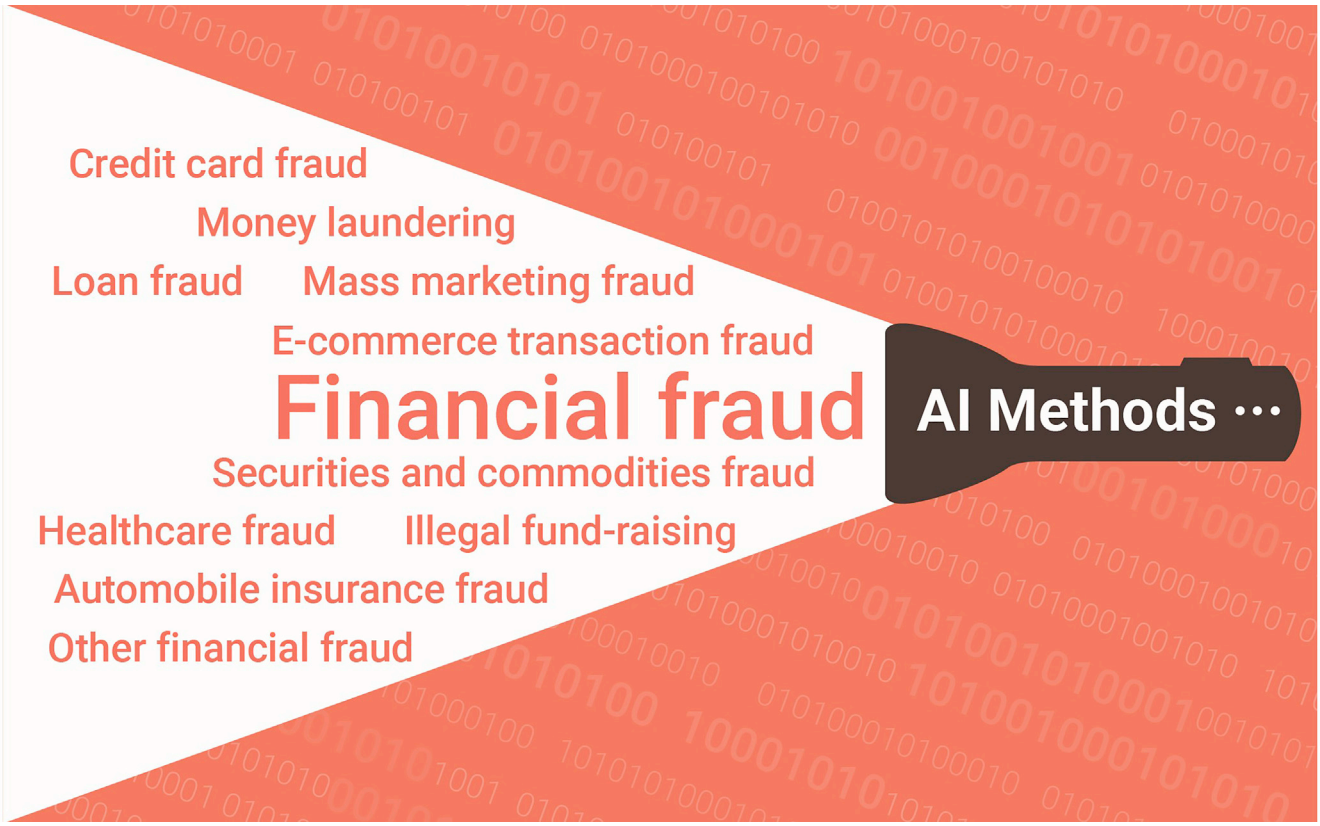
The Innovation

Review

Intelligent financial fraud detection practices in post-pandemic era

Xiaoqian Zhu,[2,4,7] Xiang Ao,[1,3,6,7] Zidi Qin,[1,3] Yanpeng Chang,[4,5] Yang Liu,[1,3] Qing He,[1,3,*] and Jianping Li[2,*]
*Correspondence: heqing@ict.ac.cn (Q.H.); ljp@ucas.ac.cn (J.L.)
Received: May 20, 2021; Accepted: October 18, 2021; Published Online: October 20, 2021; https://doi.org/10.1016/j.xinn.2021.100176

Graphical abstract



Public summary

- Financial fraud in the post-pandemic era is becoming more sophisticated and insidious
- We reiew the development of financial fraud detection from data and method perspectives
- Graph neural network methods are emphasized due to their capacity for heterogeneous data analysis
- Future directions of financial fraud detection are discussed from task, data, and model-oriented perspectives

Review

Intelligent financial fraud detection practices in post-pandemic era

Xiaoqian Zhu,[2,4,7] Xiang Ao,[1,3,6,7] Zidi Qin,[1,3] Yanpeng Chang,[4,5] Yang Liu,[1,3] Qing He,[1,3,*] and Jianping Li[2,*]

[1]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China

[2]School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China

[3]School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China

[4]Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China

[5]School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100049, China

[6]Institute of Intelligent Computing Technology, Suzhou, CAS

[7]These authors contributed equally

*Correspondence: heqing@ict.ac.cn (Q.H.); ljp@ucas.ac.cn (J.L.)

Received: May 20, 2021; Accepted: October 18, 2021; Published Online: October 20, 2021; https://doi.org/10.1016/j.xinn.2021.100176

Citation: Zhu X., Ao X., Qin Z., et al., (2021). Intelligent financial fraud detection practices in post-pandemic era. The Innovation **2(4)**, 100176.

The great losses caused by financial fraud have attracted continuous attention from academia, industry, and regulatory agencies. More concerning, the ongoing coronavirus pandemic (COVID-19) unexpectedly shocks the global financial system and accelerates the use of digital financial services, which brings new challenges in effective financial fraud detection. This paper provides a comprehensive overview of intelligent financial fraud detection practices. We analyze the new features of fraud risk caused by the pandemic and review the development of data types used in fraud detection practices from quantitative tabular data to various unstructured data. The evolution of methods in financial fraud detection is summarized, and the emerging Graph Neural Network methods in the post-pandemic era are discussed in particular. Finally, some of the key challenges and potential directions are proposed to provide inspiring information on intelligent financial fraud detection in the future.

Key words: financial fraud detection; COVID-19 pandemic; artificial intelligence

INTRODUCTION

Over the past decades, financial fraud has brought shocking losses to the global economy, threatening the efficiency and stability of capital markets.[1,2] Making things worse, the coronavirus pandemic (COVID-19) outbreak in early 2020 disrupted the international financial markets in unprecedented ways, heightening the risk of being vulnerable to financial fraud.[3] For example, in April 2020, fraud rates across all financial products in the United Kingdom soared 33% from a year earlier.[4] Meanwhile, Fidelity National Information Services, a payment services provider that assists about 3,200 U.S. banks with fraud monitoring, reported that the lost volume of fraudulent transactions leaped 35% in America compared with the previous period in 2019.[5] Financial fraud in the post-pandemic era is becoming a growing severe problem.

As defined by Black's Law Dictionary, fraud refers to a knowing misrepresentation of the truth or concealment of a material fact to induce another to act to his or her detriment.[6] The classification of financial fraud has not reached a consensus because the types of financial fraud are varied and mounting. Summarizing the previous literature,[7,8] this paper constructs a financial fraud classification framework according to the major financial institution involved. The classification framework is depicted in Figure 1. The frauds related to securities contain securities and commodities fraud, financial statement fraud, among others.[1] Insurance frauds contain health care fraud, automobile insurance fraud, corporate insurance fraud, and so on.[9,10] The frauds closely related to banks are mortgage fraud, loan default, credit card fraud, money laundering, among others.[11] Some frauds that obviously cannot be linked to the above three institutions, such as e-commerce transaction fraud, mass marketing fraud, and illegal fund-raising, are classified as others. Another common perspective is to divide fraud activities into customer level and business level, so we also take them into consideration in the framework. Financial fraud detection at the customer level is mainly related to individual financial activities, including health care insurance, automobile insurance, credit card, loans, e-commerce transaction, and so on,[8,12] whereas business-level fraud crimes, such as financial statement misconduct and money laundering, are often committed by syndicates accompanied by other crimes such as bribery, tax evasion, and even support of terrorism.[13–15]

The ongoing COVID-19 pandemic brings unexpected sudden shock to the global financial system and accelerates the use of digital financial services.[16] These changes have escalated more insidious fraud schemes, providing a breeding ground for all types of financial fraud.[17] On one hand, the economic downturns, caused by the global pandemic, bring proliferating economic pressure and stronger fraud motives to both companies and individuals. For example, in response to the expected cash flow disruption caused by the advent of the COVID-19 crisis, companies withdraw funds on a large scale from pre-existing credit lines.[18] The rising operation costs stemming from the economic shutdown threaten the survival of many companies, inducing an increase in credit fraud.[19] Furthermore, the pressure on corporate financial results intensifies the temptation to manipulate financial statements in order to meet stakeholder expectations.[20] For policyholders, poor financial conditions spawn more speculative insurance claim fraud.[21]

On the other hand, the COVID-19 outbreak significantly accelerates digital transformation and increases digital processes, which sheds new light on fraud activities. The emerging situations can be summarized into two types. The first is that the switch of the business from offline to online exacerbates information asymmetry and leads to increased difficulty in fraud detection.[3] Quarantine regulations create opportunities for online banking and remote transactions, but it is difficult for remote banking to obtain comprehensive information for customer identity verification, resulting in frequent credit fraud incidents.[22] The rise in suspected and proven insurance frauds caused by the claim process adjustment also keeps insurers up at night. The remote work not only expands workload but also hinders access to information.[21] For example, an insurance adjuster may not be able to inspect automobile repairs in detail, which provides opportunities for policyholders to exaggerate billing.

