Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Supplementary Materials

Jun ZHOU, Chao-Chao CHEN, Long-Fei LI, Zhi-Qiang ZHANG, Xiao-Lin ZHENG, 2022. FinBrain 2.0: When Finance Meets Trustworthy AI. *Front Inform Technol Electron Eng*, 23(4):617-629.
https://doi.org/10.1631/FITEE.2200039

Credit scoring

The credit scoring problem for financial operations is usually modeled as a binary classification problem based on debt repayment, which is calculated using various factors, including past performance and profiling on debt obligations. In practice, logistic regression (LR) models, a well-known statistical method, are used to assess creditworthiness due to their simplicity and interpretability. However, sophisticated machine learning models can be found in the literature to replace it. Studies have shown that the use of more complex artificial intelligence (AI) models in credit scoring can further reduce costs and improve results while allowing the assessment of the creditworthiness of customers with limited credit history (Djeundje et al., 2021). Specifically, in 2015, Lessmann et al. (2015) compared 41 classifiers in terms of six performance metrics across eight real-world credit scoring data sets and showed that some complex classifiers (neural network [NN], tree-based models, and ensemble learning) are significantly more accurate in predicting credit risk than standard LR models. Subsequently, Bhatore et al. (2020) expressed a similar view through a comprehensive literature review: "With the advent of the Big Data era, the use of complex machine learning algorithms for high-resolution risk assessment has become an industry trend." Recently, Gunnarsson et al. (2021) evaluated XGBoost (Chen and Guestrin, 2016) models and deep neural network (DNN) using a range of credit scoring data sets and performance metrics, and they concluded that XGBoost should be preferred over other credit scoring methods when classification performance is the main goal of credit scoring activities and when the computational cost is taken into account. The following is a brief overview of the research progress in terms of deep learning, tree-based models, and ensemble learning. With regard to deep learning models, in 2019, Babaev et al. (2019) proposed the embedding transactional recurrent NN (RNN), i.e., the ET-RNN methods, which use RNNs on fine-grained transaction data to calculate the credit scores of loan applicants and is applied to a credit application scenario of a European bank. While targeting networked-guarantee loans, Cheng et al. (2020a) found that existing credit risk models in the banking industry use machine learning methods to generate a credit score for each customer, which do not consider the contagion risk of the guarantee chain and require extensive feature engineering with deep domain expertise. For this reason, they proposed a new approach to rate contagion chain risk in the banking industry using DNNs, applying a spatiotemporal interchain attention network on graph-structured loan behavior data to compute contagion chain risk scores. Similarly, Lee et al. (2021) proposed a graph convolutional network-based credit default prediction model using three types of relational graphs related to borrowers, i.e., loan information, credit history information, and soft information, which can reflect the nonlinear relationship between borrowers' attributes and default risk, as well as the higher-order relationships among borrowers, and demonstrated the method's validity on a lending platform scenario. In terms of tree-based models, Xia et al. (2020) developed a strategy to dynamically assign weights to the underlying model based on the degree of overfitting and proposed a heterogeneous ensemble tree-based credit scoring model, and the proposed approach significantly outperformed the benchmark model in most cases. In addition, Li et al. (2021) identified those early defaulted borrowers by using a multilayer structured gradient boosting decision tree (GBDT) algorithm with LightGBM (Ke et al., 2017). Due to the extremely unbalanced sample distribution and the high cost of misclassification, they further applied a cost-sensitive framework to the loss function of the classification model to improve the prediction accuracy. With reference to the ensemble models, He and Fan (2021) proposed a novel hybrid ensemble model for default prediction as follows: firstly, they used LightGBM to build new feature interactions to enhance feature representation; second, they used convolutional neural network (CNN) to build new feature interactions to reflect deeper information; finally, they used an ensemble learning approach to combine deep learning classifiers with tree-based classifiers to obtain better prediction results. To reduce the adverse effects of outliers present in noise-filled credit data sets, Zhang et al. (2021a) enhanced the local outlier factor algorithm using a bagging strategy to effectively identify outliers and enhance the outlier adaptation of the underlying classifier; they then proposed a stack-based ensemble learning method with adaptive parameter optimization that outperforms the baseline model. Similarly, they also enhanced the outlier adaptation by a voting-based outlier detection method and extended the downsampling method to deal with data imbalance, and the stack-based ensemble model enhanced the prediction capability (Zhang et al., 2021b). At the same time, these studies found that through a complex scoring approach, it is possible to better serve some customers who have been overlooked in the past, which may promote financial inclusion, but also introduces issues such as opacity and bad interpretability of deep learning models and the lack of comprehensive public data sets (Dastile et al., 2020; Bhatore et al., 2020).

Financial distress prediction

Financial distress is a condition where a company faces financial difficulties. The natural and most likely outcome of financial distress is bankruptcy. In the Chinese stock market, the unique special treatment (ST) warning mechanism can signal financial distress for listed companies. The intention of a financial distress prediction is to disclose the potential operational and financial risks of a company and to alert business owners and managers of such risks before any outbreak. Such a prediction can be useful for managers, investors, and creditors. With respect to managers, a prediction provides them with early warning signals of performance deterioration in order to take corrective actions and reduce the financial distress risk. For investors, understanding the main factors leading to financial distress allows them to avoid investing in risky firms. Creditors should correctly evaluate the firmâĂŹs financial situation and be vigilant to signs of impending financial distress to avoid capital loss and costs related to counterpart risk. There are currently a significant number of credit rating agencies (such as Moody's, Standard & Poor's, and Fitch Ratings) in the industry (Table S1), and there are also a number of domestic credit rating agencies in China (Livingston et al., 2018): Chengxin, Lianhe, and Pengyuan. In academia and industry, many deep learning methods have been adopted in the field. For instance, in unstructured text, Rönnqvist and Sarlin (2017) used a combination of event and text mining to identify whether banks are facing distress. Similarly, Matin et al. (2019) developed a convolutional RNN that uses unstructured text data onto company annual reports, i.e., auditor's reports and management statements, which can learn textual descriptive representations suitable for corporate distress prediction, bringing a boost to distress prediction performance. In some specific scenarios, Gregova et al. (2020) compared traditional (LR) and new learning methods (random forest and NN) to reveal their prediction accuracy under Slovak industrial firms, and the comparison of different metrics showed that the NN models produce better results. Zhang et al. (2022) constructed a dataset of financial distress of Chinese listed companies from 2014 to 2019 and proposed a DNN approach using a genetic algorithm to automatically determine the parameter values and the number of neurons in each hidden layer, which outperforms other tested algorithms after adequate comparisons. Ensemble tree-based learning methods, however, continue to dominate research in this area in terms of the number of papers published. Huang and Yen (2019) selected a total of 16 financial variables from the financial statements of publicly listed Taiwanese companies as the input to the machine learning methods, and the results of the study showed that XGBoost provides more accurate financial distress prediction results compared with support vector machine (SVM) and deep belief network (DBN). There are many similar studies and a few examples are listed here. Du et al. (2020) achieved better performance with XGBoost as the core model for predicting the financial distress of Chinese listed companies. Mousavi and Lin (2020) used data from similar Chinese companies to answer the question: âĂŸWhich distress prediction models perform better in predicting distress?âĂŹ And the results showed that random forest is one of the best machine learning methods. Shen et al. (2020) proposed a new dynamic financial distress prediction method for unbalanced data sets and conceptual drift in data flow in wealth distress prediction and found that random forest classifier outperforms other commonly used classifiers. Based on financial data from 4,167 listed businesses in China from 2001 to 2019, Qian et al. (2022) discovered that the XGBoost-based model has higher prediction accuracy. The above text reveals that financial distress studies still focus on large enterprises as well as listed companies, and it is less common to see related studies on small- and medium-sized enterprises (SMEs), probably due to the lack of public data, data availability, and data privacy. This may harm the fairness and inclusiveness of AI.

