ISSN: 2054-7404
Archives of Business Research; Vol. 14 No. 02 (2026) (67-78)
https://doi.org/10.14738/abr.1402.19995

# The Rule-based Classification for the Embezzlement Detection using Case-based Reasoning

June-Suh Cho

1. College of Business, Hankuk University of Foreign Studies, South Korea

**Abstract:** Internal embezzlement is a major source of corporate losses. Companies are making various efforts to detect, prevent, and mitigate it. Embezzlement detection involves identifying financial irregularities through red flag behaviors, such as secretive work habits and unused vacation time. It also involves reviewing employee spending using machine learning-based transaction data analytics and forensic analysis. Furthermore, robust controls, such as double authorizations and regular audits, help detect issues such as fraudulent vendors, fake invoices, and misuse of company cards, thereby preventing significant financial losses. In this paper, we discuss a unique hybrid classification method for detecting embezzlement of corporations. We expect the proposed hybrid classification method to increase the accuracy of corporate embezzlement detection.

**Keywords:** embezzlement, Internal corruption, classification, criminology, case-based reasoning.

## INTRODUCTION

Reduced cash usage and enhanced security, such as CCTV, have sharply reduced theft and other physical economic crimes. Fraud Detection Systems (FDS) and tighter process-based controls on fund spending have also curbed external financial crimes against companies.

Despite these advances, employee embezzlement, especially by staff in finance, accounting, and bookkeeping, has increased. The damage now reaches unprecedented levels and causes serious social problems.

To prevent internal corruption, companies conduct accounting audits or use enterprise resource planning (ERP) systems and collaborative groupware to control expenditures. However, these control methods are cost-ineffective and, in practice, virtually impossible to detect.

To address these issues, developing a new intelligent internal corruption detection system is necessary. This system should detect and respond to employee misconduct, such as embezzlement and breach of trust, within a company. To this end, we aim to develop an internal embezzlement detection model based on case-based reasoning (CBR) that uses internal and external corporate data. We also propose a hybrid classification method for this purpose.

This method, developed by learning from previously collected embezzlement cases, will serve as the foundation for a system capable of detecting suspected internal corruption cases, such as embezzlement and breach of trust, in real time and minimizing corporate damage, amid the increasing incidence of financial crimes, including embezzlement and breach of trust, within companies.

In this paper, we discuss a rule-based classification method for detecting signs of embezzlement within a company.

Internal embezzlement can cause serious harm to an organization and necessitate preventive measures.

- Financial loss: Embezzlement is the theft of an organization's funds and assets, resulting in financial losses. This can threaten financial stability and limit the organization's economic opportunities.
- Loss of trust: Embezzlement can destroy internal and external trust in an organization. Employees may become aware of embezzlement and lose trust, leading to a decline in organizational morale and teamwork. Furthermore, the loss of trust with external stakeholders can negatively impact business relationships.
- Reputational damage: Embezzlement can significantly damage an organization's reputation. If an embezzlement case becomes public, stakeholders such as customers, business partners, and investors may lose trust in the organization and terminate their relationships with it.
- Legal issues: Embezzlement can result in legal consequences. Violations of relevant laws and regulations can result in legal sanctions, with serious consequences for an organization's operations and survival.

Therefore, preventing and detecting internal embezzlement is crucial for maintaining an organization's financial stability, credibility, reputation, and legal compliance.

Embezzlement detection is the process of proactively identifying anomalies, such as inconsistencies in financial records, unusual transaction patterns, and lifestyle changes, to prevent embezzlement. The main types of embezzlement indicators include:

- Unexplained discrepancies in financial records: Unexplained differences may occur in transaction details, deposits and withdrawals, and asset fluctuations.
- Abnormal transaction patterns: Typical examples include repetitive or fraudulent transactions, unauthorized access, non-compliant account movements, and financial activity at unusual times.
- Lifestyle changes: Abnormal behavior, such as sudden changes in spending patterns, frequent cash shortages, and external inflows of funds, may be detected.

Ensuring financial stability and building trust are essential for organizational health. Effectively detecting and preventing embezzlement is a critical challenge for financial stability in financial institutions and corporations.

According to the Association of Certified Fraud Examiners (ACFE), fraud is broadly categorized into asset misappropriation, corruption, and financial statement fraud, each with different causes and prevention strategies. Among these, asset misappropriation is the most controversial.

