	0.80 to 0.91
Kolmogorov–Smirnov Index, Brier Score	High
Bayesian Statistical Testing	Varies
SHAP and LIME for Model Interpretability	High
eXplainable Artificial Intelligence Methods	High
Gaussian Processes	High
Entropy-Based Analysis	High
Probabilistic Interpretation with Gaussian Processes	High
Performance Comparison and Visualization	High
Performance Metrics: ROC Curve, Kolmogorov–Smirnov Index, Brier Score	High
Multivariate Discriminant Analysis, Logit Regression, Neural Networks	Varies
Area Under ROC Curve	0.88 with 60-month prediction horizon

Table 6
Most common features used in hybrid models.

Most Used Variables	Number of Papers
Financial Ratios and Data	7
Economic and Macroeconomic Indicators	3
Profitability, Liquidity, Growth Metrics	3
Credit and Loan Quality Measures	2
Bankruptcy and Distress Indicators	2
Credit Scoring and Risk Evaluation	2
Classification and Default Rates	2
Other Specific Financial Indicators	18

towards more interpretable, robust and reliable predictive tools.

3.2. Intelligent model analysis

Table 7 explains intelligent model used for default prediction (Intelligent model means only Machine Learning and Deep Leaning – ML & DL). The analysis of intelligent models demonstrates a broad spectrum of techniques utilized in financial predictions, with Artificial Neural Networks (ANN) and Deep Learning models at the forefront in terms of both popularity and estimated accuracy (75–85 % and 80–90 %, respectively). Machine Learning, Random Forest, and Support Vector Machine (SVM) techniques also feature prominently, showcasing accuracies within specific ranges (e.g., SVMs at 78–86 %). Gradient Boosting stands out with an impressive 85–95 % accuracy. Validation techniques such as ROC Curve Analysis and Area Under Curve (AUC) Analysis further affirm the high performance of these models, with ROC analysis indicating accuracy up to 95 % for deep learning models.

In the last two columns we placed model validation methods. Validation strategies include Cross-Validation, Confusion Matrix, and Precision-Recall Analysis, among others, with estimated accuracies that vary with the model. The diversity in validation methods, from K-Fold Cross-Validation to Leave-One-Out Cross-Validation, underscores the rigorous efforts to ensure the reliability and effectiveness of these intelligent models.

Table 8 emphasize the most common variables/features used to predict default events through the intelligent techniques. This

Table 7
Intelligent model.

Technique Used	Count of Research Papers	Estimated Accuracy	Evaluation/Validation Technique	Estimated Accuracy
Artificial Neural Networks (ANN)	12	75–85 %	Cross-Validation	Varies with model
Deep Learning	15	80–90 %	ROC Curve Analysis	80–95 %
Machine Learning	10	70–85 %	Confusion Matrix	Varies with model
Random Forest	8	80–88 %	Precision-Recall Analysis	75–90 %
Support Vector Machine (SVM)	11	78–86 %	Area Under Curve (AUC) Analysis	85–95 %
Logistic Regression	9	70–80 %	Holdout Method	70–85 %
Decision Trees	10	72–82 %	Bootstrapping	75–90 %
Gradient Boosting	10	85–95 %	K-Fold Cross-Validation	80–90 %
Ensemble Learning	9	77–87 %	Leave-One-Out Cross-Validation	Varies with model
Deep Belief Networks	10	83–93 %	F1 Score Analysis	78–88 %

Table 8
Most common features used in intelligent model.

Most Used Variables	Number of Papers
Loan to Value Ratio	7
Interest Rate	6
Credit Score	5
Debt-to-Income Ratio	5
Net Profit Margin	5
Return on Assets	5
Current Ratio	5
Total Assets	5
Earnings Per Share	6
Operating Cash Flow	6
Gross Domestic Product (GDP) Growth	7
Market Capitalization	7
Liquidity Ratios	5
Asset Quality	5
Capital Adequacy Ratio	5

research analyses the significant tole of economic and financial indicators in the intelligent model's predictive accuracy. Credit Scores, Interest Rates and Loan to Value Ratio are among the most utilized variables. Moreover, the indicators of macroeconomic variables such as Market Capitalization, Gross Domestic Product (GDP) Growth are heavily featured indicating a comprehensive approach to incorporate the broader economic factors in the prediction of models. This analysis indicates the intelligent models in economic and financial models underlying the techniques diversity, the breadth of variables and the rigor of validation methods. These models characterized by nuances understanding of financial phenomena and their robustness, which represent the significant advance in the analysis of prediction in the financial sector.

Table 9 provides a comprehensive overview of recent studies on financial distress or default prediction modeling by hybrid and intelligent models. Its aim is to evaluate the impact of these models by examining their accuracy, which is influenced by factors such as the adequacy of the sample size, the efficiency of the features used, and the robustness of the research methods employed. The findings indicate that the most accurate models typically employ machine learning or deep learning techniques, achieving accuracy rates ranging from approximately 80 % to over 95 %. A total of 52 best article with highest accuracy were analyzed initially as being shown in table.

3.3. Statistical model analysis

Table 10 showcased analysis of statistical model. This analysis showcases a wide array of statistical models utilized in financial research, highlighting their versatility and effectiveness. Logistic Regression, used in 17 papers, shows a consistent accuracy range of 70–85 %, validated through ROC Curve and Confusion Matrix techniques. The exploration includes diverse methods such as Optimization and Algorithmic Methods, Data Mining, and Survival Analysis, with accuracies spanning from 60 % to as high as 95 %. Notably, Neural Networks, though classified under statistical models in some research, demonstrate an impressive accuracy range up to 98 %, indicating their potent predictive capability.

The last two columns exhibiting validation of the model. Various evaluation and validation techniques underscore the reliability of these statistical models. Techniques range from traditional p-values and Goodness-of-fit tests to more complex Cross-validation and ROC Curves, reflecting a broad spectrum of methods to ensure model accuracy. Feature Selection and Evaluation show a notable impact, suggesting a potential 5–10 % improvement in model performance, emphasizing the importance of choosing the right variables for analysis.

Table 11 explains the most common features/variables used in default prediction by adopting core statistical methods. The analysis highlights the critical role of a diverse set of variables in enhancing the predictive accuracy of statistical models. Credit Scores and Financial Ratios are among the most frequently used appearing in 12 and 15 papers respectively. The inclusion of market and demographic data and macroeconomics indicators emphasize the comprehensive approach taken by researchers to incorporate a various factor, reveals an understanding of the multifaceted nature of the financial predictions, where various type of data combine which significantly improve the outcomes of models. The statistical model's analysis in financial prediction indicates the varied and deep toolkit of variables and techniques that researchers employ to tackle predictive tasks that are complex. The thoughtful selection of variables and rigorous evaluation methods showcase the meticulous approach to enhancing the model accuracy and reliability which makes significant advancement in the financial analytics field.

3.4. Textual model analysis

Table 12 explain the textual model's analysis along the methods of evaluation. The textual model exploration in financial models indicate the rich diversity of the research techniques, with a strong emphasis on content analysis and deep learning models and various forms of document vectorization like Doc2Vec, Word2Vec and Bag of Words (BOW). The estimated accuracy for these models spans from 65 % to 95 %, revealing significant textual data potential in enhancing the predictive accuracy. The methods and models of deep

Table 9
Detailed analysis of hybrid models & intelligent model.

