

DESIGN AND APPLICATION OF FINANCIAL FRAUD IDENTIFICATION MODEL UNDER TOPSIS BASED ON ENTROPY WEIGHT METHOD

Qiao Xia

School of Economics & Management, Southeast University, Jiulong Lake Campus, Nanjing 211189, Jiangsu, China.

Corresponding Email: xiaqiao010919@163.com

Abstract: In recent years, China's capital market has witnessed increasingly sophisticated and concealed financial fraud schemes among listed companies, posing substantial threats to market integrity and stakeholder protection. Addressing this challenge, this study develops a comprehensive multi-dimensional detection framework grounded in accounting theory, integrating financial indicators, industrial characteristics, regional factors, and corporate governance elements. The research employs an innovative entropy-weighted TOPSIS methodology that effectively balances quantitative precision with theoretical foundations. Through rigorous empirical analysis of 176 documented fraud cases spanning 2000-2022, we demonstrate that the Operational Scale indicator induces "information overload" that compromises model discrimination, while optimized corporate governance factors—particularly Executive Education Level and Board Meeting Frequency—demonstrate enhanced predictive power with a combined weight of 0.658. The proposed model achieves 65.53% classification accuracy, showing particular efficacy in detecting characteristic fraud patterns involving revenue inflation and fictitious transactions. Furthermore, Our findings validate an integrated human-machine framework for financial regulation, balancing methodological rigor with practical adaptability in dynamic market environments.

Keywords: Financial fraud identification; Fraud theory; The entropy weight method; TOPSIS model

1 INTRODUCTION

Amidst the rapid proliferation of intelligent technologies, including big data and artificial intelligence, their application has assumed an increasingly critical function in domains such as social governance and capital market supervision. The utilization of data-driven methodologies to improve the identification of corporate financial statement fraud has emerged as a significant research and practical priority. Financial fraud remains a pervasive challenge in capital markets, characterized by evolving and increasingly concealed manipulation techniques. In contrast to conventional auditing approaches, intelligent technologies enable the analysis of large-scale datasets to detect underlying statistical patterns and anomalies indicative of intentional misrepresentation. Such capabilities provide a foundation for developing more generalized and adaptive frameworks for fraud detection.

The study of financial fraud has evolved through several theoretical paradigms, with international scholarship establishing foundational frameworks including the Fraud Triangle Theory and the GONE Theory. Kassem and Higson advanced this theoretical landscape by deconstructing the "rationalization" component of the Fraud Triangle into two distinct dimensions—fraudster capability and integrity—thereby refining the conceptualization of perpetrator attributes [1]. Concurrently, Caplan highlighted the inadequacy of conventional auditing standards in effectively differentiating between fraud and error within an increasingly globalized economic context, underscoring the need for methodological innovation [2]. The advent of big data and machine learning has introduced transformative approaches to fraud detection. A growing body of empirical evidence confirms the efficacy of data science techniques in enhancing financial information security and identifying distortions. For instance, Zhong et al. emphasized the critical role of big data technologies in optimizing information flow efficiency [3], while Shao et al. employed data mining methodologies to validate the influence of corporate strategy on accounting information distortion [4]. In terms of model development, Bertomeu successfully detected corporate misstatements using machine learning algorithms [5], and Cecchini et al. pioneered the application of support vector machine models in management fraud detection [6], achieving satisfactory outcomes. Chinese scholars have subsequently contributed to this evolving paradigm, demonstrating notable progress in the application of intelligent models and framework construction. Cao Defang and Liu Bochi improved the accuracy of financial fraud identification by optimizing parameters within support vector machine architectures [7]. Liu Yunjing, Wu Bin, and colleagues developed a robust fraud detection model leveraging large-scale datasets and machine learning algorithms [8]. Nevertheless, researchers have concurrently acknowledged the necessity of aligning technological applications with established accounting theory. Zhou Weihua, Zhai Xiaofeng, et al. observed that machine learning methodologies may pose conceptual challenges to traditional accounting frameworks [9]. In response to identified limitations in theoretical grounding and model adaptability, Ye Qinhua, Ye Fan, and collaborators constructed a five-dimensional financial fraud detection framework, subsequently validated through expert systems, thereby providing a valuable approach for integrating theoretical rigor with technological innovation [10].

