IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Artificial Intelligence in Accounting and Finance: Challenges and Opportunities

ZIWEI YI[1], XINWEI CAO[2], ZUYAN CHEN [3], AND SHUAI LI[4], (Senior Member, IEEE)

[1]School of Economy and Business, Jiangxi University of Science and Technology, Ganzhou 341000, China
[2]School of Business, Jiangnan University, Wuxi 214122, China
[3]Faculty of Science and Engineering, Swansea University, Swansea, SA1 8EN, UK
[4]Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu 90570, Finland

Corresponding author: XINWEI CAO (e-mail: xwcao@jiangnan.edu.cn).

ABSTRACT The rapid expansion of artificial intelligence (AI) technologies presents novel technical solutions to traditional accounting and finance problems. Despite this, scholars in accounting and finance frequently encounter difficulties navigating the extensive and intricate domain knowledge of AI and its continuously evolving literature. To address this gap, this paper conducts a qualitative survey of the implementation of AI methods in accounting and finance. The paper is structured into four sections. Firstly, we examine the conventional accounting and finance issues and their requirement for AI techniques. Secondly, to inform accounting and finance researchers about the potential of AI, we present broad categories of AI applications. Thirdly, we explore recent research on AI solutions to conventional problems. Finally, we highlight emerging trends and possible research directions.

INDEX TERMS Artificial Intelligence, Computational Finance, Accounting, Machine Learning, Computational Model.

I. INTRODUCTION

A. MOTIVATION

Nowadays, accounting and finance have assumed increasingly significant roles in economic development, from national economies to individual enterprises. As a business language, accounting supplies stakeholders with substantial information for decision-making, while also reflecting the fiduciary responsibility of enterprise managers. Accounting has expanded beyond the traditional reporting of financial data and now influences the entire spectrum of enterprise management, including forecasting, analysis, control, and decision-making [1]-[3]. Conversely, investors assess companies based on their financial data [4]. For both individuals and enterprises, finance offers access to capital and serves as a channel for asset allocation to investors, enhancing the efficiency of resource allocation [5]. However, despite the significant contributions of accounting and finance to economic growth, the industry still faces several bottlenecks that need to be addressed. Firstly, with regard to data processing, numerous simple and repetitive financial tasks are not labor-effective [6]. Furthermore, some structured data are not optimally utilized, and conventional methods typically fail to extract decision-relevant information from text, leading to the underutilization of data resources [7]. Secondly, regarding decision-making, some accounting and financial judgments rely on the expert's subjective judgment, and the information utilized to make those decisions is typically noisy, which could cause entrepreneurs to make the wrong choices. Moreover, regarding problem-solving efficiency, some traditional solutions such as multiple regression models involve many assumptions, which are not enough to solve practical problems. The era of big data has resulted in massive and complex data, and conventional statistical models are typically inadequate for handling such intricate relationships [8]. Therefore, there is a pressing need for emerging technologies that can support modern accounting in dealing with massive data with intricate complexities.

AI emerged as a critical technology in the 2010s and has increasingly become the dominant technology in the 2020s [9]. The development of AI has presented numerous opportunities to the accounting and finance industry. For instance, many firms have applied AI to optimize their business processes and enhance operational efficiencies [10], such as Facebook, Google, IBM, and Microsoft. Similarly, AI technologies have become an indispensable component of the financial industry, particularly in the automation of operations [11]. AI techniques are critical in capturing both the linear and nonlinear relationships of financial variables. This approach enables the resolution of highly nonlinear problems that are typically infeasible to solve with conventional models. Furthermore, machine learning models provide the computational power and functional flexibility required to decipher complex patterns in high-dimensional data environments, thereby providing technical support for data-driven analyses [12].

B. RELATED WORK

Numerous studies have applied AI techniques to the accounting and finance fields, significantly improving business processes. Optical character recognition [13], intelligent accounting information systems [14], and intelligent trading systems [15] are widely employed in practice, enhancing business processing efficiency. AI techniques enable feature extraction from existing data, as reported in numerous studies. For instance, machine learning methods such as Naive Bayes [16] and Support Vector Machines [17], can extract the tone of company documents. Additionally, data mining techniques have benefitted financial statement analysis. Several studies have shown that such techniques can extract future performance and the likelihood of fraud from financial statements [18]-[20]. Support vector machines [21] and neural network models [22] are also capable of this.

Stock market forecasting and quantitative modeling represent essential applications of Artificial intelligence, particularly in stock return forecasting, risk prediction, automated trading, and portfolio selection and optimization. A recent study analyzed 59 papers on intelligent methods for stock price prediction and found that support vector machines and neural networks were the most commonly employed models [23]. Moreover, numerous studies have applied AI techniques to credit risk management, particularly in bankruptcy prediction and credit rating, regarding risk prediction [24]-[26]. Additionally, portfolio optimization represents another compelling topic. Initially, researchers combined machine learning methods with traditional portfolio theory to enhance the performance of traditional models. Various optimization algorithms (e.g., genetic algorithms, particle swarm optimization algorithms) are commonly employed to solve traditional portfolio models [27], [28], and other solutions include fuzzy logic [29] and neural networks [30].

In summary, AI has made substantial strides in accounting and finance over the past few decades. Nevertheless, accounting and finance researchers may remain uncertain about the application of AI, particularly in some classic problems. Prior reviews of the application of AI in accounting and finance have focused on specific areas, such as portfolio optimization, risk management, and asset pricing [31]-[33]. They also focused on specific AI technologies, such as data mining [34], fuzzy logic [35], or deep learning [36]. Accounting and finance researchers are keen on comprehending the application of AI in classic problems and the potential problems that AI technology can solve in the future. Therefore, the objective of this paper is to review previous research on the application of AI to some classical problems in the accounting and finance fields, as well as to explore future possibilities for further AI applications in these classical problems.

The contributions of our survey paper are as follows:

- We present a taxonomy of classic problems in accounting and finance and establish their connection with relevant artificial intelligence technologies. This taxonomy would help artificial intelligence researchers navigate the complex domain of accounting and finance.
- We survey a large of existing literature on applying artificial intelligence technologies to address these classic problems and reveal the advantages and disadvantages of applying each AI method to each classical problem.
- We summarize the progress and highlight the future work present in the current implementation of artificial intelligence in accounting and finance. This provides insights to improve the existing models and make them more adaptable for practitioners and regulators.

The remainder of this paper is organized as follows. Section 2 introduces some classic accounting and finance problems, and analyzes the requirements of accounting and finance for AI technologies. Section 3 explores the opportunities and challenges presented by AI. This section concentrates on the broader issues that AI technologies can address. Section 4 conducts a qualitative review of existing research on AI in accounting and finance. This section focuses on reviewing AI applications to classical problems and analyzing unresolved issues. Finally, Section 5 summarizes the challenges inherent in classical accounting and finance problems, as well as the problems that could potentially be solved using AI techniques in the future, and calls for further research.

II. OVERVIEW OF MODERN FINANCE AND ACCOUNTING

Traditional finance and accounting encounter several challenges in data collection, data processing, and objective decision-making. This section will explicate these issues from a classical perspective and elaborate on the specific aspects of accounting and finance development that mandate the integration of AI.

A. CLASSICAL PROBLEMS AND CHALLENGES IN FINANCE AND ACCOUNTING

This section will present the challenges faced by the accounting and finance field from multiple perspectives, including but not limited to financial statement analysis, investment and financing decisions, corporate value assessment, asset pricing, risk management, financial investment, and fraud and financial crime.

1) FINANCIAL STATEMENT ANALYSIS

Financial statements are crucial materials for effective internal management and external investor decision-making [37]. By analyzing financial statements, stakeholders can gain a comprehensive understanding of the enterprise's financial position, operating results, and other information [38]. Financial statement analysis is widely employed by auditors to assess the accuracy of corporate accounting information [39], [40]. Some researchers contend that financial statement analysis plays a crucial role in predicting corporate financial distress. They argue that the level of asset impairment provisions is closely linked to corporate financial distress, and significant reductions in intangible assets, capital expenditures, or dividend payments are likewise associated with high levels of financial distress [41], [42]. Additionally, financial statement analysis can assist investors in making better investment decisions by analyzing the effectiveness of financial indicators in predicting performance or operational risks [43]. Investors also use non-financial information from financial statements to evaluate management's subjective assessments of corporate performance to mitigate information asymmetry caused by agency issues [44]. However, financial statement analysis as a traditional analytical tool may at times be ineffective in serving its users. This is because some financial ratios may be outdated or fabricated, making it challenging to perform a comprehensive and objective assessment of a company's overall financial situation [45]. Earlier studies relied on economic intuition to analyze financial ratios. For instance, auditors use more industry experience when detecting financial statement fraud [46], which entails a more subjective judgment. Moreover, as financial statement analysis of a particular issue typically utilizes a series of financial ratios, the data are often downscaled before analysis, which poses a challenge for traditional financial analysis methods.

2) CORPORATE INVESTMENT AND FINANCING DECISION-MAKING

Corporate investment is a crucial factor in assessing the potential for growth and success, while financing plays a vital role in ensuring a company's ability to sustain its operations. The net present value (NPV) method is considered one of the fundamental tools in traditional financial management theory for formulating investment plans [47]. Its formulation is as follows:

$$NPV = \sum_{t=0}^{n} \frac{NCF_t}{(1+r)^t} \tag{1}$$

Where NPV denotes the net present value, NCF_Tdenotes the net cash flow generated by the innovation project in year t, and r denotes the discount rate, respectively. The crucial aspect of investment projects lies in forecasting their associated cash flows. Previous research has revealed that cash flow forecasting is highly subjective due to the influence of decision-makers intuitive judgment and industry experience, which are in turn influenced by personal experience, cognition, and ability [48]. Since future cash flows are inherently uncertain and can be influenced by various market and political factors, predicting them with scientific accuracy is challenging. Moreover, in an uncertain environment, investment decisions may not be solely driven by NPV or IRR, and the problem of achieving multi-objective optimization needs to be addressed [49]. The ideal capital structure is the state where enterprises can make financing decisions to achieve the lowest cost of capital, maximum enterprise value, and highest stakeholder motivation [50]. Therefore, optimizing an enterprise's capital structure is a crucial topic in traditional accounting research. Previous studies have investigated factors that influence capital structure and estimated optimal capital structure using linear models [51], but the relationship between these factors and capital structure may not be linear [52]. Thus, technologies capable of handling complex relationships are demanded.

3) CORPORATE VALUATION

Company valuation is a critical component of many financial activities, which has attracted the attention of both businesses and investors [53]. In addition to its use in restructuring, mergers, and leveraged buyouts, corporate value assessment assists securities analysts in identifying undervalued stocks. The discounted cash flow method is one of the most traditional approaches for determining corporate value [54]. The method's specific formulation is as follows:

$$\begin{gathered} V_T = \frac{FCF_{T+1}}{WACC - c}, \\ FCF_{T+1} = NOPLAT_{T+1} - \left(INVCAP_{T+1} - INVCAP_T\right) \end{gathered} \tag{2}$$

Where FCF_{T+1} is free cash flow at the end of year T+1 and WACC is the nominal weighted average cost of capital of the company. Note that free cash flow equals the difference between net operating profit less adjusted taxes (NOPLAT) and the increase in invested capital (INVCAP). Initially, free cash flow forecasting relied on the industry experience and intuitive judgment of decision-makers [55]. Later, researchers employed statistical methods, such as ARIMA models, for cash flow forecasting [56]. However, this method requires stable time series data as input, and its ability to capture patterns is limited if the data are unstable. Moreover, this method only considers linear relationships, and in complex and variable environments, the predicted cash flows tend to deviate from actual values. Furthermore, statistical methods often perform poorly when dealing with multiple variables due to the difficulty in handling complex nonlinearities. Therefore, the discounted cash flow method still presents numerous challenges when used in practice.

4) FINANCIAL ASSET PRICING

Financial asset pricing and forecasting play a vital role in financial research and have a significant impact on policymakers, investors, and households [57]. However, asset pricing faces several challenges, such as determining the expected return function, measuring the risk premium, considering the impact of the external macro environment, and incorporating new behavioral finance explanations, which are difficult to address through conventional methods [58]. For instance, in the case of the external macro environment, traditional statistical models struggle to identify the state of the economy from vast amounts of macroeconomic time series data [59]. Additionally, most theoretical and empirical research in asset pricing assumes rational expectations [60]. However, examining finite rationality may offer more insightful results, but conventional approaches may not be suitable for capturing the expectation functions. Financial asset price forecasting is a crucial element in portfolio optimization for investors. Previously, regression analysis methods were utilized for predicting stock prices, but their effectiveness is limited as they do not account for the stochastic nature of stock price changes [61]. Additionally, financial data are susceptible to noise, and reducing such noise manually is challenging.

5) E: RISK MANAGEMENT

Financial risk management is a challenging task for financial institutions. Financial modeling of market risks, which aims to manage losses due to changes in financial market prices, has been a topic of study for several decades. While most prior research has examined a single or harmonized source of risk and its impact [62], market risk may involve multiple sources of risk and a large amount of data, rendering traditional methods often ineffective in regulating and evaluating market risk. In addition, the healthy development of financial markets requires the regulation and assessment of credit risk, which poses several challenges. These challenges include the incorporation of a company's credit rating, the selection of relevant proxy variables describing the company's financial position over a specific period, and the estimation of the likelihood of a company joining a high-risk category or going bankrupt in the coming years. These challenges include high-dimensional characteristics of proxy variables for a company's financial condition [63] and the possibility of fraudulent data from companies seeking more loans [64], posing a challenge to traditional risk assessment techniques.

6) FINANCIAL INVESTMENT

Investors' enthusiasm for financial investment grows in response to the rapid development of financial markets and the continuous innovation of financial products. Portfolio optimization can assist investors in risk hedging. The most famous portfolio model is the mean-variance model proposed by Markowitz in 1952 [65]. The Markowitz model is a widely recognized approach that seeks to optimize the allocation of assets in a portfolio by striking a balance between risk and return. The model assumes that investors are risk-averse and will attempt to maximize their return for a given level of risk or minimize their risk for a given level of return. The model incorporates the expected returns and covariance of assets in the portfolio to determine the optimal weight to assign to each asset. Expected return refers to the average return expected from an asset, while covariance is a measure of the extent to which the returns of one asset are related to those of other assets. The Markowitz model aims to minimize the variance of the portfolio, which is a measure of the portfolio's risk while maximizing its expected return. The model identifies the set of portfolios that lie on the efficient frontier, which refers to the set of portfolios that provide the highest expected return for a given level of risk or the lowest risk for a given level of expected return and represents the trade-off between risk and return. Despite its success and recognition through the Nobel Prize, the mean-variance model has been subject to criticism regarding its practical implementation. One shortcoming is the efficiency of the solution and the difficulty of solving the model due to the cardinality constraint. Another issue is the model's applicability, as its premise assumptions are overly idealistic and may not hold in real-world scenarios. Lastly, the stability of the model is a concern, as parameter sensitivity and estimation errors can result in an unstable solution set. Several studies have pointed out these limitations, which pose challenges to the model's use in practical settings.

7) FRAUD AND FINANCIAL CRIME DETECTION

Financial organizations and governments have long been concerned with the prevention and detection of fraud and financial crime. Each year, losses from fraud and financial crime total trillions of dollars, with considerably greater potential losses [66]. However, because of the concealed and various nature of fraud and financial crime, as well as the criminals' ever-increasing technological means, their illegal tactics have become more difficult to deal with. Corporate fraud and money laundering as major forms of fraud and financial crime [67]. Initially, detecting corporate fraud was more subjective and relied more on the auditor's industry experience. Then, researchers are beginning to use statistical methods such as multiple regression models to identify factors that may be associated with corporate fraud [68], however, these statistical methods face numerous challenges in solving the problem, such as a lack of causal inference capabilities and an inability to handle textual data. As a result, more advanced and rigorous procedures are required to identify fraudulent companies effectively.

