Journal of Risk Analysis and Crisis Response, 2026, 16(1), 87-97
https://jracr.com/
ISSN Print: 2210-8491
ISSN Online: 2210-8505

Article

Machine Learning Approaches for Detecting Irregular Financial Activities in China

Zhong-Qiang Zhou [1,2], Fei Zhang [1,2] and Ling Li [1,2,*]

[1] School of Applied Economics, Guizhou University of Finance and Economics, Guiyang (550025), Guizhou, China
[2] Digital Economy Institute, Guizhou University of Finance and Economics, Guiyang (550025), Guizhou, China
* Correspondence: 1587812159@qq.com; Tel.: +86-15971332962

Received: January 26, 2026; Received in revised form: March 28, 2026; Accepted: March 29, 2026; Available online: March 31, 2026

Abstract: Irregular financial activities (IFAs) pose serious challenges to regulators, especially in China where high-profile scams have highlighted gaps in oversight. This study develops a machine learning framework to identify such risks using a dataset of 540 financial cases from 2014 to 2024. Activities are classified as irregular or normal, and the performance of 18 algorithms—including traditional machine learning, ensemble methods, and deep learning models—is compared. Ensemble learning models demonstrate superior performance in detecting IFAs, balancing high accuracy with practical applicability. In particular, Bagging and LightGBM achieve the highest accuracy and robust F1-scores among all tested methods. These findings offer novel insights and technical tools for early warning of IFAs, contributing to the literature on financial risk detection and informing policy design. This comparison is among the first systematic evaluations of diverse machine learning algorithms for IFA detection in China, bridging a gap in the literature on regulatory technology and risk management. The proposed approach provides regulators with real-time, data-driven tools to identify irregularities before substantial losses occur.

Keywords: Irregular Financial Activities; Financial Risk Detection; Machine Learning; RegTech

1. Introduction

The rapid development of internet-based platforms and fintech applications over the past decade has significantly reduced transaction costs and expanded financial access. However, these innovations have also given rise to novel forms of financial misconduct, posing substantial challenges for traditional regulatory frameworks. Prominent examples include the 2014 Ezubao Ponzi scheme, which defrauded approximately RMB 50 billion from hundreds of thousands of investors, and the 2023 Quan Xiang Tong EV charger fraud in Hainan, involving around RMB 1 billion and affecting tens of thousands of victims. Beyond financial losses, such scandals undermine market integrity and highlight vulnerabilities in China's regulatory environment, emphasizing the urgent need for improved risk identification systems.

Recognizing these threats, China has prioritized financial risk management as a national strategic goal. The 20th Party Congress explicitly calls for comprehensive regulation and robust measures to guard against systemic financial crises. Nevertheless, traditional post-event regulatory tools are increasingly ineffective against the complexity, stealth, and cross-regional nature of irregular financial activities (IFAs). With digital innovations emerging faster than regulatory updates, many transactions remain outside effective supervision. Therefore, regulators urgently require real-time, data-driven monitoring solutions—an approach firmly embedded within regulatory technology (RegTech).

Although literature on financial risk management is extensive, previous studies predominantly focus on credit risks and market volatility. Systematic analyses of IFA risks, particularly in the context of China, are relatively scarce. Existing research often lacks unified quantitative frameworks and is insufficiently proactive in addressing regulators' need for timely alerts. Our study addresses this gap by developing a machine learning-based system capable of identifying IFAs proactively, before substantial losses occur.

Our research proceeds in two primary stages. Firstly, we construct a comprehensive feature set capturing diverse dimensions such as behavioral patterns, participant characteristics, funding dynamics, and regional contexts of IFAs. Secondly, we apply eighteen different algorithms spanning classical machine learning, ensemble learning, and deep learning methodologies to a unique dataset of 540 Chinese financial cases from 2014 to 2024. By directly comparing the predictive effectiveness of these models, we identify the most suitable methods for monitoring IFAs.

This study offers two significant contributions. First, we present the first comprehensive machine learning toolkit designed specifically for identifying IFA risks in China and demonstrate the superior performance of ensemble learning methods. Second, we empirically confirm that RegTech significantly enhances early warning capabilities, providing regulators with actionable insights to prevent financial misconduct.