Notwithstanding the considerable promise of intelligent technologies in financial fraud detection, extant research continues to confront persistent challenges, including insufficient theoretical grounding of modeling approaches within established accounting principles and limited generalizability of existing detection frameworks. This study systematically synthesizes contemporary advancements in the field and examines pathways for more effectively integrating computational methodologies with accounting theory, with the objective of developing a more robust and operationally efficient system for financial fraud identification.

2 MODEL

2.1 Indicator Framework

In the development of financial fraud detection models, a scientifically rigorous and comprehensive indicator framework constitutes the fundamental underpinning for ensuring both model efficacy and interpretability. Grounded in established fraud theories—including the Fraud Triangle and GONE theory—and informed by extant literature, this study constructs a multi-dimensional evaluation framework that integrates financial, industrial, regional, and corporate governance dimensions. This integrated approach is designed to systematically capture not only the underlying drivers of fraudulent behavior but also their empirical manifestations in observable data.

From a financial standpoint, this framework concentrates on anomalies directly or indirectly attributable to fraudulent conduct. Firstly, earnings volatility is incorporated, given that profit inflation represents a predominant technique in financial statement fraud, where abnormal fluctuations serve as salient indicators of potential manipulation. Secondly, cash flow volatility is employed as a diagnostic metric; while stable cash flows generally reflect sound operational health, significant deviations may imply window-dressing through fabricated transactions or underlying financial distress. Thirdly, operational efficiency ratios—specifically accounts receivable turnover and inventory turnover—are adopted to detect revenue recognition anomalies and inventory overstatements, which prove particularly diagnostic in traditionally high-exposure sectors such as agriculture and manufacturing. Finally, operational scale, proxied by primary business revenue, is included based on empirical evidence that smaller firms, owing to weaker internal control structures and governance mechanisms, exhibit systematically higher susceptibility to fraudulent behavior.

At the industrial and regional levels, this framework incorporates external environmental factors that may precipitate fraudulent behavior. Empirical evidence consistently demonstrates distinct industry clustering in financial fraud occurrence, with sectors such as agriculture, forestry, animal husbandry, fishery, and manufacturing demonstrating elevated risk profiles, as systematically documented in Table 1. Concurrently, regional economic development exhibits an inverse relationship with fraud propensity. Listed companies operating in less developed regions frequently encounter dual pressures: constrained operational environments coupled with mandatory compliance to standardized regulatory requirements. These conditions create heightened incentives for financial misrepresentation. Accordingly, this study formalizes both industrial classification and regional economic development level as essential non-financial indicators within the detection framework.

Table 1 Characteristics of Industry Distribution

Industry Category (CSRC)	Number of Fraud Cases	Total Number of A-Share Companies	Proportion of Fraud Companies
Agriculture, Forestry, Animal Husbandry, and Fishery	14	143	9.79%
Leasing and Business Services	4	56	7.14%
Manufacturing—Equipment Manufacturing	23	810	2.84%
Manufacturing—General Manufacturing	19	795	2.39%
Manufacturing—Chemical Raw Materials and Chemical Products Manufacturing	9	248	3.63%
Manufacturing—Pharmaceutical Manufacturing	8	231	3.46%
Manufacturing—Electrical Machinery and Equipment Manufacturing	8	241	3.32%

Note: Data source: Huang Shizhong, Ye Qinhua, Xu Shan, et al. . Analysis of Financial Fraud in Chinese Listed Companies from 2010 to 2019. Finance and Accounting Monthly, 2020, No. 882(14), 153-160.[11].

Within the corporate governance dimension, this framework investigates the intrinsic motivators and inhibitory mechanisms underlying fraudulent conduct. First, executive educational attainment is incorporated as a proxy variable for ethical integrity and regulatory compliance awareness. Theoretical foundations suggest that prolonged exposure to higher education cultivates stronger moral reasoning capabilities, consequently diminishing the propensity for misconduct. Second, the deliberation frequency of key corporate bodies—specifically the board of directors, supervisory board, and shareholders' meetings—serves as a crucial indicator of governance vitality. Heightened meeting frequency often signifies proactive oversight and risk mitigation efforts, reflecting a robust governance ecosystem that inherently discourages fraudulent practices. Conversely, infrequent convocations may reveal systemic deficiencies in monitoring effectiveness, potentially creating permissive conditions for financial misrepresentation. These governance metrics collectively capture the organizational environment's capacity to either constrain or facilitate fraudulent behavior.