Another situation engendered by the increasing digitalization is that the burgeoning of new financial products and services makes the existing detection methods difficult to adapt. To elude regulators, fraudulent behaviors and types escalate over time, which greatly lowers the effectiveness of the extant approaches. Google reports they are blocking more than 240 million COVID-themed spam emails and 18 million malicious emails related to COVID-19 each day.[23] During the crisis, fraudsters tweak their fraud schemes and add COVID-19 twists to confuse the victims, which makes fraud detection a challenging task for both individuals and detection agencies.[24] Moreover, although digital financial services, such as crowdfunding platforms and

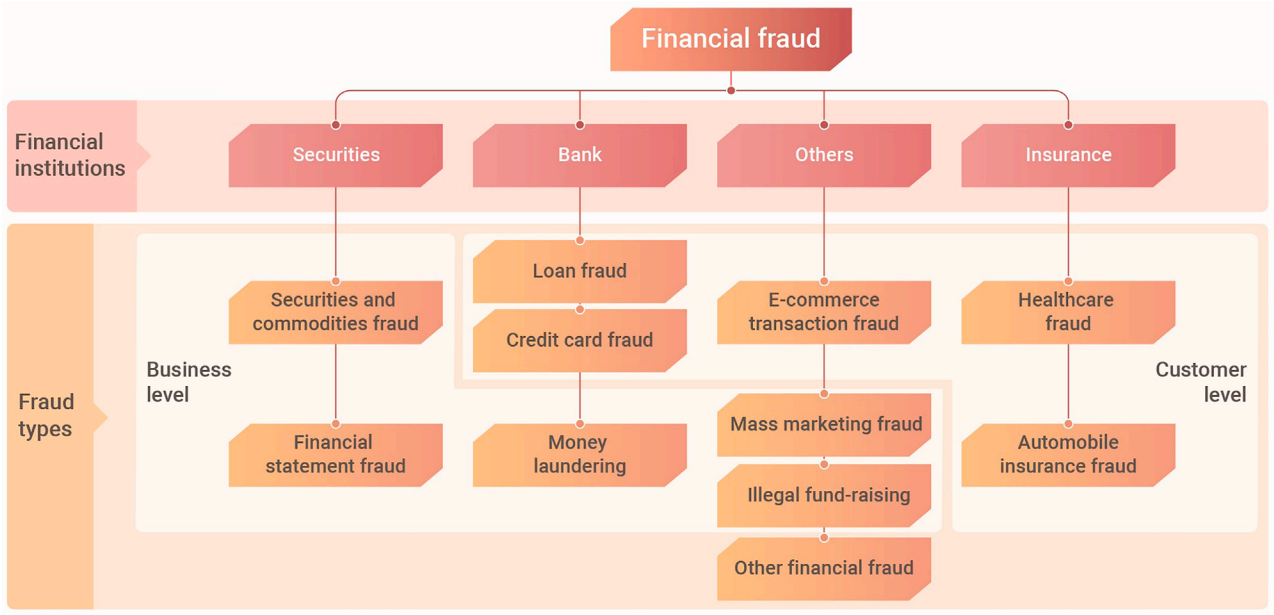


Figure 1. The classification of financial fraud types

digital payments, are quickly applied, the incomplete regulatory policies are conducive to hide fraudsters' identity information or financing history, which leads to credit fraud.[3]

Hence, financial fraud in the post-pandemic era is a critical problem with the characteristics of stronger motives, more insidious forms, and more intelligent schemes. These changes bring considerable challenges to financial fraud detection, including faster detection, better interpretability, and stronger robustness. In addition, the rapid digital transformation is not only an opportunity to obtain richer data for fraud detection but also brings more problems such as how to mine valuable information from massive but low-value-density data more effectively.

Considering the above-mentioned changes, this paper provides a comprehensive review of the development of financial fraud detection practices and highlights the new characteristics of fraud caused by the COVID-19. We first give a brief introduction to the evolution of data types used in financial fraud detection. Through the review from traditional methods to recently proposed methods, the purpose of this paper was to summarize the possible improving directions in response to more insidious fraudsters and provide insights into future algorithm design. Finally, the current challenges and potential directions are outlined to provide some inspiring information on intelligent financial fraud detection in the post-pandemic era.

The remainder of the paper is organized as follows. Section 2, financial fraud detection data evolution, presents the evolution of data used in financial fraud detection. Section 3, survey of methods, discusses the state-of-the-art fraud detection techniques according to the timeline and highlights the progress in recent years. Section 4, challenges and future directions, provides insights into problems and challenges that are still unsolved and points out the directions for future work. Last, the conclusions are summarized in Section 5.

FINANCIAL FRAUD DETECTION DATA EVOLUTION

With the rapid growth of information technology, the types of data used for financial fraud detection continue to expand, which can be roughly divided into three categories, i.e., basic quantitative structured data (a.k.a. tabular data), diverse semi-structured data, and complex unstructured data. Data types and examples are shown in Table 1.

From the very beginning, pioneered by pathfinders, such as Beaver and Altman who stated that a set of financial ratios would be investigated for bankruptcy prediction, numerous encouraging explorations on using quantitative data to predict fraud have been conducted.[35,36] The sources of these structured data consist of corporations, regulators, research teams, commercial companies, and so on. Insurance companies and banks established unique systems to collect and store the basic information of policyholders or account holders.[37] For insurers, the information used for fraud detection, such as insurance claims,[38] the characteristics of incidents,[25] and customer purchase behaviors,[39] are obtained from the claim statement or the policy.[40,41] Banks usually predict fraud with the help of transaction information, such as transactional history and payment observation.[42,43] More comprehensively, regulators collect incidents in the entire industry and issue relevant reports.[25,44] For example, the Securities and Exchange Commission (SEC) has been issuing the Accounting and Auditing Enforcement Releases (AAERs) to investigate companies for alleged accounting misconduct since 1982.[45] To enhance the availability of financial misstatements data, Dechow et al. sorted AAERs information into a numerical database.[27] Commercial companies also collect finance and market information from global institutions and build databases to meet the growing demand for data analysis. For instance, the major accounting and financial databases for researchers in the world include the Compustat North America database by Standard & Poor's and the Worldscope database by Thomson Financial.[46,47] Based on commercial databases, Beneish calculated financial indexes including the gross margin index, asset quality index, and sales growth index when detecting corporation earnings manipulation.[26]

Quantitative data are intuitive and easy to obtain, but the information contained in it is limited. As shown in Table 1, researchers seek other types of data to detect fraud with the continuous escalation of fraud patterns. For example, Law examined the organizational factors of corporate governance that are related to fraud through analyzing questionnaires and interviews from chief financial officers in Hong Kong.[28] By mining event logs for knowledge, process mining that analyzes business processes also assists in fraud detection.[48] Jans et al. developed a system that mined procurement processes to predict exposure opportunities of committing internal transaction fraud.[29] Another typical type of information is the public formatted files from corporations and regulatory authorities. For example, the SEC has required corporations to file key performance reports in the extensible Business Reporting Language (XBRL) format since 2009, which provides high-level data and improves the transparency of corporations.[49] Researchers pull out financial statements from the data repository and then predict financial misconduct through text analysis.[30,50,51]

Table 1. Types and examples of data used for fraud detection

Data type	Examples	Research
Structured	Quantitative numbers	Viaene et al. diagnose automobile insurance claims fraud by using indicators including claimant, insured driver, and lost wages.[25] Beneish detects corporation earnings manipulation by using financial indexes collected from commercial databases.[26] Dechow et al. describe the characteristics of corporation misrepresentation through sorting Accounting and Auditing Enforcement Releases information into a numerical database.[27]
Semi-structured	Interview	Law analyzes the organizational factors of corporate fraud through interviewing chief financial officers.[28]
	Business process	Jans et al. mine procurement processes to predict internal transaction fraud in companies.[29]
	Database system	The Securities and Exchange Commission requires corporations to submit reports in the eXtensible Business Reporting Language (XBRL) language, which provides public and formatted data for fraud detection.[30]
Unstructured	Text	Xiong et al. mine individual opinions on social media to detect corporate disclosure fraud.[31]
	Audio	Hobson et al. analyze the vocal and linguistic cues elicited from speech to detect misreporting.[32]
	Video	Muddy Waters Research analyzes multiple information including store traffic videos to expose Luckin Coffee of fabricating financial numbers.[33]
	Telemetry data	The China Securities Regulatory Commission detects Dalian Zhangzidao Fishery Group's financial fraud by using the BeiDou Navigation Satellite System.[34]

Nowadays, the explosion of information has brought more types of available data, which are mainly unstructured, such as text, video, and telemetry data. Typical examples are shown in Table 1. In addition to financial reports, companies' abundant email archives,[52] public corporate announcements,[53] legal proceedings published by courts,[54] and other textual information have also gradually become raw materials for fraud detection.[55] Furthermore, financial social media platforms have burgeoned in recent years.[45] By mining emotions, social relations, and other information,[31,56] the wisdom of crowds within social media is also a crucial toolkit to capture business information. Besides, multitype unstructured data, such as audio,[32,57,58] image,[59] and video,[60] are playing important roles. Recently, Muddy Waters Research analyzed multiple information, including customer receipts and store traffic videos, and accused Luckin Coffee of fabricating financial and operating numbers since the third quarter of 2019.[33] The China Securities Regulatory Commission recorded the working location and duration of the Dalian Zhangzidao Fishery Group's fishing vessels by use of the BeiDou Navigation Satellite System to expose the Chinese A-share listed fishery group that pretended that their scallops had escaped four times in 6 years to inflate profits.[34]

Notably, fraud detection, regardless of the fraud type, is faced with continuously growing data and informationthat need to be effectively mined and integrated. Reviewing the history of data types mentioned above, the data used in fraud detection practices have experienced the development from basic quantitative data to the current multi-source data. The combination of multi-source information can provide a more panoramic view of financial activities and brings opportunities for better fraud detection. It is also the general trend of scientific research in various fields.[61] However, this evolution also brings great challenges in developing intelligent methods to effectively integrate and utilize panoramic data in future detection practices.

SURVEY OF METHODS

Analogous to the evolution of data types, methods for fraud detection experienced a rapid proliferation in the past decades. Especially in the post-pandemic era, due to the intensified motives, insidious forms, and intelligent schemes of financial fraud, it is becoming more difficult to identify fraudulent behaviors accurately and efficiently. Thus, recently, researchers tend to incorporate and exploit information from as many aspects as possible for comprehensive monitoring.[62] Following these trends, in this section, we survey existing financial fraud detection methods based on the technical development routes. We highlight the research proposed in the recent 2 years to demonstrate how researchers excavate related information from multiple perspectives in the post-pandemic era. For those antiquated techniques, we merely list representative cases to clarify the historical line. Table 2 depicts the representative financial fraud detection approaches we discuss in this section.