Table. S1 Comparison of typical corporate credit rating agencies

âĂŤâĂŤâĂŤâĂŤâĂŤâĂŤâĂŤâĂŤ	Global agencies	Chinese domestic agencies
Representative enterprise	MoodyâĂŹs, Standard & PoorâĂŹs, and Fitch Ratings	Chengxin, Lianhe, and Pengyuan
Date of establishment	1900 (MoodyâĂŹs)	1992 (Chengxin)
Data sources	Financial statements	Financial statements + annual reports + legal judgments
Key factor	Country risk; industry risk; company-level: business operating risk, financial risk (e.g., cash flow, financial leverage, etc.), and corporate governance risk	Country risk; industry risk; company-level: business operating risk, financial risk (e.g., cash flow, financial leverage, etc.), and corporate governance risk; policy risk/relation risk/public sentiment risk
Model	Statistical models (e.g., scorecard)	Statistical models + machine learning
Coverage	Listed companies + bond issuer	Listed companies + bond issuer + small- and midsize enterprises (SMEs)

Bankruptcy prediction

Bankruptcy is the conclusive affirmation of the inability of a company to support and endure current operations given its current financial position and debt obligations. Predictions of corporate bankruptcy are used in various sectors of the entire economy. Companies can diagnose their current situation and formulate corresponding strategies based on predictive models. Executives can run their companies' businesses more stably by managing key indicators that affect the risk of corporate bankruptcy. Investors can modify their strategies and adjust their portfolios by studying the likelihood of corporate bankruptcy. In addition, governments can use corporate bankruptcy forecasting to improve relevant financial regulations. In these ways, bankruptcy forecasting models can help design and improve financial systems. The recent global financial crisis and the increase in credit risk highlight the critical nature of this area. If bankruptcy could be predicted with adequate precision ahead of time, managers and investors of companies may have the possibility to take action to secure their companies, reduce risk and loss of business, and even avoid bankruptcy itself. Due to its importance, bankruptcy prediction has been widely studied and can be broadly classified into statistical models and machine learning models (Kim et al., 2020; Mai et al., 2019). Researchers use statistical models in bankruptcy prediction to identify the most relevant characteristics and their relative weights, and this identification can help test bankruptcy theories and accept regulation of financial markets. Popular statistical models include discriminant analysis, LR models, and factor analysis. In contrast to the statistical models, the machine learning techniques make fewer assumptions about the data. Moreover, models that allow nonlinear decision boundaries have quickly gained popularity and are now widely applied. Among the machine learning algorithms, ensemble methods, tree-based models, and DNN algorithms are mainly used. For the sake of brevity, this section only briefly describes the latest research progress of the tree-based model and deep learning. In terms of the tree-based model, Son et al. (2019) applied XGBoost to a dataset audited by Korean credit rating agencies and achieved higher prediction accuracy compared to statistical models. Subsequently, Le et al. (2019) proposed a faster approach that utilizes graphics processing unit (GPU)-based XGBoost with the squared logistic loss for bankruptcy prediction. They utilized a histogram-based tree construction algorithm, which achieves effective improvement in the Area under the ROC Curve (AUC) (Ling et al., 2003) and processing time. Moscatelli et al. (2020) analyzed the performance of GBDT models in predicting corporate bankruptcy risk and found that when the dataset is larger and has more features, the GBDT model shows significant improvement in discriminative ability and accuracy compared to LR, while this advantage is negligible when the dataset is small. Recently, for mid- and long-term bankruptcy prediction (up to several months) for SMEs, Perboli and Arabnezhad (2021) found that the GBDT algorithm shows substantially improved prediction accuracy, compared to the state-of-the-art techniques, and outperforms LR and even some NN models. As for deep learning, many studies point out that DNNs can extract useful representations from unstructured financial data for prediction. Mai et al. (2019) collected accounting data from 11,827 US-listed companies to train a deep learning model with an average embedding layer to predict bankruptcy, demonstrating how deep learning may use both textual and numerical input to improve prediction accuracy. Hosaka (2019) used financial statements of $\geq$ 2,000 Japanese listed companies, from which a set of financial ratios was obtained and represented as grayscale images, and the images generated by this process were used to train and test CNNs. Bankruptcy predictions through the trained NN were shown to have higher performance compared to methods using decision trees, linear discriminant analysis, SVMs, AdaBoost, and AltmanâĂŹs Z-score (Altman, 2018). Recently, considering that existing methods ignore the rich relational information embodied in financial networks and many studies are mainly developed based on financial statements, Zheng et al. (2021b) proposed a novel heterogeneous graph attention-based model to facilitate the use of publicly available data to predict SME bankruptcies. However, most methods relied heavily on financial ratios or statements, which can be distorted by accounting manipulation techniques such as âĂŸwindow dressing.âĂŹ Furthermore, these data are usually not publicly accessible for SMEs. Furthermore, the most current studies have not exploited financial networks deeply, whose rich relational information can provide valuable insight in inferring firmsâĂŹ credit status.