The ACFE identified three factors contributing to embezzlement: first, opportunity refers to situations where weaknesses in a company's internal control systems, such as a lack of separate financial management functions or inadequate approval procedures, can be exploited. Second, pressure refers to the motivation or pressure that employees face to commit embezzlement, such as financial difficulties or personal issues. Last, rationalization refers to the psychological process by which executives and employees justify their actions.

Employee and customer reports are the primary means of detecting fraud, accounting for over 40% of all cases. However, they have limitations, resulting in the detection of only 4% of all cases [3]. This result stems from the limitations of existing methods as fraud techniques evolve, necessitating the development of new detection methods.

Corporate embezzlement can be detected, predicted, and prevented using financial and non-financial indicators, as well as AI-based solutions.

Key signs of embezzlement within a company include financial indicators, such as accounting manipulation, abnormal cash flow, and deteriorating financial ratios like increased accounts payable, and non-financial indicators, such as long-term tenure, employee monopolization, frequent supplier switching, low employee ethics, and behavioral changes like increased spending, frequent overtime, and refusal to take vacations.

To date, companies have mostly relied on internal control systems to detect and prevent embezzlement, but improvements are needed. Recently, digital audits using big data and AI-based analytics have emerged to detect and predict abnormal transaction patterns in real time.

This study aims to develop a method for detecting embezzlement by integrating employees' behavioral changes and patterns into big data to identify abnormal accounting data. We propose a hybrid classification approach that combines association rules and CBR as the foundation of this detection method.

Table 1 shows indicators related to fraud and embezzlement, including analytical and behavioral indicators. Although collecting and analyzing data on employee disagreements or conflicts is difficult due to privacy concerns, behavioral pattern analysis is an important tool for detecting embezzlement.

**Table 1: Indicators related to fraud and embezzlement**

| **Analytical Indicators of Fraud and Embezzlement** | **Behavioral Indicators of Fraud and Embezzlement** |
|---|---|
| Unusual fluctuations in financial ratios (e.g., operating profit margin, return on equity)<br>Discrepancies in cash flow (e.g., significant differences between accounting profit and cash flow) | Disagreements or conflicts between employees (e.g., disagreements between budget managers and financial analysts)<br>Isolating or secretive behavior (e.g., working late or alone with the office door closed) |
| Changes in profitability or spending patterns (e.g., unusually high margins, unexplained increases in spending) | Inconsistent explanations or evasive responses (e.g., vague or ambiguous answers to questions about expenses) |
| Losses or losses of assets or inventory (e.g., unreported inventory shortages or lost equipment) | Overwhelming workloads or unrealistic deadlines (e.g., managers who overwork or give impossible deadlines) |

Fraud detection programs employing AI and ML are largely designed to counter external threats, such as identity theft and hacking [2][4][5]. However, these tools do not adequately address internal corporate fraud, including embezzlement, breaches of trust, and executive or employee misconduct. Internal corruption is particularly challenging, especially in the realm of career fraud, due to limited precedent, a lack of relevant data, and continually evolving criminal tactics.

In this study, we analyzed data and cases involving experts, including lawyers and police officers, to derive rules on internal corporate corruption.

Our objective is to create a method tailored for detecting and preventing internal corruption. We leverage rules and Case-Based Reasoning (CBR), grounded in criminology, to synthesize lessons from past incidents and legal precedents in order to uncover distinct patterns of organizational misconduct.

CBR stands out from other systems in that it accurately models the human decision-making process. Its main benefit is its ability to solve complex problems through analogical or empirical reasoning, learning from past cases to improve future fraud detection and prevention. This makes CBR particularly effective for adapting to evolving corrupt practices.

Internal corruption, such as embezzlement, poses continuous and severe threats to organizations. Traditional audits have become insufficient as perpetrators adopt increasingly sophisticated methods, resulting in more frequent and more impactful incidents.

Internal corruption, including embezzlement, is often difficult to document due to restricted discussion. Nonetheless, advancements in information technology now allow detection and analysis of corruption using both internal and external data, as well as business and behavioral analytics.

In particular, internal corruption case-based inference analyzes past corruption cases within a company based on criminal cases and precedents to identify similar patterns, causes, and criminal methods. This approach is used not only to detect corruption but also to predict and prevent it.