Meanwhile, due to the enormous losses caused to the economy and society, anti-money laundering has always been a promoted activity by governments and financial organizations, and governments have long formulated various relevant laws to regulate all types of financial activities in order to reduce financial crime. Since 2003 the regulator PBC(The People’s Bank of China) has promulgated stipulations to monitor suspicious transactions related with money laundering. However, relying on the law to regulate is insufficient. In the past few decades, an operational model that is based on embedded rules was very popular among the community where the rules were simple and easy to code, and were designed by consultants and domain experts who implement their working experience into the automated decision process [69]. The problem with this technique is that as the number of scenarios increases, more rules are required to detect money laundering, yet integrating new rules with old ones degrades system performance. Also, it is very difficult to identify all the scenarios. Large capacity data sets of various sorts of structured and unstructured data also appear to be a catastrophic challenge for rule-based systems with the arrival of the big data era. As a result, more advanced technology should be used for detection of money laundering operations.

B. DEMAND FOR COMPUTATIONAL INTELLIGENCE IN FINANCE AND ACCOUNTING

1) DEMAND FOR DATA ACQUISITION

The data acquisition method of the traditional accounting and finance industry usually is manual input, which might be time-consuming and laborious. Automatic data acquisition can generate accounting information quickly, reducing data errors caused by human error, such as automatic recognition of invoices and automatic entry of vouchers. Furthermore, in the age of big data, the sources of decision-making information are diverse, and the manual method of obtaining data information has significant limitations. Traditional accounting research generally uses non-standard data, as those non-standard data are often difficult to obtain. When auditors perform audit procedures, for example, they cannot obtain sufficient audit evidence solely from financial information and must obtain specific accounting transaction records other than traditional records using technical means such as crawlers [70].

2) DEMAND FOR INTELLIGENT DATA PROCESSING

Automated processing: The acquisition of data in the traditional accounting and finance industry is typically performed through manual input, a process that can be time-consuming and labor-intensive. The implementation of automatic data acquisition techniques has the potential to generate accounting information in a timely and efficient manner, thereby minimizing data errors that may result from human intervention. In the era of big data, the sources of decision-making information are diverse, making the manual acquisition of data increasingly impractical. Furthermore, traditional accounting research often involves the use of non-standard data sources that may be difficult to obtain. In performing audit procedures, for example, auditors must often obtain specific accounting transaction records that fall outside the scope of traditional financial information. To this end, technical tools such as crawlers may be employed to supplement traditional data collection methods [70]

Computational efficiency: There is a greater demand for real-time decision-making due to the rapid growth of accounting and finance. Traditional models and algorithms may no longer be appropriate for solving some problems in real-time decision-making. Traditional models and algorithms, for example, typically only obtain optimal solutions in a given period and incur high calculation costs when solving portfolio optimization problems [71]. Other aspects also reflect the demand for real-time decision-making. Previously, futures trading was done by searching for similar cases using expert systems and then comparing them for trading. This line of reasoning, however, does not apply to real-time trading because real-time data contains far too much information. Expert systems usually cannot perform real-time analysis well and need faster algorithms to support it.

Trustworthy data pre-processing: Financial data often contain significant levels of noise, such as market microstructure noise and financial statement noise, which can make it difficult to extract meaningful information. One crucial concern is whether the information provided by a company is reliable and accurate. From high-profile cases such as Enron to more recent scandals such as Luckin Coffee, financial fraud poses a significant challenge to traditional detection methods. In addition to intentional manipulation by humans, unconscious errors may also occur, making them difficult to detect due to their hidden nature. As a result, the development of effective techniques to detect, correct, and provide early warning for such issues using AI has become a critical area of research in the accounting and finance fields.

3) DEMAND FOR INTELLIGENT DECISION-MAKING

Some solutions in accounting and financial research do not always go through the same rigorous proof process as those in science and technology, instead relying on the experience of accountants and finance professionals [72]. There is an industry trend on how to extract the experience of these people and utilize it for intelligent decision-making. The extraction of feature factors is a critical task in the detection of financial statement fraud. In practice, auditors typically rely on industry knowledge to focus on those business processes [46]. The advancement of AI technology can transform these experiences into feature factors, such as the greater a company's asset impairment provisions and financial asset investments, the greater the likelihood of financial fraud [73]. Similarly, venture capital firms require new technologies to extract factors that influence the quality of their investment decisions. They employ machine learning algorithms to determine that the most important factors to consider when investing in a startup are strategic planning and team management [74].

C. SUMMARY

This section provides an overview of the key challenges in accounting and finance, which are often inadequately addressed by conventional methods without AI models. These difficulties encompass a range of issues, including the subjective nature of decision-making, low accuracy, difficulty in processing large datasets, low reliability, and robustness in different working conditions, and limited practicality. Additionally, our comprehensive review of classical problems in finance and accounting reveals an increasing need and demand for the application of AI in the accounting and finance domains. This includes the requirements for automatic data acquisition and processing, greater computing power, better logical reasoning and querying capability, more advanced statistical models, better nonlinear information processing capability, and more intelligent decision-making.

III. EMERGING TOOLS IN AI TO ADDRESS CHALLENGES IN ACCOUNTING AND FINANCE

This section outlines the potential applications of AI in accounting and finance. Specifically, we categorize the problems that can be addressed by AI techniques into three groups: classification, clustering, and regression. The classification, clustering, and regression methods are illustrated in Figure 1, which represents the corresponding categories for each of the problems outlined in the previous section.

A. CHARACTERISTICS OF AI AND THE CORRESPONDING OPPORTUNITIES

1) AUTOMATIC OPERATION

AI technologies can improve the efficiency of accounting and finance business processes by streamlining process processing. In recent years, robotic process automation (RPA) has been the most widely discussed technology in accounting process automation [75]. AI technologies, such as invoice OCR recognition and contract NLP semantic analysis, have expanded the capabilities of RPA [76], making RPA smarter. AI technologies enable automated decision-making in addition to process automation. In [77], the authors employed neural networks to create multiple classifiers that automate consumer credit risk assessment and provide a target population for credit operations. To address the matching between the operational records of bank branches, a business analytics framework was built utilizing machine learning approaches to make an initial judgment of a business's authenticity in [78]. In addition, AI is widely used in automated trading systems, particularly in high-frequency trading. Current automated trading systems are used in conjunction with price prediction models. As a result, most price or trend forecasting models that generate buy and sell signals based on their forecasts are classified as automated trading systems. Karaoglu & Arpaci proposed a recurrent neural network with Graves LSTM layer for automated stock trading and used feature indicators related to market microstructure as an input layer in [79]. Bao et al. combined wavelet transform (WT), stacked auto-encoders (SAEs), and long short-term memory (LSTM) to build a new deep learning framework for price prediction in [80].

2) DATA-DRIVEN TECHNOLOGY

AI techniques can extract features of the data itself and thus discover some patterns. Data mining is an effective data analysis tool that can provide managers with logical and causal relationships in company data for proactive problem-solving [81]. It can also effectively predict future business trends, assisting managers in making better decisions and increasing the company's competitiveness [82]. In [83], it used data mining methods to select features without considering financial and accounting theories and then used these features to build a financial distress detection model, finding no significant differences in predictive performance between financial theory prediction models. Murray et al. extracted behavioral patterns from historical customer data by applying data mining methods in [84]. The revealed behavioral patterns and subsequent market segmentation can help managers in strategic decision-making. In [85], the author performed a textual analysis of corporate stakeholders' opinions on Twitter and proposed an alternative method to assess corporate social responsibility (CSR) fulfillment.

3) OBJECTIVE AND ACCURATE SOLUTION

Traditional accounting and finance fields rely on practitioners' problem-solving and decision-making skills. This experience is typically highly correlated with the practitioner's ability, experience, and industry knowledge, which can lead to more subjective decisions. To some extent, AI techniques can help with this problem. Frankel et al. compared the ability of a financial dictionary-based approach and a machine learning approach to capture disclosure sentiment and found that machine learning provides a more objective way to analyze sentiment in [86]. Traditional operational decisions are primarily based on managers' subjective perceptions of the overall market. However, due to the subjective nature of the perception and the difference in the perception by different managers, this strategy inevitably lacks generality. In [87], it utilized artificial intelligence techniques to synthesize a decision-making system based on production data and the experience of decision managers to help plan production objectively and efficiently.

B. CHARACTERISTICS OF PROBLEMS SOLVABLE BY AI

1) CLASSIFICATION

Classification is defined as assigning an unknown observation represented by a specific feature from a limited set of categories to a correct category. Categories for classification, unlike clusters, have already been trained and tested. After completing training for a specific classification task, a classification application can classify a newly observed object. There are two types of classification: binary classification and multivariate classification. A more common example of a binary classification problem is how banks classify loan applicants as creditworthy or non-creditworthy and then decide whether or not to approve loan applications. In this case, we can train the machine learning application using the repayment history of previous loan applicants as well as other business data. Classifiers are also useful for multi-class classification. For instance, based on data such as investors' trading history and published comments, sentiment analysis models can classify investor sentiment as highly positive, positive, neutral, negative, and highly negative. Table I shows some typical methods of applying artificial intelligence to solve classification problems, as well as the application examples.

TABLE I
TYPICAL METHODS TO SOLVE CLASSIFICATION TASKS AND APPLICATION EXAMPLES

Typical method	Application example
Naive Bayes	Corporate fraud detection [88]
	Text classification [89]
	Sentiment analysis [90]
K-nearest neighbor	Bankruptcy prediction [91]
	Market manipulation [92]
Decision trees	Corporate Fraud detection [93]
	Bankruptcy prediction [94]
Random forest	Stock market investment [95]
	Credit fraud detection [96]
Support vector machines	Bankruptcy prediction [97]
	Facial emotions detection [98]
Recurrent neural network	Bankruptcy prediction [99]
	Sentiment analysis [100]

TABLE II
TYPICAL METHODS TO SOLVE CLUSTERING TASKS AND APPLICATION EXAMPLES

Typical method	Application example
Hierarchical Clustering	Brand strategy cluster [101]
	Social impact assessment [102]
	Corporate fraud detection [103]
K-means Clustering	Portfolio optimization [104]
	Stock customers segmentation [105]
Fuzzy C-means Clustering	Portfolio optimization [106]
	Performance evaluation [107]

2) CLUSTERING

Clustering is the division of a data set into different classes or clusters based on certain criteria (e.g., distance). Clustering is used to group data with similar characteristics into the same group; the greater the similarity within the group and the greater the difference between the groups, the better the effect of clustering. According to different clustering criteria, clustering methods can be grouped into Partition Clustering methods, Density-based Clustering methods, Hierarchical Clustering methods, and other new methods. Table II shows some typical methods of applying artificial intelligence to solve clustering problems, as well as application examples

3) REGRESSION

Regression is the prediction and modeling of numerical continuous random variables. A major difference between classification and regression is that classification is used to infer rules or equations from data relevant to the current situation (e.g., detecting credit card fraud or financial fraud). A regression model, on the other hand, can forecast what will happen in the future based on current data. In stock price prediction, for example, the relevant machine learning program may examine the data to predict market signals that will influence future stock price trends. Table III shows some typical methods of applying artificial intelligence to solve regression problems, as well as application examples.

TABLE III
TYPICAL METHODS TO SOLVE REGRESSION TASKS AND APPLICATION EXAMPLES

Typical method	Application example
Decision tree	Corporate performance prediction [108]
	Stock price prediction [109]
K-nearest neighbor	Stock index prediction [110]
	Liquidity prediction [111]
Long Short-Term Memory	Stock market prediction [112]
Convolutional neural networks	Bitcoin price prediction [113]

C. SUMMARY

This paper summarizes some of the opportunities that AI presents to the accounting and finance fields in this section. We discovered that AI can assist in the automation of many simple and repetitive processes. Simultaneously, AI technology can extract critical information from large amounts of data in order to make more accurate and objective decisions. It can also handle highly complex relationships, allowing it to provide a more dependable solution to a variety of problems. Furthermore, we examine the general AI problems of classification, clustering, and regression and summarize some intelligent techniques commonly used to solve these problems.

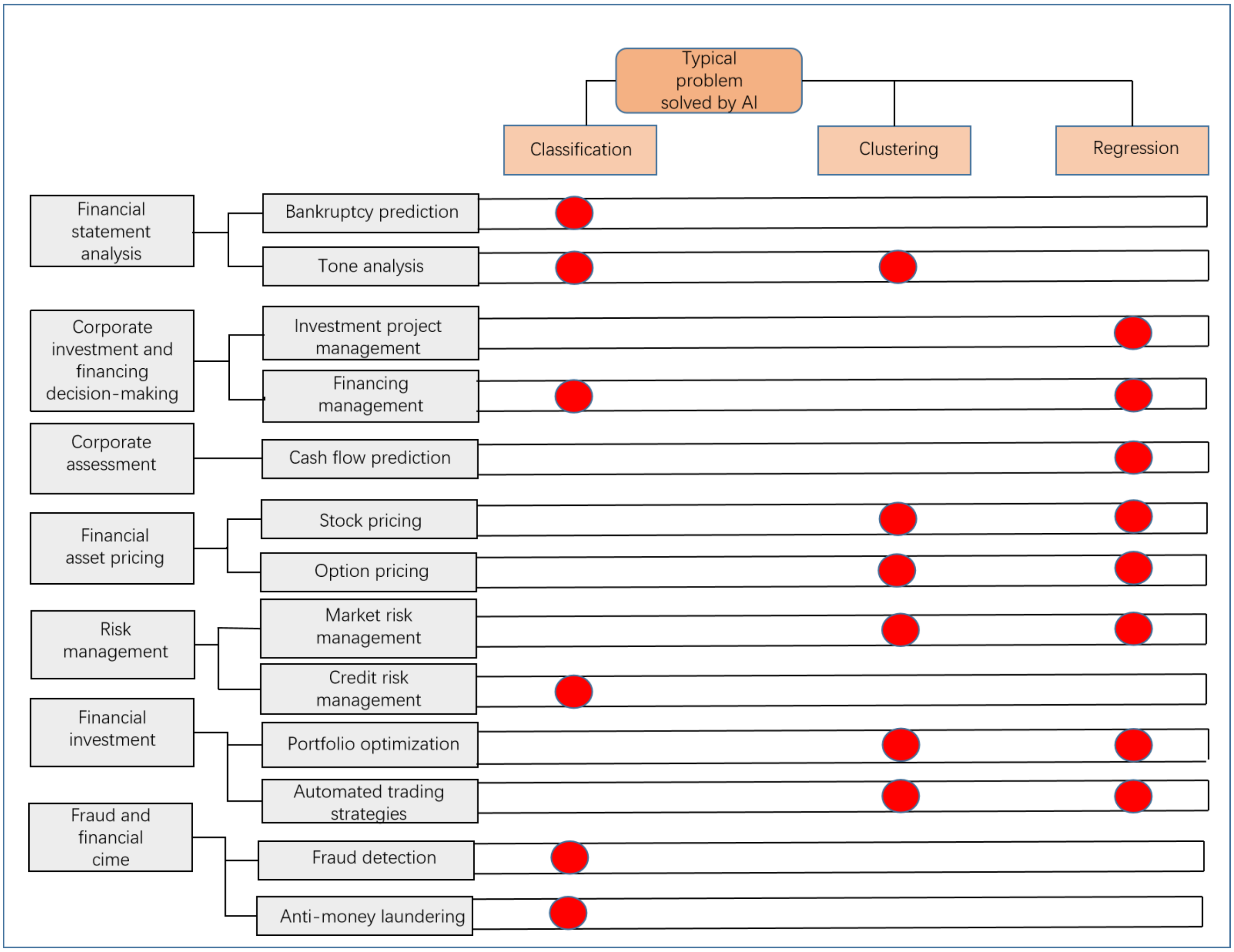


FIGURE 1. **Categories of problems solvable by AI and classification of the classical problems in accounting and finance**