#	Journal Name	Author and Year	Citations	Country	Sample	IV Variables	DV Variable	Impact Size
1		[56]	26	Korea	454,752 monthly observations of non-financial firms	Financial Ratios	Bankruptcy Probabilities	A default probability of less than 10 % defaults within 1 year
2		[57]	1	USA	109 companies from machine-building sectors	Financial Ratios	credit risks	An accuracy of predictive model is 50 %, the gradient boosting model is 47 %
3		[58]	23	UAE	German Credit, Australian Credit	–	Default Risk	The proposed model accuracy attaining 96.87 %, 84.74 % and 94.20 %
4		[59]	0	USA	24 companies engaged in the IT field	Financial Ratios	Probability to Default	Compared to financial ratio-based models, the multi-source framework provides higher accuracy
5		[60]	8	Malaysia	98 public listed companies in Bursa	Financial Ratios	credit risks	LSTMs and gated recurrent units (GRUs) achieved 90 % accuracy and 93 % accuracy, respectively
6		[61]	2	Turkey	The portfolios are equally weighted and updated every month	Financial Ratios	Stock Return	Approximately 86 % of the returns of stocks are correctly classified
7		[62]	119	France	133 companies that filed for bankruptcy in 2017	Financial Ratios	Bankruptcy	Among all the models, CatBoost's AUC was significantly higher at 10 % than others
8		[63]	3	Indonesia	54 financial services companies	Financial Ratios	Financial distress	The financial services industry had an accuracy rate of 81.03 percent, while the manufacturing industry had an accuracy rate of 96.6 percent
9		[64]	16	China	226 financial enterprises	Investment Activities	Financial Risk	Deep belief network (DBN) is an accurate model for assessing financial risks. Its accuracy rate exceeds 91 %.
10		[65]	83	Belgium	Four data sets, Bene 1, Bene2, Bene 3 and UK from financial institutions in the Benelux and the UK	–	Credit Score	A credit score based on XGBoost is preferable to other credit scores
11		[66]	21	Canada	30,000 imbalanced dataset	Characteristics of client	Default	ALL-KNN sampling technique produced a 98.6 % accuracy, and the cross entropy loss measurement was 0.028.
12		[67]	84	Spain	17 developed countries, 2499 country-year observations,	Macroeconomics Variables	Crisis Risk	The inversion or flattening of the yield curve can be a cause for concern at times of low nominal interest rates and high credit growth.
13		[68]	63	China	2893 samples	Financial Ratios	FDP Performance	There are high levels of class inequality in four financial states: 77.17 %, 12.18 %, 4.98 %, and 5.67 %
14		[69]	36	Taiwan	344 OTC companies in Taiwan	Financial Ratios	Financial Distress	Its prediction accuracy rate is 94.23 %, which is the highest of all the financial distress models
15		[70]	28	China	Default Data dataset contains 30,000 instances. PD dataset contains 55,596 instances.	–	Probability of default, Credit Scoring	Overall performance has been significantly improved by each component of the proposed model
16		[71]	4	Turkey	126 businesses in manufacturing industry	Financial Ratios	Prediction of Default	In 2006, the accuracy of CART was 84.21 %, ANN was 81.58 %, and C5.0 was 76.32 %.

(continued on next page)

Table 9 (*continued*)

#	Journal Name	Author and Year	Citations	Country	Sample	IV Variables	DV Variable	Impact Size
17		[72]	59	UK	553 firms from financial and utility sectors	Financial Ratios	Credit Default	It is considered one of the best machine learning DPMs to use bagging CART and random forest
18		[73]	29	USA	132 observations	Financial Ratios	Bankruptcy	To predict corporate bankruptcy accurately, SAE's softmax classifier outperforms other existing classifiers
19		[74]	7	Cyprus	175,000 records	CAR, Macroeconomics Variables, Financial Ratios	Risk	Methods that simulate real-world situations more accurately and efficiently than traditional approaches
20		[35]	16	USA	600,000 publicly available single-family US mortgages	Loan Level Variables, Macroeconomics Variables	Risks	A model adapted from the Cox proportional hazard model, and one adapted from Fine and Grey forecasted prepayment and default and risk more accurately than statistical benchmarks.
21		[75]	55	France	80 banks from the GCC	Financial Ratios	credit risks	In terms of F1 score, specificity, and accuracy scores, arbitrarily chosen forests showed the greatest precision
22		[76]	36	Brazil	1890 records was accessed and retrieved	Predictor Attributes	Commercial Risks, Credit Risk	With 80 % mean accuracy, the best results were obtained when linear regression was equal to 0.2.
23		[77]	71	Germany	478 companies	Financial and Non-Financial Variables	Bankruptcy Prediction	With an accuracy of 98.3 %, the model has been proven to be accurate.
24		[78]	81	Taiwan	764 bankrupt firms	Financial Ratios, Corporate Governance	Financial Distress	There is a significant improvement in performance over the baseline model if there is a high cost associated with misclassifying bankruptcy companies.
25		[79]	1	India	Financial data of private sector banks of India	Financial Ratios	Z-Score	The artificial neural network model is highly accurate
26		[80]	67	Poland	five-real world datasets of Polish companies.	Financial Ratios	Probability to Default	In comparison with other banksruptcy prediction methods and machine learning methods, BSM-SAES has the highest AUC rate.
27		[81]	16	UAE	613 sample data collected from the companies in MENA	Financial Ratios	credit risks	It is more accurate to predict events with hybrid models than with classical machine learning models
28		[2]	42	Australia	4,631,655 firm years for private companies classified as active and trading (2) 23,405 firm years	Financial Ratios, Macroeconomics Variables	credit risks	Additionally, the three- and five-state failure models performed well one year before failure, with AUCs of 0.951 and 0.912, respectively
29		[82]	40	Spain	Total 808 Firms	Audit Report Variables	credit risks	A total of 68.2 %, 76.5 %, and 80.8 % accuracy were determined for Models 1, 2, and 3.
30		[83]	88	Greece	175,000 records	CAMELS	Insolvency Risks	Random Forests (RF) have superior out-of-sample and out-of-time predictive performance compared to

(*continued on next page*)

Table 9 (*continued*)

#	Journal Name	Author and Year	Citations	Country	Sample	IV Variables	DV Variable	Impact Size
								other widely used bank failure models, as well.
31		[84]	115	UK	The UK regulated mortgages collected by the UK Financial Conduct Authority (FCA). Starting from June 2015, updated every six months	Loan Characteristics	Default Probabilities	A shift towards higher default probabilities is predicted by both models.
32		[85]	30	USA	7000 to 9000 companies	Stock Prices	Default Return	When compared to SVM, DBN performs better for all 30-180-360-day prior to default data sets
33		[86]	24	SWEDEN	1489 observations of publicly listed companies	CDS	Credit Score	Based on the proposed approach, predictions are more accurate than those from single classifiers
34		[87]	91	UK	7996 observations	Financial Ratios	Default Risk	MCMC provides better results than Gauss-Newton for estimating Bayesian regularization parameters
35		[88]	21	Greece	150 individual firms and 450 firm-year observations	Multiple Accounting Ratios	Default Probability	Compared to other algorithms, deep dense networks perform better
36		[89]	31	Pakistan	273 ongoing non-financial firms listed at Pakistan Stock Exchange	Financial Ratios, Stock Market Prices	Default Probability	For predicting MSFD, the proposed model had an accuracy score of 84.06 %
37		[90]	64	China	25,474 customers	Financial Ratios	credit risks	There is a 91.5 % accuracy rate for the comprehensive prediction model
38		[91]	67	Tunisia	several Tunisian banks	Loan and Contract Characteristics	credit risks	Based on credit risk profiles, two classes were determined
39		[92]	28	Turkey	default characteristics of 79254 granted loans in a Turkish financial institution	Characteristics of Loan	Default Prediction	Regression-based models might not yield as good results as deep learning models
40		[93]	138	China	690 in which, 307 good applicants and 383 bad applicants	Characteristics of Applicants	Credit Risks	Credit scoring models perform better if they are integrated at either the clustering or consensus stages
41		[94]	53	Italy	250 instances and 6 attributes		Financial Risks	An accuracy measurement of 99.96 %
42		[95]	55	Australia	12,156 active firm years and 3348 distressed firm years	Accounting Variables and Financial Ratios	Probability to Default	The accuracy in predicting distress in binary TreeNet model is 93.74 %
43		[96]	8	China	More than 50,000 historical credit records gathered from the credit scoring system of the bank	Financial ratios and Characteristics of Customers	Loan Default	F-measure performance in several real credit datasets shows RobustFM to be significantly better than state-of-the-art methods
44		[97]	39	India	850 customers	Attributes	Credit Risk	The gradient boosting model revealed that ensemble model works better than individual models
45		[98]	232	China	Chinese public listed companies	Financial Ratios	Financial Distress	In the case of minority financial distress samples, both a simple integration model and an embedding integration model significantly enhance the recognition capabilities, and the embedding integration model performs better than the simple integration mode as well.
46		[99]	123	Taiwan	30,000 observations	Default Payment	Default Risk	The boosting classification ability is superior to other