Rule-based expert systems

In the early stages, data used for fraud detection are usually highly structured, e.g., transaction logs or well-designed financial metrics, and the means for detecting fraud are undecorated. A number of rules and static thresholds can be used to filter out misbehavior. A straightforward case is that a system will alert if important indexes like liquidity or profitability are unusually high or low.[26] Then, expert systems were designed to facilitate the work of human auditors. They generally use symbolic rules to encode knowledge created by human experts, which was an important part of artificial intelligence during the 1970s and 1980s. This encoded knowledge base is then queried to yield a result through reasoning.[105] For example, Quinlan et al. and Cohen et al. introduced a set of if-then statements to recognize fraud records in multiple fields.[98,99] Moreover, association rules,[63] fuzzy rules,[64] and manual trial-and-error rules[65] are applied to settle the problems of credit card fraud detection as well.

Nevertheless, these manual and rule-based approaches have become particularly costly and ineffective at present.[106] As fraudsters begin to employ trickier strategies to elude regulators, rich financial-related information is required to be analyzed, which undoubtedly exacerbates difficulties in extracting and summarizing effective rules. Small sets of human-summarized rules are no longer sufficient to meet the demand, motivating to build and maintain a large set of rules.[107] However, managing a large ruleset requires more computing resources and is challenging to evaluate and understand.[105]

Traditional machine learning algorithms

Considering the defects of rule-based approaches, growing numbers of machine learning-based methods have been developed. They usually start with extracting statistical features relevant to the given task, such as user profiles, credit history, and historical transactions.[94] After performing feature engineering, a classifier can be trained with these features.[71] Next, we introduce several typical algorithms and their corresponding applications in financial fraud detection.

Naive Bayes, Logistic Regression (LR), and Support Vector Machine (SVM) are standard linear classifiers that have shown excellent performance in various applications.[25,108–110] Naive Bayes is a simple probabilistic classifier based on the "Bayes" theorem under the assumption of strong (naive) independence of the attributes. Panigrahi et al. proposed a well-designed model for credit card fraud detection, combining a Dempster-Shafer adder with a Bayesian learner.[69] Deng designed a fraudulent financial statements detection model based on a Naive Bayes classifier to facilitate human auditors.[82] LR classifies the existing data by establishing regression equations classification boundaries, mainly used for binary classification problems.[111,112] Artıs et al. applied LR model to detect fraudulent insurance claims based on the Spanish market and estimated the error rate.[87] Viaene et al. considered the damages and audit costs and applied LR model to decide suspicious

Table 2. Financial fraud detection practices discussed in the section "survey of methods"

Fraud type		Data type	Algorithm	XAI	Research
Credit fraud	Customer level	Structured	Expert system	●	Brause et al.,[63] HaratiNik et al.,[64] Correia et al.[65]
			SVM	●	Dheepa et al.[66]
			RF	●	Noghani et al.[67]
			CNN	*	Fu et al.[68]
		Semi-structured	Naive bayes	●	Panigrahi et al.[69]
			CNN	*	Zheng et al.[70]
		Unstructured	FNN, Att.	○	HACUD[71]
			LSTM, Att.	○	MAHINDER[72]
			GNN	*	PC-GNN[73]
			GNN, Att.	○	AMG-DP,[74] SemiGNN[75]
			GNN, LSTM, Att.	*	TemGNN[76]
Money laundering	Business level	Unstructured	Graph AD	●	FlowScope[77]
			Supervised network analysis	●	Savage et al.[78]
			GNN	*	Weber et al.[79]
Loan fraud	Customer level	Unstructured	GNN, GRU, Att.	*	DGANN[80]
			GNN, LSTM, Att.	*	ST-GNN[81]
Financial statement fraud	Business level	Structured	Naive bayes	●	Deng[82]
			SVM	●	Ravisankar et al.[83]
			RF, GBT, Rule ensembles	●	Whiting et al.[84]
			FNN	*	Green and Choi,[85] Fanning and Cogger[86]
Insurance fraud	Customer level	Structured	LR	●	Artís et al.,[87] Viaene et al.[88]
			GBT	●	Guelman[89]
		Unstructured	GNN	*	Liang et al.[90]
E-commerce transaction fraud	Customer level	Semi-structured	LSTM	*	Jurgovsky et al.[91]
			GRU	*	Branco et al.[92]
		Unstructured	RNN	*	CLUE[93]
			LSTM, Att.	○	LIC Tree-LSTM[94]
			FNN, Att., FM	○	HEN[95]
			FNN, Att., FM	*	NHFM,[96] DIFM[97]
Others		Structured	Expert system	●	Quinlan et al.,[98] Cohen et al.[99]
		Unstructured	Graph AD	●	Li et al.[100]
			GNN	*	CARE-GNN[101]
			GNN, Att.	*	Player2Vec,[102] GraphConsis,[103] PIdentifier[104]

AD, anomaly detection; Att., attention; XAI, explainable artificial intelligence; ● represents non-deep method and is generally considered to be interpretable; ○ represents the method claims to be interpretable; * indicates that it is hard to evaluate.

claims.[88] SVM is also a linear classifier that separates all data samples into correct classes by finding the maximum margin hyperplane. Kernel techniques and margin optimization are two critical properties of SVM.[43,113] With these two tricks, SVM is capable of solving complex fraud detection problems. To name some, Ravisankar et al. tested SVM techniques on data from 202 Chinese companies to find out a fraudulent financial statement.[83] Dheepa and Dhanapal employed behavior-based SVM to predict suspicious transactions.[66]

Tree-based classifiers attempt to separate data into exclusive categories. Each leaf node represents a specific class, and each tree branch represents a possible attribute value.[114] Decision Tree is the most fundamental one; however, it is likely to be unstable and easily over-fitting. Therefore, more advanced tree-based classifiers such as Random Forest (RF),[115] XGBoost,[116] or LightGBM[117] apply ensemble strategies such as bagging and boosting to improve performance. In the financial detection area, tree-based models have shown performance superior to other learning algorithms like SVM.[118–120] For example, Guelman researched Gradient Boosting Tree (GBT) in modeling auto insurance loss cost based on data from a Canadian company.[89] Whiting et al. reported the performance of methods including RF, GBT, and rule ensembles when applying to financial fraud detection.[84] Taking feature selection and decision cost into account, Noghani and Moattar proposed an advanced RF-based model, which yielded certain performance improvements.[67]

Furthermore, some applications represent transaction data as graphs, using nodes to represent financial entities and edges to represent money transfer.[121] After extracting features through feature engineering and graph-embedding techniques to preserve topological and structural properties,[122–124] machine learning models are built afterward. For example, Savage et al. extracted meaningful communities from the network and performed classification to detect money-laundering activities.[78] A few works consider graph anomaly detection skills, as fraud can be seen as unusual events different from normal behaviors.[125–127] For instance, Li et al. spotted potential fraudulent cases in trading networks by finding the black hole and volcano patterns.[100] Li et al. modeled the laundering as densest and multi-step money flow and proposed an algorithm FlowScope to search dense flow accurately and efficiently in large transaction graphs.[77]

Deep-learning-based approaches

Deep Learning (DL) is becoming a particular type of machine learning, as it achieves great success in various domains. At its heart, the most essential advantages of DL models are that they can extract features directly from raw data without hard-coding task-specific knowledge or tedious feature engineering.[128] With the increasingly complex fraud in the financial scenario, researchers try their best to use these massive and various data to uncover these concealed miscreants. Thus, DL techniques for fraud detection have gained popularity over recent years, especially in the post-pandemic era where digital transformation has become the new normal. In this section, we discuss the surveyed approaches according to the different types of input data.

Modeling tabular data. In the first few years, researchers merely used the basic feedforward neural networks (FNNs), also known as multi-layer perceptrons (MLPs), as classifiers based on static tabular data.[129] for example, Green and Choi presented a neural network classifier employing variables related to the financial statement.[85] Fanning and Cogger also used an artificial neural network for management of fraud prediction.[86] Their input vectors mainly consist of financial ratios and qualitative variables derived from financial statements. Though many attempts using MLP in financial fraud detection have shown better performance than rule-based systems and other classification methods like LR,[130–133] these networks are acyclic and incapable of modeling sequential data that might be essential to discover anomaly users or transactions.[129]

Modeling sequential data. Hence, for better excavating and utilizing sequential data, more complex and elaborate network structures are designed. Convolutional Neural Networks (CNNs), with the convolutional operations, are capable of capturing short-term contextual information and can be applied in financial fraud detection. For example, Fu et al. recombined transaction data to feature matrices and performed a CNN-based approach to identify latent fraud behaviors.[68] Zheng et al. formulated a meta-learning-based classifier, including a feature extraction module, a K-Tuplet Network based on ResNet-34, which is a typical CNN structure.[70]

Besides CNN, cyclic DL models, e.g., Recurrent Neural Networks (RNNs), are further proposed and developed for sequence prediction.[134–136] In RNNs, the output of the last hidden layer is also the input of the current hidden layer, which renders it suitable to encode variable sequences of inputs. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are typical architectures of RNNs. They introduce "Gates" to optionally let information through to avoid the problems of gradient vanishing and exploding.[137–139] As temporal information is a crucial factor in financial data analysis, the RNN models significantly outperform basic MLPs due to their abilities to encode sequential data.[140] To name some, Wang et al. presented a novel deep-learning-based system, namely CLUE, to detect transaction fraud at JD.com, one of China's largest e-commerce platforms.[93] Jurgovsky et al. considered the fraud detection problem in e-commerce as a sequence classification task and employed LSTM networks to incorporate the historical behavior of the users for detecting fraud on new incoming transactions.[91] Branco et al. introduced a GRU-based framework to detect payment card fraud, in which the payments are treated as an interleaved sequence.[92] Liu et al. devised a behavior tree and introduced a Local Intention Calibrated Tree-LSTM (LIC Tree-LSTM) for fraud transaction detection.[94] The behavior tree is built by splitting and reorganizing the sequential behavioral data, and its branch represents a specific user intention.