IV. AI Solutions for Finance and Accounting

The taxonomy shown in Fig. 2 can be used to organize accounting and finance tasks into different categories based on their characteristics and objectives. By categorizing these tasks, it becomes easier to identify relevant AI techniques that can be applied to solve them. The taxonomy includes categories such as financial statement analysis, fraud detection, credit risk assessment, and portfolio optimization, among others. By examining each of these categories in more detail, we can gain a better understanding of the specific challenges and opportunities that exist within each area, and the AI techniques that are best suited to address them. The following sections will provide a comprehensive overview of each category, including specific examples of relevant AI techniques and their applications.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3333389

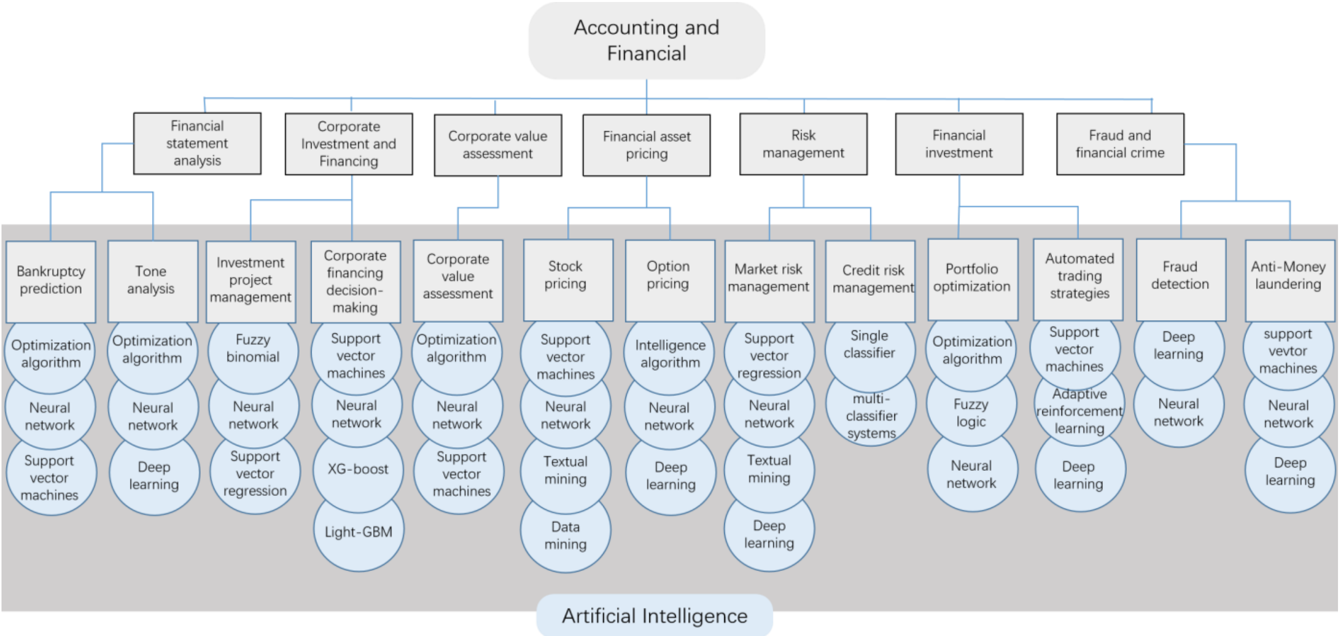


FIGURE 2. **Artificial intelligence for accounting and finance: task taxonomy**

A. FINANCIAL STATEMENT ANALYSIS

1) BANKRUPTCY PREDICTION

How to predict corporate financial risk using disclosed financial data and determine the likelihood of financial distress is a pressing issue for investors and regulators to address. In financial insolvency prediction models, the selection of financial ratios and the design of classifiers are crucial. In [114], the authors used variables such as profitability, leverage level, efficiency, and valuation ratio to construct a probabilistic model of financial distress, which has an accuracy of 87.93%. In addition to financial ratios, corporate governance indicators are important input variables [115]. In [116] examined the relative importance of various bankruptcy predictors commonly used in the existing literature and found that financial ratios do not contain all information for bankruptcy prediction. Textual information is equally valuable for predicting corporate financial distress because the textual content of financial statements may contain emotions or tones. Several researchers have worked on using textual disclosure information in bankruptcy prediction models. A deep learning model is developed for corporate bankruptcy prediction that includes textual disclosure information and confirmed the higher accuracy of bankruptcy prediction models with textual information using a large sample dataset in [117].

Previously, discriminant analysis and logistic regression were the most commonly used models to predict financial distress. However, the machine learning method, according to the report, is more accurate than traditional bankruptcy prediction models [118]. Considering the low explanatory power and poor stability of traditional statistical models in prediction, a support vector machine has been proposed to solve these problems and find the optimal parameter values of the SVM kernel function by 5-fold cross-validation [97]. Subsequent studies have combined some optimization algorithms with support vector machines for bankruptcy prediction, such as genetic algorithms [119], particle swarm optimization algorithms[120], and switching particle swarm optimization algorithms [121]. In [122], the authors compared the performance of logit, probit, linear discriminant analysis, neural networks, support vector machines, and other classifiers, and generalized boosting, AdaBoost, and random forests stood out among these methods due to their relative ease of estimation and implementation. And the researchers examined the performance of 11 classifiers using different datasets and found that each classifier has a varied representation in [123]. Therefore, we need to find integrated models to apply to various datasets. To address the difficulties of traditional models in dealing with time-series data, in [99], proposed an integrated model combining recurrent neural network (RNN) and long short-term memory (LSTM) for predicting the probability of corporate bankruptcy in a year, but the drawback of this approach is that cannot obtain the sensitivity coefficients of individual variables. Table IV presents the methods of these related works and their advantages.

2) TONE ANALYSIS

Sentiment analysis in financial report narratives has received increased attention in recent studies, and financial statement analysis is not limited to financial ratio analysis. The most common sentiment measure in the previous literature is based on dictionaries [124]. However, in their study, the dictionaries are built on publicly available data from the United States, which makes the lexicon analysis less scalable due to the different institutional and market environments between countries. Therefore, some researchers transformed the narrative part of financial statements into quantitative metrics and performed tone analysis through AI techniques. In [125], it used a text mining approach to identify the tone of 10-K narratives for all US companies from 1996-2010 and found that the tone changes depending on the company's revenue. particle swarm optimization (PSO) as a search process for relevance-based feature selection based on word frequency analysis is used to improve the accuracy of text sentiment classification in [126]. In addition, neural networks may outperform many other machine learning techniques in sentiment analysis [100]. In [127], the authors used recurrent neural networks (RNN) to measure the sentiment of a large sample of 10-k narratives with an accuracy of over 90%. In addition, deep learning is superior to dictionary-based methods, but its accuracy is lower than trained manual coding [128]. Actually, Deep learning refers to a neural network containing a multi-layer network structure [129]. Shallow neural networks are those shown in Figure 1, typically with only one or a few hidden layers. They have value and applications in accounting and finance, particularly when the amount of data is minimal or the model interpretability requirements are high. Deep learning, on the other hand, encompasses a broader range of approaches and deeper network architectures. Deep learning has outperformed traditional methods in many areas as computing power and data have increased.

Besides, the feasibility of manual coding is not high for big data, which is the difficulty of AI-enabled annual report tone analysis, so we need to develop intelligent tone analysis systems with higher accuracy. In [130], AI algorithms are applied to judge and identify the positive and negative aspects of textual information when considering the impact of textual information on bankruptcy prediction. Table IV presents the methods of these related works and their advantages.

TABLE IV AI METHODS APPLIED IN FINANCIAL STATEMENT ANALYSIS AND THEIR ADVANTAGES

Method	Application	Advantage
Cox Proportional Hazard model	bankruptcy prediction[114]	The method has a good fit and a prediction accuracy of 87.93%
Support vector machines	bankruptcy prediction[115], [97]	Compared to Discriminant Analysis, this method gives more robust results.
Average Embedding Model	bankruptcy prediction[117]	The model yields superior prediction performance in forecasting bankruptcy using textual disclosures, and outperforms convolutional neural networks
GA-SVM	bankruptcy prediction[118], [119]	The model was effective in finding the optimal feature subset and parameters of SVM, and that it improved the prediction of bankruptcy.
EABC-PGSVM	bankruptcy prediction[120]	The method could not only save CPU time but also enhance the hit ratio
SPSO-PGSVM	bankruptcy prediction[121]	Compared to traditional support vector machines, this method greatly improves the explanatory power and stability of the predictive model
AdaBoost	bankruptcy prediction[122]	The method is robust to outliers and missing values and appears largely insensitive to the monotone transformation of input variables.
RNN-LSTM	bankruptcy prediction[99]	The model is better than other methods in dealing with highly precise tasks
Cluster – robust regression model	tone analysis[124]	This method is simpler and more readable
PSO	tone analysis[126]	This method allows for more efficient searching of text features
LTSM	tone analysis[127]	Compared to Bayesian and other neural networks, this method obtains higher accuracy
CatBoost model	tone analysis[130]	This method is very good at capturing complex feature relationships

B. CORPORATE INVESTMENT AND FINANCING DECISION-MAKING

1) INVESTMENT PROJECT MANAGEMENT

The difficulty of investment project management is how to estimate the cash flow and capital cost, and then decide whether to invest or not. Traditionally, investment project selection performs under the assumption of certainty; however, in the real world, decision-makers in project selection may encounter both stochasticity and fuzziness. To address these issues, the researchers treat the investment expenditure and annual net cash flow as stochastic fuzzy variables and develop a constrained model with stochastic and fuzzy parameters based on the net present value method in [131]. Then they combine a genetic algorithm (GA) and stochastic fuzzy simulation to solve this stochastic fuzzy planning problem. Meanwhile, the traditional NPV approach ignores future management flexibility, which may lead to the undervaluation of the project. However, statistical models do not estimate management flexibility effectively. Therefore to evaluate projects more accurately, in [132], the authors considered flexibility as a potential option and calculated it using a fuzzy binomial valuation technique. In addition to the net present value method, some studies attempted to develop project revenue forecasting models with multiple factors to aid project decisions [133], [134]. In addition, investment project selection is often multi-objective optimization, and considering NPV maximization alone may miss some good investment projects [135]. The difficulty of planning problems with multi-objective optimization lies in computing the optimal solution, and some optimization algorithms in AI can solve the issue. Time estimation is equally essential for investment project management, where failure to complete the project on time can lead to investment failure. For example, new product development projects must consider time, so an integrated support-vector regression model is introduced to improve the accuracy of time estimation in new product development projects while using ICA to enhance the speed of computation in [136].

Another important topic in investment project management is the assessment of investment risk, which is the instability of returns. Higher-than-expected returns, on the other hand, do not elicit the same emotions in decision-makers as lower-than-expected returns. As a result, simply using variance to measure risk is biased. In [137], the authors considered project costs and use artificial neural networks to output the associated costs and risks. In [138], a new OFDI metric is employed to predict future overseas investment risk by training the previous year's ICRG data and the next year's OFDI data with a KNN algorithm to obtain predictive power. Table V presents the methods of these related works and their advantages.

2) CORPORATE FINANCING DECISION-MAKING

In general, corporate financing channels are diverse, and the proportion of each financing determines the capital structure. Many scholars have conducted a series of studies on the economic consequences of capital structure, including firm performance, investment efficiency, and firm value, since the introduction of MM theory. With more research, the emphasis gradually shifted to the factors influencing capital structure, and early studies on the factors influencing capital structure typically employed multiple regression models. Traditional multiple regression models, on the other hand, frequently contain assumptions, such as the normal distribution of independent variables, that cannot be met in practice. As a result, some studies are looking for new methods. In [139], it used generalized regression neural networks (GNNs) to test the influencing factors of capital structure and found that this method has less error than traditional multiple regressions. Meanwhile, determining the optimal capital structure is a hot topic. Traditional linear panel data models are ineffective because the relationship between optimal capital structure and related determinants is usually complex or nonlinear. To capture biases and achieve better forecasting results than multiple regression models, artificial neural network models have been widely used to fit complex relationships between dependent and independent variables [140], [141]. However, these technical methods cannot obtain the sensitivity coefficients between single factors and provide managers and other stakeholders with obvious recommendations. In [142], the authors collected 224 financial and non-financial indicators to test their relationship with heavy polluters' debt financing ability and used the XGboost method to find the top six indicators that affect heavy polluters' debt financing ability in a high-dimensional condition, which provides a reference for corporate financing decisions.

Decision-makers must consider the financial risk as a result of changes in financing decisions because excessive financial risk can lead to financial distress. Traditional finance theory employs leverage to assess a company's financial risk, but financial institutions provide funding that takes into account not only accounting data in financial reports, but also the strategic level of corporate integrity, competitive environment, social responsibility, and development plans. Therefore, technical is needed to assess financial risk more correctly. In [143], the authors compared the effectiveness of several commonly-used neural network models in Chinese SME datasets and found that probabilistic neural networks (PNN) have higher prediction accuracy for financial risks. Support vector machines (SVM) have a strong generalization ability to map nonlinear relationships to linear relationships in a high-dimensional space and thus can be better applied to early financial risk warnings [144]. In [145], the big data mining techniques combined with information fusion methods for early financial risk warning, and the study found that this model outperformed the support vector machine. A Light GBM-based financial risk prediction model is proposed and the study compared it with K-nearest neighbors, decision trees, and random forests, which had better prediction accuracy in [146]. Table V presents the methods of these related works and their advantages.

TABLE V AI METHODS APPLIED IN CORPORATE INVESTMENT AND FINANCING MANAGEMENT AND THEIR ADVANTAGES

Method	Application	Advantage
GA	Investment project management[131]	This method can solve the stochastic fuzzy planning problem well, but it is time-consuming
Fuzzy binomial approach	Investment project management[132]	The methodology facilitates investment decisions in uncertain environments
The least squares support vector machine	Investment project management[133]	Compared with back propagation neural network and radial basis function neural network, the method has higher performance and accuracy
Support vector regression	Investment project management[136]	Compared with nonlinear regression, back-propagation neural networks, and general regression neural networks, the model achieves high estimation accuracy
Artificial neural network	Investment project management[137]	The method can apply to any risky economic project analysis
	Corporate financing decision-making [140],[40]	Capable of capturing complex non-linear relationships well
K-nearest neighbor	Investment project management[138]	Lower requirements for data volume
Generalized regression neural networks	Corporate financing decision-making[139]	Compared to multiple regressions, This method is better at explaining the non-

		linear relationship between variables
XGBoost	Corporate financing decision-making[142]	The method can recognize features in high-dimensional conditions
Probabilistic neural network	Corporate financing decision-making[143]	Compared to other neural network models, the method allows for a more accurate prediction of financing risk
DEA-SVM	Corporate financing decision-making[144]	Compared to SVM, This method has higher accuracy
The information fusion-based model	Corporate financing decision-making[145]	Compared to SVM and Logistic regression, this method obtains higher accuracy
LightGBM	Corporate financing decision-making[146]	Compared to k-nearest-neighbors algorithm, decision tree algorithm, and random forest algorithm, this method obtains higher accuracy

C. CORPORATE VALUE ASSESSMENT

Forecasting free cash flow is an important part of determining enterprise value. Choosing a scientific and reasonable method for forecasting free cash flow is a necessary prerequisite for an accurate evaluation of enterprise value. The ARIMA model is widely used to predict cash flow due to it can handle time-series data better. However, it is only suitable for short-term data and does not take into account factors such as periodicity and the number of exogenous variables. Such issues can be addressed by Facebook's Prophet forecasting tool, which was open-sourced in 2017. Recent studies have shown that the cash flow forecasting accuracy of the Prophet is higher than the ARIMA model [147], [148]. Among AI methods, Support vector machines can accurately predict time series data, especially when the data are nonlinear and non-stationary. Response Surface, Back Propagation Neural Network, Radial Basis Functions Neural Network, and Support Vector Machine models respectively by using the sliding window technique are built in [149], and the study found that support vector machine has the highest prediction accuracy in cash flow prediction. Since cash flow data are time-dependent, data collected at different time points may lead to inconsistent parameters of the trained models, resulting in instability of the prediction results. To address such problems, least squares support vector machines (LS-SVM) and adaptive time functions (ATF) are integrated in [150], and by doing so, LS-SVMAT can better handle the dynamic properties of time series. In [151], an integrated learning algorithm that combines the adaptive population activity particle swarm optimization algorithm and least squares method is proposed to optimize the parameters of the network-based fuzzy logic system (ANFIS) model and verified the effectiveness of the method for cash flow time series forecasting in the bank through simulation. In [152], a BP neural network model based on an improved genetic algorithm is used to predict the free cash flow of a firm and found that the average relative error of this method is less compared with other related forecasting models. Table VI presents the methods of these related works and their advantages.