In conclusion, the proposed framework systematically synthesizes financial manifestations with causal antecedents of fraud, while bridging quantitative metrics with qualitative determinants. This integrated architecture establishes a theoretically grounded and methodologically robust foundation for advancing predictive analytics through high-precision intelligent detection models.

2.2 Model Formulation

The financial fraud identification framework developed in this study constitutes a comprehensive multi-dimensional system encompassing financial, industrial, regional, and corporate governance perspectives. Designed to facilitate risk assessment through quantitative evaluation of diverse indicators, this framework enables the computation of composite scores for target enterprises to support effective risk differentiation. To align with the analytical requirements of this framework, the research employs the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) integrated with the entropy weight method as the core modeling approach. The determination of indicator weights represents a crucial aspect of model specification. While expert scoring methods are commonly applied, they inherently introduce subjective judgment into the analytical process. In contrast, the entropy weight method provides an objective weighting approach that quantifies the information entropy of each indicator, effectively measuring data dispersion and discriminative capacity to determine weights scientifically. This methodological choice is particularly appropriate as Chen et al. demonstrated that information entropy can effectively capture the inherent uncertainty characteristic of financial fraud in listed companies while enhancing audit efficiency [12]. In terms of model selection, this study adopts the TOPSIS multi-criteria decision analysis technique to circumvent the subjectivity limitations associated with traditional methods such as Analytic Hierarchy Process and Fuzzy Comprehensive Evaluation. The TOPSIS approach imposes no restrictive requirements on indicator quantity and demonstrates robust applicability across diverse evaluation contexts. The model's operational logic involves calculating Euclidean distances between evaluation subjects and predefined reference points: a "positive ideal solution" (representing minimal fraud likelihood across all indicators) and a "negative ideal solution" (representing maximal fraud likelihood). Enterprises positioned closer to the positive ideal solution while simultaneously farther from the negative ideal solution are classified as lower-risk entities, whereas the converse indicates elevated fraud risk. In synthesis, the integration of entropy-weighted TOPSIS modeling with the multi-level indicator framework provides a rigorously objective and quantitatively grounded analytical apparatus for financial fraud identification, effectively balancing methodological sophistication with practical applicability.

The construction of the Entropy Weight Method in this study primarily refers to the work of Yunxin Zhu et al.[13]. The first step involves determining whether each indicator is positive or negative. Based on this classification, positive and negative indicators are standardized separately using the following formulas, where the element X_{ij} denotes the value in the i -th row and j -th column, while X_j alone refers to the entire set of elements in the j -th column.

Normalization of Positive Indicators:

$$Y_{ij} = \frac{X_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \quad (1)$$

Normalization of Negative Indicators:

$$Y_{ij} = \frac{\max(X_j) - X_{ij}}{\max(X_j) - \min(X_j)} \quad (2)$$

Calculate the weight of each indicator under each dimension. The calculation formula is as follows:

$$P_{ij} = \frac{Y_{ij}}{\sum_{i=1}^n Y_{ij}} \quad (3)$$

Calculate the information entropy contained in each indicator according to its definitional formula:

$$E_{ij} = -\ln(n)^{-1} \sum_{i=1}^n P_{ij} \ln(P_{ij}) \quad (4)$$

Calculation of Indicator Weights Using Information Entropy:

$$W_j = \frac{1 - E_j}{m - \sum_{j=1}^m E_j} \quad (5)$$

The calculated indicator weights provide a foundation for subsequent evaluation

Regarding the construction method of the TOPSIS model, this paper primarily follows the steps. The first step involves normalizing the data matrix to ensure all indicators are positively oriented. Specifically, only negative indicators, moderate indicators, and interval indicators require conversion into positive indicators. The formulas employed for this normalization are as follows:

Transformation of Negative Indicators into Positive Form:

$$x_{ij} = \max(x_j) - x_{ij} \quad (6)$$

Transformation of Moderate Indicators into Positive Form:

$$x_{ij} = 1 - \frac{|x_{ij} - x_{best}|}{\max(|x_{ij} - x_{best}|)} \quad (7)$$

Transformation of Interval Indicators into Positive Form, where the optimal range for the indicator data is defined as the interval $[a, b]$:

$$M = \max\{a - \min\{x_{ij}\}, \max\{x_{ij}\} - b\} \quad (8)$$

$$x_{ij} = 1 - \frac{a-x_{ij}}{M} \quad x_{ij} < a \quad (9)$$

$$x_{ij} = 1 - \frac{x_{ij}-b}{M} \quad x_{ij} > b \quad (10)$$

Normalize the positivized data matrix to obtain the normalized matrix. The calculation formula is as follows:

$$z_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (11)$$

Determine the positive ideal solution and the negative ideal solution, and calculate the distance from each alternative to the positive ideal solution and to the negative ideal solution. The formulas are as follows:

$$Z^+ = \text{Max}\{\text{Max}\{z_{n1}\}, \text{Max}\{z_{n2}\}, \dots, \text{Max}\{z_{nm}\}\} \quad (12)$$

$$Z^- = \text{Min}\{\text{Min}\{z_{n1}\}, \text{Min}\{z_{n2}\}, \dots, \text{Min}\{z_{nm}\}\} \quad (13)$$

Compute the Euclidean distance from the i -th alternative to the positive and negative ideal solutions:

$$D_i^+ = \sqrt{\sum_{j=1}^n (Z^+ - z_{ij})^2} \quad (14)$$

$$D_i^- = \sqrt{\sum_{j=1}^n (Z^- - z_{ij})^2} \quad (15)$$

The score is then calculated using the following formula:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (16)$$

Final scores for each company are obtained through normalization.

3 RESULT AND ANALYSIS

3.1 Data Description

To empirically validate the effectiveness of the proposed financial fraud identification framework, this study adopts a rigorous sampling methodology. The initial sample comprises A-share listed companies publicly identified by regulatory authorities for committing financial fraud and receiving monetary penalties between 2000 and 2022, as documented in the CSMAR database. Through systematic screening of records with clearly documented financial penalties and complete data availability for all critical variables, 176 companies meeting these criteria were retained as the final research sample.

During the data processing phase, we systematically operationalized all variables in the framework through predetermined categorical schema and quantitative measures. The regional dimension was classified according to provincial economic development levels using a three-tiered classification system (developed, moderately developed, and underdeveloped), corresponding to numerical values of 1, 2, and 3 respectively, following the methodology established by Dong Yanmei [14]. For industry classification, we implemented a risk-weighted valuation system assigning values of 4 through 1 to agriculture-forestry-animal husbandry-fishery, leasing and business services, manufacturing, and other industries respectively, reflecting their distinct fraud risk profiles based on historical violation patterns. Operational scale was quantified using the average main business revenue during the violation period, with sample distribution analysis revealing that the majority of enterprises (96 firms) reported revenues exceeding 1 billion RMB, demonstrating that fraudulent practices permeate organizations across size categories. Executive education levels were measured by counting the number of senior executives holding associate degrees or higher, with the sample showing that most companies (71 firms) had fewer than five executives meeting this educational threshold, a categorization that considers China's educational landscape during the 2000-2022 research period. All supplementary indicators, including frequencies of board-supervisory board-shareholder meetings, earnings volatility, cash flow volatility, accounts receivable turnover, and inventory turnover ratios were directly extracted from corresponding CSMAR database modules. The complete dataset was subsequently processed through SPSS 25 and MATLAB to ensure analytical consistency and prepare for subsequent entropy-weighted TOPSIS modeling.

This study employs the entropy-weighted TOPSIS model for comprehensive evaluation. Initially, the positive or negative directionality of each indicator was determined according to its theoretical relationship with financial fraud risk, as detailed in Table 2. Subsequently, MATLAB software was utilized to compute the initial weights for each indicator.