Portfolio management

The process of continuously reallocating funds into financial assets with the aim of increasing the expected return on investments and minimizing risk is known as portfolio management. It usually exhibits complicated behavior that is intrinsically nonlinear, uncertain, and nonstationary due to external influences such as the global economy and political atmosphere. Because machine learning models can monitor thousands of risk factors daily and test portfolio performance under thousands of market/economic scenarios, AI technologies can enhance risk management for asset managers and other large institutional investors. Feeding machine learning models with big data can provide recommendations to asset managers that influence decisions around portfolio allocation or stock selection. Portfolio management includes the following closely related areas: stock forecasting, portfolio selection, portfolio optimization, portfolio allocation (sometimes these three terms are used interchangeably), robo-advisors, etc. Traditional studies use human-designed features to predict stock movements. However, stock markets are highly stochastic and it is almost impossible to predict stock movements based on historical market prices alone. Therefore, Hu et al. (2018) proposed to predict stock movements based on continuous news, and similarly, Xu and Cohen (2018) proposed to predict stock movements using continuous tweets and market data. Hou et al. (2021) used multigranularity market data and contrastive learning to improve the accuracy of stock trend prediction. Lin et al. (2021a) proposed temporal routing adapters and designed an optimal transmission-based learning algorithm to obtain optimal sample-to-predictor assignments to enhance existing stock prediction models to learn multiple stock trading patterns. Recently, a large number of technical approaches based on historical prices, technical indicators, texts, knowledge graphs, and images to synthesize forecasts have emerged, e.g., Sezer and Ozbayoglu (2018), Li et al. (2019), Long et al. (2020), and Cheng et al. (2020c). In portfolio selection, recent advances in deep learning have stimulated increasing interest in the use of DNNs for portfolios. Soleymani and Paquet (2020) combined restricted stacked autoencoders and CNNs into a deep reinforcement learning (RL) framework to develop a portfolio management approach. Similarly, Xu et al. (2021b) proposed portfolio selection under an RL paradigm and designed a novel relationship-aware transformer to process asset price series. Recently, Shi et al. (2021) designed a new strategy network that uses a temporal CNN to extract multiple temporal features of time series and then used global average pooling and a fully connected layer to integrate global feature maps to handle asset correlations in portfolio management. On behalf of human investment professionals, robo-advisors provide an online investment advisory service for portfolio management according to individual investment propensity and investment purpose with Big Data analysis and advanced algorithm-based automation system. In fact, robo-advisors are not an entirely new technique, but the terminology is new. Traditional wealth management is both expensive and exclusive. Automated investment management in the form of robo-advisors seeks to change this and brings wealth management to an affordable price. It incorporates approaches such as stock forecasting, portfolio selection, portfolio optimization, and predictive analysis in a holistic way. For instance, Rezaei et al. (2021) introduced a CNN-long short-term memory (LSTM) hybrid deep learning model for stock forecasting and incorporated these forecasts as investorsâĂŹ views in the Black-Litterman (Cheung, 2010) asset allocation model. Similarly, Leow et al. (2021) used the bidirectional encoder representations from transformers (BERT) model to capture the latest market conditions through Twitter sentiment and used genetic algorithms to optimize the model for different objectives.

Algorithmic trading (or quantitative trading)

It is defined as buy sell decisions made solely using algorithmic models. Most of the algo-trading applications are coupled with price prediction models for market timing purposes. As a result, the majority of the price or trend forecasting models that trigger buy sell signals based on their predictions are also considered as algo-trading systems (Ozbayoglu et al., 2020). Machine learning techniques for algo-trading can be divided into deep learning and RL algorithm-based methods (Aloud and Alkhamees, 2021). Ozbayoglu et al. (2020) delved into the current state of deep learning-based algorithmic trading research and found that most research focuses on stock or index forecasting and that LSTMs are the most popular deep learning models among these implementations. The core of AI-powered quantitative trading lies in the training and adjusting of models to adapt to the changing market conditions, which is what RL excels at; so, we focus on the latest RL algorithm innovations in this field. Jia et al. (2019) proposed an LSTM-based agent to learn temporal patterns and automatically trade based on current market conditions and historical data using deep RL. Lei et al. (2020) proposed a time-driven feature-aware joint deep RL model, which integrates deep learning models and RL models to improve the learning of financial signal representations and action decisions in algorithmic trading. Recently, Théate and Ernst (2021) proposed a deep RL-based approach to solve the algorithmic trading problem of determining the best trading position at any point in time during stock market trading activities. There are many similar studies, such as those by Li et al. (2019), Park and Lee (2021), and Aloud and Alkhamees (2021). Some practical algo-trading systems have also emerged; for instance, Yin et al. (2022) combined LSTM, attention mechanism, Hawkes process, and graph for learning correlations among stocks and automatically predicting their future prices to achieve a high-frequency quantitative system capable of obtaining stable returns for the Chinese A-share market.

Recommendation

Personalized recommendations are an important feature of the next generation of financial services. NN models and other advanced architectures have produced significant improvements in recommendations. In particular, performance has been significantly improved on platforms that have access to huge amounts of data, such as e-commerce platforms. However, because e-commerce products are different from financial products, we face special challenges in designing personalized systems for the financial industry. For example, the prices of products on e-commerce platforms usually do not change constantly, while the prices of financial assets on financial markets usually change constantly. In fact, in financial markets, the prices of stocks, bonds, and funds change daily; they can even change repeatedly in a second. Furthermore, on e-commerce platforms, product specifications generally remain the same. For example, an iPhone 13 Pro with an Apple A15 Bionic chip will not change to use an Apple A14 Bionic chip (in most cases). However, in the financial markets, a company's business may change every quarter. Since companies are the underlying assets for financial instruments such as stocks and bonds, an opinion about the iPhone 13 Pro may still have value a year later, while opinions about Apple Inc. (AAPL) stock may have no value after the same year. Not all general recommendation methods can be used in the financial sector, but some recommendation models can be used in both e-commerce platforms and financial markets. For example, to make better use of graph structures, Sun et al. (2019) combined online analytical processing techniques with social networks to propose an insurance recommendation framework based on graph mining. To promote fairness in microfinance recommendations, Liu et al. (2019) proposed a fairness-conscious re-ranking algorithm to balance ranking quality and fairness on the borrower's side. To solve the insurance cold-start problem, Bi et al. (2020b) proposed a cross-domain insurance recommendation system based on heterogeneous information networks; meanwhile, given the complexity of insurance products, they designed a meta-path-based approach on the insurance product knowledge graph and proposed the DCDIR, which is another deep cross-domain insurance recommendation system for cold-start users (Bi et al., 2020a). In the click-through rate (CTR) prediction scenario, Zhang et al. (2021) proposed a multi-interaction attention network to comprehensively extract potential relationships among various fine-grained features (e.g., gender, age, and occupation in user profiles) to improve the CTR. To overcome the problem of weakening transferability due to differences between source and target data distributions, Huan et al. (2021) proposed a novel adaptive clustering transfer learning method to improve the conversion rate (CVR). However, the characteristics of the financial domain (risk, return, etc.) must still be taken into account to improve performance. For example, to increase the total revenue of insurance products, Li et al. (2021a) proposed a multitask network for online insurance recommendations, which uses an adaptive attention mechanism that allows efficient feature sharing among complex insurance products and sales scenarios. To predict the future medical behavior of patients and recommend whether patients need medical migration, Cheng et al. (2020d) proposed a medical migration prediction model with attention-based bidirectional gated recurrent units. To improve lenders' investment decisions based on two investment objectives (return, risk, etc.), Babaei and Bamdad (2020) used NNs and LR to estimate the return and the default probability of each investment (loan) and formulated the investment recommendation problem as a multiobjective portfolio optimization problem based on the mean-variance theory.