The remainder of the paper is structured as follows: Section 2 reviews existing literature and theoretical foundations. Section 3 outlines data collection, feature engineering, and model specifications. Section 4 reports the empirical results, and Section 5 concludes.

2. Literature Review

The concept of IFAs has evolved significantly in recent years. Initially, IFAs were defined narrowly as explicit violations of securities laws, such as illegal fundraising and insider trading [1]. This definition has since broadened to encompass a wider spectrum of financial behaviors characterized by structural complexity, opacity, or incentive mechanisms that diverge markedly from established regulatory norms. Under this expanded perspective, IFAs now include practices such as market manipulation, fraudulent financial disclosures, and novel high-risk schemes like crowdfunding and crypto-asset speculation [2-4].

In China, illegal fundraising remains the most prevalent form of IFAs. Such schemes often employ Ponzi-like structures, promising unsustainable returns and ultimately inflicting severe losses on retail investors when they collapse [5-7]. Empirical evidence suggests that analogous schemes have been documented globally, with significant welfare implications for local economies [8-9].

Multiple factors drive the persistence and proliferation of IFAs, each compounding the regulatory challenges involved. These factors include information asymmetry, regulatory lag, behavioral biases, and credit market constraints. First, information asymmetry severely undermines investors' ability to distinguish legitimate financial innovations from fraudulent schemes, a problem that intensifies during economic downturns as investors chase unusually high returns in opaque channels [10-11]. Second, regulatory lag further complicates effective oversight: fintech platforms can rapidly scale nationwide, often outpacing the capacity of traditional statutory and supervisory frameworks [12-13]. In line with this, cross-country evidence indicates that jurisdictions with stronger regulatory enforcement tend to experience fewer IFAs [14]. Third, behavioral biases—including over-optimism and the so-called "greater fool" mentality—encourage investors to participate in dubious financial schemes despite the evident risks [15]. Finally, credit exclusion pushes smaller enterprises and households toward IFAs: small and medium-sized enterprises (SMEs) and individuals without sufficient collateral or credit history may resort to alternative financing channels when traditional credit is inaccessible [16-18]. Together, these factors create a self-reinforcing cycle that amplifies the supervisory challenges posed by IFAs.

Consequently, the socioeconomic costs of IFAs are evident at both individ1q aual and systemic levels. At the individual level, retail investors often suffer substantial financial losses and a concomitant erosion of trust in formal financial institutions [19-20]. At the aggregate level, IFAs distort market price discovery and amplify financial volatility, thereby elevating systemic risk [21-22]. Empirical evidence further suggests that major fraud incidents can have enduring adverse effects: Gurun et al. [3] document long-term declines in local market participation following prominent frauds, and others highlight the role of IFAs in precipitating broader financial crises through contagion effects [23-24]. These findings underscore the systemic importance of addressing IFAs.

Given these consequences, policymakers have pursued initiatives on three main fronts. In particular, legal reforms have imposed stricter penalties for financial fraud, though their efficacy depends critically on robust enforcement [25-26]. Concurrently, investor education programs aim to improve financial literacy and thereby reduce individual investors' susceptibility to IFAs [27]. Lastly, advances in regulatory technology (RegTech) offer data-driven tools to narrow the supervisory gap. For instance, the People's Bank of China has implemented real-time monitoring systems powered by machine-learning algorithms to flag anomalous transactions [28].

In parallel, academic research has increasingly aligned with these regulatory developments, evolving from rudimentary early-warning indices based on macro-financial indicators [29-30] to a systematic multi-dimensional quantitative risk monitoring model for illegal financial activities [31], and further to sophisticated predictive models employing machine-learning techniques. Examples include support vector machines [24,32], dynamic Bayesian networks [33], and long short-term memory (LSTM) neural networks [34], all of which aim to forecast financial distress and default with high accuracy.

Despite these advancements, significant gaps remain in the existing literature. Current early-warning models often focus on broad financial stress or isolated market-specific incidents, seldom capturing the nuanced complexity and rapid evolution characteristic of IFAs. In addition, comparative evaluations of different algorithmic approaches using a common dataset are scarce. This paper addresses these deficiencies by systematically benchmarking eighteen machine-learning algorithms—including classical, ensemble, and deep learning models—on a novel dataset of 540 financial cases from China, thereby providing critical guidance for effective

RegTech integration and proactive supervision of IFAs.