TABLE VI AI METHODS APPLIED IN CORPORATE VALUE ASSESSMENT AND THEIR ADVANTAGES

Method	Application	Advantages
Support vector machines	Corporate value assessment[149]	The method can be adapted to cash flow forecasts based on small samples
Least Squares Support Vector Machine	Corporate value assessment[150]	The method can better deal with the dynamic nature of the time series and has higher accuracy than SVM
APAPSO prediction algorithm	Corporate value assessment[151]	The method improved the premature convergence of PSO algorithm
BP neural network based on improved genetic algorithm	Corporate value assessment[152]	Better training and learning capabilities in non-linear environments with short learning times

D. FINANCIAL ASSET PRICING

1) STOCK PRICING

Capital asset pricing models, discounted cash flow models and factor models remain the primary theories for international stock pricing (especially in more market-oriented countries). However, systematic risk may not be the most important factor influencing stock prices. Other non-diversifiable risks include company size, P/E ratio, and other company characteristics, so these pricing models need to be revised and improved. Meanwhile, as AI technology advances, new tools such as artificial neural networks for stock pricing have become commonplace. A multilayer feed-forward neural network is used to train and simulate forecast data generated by a CAPM model, which has 98-99% lower errors than OLS and GARCH in [153]. This study validates the theoretical validity of capital asset pricing models, but traditional regression and econometric methods are not enough to predict accurately. In [154], machine learning techniques are applied to forecast market equity risk premiums from CAPM models, which include generalized linear models, dimensionality reduction techniques, augmented regression trees, and shallow neural networks. To solve the equilibrium price-dividend function of the CAPM model, an improved trigonometric neural network is proposed in [155], which has the highest accuracy and the fastest speed compared with several methods for asset pricing models. In the practical application of factor models, the factor selection process based on regression methods is prone to overfitting problems. To solve this problem, neural networks are applied to detect interactions between factors in [156], while this method can cope with different degrees of nonlinearity in historical financial data.

Among the classic financial market prediction techniques, of particular note are the following: technical analysis with indicators calculated from past prices to indicate bearish or bullish trends [157] and fundamentalist analysis, which seeks economic factors that influence market trends [158]. Traditional fundamentalist analysis usually focuses on regression analysis of financial data disclosed by companies to predict stock prices, but regression methods usually cannot accommodate the stochastic nature of stock price changes. Stock price prediction is a nonlinear and time-dependent problem because financial data is influenced by numerous and constantly changing correlation factors. Deep neural networks (DNNs) combine the benefits of deep learning (DL) and neural networks and can be used to solve nonlinear problems more satisfactorily than traditional machine learning algorithms [159]. With the advancement of technology, structured data have become an important data source for fundamental analysis. Text mining techniques are used to extract features from financial news for stock price prediction in [17], which improves prediction accuracy due to the ability to select semantically relevant indicators compared to previous methods. The study of [160] is based on state-of-the-art text mining techniques to verify whether stock price movements can be predicted more accurately by including irrational indicators. Stock prices contain a lot of noise, which can bias the prediction results. A denoising scheme, which is an integrated independent component analysis (ICA) based on neural networks is proposed in [161]. The method first uses ICA for stock prices to generate independent components (ICs), then identify and removes the independent components containing noise, and finally uses the reconstructed variables containing less noisy information as input variables for the neural network model. The key to technical analysis is to identify technical indicators that strongly correlate with stock prices. In [162] a short-term stock price forecasting method based on technical indicators is proposed, where SVR and MLP are used to explain the relationship between stock prices and technical indicators. In [163], the authors used a hybrid artificial neural network (ANN) model containing harmony search (HS) and genetic algorithm (GA) to select the most relevant technical indicators, and improve the prediction accuracy. Table VII presents the methods of these related works and their advantages.

2) OPTION PRICING

Options are a significant component of the derivatives market. Researchers, speculators, and other traders all want to get a fair price for their options. There are only a few options we can get an exact price, and for the rest, we must use numerical methods. The Black-Scholes model is arguably the most classic mathematical model used to price and hedge options, first proposed by Black and Scholes to price European options and later modified by Merton to work with dividends in [164]. Scholes's model is an idealized model that does not perform as well in practice. For example, the model assumes that stock prices follow a continuous geometric Brownian motion, whereas, in reality, stock prices can jump. The neural network approach to this problem is very different that it does not make any distributional assumptions about the underlying variables. A neural network model with three output nodes is conducted in [165] to predict option price changes, an approach that produces better hedging parameters than standard option pricing models. In [166], the authors integrated a new hybrid asymmetric volatility approach into an artificial neural network option pricing model to improve the ability to predict the price of financial derivatives. A hybrid model, consisting of a parametric option pricing model and a nonparametric machine learning technique is used to price European index options homogeneously in [167]. In [168], the authors argued that generative bayesian learning can achieve better calibration and prediction performance than classical US option models.

In addition to neural networks, machine learning techniques such as XGBoost and LightGBM have also been applied to option pricing [169]. To highlight the power of these algorithms, they compare them with classical methods such as Black Scholes and Corrado-Su's history and implied parameters and find that both machine learning algorithms perform much better than classical methods. To overcome the drawbacks of expensive computational costs and unrealistic model assumptions of numerical methods, a deep learning model is used to predict Asian option prices in [170] and found that the trained deep learning model was very efficient compared to the three traditional methods. In addition, the liquidity of options leads to unbalanced option data, which makes option pricing more difficult. In [171], the authors proposed an option pricing model based on deep learning to overcome the problem of data scarcity and partial data, which can be used for any European call option and put option prices, regardless of money. Table VII presents the methods of these related works and their advantages.

TABLE VII AI METHODS APPLIED IN FINANCIAL ASSET PRICING AND THEIR ADVANTAGES

Method	Application	Advantages
Multilayer feed-forward neural network	Stock pricing [153]	The method has 98-99% lower errors than OLS and GARCH
Shallow neural network	Stock pricing [154]	The method can perform good predictions with small amounts of data
The improved trigonometric neural network	Stock pricing [155]	the model yields superior prediction performance in forecasting bankruptcy using textual disclosures, and and outperforms convolutional neural networks
Improved trigonometric neural network	Stock pricing [156]	The method can obtain the price-dividend function precisely, quickly, and feasibly.
Deep neural networks	Stock pricing [159]	The method can solve nonlinear problems more satisfactorily compared to conventional machine learning algorithms.

Text mining techniques	Stock pricing [160]	This method extracts useful features from text and improves prediction accuracy
Neural network based on independent component analysis	Stock pricing [161]	This method reduces stock price noise and improves data smoothness
Support vector regression	Stock pricing [162]	The method is better able to deal with non-linear and highly correlated relationships between indicators
Artificial neural network based on Harmony Search and Genetic Algorithm	Stock pricing [163]	The model mitigates the well-known problem of overfitting and underfitting ANN
Multilayer perceptron	Option pricing [165]	Helping the Black-Scholes model fit better into reality
Nonlinear neural network	Option pricing [166]	The model has the functional flexibility to capture the nonlinearities in financial data
Support vector regression based on homogeneity hint	Option pricing [167]	The model can tackle the challenges posed by the option data's extensive spectrum, noise, and multivariate and stochastic nature.
Bayesian neural network	Option pricing [168]	It can achieve better calibration and prediction performance than classical US option models
XGBoost and LightGBM	Option pricing [169]	Able to better handle the stochasticity and nonlinearity of the underlying asset
Back Propagation neural network	Option pricing [170]	Compared to the traditional method, it solves the problem more quickly and accurately
Deep feedforward neural network	Option pricing [171]	It can overcome the problem of data scarcity and partial data

E. RISK MANAGEMENT

1) MARKET RISK MANAGEMENT

Volatility, defined as the standard deviation of an asset's return, is a key indicator of market risk. This risk is typically managed through options, financial derivatives, or the construction of various portfolios. However, we cannot predict future market volatility. Volatility estimation is thus a critical issue in market risk management. Traditional models used for volatility forecasting include exponentially weighted moving average ((EWMA), autoregressive conditional heteroskedasticity model (ARCH), generalized autoregressive conditional heteroskedasticity model (GARCH), and stochastic volatility model (SV). These models then do not take into account the asymmetry of past return correlations and the fact that this model does not apply to multivariate time series analysis [172].

Hybrid machine learning techniques can improve the performance of traditional GARCH and stochastic volatility methods [173]-[175]. Artificial neural networks combined with GARCH and stochastic volatility models can produce more accurate forecasts. In [176], a hybrid artificial neural network with a traditional GARCH model is proposed and the study tested it using three Latin American stock exchange indices from Brazil, Chile, and Mexico. In their study, the model can achieve higher performance than the GARCH model. Combining recurrent neural networks with stochastic volatility models can also significantly improve forecasting performance [177]. In [175], a hybrid model based on the GARCH model is proposed, and they used support vector machines as a predictor for performing meta-learning. In addition, some researchers exploit sequential learning methods based on neural networks for volatility prediction. Long short-term memory (LSTM) is widely used in sequence analysis, and its performance is better in predicting long-series data compared to support vector machines when predicting volatility [178].

However, none of the preceding approaches took unstructured data into account. Many studies have shown that unstructured data (financial news, political news, etc.) contain many descriptions of uncertainty, which can be used to forecast volatility [179]. These textual data about financial markets may come from a variety of sources. One source is online stock message boards, which reflect public perceptions of stock longs and shorts. Bayesian algorithms have been used to accurately predict trading volume and volatility based on these message postings [180]. The implied volatility of news, such as uncertainty generated by disaster news (e.g., war, market crashes), may also be a strong predictor of return volatility. Support vector regression can estimate such predictors [181]. We observe that media news stories about companies also are good predictors of potential stock volatility. A hybrid model combining text mining techniques and GARCH models can help analysts automatically classify news that leads to greater stock market volatility in [182]. Table VIII presents the methods of these related works and their advantages.

2) CREDIT RISK MANAGEMENT

Credit risk research is broadly divided into two types of issues: credit scoring and bankruptcy (failure) forecasting. Credit scoring is the risk classification of retail borrowers (including personal loans and mortgages), whereas bankruptcy forecasting is the risk assessment of institutional borrowers (e.g. small businesses). Because corporate bankruptcy forecasting was discussed in the previous section of this paper, we will concentrate on credit scoring in this section. Earlier studies generally used statistical models for credit scoring of lenders, and these include linear discriminant analysis (LDA), logistic regression (LR), multivariate discriminant analysis (MDA), quadratic discriminant analysis (QDA), factor analysis (FA), risk index models, and conditional probability models, etc. However, due to some strict assumptions, such as linear divisibility, multivariate normality, independence of influencing factors, and already existing functional forms, traditional statistical models tend to ignore the complexity, boundaries, and interrelationships of various types of influencing factors. Therefore, a large number of classifiers based on machine learning are widely used in credit scoring.

Because of their superior ability to extract important information from complex data and handle nonlinear relationships, artificial neural networks play an important role in credit risk assessment. Multi-Layer Perceptron, Expert Mixture, Radial Basis Function, Learning Vector Quantization, and Fuzzy Adaptive Resonance model were used for credit rating, and the performance of the radial basis function was better after 10-fold cross-validation using the empirical data set [183]. In [184], a neural network model based on a back propagation-learning algorithm is proposed and the study trained it using real credit applications from the Australian credit approval dataset, which can effectively help in the automatic processing of credit applications. A new method of average random selection is used to optimize the data distribution in the dataset to improve the accuracy of the credit scoring model based on the MLP neural network in [185]. Support vector machines can help banks in credit rating their customers due to their better generalization ability [186]. Meanwhile, the clustering support vector machine is developed to address the disadvantage of the high computational cost of traditional support vector machines. In addition, we observed that some studies proposed hybrid-learning models which combine optimization algorithms and a single classifier for the credit rating of individuals or firms, improving the performance of a single classifier [187],[188]. In addition to single classifiers, multi-classifier systems are widely used to solve various classification problems, which mainly combine several individual classifiers to provide the final classification decision. In [189], an integrated model based on decision trees, combined with bagging techniques is conducted, aiming at the stability and accuracy of a single classifier. At the same time, in order to ensure the privacy of financial data, federated learning is used for credit rating in [190]. Table VIII presents the methods of these related works and their advantages.

TABLE VIII AI METHODS APPLIED IN RISK MANAGEMENT AND THEIR ADVANTAGES

Method	Application	Advantages
Support vector machines	Market risk management[173],[175]	The method can efficiently work with high-dimensional inputs to account for volatility long-memory and multiscale effects
	Credit scoring[186]	Have better generalization skills
Random forests algorithm	Market risk management[174]	Ability to adapt to high-frequency data
Hybrid Neural Networks-GARCH model	Market risk management[175]	The method is superior in handling time series data and improves the prediction performance of the GARCH model
Recurrent neural networks	Market risk management[177]	The method improved the predictive power of stochastic volatility models
Long short-term memory (LSTM)	Market risk management[178]	its performance is better in predicting long-series data compared to support vector machines when predicting volatility
Bayesian algorithms	Market risk management[179]	The method is insensitive to missing data and excels in extracting textual information
Support vector regression	Market risk management[181]	this method over ordinary least squares (OLS) is its ability to deal with a large feature space
Text mining techniques	Market risk management[182]	It can help analysts automatically classify news that leads to greater stock market volatility
The mixture-of-experts neural network	Credit scoring[183]	the ability to partition the input space
Neural network based on the backpropagation learning algorithm	Credit scoring[184]	Save computing time
Multi-layer perception	Credit scoring[185]	It can optimize the structure of the dataset, which improve the classification accuracy
Multi-stage hybrid model	Credit scoring[187]	Combined feature selection and classifier selection for a superior binary classification prediction performance
Automatic credit scoring strategy	Credit scoring[188]	It balances automation and accuracy
Ensemble model	Credit scoring[189]	It is more robust compared to individual classifiers
Federated Learning	Credit scoring [190]	It can ensure the accuracy of classification without data leakage

F. FINANCIAL INVESTMENT

1) PORTFOLIO OPTIMIZATION

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3333389

The mean-variance model's proposal laid the groundwork for portfolio theory research and practical application. Despite the traditional mean-variance model's great success, the optimal solution derived from its practical application frequently deviates from the actual value. In general, the Markowitz mean-variance model has three shortcomings in its application. The first is the efficiency of the solution and the difficulty of the model solution due to the cardinality constraint. The second issue is applicability, and because the model's premise assumptions are overly ideal, its applicability in a realistic setting will be limited. The third factor is stability, which means that parameter sensitivity and estimation errors cannot be completely avoided, resulting in an unstable solution set for the model. As AI technology advances, a plethora of advanced techniques are being used to improve the performance of Markowitz models.

For problem one, we can use algorithms to achieve the goal of simplifying computation. Early studies generally used Monte Carlo, branch-and-bound, and iterative methods to solve the optimization problem. Recent studies use some heuristic algorithms to find the approximate optimal solution of the model, such as rotation algorithms for inequality sets, genetic algorithms, and other methods. In [191], the authors combined the features of the firefly algorithm (FA) and genetic algorithm (GA) and proposed a hybrid solution method called FA-GA, which can significantly improve the efficiency of solving mean-variance models. Multi-objective particle swarm optimization can also solve this critical problem effectively [192]. In [193], an iterative method based on backward recursive programming is proposed to obtain the optimal solution for multi-period portfolio optimization. In their study, one iteration takes only a few seconds, which improves the computational speed greatly.