In addition to CNNs and RNNs, quite a few other techniques can be employed to model sequential data. For example, Zhu et al. proposed a Hierarchical Explainable Network (HEN) to model users' behavior sequences.[95] In HEN, a field-level extractor encodes both first- and second-order information through Factorization Machines (FM).[141] Then, an event-level extractor captures higher-order feature interactions for better sequence representation. Similarly, Xi et al. designed a Neural Hierarchical Factorization Machine (NHFM) model, a two-level architecture capturing feature interactions and representations of users' historical events.[96] They further presented the Dual Importance-aware Factorization Machines, which exploits users' historical behavior in dual perspectives.[97]

Modeling relational data. Though sequential data demonstrate effectiveness in detecting fraud among users and transactions, the changes in the post-pandemic era propel modeling relational data in urgent demand. As we mentioned before, the intensified motives, insidious forms, and intelligent schemes impel comprehensive data analysis and considering interaction relations among users. It thus motivates graph-based DL and Graph Neural Networks (GNNs) widely applied in the financial fraud detection areas since the graph is a natural choice for presenting relational data.[122,126,127,142,143]

Homogeneous relations. Initially, most graph-based methods only consider homogeneous graphs, in which the node and edge types are undifferentiated. Even so, they have yielded promising performance in financial crime and fraud detection, especially GNNs, which have the potential to improve structural representations and causal reasoning.[121] They broadly follow a recursive message passing schema, in which each node computes its new representation through aggregating feature vectors of its neighbors.[144] For instance, Weber et al. applied Graph Convolutional Network (GCN), a typical GNN model, in anti-money laundering.[79] Liang et al. introduced a device-sharing network among claimants and developed a GNN-based solution to uncover groups of organized fraudsters for return-freight insurance on the e-commerce platform.[90]

Furthermore, as most real-world graphs are dynamic, a few models consider an additional time dimension based on previous studies. By combining GNN and RNN in different ways, dynamic GNNs are proposed to mine structural and temporal information simultaneously. For example, DGANN is a dynamic graph-based attention neural network for risk guarantee relationship prediction.[80] Each node in the graph represents a company, and each edge represents a guarantee. In the model, a GCN layer with structural attention can process each snapshot, a Graph Recurrent Network with temporal attention is applied to exploit the temporal relationships between snapshots. Similarly, Yang et al. proposed a Spatial Temporal GNN (ST-GNN) to mine credible supply chain relationships, including risk analysis of small and medium-sized enterprises.[81] Wang et al. proposed a Temporal-Aware GNN (TemGNN) to model the credit risk prediction on dynamic graphs.[76] Considering the time interval irregularity between dynamic snapshots, TemGNN adopts

an interval-decayed attention mechanism and can assemble short- and long-term temporal-structural information.

Heterogeneous relations. Although previous homogeneous works offer practical solutions for modeling relational data, they still have limited ability to capture information from realistic situations, especially the multi-relational data that emerged in the post-virus era. As remote businesses and transactions hinder access to comprehensive user information for identity verification, this switch from offline to online exacerbates information asymmetry and makes data completeness and data quality a major concern. Multi-source data are thus collected to alleviate this problem and better model user profiles. Researchers start focusing on heterogeneous graphs, as they contain multiple types of nodes and links to represent different entities and relations, which mimic the data flows more closely in the real-world network.[74,145–148] For instance, under problem formulation in Zhong et al., a node can be a customer, a merchant, or a device.[72] In the graph constructed in Hu et al., an edge implies social connection, money transaction, or device ownership, and so forth.[74] The heterogeneous graph also can be termed as a Heterogeneous Information Network. Meanwhile, in the recent scenario of financial fraud detection, DL solutions for the heterogeneous graph are often proposed under the Attributed Heterogeneous Information Network (AHIN), where both nodes and edges may contain attributes (or named features). Thus, we discuss the heterogeneous graph-based methods under the concept of AHIN.

While AHIN is a powerful information modeling method for characterizing data heterogeneity,[74] it brings about extra challenges in designing algorithms because of its complex topology and higher feature dimensions. One intuitive solution for AHIN is decomposing the heterogeneous graph as a combination of series of homogeneous graphs and fusing the homogeneous representations. For example, Hu et al. devised AMG-DP that employs relation-specific receptive layers to distinguish neighbors by relation attributes.[74] After aggregating the neighbor information following a typical GNN schema, representations incorporating rich semantics derived from multiplex relations are learned. Then, they implement a relation-specific attention mechanism to integrate multiple representations adaptively for loan default prediction. Zhang et al. proposed Player2Vec to identify key players in online underground forums.[102] In the model, GCN is employed to learn embedding from each single-view attributed graph. Then, an attention mechanism fuses the learned embedding based on different single-view attributed graphs to get the final representations. Similarly, Wang et al. proposed a semi-supervised attentive GNN, named SemiGNN, which applies a hierarchical attention mechanism to correlate different neighbors and different views better.[75]

The aforementioned heterogeneous GNNs reveal illegal acts through aggregating nodes' neighborhood information across different relations. However, under the fraud detection scenario, some inherent characteristics of the data will hamper the performance of GNN-based fraud detectors, so a few methods are proposed to alleviate these issues. For example, to escape regulation, fraudsters will camouflage themselves through adjusting their behavior to act like benign users or connecting themselves to benign users, which we call the feature and relation camouflage. Dou et al. propose CARE-GNN, consisting of three neural modules against the camouflage.[101] A label-aware similarity measure and a similarity-aware neighbor selector are leveraged to find informative neighboring nodes. A relation-aware neighbor aggregator combines neighborhood information across different relations with trainable weights. Sharing a similar idea, Liu et al. introduced a GNN framework, namely GraphConsis, to alleviate the problems of context, feature, and relation inconsistency.[103] Besides, class imbalance also has negative influence on models, which means the label distribution of samples is heavily skewed. Liu et al. proposed a Pick and Choose GNN (PC-GNN) to remedy this challenge.[73] In PC-GNN, first, nodes and edges are picked with a devised label-balanced sampler to construct sub-graphs for mini-batch training. Next, for each node in the sub-graph, the neighbor candidates are chosen by a proposed neighborhood sampler. Finally, information from the selected neighbors and different relations is aggregated to obtain the final representation of a target node.

Another route for modeling AHIN is encoding nodes' or links' attributes via meta-path sampling. Meta-path is a path sampled over graphs according to preset rules, refined from prior experience about specific fraud patterns.[149] For example, "User $\xrightarrow{Transaction}$ Merchant $\xrightarrow{Transaction}$ User" represents all paths starting from a user node, passing a merchant node, ending in a user node via two "Transaction" edges. The interaction relations among users can be explored according to the guidance of predefined meta-paths. Hu et al. proposed HACUD, which picks meta-path-aware neighborhoods for each node, then aggregates features with a hierarchical attention mechanism to classify whether a user is cash-out or not.[71] Zhong et al. proposed MAHINDER for financial defaulter detection, which implements meta-path sampling and considers multi-view decomposing.[72] Unlike HACUD, MAHINDER models each meta-path by an LSTM-based encoder to capture local structural patterns and then adopts attention mechanisms on the node, link, and meta-path levels to learn fusion weights. The works of Hu et al. and Zhong et al. are typical meta-path-based algorithms in AHIN, although do not follow GNNs' typical message passing schema.[71,72] Fan et al. further proposed PIdentifier to detect illicit trade in the underground market, which upgrades kernel of meta-path to meta-graph, a graphlet composed of meta-paths.[104] For each sampled meta-graph, a representation is learned based on a meta-graph-guided search. Multi-head attention is computed to construct embedding for buyer nodes and products nodes separately.

The global coronavirus pandemic makes it harder to detect suspects for the following reasons: economic fallout brings stronger fraud motive, social distancing hinders information collection, and accelerated digital transformation affects existing detection methods. Reviewing and summarizing the representative cases mentioned above, we see that in response to the problem, the anti-fraud systems begin excavating deeper user-related information, like sequential and relational data, and gather information from multiple sources to better model real-world activities. Consequently, the data are getting more irregular, from numerical indicators to transaction networks, from Euclidean to non-Euclidean data. In this case, DL techniques are becoming increasingly popular, as they can identify and combine crucial features from unstructured data to achieve high performance without any domain knowledge. In addition, graph-based, especially heterogeneous graph-based fraud detection, has been focused on recently, as graphs can capture rich behavioral interactions.

CHALLENGES AND FUTURE DIRECTIONS

Although data-driven artificial intelligent techniques have achieved excellent performance in the financial fraud detection domain, there are still key issues remaining unsolved, as financial fraud schemes are rapidly evolving to adapt to this new digital environment. In this section, we provide the major challenges and suggest directions for future work from task-oriented, data-oriented, and model-oriented perspectives.

Financial fraud is harder to identify due to its increasing secretiveness and complexity

One of the severe difficulties for financial fraud detection is that the fraud is hidden in complex financial activities. The increased motives and the accelerated digital transformation caused by the pandemic even lead to more intelligent fraud schemes, which makes fraud more difficult to identify. These issues bring two essential challenges for detection.

The secretiveness of financial fraud leads to the natural error in samples. Fraud detection, in many cases, can be regarded as a classification task essentially, which requires fraud samples and non-fraud samples as training data. However, as the fraud activities are increasingly hidden, in most portions of practices, fraud usually cannot be fully identified by regulators and market participants. Consequently, the non-fraud samples used for training may contain some unrecognized fraud samples, leading to natural errors among training samples. When the natural error rate of samples is serious, the basic features of fraud and non-fraud samples captured by the detection model may have fundamental errors, but the users of the model are not aware of them, thus seriously threatening the accuracy of detection results.

The complexity of financial activities leads to massive information involved. The financial activities are related to a wider range of business. Therefore, the involved information is massive but heterogeneous, accompanied by lower-value density. The multi-source information will be difficult to play its role if it is not well integrated. Some researchers have explored models for storing and analyzing massive data, among which the knowledge graph is most suitable for solving this problem. The knowledge graph is a knowledge system connecting all data through the relationships between the data.[150,151] This knowledge-based system, if possible, will contain information about every entity related to fraud in the real world,[152,153] which provides a panoramic perspective. Furthermore, the logic consistency analysis between different nodes of the knowledge graph can help verify the authenticity of information and correct inconsistent information.[154] Powerful knowledge reasoning technologies based on knowledge graphs can help mine the secret relationship between entities connected with fraud and provide potential evidence to make up for missing information.[155,156] Thus, the knowledge graph will be one of the most important and promising tools that mine valuable information for comprehensive detection in the future.