Marketing

Deep learning has been strategically applied to marketing-related activities in various industries, and the financial sector is no exception. In general, deep learning can help financial services industry target suitable customers and launch suitable marketing campaigns to ensure the effectiveness of its marketing activities in the face of the fierce competition experienced by the financial services industry today. Researchers have experimented with applying deep learning techniques to personalized marketing in the financial services industry, where the most suitable customer group is identified with the appropriate marketing campaigns. For example, to achieve personalized fund recommendations, Chou et al. (2022) proposed a graph-based deep collaborative filtering algorithm. To accomplish the audience expansion task for marketing campaigns of financial institutions, Zhuang et al. (2020) used a hybrid online offline architecture, proposing an adaptive and disentangled graph NN (GNN) in the offline phase and developing a novel audience expansion model with a knowledge distillation mechanism in the online phase to absorb knowledge from the offline network and mitigate coverage bias. To achieve incentive optimization on a limited budget, Liu et al. (2019) proposed a graph representation learning method on transaction networks for merchant incentive optimization in mobile payment marketing; meanwhile, to capture the preferences of inactive users of financial companies and thus recommend suitable content, they defined heterogeneous graphs on social networks and used GNN algorithms to learn inactive users' representation (Liu et al., 2021). To assign approximate optimal incentive values to customers and merchants to encourage mobile payment activities, Yu et al. (2021) used a GNN to model potential two-sided influences and model hazard rates with a piecewise nonlinear function to capture changes in responses to different incentive values.

Customer services

It is another main usage of intelligent technology in the financial sector. The advancements of AI bring us intelligent chatbots/service bots, which offer 24/7 customer services and improve customer satisfaction. For instance, according to Hassani et al. (2020), there have been many successful use cases of chat-bot across the world, e.g., Erica (the virtual assistant of Bank of America), COIN (contract intelligence platform of JPMorgan), AmEx (by American Express), and POSB (DBS Bank). These AI-driven chatbots typically interact with customers via voice or text, as well as on clicking on the options on the screen. Deep learning methods have been widely applied to customer services. For example, Chen et al. (2019a) formulated user intent prediction as a continuous decision process using RL to solve this problem. Subsequently, Xu et al. (2020) proposed an attention-based deep multi-instance sequential intersection network to overcome the problem of feature drift and class imbalance in user intent prediction. Recently, Yang et al. (2021) added a tag recommendation function to reduce the workload on customer service in SMEs and then proposed an intelligent cloud customer service system. In addition to enhancing the customer experience, chatbots have been applied in other fields. In the financial anti-fraud field, Liu et al. (2020) proposed a goal-oriented chat-bot using hierarchical RL and applied it to financial fraud detection conversations. Similarly, Wang et al. (2021b) used RL to learn conversation strategies from real-world human-to-human chat records and designed a voice-enabled bot that seeks additional information from customers through natural conversations to confirm whether they have been defrauded and determine the actual type of fraud. Moreover, chatbots can be found in robo-advisors. For example, Day et al. (2018) developed a comprehensive knowledge-based and generative model for conversational robo-advisors aimed at optimizing portfolios.

Know your customer (KYC)

It is a verification process that the financial institutions need to execute before they can start conducting business with new customers. The increasing level of regulations imposed on this process makes it burdensome. Generally, most KYC efforts use a rule-based approach that is slow and manual. Recently, some work using machine learning approaches for KYC has also emerged. Suzumura et al. (2019) used a federated graph learning platform to share key information among various agencies to combat financial crimes. Huesca et al. (2020) used text and financial transaction data for customer due diligence and for identifying financial crimes using a multiview semisupervised risk classification approach. More recently, Chen (2020) proposed that bank branches are the best level to determine the level of risk and can provide insights into suspicious transaction patterns, showing how financial institutions can use a small set of labeled customer data onto the knowledge discovery process to assess customer risk and combat financial crime, suggesting that machine learning techniques will improve KYC processes and help protect the financial system from illicit activities.

Anti-money laundering (AML)

Broadly speaking, AML refers to all efforts involved in preventing money laundering, such as stopping criminals from becoming customers and monitoring transactions for suspicious activity. The financial services industry and academia agree that machine learning and graph mining could have a significant impact on monitoring currency transaction tools to combat money laundering (Chen et al., 2018; Sobreira Leite et al., 2019). For example, to find suspected money laundering groups, Li et al. (2017) proposed a time-oriented

Louvain algorithm. To address existing AML systems' failure to effectively and efficiently identify hidden and complex money laundering activities, Han et al. (2018) enhanced a new framework of next-generation AML in a distributed and scalable manner by applying visual deep learning-driven natural language processing techniques to provide intelligent and dynamic reporting, reducing the time and cost by about 30%. To discover suspicious behavior in the Bitcoin blockchain, Alarab et al. (2020) used an ensemble learning approach to predict licit and illicit transactions in the network. Recently, to reduce blocked transactions due to false alarms, Alkhalili et al. (2021) adopted SVM algorithms with the polynomial kernel to automatically check blocked transactions in a watchlist filtering system to minimize the workload of compliance officers. To detect the money laundering technique of smurfing, Starnini et al. (2021) found that the speed characteristics of smurfing can help to find smurfs by using a standard database connection. Unlike dense subgraph detection, a flow-based method called FlowScope (Li et al., 2020) was proposed to detect money laundering behavior in the transaction chains.

Representative financial AI practices discussed in Section 3 on "Research and Applications" in the main paper

Table S2 depicts the representative AI approaches discussed in this paper.