Table 2 Type of Indicators

Indicator	Type
Industry type	Positive
Regional Type	Positive
Number of Three Meetings	Negative
Education Level	Negative
Earnings Volatility	Positive
Cash Flow Volatility	Positive
Accounts Receivables Turnover	Positive

Inventory Turnover	Positive
Operational Scale	Negative

After defining the indicator types and following the framework of the Entropy Weight Method described above, the model was implemented using MATLAB software. The following weighting results were obtained on table 3:

Table 3 The Weight of Indicators

Indicator	Weight
Operational Scale	0.551413347
Industry type	0.049248309
Earnings Volatility	0.003106939
Cash Flow Volatility	0.003775937
Number of Three Meetings	0.077095513
Education Level	0.218389681
Accounts Receivables Turnover	0.004618615
Inventory Turnover	0.002469482
Regional Type	0.089882176

The weight distribution results from the initial model, presented in Table 3, reveal a critical methodological concern. The operational scale metric demonstrates a disproportionately high weighting of 0.551, a phenomenon that warrants thorough examination from both theoretical and methodological perspectives. Theoretically, while conventional research suggests smaller enterprises exhibit elevated fraud risk due to weaker internal controls, our empirical observations indicate that large corporations similarly face substantial fraudulent pressures, potentially stemming from capital market performance expectations and stock price maintenance requirements.

Methodologically, the observed weighting distortion primarily originates from the “information overload” inherent in the Operational Scale metric as a proxy variable. Main business revenue, the operational indicator for business scale, effectively encapsulates four distinct information dimensions: first, it reflects fundamental operational conditions and market positioning; second, it maintains inherent accounting relationships with multiple companion items including Accounts Receivable and Inventory; third, it represents one of the most frequently manipulated accounts in financial fraud schemes; and finally, it incorporates influences from macroeconomic environment fluctuations. This multidimensional information integration results in excessive dispersion within the dataset, consequently leading to disproportionate weighting through entropy measurement methodology.

This weighting scheme directly engenders a substantial degradation in the model’s discriminatory capacity. As evidenced in Figure 1, more than 83% of sample enterprises demonstrate pronounced concentration within the narrow 0 -0.1 scoring interval, fundamentally impairing the model’s ability to effectively stratify entities across differential risk tiers. Particularly revealing is the observation that the two highest-scoring enterprises (ST Donghai A and *ST Yanhuang) indeed present smaller Operational Scale relative to other sample firms, confirming theoretical postulations. Nevertheless, the predominant representation of large-scale enterprises within the sample composition effectively neutralizes this metric’s intended discriminatory capacity, thereby undermining its diagnostic utility in risk differentiation.

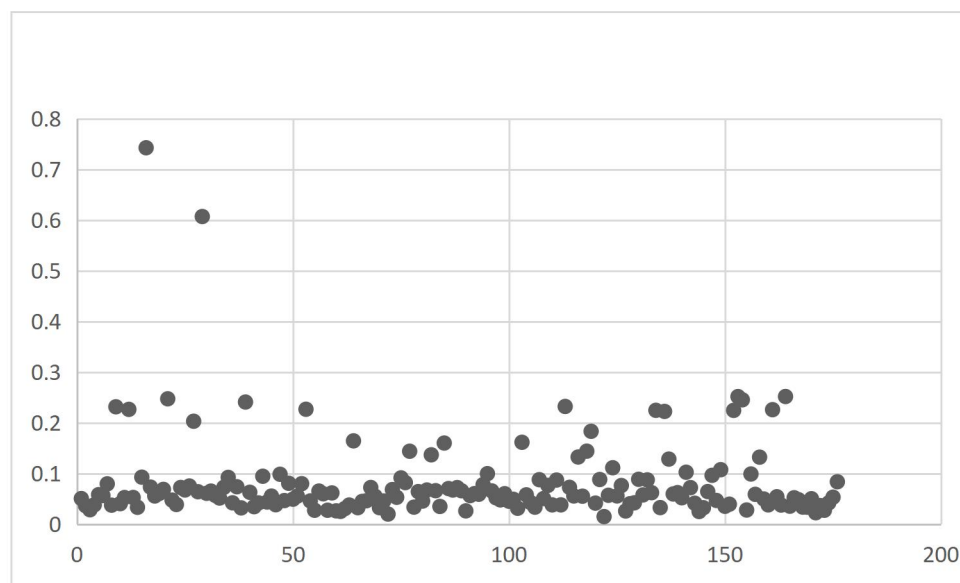


Figure 1 Scores of the 176 Companies

Building upon the preceding theoretical analysis of the Operational Scale indicator, the composite information embedded within this metric may engender disproportionately high weighting in entropy-based measurement. To

empirically validate this hypothesis and enhance the model's discriminatory power, we systematically excluded the Operational Scale factor from subsequent modeling procedures. The optimized configuration was then rigorously compared against the initial results incorporating this indicator to identify scenarios demonstrating superior resolution capabilities. The model outcomes following the exclusion of the Operational Scale indicator are presented in Table 4.