Financial data for fraud detection is massive but scattered

In the information explosion era, the multi-source data are massive but usually scattered across different institutions. At the same time, detecting fraud activities increasingly requires the support of panoramic data to gain a comprehensive understanding of miscreant activities. It thus remains challenging in integrating these scattered data and processing the massive data efficiently.

Data isolation is difficult to resolve. Although the amount of data used in fraud detection is much more tremendous than before, most of the data exist in the form of isolated islands, i.e., scattered in different institutions or even different countries.[157] It will be hard to provide a comprehensive view of financial activities due to the difficulties in data aggregation, which will further greatly affect the effectiveness of detection methods. Google proposes the federated learning framework, which helps to construct a complete and powerful model through joint modeling of multiple institutions.[158,159] However, some key issues remain to be studied, such as the data formats of different institutions are inconsistent, and the network connections between institutions are unstable.

Large-scale data processing brings great challenges to model training. The increase of digital services records more user footprints and information, but it also brings more challenges to massive data processing. Many detection methods require plenty of time to optimize parameters, and the time grows nonlinearly with the expansion of the sample size. Time-consuming modeling cannot obtain the detection model quickly, so it is difficult to update the detection model in time. For example, DL can be applied to process large amounts of data, but training parameters are extremely time-consuming.[160,161] Further research is required to fully develop and apply more advanced technologies to solve these practical fraud detection problems.

Financial fraud detection models need to be more flexible and interpretable

Nowadays, though the emerging research and application of GNN and other models have helped improve financial fraud detection efficiency by utilizing multiple types of information, there are still many challenges with the practicality of the detection models such as model bias, robustness, and interpretability.

Model bias issue needs to be taken into account. Model bias is a significant issue in the machine learning field, which refers to the difference between the model prediction and the actual value we are trying to predict. In fraud detection practice, there are roughly two reasons for model bias: one is the problem of the data samples; the other is from the models themselves. Class imbalance is a crucial factor to high model bias and is overwhelmingly observed in fraud detection, as regularly fraudsters are far fewer than regular users. Models performing poorly on the minority may lead to undesirable results, as people are more concerned about the minority classes, i.e., the fraudsters. The class imbalance problem on feature-based neural methods has been studied in depth, such as re-sampling,[162–164] re-weighting,[165–168] and transfer learning.[95,169] Whereas in the GNNs works, the noisy information, few interactions among fraudsters, and desalination of the minority class's features caused by the message aggregation of GNNs are three major challenges in designing class imbalanced GNNs for fraud detection.[73,170] Future studies that follow-up on these directions would be beneficial. There are other model biases caused by samples that need to be addressed, such as the under-representation, ignoring sensitive attributes, and the social feedback loops.[171,172] As for the detection models themselves, the initial design flaw of the model is also one of the essential factors leading to model bias, which is very hard to avoid. However, there is also recent research progress working on calibrating such a kind of bias.[173]

Robustness needs to be strengthened. The fraudster and the anti-fraud party are always in a dynamic game. With the new technology, the game among fraudsters, financial institutions, and regulators is upgrading, presenting high confrontational features. The robustness and adversarial issues based on conventional DL models have attracted extensive attention from researchers[174–176]; however, the study for GNNs is still in its nascent stages.[177,178] Moreover, due to the message propagation mechanism of GNNs, the effects caused by small perturbations can spread, resulting in even worse performance than non-GNNs. In financial scenarios, attackers always aim for interference with defense models to seek exorbitant profits. Hence, how to detect and defend against harmful perturbations and design robust models, especially for GNN, are becoming major implementation goals.

Interpretability needs to be improved. A key factor in the success of deep neural networks is the fact that networks can be seen as a very large number of nonlinear functions, rendering them possible to learn features at various levels of abstraction with the cost of interpretability and explainability.[179–181] As a result, they cannot be fully trusted in critical applications such as financial fraud detection. Although several post hoc explanatory methods have been developed recently to understand DL models, research has shown that many interpretation methods may produce unfaithful results.[182–185] Especially for a graph-based neural network, its unique non-Euclidean structure brings more challenges, as gradient or backpropagation-related methods cannot be directly applied. Although researchers have made explorations on interpreting GNNs, most of them are still working on toy examples and cannot solve problems in real-world financial scenarios.[186–188] Hence, further research is required to understand not only conventional GNNs but also more complex structures, such as models on AHIN.

CONCLUSIONS

In this survey, we provided a comprehensive overview of financial fraud detection practices from three aspects: the impact of the pandemic, the evolution of the data, and the advancement of methods. The unprecedented pandemic shocked the global financial system and accelerated digital transformation, which brings stronger motives, more insidious forms, and more intelligent schemes of financial fraud activities.

As for the data, applying more panoramic data to comprehensively detect fraud activities is the prevailing trend. The data used in fraud detection practices have experienced the development from basic quantitative data to the current multi-source unstructured data. In the post-pandemic era, explosive data provide more information than before, and fraud detection is inclined to use multi-source data to obtain a comprehensive understanding of financial activities.

As for the model, DL systems have been popular recently for their versatility and revolutionary success in financial fraud detection. The graph-based detection approach is an emerging direction to analyze multi-source data of fraud activities. With the rapid development of technology, financial scenarios and behaviors are becoming more intelligent and sophisticated. Graph-based detection, such as GNN, attracts more attention since the graph can gather information from multiple sources to better model real-world activities and detect hidden anomalies more effectively.

Although the data-driven DL models have been proven to be helpful in fraud detection problems, there are still many challenges to be solved for future development. Complex and hidden fraud activities bring greater challenges to a comprehensive understanding and accurate identification. Achieving efficient integration and processing of massive but scattered financial data is one of the important foundations for panoramic fraud detection. Finally, the flexibility, robustness, and interpretability challenges of models need to be considered more seriously in the context of financial fraud.