Table. S2 Representative financial AI practices (P, privacy; R, robustness; I, interpretability; S, strong; M, medium; W, weak)

Financial services	âĂŤâĂŤ	Data type	Method	P	R	I	Paper
Risk management	Credit scoring	Semistructured	LR	W	W	S	Djeundje et al. (2021)
		Unstructured	RNN, GRU, GNN	W	W	W	Babaev et al. (2019); Cheng et al. (2020a); Lee et al. (2021)
	Financial distress prediction	Unstructured	LSTM, CNN	W	W	W	Matin et al. (2019); Gregova et al. (2020); Zhang et al. (2022)
		Semistructured	tree-based	W	W	S	Huang and Yen (2019); Du et al. (2020); Mousavi and Lin (2020); Shen et al. (2020); Qian et al. (2022)
	Bankruptcy prediction	Unstructured	CNN, GNN, attention	W	W	M	Mai et al. (2019); Hosaka (2019); Zheng et al. (2021b)
		Structured	tree-based	W	W	S	Son et al. (2019); Le et al. (2019); Moscatelli et al. (2020); Perboli and Arabnezhad (2021)
Fraud detection	Credit card fraud detection	SemiStructured	GAN, CNN, GNN	W	M	W	Fiore et al. (2019); Cheng et al. (2020b); Hu et al. (2019); Zhang et al. (2021c); Forough and Momtazi (2021); Zheng et al. (2021a); Zhu et al. (2020)
		Structured	LR, tree-based	W	W	S	Baesens et al. (2021)
	E-commerce transaction fraud detection	Unstructured	GNN, RNN	W	W	W	Liu et al. (2018); Cao et al. (2019); Wang et al. (2021a); Li et al. (2021b); Lin et al. (2021b)
	Loans fraud detection	Unstructured	GNN, attention	W	W	M	Wang et al. (2019); Zhong et al. (2020); Xu et al. (2021a)
	Insurance fraud detection	Unstructured	GNN	W	W	W	Chen et al. (2019b); Liang et al. (2019); Cui et al. (2020)
Wealth management	Portfolio management	SemiStructured	RL, LSTM, BERT	W	W	W	Soleymani and Paquet (2020); Xu et al. (2021b); Shi et al. (2021); Rezaei et al. (2021); Leow et al. (2021); Hu et al. (2018); Xu and Cohen (2018)
	Algorithmic Trading	Unstructured	RL	W	W	W	Jia et al. (2019); Lei et al. (2020); Théate and Ernst (2021); Yin et al. (2022)
Personalized services	Recommendation	Unstructured	Graph, attention	W	W	M	Sun et al. (2019); Bi et al. (2020b); Zhang et al. (2021); Huan et al. (2021); Li et al. (2021a); Cheng et al. (2020d); Babaei and Bamdad (2020)
	Marketing	Unstructured	GNN	W	W	W	Chou et al. (2022); Zhuang et al. (2020); Liu et al. (2019, 2021); Yu et al. (2021)
	Customer services	Unstructured	RL	W	W	W	Chen et al. (2019a); Liu et al. (2020); Wang et al. (2021b); Day et al. (2018)
		Unstructured	NN, attention	W	W	M	Xu et al. (2020); Yang et al. (2021)

*Notes: BERT, bidirectional encoder representations from transformers; CNN, convolutional neural network; GAN, generative adversarial network; GNN, graph neural network; GRU, gated recurrent unit; LR, logistic regression; LSTM, long short-term memory; NN, neural network; RL, reinforcement learning; RNN, recurrent neural network.

References

Alarab I, Prakoonwit S, Nacer MI, 2020. Comparative analysis using supervised learning methods for anti-money laundering in bitcoin. Proceedings of the 2020 5th International Conference on Machine Learning Technologies, p.11-17. https://doi.org/10.1145/3409073.3409078

Alkhalili M, Qutqut MH, Almasalha F, 2021. Investigation of applying machine learning for watch-list filtering in anti-money laundering. *IEEE Access*, 9:18481-18496. https://doi.org/10.1109/ACCESS.2021.3052313

Aloud ME, Alkhamees N, 2021. Intelligent algorithmic trading strategy using reinforcement learning and directional change. *IEEE Access*, 9:114659-114671. https://doi.org/10.1109/ACCESS.2021.3105259

Altman EI, 2018. A fifty-year retrospective on credit risk models, the altman z-score family of models and their applications to financial markets and managerial strategies. *Journal of Credit Risk*, 14(4). https://doi.org/10.21314/JCR.2018.243

Babaei G, Bamdad S, 2020. A multi-objective instance-based decision support system for investment recommendation in peer-to-peer lending. *Expert Systems with Applications*, 150:113278. https://doi.org/10.1016/j.eswa.2020.113278

Babaev D, Savchenko M, Tuzhilin A, et al., 2019. Et-rnn: Applying deep learning to credit loan applications. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, p.2183-2190. https://doi.org/10.1145/3292500.3330693

Baesens B, Höppner S, Verdonck T, 2021. Data engineering for fraud detection. *Decision Support Systems*, 150:113492. https://doi.org/10.1016/j.dss.2021.113492

Bhatore S, Mohan L, Reddy YR, 2020. Machine learning techniques for credit risk evaluation: a systematic literature review. *Journal of Banking and Financial Technology*, 4(1):111-138. https://doi.org/10.1007/s42786-020-00020-3

Bi Y, Song L, Yao M, et al., 2020a. Dcdir: A deep cross-domain recommendation system for cold start users in insurance domain. Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, p.1661-1664. https://doi.org/10.1145/3397271.3401193

Bi Y, Song L, Yao M, et al., 2020b. A heterogeneous information network based cross domain insurance recommendation system for cold start users. Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, p.2211-2220. https://doi.org/10.1145/3397271.3401426

Cao S, Yang X, Chen C, et al., 2019. Titant: online real-time transaction fraud detection in ant financial. *Proceedings of the VLDB Endowment*, 12(12):2082-2093. https://doi.org/10.14778/3352063.3352126

Chen C, Fu C, Hu X, et al., 2019a. Reinforcement learning for user intent prediction in customer service bots. Proceedings of the 42Nd International ACM SIGIR Conference on Research and Development in Information Retrieval, p.1265-1268. https://doi.org/10.1145/3331184.3331370

Chen C, Liang C, Lin J, et al., 2019b. Infdetect: A large scale graph-based fraud detection system for e-commerce insurance. 2019 IEEE International Conference on Big Data (Big Data), p.1765-1773. https://doi.org/10.1109/BigData47090.2019.9006115

Chen T, Guestrin C, 2016. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, p.785-794. https://doi.org/10.1145/2939672.2939785

Chen TH, 2020. Do you know your customer? bank risk assessment based on machine learning. *Applied Soft Computing*, 86:105779. https://doi.org/10.1016/j.asoc.2019.105779

Chen Z, Van Khoa LD, Teoh EN, et al., 2018. Machine learning techniques for anti-money laundering (aml) solutions in suspicious transaction detection: a review. *Knowledge and Information Systems*, 57(2):245-285. https://doi.org/10.1007/s10115-017-1144-z

Cheng D, Niu Z, Zhang Y, 2020a. Contagious chain risk rating for networked-guarantee loans. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, p.2715-2723. https://doi.org/10.1145/3394486.3403322

Cheng D, Xiang S, Shang C, et al., 2020b. Spatio-temporal attention-based neural network for credit card fraud detection. Proceedings of the AAAI conference on artificial intelligence, 34(01):362-369.