For problem second, many researchers have improved the Markowitz model for portfolio optimization by continuously relaxing the assumptions to make the model more realistic. The traditional model does not consider the transaction cost, and the adjustment of asset structure will cause an unnecessary increase in transaction cost. Therefore, many researchers considered transaction costs in improving the model and optimizing the objective function [194]-[196]. At the same time, the consideration of transaction costs leads to further computational difficulties, and fuzzy logic and beetle antennae search algorithm (BAS) are used for the solution of the problem in their study. In addition to transaction costs, the variance has been used in traditional models to measure the risk of assets. Variance is meaningful as a measure of risk to some extent, but it is unfair to measure risk using variance due to investors' biased perceptions of price downside and price upside. Mean-VAR and Mean-CVAR are proposed to solve this problem. However, the Mean-VAR model often leads to a non-convex NP-hard problem, which makes the solution more difficult. In [197] a hybrid multi-objective optimization algorithm is proposed to solve the optimal solution of the Mean-VAR model and validated the effectiveness of the algorithm with the S & P 100 index and S & P 500 index datasets. In [198], the Mean-CVAR model is applied to define a portfolio optimization problem with basis constraints. Considering the computational efficiency issue, they propose a high-performance algorithm named the bilevel cutting-plane algorithm.

For problem three, many studies have attempted to build more stable Mean-Variance portfolio selection models using robust techniques to reduce estimation errors, such as Autoregressive, Volatility clustering, and Skewness techniques [199]. In [200], the authors used performance-based regularization to bind the estimation error and improve the stability of Mean-Variance models. A Bayesian posterior approach is applied to predict the future distribution of asset returns in [201], where the parameters of the posterior predictive distribution are a function of the observed data values, thus avoiding reliance on the estimated unknown quantity and reducing the estimation error. Table IX presents the methods of these related works and their advantages.

2) AUTOMATED TRADING STRATEGIES

Making the right investment decisions is a difficult task for investors, and the 2008 global financial crisis prompted the use of automated technology to assist investors in making the right investment decisions in the security secondary market. Most traditional trading models are based on mathematics and financial theory, and they rely on experts to understand and apply them, which can result in decisions that are overly subjective and susceptible to investor emotions [202]. Researchers have tried to combine machine learning techniques with human expert knowledge (supervised algorithms) to build more sophisticated models to support traders' decision-making processes [158]. For instance, support vector machine (SVM) techniques are used to classify stock market assets, and genetic algorithms (GA) are used to select the best-traded assets [203]. In [204], the authors analyzed the role of simple machine learning models in achieving profitable trades through a series of forex market trading simulations. However, equity markets are increasingly complex and dynamic, resulting in automated trading systems that are often less able to capture market changes and thus make incorrect decisions. Deep learning and reinforcement learning provide a solution to this problem. Adaptive reinforcement learning has been widely used to develop automated trading strategies and was first used for investment decision-making in the foreign exchange market [205]. In [206], long and short-term memory networks (LSTM) are used for automated stock trading to address the drawbacks of previous methods that relied too much on training datasets. An adaptive stock trading strategy based on deep reinforcement learning methods is proposed in [207] and used Gated Recurrent Units (GRU) to extract time series properties, which can be better applied to out-of-sample data.

The explosive growth of high-frequency trading in recent years has placed a higher demand on automated trading systems, where the speed of data collection, processing, and delivery is a key factor in generating profits in the context of high-frequency trading [208]. As a result, researchers have also attempted to develop automated trading systems in the context of high-frequency trading. In [209], a sparse coding-inspired optimal trading (SCOT) system for real-time high-frequency financial signal representation and trading is conducted, which can capture real-time market signals. And deep learning and reinforcement learning algorithms are used to predict short-term trends in the forex market and apply a grid trading engine to enable high-frequency trading in [210]. Table IX presents the methods of these related works and their advantages.

TABLE IX AI METHODS APPLIED IN FINANCIAL INVESTMENT AND THEIR ADVANTAGES

Method	Application	Advantages
Integrated method Firefly algorithm and genetic algorithm	Portfolio optimization[191]	The balance between exploration and utilization is achieved.
Multi-objective particle swarm algorithm	Portfolio optimization[192]	Compared with other multi-objective evolutionary algorithms, the method can effectively solve the problem
Backward recursive algorithm	Portfolio optimization[193]	Guaranteed convergence
Multi-objective optimization algorithm	Portfolio optimization[197]	This method provides greater flexibility and does not depend on specific distributional assumptions
Bilevel cutting-plane algorithm	Portfolio optimization[198]	Suitable for optimization problems with a large number of investment assets in the portfolio
Performance-based regularization	Portfolio optimization[200]	It can effectively capture the complex dynamics of the market and improve forecasting capabilities by extracting textual information
Bayesian posterior approach	Portfolio optimization[201]	This method directly accounts for the uncertainty of the parameters, thus providing a more robust portfolio selection.
Support vector machines and real-coded genetic algorithm	Portfolio optimization[203]	The method can flexibly handle complex portfolio optimization problems and is adaptive
Adaptive reinforcement learning	Portfolio optimization[205]	Ability to adapt to market changes
Long and short-term memory networks	Portfolio optimization[206],[210]	This method is particularly suitable for dealing with time series data, such as stock prices
Deep reinforcement learning methods	Portfolio optimization[207]	This method is particularly suitable for dealing with time series data, and adapting to changing market conditions
Sparse coding-inspired optimal trading system	Portfolio optimization[209]	This method can effectively compress high-dimensional data and is suitable for high-frequency trading

G. Fraud and financial crime detection

1) CORPORATE FRAUD DETECTION

In the age of AI, Corporate fraud detection has more possibilities. Initially, financial statement fraud detection relied on auditors' subjective judgments of the relevant financial ratios, but previous research indicates that auditors failed to detect significant fraud [211]. Furthermore, manual detection is thought to be time-consuming, costly, and inaccurate. AI technology methods, on average, are more consistent and accurate than human judgment. Early research can use logistic regression models to investigate the relationship between the likelihood of financial statement fraud and specific fraud factors such as board structure, insider trading, and the presence of an independent audit committee [212]. Spathis (2002) used two financial ratios and predicted the likelihood of falsification using a logistic regression model with a detection accuracy of over 84% in [213]. Due to the financial ratios provided by experts, a common drawback in the study is the lack of data availability, which leads to lower accuracy. As a result, data mining techniques are widely used in extracting financial statement falsification features, with decision trees being one of the most popular feature selection methods [214]. In [215], the authors applied Z-scores as the first-level allocator of a decision tree model, which significantly reduced the type one error of classifying fraudulent firms as non-fraudulent. In their study, the bayesian belief network was also used for prediction and was found to have the best overall performance. In a recent study, the researchers also argued that Bayesian Belief Network (BBN) performs best in corporate fraud detection in [88]. In addition, neural networks are another popular classifier that has been applied successfully to detect financial statement fraud. In [216], it used an integrated fuzzy neural network (FNN) for fraud detection, and the fuzzy neural network developed in this study outperformed most artificial neural networks (ANNs) reported in previous studies. However, the feature factors used in their study are mostly financial ratios, ignoring some of the implicit discussion of some issues in the financial statements' textual information. Craja et al. extracted textual information features from the management discussion and analysis (MD&A) section of the annual report through a hierarchical attention network (HAN) to complement the information provided by the financial ratios, thus improving the predictive accuracy of the financial statement fraud model in [217]. In addition, financial ratios are calculated from raw data, which may lead to distorted information. To overcome this problem, an integrated learning model based on logistic regression is used directly to the raw data set rather than to the financial ratios, achieving an accuracy rate better than previous studies in [218]. Table X presents the methods of these related works and their advantages.

2) ANTI-MONEY LAUNDERING

Since there are several ways to launder money and a wide range of tools at the disposal of criminals, money laundering may occur anywhere. Therefore, combating fraud and financial crime has always required effective anti-money laundering solutions. Most financial institutions have relied on rule-based systems to address anti-money laundering in recent years that eliminate as many questionable transactions as possible based on static rules established in advance by professionals [219]. Fuzzy Computing has been used to create natural language-based rules [220], which offers us an alternative to traditional knowledge-driven reasoning systems and overcomes their major shortcomings in terms of the rigidity of rule structure. Static rules, however, are difficult to adapt to changing methods of money laundering. In [221], the authors use intelligent agents, which can learn from their surroundings and adjust to changes in their surroundings, to prevent and regulate money laundering. This kind of strategy, nonetheless, has some processing of data constraints and issues with practical use. Given this fact, using machine learning algorithms to find potential evidence of aberrant activity from current data is the most economical choice.

Scholars have made numerous attempts at this topic. Tang and Yin (2005) use support vector machines and integrate them with learning theory to assist financial firms in identifying anomalous client behavior in [222]. In [223], A radial basis function neural network model based on the APC-III clustering algorithm and recursive least squares outperforms support vector machines in anti-money laundering (AML). In [224], the authors used five commonly used machine learning algorithms to identify money laundering after processing the samples using sampling techniques, and discovered that neural network performs the best among these models. In [225], the authors created a machine learning component that connects with existing watch-list filtering systems and compared the performance of SVMs, Decision Trees, and Naive Bayes. In [226], the authors employed four unsupervised algorithms to enhance the clustering process, which was successful in lowering the percentage of false positives and raising accuracy. They also used innovative trading anomaly indicators as new proof of money laundering. Similarly, deep learning is widely used in money laundering detection. For example, Weber et al. used some advanced techniques, including logistic regression, random forest, multilayer perceptron, and graph convolutional networks, to predict illegal activities such as money laundering in [227]. They validated the results on an elliptical dataset and discovered that the graph convolutional network has a significant advantage. Cheng et al. created a method based on a group-aware graph neural network approach (GAGNN), which allows the model to learn from the user transaction graph and increase detection capability in [228]. In addition, federated learning is also applied for abnormal business detection in [229], providing safe and privacy-aware learning and reasoning for financial crime detection. Table X presents the methods of these related works and their advantages.

TABLE X AI METHODS APPLIED IN FRAUD AND FINANCIAL CRIME DETECTION AND THEIR ADVANTAGES

Method	Application	Advantages
Logistic regression	Corporate fraud detection[213]	Simple and easy to understand
Bayesian belief network	Corporate fraud detection[88],[215]	Flexible application to datasets of different types and sizes
Fuzzy neural network	Corporate fraud detection[216]	Outperformed most artificial neural networks (ANNs)
Hierarchical attention network	Corporate fraud detection[217]	This approach automatically learns fraud-related features and scales well to large datasets
RUSBoost	Corporate fraud detection[218]	Especially suitable for dealing with the imbalance between positive and negative samples
Fuzzy computing	Anti-money laundering[220]	It can effectively capture the complex dynamics of the market and improve forecasting capabilities by extracting textual information
Intelligent agents	Anti-money laundering[221]	It can make the system more flexible to adapt to changing regulatory environment
Support vector machines	Anti-money laundering[222]	Good generalization ability
Radial basis function neural network	Anti-money laundering[223]	Outperform support vector machines
Artificial neural network	Anti-money laundering[224]	Outperform decision trees and random forest
Machine learning component that connects with	Anti-money laundering[225]	Outperform support vector machines, decision Trees, and naïve bayes

existing watch-list filtering systems		
Strict competitive learning, self-organizing-map, C-means and neural gas	Anti-money laundering[226]	Identify money laundering in complex financial data.
Graph convolutional network	Anti-money laundering[227]	Outperform logistic regression, random forest, and multilayer perceptron
Group-aware graph neural network	Anti-money laundering[228]	This method can capture complex relationships and automatically extract and learn important features in data
Federated learning	Anti-money laundering[229]	providing safe and privacy-aware learning and reasoning for financial crime detection

H. SUMMARY

We have reviewed recent AI applications in accounting and finance. We identified areas that have been well-studied and also areas that require further research efforts. The well-studied areas include bankruptcy prediction, fraud detection, market risk management, credit risk management, and portfolio optimization. In these tasks, advanced AI technologies, including machine learning and deep-learning models, have been extensively used. On the other hand, areas such as tune analysis, investment project management, corporate finance decision-making, or Automated trading strategies have not attracted an equal level of attention.

In terms of models, although important research challenges still exist in the more traditional statistical models, more high-value open problems involve the more advanced machine-learning models. The accounting and finance domain has much to gain from leveraging recent breakthroughs in AI technologies, particularly deep learning, applied to other domains. These include the new uncertainty estimation methods in computer vision, and robust algorithms for small, noisy, or nonstationary data.

V. CONCLUSION AND FUTURE WORK

We have reviewed the application of AI to solve some classic accounting and finance problems. First, this study examines some classic problems in accounting and finance, as well as some challenges encountered in the research of these classic problems, such as high subjectivity in decision-making, low accuracy of traditional methods, inability to handle large data sets and too many traditional theoretical assumptions that do not match reality. Then we discuss some opportunities presented by AI as technology advances. AI can replace some repetitive tasks, mine the information implied in the data better, and provide more accurate and objective solutions. At the same time, this paper provides an overview of typical problems that AI can solve, including classification, clustering and regression (prediction). Finally, the paper summarizes the solutions provided by AI based on the classical problems discussed in Section 2 and analyses which challenges exist that will necessitate the use of AI techniques to solve them in the future.

These challenges include data security and privacy, interpretability, and data dependency, etc.

- Data security and privacy: Data security and privacy issues are particularly sensitive and important in the accounting and financial fields. Although existing technologies have solved some problems, they still face bottlenecks in efficiency and scalability.
- Interpretability: Many sophisticated models, such as kernel techniques and neural networks, operate in high-dimensional domains and employ nonlinear functions. The model's decision boundary is difficult for humans to understand intuitively due to its nonlinearity and high dimensionality.
- Data dependency: The application of artificial intelligence models in the accounting and financial fields often requires a large amount of data for training. Therefore, problems that may exist in the data, such as data imbalance and data bias, will affect the performance of the model.

Based on the challenges mentioned in the previous section, we can highlight some key research directions that can highly benefit artificial intelligence applications within the accounting and finance domain.

Future studies could focus on increasing the robustness of the solutions. Because of the enormous influence of data flaws and population drift, robustness is a critical factor for AI systems. Small samples, for example, can be addressed in deep learning regularization, data augmentation, early stopping, and increasing the group in cross-validation, which improves the robustness of the model. Another example is that when solving classification issues like financial fraud, dealing with unbalanced data samples is essential for boosting model robustness.

Furthermore, beyond the issue of robustness, the interpretability of AI models and the implementation of their results in practice are critical. For example, using machine learning approaches in risk management can increase credit rating accuracy. However, it has the potential to automatically discriminate against clients from particularly vulnerable groups and fails to sufficiently explain credit denials, resulting in a breakdown in financial inclusion. Future research can focus on the application of explainable AI in this field [230].

Finally, financial and fiscal data have a certain level of privacy; for example, financial organizations have trouble sharing their business data, which might make it difficult for some methods to find adequate training datasets, lowering model accuracy. Future studies could potentially look into the use of federated learning [231] in accounting and finance, which does not involve the exchange of underlying financial data and so avoids data leaking.

REFERENCES

[1] S. H. Penman, "Financial forecasting, risk and valuation: Accounting for the future: Financial forecasting, risk and valuation," Abacus, vol. 46, no. 2, pp. 211–228, 2010.

[2] D. Appelbaum, A. Kogan, M. Vasarhelyi, and Z. Yan, "Impact of business analytics and enterprise systems on managerial accounting," Int. J. Acc. Inf. Syst., vol. 25, pp. 29–44, 2017.