(continued on next page)

Table 9 (*continued*)

#	Journal Name	Author and Year	Citations	Country	Sample	IV Variables	DV Variable	Impact Size
								methods of machine-learning such as neural networks
47		[100]	57	Bangladesh	Six real-world credit datasets	–	Credit Default	Model SVM is marginally more accurate to CART with DA
48		[101]	78	Spain	Nine real-life data sets	–	Credit Risks and Corporate Failure	In order to determine credit risk and bankruptcy, the associative memory here proposed provides a useful solution
49		[102]	49	Iran	117 firms	Financial Ratios	Financial Distress and Bankruptcy	ANN outperforms other techniques
50		[103]	72	Iran	180 companies, 90 financially distressed companies and 90 financially healthy	Financial Ratios	Probability to Default	A combined SVM model using the SFFS approach can achieve higher accuracy compared to other methods applied to foreign and domestic datasets.
51		[104]	16	Italy	6929 companies Data gathered from Thomson Reuters	Financial Ratios	Z-Score	Achieved consistent success rates of approximately 90 % for both one-year and two-year forecast predictions.
52		[105]	500	China	107 listed companies from CSMAR	Financial Ratios	Financial Distress	Neural networks deliver superior accuracy compared to other classifiers, such as support vector machines and decision trees.

Table 10
Statistical models.

Technique Used	Count of Research Papers	Estimated Accuracy	Evaluation/Validation Technique	Estimated Accuracy Range
Logistic Regression	17	70–85 %	ROC Curve, Confusion Matrix	70–85 %
Statistical Models and Tests	11	60–80 %	p-values, Goodness-of-fit tests	60–80 %
Optimization and Algorithmic Methods	10	80–95 %	Cross-validation, ROC Curve	80–95 %
Data Mining and Analysis	8	70–90 %	K-fold Cross-validation, Information Gain	70–90 %
Survival Analysis	6	70–85 %	Kaplan-Meier, Log-rank test	70–85 %
Machine Learning Algorithms	5	80–95 %	Cross-validation, Precision-Recall	80–95 %
Neural Networks	4	75–98 %	Holdout method, Confusion Matrix	75–98 %
Random Forest and Decision Trees	4	80–90 %	Out-of-bag error, Feature Importance	80–90 %
Feature Selection and Evaluation	4	70–90 %	Feature Importance, Wrapper methods	5–10 % improvement
Discriminant Analysis	2	70–85 %	Wilks' Lambda, Cross-validation	70–85 %
Other Models and Techniques	25	Varies widely	Varies with model	60–90 %

Table 11
Most common features used in statistical model.

Most Used Variables	Number of Papers
Financial Ratios	15
Credit Scores	12
Macroeconomic Indicators	10
Demographic Data	8
Market Data	7
Transaction Data	5
Customer Behavior Metrics	4
Historical Performance Data	4
Time-to-Event Data	3
Company Financial Ratios	3

Table 12
Textual model.

Research Technique	Estimated Accuracy Range of AI Models	Common Model Evaluation Techniques	Estimated Accuracy Range of Evaluation Model	Number of Papers	Research Variable	Number of Papers
Deep Learning Models	65 %–95 %	1 - Precision, Recall, and F1 Score 2 - HARA 3 - Deep Dense Multilayer Perceptron (DDMP) 4 - Content Analysis 5 - Textual Analysis	85 %–95 %	6	Textual Disclosures	6
Content Analysis	65 %–95 %			3	Risk Disclosure	2
Document Vectorization (BOW, Word2Vec, Doc2Vec)	65 %–95 %			3	Qualitative Information in Reporting	3
Transformer Encoder	65 %–95 %			2	Loan Descriptions	2
Machine Learning	65 %–95 %			2	Alternative Data Sources	2
Natural Language Processing	65 %–95 %			2	Loan Assessment Text	2
Financial Sentiment Analysis	65 %–95 %			1	Financial Sentiment	1
Internal Ratings-Based Approach	65 %–95 %			1	Regulatory Reporting Systems	1
Dynamic Probit Model	65 %–95 %			1	Financial and Nonfinancial Variables	1
Logistic Lasso Regression	65 %–95 %			1	Localization and Economic Conditions	1
XGBoost Algorithm	65 %–95 %			1	SME Failure Factors	1
BERT Model	65 %–95 %			1	Text Length and Keywords	1
Convolutional Recurrent Neural Network	70 %–90 %			1	Auditors' Reports and Management Statements	1
Document to Vector (Doc2Vec)	70 %–90 %			1	Annual Country Reports	1
Bidirectional Long Short-Term Memory (BiLSTM)	70 %–90 %			1	Investor Comments	1
Deep Neural Network	70 %–96 %			1	Credit Risk of Contagion Chains	1
Sentiment Analysis	70 %–90 %			1	Emotions and Credit Risk	1
Topic Modeling	70 %–90 %			1	Descriptive Text Concerning Loans	1
Doc2Vec	70 %–90 %			1	Financial News	1
Others	60 %–85 %			2	Others	2

learning employing the XGBoost Algorithm and Transformer Encoders share a high range of accuracy, revealing their effectiveness in extracting meaningful insights from the complex data of textual.

Moreover, the common model techniques of evaluation include specialized approaches like DDMP Deep Dense Multilayer Perceptron and HARA High Accuracy Rate Analysis, F1 Score and Precision Recall. These methods of evaluation corroborate the effectiveness of models, with some reaching up to the accuracy of 95 %. This indicates the meticulous efforts in the performance validating of the models of textual, ensuring their applicability and reliability in the financial analysis.

The significance analysis underscores various variables derived from the textual data. The risk and textual disclosure are prominent, reflecting an interest in the information of these provided source data. Other variables such as financial sentiment, qualitative information and loan description highlight the textual information breadth leveraged for the purpose of predictive, ranging from the regulatory reports to financial news and investors comments.

The advance model application models like Convolution Recurrent Neural Network (CRNN) and BERT (Bidirectional Encoder Representations from Transformers) to specific textual analysis such as evaluating auditors' reports, management statement or assessing the text length and keywords. These applications highlight the innovative ways in which the data of textual is being used to assess the non-financial and financial variables, SME failure factors and even the emotional content that is related to the credit risk.

The model of textual analysis shows a dynamic field where AI advanced techniques are applied to a various textual data source to forecast the financial outcomes. The high rates of accuracy, underlying the growing effectiveness and importance, coupled with rigorous evaluation methods in financial research. This approach expands the toolkit for financial forecasting and also highlighted the textual data critical role in uncovering the deep insights into risk assessment and financial health.