Cheng D, Yang F, Wang X, et al., 2020c. Knowledge graph-based event embedding framework for financial quantitative investments. Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, p.2221-2230. https://doi.org/10.1145/3397271.3401427

Cheng L, Shi Y, Zhang K, 2020d. Medical treatment migration behavior prediction and recommendation based on health insurance data. *World Wide Web*, 23(3):2023-2042. https://doi.org/10.1007/s11280-020-00781-3

Cheung W, 2010. The black–litterman model explained. *Journal of Asset Management*, 11(4):229-243.
https://doi.org/10.1057/jam.2009.28

Chou YC, Chen CT, Huang SH, 2022. Modeling behavior sequence for personalized fund recommendation with graphical deep collaborative filtering. *Expert Systems with Applications*, 192:116311.
https://doi.org/10.1016/j.eswa.2021.116311

Cui L, Seo H, Tabar M, et al., 2020. Deterrent: Knowledge guided graph attention network for detecting healthcare misinformation. Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, p.492-502.
https://doi.org/10.1145/3394486.3403092

Dastile X, Celik T, Potsane M, 2020. Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91:106263.
https://doi.org/10.1016/j.asoc.2020.106263

Day MY, Cheng TK, Li JG, 2018. Ai robo-advisor with big data analytics for financial services. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), p.1027-1031.
https://doi.org/10.1109/ASONAM.2018.8508854

Djeundje VB, Crook J, Calabrese R, et al., 2021. Enhancing credit scoring with alternative data. *Expert Systems with Applications*, 163:113766.
https://doi.org/10.1016/j.eswa.2020.113766

Du X, Li W, Ruan S, et al., 2020. Cus-heterogeneous ensemble-based financial distress prediction for imbalanced dataset with ensemble feature selection. *Applied Soft Computing*, 97:106758.
https://doi.org/10.1016/j.asoc.2020.106758

Fiore U, De Santis A, Perla F, et al., 2019. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479:448-455.
https://doi.org/10.1016/j.ins.2017.12.030

Forough J, Momtazi S, 2021. Ensemble of deep sequential models for credit card fraud detection. *Applied Soft Computing*, 99:106883.
https://doi.org/10.1016/j.asoc.2020.106883

Gregova E, Valaskova K, Adamko P, et al., 2020. Predicting financial distress of slovak enterprises: Comparison of selected traditional and learning algorithms methods. *Sustainability*, 12(10):3954.
https://doi.org/10.3390/su12103954

Gunnarsson BR, Vanden Broucke S, Baesens B, et al., 2021. Deep learning for credit scoring: Do or donâĂŹt? *European Journal of Operational Research*, 295(1):292-305.
https://doi.org/10.1016/j.ejor.2021.03.006

Han J, Barman U, Hayes J, et al., 2018. Nextgen aml: Distributed deep learning based language technologies to augment anti money laundering investigation.
https://doi.org/10.18653/v1/P18-4007

Hassani H, Huang X, Silva E, et al., 2020. Deep learning and implementations in banking. *Annals of Data Science*, 7(3):433-446.
https://doi.org/10.1007/s40745-020-00300-1

He H, Fan Y, 2021. A novel hybrid ensemble model based on tree-based method and deep learning method for default prediction. *Expert Systems with Applications*, 176:114899.
https://doi.org/10.1016/j.eswa.2021.114899

Hosaka T, 2019. Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert systems with applications*, 117:287-299.
https://doi.org/10.1016/j.eswa.2018.09.039

Hou M, Xu C, Liu Y, et al., 2021. Stock trend prediction with multi-granularity data: A contrastive learning approach with adaptive fusion. Proceedings of the 30th ACM International Conference on Information & Knowledge Management, p.700-709.
https://doi.org/10.1145/3459637.3482483

Hu B, Zhang Z, Shi C, et al., 2019. Cash-out user detection based on attributed heterogeneous information network with a hierarchical attention mechanism. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01):946-953.
https://doi.org/10.1609/aaai.v33i01.3301946

Hu Z, Liu W, Bian J, et al., 2018. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. Proceedings of the eleventh ACM international conference on web search and data mining, p.261-269.
https://doi.org/10.1145/3159652.3159690

Huan Z, Wang Y, He Y, et al., 2021. Learning to select instance: Simultaneous transfer learning and clustering. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, p.1950-1954.
https://doi.org/10.1145/3404835.3462992

Huang YP, Yen MF, 2019. A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing*, 83:105663.
https://doi.org/10.1016/j.asoc.2019.105663