3. Research Design

3.1. Data and Pre-processing

3.1.1. Sample

Our empirical analysis is based on 540 financial activity cases in China between January 2014 and August 2024. Of these, 324 cases involve irregular financial activities—including 69 highly representative flagship cases—while the remaining 216 cases reflect normal financial behavior and serve as the control group. This composition enables the models to simultaneously learn the salient characteristics of both positive (irregular) and negative (normal) observations, thereby enhancing classification accuracy. All cases were obtained from authoritative public sources, including China Judgments Online, Tonghuashun Financial News, corporate websites, and the WIND database, ensuring the authenticity, reliability, and representativeness of the dataset.

3.1.2. Data Cleaning and Encoding

After data collection, the raw dataset underwent a four-step cleansing pipeline. First, any fields irrelevant to risk prediction (e.g., case identifiers, investor testimonials, and promotional materials) were removed to reduce noise. Second, missing values were imputed: numerical attributes were filled using either the mean or median, while categorical attributes were filled using the mode (most frequent value) or a dedicated 'missing' category flag. Third, all numerical variables were standardized to have zero mean and unit variance, thereby eliminating scale heterogeneity that could otherwise disrupt training stability. Fourth, categorical variables were transformed via one-hot encoding; for example, dummy variables were created for regions such as Beijing, Shanghai, Guangzhou, and other major cities. Together, these steps yield a cleaned and structured dataset that provides a robust foundation for subsequent modeling.

3.1.3. Dataset Partitioning and Class-Imbalance Treatment

To obtain unbiased estimates of out-of-sample performance, the dataset was split into a training set (60%), a validation set (20%), and a test set (20%). The validation set was used for hyper-parameter tuning and early stopping, whereas the test set was reserved for final evaluation. Because IFAs are less numerous than normal cases, the training set exhibits a class imbalance; directly fitting a classifier to such imbalanced data can cause the algorithm to favor the majority class while overlooking the minority [34]. To address this imbalance, we applied the Synthetic Minority Over-sampling Technique (SMOTE) to the training set, synthetically generating additional minority-class examples in feature space. Compared to naive duplication, SMOTE introduces greater diversity, mitigates overfitting, and improves recall for the IFA class. After resampling, the training set attained approximate parity between the positive and negative classes.

3.2. Feature Engineering

To capture the multifaceted drivers of IFAs, we construct features in six dimensions: liquidity risk, financial soundness, credit risk, regional risk background, trading behavior, and violation type. Specifically, liquidity risk is measured by the delay (in days) between withdrawal requests and receipt of funds; prolonged delays indicate funding stress. Financial soundness is captured by a binary indicator that flags whether key financial ratios (e.g., leverage, current ratio) deviate from normal ranges, potentially signaling deterioration in stability. Credit risk is indicated by the presence of prior defaults or misconduct records involving key insiders (e.g., executives, owners), reflecting elevated credit or reputational risk. Regional risk background is measured by the count of past irregular cases in the same region, capturing latent systemic risk. Trading behavior is summarized by metrics such as the frequency and volatility of large-value transactions, since abnormal spikes often precede market stress. Finally, violation type is captured by categorical flags indicating the specific nature of irregular behavior (e.g., Ponzi scheme, cryptocurrency speculation, illicit lending). Collectively, these features cover dimensions of liquidity, solvency, reputation, geography, transaction dynamics, and violation category, yielding a rich information set for our machine-learning classifiers. Table 1 summarizes the definitions and interpretations of each feature category.

Table 1. Feature categories and economic interpretations.

Feature category	Economic interpretation
Liquidity risk	Measures the efficiency of fund withdrawals; prolonged delays indicate funding stress and elevated liquidity risk.
Financial soundness	Flags whether key financial ratios (e.g., leverage, current ratio) deviate from benchmarks, signaling potential deterioration in financial stability.
Credit risk	Flags prior defaults or misconduct by key insiders (e.g., executives, owners), indicating elevated credit risk or reputational concern.
Regional risk background	Counts historical irregular cases in the same region, capturing latent systemic or contagion risk in that area.
Trading behavior feature	Captures trading dynamics (e.g., frequency and volatility of large transactions) to detect abnormal patterns that often precede market stress.
Behavior type label	Categorical flags indicating specific irregular behavior (e.g., Ponzi schemes, crypto speculation, illicit lending), explicitly capturing the violation type.