Based on above given analysis we may infer that within the Intelligent Model category, deep learning algorithms, particularly those employing LightGBM and XGBoost, demonstrate a clear trend towards higher citation scores and accuracy. The study utilized a Neural Network model and achieved a high accuracy rate. This was evidenced by its strong performance across various evaluation metrics, including Precision, Recall, F1 Score, and AUC [105]. This highlights the potential of deep learning techniques in extracting complex patterns from financial data for improved default prediction. In contrast, while statistical models like Logistic Regression still show respectable accuracy (70–85 %), their lower citation scores suggest a shift in research focus towards more sophisticated AI-driven approaches.

3.5. Systematic literature review analysis

Table 13 and Fig. 2 explains that how we extracted key grey areas by analyzing previous papers based on the review research. This analysis finds significant research gaps and directions from systematic literature reviews in the field of credit risk assessment and default prediction, outlining areas for future exploration:

There's a call for the exploration of Notch Distance Analysis for credit rating comparisons and the development of Value Erosion Models to address bankruptcy protection limitations. The need to combine qualitative and quantitative variables for enhanced bankruptcy prediction models (BPMs) is emphasized, alongside enhancing model interpretability to better understand bankruptcy impacts. There's also an interest in evaluating various AI/ML algorithms for bankruptcy prediction, highlighting the potential of AI for developing sustainable business models.

Suggestions include expanding reviews to cover more articles and examine primary data, developing comprehensive models that include aspects such as financial reporting and management quality, and exploring the impact of misclassification, specifically the repercussions of false negatives in credit risk assessment. There's an urge to develop tools that accommodate changing data domains and model flexibility, and to investigate machine learning methodologies, including feature selection, engineering, and tuning.

The analysis points to exploring deep learning applications in banking beyond current literature, understanding the influence of corporate governance measures on financial distress prediction, and providing overviews and methodologies for data mining in finance. Conducting sensitivity analysis with larger datasets to discern between data versus model improvements, diversifying data sources, and using advanced classification techniques for heterogeneous data prediction are also noted as areas needing attention.

Performing industry-specific analyses that combine quantitative and qualitative information, implementing data sampling approaches for balanced datasets, and studying model effectiveness in different regulatory environments are highlighted as crucial for advancing the field.

This comprehensive analysis encapsulates the diverse avenues for future research, focusing on advancing technical methodologies, developing innovative models, and exploring the practical applications and impacts of these models in the evolving landscape of financial risk assessment.

Table 14 exhibits the top ten authors and articles, we observed that only 1 article belongs to MDPI with SLR study, the rest of 9 articles are from Elsevier publishing company. However, we further noticed that 3 articles were taken from Intelligent modeling and 3 were from SLR studies in the top ten article list. We developed the list on the basis of overall citations. We observed that from 3 articles of Intelligent modeling 2 belongs to deep learning algorithmic systems using LightGBM and XGboost algorithms and so on. Whereas the top article was developed by using a normal machine learning algorithm. We have seen trends of evaluation in each intelligent and hybrid model with common tools of evaluation/validation by using Precision, Recall, F1, Confusion Matrices, Accuracy, ROC in financial distress modeling. While conducting this study we noticed that Expert Systems with Applications (Elsevier) is the only journal that entertains default studies a lot as compared to others.

While our review highlights common evaluation metrics such as Precision, Recall, and ROC AUC, it's crucial to acknowledge their context-dependent appropriateness and the evolving landscape of evaluation standards in financial distress prediction. The choice of suitable metrics should align with the specific research objectives and data characteristics. For instance, in imbalanced datasets, relying solely on accuracy can be misleading, necessitating the use of metrics like F1-score or AUC for a more comprehensive

Table 13
Detailed analysis of Systematic Literature in between 2015 and 2024.

Research Gap/Direction	Details
Notch Distance Analysis	Explore accuracy of Notch Distance in credit rating comparisons.
Value Erosion Model Development	Develop models to address current bankruptcy protection limitations.
Combining Data Types	Combine qualitative and quantitative variables for better BPMs.
Model Interpretability	Enhance interpretability to understand bankruptcy impact.
Artificial Intelligence and Machine Learning Algorithms	Evaluate various AI/ML algorithms for bankruptcy prediction.
Use of AI for Sustainable Business Models	Expand reviews to include more articles and examine primary data.
Comprehensive Model Development	Develop models including financial reporting, management quality, etc.
Impact of Misclassification	Examine repercussions of false negatives in credit risk assessment.
Machine Learning Methodology	Investigate ML methods like feature selection, engineering, and tuning.
Diverse Data Domains	Develop tools for changing data domains and model flexibility.
Deep Learning Applications	Explore deep learning in banking beyond the current literature.
Corporate Governance Measures	Research influence of governance measures on financial distress prediction.
Data Mining in Finance	Provide overviews and methodologies for data mining in finance.
Sensitivity Analysis with Larger Datasets	Conduct sensitivity analysis to discern data vs. model improvements.
Diversification of Data Sources	Use advanced classification techniques for heterogeneous data prediction.
Industry-Specific Analysis	Perform industry-specific analysis combining quantitative and qualitative info.
Sampling and Data Balance	Implement data sampling approaches for balanced datasets.
Transparency and Regulations	Study model effectiveness in different regulatory environments.

Note: This table encapsulates the various areas where future research can build upon existing work, addressing identified limitations and gaps to advance the field of credit risk assessment and default prediction.

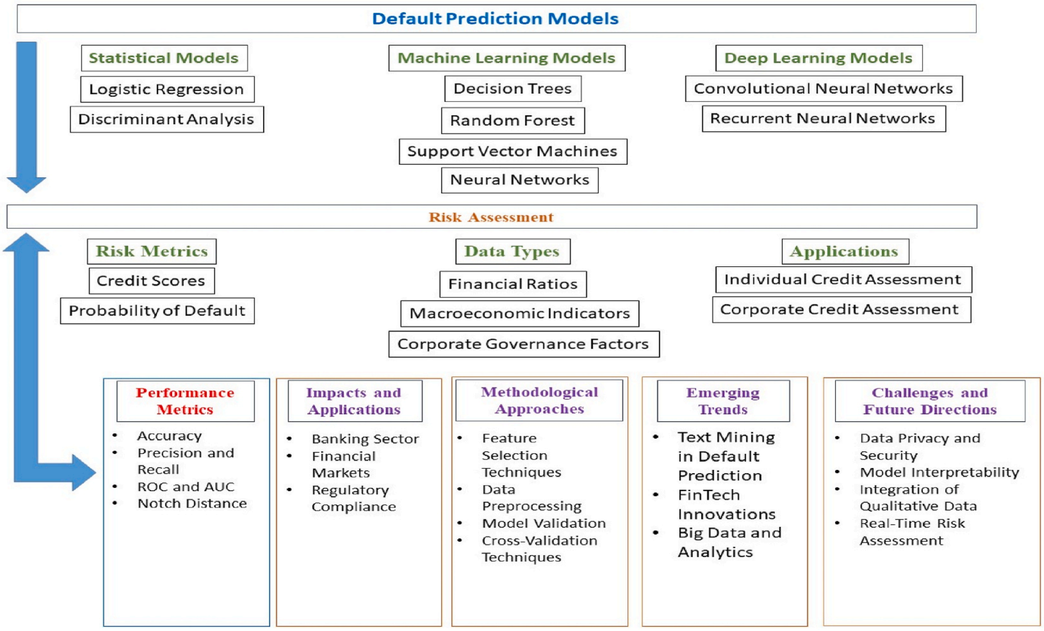

Fig. 2. Process systematic literature review Paper's analysis.

assessment.