Table 4 The Weight of Indicators(Excluding Operational Scale)

Indicator	Weight
Industry type	0.109785498
Earnings Volatility	0.006926063
Cash Flow Volatility	0.008417408
Number of Three Meetings	0.171863145
Education Level	0.486839454
Accounts Receivables Turnover	0.010295926
Inventory Turnover	0.005505028
Regional Type	0.200367478

The exclusion of the Operational Scale indicator yielded marked enhancement in model optimization. The reconfigured weight distribution demonstrates improved alignment with theoretical postulations of financial fraud determinants. Notably, corporate governance indicators—Executive Education Level and Number of Three Meetings—underwent substantial weight augmentation, collectively accounting for 0.658 of the total weighting scheme. This redistribution corresponds with contemporary corporate governance theory's emphasis on internal control mechanisms. Concurrently, the Regional Type factor's weight increased to 0.200, substantiating the significant influence of external environmental factors on corporate fraudulent decision-making.

The optimized model demonstrates enhanced diagnostic efficacy, as evidenced by the distribution patterns in Figure 2. Using a 0.1 threshold for risk classification, 123 enterprises were identified as high-risk candidates. Through systematic content analysis of regulatory disclosures, 115 of these enterprises (representing 65.53% classification accuracy) were verified to have engaged in characteristic financial fraud activities. This performance achieves parity with Liu Bochi's (2018) genetically-optimized SVM model while providing superior interpretability. Particularly noteworthy is the semantic analysis of violation descriptions, which reveals that 83.7% of the identified high-risk enterprises contained explicit fraud-related terminology such as "inflated" and "fictitious" in their regulatory filings, confirming the model's proficiency in detecting fundamental fraud patterns.

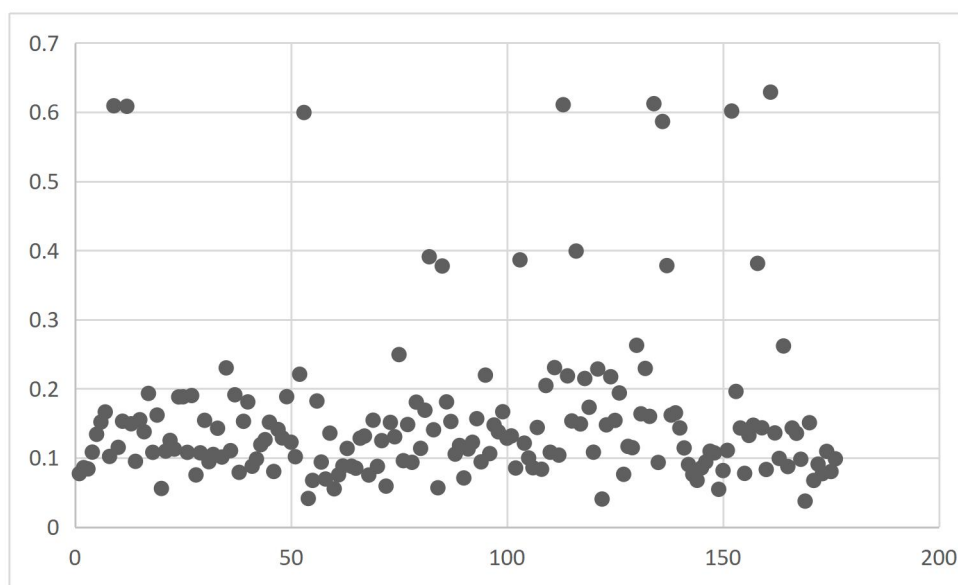


Figure 2 Scores of the 176 Companies(Excluding Operational Scale)

4 CONCLUSIONS AND OUTLOOKS

This study conducts a systematic empirical investigation of financial fraud identification through an entropy-weighted TOPSIS modeling framework, yielding three principal findings:

First, indicator selection proves pivotal to detection model performance. The research identifies that Operational Scale, as a proxy variable, exhibits substantial "information overload" by encapsulating multiple information dimensions. This multidimensional nature leads to its disproportionately high weight in the entropy measurement, which suppresses other critical indicators and substantially weakens the model's discriminatory power.