Huesca JMG, van der Zon S, van Ipenburg W, et al., 2020. Multi-view risk classification for customer due diligence. *AAAI-20 Workshop on Knowledge Discovery from Unstructured Data in Financial Services*, .
Jia W, Chen W, Xiong L, et al., 2019. Quantitative trading on stock market based on deep reinforcement learning. 2019 International Joint Conference on Neural Networks (IJCNN), p.1-8.
https://doi.org/10.1109/IJCNN.2019.8851831
Ke G, Meng Q, Finley T, et al., 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
https://doi.org/doi/10.5555/3294996.3295074
Kim H, Cho H, Ryu D, 2020. Corporate default predictions using machine learning: Literature review. *Sustainability*, 12(16):6325.
https://doi.org/10.3390/su12166325
Le T, Vo B, Fujita H, et al., 2019. A fast and accurate approach for bankruptcy forecasting using squared logistics loss with gpu-based extreme gradient boosting. *Information Sciences*, 494:294-310.
https://doi.org/10.1016/j.ins.2019.04.060
Lee JW, Lee WK, Sohn SY, 2021. Graph convolutional network-based credit default prediction utilizing three types of virtual distances among borrowers. *Expert Systems with Applications*, 168:114411.
https://doi.org/10.1016/j.eswa.2020.114411
Lei K, Zhang B, Li Y, et al., 2020. Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. *Expert Systems with Applications*, 140:112872.
https://doi.org/10.1016/j.eswa.2019.112872
Leow EKW, Nguyen BP, Chua MCH, 2021. Robo-advisor using genetic algorithm and bert sentiments from tweets for hybrid portfolio optimisation. *Expert Systems with Applications*, 179:115060.
https://doi.org/10.1016/j.eswa.2021.115060
Lessmann S, Baesens B, Seow HV, et al., 2015. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1):124-136.
https://doi.org/10.1016/j.ejor.2015.05.030
Li X, Liu S, Li Z, et al., 2020. Flowscope: Spotting money laundering based on graphs. Proceedings of the AAAI conference on artificial intelligence, 34(04):4731-4738.
https://doi.org/10.1609/aaai.v34i04.5906
Li X, Cao X, Qiu X, et al., 2017. Intelligent anti-money laundering solution based upon novel community detection in massive transaction networks on spark. 2017 fifth international conference on advanced cloud and big data (CBD), p.176-181.
https://doi.org/10.1109/CBD.2017.38
Li Y, Zheng W, Zheng Z, 2019. Deep robust reinforcement learning for practical algorithmic trading. *IEEE Access*, 7:108014-108022.
https://doi.org/10.1109/ACCESS.2019.2932789
Li Y, Zhang Y, Gan L, et al., 2021a. Revman: Revenue-aware multi-task online insurance recommendation. Proceedings of the AAAI Conference on Artificial Intelligence, 35(1):303-310.
Li Z, Hui P, Zhang P, et al., 2021b. What happens behind the scene? towards fraud community detection in e-commerce from online to offline. Companion Proceedings of the Web Conference 2021, p.105-113.
https://doi.org/10.1145/3442442.3451147
Li Z, Yang D, Zhao L, et al., 2019. Individualized indicator for all: Stock-wise technical indicator optimization with stock embedding. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, p.894-902.
https://doi.org/10.1145/3292500.3330833
Li Z, Zhang J, Yao X, et al., 2021. How to identify early defaults in online lending: a cost-sensitive multi-layer learning framework. *Knowledge-Based Systems*, 221:106963.
https://doi.org/10.1016/j.knosys.2021.106963
Liang C, Liu Z, Liu B, et al., 2019. Uncovering insurance fraud conspiracy with network learning. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, p.1181-1184.
https://doi.org/10.1145/3331184.3331372
Lin H, Zhou D, Liu W, et al., 2021a. Learning multiple stock trading patterns with temporal routing adaptor and optimal transport. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, p.1017-1026.
https://doi.org/10.1145/3447548.3467358
Lin W, Sun L, Zhong Q, et al., 2021b. Online credit payment fraud detection via structure-aware hierarchical recurrent neural network. IJCAI, p.3670-3676.
https://doi.org/10.24963/ijcai.2021/505
Ling CX, Huang J, Zhang H, et al., 2003. Auc: a statistically consistent and more discriminating measure than accuracy. Ijcai, 3:519-524.
Liu J, Pan F, Luo L, 2020. Gochat: Goal-oriented chatbots with hierarchical reinforcement learning. Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, p.1793-1796.
https://doi.org/10.1145/3397271.3401250

Liu W, Guo J, Sonboli N, et al., 2019. Personalized fairness-aware re-ranking for microlending. Proceedings of the 13th ACM Conference on Recommender Systems, p.467-471.
https://doi.org/10.1145/3298689.3347016

Liu Z, Chen C, Yang X, et al., 2018. Heterogeneous graph neural networks for malicious account detection. Proceedings of the 27th ACM International Conference on Information and Knowledge Management, p.2077-2085.
https://doi.org/10.1145/3269206.3272010

Liu Z, Wang D, Yu Q, et al., 2019. Graph representation learning for merchant incentive optimization in mobile payment marketing. Proceedings of the 28th ACM International Conference on Information and Knowledge Management, p.2577-2584.
https://doi.org/10.1145/3357384.3357835

Liu Z, Shen Y, Cheng X, et al., 2021. Learning representations of inactive users: A cross domain approach with graph neural networks. Proceedings of the 30th ACM International Conference on Information & Knowledge Management, p.3278-3282.
https://doi.org/10.1145/3459637.3482131

Livingston M, Poon WP, Zhou L, 2018. Are chinese credit ratings relevant? a study of the chinese bond market and credit rating industry. *Journal of Banking & Finance*, 87:216-232.
https://doi.org/10.1016/j.jbankfin.2017.09.020

Long J, Chen Z, He W, et al., 2020. An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in chinese stock exchange market. *Applied Soft Computing*, 91:106205.
https://doi.org/10.1016/j.asoc.2020.106205

Mai F, Tian S, Lee C, et al., 2019. Deep learning models for bankruptcy prediction using textual disclosures. *European journal of operational research*, 274(2):743-758.
https://doi.org/10.1016/j.ejor.2018.10.024

Matin R, Hansen C, Hansen C, et al., 2019. Predicting distresses using deep learning of text segments in annual reports. *Expert Systems with Applications*, 132:199-208.
https://doi.org/10.1016/j.eswa.2019.04.071

Moscatelli M, Parlapiano F, Narizzano S, et al., 2020. Corporate default forecasting with machine learning. *Expert Systems with Applications*, 161:113567.
https://doi.org/10.1016/j.eswa.2020.113567

Mousavi MM, Lin J, 2020. The application of promethee multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. *Expert Systems with Applications*, 159:113438.
https://doi.org/10.1016/j.eswa.2020.113438

Ozbayoglu AM, Gudelek MU, Sezer OB, 2020. Deep learning for financial applications: A survey. *Applied Soft Computing*, 93:106384.
https://doi.org/10.1016/j.asoc.2020.106384

Park DY, Lee KH, 2021. Practical algorithmic trading using state representation learning and imitative reinforcement learning. *IEEE Access*, 9:152310-152321.
https://doi.org/10.1109/ACCESS.2021.3127209

Perboli G, Arabnezhad E, 2021. A machine learning-based dss for mid and long-term company crisis prediction. *Expert Systems with Applications*, 174:114758.
https://doi.org/10.1016/j.eswa.2021.114758

Qian H, Wang B, Yuan M, et al., 2022. Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. *Expert Systems with Applications*, 190:116202.
https://doi.org/10.1016/j.eswa.2021.116202

Rezaei H, Faaljou H, Mansourfar G, 2021. Intelligent asset allocation using predictions of deep frequency decomposition. *Expert Systems with Applications*, 186:115715.
https://doi.org/10.1016/j.eswa.2021.115715

Rönnqvist S, Sarlin P, 2017. Bank distress in the news: Describing events through deep learning. *Neurocomputing*, 264:57-70.
https://doi.org/10.1016/j.neucom.2016.12.110

Sezer OB, Ozbayoglu AM, 2018. Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing*, 70:525-538.
https://doi.org/10.1016/j.asoc.2018.04.024

Shen F, Liu Y, Wang R, et al., 2020. A dynamic financial distress forecast model with multiple forecast results under unbalanced data environment. *Knowledge-Based Systems*, 192:105365.
https://doi.org/10.1016/j.knosys.2019.105365

Shi S, Li J, Li G, et al., 2021. Xpm: An explainable deep reinforcement learning framework for portfolio management. Proceedings of the 30th ACM international conference on information & knowledge management, p.1661-1670.
https://doi.org/10.1145/3459637.3482494