Furthermore, the field has witnessed a shift towards emphasizing model interpretability and explainability. While traditional metrics focus on predictive performance, techniques like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are gaining traction as they offer insights into the factors driving model predictions. This evolution reflects the growing need for transparency and trust in financial models, especially in high-stakes applications like credit risk assessment.

Table 15 represents all the databases which were used in this SLR study. We observed that studies based on personal loan data were conducted by using The Lending Club, Machine Learning Repository and Chinese P2P Lending Platform in most of the cases. However, we further noticed that studies with highest citation scores used databases of Bloomberg, rest of the databases are the common databases in substantial dataset of articles.

4. Conclusion

After the global financial crisis in 2008-09, financial default prediction modeling extreme attention of researchers. It has always been a typical problem that which tools/technique/method can be better to predict financial distress/financial default/bankrupt/credit default not only suitable technique with respect to time and place (year of study and region of study) but also the most efficient features which have potential to report the defaults in future. The core of this SLR study was to gauge the answers through in depth analysis of past studies as ranges in between 2015 and 2024. The set environment of this SLR was based on 5 critical domains in default prediction models as mentioned above in detail. Findings of the SLR study revealed that mostly high cited studies are being done by using intelligent models which are further divided in machine learning and deep learning modeling. In which, we further explored that research based on deep learning algorithms are being cited a lot that witnessed that recent trend and quality working of these researches. We have noticed that Elsevier, being a publishing house, is ruling this domain by holding around 66.59 % citations alone. We further observed that in author and journal ranking 9 out of 10 belongs to Elsevier and 'Expert Systems with Applications' is the journal which entertains quality articles based on these intelligent techniques. Interestingly, the article based on deep learning algorithms holds highest citations and ranked 1 in all.

In machine learning algorithms we found the most used tools are as follows (1) Neural Network (NN), (2) Random Forest (RF), (3) Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN), and Naïve Bayes (NB). Whereas we witnessed Convolution Neural Networks (CNN, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN) and Structure of the GRU Unit are being used for deep learning financial distress studies. Moreover, every intelligent algorithm needs to be validated or evaluated through a systematic methodology by using evaluation tools of AI. We found Precision, Recall, F1, Confusion Matrices, Accuracy, ROC are common evaluation methods in AI research.

Consequently, our findings regarding the superior performance of Intelligent Models, particularly those employing deep learning algorithms, align with the growing body of literature highlighting the effectiveness of AI in financial distress prediction. For example [107], advocated for a framework incorporating big data analytics and machine learning for predicting failure in construction firms, echoing our observation of the increasing dominance of AI-driven approaches. Similarly [113], in their comprehensive review of machine learning in banking risk management, emphasized the potential of these techniques for enhanced accuracy in default

Table 14
Top 10 Authors and Rankings based on Citation.

Ranks	Year	Author	Citation	META	Country	Publisher	Journal	Title	Method/Model	Evaluation measures
1	2015	[105]	395	Intelligent Model	China	Elsevier	European Journal of Operational Research	Prediction of financial distress: An empirical study of listed Chinese companies using data mining	NN, DT and SVM	Precision, Recall, F1, Confusion Matrices, Accuracy, ROC
2	2016	[106]	254	Textual Model	Germany	Elsevier	Journal of Banking & Finance	Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms	Probit regressions	ROC and AUC
3	2018	[107]	236	Systematic Review	UK	Elsevier	Expert Systems With Applications	Systematic review of bankruptcy prediction models: Towards a framework for tool selection. Expert Systems with Applications	Systematic Review	NA
4	2015	[108]	218	Hybrid Model	USA	Elsevier	Expert Systems with Applications	Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks	DA, LR, RF, MLP, SVP, MLP-SOM	Precision, Recall, F1, Confusion Matrices, Accuracy, ROC
5	2019	[109]	185	Textual Model	USA	Elsevier	European Journal of Operational Research	Deep learning models for bankruptcy prediction using textual disclosures.	NN, LR, RF and SVM	Predictive Performance Model, Model Validation and Model Monitoring
6	2018	[110]	179	Intelligent Model	China	Elsevier	Electronic Commerce Research and Applications	Study on A Prediction of P2P Network Loan Default Based on the Machine Learning LightGBM and XGboost Algorithms according to Different High Dimensional Data Cleaning	LightGBM and XGboost algorithms	ROC and AUC
7	2015	[111]	174	Statistical Model	Taiwan	Elsevier	Knowledge-Based Systems	The effect of feature selection on financial distress prediction	Discriminant Analysis, *t*-Test, Logistic regression, Wrapper based feature selection, Genetic algorithms, Particle swarm optimization	NA
8	2020	[112]	173	Systematic Review	Malaysia	Elsevier	Journal of Business Research	Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review.	Systematic Review	NA
9	2018	[98]	161	Intelligent Model	China	Elsevier	Information Fusion	Class-imbalanced dynamic Financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting	ADASVM-TW, SMOTE,	ROC and AUC
10	2019	[113]	156	Systematic Review	Australia	MDPI	Risks	Machine Learning in Banking Risk Management: A Literature Review	Systematic Review	NA
		Total Citations	**2131**							

Table 15
Top International Databases used in Default Prediction Models.

S. no	List of Databases	S. no	List of Databases	S. no	List of Databases
1	AIDA database	24	EMIS Database	47	Pakistan Stock Exchange
2	Analyzing Categorical Data – USA	25	FAME Bureau Van Dijk	48	Postal Savings Bank of China
3	Bank for International Settlements and the International Monetary Fund	26	Federal Deposit Insurance Corporation (FDIC)	49	SAS Compute Database
4	Bloomberg	27	Freddie Mac	50	SCF Dataset
5	Bombay Stock Exchange	28	Funding Circle	51	Shanghai and Shenzhen Stock Exchanges
6	Bondora	29	Iberian Statements Analysis System	52	SIRENE Directory
7	Borsa Stock Exchange	30	Indonesia Development Forum	53	State Bank of Pakistan
8	Bureau van Dyke ORBIS database	31	Insight database of Dion Global	54	STX
9	Bursa Malaysia	32	International ISI publication	55	Taiwan Economic Journal (TEJ)
10	Center for Research in Security Prices (CRSP)	33	Irvine (UCI) Machine Learning Repository	56	Taiwan Stock Exchange
11	Centre for Monitoring Indian Economy (CMIE)	34	Kaggle	57	Tehran Stock Exchange
12	China Security Market Accounting Research (CSMAR)	35	Lending platform (Eloan) of China	58	The Central Bank of the Republic of Turkey (CBRT)
13	Chinese P2P Lending Platform	36	Macrohistory Database	59	The Lending Club
14	Chinese Stock Exchanges	37	MARKIT Database	60	The New York Stock Exchnage
15	CMIE Prowess Database	38	MBCs	61	Tunisian Central Bank
16	COMPUSTAT	39	Ministry of Finance of the Slovak Republic	62	UCI Irvine Machine Learning Repository
17	CSMAR Database	40	Ministry of Finance	63	UCI Machine Learning Library
18	Customer Information & Transactional Database	41	Moody's	64	UCI Machine Learning Repository
19	Czech Stock Exchange	42	Nairobi Securities Exchange	65	University of California Irvine Machine Learning Database Repository.
20	Datastream	43	National Bank of Ethiopia	66	VOITTO Database
21	Dhaka Stock Exchange	44	National Bank of Greece	67	Web Crawler Software
22	DIANE Database	45	Nile Stock Exchange (NSE)	68	Wharton Research Data Services (WRDS)
23	Egyptian Stock Exchange	46	Osiris Database		

prediction.