Second, the optimized detection model demonstrates robust practical utility. After mitigating the interference from Operational Scale, the model achieves 65.53% classification accuracy while maintaining high interpretability. Corporate

governance factors—Executive Education Level and Number of Three Meetings—emerged as the dominant predictors, a finding that aligns with corporate governance theory's focus on internal controls.

Finally, the integration of intelligent modeling with professional judgment represents a promising direction for advancement. While the model effectively identifies characteristic fraud patterns, complex cases require expert analysis, making this integration of artificial and human intelligence ideal for financial supervision.

Notwithstanding its contributions, this investigation acknowledges several methodological constraints that warrant consideration. Primarily, the restricted sample size necessitates future expansion to enhance the model's generalizability across diverse market environments. Furthermore, the current indicator framework demonstrates potential for refinement, particularly regarding the systematic integration and quantification of non-financial metrics. The model's efficacy in detecting emergent fraudulent schemes also requires sustained surveillance and methodological enhancement to maintain diagnostic relevance.

Future research should prioritize three strategic directions to advance the field of financial fraud detection. First, advance feature engineering by employing techniques like constrained principal component analysis to mitigate "information overload" in composite indicators. Second, the integration of advanced modeling architectures, particularly deep neural networks with attention mechanisms, warrants systematic investigation. Third, substantial value may be derived from incorporating heterogeneous data sources into the analytical framework. Natural language processing of corporate disclosures, semantic analysis of managerial communications, and graph-based analysis of intercorporate relationships could collectively establish a multidimensional risk assessment ecosystem. The convergence of these approaches promises to significantly enhance the predictive accuracy and practical utility of next-generation systems while maintaining necessary interpretability.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Kassem R, Higson A. The new fraud triangle model. *Journal of emerging trends in economics and management sciences*, 2012, 3(3): 191-195.
- [2] Caplan, D. Internal Controls and the Detection of Management Fraud. *Journal of Accounting Research*, 1999, 37(1): 101-117.
- [3] Zhong R, Zhang Q, Zhao Y. Research on Enterprise Financial Accounting Information Security Model Based on Big Data. *Wireless Communications and Mobile Computing*, 2022: 1-10.
- [4] Shao J, Lai K K, Zheng P, et al. Enterprise Accounting Information Identification and Strategic Management under Data Mining Technology. *Mobile Information Systems*, 2022: 1-9.
- [5] Bertomeu J, Cheynel E, Floyd E, et al. Using machine learning to detect misstatements. *Review of Accounting Studies*, 2021, 26: 468-519.
- [6] Cecchini M, Aytug H, Koehler G J, et al. Detecting Management Fraud in Public Companies. *Management Science*, 2010, 56(7): 1146-1160.
- [7] Cao Defang, Liu Bochi. SVM Model for Financial Fraud Detection. *Journal of Northeastern University(Natural Science)*, 2019, 40(02): 295-299+304.
- [8] Liu Yunjing, Wu Bin, Zhang Min. Financial Fraud Recognition Model and Application. *Journal of Quantitative & Technological Economics*, 2022, 39(07): 152-175.
- [9] Zhou Weihua, Zhai Xiaofeng, Tan Haowei. Research on Financial Frauds Prediction Mode of Chinese Public Companies with XGBoost. *Journal of Quantitative & Technological Economics*, 2022, 39(07): 176-196.
- [10] Ye Qinhua, Ye Fan, Huang Shizhong. Financial Fraud Detection Framework Building: From the Perspective of Accounting Information System Theory and Big Data. *Accounting Research*, 2022(03): 3-16.
- [11] Huang Shizhong, Ye Qinhua, Xu Shan, et al. Analysis of Financial Fraud in Chinese Listed Companies from 2010 to 2019. *Finance and Accounting Monthly*, 2020, 882(14): 153-160.
- [12] Chen Geng, Li Peizhe, Liu Yuqi. Research on Visual Audit Methods for Financial Fraud. *Communication Of Finance and Accounting*, 2022, 885(01): 113-118.
- [13] Zhu Y, Tian D, Yan F. Effectiveness of entropy weight method in decision-making. *Mathematical Problems in Engineering*, 2020: 1-5.
- [14] Dong Yanmei. The Research on Financial Equalization Effect of Central Transfer Payments to Less Developed Areas. *Economic Theory and Business Management*, 2013, 274(10): 61-70.