REFERENCES

1. Amiram, D., Bozanic, Z., Cox, J.D., et al. (2018). Financial reporting fraud and other forms of misconduct: a multidisciplinary review of the literature. Rev. Account. Stud. ***23***, 732–783.
2. Li, C., Lou, C., Luo, D., et al. (2021). Chinese corporate distress prediction using LASSO: the role of earnings management. Int. Rev. Financ. Anal. ***76***, 101776.
3. Karpoff, J.M. (2020). The future of financial fraud. J. Corp. Financ. ***66***, 101694.
4. Experian; National Hunter Fraud Prevention Service (2020). Fraud rate rises 33% during Covid-19 lockdown. https://www.experianplc.com/media/news/2020/fraud-rate-rises-33-during-covid-19-lockdown.
5. Andriotis, A., and McCaffrey, O. (2020). Borrower, beware: credit-card fraud attempts rise during the coronavirus crisis. Wall Street J. https://www.wsj.com/articles/borrower-beware-credit-card-fraud-attempts-rise-during-the-coronavirus-crisis-11590571800.
6. Garner, B.A. (2009). Black's Law Dictionary, 9th Edition (West Group Publishing House), p. 731.
7. Ngai, E.W.T., Hu, Y., Wong, Y.H., et al. (2011). The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. Decis. Support Syst. ***50***, 559–569.
8. West, J., and Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. Comput. Secur. ***57***, 47–66.
9. Kose, I., Gokturk, M., and Kilic, K. (2015). An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Appl. Soft Comput. ***36***, 283–299.
10. Yan, C., Li, Y., Liu, W., et al. (2020). An artificial bee colony-based kernel ridge regression for automobile insurance fraud identification. Neurocomputing ***393***, 115–125.
11. Al-Hashedi, K.G., and Magalingam, P. (2021). Financial fraud detection applying data mining techniques: a comprehensive review from 2009 to 2019. Comput. Sci. Rev. ***40***, 100402.
12. Modi, K., and Dayma, R. (2017). Review on fraud detection methods in credit card transactions. In 2017 International Conference on Intelligent Computing and Control (I2C2) (IEEE), pp. 1–5.
13. Canhoto, A.I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: an affordances perspective. J. Bus. Res. ***131***, 441–452.
14. Islam, S.R., Khaled Ghafoor, S., and Eberle, W. (2018). Mining illegal insider trading of stocks: a proactive approach. In Proc. 2018 IEEE Int. Conf. Big Data, pp. 1397–1406.
15. Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: a decade review from 2004 to 2015. J. Data Sci. ***14***, 553–569.
16. Li, J., Guo, K., Viedma, E.H., et al. (2020). Culture versus policy: more global collaboration to effectively combat COVID-19. Innovation ***1***, 100023. https://doi.org/10.1016/j.xinn.2020.100023.
17. Agur, I., Peria, S.M., and Rochon, C. (2020). Digital Financial Services and the Pandemic: Opportunities and Risks for Emerging and Developing Economies (International Monetary Fund).
18. Li, L., Strahan, P.E., and Zhang, S. (2020). Banks as lenders of first resort: evidence from the COVID-19 crisis. Rev. Corp. Financ. Stud. ***9***, 472–500.
19. Hasan, I., Politsidis, P.N., and Sharma, Z. (2021). Global syndicated lending during the COVID-19 pandemic. J. Bank Financ. ***16***, 106121.
20. Deloitte (2020). Forensic Focus on COVID-19 Financial Statement Fraud.
21. FRISS (2020). Insurance Fraud Report 2020.
22. Rosen, L.W. (2020). COVID-19 and Emerging Global Patterns of Financial Crime (Congressional Research Service).
23. Muncaster, P. (2020). Google: We Block 240 Million Daily #COVID19 Spam Messages (Infosecurity Magazine). https://www.infosecurity-magazine.com/news/google-block-240-million/.
24. De, R., Pandey, N., and Pal, A. (2020). Impact of digital surge during Covid-19 pandemic: a viewpoint on research and practice. Int. J. Inform. Manage. ***55***, 102171.
25. Viaene, S., Derrig, R.A., and Dedene, G. (2004). A case study of applying boosting naive bayes to claim fraud diagnosis. IEEE Trans. Knowl. Data Eng. ***16***, 612–620.
26. Beneish, M.D. (1999). The detection of earnings manipulation. Financ. Anal. J. ***55***, 24–36.
27. Dechow, P.M., Ge, W., Larson, C.R., et al. (2011). Predicting material accounting misstatements. Contemp. Account. Res. ***28***, 17–82.
28. Law, P. (2011). Corporate governance and no fraud occurrence in organizations. Manag. Audit. J. ***26***, 501–518.
29. Jans, M., Van Der Werf, J.M., Lybaert, N., et al. (2011). A business process mining application for internal transaction fraud mitigation. Expert Syst. Appl. ***38***, 13351–13359.
30. Humpherys, S.L., Moffitt, K.C., Burns, M.B., et al. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. Decis. Support Syst. ***50***, 585–594.
31. Xiong, F., Chapple, L., and Yin, H. (2018). The use of social media to detect corporate fraud: a case study approach. Bus. Horizons ***61***, 623–633.
32. Hobson, J.L., Mayew, W.J., and Venkatachalam, M. (2012). Analyzing speech to detect financial misreporting. J. Account. Res. ***50***, 349–392.
33. Muddy Waters Research (2020). Luckin Coffee: Fraud + Fundamentally Broken Business.
34. Zhang, P. (2020). Chinese Securities Regulator Uses BeiDou Satellites to Expose Firm's Financial Fraud. https://cntechpost.com/2020/06/24/chinese-securities-regulator-uses-beidou-satellites-to-expose-firms-financial-fraud/.
35. Beaver, W.H. (1966). Financial ratios as predictors of failure. J. Account. Res. ***4***, 71–111.
36. Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Financ. ***23***, 589–609.
37. Yeh, I.C., and Lien, C.H. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert Syst. Appl. ***36***, 2473–2480.
38. Ghani, R., and Kumar, M. (2011). Interactive learning for efficiently detecting errors in insurance claims. In Proc.17th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 325–333.
39. Lin, W., Wu, Z., Lin, L., et al. (2017). An ensemble random forest algorithm for insurance big data analysis. IEEE Access ***5***, 16568–16575.
40. Artıis, M., Ayuso, M., and Guillen, M. (1999). Modelling different types of automobile insurance fraud behaviour in the Spanish market. Insur. Math. Econ. ***24***, 67–81.
41. Viaene, S., Dedene, G., and Derrig, R.A. (2005). Auto claim fraud detection using Bayesian learning neural networks. Expert Syst. Appl. ***29***, 653–666.
42. Chan, P.K., Fan, W., Prodromidis, A.L., et al. (1999). Distributed data mining in credit card fraud detection. IEEE Intell. Syst. Appl. ***14***, 67–74.
43. Bhattacharyya, S., Jha, S., Tharakunnel, K., et al. (2011). Data mining for credit card fraud: a comparative study. Decis. Support Syst. ***50***, 602–613.
44. Weisberg, H.I.D., and Richard, A. (1991). Fraud and automobile insurance: a report on bodily injury liability claims in Massachusetts. J. Insur. Regul. ***9***, 497–542.
45. Dong, W., Liao, S., and Zhang, Z. (2018). Leveraging financial social media data for corporate fraud detection. J. Manag. Inform. Syst. ***35***, 461–487.
46. Ulbricht, N., and Weiner, C. (2005). Worldscope meets Compustat: a comparison of financial databases.
47. Guenther, D.A., and Rosman, A.J. (1994). Differences between COMPUSTAT and CRSP SIC codes and related effects on research. J. Account. Econ. ***18***, 115–128.
48. Van Dongen, B.F., de Medeiros, A.K.A., Verbeek, H., et al. (2005). The ProM framework: a new era in process mining tool support. International conference on application and theory of petri nets. In Proc. 26th Int. Conf. Appl. Theor. Petri Nets, pp. 444–454.
49. Debreceny, R., Farewell, S., Piechocki, M., et al. (2010). Does it add up? Early evidence on the data quality of XBRL filings to the SEC. J. Account. Public Policy ***29***, 296–306.
50. Cecchini, M., Aytug, H., Koehler, G.J., et al. (2010). Making words work: using financial text as a predictor of financial events. Decis. Support Syst. ***50***, 164–175.
51. Wei, L., Li, G., Zhu, X., et al. (2019). Discovering bank risk factors from financial statements based on a new semi-supervised text mining algorithm. Account. Financ. ***59***, 1519–1552.
52. Holton, C. (2009). Identifying disgruntled employee systems fraud risk through text mining: a simple solution for a multi-billion dollar problem. Decis. Support Syst. ***46***, 853–864.
53. Chakravarthy, J., DeHaan, E., and Rajgopal, S. (2014). Reputation repair after a serious restatement. Account. Rev. ***89***, 1329–1363.
54. Donelson, D.C., Kartapanis, A., McInnis, J.M., et al. (2021). Measuring accounting fraud and irregularities using public and private enforcement. Account. Rev., Forthcoming.
55. Dimmock, S.G., and Gerken, W.C. (2012). Predicting fraud by investment managers. J. Financ. Econ. ***105***, 153–173.
56. Ru, Y., Xue, J., Zhang, Y., and Zhou, X. (2020). Social connections between media and firm executives and the properties of media reporting. Rev. Account. Stud. ***25***, 963–1001.
57. Hassan, T.A., Hollander, S., Van Lent, L., et al. (2020). Firm-level Exposure to Epidemic Diseases: Covid-19, SARS, and H1N1. In National Bureau of Economic Research (Working Papers 26971).
58. Hassan, T.A., Hollander, S., Van Lent, L., and Tahoun, A. (2019). Firm-level political risk: measurement and effects. Q. J. Econ. ***134***, 2135–2202.
59. Dong, W., and Soh, Y. (2006). Image-based fraud detection in automatic teller machine. Int. J. Compu. Sci. Netw. Secur. ***6***, 13–18.