[3] S. Askary, N. Abu-Ghazaleh, and Y. A. Tahat, "Artificial intelligence and reliability of accounting information," in Lecture Notes in Computer Science, Cham: Springer International Publishing, 2018, pp. 315–324.

[4] Y.-J. Chen, Y.-M. Chen, and C. L. Lu, "Enhancement of stock market forecasting using an improved fundamental analysis-based approach," Soft Comput., vol. 21, no. 13, pp. 3735–3757, 2017.

[5] P. Bond, A. Edmans, and I. Goldstein, "The real effects of financial markets," Annu. Rev. Fin. Econ., vol. 4, no. 1, pp. 339–360, 2012.

[6] H. Issa, T. Sun, and M. A. Vasarhelyi, "Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation," J. Emerg. Technol. Acc., vol. 13, no. 2, pp. 1–20, 2016.

[7] M. López de Prado, "Beyond econometrics: A roadmap towards financial machine learning," SSRN Electron. J., 2019.

[8] P. Goulet Coulombe, M. Leroux, D. Stevanovic, and S. Surprenant, "How is machine learning useful for macroeconomic forecasting?," J. Appl. Econ. (Chichester Engl.), vol. 37, no. 5, pp. 920–964, 2022.

[9] Y. Hilpisch, Artificial Intelligence in Finance: A Python-Based Guide. Sebastopol, CA: O'Reilly Media, 2020.

[10] D. Mhlanga, "Industry 4.0 in finance: The impact of artificial intelligence (AI) on digital financial inclusion," Int. J. Fin. Stud., vol. 8, no. 3, p. 45, 2020.

[11] F. Königstorfer and S. Thalmann, "Applications of Artificial Intelligence in commercial banks – A research agenda for behavioral finance," J. Behav. Exp. Finance, vol. 27, no. 100352, p. 100352, 2020.

[12] S. Garg, S. Sinha, A. K. Kar, and M. Mani, "A review of machine learning applications in human resource management," Int. j. product. perform. manag., vol. 71, no. 5, pp. 1590–1610, 2022.

[13] F. Yi et al., "Dual model medical invoices recognition," Sensors (Basel), vol. 19, no. 20, p. 4370, 2019.

[14] J. Qiu, "Analysis of human interactive accounting management information systems based on Artificial Intelligence," J. Glob. Inf. Manag., vol. 30, no. 7, pp. 1–13, 2021.

[15] S. Karaoglu and U. Arpaci, "A deep learning approach for optimization of systematic signal detection in financial trading systems with big data," Int. j. Intellnoi. syst. appl. eng., vol. Special Issue, no. Special Issue, pp. 31–36, 2017.

[16] A. H. Huang, A. Y. Zang, and R. Zheng, "Evidence on the information content of text in analyst reports," account. rev., vol. 89, no. 6, pp. 2151–2180, 2014.

[17] M. Hagenau, M. Liebmann, and D. Neumann, "Automated news reading: Stock price prediction based on financial news using context-capturing features," Decis. Support Syst., vol. 55, no. 3, pp. 685–697, 2013

[18] T.-M. Chang, M.-F. Hsu, and S.-J. Lin, "Integrated news mining technique and AI-based mechanism for corporate performance forecasting," Inf. Sci. (Ny), vol. 424, pp. 273–286, 2018.

[19] R. Santhosh and M. Mohanapriya, "Generalized fuzzy logic based performance prediction in data mining," Mater. Today, vol. 45, pp. 1770–1774, 2021.

[20] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," Comput. Sci. Rev., vol. 40, no. 100402, p. 100402, 2021.

[21] P.-F. Pai, M.-F. Hsu, and M.-C. Wang, "A support vector machine-based model for detecting top management fraud," Knowl. Based Syst., vol. 24, no. 2, pp. 314–321, 2011.

[22] M. Krambia-Kapardis, C. Christodoulou, and M. Agathocleous, "Neural networks: the panacea in fraud detection?," Manag. Audit. J., vol. 25, no. 7, pp. 659–678, 2010.

[23] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Literature review: Machine learning techniques applied to financial market prediction," Expert Syst. Appl., vol. 124, pp. 226–251, 2019.

[24] P. Pławiak, M. Abdar, and U. Rajendra Acharya, "Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring," Appl. Soft Comput., vol. 84, no. 105740, p. 105740, 2019.

[25] J. Horak, J. Vrbka, and P. Suler, "Support Vector Machine methods and artificial neural networks used for the development of bankruptcy prediction models and their comparison," J. Risk Fin. Manag., vol. 13, no. 3, p. 60, 2020.

[26] M. S. Park, H. Son, C. Hyun, and H. J. Hwang, "Explainability of machine learning models for bankruptcy prediction," IEEE Access, vol. 9, pp. 124887–124899, 2021.

[27] C. B. Kalayci, O. Polat, and M. A. Akbay, "An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization," Swarm Evol. Comput., vol. 54, no. 100662, p. 100662, 2020.

[28] A. H. Khan et al., "Optimal portfolio management for engineering problems using nonconvex cardinality constraint: A computing perspective," IEEE Access, vol. 8, pp. 57437–57450, 2020.

[29] M. Rahiminezhad Galankashi, F. Mokhatab Rafiei, and M. Ghezelbash, "Portfolio selection: a fuzzy-ANP approach," Financ. innov., vol. 6, no. 1, 2020.

[30] X. Cao et al., "A novel recurrent neural network based online portfolio analysis for high frequency trading," SSRN Electron. J., 2022.

[31] O. Ertenlice and C. B. Kalayci, "A survey of swarm intelligence for portfolio optimization: Algorithms and applications," Swarm Evol. Comput., vol. 39, pp. 36–52, 2018.

[32] M. Leo, S. Sharma, and K. Maddulety, "Machine learning in banking risk management: A literature review," Risks, vol. 7, no. 1, p. 29, 2019.

[33] S. Collot and T. Hemauer, "A literature review of new methods in empirical asset pricing: omitted-variable and errors-in-variable bias," Fin. Mark. Portf. Manag., vol. 35, no. 1, pp. 77–100, 2021.

[34] F. A. Amani and A. M. Fadlalla, "Data mining applications in accounting: A review of the literature and organizing framework," Int. J. Acc. Inf. Syst., vol. 24, pp. 32–58, 2017.

[35] Sanchez-Roger, Oliver-Alfonso, and Sanchís-Pedregosa, "Fuzzy logic and its uses in finance: A systematic review exploring its potential to deal with banking crises," Mathematics, vol. 7, no. 11, p. 1091, 2019.

[36] A. M. Ozbayoglu, M. U. Gudelek, and O. B. Sezer, "Deep learning for financial applications : A survey," Appl. Soft Comput., vol. 93, no. 106384, p. 106384, 2020.

[37] A. S. Ahmed and I. Safdar, "Dissecting stock price momentum using financial statement analysis," Accounting Finance, vol. 58, pp. 3–43, 2018.

[38] E. H. Feroz, S. Kim, and R. L. Raab, "Financial statement analysis: A data envelopment analysis approach," J. Oper. Res. Soc., vol. 54, no. 1, pp. 48–58, 2003.

[39] M. Firth, P. L. L. Mo, and R. M. K. Wong, "Financial statement frauds and auditor sanctions: An analysis of enforcement actions in China," J. Bus. Ethics, vol. 62, no. 4, pp. 367–381, 2005.

[40] L. L. Lisic, S. (dino) Silveri, Y. Song, and K. Wang, "Accounting fraud, auditing, and the role of government sanctions in China," J. Bus. Res., vol. 68, no. 6, pp. 1186–1195, 2015.

[41] W. H. Beaver, M. F. McNichols, and J.-W. Rhie, "Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy," Rev. Acc. Stud., vol. 10, no. 1, pp. 93–122, 2005.

[42] S. Jones, D. Johnstone, and R. Wilson, "An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes," J. Bank. Financ., vol. 56, pp. 72–85, 2015.

[43] J. Lewellen, "Predicting returns with financial ratios," J. financ. econ., vol. 74, no. 2, pp. 209–235, 2004.

[44] H.-T. Tan, E. Ying Wang, and B. O. Zhou, "When the use of positive language backfires: The joint effect of tone, readability, and investor sophistication on earnings judgments: When the use of positive language backfires," J. Acc. Res., vol. 52, no. 1, pp. 273–302, 2014.

[45] J. Christensen, "Conceptual frameworks of accounting from an information perspective," Acc. Bus. Res., vol. 40, no. 3, pp. 287–299, 2010.
[46] K. J. Reichelt and D. Wang, "National and office-specific measures of auditor industry expertise and effects on audit quality," J. Acc. Res., vol. 48, no. 3, pp. 647–686, 2010.
[47] W. Wiesemann, D. Kuhn, and B. Rustem, "Maximizing the net present value of a project under uncertainty," Eur. J. Oper. Res., vol. 202, no. 2, pp. 356–367, 2010.
[48] M.-Y. Cheng, H.-C. Tsai, and C.-L. Liu, "Artificial intelligence approaches to achieve strategic control over project cash flows," Autom. Constr., vol. 18, no. 4, pp. 386–393, 2009.
[49] A. Marchioni and C. A. Magni, "Investment decisions and sensitivity analysis: NPV-consistency of rates of return," Eur. J. Oper. Res., vol. 268, no. 1, pp. 361–372, 2018
[50] M. Bradley, G. A. Jarrell, and E. H. Kim, "On the existence of an optimal capital structure: Theory and evidence," J. Finance, vol. 39, no. 3, p. 857, 1984.
[51] B. P. Singh and M. Kannadhasan, "Corruption and capital structure in emerging markets: A panel quantile regression approach," J. Behav. Exp. Finance, vol. 28, no. 100417, p. 100417, 2020.
[52] R. Zeitun and M. Goaied, "The nonlinear effect of foreign ownership on capital structure in Japan: A panel threshold analysis," Pac.-basin finance j., vol. 68, no. 101594, p. 101594, 2021.
[53] R. C. Higgins, Analysis for Financial Management, 7th ed. Maidenhead, England: McGraw Hill Higher Education, 2003.
[54] Ref.2.1.18 L. P. Jennergren, "Continuing value in firm valuation by the discounted cash flow model," Eur. J. Oper. Res., vol. 185, no. 3, pp. 1548–1563, 2008.
[55] S.-C. Chang, S.-S. Chen, A. Hsing, and C. W. Huang, "Investment opportunities, free cash flow, and stock valuation effects of secured debt offerings," Rev. Quant. Fin. Acc., vol. 28, no. 2, pp. 123–145, 2007.
[56] K. S. Lorek, "Trends in statistically based quarterly cash-flow prediction models," Account. forum, vol. 38, no. 2, pp. 145–151, 2014.
[57] R. S. J. Koijen and M. Yogo, "A demand system approach to asset pricing," J. Polit. Econ., vol. 127, no. 4, pp. 1475–1515, 2019.
[58] M. Brunnermeier et al., "Review article: Perspectives on the future of asset pricing," Rev. Financ. Stud., vol. 34, no. 4, pp. 2126–2160, 2021.
[59] L. Chen, M. Pelger, and J. Zhu, "Deep learning in asset pricing," 2019.
[60] N. Augenblick and E. Lazarus, "Restrictions on asset-price movements under rational expectations: Theory and evidence," SSRN Electron. J., 2018.
[61] R. Efendi, N. Arbaiy, and M. M. Deris, "A new procedure in stock market forecasting based on fuzzy random auto-regression time series model," Inf. Sci. (Ny), vol. 441, pp. 113–132, 2018.
[62] I.-C. Tsai, "The source of global stock market risk: A viewpoint of economic policy uncertainty," Econ. Model., vol. 60, pp. 122–131, 2017.
[63] S. Oreski, D. Oreski, and G. Oreski, "Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment," Expert Syst. Appl., vol. 39, no. 16, pp. 12605–12617, 2012.
[64] D. Yang, P. Chen, F. Shi, and C. Wen, "Internet finance: Its uncertain legal foundations and the role of big data in its development," Emerg. Mark. Fin. Trade, vol. 54, no. 4, pp. 721–732, 2018.
[65] H. Markowitz, "Portfolio Selection," J. Finance, vol. 7, no. 1, p. 77, 1952.
[66] J. A. Cassara, Money laundering and illicit financial flows: Following the money and value trails. Independently Published, 2020.
[67] F. Fauzi, K. Szulczyk, and A. Basyith, "Moving in the right direction to fight financial crime: prevention and detection," J. Financ. Crime, vol. 25, no. 2, pp. 362–368, 2018.
[68] J. Khaksar, M. Salehi, and M. Lari DashtBayaz, "The relationship between auditor characteristics and fraud detection," J. Facil. Manag., vol. 20, no. 1, pp. 79–101, 2022.
[69] Z. Chen, L. D. Van Khoa, E. N. Teoh, A. Nazir, E. K. Karuppiah, and K. S. Lam, "Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review," Knowl. Inf. Syst., vol. 57, no. 2, pp. 245–285, 2018.
[70] W. R. Knechel and J. L. Payne, "Additional evidence on audit report lag," Audit. J. Pract. Theory, vol. 20, no. 1, pp. 137–146, 2001.
[71] P. Filipiak and P. Lipinski, "Dynamic portfolio optimization in ultra-high frequency environment," in Applications of Evolutionary Computation, Cham: Springer International Publishing, 2017, pp. 34–50.
[72] A.-D. Socea, "Managerial decision-making and financial accounting information," Procedia Soc. Behav. Sci., vol. 58, pp. 47–55, 2012.
[73] J. Zhu and S. S. Gao, "Fraudulent financial reporting: Corporate behavior of Chinese listed companies," in Research in Accounting in Emerging Economies, Emerald Group Publishing Limited, 2011, pp. 61–82.
[74] S. Bai and Y. Zhao, "Startup investment decision support: Application of venture capital scorecards using machine learning approaches," Systems, vol. 9, no. 3, p. 55, 2021.
[75] J. Kokina and S. Blanchette, "Early evidence of digital labor in accounting: Innovation with Robotic Process Automation," Int. J. Acc. Inf. Syst., vol. 35, no. 100431, p. 100431, 2019.
[76] T. Ko, H. D. Jeong, and G. Lee, "Natural language processing–driven model to extract contract change reasons and altered work items for advanced retrieval of change orders," J. Constr. Eng. Manag., vol. 147, no. 11, p. 04021147, 2021.
[77] S. Finlay, "Multiple classifier architectures and their application to credit risk assessment," Eur. J. Oper. Res., vol. 210, no. 2, pp. 368–378, 2011.
[78] I. González-Carrasco, J. L. Jiménez-Márquez, J. L. López-Cuadrado, and B. Ruiz-Mezcua, "Automatic detection of relationships between banking operations using machine learning," Inf. Sci. (Ny), vol. 485, pp. 319–346, 2019.
[79] S. Karaoglu and U. Arpaci, "A deep learning approach for optimization of systematic signal detection in financial trading systems with big data," Int. j. intell. syst. appl. eng., vol. Special Issue, no. Special Issue, pp. 31–36, 2017.
[80] W. Bao, J. Yue, and Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory," PLoS One, vol. 12, no. 7, p. e0180944, 2017.
[81] O. M. Yigitbasioglu and O. Velcu, "A review of dashboards in performance management: Implications for design and research," Int. J. Acc. Inf. Syst., vol. 13, no. 1, pp. 41–59, 2012.
[82] M. Xiao, H. Xiaoli, and L. Gaojin, "Research on application of data mining technology in financial decision support system," in 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering, 2010.
[83] L. Zhou, D. Lu, and H. Fujita, "The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches," Knowl. Based Syst., vol. 85, pp. 52–61, 2015.
[84] P. W. Murray, B. Agard, and M. A. Barajas, "Market segmentation through data mining: A method to extract behaviors from a noisy data set," Comput. Ind. Eng., vol. 109, pp. 233–252, 2017.
[85] Barbeito-Caamaño and R. Chalmeta, "Using big data to evaluate corporate social responsibility and sustainable development practices," Corp. Soc. Responsibility Environ. Manage., vol. 27, no. 6, pp. 2831–2848, 2020.
[86] R. Frankel, J. Jennings, and J. Lee, "Disclosure sentiment: Machine learning vs. Dictionary methods," Manage. Sci., vol. 68, no. 7, pp. 5514–5532, 2022.
[87] G. González Rodríguez, J. M. Gonzalez-Cava, and J. A. Méndez Pérez, "An intelligent decision support system for production planning based on machine learning," J. Intell. Manuf., vol. 31, no. 5, pp. 1257–1273, 2020.
[88] P. Hajek and R. Henriques, "Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods," Knowl. Based Syst., vol. 128, pp. 139–152, 2017.
[89] S.-B. Kim, K.-S. Han, H.-C. Rim, and S. H. Myaeng, "Some effective techniques for naive Bayes text classification," IEEE Trans. Knowl. Data Eng., vol. 18, no. 11, pp. 1457–1466, 2006.
[90] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J.-H. Jeng, and J.-G. Hsieh, "Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using Naïve Bayes," Information (Basel), vol. 12, no. 5, p. 204, 2021.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2023.3333389