Identifying internal corporate corruption is critical for organizational survival. Sustained vigilance is required to detect early signs of embezzlement, insider trading, and other internal threats. Utilizing advances in AI and data analysis enables companies to proactively protect their reputation and assets by predicting or uncovering misconduct.

This paper presents a hybrid rule-based classification method that uses case-based reasoning (CBR) to detect internal corruption, such as embezzlement.

## BACKGROUND

Financial fraud is a serious problem for businesses and society, impacting people's lives [11][17]. It also undermines market trust and stability, causing direct economic impacts [16]. [1] noted that financial fraud, such as embezzlement, takes various forms, including asset misappropriation, expense reimbursement fraud, and financial statement manipulation.

Machine learning (ML) has improved financial fraud detection and prevention by helping detect new threats, minimize damage, and reduce the negative impact of financial fraud on organizations [4]. These ML methods find patterns and anomalies in large-scale financial data related to fraud. However, their high cost, concerns about inaccurate classification, current detection methods, and privacy issues hinder fraud detection [10][15][18].

Threats to companies can be external or internal. External threats are often disguised as systems, while internal threats, such as corruption, are among the most damaging [6].

In criminology, embezzlement is a clear example of internal corporate fraud. The specific causes of corporate embezzlement include:

Professional ethics are essential. While most employees work to benefit the company, a small minority commit embezzlement. The workplace supports families and helps employees find fulfillment.

Stronger management oversight prevents embezzlement. Consistent procedures and document review deter fraud. Strict supervision discourages criminal activity.

Finance and accounting staff should rotate regularly for cross-checking. Phased approvals and strict supervision prevent workplace crime.

Management negligence leads to workplace crime. Headquarters must combat embezzlement through strong management and ongoing training. Ethics instruction and moral reinforcement help prevent these crimes.

In this paper, we discuss rule-based classification. Rule-based classification is a simple, easy-to-understand, and easy-to-implement method.

Rules in rule-based classification methods are based on data characteristics and the problem domain. One of their advantages is that they are easy to understand and explain. Furthermore, for well-defined, specific problems, they tend to be more accurate than machine learning algorithms.

However, rule-based classification methods have several limitations. They can easily generate errors when presented with new examples that do not conform to predefined rules, and they tend to be less accurate on complex, diverse datasets. However, rule-based classification methods are simple, easy to understand, and easy to implement.

Rule-based classifiers tend to perform poorly on complex, varied datasets and can be difficult to scale to large datasets. Additionally, they require significant expertise and time to define the rules and keep them up to date.

Therefore, a hybrid classification method is derived by utilizing CBR and association rules.

Case-based reasoning (CBR) is a prominent artificial intelligence technique. Its core principle is that human experts employ analogical or heuristic reasoning to address complex problems and derive knowledge from past experiences. CBR systematically searches a case repository to retrieve cases most analogous to the current issue.

CBR has modeled decision-making processes. [14] developed a CBR system to evaluate EDP controls and generate recommendations for information system controls. The system uses prior cases to compare controls, infer failure scenarios, and explain recommendations. Morris validated it by comparing its performance with that of human subjects, showing that it generated superior recommendations.

[13] observed that auditors use analogies from prior experience. [8] noted that CBR can enhance memory and machine learning in intelligent systems. This applies to CBR and to machine learning methods using crime-related behavioral data [12].

Effective detection and prediction of internal corporate embezzlement requires a multifaceted approach to preventing internal corruption. Companies can detect and prevent internal corruption through regular transaction monitoring, analysis of diverse employee data, regular audits, and the adoption of cutting-edge technologies.

## RULE-BASED CLASSIFICATION USING CASE-BASED REASONING

With recent rapid advances in digital technologies such as artificial intelligence and machine learning, companies are increasingly leveraging them to develop more intelligent detection methods. These methods utilize large amounts of data to identify abnormal patterns or suspicious employee behaviors through analysis.

This study will discuss a hybrid classification method that combines association rules and CBR to support these detection methods.

### Hybrid Classification

In this paper, we combine rule-based and CBR methods to leverage the strengths of rule-based methods, such as their speed and efficiency in classifying transactions that clearly match established rules, as well as transparency and explainability. And, the weaknesses of the rule-based classification are as follows;

- Cannot detect new, evolving, or sophisticated fraud patterns that do not violate an existing rule.
- Requires constant, manual updates by human experts as fraudsters find ways to bypass existing rules.
- Rigid rules can sometimes flag legitimate transactions, leading to alert fatigue.