60. Dhiran, A., Kumar, D., Abhishek, and Arora, A. (2020). Video fraud detection using blockchain. In Proc. 2020 Second Int. Conf. on Inventive Res. Comput. Appl. (ICIRCA), pp. 102–107.
61. Wu, L., Wang, L., Li, N., et al. (2020). Modeling the COVID-19 outbreak in China through multi-source information fusion. The Innovation ***1***, 100033. https://doi.org/10.1016/j.xinn.2020.100033.
62. Wang, S., and Zhu, D. (2020). Interpretable multimodal learning for intelligent regulation in online payment systems. In Proc. 29th Int. Jt. Conf. Artif. Intell., pp. 4675–4681.
63. Brause, R., Langsdorf, T., and Hepp, M. (1999). Neural data mining for credit card fraud detection. In Proc. 11th Int. Conf. Tools Artif. Intell., pp. 103–106.
64. HaratiNik, M.R., Akrami, M., Khadivi, S., and Shajari, M. (2012). FUZZGY: a hybrid model for credit card fraud detection. In 6th Int. Symp. Telecommun., pp. 1088–1093.
65. Correia, I., Fournier, F., and Skarbovsky, I. (2015). The uncertain case of credit card fraud detection. In Proc. 9th ACM Int. Conf. Distrib. Event- Based Syst., pp. 181–192.
66. Dheepa, V., and Dhanapal, R. (2012). Behavior based credit card fraud detection using support vector machines. ICTACT J. Soft Comput. ***2***, 2012.
67. Fadaei Noghani, F., and Moattar, M. (2017). Ensemble classification and extended feature selection for credit card fraud detection. J. AI Data Min. ***5***, 235–243.
68. Fu, K., Cheng, D., Tu, Y., and Zhang, L. (2016). Credit card fraud detection using convolutional neural networks. In Neural Inf. Process., pp. 483–490.
69. Panigrahi, S., Kundu, A., Sural, S., and Majumdar, A.K. (2009). Credit card fraud detection: a fusion approach using Dempster-Shafer theory and Bayesian learning. Inf. Fusion ***10***, 354–363.
70. Zheng, W., Yan, L., Gou, C., and Wang, F.-Y. (2020). Federated meta-learning for fraudulent credit card detection. In Proc. 29th Int. Jt. Conf. Artif. Intell., pp. 4654–4660.
71. Hu, B., Zhang, Z., Shi, C., et al. (2019). Cash-Out user detection based on attributed heterogeneous information network with a hierarchical attention mechanism. Proc. AAAI Conf. Artif. Intell. ***33***, 946–953.
72. Zhong, Q., Liu, Y., Ao, X., et al. (2020). Financial defaulter detection on online credit payment via multi-view attributed heterogeneous information network. Proc. Web Conf. ***2020***, 785–795.
73. Liu, Y., Ao, X., Qin, Z., et al. (2021). Pick and choose: a GNN-based imbalanced learning approach for fraud detection. Proc. Web Conf. ***2021***, 3168–3177.
74. Hu, B., Zhang, Z., Zhou, J., et al. (2020). Loan default analysis with multiplex graph learning. In Proc. 29th ACM Int. Conf. Inf. Knowl. Manag., pp. 2525–2532.
75. Wang, D., Lin, J., Cui, P., et al. (2019). A semi-supervised graph attentive network for financial fraud detection. In 2019 IEEE Int. Conf. Data Min., pp. 598–607.
76. Wang, D., Zhang, Z., Zhou, J., et al. (2021). Temporal-aware graph neural network for credit risk prediction. In Proc. 2021 SIAM Int. Conf. Data Min., pp. 702–710.
77. Li, X., Liu, S., Li, Z., et al. (2020). Flowscope: spotting money laundering based on graphs. In Proc. AAAI Conf. Artif. Intell., *34*, pp. 4731–4738.
78. Savage, D., Wang, Q., Zhang, X., et al. (2017). Detection of money laundering groups: supervised learning on small networks. In AAAI Workshops.
79. Weber, M., Chen, J., Suzumura, T., et al. (2018). Scalable graph learning for anti-money laundering: A first look. arXiv:1812.00076, 1–7. https://arxiv.org/abs/1812.00076.
80. Cheng, D., Wang, X., Zhang, Y., and Zhang, L. (2020). Risk guarantee prediction in networked-loans. In Proc. 29th Int. Jt. Conf. Artif. Intell., pp. 4483–4489.
81. Yang, S., Zhang, Z., Zhou, J., et al. (2020). Financial risk analysis for SMEs with graph-based supply chain mining. In Proc. 29th Int. Jt. Conf. Artif. Intell., pp. 4661–4667.
82. Deng, Q. (2010). Detection of fraudulent financial statements based on Naïve Bayes classifier. In 2010 5th Int. Conf. Comput. Sci. Educ., pp. 1032–1035.
83. Ravisankar, P., Ravi, V., Rao, G.R., et al. (2011). Detection of financial statement fraud and feature selection using data mining techniques. Decis. Support Syst. ***50***, 491–500.
84. Whiting, D.G., Hansen, J.V., McDonald, J.B., et al. (2012). Machine learning methods for detecting patterns of management fraud. Comput. Intell. ***28***, 505–527.
85. Green, B.P., and Choi, J.H. (1997). Assessing the risk of management fraud through neural network technology. Auditing ***16***, 14–28.
86. Fanning, K.M., and Cogger, K.O. (1998). Neural network detection of management fraud using published financial data. Int. Syst. Account. Financ. Manag. ***7***, 21–41.
87. Art'ıs, M., Ayuso, M., and Guill'en, M. (2002). Detection of automobile insurance fraud with discrete choice models and misclassified claims. J. Risk Insur. ***69***, 325–340.
88. Viaene, S., Ayuso, M., Guillen, M., et al. (2007). Strategies for detecting fraudulent claims in the automobile insurance industry. Eur. J. Oper. Res. ***176***, 565–583.
89. Guelman, L. (2012). Gradient boosting trees for auto insurance loss cost modeling and prediction. Expert Syst. Appl. ***39***, 3659–3667.
90. Liang, C., Liu, Z., Liu, B., et al. (2019). Uncovering insurance fraud conspiracy with network learning. In Proc. 42nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., pp. 1181–1184.
91. Jurgovsky, J., Granitzer, M., Ziegler, K., et al. (2018). Sequence classification for credit-card fraud detection. Expert Syst. Appl. ***100***, 234–245.
92. Branco, B., Abreu, P., Gomes, A.S., et al. (2020). Interleaved sequence RNNs for fraud detection. In Proc. 26th ACM SIGKDD Int. Conf.. Knowl. Discov. Data Min., pp. 3101–3109.
93. Wang, S., Liu, C., Gao, X., et al. (2017). Session-based fraud detection in online e-commerce transactions using recurrent neural networks. In Mach. Learn. Knowl. Discov. Databases, pp. 241–252.
94. Liu, C., Zhong, Q., Ao, X., et al. (2020). Fraud transactions detection via behavior tree with local intention calibration. In Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 3035–3043.
95. Zhu, Y., Xi, D., Song, B., et al. (2020). Modeling users' behavior sequences with hierarchical explainable network for cross-domain fraud detection. Proc. Web Conf. ***2020***, 928–938.
96. Xi, D., Zhuang, F., Song, B., et al. (2020). Neural hierarchical factorization machines for user's event sequence analysis. In Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., pp. 1893–1896.
97. Xi, D., Song, B., Zhuang, F., et al. (2021). Modeling the field value variations and field interactions simultaneously for fraud detection. Proc. AAAI Conf. Artif. Intell. ***35***, 14957–14965.
98. Quinlan, J.R. (1990). Learning logical definitions from relations. Mach. Learn. ***5***, 239–266.
99. Cohen, W.W. (1995). Fast effective rule induction. In Machine Learning Proceedings 1995, A. Prieditis and S. Russell, eds. (Morgan Kaufmann), pp. 115–123.
100. Li, Z., Xiong, H., Liu, Y., and Zhou, A. (2010). Detecting blackhole and volcano patterns in directed networks. In 2010 IEEE Int. Conf. Data Min., pp. 294–303.
101. Dou, Y., Liu, Z., Sun, L., et al. (2020). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In Proc. 29th ACM Int. Conf. Inf. Knowl. Manag., pp. 315–324.
102. Zhang, Y., Fan, Y., Ye, Y., et al. (2019). Key player identification in underground forums over attributed heterogeneous information network embedding framework. Proc. 28th ACM Int. Conf. Inf. Knowl. Manag. 549–558.
103. Liu, Z., Dou, Y., Yu, P.S., et al. (2020). Alleviating the inconsistency problem of applying graph neural network to fraud detection. In Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., pp. 1569–1572.
104. Fan, Y., Ye, Y., Peng, Q., et al. (2020). Metagraph aggregated heterogeneous graph neural network for illicit traded product identification in underground market. In 2020 IEEE Int. Conf. Data Min., pp. 132–141.
105. Ryman-Tubb, N.F., Krause, P., and Garn, W. (2018). How Artificial Intelligence and machine learning research impacts payment card fraud detection: a survey and industry benchmark. Eng. Appl. Artif. Intell. ***76***, 130–157.
106. Dazeley, R.P. (2006). To the Knowledge Frontier and beyond: A Hybrid System for Incremental Contextual-Learning and Prudence Analysis, PhD thesis (University of Tasmania).
107. Cao, S., Yang, X., Chen, C., et al. (2019). TitAnt: online real-time transaction fraud detection in ant financial. In Proc. VLDB Endow., *12*, pp. 2082–2093.
108. Jin, Y., Rejesus*, R.M., and Little, B.B. (2005). Binary choice models for rare events data: a crop insurance fraud application. Appl. Econ. ***37***, 841–848.
109. Chen, R.C., Chiu, M.L., Huang, Y.L., and Chen, L.T. (2004). Detecting credit card fraud by using questionnaire-responded transaction model based on support vector machines. Intell. Data Eng. Autom. Learn. ***3177***, 800–806.
110. Chen, R.C., Luo, S.T., Liang, X., and Lee, V.C. (2005). Personalized approach based on SVM and ANN for detecting credit card fraud. In 2005 Int. Conf. Neural Netw. Brain, *2*, pp. 810–815.
111. Hosmer, D.W., Jr., Lemeshow, S., and Sturdivant, R.X. (2013). Applied Logistic Regression (John Wiley and Sons).
112. Zhou, X., Cheng, S., Zhu, M., et al. (2018). A state of the art survey of data mining-based fraud detection and credit scoring. MATEC Web Conf. ***189***, 03002.
113. Cortes, C., and Vapnik, V. (1995). Support-vector networks. Mach. Learn. ***20***, 273–297.
114. Quinlan, J.R. (1986). Induction of decision trees. Mach. Learn. ***1***, 81–106.
115. Breiman, L. (2001). Random forests. Mach. Learn. ***45***, 5–32.
116. Chen, T., and Guestrin, C. (2016). XGBoost: a scalable tree boosting system. In Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 785–794.
117. Ke, G., Meng, Q., Finley, T., et al. (2017). LightGBM: a highly efficient gradient boosting decision tree. In Proc. 31st Int. Conf. Neural Inf. Process. Syst., pp. 3149–3157.
118. Caruana, R., and Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In Proc. 23rd Int. Conf. Mach. Learn., pp. 161–168.
119. Caruana, R., Karampatziakis, N., and Yessenalina, A. (2008). An empirical evaluation of supervised learning in high dimensions. In Proc. 25th Int. Conf. Mach. Learn., pp. 96–103.
120. Khoshgoftaar, T.M., Golawala, M., and Van Hulse, J. (2007). An empirical study of learning from imbalanced data using random forest. In 19th IEEE Int. Conf. Tools Artif. Intell., *2*, pp. 310–317.