Whilst, It has been further observed that the features considered in the majority of these studies are predominantly Financial Ratios (over 80 %), with the remainder involving off-balance sheet items. Importance of financial ratios as key features in our analysis resonates with the findings of Gen, who demonstrated the effectiveness of data mining techniques utilizing financial ratios for predicting financial distress in Chinese listed companies. However, our results diverge from those of [106] who found that textual disclosures in peer-to-peer lending platforms provided significant predictive power. This discrepancy highlights the importance of considering contextual factors and data characteristics when interpreting model performance. For instance, the focus on publicly traded companies in our study, as opposed to the peer-to-peer lending context of [106], might explain the differing importance of financial ratios versus textual data.

Last not Least, we found a drastic trend in SLR studies in default modeling which indicates a solid trend in default modeling. We further observed severe passion in contextual feature selection studies by using intelligent algorithms in this domain, moreover we observed healthy citation rate in Textual model studies which indicates further area of research in this domain. While studying all these textual models we found it a difficult and time consuming technique which increases the efficiency rate of predicting defaults [42,43]. Indeed, statistical models are the core and foundation of default studies [10,11]. On other hand, we found a mild trend in this domain with less citation score which indicates that researchers and readers are curious to see some advanced workout to predict financial defaults. Lastly, Intelligent AI algorithms are the focused area in recent 9 years which is witnessed and reported on the basis of total article produced with per article citation, it further encourages researchers to develop models of default prediction based on AI algorithms.

This systematic review highlights the growing dominance of Intelligent Models, particularly those employing deep learning algorithms, in the field of financial distress prediction. Our analysis reveals a consistent trend towards higher accuracy and citation scores for these models, suggesting their superior ability to capture complex relationships in financial data. Furthermore, the study underscores the importance of careful feature selection and data source considerations for enhancing model performance. These findings provide valuable insights for researchers and practitioners seeking to develop and implement more effective financial distress prediction models, ultimately contributing to greater financial stability and informed decision-making.

Main limitation of this study is that we could not cover more than 9 years of data with all publishing houses (rest of top-tier publishers). We focused on the top tier publishing houses with relevant articles; we did not include the articles based on Credit Default Swaps (CDRs), Sovereign Ratings Models (SRM), and Rating Models for Financial Intuitions. this study provides a holistic view of the developments in financial default prediction over a contemporary time span. By mapping out tools, data sources, evaluation techniques, and geographical hubs of research, it contributes to the practical, theoretical, and methodological understanding of the domain. Future research can further expand upon these findings, diving deeper into the areas the current study highlighted or filling the existing gaps.

Registration and protocol

- This review was not registered.
- No formal protocol was prepared for this systematic review.
- Not applicable, as no protocol was registered or prepared.

CRediT authorship contribution statement

Jahanzaib Alvi: Writing – original draft. **Imtiaz Arif:** Supervision. **Kehkashan Nizam:** Data curation.

Data availability statement

Data will be furnished upon a formal request to corresponding author.

Funding details

- This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Dr. Jahanzaib Alvi reports was provided by Iqra University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We express our deep gratitude to the honorable editor of the journal and the esteemed anonymous reviewers for their valuable insights. We extend special thanks to Prof. Dr. Imtiaz Arif for his guidance throughout the entire process of writing this article.

References

[1] J. Alvi, I. Arif, Unraveling the mystery of default prediction: a study on the textile industry in Pakistan, Journal of Quantitative Methods 7 (2) (2023).
[2] S. Jones, A literature survey of corporate failure prediction models, J. Account. Lit. 45 (2) (2023) 364–405.
[3] Z. ltüzer, Predicting stock returns with financial ratios: a new methodology incorporating machine learning techniques to beat the market, Asia-Pacific Journal of Accounting & Economics 30 (3) (2023) 619–632.
[4] A. Hasnain, et al., Assessing the ambient air quality patterns associated to the COVID-19 outbreak in the Yangtze River Delta: a random forest approach, Chemosphere 314 (2023) 137638.
[5] Y. Xu, et al., Profit-and risk-driven credit scoring under parameter uncertainty: a multiobjective approach, Omega 125 (2024) 103004.
[6] A. Saha, L. Hock Eam, S. Goh Yeok, Housing loan default in Malaysia: an analytical insight and policy implications, Int. J. Hous. Mark. Anal. 16 (2) (2023) 273–291.
[7] M. Crouhy, D. Galai, R. Mark, A comparative analysis of current credit risk models, J. Bank. Finance 24 (1–2) (2020) 59–117.
[8] M.B. Gordy, A comparative anatomy of credit risk models, J. Bank. Finance 24 (1–2) (2000) 119–149.
[9] D. Kok, Stakeholder-centric Approach to Applying Machine Learning to Probability of Default Models, University of Twente, 2024. *Master's thesis*.
[10] E.I. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, J. Finance 23 (4) (1968) 589–609.
[11] J.A. Ohlson, Financial ratios and the probabilistic prediction of bankruptcy, J. Account. Res. 18 (1) (1980) 109.
[12] A.S. Koyuncugil, N. Ozgulbas, Financial early warning system model and data mining application for risk detection, Expert Syst. Appl. 39 (6) (2012) 6238–6253.
[13] W. Beaver, M. McNichols, J.W. Rhie, Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy, Rev. Account. Stud. 10 (1) (2005) 93–122.
[14] M.E. Zmijewski, Methodological issues related to the estimation of financial distress prediction models, J. Account. Res. 59 (82) (1984).
[15] T. Shumway, Forecasting bankruptcy more accurately: a simple hazard model, J. Bus. 74 (1) (2001) 101–124.
[16] S.T. Bharath, T. Shumway, Forecasting default with the Merton distance to default model, Rev. Financ. Stud. 21 (3) (2008) 1339–1369.
[17] J.Y. Campbell, S.B. Thompson, Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Financ. Stud. 21 (4) (2008) 1509–1531.
[18] I.M. Premachandra, G.S. Bhabra, T. Sueyoshi, DEA as a tool for bankruptcy assessment: a comparative study with logistic regression technique, Eur. J. Oper. Res. 193 (2) (2009) 412–424.
[19] F.M. Tseng, Y.C. Hu, Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks, Expert Syst. Appl. 37 (2010) 1846–1853.
[20] K.Y. Tam, M.Y. Kiang, Managerial applications of neural networks: the case of bank failure predictions, Manag. Sci. 38 (7) (1992) 926–947.
[21] G.P. Naidu, K. Govinda, G.P. Naidu, K. Govinda, Bankruptcy prediction using neural networks, in: 2018 2nd International Conference on Inventive Systems and Control (ICISC) 2018, 2018, January, pp. 248–251.
[22] E.I. Altman, G. Marco, F. Varetto, Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience), J. Bank. Finance 18 (3) (1994) 505–529.
[23] T.C. Jh Zhang, A simple statistical parameter for use in evaluation and validation of high throughput screening assays, J. Biomol. Screen (1999) 67–73.
[24] D. Fletcher, E. Goss, Forecasting with neural networks. An application using bankruptcy data, Inf. Manag. 24 (3) (1993) 159–167.
[25] R. Cifuentes, G. Ferrucci, H.S. Shin, Liquidity risk and contagion, J. Eur. Econ. Assoc. 3 (2–3) (2005) 556–566.
[26] G. Kumar, et al., Predicting systemic risk of banks: a machine learning approach, J. Model. Manag. 19 (2) (2024) 441–469.
[27] X. Liu, J. Zhou, Short-term wind power forecasting based on multivariate/multi-step LSTM with temporal feature attention mechanism, Appl. Soft Comput. 150 (2024) 111050.
[28] M.H. Khan, et al., Corruption and bank risk-taking in dual banking systems, Corp. Govern. Int. Rev. (2024).