To overcome these weaknesses, we use CBR together. The combination of Rule-Based Classification and Case-Based Reasoning (CBR) provides a more robust and flexible system than either method could provide on its own.

CBR is an AI methodology that solves new problems by adapting solutions from previously solved similar problems. It essentially mimics human experiential reasoning.

The core principle of CBR is that a new, suspicious activity is compared to a library of stored, categorized past embezzlement or fraud cases.

The main processes of the CBR are as follows.

- Find the past cases in the case base that are most similar to the new target case.

- Propose a solution for the new case based on the classification of the retrieved similar cases.
- If the proposed solution is incorrect or the case is new, it will be reviewed by an expert.
- Store the new problem and its validated solution as a new case in the case base, allowing the system to learn and adapt over time.

The hybrid method quickly processes routine transactions and flags clear violations using rule-based methods, while CBR handles complex, borderline, or new cases that fixed rules cannot classify. By finding analogous past cases, CBR can provide a reasoned classification even when no rule is directly violated.

In the context of embezzlement detection, this hybrid model can analyze financial indicators such as internal financial transactions, journal entries, and vendor setups, as well as non-financial indicators, including big data and employee behavioral data.

Figure 1 shows the hybrid classification process for embezzlement detection. There are four steps in the process as follows.

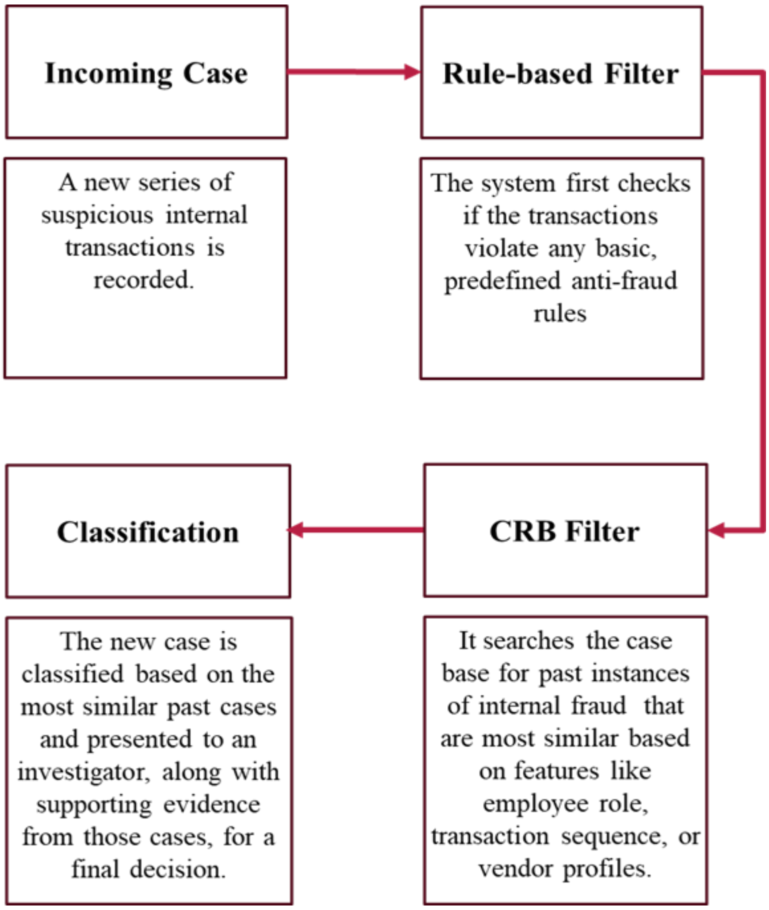
Incoming Case
Rule-based Filter
A new series of suspicious internal transactions is recorded.
The system first checks if the transactions violate any basic, predefined anti-fraud rules
Classification
CRB Filter
The new case is classified based on the most similar past cases and presented to an investigator, along with supporting evidence from those cases, for a final decision.
It searches the case base for past instances of internal fraud that are most similar based on features like employee role, transaction sequence, or vendor profiles.

**Figure 1: The Classification Process**

First, the Incoming Case is that a new series of suspicious internal transactions has been recorded.

Second, the Rule-based Filter checks whether transactions violate predefined anti-fraud rules.