3.3. Model Selection

To compare alternative modeling approaches, we considered eighteen algorithms drawn from three broad families: classical machine-learning models (e.g., logistic regression, support vector machines, k-nearest neighbors), ensemble-learning methods (random forests, AdaBoost, bagging, gradient boosting, LightGBM), and deep-learning architectures (multilayer perceptrons, recurrent networks such as LSTM).

Classical models offer a simple, transparent structure that is easy to interpret but often fail to capture complex nonlinear relationships. Ensemble methods typically achieve higher predictive accuracy and robustness by aggregating multiple base learners; however, this improved performance comes at the cost of reduced interpretability. Deep neural networks excel at learning hierarchical representations from high-dimensional or sequential data, enabling outstanding performance on large complex datasets; however, they require very large training samples and careful regularization. Table 2 summarizes the key advantages and principal limitations of each model category.

Table 2. Model Categories: Key Advantages and Principal Limitations.

Model category	Key advantages	Principal limitations
Classical machine-learning models (e.g., logistic regression, SVM, kNN)	Simple, well-established structure; training and tuning are relatively straightforward. Highly interpretable, making decision logic easy to understand. Deliver stable performance on small- and medium-sized datasets.	Limited capacity to capture complex nonlinear relationships; performance may degrade in high-dimensional or highly nonlinear settings. Accuracy improvements typically plateau as data volume and feature complexity grow.
Ensemble-learning models (e.g., Random Forest, AdaBoost, Bagging, Gradient Boosting, LightGBM)	Combine multiple base learners to boost predictive accuracy and robustness by reducing variance. Less sensitive to noise and outliers, yielding strong overall resilience. Often excel on complex tasks.	More complex model structure makes individual decision paths difficult to interpret (reduced transparency). High computational cost for training and inference; resource requirements grow with data size and model complexity.
Deep-learning models (e.g., multilayer perceptron, RNN/LSTM)	Deep, multi-layer architectures learn hierarchical representations of high-dimensional inputs, automatically extracting complex features. Achieve outstanding performance on large datasets; can capture long-range temporal dependencies in sequential data.	Require large amounts of labeled data and heavy computational resources for training. Operate as "black boxes" with limited interpretability of internal representations. Prone to overfitting when data are limited, and generalization can be unstable without careful regularization.

3.4. Training Strategy and Hyper-parameter Optimization

All models were trained using mini-batch stochastic gradient descent (SGD) on the training set, with early stopping applied based on validation performance (i.e. stopping training once validation loss failed to improve). For deep neural networks, we employed the Adam optimizer, an efficient adaptive method well-suited to large and noisy datasets. Hyperparameters of the classical and ensemble models were tuned via exhaustive grid search within a five-fold cross-validation framework.

Key hyperparameters included the kernel type and regularization coefficient for SVM; the number and maximum depth of trees for random forests; and the learning rate, maximum depth, and number of leaves for LightGBM. For the LSTM models, we varied the number of layers, the number of hidden units per layer, and the dropout rate to balance model complexity and generalization.

Training proceeded until the validation loss plateaued or began to deteriorate, at which point the final models were evaluated on the held-out test set. Performance was assessed using standard classification metrics: **Accuracy** (overall correct classification rate), **Precision** (true positives as a fraction of positive predictions), **Recall** (true positives as a fraction of actual IFAs), and the **F1 score** (the harmonic mean of precision and recall). This iterative training and tuning procedure yielded a well-calibrated set of classifiers for the empirical analysis.

4. Results and Analysis

Using the full sample of 540 cases, we evaluated 18 models for the task of IFA detection. Performance was measured by four standard metrics. Table 3 (and Figure 1) shows pronounced heterogeneity in model performance. Ensemble methods consistently dominate all metrics. On average, ensembles reached roughly 67% accuracy (e.g. Bagging: 69.06%, Voting: 67.56%), whereas traditional classifiers averaged about 59% and deep networks about 53%. The F1 scores show a similar gap: ensembles average ~66%, traditional models ~58%, and deep models ~45%. The disparity is most stark in recall, the costliest error: the best ensemble (LightGBM) recalls 61.31% of IFAs, versus deep models like DNN or RNN at only 27.77%. The hierarchy is unambiguous—ensemble learners top every metric, traditional models occupy the middle ground, and deep neural networks lag at the bottom with especially low recall.