[29] I.O. Oyewale, E. Alabi, Financial inclusion as a veritable tool for the performance of small and medium enterprises: empirical evidences from osun state, Nigeria, European Journal of Arts, Humanities and Social Sciences 1 (2) (2024) 128–140.
[30] H. Liu, et al., Research on deep learning model of feature extraction based on convolutional neural network, arXiv preprint arXiv:2406.08837 (2024).
[31] A.K. Sahoo, D.P. Tripathy, S. Jayanthu, Application of machine learning techniques in slope stability analysis: a comprehensive overview, Journal of Mining and Environment 15 (3) (2024) 907–921.
[32] Z.Y. Chen, Data-driven risk-averse newsvendor problems: developing the CVaR criteria and support vector machines, Int. J. Prod. Res. 62 (4) (2024) 1221–1238.
[33] X. Zhang, L. Yu, Consumer credit risk assessment: a review from the state-of-the-art classification algorithms, data traits, and learning methods, Expert Syst. Appl. (2023) 121484.
[34] A.B. Musa, et al., IMPACT OF CAPITAL STRUCTURE ON ECONOMIC GROWTH: EVIDENCE FROM NON-FINANCIAL FIRMS IN NIGERIA, 2022.
[35] G. Blumenstock, S. Lessmann, H.V. Seow, Deep learning for survival and competing risk modelling, J. Oper. Res. Soc. 73 (1) (2022) 26–38.
[36] S. Pol, S.S. Ambekar, Predicting credit ratings using deep learning models–an analysis of the indian it industry, Australasian Accounting, Business and Finance Journal 16 (5) (2022) 38–51.
[37] N. Thakur, S. Karmakar, S. Soni, Time series forecasting for uni-variant data using hybrid GA-OLSTM model and performance evaluations, Int. J. Inf. Technol. 14 (4) (2022) 1961–1966.
[38] M. Jacobs Jr, Quantification of model risk with an application to probability of default estimation and stress testing for a large corporate portfolio, Journal of Risk Model Validation 16 (3) (2022).
[39] M. Perez, Artificial neural networks and bankruptcy forecasting: a state of the art, Neural Comput. Appl. 15 (2) (2006) 154–163.
[40] A. Verikas, et al., Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: a survey, Soft Comput. 14 (9) (2010) 995–1010.
[41] W.C. Lin, Y.H. Lu, C.F. Tsai, Feature selection in single and ensemble learning-based bankruptcy prediction models, Expet Syst. 36 (1) (2019) 1–8.
[42] L. Bornmann, H.D. Daniel, What do citation counts measure? A review of studies on citing behavior, J. Doc. 64 (1) (2008) 45–80.
[43] M.E. Falagas, et al., Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses, Faseb. J. 22 (2) (2008) 338–342.
[44] A. Ampatzoglou, et al., The financial aspect of managing technical debt: a systematic literature review, Inf. Software Technol. 64 (2015) 52–73.
[45] F. Pearson, Systematic approaches to a successful literature review, Educ. Psychol. Pract. 30 (2) (2014) 205–206.
[46] R. Mallett, et al., The benefits and challenges of using systematic reviews in international development research, J. Dev. Effect. 4 (3) (2012) 445–455.
[47] B. Kitchenham, et al., Systematic literature reviews in software engineering - a systematic literature review, Inf. Software Technol. 51 (1) (2009) 7–15.
[48] C.E. Flodmark, C. Marcus, M. Britton, Interventions to prevent obesity in children and adolescents: a systematic literature review, Int. J. Obes. 30 (4) (2006) 579–589.
[49] S. Michie, S. Williams, Reducing work related psychological ill health and sickness absence: a systematic literature review, Occup. Environ. Med. 60 (1) (2003) 3–9.
[50] T.M. Connolly, et al., A systematic literature review of empirical evidence on computer games and serious games, Comput. Educ. 59 (2) (2015) 661–686.
[51] K.O. Appiah, A. Chizema, J. Arthur, Predicting corporate failure: a systematic literature review of methodological issues, International Journal of Law and Management 57 (5) (2015) 461–485.
[52] S.C. Gilson, Management turnover and financial distress, J. Financ. Econ. 25 (2) (1989) 241–262.
[53] T.C. Opler, S. Titman, Financial distress and corporate performance, J. Finance 49 (3) (1994) 1015–1040.
[54] G. Andrade, S.N. Kaplan, How costly is financial (not economic) distress? Evidence from highly leveraged transactions that became distressed, J. Finance 53 (5) (1998) 1443–1493.
[55] T. Hoshi, A. Kashyap, D. Scharfstein, The role of banks in reducing the costs of financial distress in Japan, J. Financ. Econ. 27 (1) (1990) 67–88.
[56] H. Kim, H. Cho, D. Ryu, Corporate bankruptcy prediction using machine learning methodologies with a focus on sequential data, Comput. Econ. 59 (3) (2021) 1231–1249.
[57] S. Grishunin, A. Egorova, Comparative analysis of the predictive power of machine learning models for forecasting the credit ratings of machine-building companies, Journal of Corporate Finance Research 16 (1) (2022) 99–112.
[58] N. Metawa, et al., Computational intelligence-based financial crisis prediction model using feature subset selection with optimal deep belief network, Big Data 9 (2) (2021) 100–115.
[59] S. Jiang, Multi-Source Default Probability Prediction Framework Applying Attention Mechanism (Doctoral Dissertation, Stevens Institute of Technology, 2021.
[60] Z. Halim, S.M. Shuhidan, Z.M. Sanusi, Corporation financial distress prediction with deep learning: analysis of public listed companies in Malaysia, Bus. Process Manag. J. 27 (4) (2021) 1163–1178.
[61] Z. İltüzer, Predicting stock returns with financial ratios: a new methodology incorporating machine learning techniques to beat the market, Asia-Pacific Journal of Accounting & Economics 30 (3) (2023) 619–632.
[62] S.B. Jabeur, et al., CatBoost model and artificial intelligence techniques for corporate failure prediction, Technol. Forecast. Soc. Change (2021) 166.
[63] F. Dube, N. Nzimande, P.-F. Muzindutsi, Application of artificial neural networks in predicting financial distress in the JSE financial services and manufacturing companies, Journal of Sustainable Finance & Investment 13 (1) (2023) 723–743.
[64] Y. Song, R. Wu, The impact of financial enterprises' excessive financialization risk assessment for risk control based on data mining and machine learning, Comput. Econ. 60 (4) (2021) 1245–1267.
[65] B.R. Gunnarsson, et al., Deep learning for credit scoring: do or don't? Eur. J. Oper. Res. 295 (1) (2021) 292–305.
[66] M. Mahbobi, S. Kimiagari, M. Vasudevan, Credit risk classification: an integrated predictive accuracy algorithm using artificial and deep neural networks, Ann. Oper. Res. (2021).
[67] K. Bluwstein, et al., Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach, J. Int. Econ. (2023) 103773.
[68] J. Sun, et al., Multi-class financial distress prediction based on support vector machines integrated with the decomposition and fusion methods, Inf. Sci. 559 (2021) 153–170.
[69] C.L. Jan, Financial information asymmetry: using deep learning algorithms to predict financial distress, Symmetry 13 (3) (2021).
[70] H. He, Y. Fan, A novel hybrid ensemble model based on tree-based method and deep learning method for default prediction, Expert Syst. Appl. 176 (November 2019) (2021) 114899.
[71] B. Aksoy, D. Boztosun, Comparison of machine learning methods in prediction of financial failure of businesses in the manufacturing industry: evidence from borsa İstanbul, Anadolu Üniversitesi Sosyal Bilimler Dergisi 20 (4) (2020) 237–268.
[72] M.M. Mousavi, J. Lin, The application of PROMETHEE multi-criteria decision aid in financial decision making: case of distress prediction models evaluation, Expert Syst. Appl. 159 (2020) 113438.
[73] M. Soui, et al., Bankruptcy prediction using stacked auto-encoders, Appl. Artif. Intell. 34 (1) (2019) 80–100.
[74] A. Petropoulos, et al., A deep learning approach for dynamic balance sheet stress testing, in: Paper Presented at the Proceedings of the Third ACM International Conference on AI in Finance, 2022.
[75] J.P. Li, et al., Machine learning and credit ratings prediction in the age of fourth industrial revolution, Technol. Forecast. Soc. Change 161 (September) (2020) 120309.
[76] G. Teles, et al., Artificial neural network and Bayesian network models for credit risk prediction, Journal of Artificial Intelligence and Systems 2 (1) (2020) 118–132.
[77] H. Faris, et al., Improving financial bankruptcy prediction in a highly imbalanced class distribution using oversampling and ensemble learning: a case from the Spanish market, Progress in Artificial Intelligence 9 (2020) 31–53.
[78] D. Liang, et al., Combining corporate governance indicators with stacking ensembles for financial distress prediction, J. Bus. Res. 120 (2020) 137–146.
[79] S. Dhama, Bankruptcy in Indian private sector banks: a neural network analysis, Ushus Journal of Business Management 19 (3) (2020) 89–114.
[80] S. Smiti, M. Soui, Bankruptcy prediction using deep learning approach based on borderline SMOTE, Inf. Syst. Front 22 (5) (2020) 1067–1083.