Third, the CRB Filter searches the case base for the most similar past instances of internal fraud, based on features such as employee role, transaction sequence, or vendor profiles.

Fourth, classification is the process by which a new case is assigned to the most similar past cases and presented to an investigator, along with supporting evidence from those cases, for a final decision.

Table 2 shows an example of deriving rules for embezzlement crimes by analyzing criminal case law. We derived victimization cases from criminal methods identified in crime cases. We derived association rules based on the given basic resource data.

In Case-Based Reasoning (CBR), the system must determine how closely a current suspicious transaction resembles historical examples of embezzlement stored in its database.

**Table 2: An example of an embezzlement method derived from case law**

| **criminal method** | **Split the loan under a third party's name** |
|---|---|
| damage cases | Borrowers who are unable to obtain a loan due to poor credit or insufficient collateral obtain multiple loans under third-party names. |
| basic source data | Borrower's personal information<br>Interest payment account holder's name or depositor's name |
| rules | If the borrower's name differs but the personal information (address, phone number, email address, etc.) remains the same, an alarm will be raised.<br>If the borrower's name is different from the name of the person who paid the interest, an alarm will be raised. |

### Rule-based classification using Case-Based Reasoning

The company's embezzlement detection method uses case-based reasoning, which involves analyzing specific past embezzlement cases within the company, drawing on criminal precedents, and leveraging big data and knowledge to detect and identify potential patterns and causes of internal embezzlement.

[9] discussed the system process of internal corruption detection in Figure 2. In this paper, we focus on the classification step in embezzlement detection.

Figure 2 shows the system structure of internal corruption detection, which includes a classification step. The classification process, the previous step of model training, uses preprocessed historical case and precedent data. This step is necessary to improve the accuracy of embezzlement detection. The rules are derived from an analysis of data on embezzlement crimes.

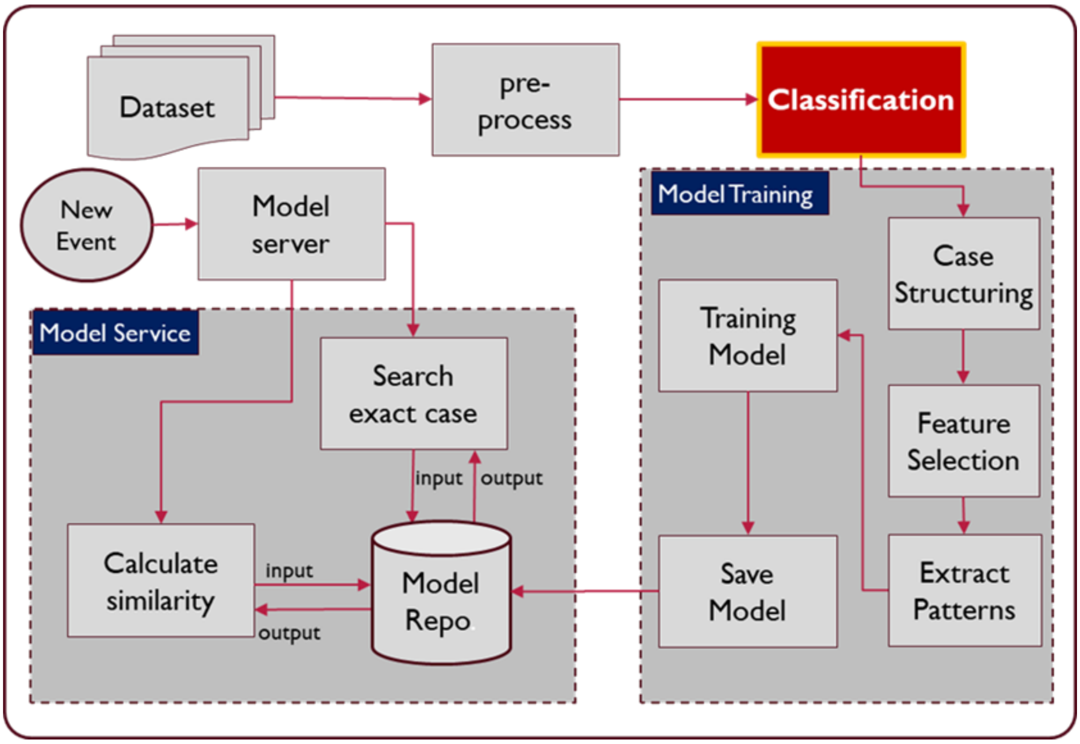

**Figure 2: The Structure of Internal Corruption Detection**

Applying machine learning to rule-based classification and case-based reasoning helps companies analyze past embezzlement, identify vulnerabilities, and develop detection methods to minimize internal corruption risks. This proactive approach reduces reputational damage and prevents financial losses.