Table 3. Out-of-Sample Classification Performance of 18 Machine-Learning Models.

Model		Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Traditional Models	Support Vector Machine (SVM)	62.43	60.42	65.91	63.04
	Naive Bayes (NB)	48.07	45.31	32.95	38.16
	Decision Tree (DT)	59.67	57.89	62.50	60.11
	K-Nearest Neighbor (kNN)	50.28	49.17	67.05	56.73
	Conditional Random Field (CRF)	66.11	69.58	61.11	66.08
	Random Forest (RF)	66.85	64.58	70.45	67.39
Ensemble Models	Stacking Ensemble	66.85	65.91	65.91	65.91
	Voting Ensemble	67.96	66.67	68.18	67.42
	Extra Trees	68.51	66.32	71.59	68.85
	AdaBoost	64.64	65.00	59.09	61.90
	Bagging Ensemble	69.06	68.60	67.05	67.82
	LightGBM	69.06	71.05	61.36	65.85
Deep-Learning Models	Deep Neural Network (DNN)	52.49	52.17	27.27	35.82
	Bidirectional LSTM (BiLSTM)	46.96	45.24	43.18	44.19
	Recurrent Neural Network (RNN)	53.04	51.95	45.45	48.48
	RNN + Dropout	52.49	51.14	51.14	51.14
	LSTM + Dropout + L2	54.14	53.42	44.32	48.45
	Multilayer Perceptron (MLP)	56.91	60.42	32.95	42.65

The superiority of ensembles derives from combining multiple learners to stabilize predictions. In our experiments, the Bagging ensemble and gradient-boosted LightGBM attained the highest accuracy (69.06% each) and strong F1 scores (Bagging F1 = 67.82%, LightGBM F1 = 65.86%). Other ensemble strategies also outperform single models: for instance, Voting yields 67.96% accuracy and 67.42% F1, whereas a single SVM achieves only 62.43% accuracy and 63.04% F1. Even simpler ensembles like AdaBoost (64.06% accuracy, F1 = 63.81%) improve on their base trees. These results confirm that model fusion substantially boosts detection performance in this complex task. Although ensembles are structurally more complex, their markedly higher hit rate on IFAs (e.g. ≥61% recall for ensembles vs. <45% for deep networks) justifies the added complexity.

Several established models—including Random Forests and Conditional Random Fields—also perform strongly, approaching the leading ensemble methods in terms of accuracy and precision. With careful feature engineering and parameter tuning, these traditional approaches remain highly competitive. They are straightforward to train, computationally efficient, and—crucially—highly interpretable. For example, decision-tree variants provide explicit feature-importance rankings, and CRFs model the dependency structure in sequential data, clarifying why a given case is flagged. Such transparency appeals to regulators who must understand and defend any automated warning signals. However, traditional models continue to lag ensemble methods in recall and other key metrics, suggesting limits to their ability to capture the most complex risk patterns.

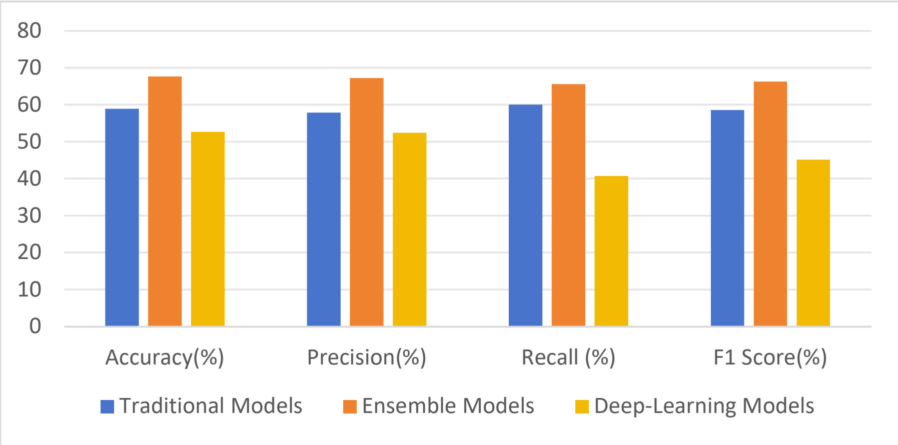

Figure 1. Comparison of performance for traditional, ensemble, and deep-learning models.