121. Kurshan, E., Shen, H., and Yu, H. (2020). Financial crime and fraud detection using graph computing: application considerations and outlook. In 2020 2nd Int. Conf. Transdiscip. AI, pp. 125–130.
122. Pourhabibi, T., Ong, K.-L., Kam, B.H., and Boo, Y.L. (2020). Fraud detection: a systematic literature review of graph-based anomaly detection approaches. Decis. Support Syst. ***133***, 113303.
123. Goyal, P., and Ferrara, E. (2018). Graph embedding techniques, applications, and performance: a survey. Knowl. Based Syst. ***151***, 78–94.
124. Cui, P., Wang, X., Pei, J., and Zhu, W. (2019). A survey on network embedding. IEEE Trans. Knowl. Data Eng. ***31***, 833–852.
125. Irofti, P., Patrascu, A., and Baltoiu, A. (2021). Quick survey of graph-based fraud detection methods. arXiv:1910.11299, 1–13. https://arxiv.org/abs/1910.11299.
126. Hogan, A., Blomqvist, E., Cochez, M., et al. (2021). Knowledge graphs. ACM Comput. Surv. ***54***, 1–37.
127. Qiu, J., Tang, J., Ma, H., et al. (2018). DeepInf: social influence prediction with deep learning. In Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 2110–2119.
128. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature ***521***, 436–444.
129. Schmidhuber, J. (2015). Deep learning in neural networks: an overview. Neural Netw. ***61***, 85–117.
130. Patidar, R., and Sharma, L. (2011). Credit card fraud detection using neural network. Int. J. Soft Comput. Eng. ***1***, 32–38.
131. G'omez, J.A., Ar'evalo, J., Paredes, R., and Nin, J. (2018). End-to-end neural network architecture for fraud scoring in card payments. Pattern Recognit. Lett. ***105***, 175–181.
132. Sohony, I., Pratap, R., and Nambiar, U. (2018). Ensemble learning for credit card fraud detection. In Proc. ACM India Jt. Int. Conf. Data Sci. Manag. Data, pp. 289–294.
133. Rushin, G., Stancil, C., Sun, M., et al. (2017). Horse race analysis in credit card fraud-deep learning, logistic regression, and gradient boosted tree. In 2017 Syst. Inf. Eng. Des. Symp., pp. 117–121.
134. Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning representations by back-propagating errors. Nature ***323***, 533–536.
135. Elman, J.L. (1990). Finding structure in time. Cogn. Sci. ***14***, 179–211.
136. Shan, G., Wang, H., Liang, W., and Chen, K. (2020). Robust encoder-decoder learning framework for offline handwritten mathematical expression recognition based on a multi-scale deep neural network. Sci. China Inf. Sci. ***64***, 139101.
137. Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neural Comput. ***9***, 1735–1780.
138. Cho, K., van Merriënboer, B., Gulcehre, C., et al. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proc. 2014 Conf. Empir. Methods Nat. Lang. Process., pp. 1724–1734.
139. Pascanu, R., Mikolov, T., and Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In Proc. 30th Int. Conf. Mach. Learn., *28*, pp. 1310–1318.
140. Roy, A., Sun, J., Mahoney, R., et al. (2018). Deep learning detecting fraud in credit card transactions. In 2018 Syst. Inf. Eng. Des. Symp., pp. 129–134.
141. Rendle, S. (2010). Factorization machines. In Proc. 2010 IEEE Int. Conf. Data Min., pp. 995–1000.
142. Wu, Z., Pan, S., Chen, F., et al. (2021). A comprehensive survey on graph neural networks. IEEE Trans. Neural Netw. Learn. Syst. ***32***, 4–24.
143. Scarselli, F., Gori, M., Tsoi, A.C., et al. (2009). The graph neural network model. IEEE Trans. Neural Netw. ***20***, 61–80.
144. Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2019). How powerful are graph neural networks? In Int. Conf. Learn. Represent., pp. 1–17.
145. Liu, Z., Chen, C., Yang, X., et al. (2018). Heterogeneous graph neural networks for malicious account detection. In Proc. 27th ACM Int. Conf. Inf. Knowl. Manag., pp. 2077–2085.
146. Li, A., Qin, Z., Liu, R., et al. (2019). Spam review detection with graph convolutional networks. In Proc. 28th ACM Int. Conf. Inf. Knowl. Manag., pp. 2703–2711.
147. Nguyen, V.H., Sugiyama, K., Nakov, P., and Kan, M.Y. (2020). Fang: leveraging social context for fake news detection using graph representation. In Proc. 29th ACM Int. Conf. Inf. Knowl. Manag., pp. 1165–1174.
148. Cui, L., Seo, H., Tabar, M., et al. (2020). Deterrent: knowledge guided graph attention network for detecting healthcare misinformation. In Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 492–502.
149. Sun, Y., Han, J., Yan, X., et al. (2011). Pathsim: meta path-based top-k similarity search in heterogeneous information networks. Proc. VLDB Endow. ***4***, 992–1003.
150. Zhan, Q., and Yin, H. (2018). A loan application fraud detection method based on knowledge graph and neural network. In Proc. 2nd Int. Conf. Innov. Artif. Intell., pp. 111–115.
151. Wang, Z., Guo, M., Li, Z., et al. (2020). Knowledge graph construction for payment data risk control. Proc. Innov. Comput. ***2020***, 1901–1907.
152. Li, Z., Ding, X., and Liu, T. (2018). Constructing narrative event evolutionary graph for script event prediction. In Proc. 27th Int. Joint Conf. Artif. Intell., pp. 4201–4207.
153. Paulheim, H. (2017). Automatic knowledge graph refinement: a survey of approaches and evaluation methods. Semant. Web ***8***, 489–508.
154. Guo, M., Liu, Y., Li, J., et al. (2014). A knowledge based approach for tackling mislabeled multi-class big social data. In The Semantic Web: Trends and Challenges, V. Presutti, et al., eds. (Springer International Publishing), pp. 349–363.
155. Xiong, W., Hoang, T., and Wang, W.Y. (2017). DeepPath: a reinforcement learning method for knowledge graph reasoning. In Proc. 2017 Conf. Empir. Method. Nat. Lang. Process., pp. 564–573.
156. Wang, Z., Chen, T., Ren, J., et al. (2018). Deep reasoning with knowledge graph for social relationship understanding. In Proc. 27th Int. Joint Conf. Artif. Intell., pp. 1021–1028.
157. Yang, Q., Liu, Y., Chen, T., et al. (2019). Federated machine learning: concept and applications. ACM T. Intel. Syst. Technol. ***10***, 1–19.
158. Konecny, J., McMahan, H.B., Yu, F.X., et al. (2016). Federated learning: Strategies for improving communication efficiency. arXiv:1610.05492, 1–10. https://arxiv.org/abs/1610.05492.
159. Kairouz, P., McMahan, H.B., Avent, B., et al. (2019). Advances and open problems in federated learning. arXiv:1912.04977, 1–121. https://arxiv.org/abs/1912.04977.
160. Pavlo, A., Paulson, E., Rasin, A., et al. (2009). A comparison of approaches to large-scale data analysis. In Proc. 2009 ACM SIGMOD Int. Conf. Manag. Data, pp. 165–178.
161. Perez, L., and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv:1712.04621, 1–8. https://arxiv.org/abs/1712.04621.
162. Chawla, N.V., Bowyer, K.W., Hall, L.O., et al. (2002). SMOTE: synthetic minority oversampling technique. J. Artif. Intell. Res. ***16***, 321–357.
163. Liu, X.Y., Wu, J., and Zhou, Z.H. (2008). Exploratory undersampling for class-imbalance learning. IEEE Trans. Syst. Man, Cybern. B (Cybern.) ***39***, 539–550.
164. Peng, M., Zhang, Q., Xing, X., et al. (2019). Trainable undersampling for class-imbalance learning. In Proc. AAAI Conf. Artif. Intell., *33*, pp. 4707–4714.
165. Cao, K., Wei, C., Gaidon, A., et al. (2019). Learning imbalanced datasets with label-distribution-aware margin loss. In Proc. 33rd Int. Conf. Neural Inform. Process. Syst., pp. 1567–1578.
166. Li, X., Sun, X., Meng, Y., et al. (2019). Dice loss for data-imbalanced NLP tasks. In Proc. 58th Annu. Meet. Assoc. Comput. Linguist., pp. 465–476.
167. Ren, M., Zeng, W., Yang, B., et al. (2018). Learning to reweight examples for robust deep learning. In Proc. 35th Int. Conf. Mach. Learn., pp. 4334–4343.
168. Shu, J., Xie, Q., Yi, L., et al. (2019). Meta-weight-net: learning an explicit mapping for sample weighting. In Proc. 33rd Int. Conf. Neural Inform. Process. Syst., pp. 1919–1930.
169. Yin, X., Yu, X., Sohn, K., et al. (2019). Feature transfer learning for face recognition with under-represented data. In Proc. IEEE/CVF Conf. Computer Vision Pattern Recogn.(CVPR), L. O'Conner, ed. (IEEE), pp. 5704–5713.
170. Shi, M., Tang, Y., Zhu, X., et al. (2020). Multi-class imbalanced graph convolutional network learning. In Proc. 29th Int. Joint Conf. Artif. Intel., pp. 2879–2885.
171. Buolamwini, J., and Gebru, T. (2018). Gender shades: intersectional accuracy disparities in commercial gender classification. In Proc. 1st Conf. Fairness Account. Transp., pp. 77–91.
172. Zhao, J., Wang, T., Yatskar, M., and Vicente. (2017). Men also like shopping: reducing gender bias amplification using Corpus-level constraints. In Proc. 2017 Conf. Empir. Method. Nat. Lang. Process., pp. 2979–2989.
173. Pan, F., Ao, X., Tang, P., et al. (2020). Field-aware calibration: a simple and empirically strong method for reliable probabilistic predictions. Proc. Web Conf. ***2020***, 729–739.
174. Goodfellow, I.J., Shlens, J., and Szegedy, C. (2015). Explaining and harnessing adversarial examples. Int. Conf. Learn. Represent. 1–11.
175. Carlini, N., and Wagner, D. (2017). Towards evaluating the robustness of neural networks. In 2017 IEEE Symposium Secur. Priv., pp. 39–57.
176. Madry, A., Makelov, A., Schmidt, L., et al. (2017). Towards deep learning models resistant to adversarial attacks. Int. Conf. Learn. Represent. 1–23.
177. Zhu, D., Zhang, Z., Cui, P., et al. (2019). Robust graph convolutional networks against adversarial attacks. In Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 1399–1407.
178. Tang, X., Li, Y., Sun, Y., et al. (2020). Transferring robustness for graph neural network against poisoning attacks. In Proc. 13th Int. Conf. Web Search Data Min., pp. 600–608.
179. Telgarsky, M. (2016). Benefits of depth in neural networks. In Proc. 29th Annu. Conf. Learn. Theor., pp. 1517–1539.
180. Eldan, R., and Shamir, O. (2016). The power of depth for feedforward Nneural networks. In Proc. 29th Annu. Conf. Learn. Theor., pp. 907–940.
181. Camburu, O.M. (2020). Explaining deep neural networks. PhD thesis (University of Oxford).
182. Adebayo, J., Gilmer, J., Muelly, M., et al. (2018). Sanity checks for saliency maps. In Proc. 32nd Int. Conf. Neural Inform. Process. Syst., pp. 9525–9536.

183. Kindermans, P.J., Hooker, S., Adebayo, J., et al. (2019). The (un) reliability of saliency methods. In Explainable AI: Interpreting, Explaining and Visualizing Deep Learning (Springer), pp. 267–280.
184. Jain, S., and Wallace, B.C. (2019). Attention is not explanation. In Proc. 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol., pp. 3543–3556.
185. Wiegreffe, S., and Pinter, Y. (2019). Attention is not not explanation. In Proc. 2019 Conf. Empir. Methods Nat. Lang. 9th Int. Joint Conf. Nat. Lang. Process. (Emnlp-ijcnlp), pp. 11–20.
186. Ying, R., Bourgeois, D., You, J., et al. (2019). Gnnexplainer: generating explanations for graph neural networks. In Proc. Adv. Neural Inf. Process. Syst., pp. 9240–9251.
187. Yuan, H., Tang, J., Hu, X., et al. (2020). Xgnn: towards model-level explanations of graph neural networks. In Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 430–438.
188. Luo, D., Cheng, W., Xu, D., et al. (2020). Parameterized explainer for graph neural network. In Proc. 34th Conf. Neural Inform. Process. Syst., pp. 19620–19631.

ACKNOWLEDGMENTS

This work is supported by grants from the National Key Research and Development Program of China (no. 2017YFB1002104), the National Natural Science Foundation of China (no. 92046023, 92046003, 71971207, 71601178), the Youth Innovation Promotion Association of Chinese Academy of Sciences, and Beijing Nova Program (no. Z201100006820062).

AUTHOR CONTRIBUTIONS

X.Z. and X.A. conceived, organized, and revised the manuscript. Z.Q. and Y.C. wrote the draft manuscript. Y.L. helped with the writing of the manuscript. Q.H. and J.L. supervised and instructed on the manuscript. All authors discussed and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

LEAD CONTACT WEBSITE

Xiaoqian Zhu: http://www.casisd.cn/sourcedb_ipm/zw/zjrc/201801/t20180116_4933514.html
Xiang Ao: https://aoxaustin.github.io/
Qing He: http://people.ucas.ac.cn/~0000964.
Jianping Li: http://people.ucas.ac.cn/http://people.ucas.ac.cn/~lijianping