[81] M. Sreedharan, A.M. Khedr, M. El Bannany, A multi-layer perceptron approach to financial distress prediction with genetic algorithm, Automat. Control Comput. Sci. 54 (2020) 475–482.
[82] N. Muñoz-Izquierdo, et al., Is the external audit report useful for bankruptcy prediction? Evidence using artificial intelligence, Int. J. Financ. Stud. 7 (2) (2019).
[83] A. Petropoulos, et al., A Deep Learning Approach for Dynamic Balance Sheet Stress Testing, 2020.
[84] P. Bracke, et al., Machine Learning Explainability in Finance: an Application to Default Risk Analysis, 2019.
[85] S.H. Yeh, C.J. Wang, M.F. Tsai, Deep belief networks for predicting corporate defaults, in: 2015 24th Wireless and Optical Communication Conference, vol. 2015, WOCC, 2015, pp. 159–163.
[86] C. Luo, A comprehensive decision support approach for credit scoring, Ind. Manag. Data Syst. 120 (2) (2020) 280–290.
[87] E. Sariev, G. Germano, Bayesian regularized artificial neural networks for the estimation of the probability of default, Quant. Finance 20 (2) (2019) 311–328.
[88] S.-A.N. Alexandropoulos, et al., A deep dense neural network for bankruptcy prediction, in: Engineering Applications Of Neural Networks: 20th International Conference, EANN 2019, Xersonisos, Crete, Greece, May 24-26, 2019, Proceedings 20, Springer, 2019.
[89] U. Farooq, J. Qamar, Predicting multistage financial distress: reflections on sampling, feature and model selection criteria, J. Forecast. 38 (7) (2019) 632–648.
[90] L. Tang, F. Cai, Y. Ouyang, Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China, Technol. Forecast. Soc. Change 144 (October 2017) (2019) 563–572.
[91] K. Masmoudi, L. Abid, A. Masmoudi, Credit risk modeling using bayesian network with a latent variable, Expert Syst. Appl. 127 (2019) 157–166.
[92] S. Bayraci, O.J.T. Susuz, A. Economics, A Deep Neural Network (DNN) based classification model in application to loan default prediction 4 (621) (2019) 75–84.
[93] W. Bao, N. Lianju, E.S.w.A. Yue, Integration of Unsupervised and Supervised Machine Learning Algorithms for Credit Risk Assessment, vol. 128, 2019, pp. 301–315.
[94] S. Lahmiri, S. Bekiros, Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design, Quant. Finance 19 (9) (2019) 1569–1577.
[95] C. Jiang, et al., Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending, Ann. Oper. Res. 266 (1–2) (2018) 511–529.
[96] W. Ni, et al., Robust factorization machines for credit default prediction, in: PRICAI 2018: Trends in Artificial Intelligence: 15th Pacific Rim International Conference on Artificial Intelligence, Nanjing, China, August 28–31, 2018, Proceedings, Part I 15, Springer International Publishing, 2018, pp. 941–953.
[97] A. Chopra, P. Bhilare, Application of ensemble models in credit scoring models, Business Perspectives and Research 6 (2) (2018) 129–141.
[98] J. Sun, et al., Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting, Inf. Fusion 54 (2020) 128–144.
[99] S. Hamori, et al., Ensemble learning or deep learning? Application to default risk analysis, J. Risk Financ. Manag. 11 (1) (2018) 12.
[100] F.E. Moula, C. Guotai, M.Z. Abedin, Credit default prediction modeling: an application of support vector machine, Risk Manag. 19 (2017) 158–187.
[101] L. Cleofas-Sánchez, et al., Financial distress prediction using the hybrid associative memory with translation, Appl. Soft Comput. 44 (2016) 144–152.
[102] M. Salehi, M. Mousavi Shiri, M. Bolandraftar Pasikhani, Predicting corporate financial distress using data mining techniques: an application in Tehran Stock Exchange, International Journal of Law and Management 58 (2) (2016) 216–230.
[103] S. Fallahpour, E.N. Lakvan, M.H. Zadeh, Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem, J. Retailing Consum. Serv. 34 (2017) 159–167.
[104] U. Dellepiane, et al., Bankruptcy prediction using support vector machines and feature selection during the recent financial crisis, Int. J. Econ. Finance 7 (8) (2015).
[105] R. Geng, I. Bose, X. Chen, Prediction of financial distress: an empirical study of listed Chinese companies using data mining, Eur. J. Oper. Res. 241 (1) (2015) 236–247.
[106] G. Dorfleitner, et al., Description-text related soft information in peer-to-peer lending - evidence from two leading European platforms, J. Bank. Finance 64 (2016) 169–187.
[107] H.A. Alaka, et al., A framework for big data analytics approach to failure prediction of construction firms, Appl. Comput. Inform. 16 (1–2) (2019) 207–222.
[108] F.J. López Iturriaga, I.P. Sanz, Bankruptcy visualization and prediction using neural networks: a study of U.S. commercial banks, Expert Syst. Appl. 42 (6) (2015) 2857–2869.
[109] F. Mai, et al., Deep learning models for bankruptcy prediction using textual disclosures, Eur. J. Oper. Res. 274 (2) (2019) 743–758.
[110] X. Ma, et al., Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning, Electron. Commer. Res. Appl. 31 (2018) 24–39.
[111] D. Liang, C.F. Tsai, H.T. Wu, The effect of feature selection on financial distress prediction, Knowl. Base Syst. 73 (1) (2015) 289–297.
[112] A. Di Vaio, et al., Artificial intelligence and business models in the sustainable development goals perspective: a systematic literature review, J. Bus. Res. 121 (September) (2020) 283–314.
[113] M. Leo, S. Sharma, K. Maddulety, Machine learning in banking risk management: a literature review, Risks 7 (1) (2019).