Deep neural networks, despite their theoretical power for modeling nonlinear interactions, fail to outperform simpler methods given the current data constraints. Baseline DNN and BiLSTM models achieve accuracies of only around 52% and F1 scores near 36%, far below the ensemble benchmarks. This shortfall likely stems from the need for very large sample sizes and meticulous hyperparameter tuning required for deep networks to fully realize their representational power; with only modest datasets, they are prone to underfitting or overfitting. They also demand long training times and considerable computational resources, reducing their operational appeal. Regularization helps to some extent—dropout improves RNN recall and F1, and dropout combined with L2 regularization slightly improves LSTM performance—but even the best-regularized deep models remain well behind ensemble methods. Moreover, deep networks function as “black boxes,” offering little insight into why a case is labeled as high-risk. This opacity undermines trust in regulatory settings where supervisors must justify each intervention.

Taken together, these performance differences reflect each modeling family’s inherent bias–variance trade-offs and data requirements. Ensemble methods reduce both bias and variance by combining diverse weak learners, yielding the most robust predictions. Traditional models, when fully tuned, achieve respectable accuracy with relatively minimal calibration effort and strong interpretability. Deep networks possess formidable expressive power, but they incur steep costs in terms of data requirements and transparency—costs that are rarely offset in small to medium-sized samples. The resulting interpretability gap further widens the practical divide: regulators tend to favor models (such as decision trees and their ensembles) that can explain their decision logic.

For supervisory applications, recall is paramount: every missed IFA can entail substantial financial and reputational damage. Ensemble classifiers and the best traditional models maintain high recall rates, keeping false negatives to a minimum. In contrast, the recall shortfall of deep networks implies an unacceptably high rate of missed detections. The second critical criterion is transparency: decision-tree-based models provide clear decision rules and feature-importance metrics, aligning with compliance mandates. Accordingly, ensemble tree methods—such as Bagging and LightGBM—emerge as the most pragmatic choices for real-time IFA monitoring, with Random Forests and CRFs serving as interpretable, resource-efficient alternatives. Selecting among these options allows regulators to reliably detect latent misconduct while retaining full visibility into each alert's rationale—a prerequisite for effective and accountable RegTech deployment.

5. Conclusion

This study addresses the surge in irregular financial activities (IFAs) and the challenges of early detection by traditional supervision through a machine-learning–based risk identification framework and comprehensive algorithm benchmarking. Using 540 Chinese cases from 2014 to 2024, models from traditional machine learning, ensemble learning, and deep learning paradigms were trained and evaluated. Ensemble methods yielded the strongest performance, with Bagging and LightGBM standing out in particular. Traditional models lagged slightly in predictive power but offered a favorable balance of accuracy, computational speed, and interpretability, making them attractive for routine regulatory tasks. Deep learning architectures underperformed given the current data limitations, underscoring their need for larger sample sizes and extensive fine-tuning.

These findings have both theoretical and practical implications. Theoretically, this study broadens the scope of financial risk research by focusing on IFAs—an area largely underexplored in China—and provides the first quantitative, multi-algorithm comparison for their detection. Practically, the proposed models equip regulators with an intelligent early-warning tool that complements traditional oversight methods and, in some cases, even surpasses them. The policy implication is clear: supervisory agencies should adopt advanced ensemble algorithms to enhance the timeliness and effectiveness of IFA monitoring.

This study has certain limitations. The sample size is modest, and the findings would benefit from validation on larger, multi-year, cross-border datasets. Future research should broaden the temporal and geographic scope, incorporate novel algorithms, and explore more sophisticated model architectures to further improve detection capabilities. As financial innovation accelerates, regulators must continuously monitor the evolving IFA landscape and regularly update their models and strategies to safeguard systemic stability.