[91] H.-L. Chen et al., "A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method," Knowl. Based Syst., vol. 24, no. 8, pp. 1348–1359, 2011.
[92] A. Li, J. Wu, and Z. Liu, "Market manipulation detection based on classification methods," Procedia Comput. Sci., vol. 122, pp. 788–795, 2017.
[93] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data Mining techniques for the detection of fraudulent financial statements," Expert Syst. Appl., vol. 32, no. 4, pp. 995–1003, 2007.
[94] H. Qian, B. Wang, M. Yuan, S. Gao, and Y. Song, "Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree," Expert Syst. Appl., vol. 190, no. 116202, p. 116202, 2022.
[95] T. K. Lee, J. H. Cho, D. S. Kwon, and S. Y. Sohn, "Global stock market investment strategies based on financial network indicators using machine learning techniques," Expert Syst. Appl., vol. 117, pp. 228–242, 2019.
[96] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018.
[97] J. Min and Y. Lee, "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters," Expert Syst. Appl., vol. 28, no. 4, pp. 603–614, 2005.
[98] M. P. Kumar and M. K. Rajagopal, "Detecting facial emotions using normalized minimal feature vectors and semi-supervised twin support vector machines classifier," Appl. Intell., vol. 49, no. 12, pp. 4150–4174, 2019.
[99] H. Kim, H. Cho, and D. Ryu, "Corporate bankruptcy prediction using machine learning methodologies with a focus on sequential data," Comput. Econ., vol. 59, no. 3, pp. 1231–1249, 2022.
[100] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," IEEE Trans. Cybern., vol. 49, no. 3, pp. 839–847, 2019.
[101] M. Vernuccio, "Communicating corporate brands through social media: An exploratory study," Int. J. Bus. Commun., vol. 51, no. 3, pp. 211–233, 2014.
[102] C. Grieco, L. Michelini, and G. Iasevoli, "Measuring value creation in social enterprises: A cluster analysis of Social Impact Assessment models," Nonprofit Volunt. Sect. Q., vol. 44, no. 6, pp. 1173–1193, 2015.
[103] W. W. Powell and K. W. Sandholtz, "Amphibious entrepreneurs and the emergence of organizational forms," Strat. Entrep.. J., vol. 6, no. 2, pp. 94–115, 2012.
[104] S. Tolun Tayalı, "A novel backtesting methodology for clustering in mean–variance portfolio optimization," Knowl. Based Syst., vol. 209, no. 106454, p. 106454, 2020.
[105] H. W. Shin and S. Y. Sohn, "Segmentation of stock trading customers according to potential value," Expert Syst. Appl., vol. 27, no. 1, pp. 27–33, 2004.
[106] S. R. Nanda, B. Mahanty, and M. K. Tiwari, "Clustering Indian stock market data for portfolio management," Expert Syst. Appl., vol. 37, no. 12, pp. 8793–8798, 2010
[107] C. Bai, D. Dhavale, and J. Sarkis, "Integrating Fuzzy C-Means and TOPSIS for performance evaluation: An application and comparative analysis," Expert Syst. Appl., vol. 41, no. 9, pp. 4186–4196, 2014.
[108] B. Yeo and D. Grant, "Predicting service industry performance using decision tree analysis," Int. J. Inf. Manage., vol. 38, no. 1, pp. 288–300, 2018.
[109] R. K. Lai, C.-Y. Fan, W.-H. Huang, and P.-C. Chang, "Evolving and clustering fuzzy decision tree for financial time series data forecasting," Expert Syst. Appl., vol. 36, no. 2, pp. 3761–3773, 2009.
[110] Y. Chen and Y. Hao, "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction," Expert Syst. Appl., vol. 80, pp. 340–355, 2017.
[111] K. Cortez, M. del P. Rodríguez-García, and S. Mongrut, "Exchange market liquidity prediction with the K-nearest neighbor approach: Crypto vs. Fiat currencies," Mathematics, vol. 9, no. 1, p. 56, 2020.
[112] W. Jiang, "Applications of deep learning in stock market prediction: Recent progress," Expert Syst. Appl., vol. 184, no. 115537, p. 115537, 2021.
[113] S. Cavalli and M. Amoretti, "CNN-based multivariate data analysis for bitcoin trend prediction," Appl. Soft Comput., vol. 101, no. 107065, p. 107065, 2021.
[114] M.-C. Lee, "Business bankruptcy prediction based on survival analysis approach," Int. j. comput. sci. inf. technol., vol. 6, no. 2, pp. 103–119, 2014.
[115] D. Liang, C.-C. Lu, C.-F. Tsai, and G.-A. Shih, "Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study," Eur. J. Oper. Res., vol. 252, no. 2, pp. 561–572, 2016.
[116] S. Tian, Y. Yu, and H. Guo, "Variable selection and corporate bankruptcy forecasts," J. Bank. Financ., vol. 52, pp. 89–100, 2015.
[117] F. Mai, S. Tian, C. Lee, and L. Ma, "Deep learning models for bankruptcy prediction using textual disclosures," Eur. J. Oper. Res., vol. 274, no. 2, pp. 743–758, 2019.
[118] F. Barboza, H. Kimura, and E. Altman, "Machine learning models and bankruptcy prediction," Expert Syst. Appl., vol. 83, pp. 405–417, 2017.
[119] S.-H. Min, J. Lee, and I. Han, "Hybrid genetic algorithms and support vector machines for bankruptcy prediction," Expert Syst. Appl., vol. 31, no. 3, pp. 652–660, 2006.
[120] T.-J. Hsieh, H.-F. Hsiao, and W.-C. Yeh, "Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm," Neurocomputing, vol. 82, pp. 196–206, 2012.
[121] Y. Lu, N. Zeng, X. Liu, and S. Yi, "A new hybrid algorithm for bankruptcy prediction using switching particle swarm optimization and support vector machines," Discrete Dyn. Nat. Soc., vol. 2015, pp. 1–7, 2015.
[122] S. Jones, D. Johnstone, and R. Wilson, "Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks: An evaluation of alternative statistical frameworks," J. Bus. Finance Account., vol. 44, no. 1–2, pp. 3–34, 2017.
[123] D. Boughaci and A. A. K. Alkhawaldeh, "Appropriate machine learning techniques for credit scoring and bankruptcy prediction in banking and finance: A comparative study," Risk Decis. Anal., vol. 8, no. 1–2, pp. 15–24, 2020.
[124] T. Loughran and B. Mcdonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-ks," J. Finance, vol. 66, no. 1, pp. 35–65, 2011.
[125] T. Kang, D.-H. Park, and I. Han, "Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics," Telemat. inform., vol. 35, no. 2, pp. 370–381, 2018.
[126] P. Hájek, "Combining bag-of-words and sentiment features of annual reports to predict abnormal stock returns," Neural Comput. Appl., vol. 29, no. 7, pp. 343–358, 2018.
[127] M. Azimi and A. Agrawal, "Is positive sentiment in corporate annual reports informative? Evidence from deep learning," rev. asset pricing stud., vol. 11, no. 4, pp. 762–805, 2021.
[128] W. van Atteveldt, M. A. C. G. van der Velden, and M. Boukes, "The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms," Commun. Methods Meas., vol. 15, no. 2, pp. 121–140, 2021.
[129] Y. Bengio, Deep Learning. London, England: MIT Press, 2016.
[130] S. Zhao, K. Xu, Z. Wang, C. Liang, W. Lu, and B. Chen, "Financial distress prediction by combining sentiment tone features," Econ. Model., vol. 106, no. 105709, p. 105709, 2022.
[131] X. Huang, "Optimal project selection with random fuzzy parameters," Int. J. Prod. Econ., vol. 106, no. 2, pp. 513–522, 2007.
[132] S.-H. Liao and S.-H. Ho, "Investment project valuation based on a fuzzy binomial approach," Inf. Sci. (Ny), vol. 180, no. 11, pp. 2124–2133, 2010.
[133] B. Vahdani, S. M. Mousavi, H. Hashemi, M. Mousakhani, and S. Ebrahimnejad, "A new hybrid model based on least squares support vector machine for project selection problem in construction industry," Arab. J. Sci. Eng., vol. 39, no. 5, pp. 4301–4314, 2014.
[134] Y. Li, L. Yang, B. Yang, N. Wang, and T. Wu, "Application of interpretable machine learning models for the intelligent decision," Neurocomputing, vol. 333, pp. 273–283, 2019.
[135] A. RezaHoseini, S. F. Ghannadpour, and M. Hemmati, "A comprehensive mathematical model for resource-constrained multi-objective project portfolio selection and scheduling considering

sustainability and projects splitting," J. Clean. Prod., vol. 269, no. 122073, p. 122073, 2020.
[136] S. Meysam Mousavi, R. Tavakkoli-Moghaddam, B. Vahdani, H. Hashemi, and M. J. Sanjari, "A new support vector model-based imperialist competitive algorithm for time estimation in new product development projects," Robot. Comput. Integr. Manuf., vol. 29, no. 1, pp. 157–168, 2013.
[137] A. Taskin and A. F. Guneri, "Economic analysis of risky projects by ANNs," Appl. Math. Comput., vol. 175, no. 1, pp. 171–181, 2006.
[138] J. Xu et al., "Assessing China's investment risk of the maritime silk road: A model based on multiple machine learning methods," Energies, vol. 15, no. 16, p. 5780, 2022.
[139] H. A. Abdou, A. Kuzmic, J. Pointon, and R. J. Lister, "Determinants of capital structure in the UK retail industry: A comparison of multiple regression and generalized regression neural network: Determinants of capital structure in UK retail industry," Intell. Syst. Account. Finance Manag., vol. 19, no. 3, pp. 151–169, 2012.
[140] H.-T. Pao, "A comparison of neural network and multiple regression analysis in modeling capital structure," Expert Syst. Appl., vol. 35, no. 3, pp. 720–727, 2008
[141] H. I. Hussain, F. Kamarudin, H. M. T. Thaker, and M. A. Salem, "Artificial neural network to model managerial timing decision: Non-linear evidence of deviation from target leverage," Int. J. Comput. Intell. Syst., vol. 12, no. 2, p. 1282, 2019.
[142] B. Lin and R. Bai, "Machine learning approaches for explaining determinants of the debt financing in heavy-polluting enterprises," Fin. Res. Lett., vol. 44, no. 102094, p. 102094, 2022.
[143] X. Huang, X. Liu, and Y. Ren, "Enterprise credit risk evaluation based on neural network algorithm," Cogn. Syst. Res., vol. 52, pp. 317–324, 2018.
[144] B. Liu, "Influence of enterprise cloud computing database application and enterprise debt financing based on SVM algorithm: internal control quality analysis," J. Ambient Intell. Humaniz. Comput., 2021.
[145] H. Yue, H. Liao, D. Li, and L. Chen, "Enterprise Financial Risk Management using information fusion technology and Big Data Mining," Wirel. Commun. Mob. Comput., vol. 2021, pp. 1–13, 2021.
[146] D.-N. Wang, L. Li, and D. Zhao, "Corporate finance risk prediction based on LightGBM," Inf. Sci. (Ny), vol. 602, pp. 259–268, 2022.
[147] H. Weytjens, E. Lohmann, and M. Kleinsteuber, "Cash flow prediction: MLP and LSTM compared to ARIMA and Prophet," Electron. Commer. Res., vol. 21, no. 2, pp. 371–391, 2021.
[148] W. Tchoudi and G. Sergeenko, "Forecasting financial valuation comparing ARIMA and prophet," AfricaArXiv, 2022.
[149] L. Hongjiu, R. Rieg, and H. Yanrong, "Performance comparison of artificial intelligence methods for predicting cash flow," Neural Netw. World, vol. 22, no. 6, pp. 549–564, 2012.
[150] M.-Y. Cheng, N.-D. Hoang, and Y.-W. Wu, "Cash flow prediction for construction project using a novel adaptive time-dependent Least Squares Support Vector Machine inference model," J. Civ. Eng. Manag., vol. 21, no. 6, pp. 679–688, 2015
[151] J.-S. Wang and C.-X. Ning, "ANFIS based time series prediction method of bank cash flow optimized by adaptive population activity PSO algorithm," Information (Basel), vol. 6, no. 3, pp. 300–313, 2015. Aras
[152] L. Zhu, M. Yan, and L. Bai, "Prediction of enterprise free cash flow based on a Backpropagation neural network model of the improved genetic algorithm," Information (Basel), vol. 13, no. 4, p. 172, 2022.
[153] E. Giovanis, "Application of ARCH-GARCH models and feed-forward neural networks with Bayesian regularization in capital asset pricing model: The case of two stocks in Athens exchange stock market," SSRN Electron. J., 2009.
[154] S. Gu, B. Kelly, and D. Xiu, "Empirical asset pricing via machine learning," Rev. Financ. Stud., vol. 33, no. 5, pp. 2223–2273, 2020.
[155] M. Ma, L. Zheng, and J. Yang, "A novel improved trigonometric neural network algorithm for solving price-dividend functions of continuous time one-dimensional asset-pricing models," Neurocomputing, vol. 435, pp. 151–161, 2021.
[156] H. T. Anis and R. H. Kwon, "A sparse regression and neural network approach for financial factor modeling," Appl. Soft Comput., vol. 113, no. 107983, p. 107983, 2021.
[157] Y.-S. Chen, C.-H. Cheng, and W.-L. Tsai, "Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting," Appl. Intell., vol. 41, no. 2, pp. 327–347, 2014.
[158] R. C. 211, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega, and A. L. I. Oliveira, "Computational intelligence and financial markets: A survey and future directions," Expert Syst. Appl., vol. 55, pp. 194–211, 2016.
[159] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," Neural Comput. Appl., vol. 32, no. 6, pp. 1609–1628, 2020.
[160] E. Junqué de Fortuny, T. De Smedt, D. Martens, and W. Daelemans, "Evaluating and understanding text-based stock price prediction models," Inf. Process. Manag., vol. 50, no. 2, pp. 426–441, 2014.
[161] C.-J. Lu, "Integrating independent component analysis-based denoising scheme with neural network for stock price prediction," Expert Syst. Appl., vol. 37, no. 10, pp. 7056–7064, 2010.
[162] H. Ince and T. B. Trafalis, "Short term forecasting with support vector machines and application to stock price prediction," Int. J. Gen. Syst., vol. 37, no. 6, pp. 677–687, 2008.
[163] M. Göçken, M. Özçalıcı, A. Boru, and A. T. Dosdoğru, "Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction," Expert Syst. Appl., vol. 44, pp. 320–331, 2016.
[164] A. Bick, "On the consistency of the black-Scholes model with a general equilibrium framework," J. Fin. Quant. Anal., vol. 22, no. 3, p. 259, 1987.
[165] A. P. Carverhill and T. H. F. Cheuk, "Alternative neural network approach for option pricing and hedging," SSRN Electron. J., 2003.
[166] Y.-H. Wang, "Nonlinear neural network forecasting model for stock index option price: Hybrid GJR–GARCH approach," Expert Syst. Appl., vol. 36, no. 1, pp. 564–570, 2009.
[167] S. P. Das and S. Padhy, "A new hybrid parametric and machine learning model with homogeneity hint for European-style index option pricing," Neural Comput. Appl., vol. 28, no. 12, pp. 4061–4077, 2017.
[168] H. Jang and J. Lee, "Generative Bayesian neural network model for risk-neutral pricing of American index options," Quant. Finance, vol. 19, no. 4, pp. 1–17, 2018.
[169] C.-F. Ivașcu, "Option pricing using Machine Learning," Expert Syst. Appl., vol. 163, no. 113799, p. 113799, 2021.
[170] L. Gan, H. Wang, and Z. Yang, "Machine learning solutions to challenges in finance: An application to the pricing of financial products," Technol. Forecast. Soc. Change, vol. 153, no. 119928, p. 119928, 2020.
[171] J. H. Jang, J. Yoon, J. Kim, J. Gu, and H. Y. Kim, "Deep Option: A novel option pricing framework based on deep learning with fused distilled data from multiple parametric methods," Inf. Fusion, vol. 70, pp. 43–59, 2021.
[172] A. Mashrur, W. Luo, N. A. Zaidi, and A. Robles-Kelly, "Machine learning for financial risk management: A survey," IEEE Access, vol. 8, pp. 203203–203223, 2020.
[173] V. V. Gavrishchaka and S. Banerjee, "Support vector machine as an efficient framework for stock market volatility forecasting," Comput. Manag. Sci., vol. 3, no. 2, pp. 147–160, 2006.
[174] C. Luong and N. Dokuchaev, "Forecasting of realised volatility with the random forests algorithm," J. Risk Fin. Manag., vol. 11, no. 4, p. 61, 2018.
[175] S. Aras, "On improving GARCH volatility forecasts for Bitcoin via a meta-learning approach," Knowl. Based Syst., vol. 230, no. 107393, p. 107393, 2021.
[176] W. Kristjanpoller, A. Fadic, and M. C. Minutolo, "Volatility forecast using hybrid Neural Network models," Expert Syst. Appl., vol. 41, no. 5, pp. 2437–2442, 2014.
[177] R. Luo, W. Zhang, X. Xu, and J. Wang, "A neural stochastic volatility model," Proc. Conf. AAAI Artif. Intell., vol. 32, no. 1, 2018.
[178] Y. Liu, "Novel volatility forecasting using deep learning–Long Short Term Memory Recurrent Neural Networks," Expert Syst. Appl., vol. 132, pp. 99–109, 2019.
[179] S. S. Groth and J. Muntermann, "An intraday market risk management approach based on textual analysis," Decis. Support Syst., vol. 50, no. 4, pp. 680–691, 2011.
[180] W. Antweiler and M. Z. Frank, "Is all that talk just noise? The information content of internet stock message boards," J. Finance, vol. 59, no. 3, pp. 1259–1294, 2004.