Sobreira Leite G, Bessa Albuquerque A, Rogerio Pinheiro P, 2019. Application of technological solutions in the fight against money launderingâĂŤa systematic literature review. *Applied Sciences*, 9(22):4800.
https://doi.org/10.3390/app9224800

Soleymani F, Paquet E, 2020. Financial portfolio optimization with online deep reinforcement learning and restricted stacked autoencoderâĂŤdeepbreath. *Expert Systems with Applications*, 156:113456.
https://doi.org/10.1016/j.eswa.2020.113456

Son H, Hyun C, Phan D, et al., 2019. Data analytic approach for bankruptcy prediction. *Expert Systems with Applications*, 138:112816.
https://doi.org/10.1016/j.eswa.2019.07.033

Starnini M, Tsourakakis CE, Zamanipour M, et al., 2021. Smurf-based anti-money laundering in time-evolving transaction networks. p.171-186.
https://doi.org/10.1007/978-3-030-86514-6_11

Sun S, Wu B, Zhang Z, et al., 2019. A hierarchical insurance recommendation framework using grapholam approach. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, p.757-764.
https://doi.org/10.1145/3341161.3345643

Suzumura T, Zhou Y, Baracaldo N, et al., 2019. Towards federated graph learning for collaborative financial crimes detection. *arXiv* (in press).

Théate T, Ernst D, 2021. An application of deep reinforcement learning to algorithmic trading. *Expert Systems with Applications*, 173:114632.
https://doi.org/10.1016/j.eswa.2021.114632

Wang D, Lin J, Cui P, et al., 2019. A semi-supervised graph attentive network for financial fraud detection. IEEE International Conference on Data Mining, p.598-607.
https://doi.org/10.1109/ICDM.2019.00070

Wang L, Li P, Xiong K, et al., 2021a. Modeling heterogeneous graph network on fraud detection: A community-based framework with attention mechanism. Proceedings of the 30th ACM International Conference on Information & Knowledge Management, p.1959-1968.
https://doi.org/10.1145/3459637.3482277

Wang Z, Wang F, Zhang H, et al., 2021b. 'could you describe the reason for the transfer?' a reinforcement learning based voice-enabled bot protecting customers from financial frauds. Proceedings of the 30th ACM International Conference on Information & Knowledge Management, p.4214-4223.
https://doi.org/10.1145/3459637.3481906

Xia Y, Zhao J, He L, et al., 2020. A novel tree-based dynamic heterogeneous ensemble method for credit scoring. *Expert Systems with Applications*, 159:113615.
https://doi.org/10.1016/j.eswa.2020.113615

Xu B, Shen H, Sun B, et al., 2021a. Towards consumer loan fraud detection: Graph neural networks with role-constrained conditional random field. Proceedings of the AAAI Conference on Artificial Intelligence, 35(5):4537-4545.

Xu K, Zhang Y, Ye D, et al., 2021b. Relation-aware transformer for portfolio policy learning. Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, p.4647-4653.
https://doi.org/10.24963/ijcai.2020/641

Xu K, Fu C, Zhang X, et al., 2020. admscn: A novel perspective for user intent prediction in customer service bots. Proceedings of the 29th ACM International Conference on Information & Knowledge Management, p.2853-2860.
https://doi.org/10.1145/3340531.3412683

Xu Y, Cohen SB, 2018. Stock movement prediction from tweets and historical prices. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), p.1970-1979.
https://doi.org/10.18653/v1/P18-1183

Yang M, Cao S, Hu B, et al., 2021. Intellitag: An intelligent cloud customer service system based on tag recommendation. IEEE 37th International Conference on Data Engineering, p.2559-2570.
https://doi.org/10.1109/ICDE51399.2021.00287

Yin T, Liu C, Ding F, et al., 2022. Graph-based stock correlation and prediction for high-frequency trading systems. *Pattern Recognition*, 122:108209.
https://doi.org/10.1016/j.patcog.2021.108209

Yu L, Wu Z, Cai T, et al., 2021. Joint incentive optimization of customer and merchant in mobile payment marketing. The AAAI Conference on Artificial Intelligence, 35(17):15000-15007.

Zhang K, Qian H, Cui Q, et al., 2021. Multi-interactive attention network for fine-grained feature learning in ctr prediction. The 14th ACM International Conference on Web Search and Data Mining, p.984-992.
https://doi.org/10.1145/3437963.3441761

Zhang R, Zhang Z, Wang D, et al., 2022. Financial distress prediction with a novel diversity-considered ga-mlp ensemble algorithm. *Neural Processing Letters*, 54(2):1175-1194.
https://doi.org/10.1007/s11063-021-10674-9

Zhang W, Yang D, Zhang S, et al., 2021a. A novel multi-stage ensemble model with enhanced outlier adaptation for credit scoring. *Expert Systems with Applications*, 165:113872.
https://doi.org/10.1016/j.eswa.2020.113872

Zhang W, Yang D, Zhang S, 2021b. A new hybrid ensemble model with voting-based outlier detection and balanced sampling for credit scoring. *Expert Systems with Applications*, 174:114744.
https://doi.org/10.1016/j.eswa.2021.114744

Zhang X, Han Y, Xu W, et al., 2021c. Hoba: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture. *Information Sciences*, 557:302-316.
https://doi.org/10.1016/j.ins.2019.05.023

Zheng W, Yan L, Gou C, et al., 2021a. Federated meta-learning for fraudulent credit card detection. The 29th International Conference on International Joint Conferences on Artificial Intelligence, p.4654-4660.
https://doi.org/10.5555/3491440.3492082

Zheng Y, Lee V, Wu Z, et al., 2021b. Heterogeneous graph attention network for small and medium-sized enterprises bankruptcy prediction. Pacific-Asia Conference on Knowledge Discovery and Data Mining, p.140-151.
https://doi.org/10.1007/978-3-030-75762-5_12

Zhong Q, Liu Y, Ao X, et al., 2020. Financial defaulter detection on online credit payment via multi-view attributed heterogeneous information network. Proceedings of The Web Conference 2020, p.785-795.
https://doi.org/10.1145/3366423.3380159

Zhu Y, Xi D, Song B, et al., 2020. Modeling usersâĂŹ behavior sequences with hierarchical explainable network for cross-domain fraud detection. Proceedings of The Web Conference 2020, p.928-938.
https://doi.org/10.1145/3366423.3380172

Zhuang C, Liu Z, Zhang Z, et al., 2020. Hubble: An industrial system for audience expansion in mobile marketing. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, p.2455-2463.
https://doi.org/10.1145/3394486.3403295