Funding: This research was supported by the Guizhou Province Philosophy and Social Sciences Planning Project, Grant No. 22GZYB15.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] Liu, X., Huang, F., & Yeung, H. (2018). The regulation of illegal fundraising in China. Asia Pacific Law Review, 26(1), 77-100. https://doi.org/10.1080/10192557.2018.1511086

[2] Siering, M., Koch, J. A., & Deokar, A. V. (2016). Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. Journal of Management Information Systems, 33(2), 421-455. https://doi.org/10.1080/07421222.2016.1205930

[3] Gurun, U. G., Stoffman, N., & Yonker, S. E. (2018). Trust Busting: The Effect of Fraud on Investor Behavior. The Review of Financial Studies, 31(4), 1341-1376. https://doi.org/10.1093/rfs/hhx058

[4] Fratric, P., Sileno, G., Klous, S., & van Engers, T. (2022). Manipulation of the Bitcoin market: an agent-based study. Financial Innovation, 8, 60. https://doi.org/10.1186/s40854-022-00364-3

[5] Li, Y. X., Chen, F., & Jin, Y. F. (2014) A Study on the Regulation of the Internet Finance. Journal of Zhejiang University (Humanities and Social Sciences), 44(04), 87-97. https://link.cnki.net/urlid/10.3785/j.issn.1008-942X.2014.02.121

[6] Li, C. S., & Shen, Y. (2019). Identification, Measurement, and Control of Risks in New Financial Business Models under the Digital Economy. Journal of Management World, 35(12), 53-69. https://link.cnki.net/doi/10.19744/j.cnki.11-1235/f.2019.0163

[7] Fang, Y., Shao, Z. Q., Chen, J. H., et al. (2020). Research on the Identification and Monitoring Framework for Illegal Financial Activities. Study and Practice, 9(09), 54-65. https://link.cnki.net/doi/10.19624/j.cnki.cn42-1005/c.2020.09.007

[8] Cortes, D., Santamaria, J., & Vargas, J. F. (2016). Economic Shocks and Crime: Evidence from the Crash of Ponzi Schemes. Journal of Economic Behavior & Organization, 131, 263-275. https://doi.org/10.1016/j.jebo.2016.07.024

[9] Brinster, L., & Tykvova, T. (2021). Connected VCs and Strategic Alliances: Evidence from Biotech Companies. Journal of Corporate Finance, 66, 1-17. https://doi.org/10.1016/j.jcorpfin.2020.101835

[10] Borio, C. (2014). The Financial Cycle and Macroeconomics: What Have We Learnt? Journal of Banking & Finance, 45, 182-198. https://doi.org/10.1016/j.jbankfin.2013.07.031

[11] Gong, Q., & Wang, L. Y. (2018). Financial Inclusion, Risk Reserve, and Investor Protection: A Case Study of P2P Loan Mode. China Economic Quarterly, 17(04), 1581-1598. https://link.cnki.net/doi/10.13821/j.cnki.ceq.2018.03.13

[12] Jiang, J. J, Gao, M., & Lu, R. C. (2020). Risk of Online Lending Marketplace: Macro Factors and Regulation. Journal of Financial Research, (06), 152-170.

[13] Buchak, G., Matvos, G., Piskorski, T., et al. (2018). Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks. Journal of Financial Economics, 130(3), 453-483. https://doi.org/10.1016/j.jfineco.2018.03.011

[14] La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2006). What Works in Securities Laws? The Journal of Finance, 61(1), 1-32. https://doi.org/10.1111/j.1540-6261.2006.00828.x

[15] Bushman, R. M., & Williams, C. D. (2012). Accounting Discretion, Loan Loss Provisioning, and Discipline of Banks' Risk-taking. Journal of Accounting and Economics, 54(1), 1-8. https://doi.org/10.1016/j.jacceco.2012.04.002

[16] Yang, F. L., & Huang, Y. H. (2006). Corporate Governance Structure, Information Asymmetry, and Financing of Small and Medium Enterprises. Journal of Financial Research, (05), 159-166.

[17] Liang, P. H., & Jiang, H. Z. (2020). Financial Access and Internet Financial Risk Alleviation—An Empirical Analysis Based on Internet Pyramid Scheme Judicial Documents. China Industrial Economics, (04), 116-134. https://link.cnki.net/doi/10.19581/j.cnki.ciejournal.2020.04.006