[181] A. Manela and A. Moreira, “News implied volatility and disaster concerns,” J. financ. econ., vol. 123, no. 1, pp. 137–162, 2017.

[182] P. S. M. Nizer and J. C. Nievola, “Predicting published news effect in the Brazilian stock market,” Expert Syst. Appl., vol. 39, no. 12, pp. 10674–10680, 2012.

[183] D. West, “Neural network credit scoring models,” Comput. Oper. Res., vol. 27, no. 11–12, pp. 1131–1152, 2000.

[184] A. Khashman, “A neural network model for credit risk evaluation,” Int. J. Neural Syst., vol. 19, no. 4, pp. 285–294, 2009.

[185] Z. Zhao, S. Xu, B. H. Kang, M. M. J. Kabir, Y. Liu, and R. Wasinger, “Investigation and improvement of multi-layer perceptron neural networks for credit scoring,” Expert Syst. Appl., vol. 42, no. 7, pp. 3508–3516, 2015.

[186] C.-L. Huang, M.-C. Chen, and C.-J. Wang, “Credit scoring with a data mining approach based on support vector machines,” Expert Syst. Appl., vol. 33, no. 4, pp. 847–856, 2007.

[187] W. Zhang, H. He, and S. Zhang, “A novel multi-stage hybrid model with enhanced multi-population niche genetic algorithm: An application in credit scoring,” Expert Syst. Appl., vol. 121, pp. 221–232, 2019.

[188] F. Yang, Y. Qiao, C. Huang, S. Wang, and X. Wang, “An Automatic Credit Scoring Strategy (ACSS) using memetic evolutionary algorithm and neural architecture search,” Appl. Soft Comput., vol. 113, no. 107871, p. 107871, 2021

[189] A. Chopra and P. Bhilare, “Application of ensemble models in credit scoring models,” Bus. Perspect. Res., vol. 6, no. 2, pp. 129–141, 2018

[190] Z. Wang, L. Wang, J. Yao, Y. Chen, and H. Jiang, “A federated learning approach with knowledge transfer for credit scoring,” SSRN Electron. J., 2022.

[191] W. Chen, Y. Wang, P. Gupta, and M. K. Mehlawat, “A novel hybrid heuristic algorithm for a new uncertain mean-variance-skewness portfolio selection model with real constraints,” Appl. Intell., vol. 48, no. 9, pp. 2996–3018, 2018.

[192] S. K. Mishra, G. Panda, and B. Majhi, “Prediction based mean-variance model for constrained portfolio assets selection using multiobjective evolutionary algorithms,” Swarm Evol. Comput., vol. 28, pp. 117–130, 2016.

[193] F. Cong and C. W. Oosterlee, “Multi-period mean–variance portfolio optimization based on Monte-Carlo simulation,” J. Econ. Dyn. Control, vol. 64, pp. 23–38, 2016.

[194] Y. Fang, K. K. Lai, and S.-Y. Wang, “Portfolio rebalancing model with transaction costs based on fuzzy decision theory,” Eur. J. Oper. Res., vol. 175, no. 2, pp. 879–893, 2006.

[195] V. N. Katsikis, S. D. Mourtas, P. S. Stanimirović, S. Li, and X. Cao, “Time-varying mean-variance portfolio selection under transaction costs and cardinality constraint problem via beetle antennae search algorithm (BAS),” SN Oper. Res. Forum, vol. 2, no. 2, 2021.

[196] X. Mei, H. Zhu, and C. Chen, “Mean-variance portfolio selection with estimation risk and transaction costs,” Appl. Econ., pp. 1–18, 2022.

[197] K. T. Lwin, R. Qu, and B. L. MacCarthy, “Mean-VaR portfolio optimization: A nonparametric approach,” Eur. J. Oper. Res., vol. 260, no. 2, pp. 751–766, 2017.

[198] K. Kobayashi, Y. Takano, and K. Nakata, “Bilevel cutting-plane algorithm for cardinality-constrained mean-CVaR portfolio optimization,” J. Glob. Optim., vol. 81, no. 2, pp. 493–528, 2021.

[199] R. K. Yew Low, R. Faff, and K. Aas, “Enhancing mean–variance portfolio selection by modeling distributional asymmetries,” J. Econ. Bus., vol. 85, pp. 49–72, 2016.

[200] G.-Y. Ban, N. El Karoui, and A. E. B. Lim, “Machine learning and portfolio optimization,” Manage. Sci., vol. 64, no. 3, pp. 1136–1154, 2018.

[201] D. Bauder, T. Bodnar, N. Parolya, and W. Schmid, “Bayesian mean–variance analysis: optimal portfolio selection under parameter uncertainty,” Quant. Finance, vol. 21, no. 2, pp. 221–242, 2021.

[202] R. AbdelKawy, W. M. Abdelmoez, and A. Shoukry, “A synchronous deep reinforcement learning model for automated multi-stock trading,” Prog. Artif. Intell., vol. 10, no. 1, pp. 83–97, 2021.

[203] P. Gupta, M. K. Mehlawat, and G. Mittal, “Asset portfolio optimization using support vector machines and real-coded genetic algorithm,” J. Glob. Optim., vol. 53, no. 2, pp. 297–315, 2012.

[204] E. A. Gerlein, M. McGinnity, A. Belatreche, and S. Coleman, “Evaluating machine learning classification for financial trading: An empirical approach,” Expert Syst. Appl., vol. 54, pp. 193–207, 2016.

[205] M. A. H. Dempster and V. Leemans, “An automated FX trading system using adaptive reinforcement learning,” Expert Syst. Appl., vol. 30, no. 3, pp. 543–552, 2006.

[206] D. Fister, J. C. Mun, V. Jagrič, and T. Jagrič, “Deep learning for stock market trading: A superior trading strategy?,” Neural Netw. World, vol. 29, no. 3, pp. 151–171, 2019.

[207] X. Wu, H. Chen, J. Wang, L. Troiano, V. Loia, and H. Fujita, “Adaptive stock trading strategies with deep reinforcement learning methods,” Inf. Sci. (Ny), vol. 538, pp. 142–158, 2020.

[208] B. Huang, Y. Huan, L. D. Xu, L. Zheng, and Z. Zou, “Automated trading systems statistical and machine learning methods and hardware implementation: a survey,” Enterp. Inf. Syst., vol. 13, no. 1, pp. 132–144, 2019.

[209] Y. Deng, Y. Kong, F. Bao, and Q. Dai, “Sparse coding-inspired optimal trading system for HFT industry,” IEEE Trans. Industr. Inform., vol. 11, no. 2, pp. 467–475, 2015.

[210] F. Rundo, “Deep LSTM with Reinforcement Learning layer for financial trend prediction in FX high frequency trading systems,” Appl. Sci. (Basel), vol. 9, no. 20, p. 4460, 2019.

[211] A. Dyck, A. Morse, and L. Zingales, “Who blows the whistle on corporate fraud?,” J. Finance, vol. 65, no. 6, pp. 2213–2253, 2010.

[212] L. J. Abbott, Y. Park, and S. Parker, “The effects of audit committee activity and independence on corporate fraud,” Manag. finance, vol. 26, no. 11, pp. 55–68, 2000.

[213] C. T. Spathis, “Detecting false financial statements using published data: some evidence from Greece,” Manag. Audit. J., vol. 17, no. 4, pp. 179–191, 2002.

[214] V. Sugumaran, V. Muralidharan, and K. I. Ramachandran, “Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing,” Mech. Syst. Signal Process., vol. 21, no. 2, pp. 930–942, 2007.

[215] E. Kirkos, C. Spathis, and Y. Manolopoulos, “Data Mining techniques for the detection of fraudulent financial statements,” Expert Syst. Appl., vol. 32, no. 4, pp. 995–1003, 2007.

[216] J. W. Lin, M. I. Hwang, and J. D. Becker, “A fuzzy neural network for assessing the risk of fraudulent financial reporting,” Manag. Audit. J., vol. 18, no. 8, pp. 657–665, 2003.

[217] P. Craja, A. Kim, and S. Lessmann, “Deep learning for detecting financial statement fraud,” Decis. Support Syst., vol. 139, no. 113421, p. 113421, 2020.

[218] Y. Bao, B. Ke, B. Li, Y. J. Yu, and J. Zhang, “Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach,” J. Acc. Res., vol. 58, no. 1, pp. 199–235, 2020.

[219] B. Unger and F. van Waarden, “How to dodge drowning in data? Rule- and risk-based anti money laundering policies compared,” Rev. Law Econ., vol. 5, no. 2, pp. 953–985, 2009.

[220] N. Heidarinia, A. Harounabadi, and M. Sadeghzadeh, “An intelligent anti-money laundering method for detecting risky users in the banking systems,” Int. J. Comput. Appl., vol. 97, no. 22, pp. 35–39, 2014.

[221] S. Gao and D. Xu, “Conceptual modeling and development of an intelligent agent-assisted decision support system for anti-money laundering,” Expert Syst. Appl., vol. 36, no. 2, pp. 1493–1504, 2009.

[222] J. Tang and J. Yin, “Developing an intelligent data discriminating system of anti-money laundering based on SVM,” in 2005 International Conference on Machine Learning and Cybernetics, 2005.

[223] L.-T. Lv, N. Ji, and J.-L. Zhang, “An RBF neural network model for anti-money laundering,” in 2008 International Conference on Wavelet Analysis and Pattern Recognition, 2008.

[224] Y. Zhang and P. Trubey, “Machine learning and sampling scheme: An empirical study of money laundering detection,” Comput. Econ., vol. 54, no. 3, pp. 1043–1063, 2019.

[225] M. Alkhalili, M. H. Qutqut, and F. Almasalha, “Investigation of applying machine learning for watch-list filtering in anti-money laundering,” IEEE Access, vol. 9, pp. 18481–18496, 2021.

[226] J.-J. Rocha-Salazar, M.-J. Segovia-Vargas, and M.-M. Camacho-Miñano, “Money laundering and terrorism financing detection using neural networks and an abnormality indicator,” Expert Syst. Appl., vol. 169, no. 114470, p. 114470, 2021.

[227]M. Weber et al., "Anti-money laundering in Bitcoin: Experimenting with Graph Convolutional Networks for financial forensics," 2019.

[228]D. Cheng, Y. Ye, S. Xiang, Z. Ma, Y. Zhang, and C. Jiang, "Anti-money laundering by group-aware deep graph learning," IEEE Trans. Knowl. Data Eng., pp. 1–13, 2023.

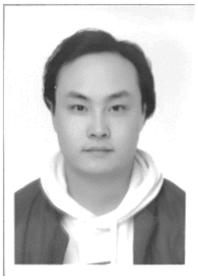
[229]H. Zhang, J. Hong, F. Dong, S. Drew, L. Xue, and J. Zhou, "A privacy-preserving hybrid federated learning framework for financial crime detection," arXiv [cs.LG], 2023.

[230]S. Wang, M. A. Qureshi, L. Miralles-Pechuán, Thien Huynh-The, T. R. Gadekallu, and M. Liyanage, "Applications of explainable AI for 6G: Technical aspects, use cases, and research challenges," arXiv [cs.NI], 2021.

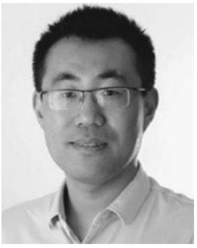
[231]M. Alazab, S. P. Rm, Parimala, P. K. R. Maddikunta, T. R. Gadekallu, and Q.-V. Pham, "Federated Learning for Cybersecurity: Concepts, Challenges, and Future Directions," IEEE Trans. Industr. Inform., vol. 18, no. 5, pp. 3501–3509, 2022

ZIWEI YI received a bachelor's degree from the School of Economy and Business of Jiangxi University of Science and Technology, China, in 2020, and a master's degree from Jiangxi University of Science and Technology in 2023, all in management. Her research interests include corporate finance, computational and quantitative finance, and financial markets.

XINWEI CAO received the bachelor's degree from Shandong University, China, in 2009, the master's degree from Tongji University, China, in 2012, and the Ph.D. degree from Fudan University, China, in 2017, all in management. She is currently a professor with School of Business, Jiangnan University, China. Her research interests include operational management, computational and quantitative finance, and financial market.



ZUYAN CHEN received the B.E. degree in communication engineering from the Jiangxi University of Science and Technology, China. He is currently pursuing a Master by Research degree in electronic and electrical engineering with Swansea University, UK. His research interests include robot force control and machine learning.



SHUAI LI (Senior Member, IEEE) received the B.E. degree in precision mechanical engineering from the Hefei University of Technology, Hefei, China, in 2005, the M.E. degree in automatic control engineering from the University of Science and Technology of China, Hefei, in 2008, and the Ph.D. degree in electrical and computer engineering from the Stevens Institute of Technology, Hoboken, NJ, USA, in 2014. He is currently a Full Professor with the Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland and also with VTT-Technical National Research Center of Finland, Oulu, Finland. His current research interests include dynamic neural networks, robotics, machine learning, and autonomous systems.