[18] Allen, F., Qian, J., & Qian, M. (2005). Law, Finance, and Economic Growth in China. Journal of Financial Economics, 77(1), 57-116. https://doi.org/10.1016/j.jfineco.2004.06.010
[19] Hu, J. X., & Wu, Y. M. (2013). The Development and Regulation of Crowdfunding Financing. Securities Market Herald, (12), 60-65.
[20] Zhang, T. W., Wang, W., & Yu, C. T. (2016). Financial Knowledge, Risk Awareness and Investment Income—Based on the Small Investors' Rights and Interests Protection Questionnaire. Accounting Research, (07), 66-73+97.
[21] Xu, L. B., Li, C., & Chen, Q. W. (2021) Information-based Market Manipulation and Wealth Transfer Effect: Evidence from Insiders' Shares Selling. Journal of Finance and Economics, 47(05), 4-18+48. https://link.cnki.net/doi/10.16538/j.cnki.jfe.20210313.101
[22] Gennaioli, N., Shleifer, A., & Vishny, R. (2012). Neglected Risks, Financial Innovation, and Financial Fragility. Journal of Financial Economics, 104(3), 452-468. https://doi.org/10.1016/j.jfineco.2011.05.005
[23] Lang, W. W., & Jagtiani, J. A. (2010). The Mortgage and Financial Crises: The Role of Credit Risk Management and Corporate Governance. Atlantic Economic Journal, 38(2), 123-144. https://doi.org/10.1007/s11293-010-9221-7
[24] Lin, L., Cao, Y., & Xiao, H. (2016). The Fragility of Financial System under the Growing Shadow Banking System in China. China Economic Quarterly, 15(03), 1113-1136. https://link.cnki.net/doi/10.13821/j.cnki.ceq.2016.02.12
[25] Rose, A. M. (2010). The Multi-Enforcer Approach To Securities Fraud Deterrence: A Critical Analysis, University of Pennsylvania Law Review, 158, 2173-2231. https://doi.org/10.2139/ssrn.1504158
[26] Bai, J. (2017). Risks and Legal Regulation of Equity Crowdfunding in China. Oriental Law, (01), 14-28. https://link.cnki.net/doi/10.19404/j.cnki.dffx.2017.01.002
[27] Liu, G. Q. (2018). A Study of Financial Literacy in China: Based on the 2017 Financial Literacy Survey. Journal of Financial Research, (03), 1-20.
[28] Zhao, J. Q., Jiang, H. Y., Hu, S. Y., et al. (2023). Fintech, Small Enterprise Credit and Credit Risk—Evidence from Bank Digitization in China. China Economic Quarterly, 23(05), 1686-1703. https://link.cnki.net/doi/10.13821/j.cnki.ceq.2023.05.03
[29] Illing, M., & Liu, Y. (2006). Measuring Financial Stress in a Developed Country: An Application to Canada. Journal of Financial Stability, 2(3), 243-265. https://doi.org/10.1016/j.jfs.2006.06.002
[30] Balakrishnan, R., Danninger, S., Elekdag, S., et al. (2011). The Transmission of Financial Stress from Advanced to Emerging Economies. Emerging Markets Finance and Trade, 47, 40-68. https://doi.org/10.2753/REE1540-496X4703S203
[31] Zhou, Z. Q., Li, L., & Huang, P. (2024). A Comprehensive Multi-Dimensional Risk Monitoring Model for Illegal Financial Activities. Journal of Risk Analysis and Crisis Response, 14(4), 472-488. https://doi.org/10.54560/jracr.v14i4.558
[32] Ahn, J. J., Oh, K. J, Kim, T. Y., et al. (2011). Usefulness of Support Vector Machine to Develop an Early Warning System for Financial Crisis. Expert Systems with Applications, 38(4), 2966-2973. https://doi.org/10.1016/j.eswa.2010.08.085
[33] Dabrowski, J. J., Beyers, C., & de Villiers, J. P. (2016). Systemic Banking Crisis Early Warning Systems Using Dynamic Bayesian Networks. Expert Systems with Applications, 62, 225-242. https://doi.org/10.1016/j.eswa.2016.06.024
[34] Elhoseny, M., Alhashmi, S. M., Deif, M. A., et al. (2025). Hybrid deep learning for bankruptcy prediction: An optimized LSTM model with harmony search algorithm. Journal of Intelligent Systems, 34(1), 20250058. https://doi.org/10.1515/jisys-2025-0058
[35] Krawczyk, B. (2016). Learning from Imbalanced Data: Open Challenges and Future Directions. Progress in Artificial Intelligence, 5(4), 221-232. https://doi.org/10.1007/s13748-016-0094-0

(Executive Editor: Qun Niu)