Building on this approach, case-based reasoning detects internal corporate corruption by analyzing past cases and comparing them to current transactions or behavior patterns. When combined with rule-based classification, as presented in this study, the hybrid method further improves detection and prediction accuracy.

To ensure comprehensive dembezzlement detection, it is important to use an integrated approach that combines rule-based classification, data analysis, pattern recognition, and continuous monitoring. Specifically, rule-based classification and case-based inference detection methods using machine learning offer further improvements in accuracy and efficiency. One of the prominent artificial intelligence techniques is case-based reasoning (CBR). Its core principle is that experts use analogical or heuristic reasoning to solve complex problems and draw on historical experiences. CBR systematically searches a case repository to retrieve cases most analogous to the current issue.

In this study, we extract rules from criminal cases and various financial data and identify patterns in legal cases.

## EVALUATION

We analyzed criminal methods across cases and derived rules based on data correlations.

We derived 10 rules from 300 criminal cases and evaluated them against 500 sample data from companies. In particular, many experts have sought to derive methods from criminal cases and related rules.

First, we evaluated with rule-based classification. When tested on refined data, the rules were correctly classified. However, when classification was performed using unrefined data, false alarms occurred.

False alarms can occur in a variety of ways, including one rule that detects problems by comparing bank deposit details with the name on the tax invoice. The algorithm must determine the similarity between the bank deposit details and the name on the tax invoice, but this error occurs because the algorithm is inaccurate.

The system issues an alarm indicating that a deposit has been made on the tax invoice, but the deposit has not been made. Tests using supervised learning achieved an accuracy of over 92%, whereas those using unsupervised learning achieved less than 40% due to many false alarms.

Second, we evaluated a hybrid classification that combines association rules and Case-based Reasoning. When we have applied a hybrid classification, unsupervised learning improved to 61%. We need to improve accuracy, and to do that, we need to derive more rules and learn more data to find similar crime patterns. Additionally, efforts are needed to improve accuracy by applying neural networks.

In conclusion, rule-based classification has limitations, and to improve classification accuracy, ML utilizing CBR must be applied.

To further improve accuracy, more rules must be derived and more data trained to identify similar patterns.

In addition, future work will focus on analyzing employee behavioral data to detect and predict unique patterns of embezzlement.

## CONCLUSION

This paper presents a hybrid classification method that combines rule-based methods with case-based reasoning (CBR), which serves as the foundation for embezzlement detection.

This unique hybrid method combines association rules and case-based reasoning with diverse data, including financial, accounting, and employee behavior data. It draws on a criminological approach derived from embezzlement-related data, case studies, criminal techniques, and precedents.

The proposed method is expected to serve as a foundation for detecting internal corporate corruption, improving not only embezzlement detection but also prediction accuracy.

Furthermore, detecting internal corruption, such as embezzlement, is crucial for enhancing corporate transparency and ethical behavior and preventing and detecting fraud.

This study has several limitations. First and foremost, the limitation of rule derivation posed significant difficulties in obtaining real corporate data for testing.

Future embezzlement detection research will focus on deriving more rules from a broader range of case studies and criminal techniques, and on utilizing CBR to derive more patterns from these rules. This method will enhance the accuracy of detection and prediction.

## ACKNOWLEDGMENT

This study was supported by Hankuk University of Foreign Studies Research Fund of 2025.

## REFERENCES

[1]. Abdallah, A., et al., Fraud detection system: a survey. Journal of Network and Computer Applications, 2016, Vol. 68.

[2]. Achakzai, M. and Juan, P., Using machine learning meta-classifiers to detect financial fraud. Finance Research Letters, 2022, Vol. 48.

[3]. ACFE, Occupational Fraud 2024: A Report to the Nations, 2024.

[4]. Ahmed, M., et al., A survey of anomaly detection techniques in the financial domain. Future Gener Comput System, 2026, 55.

[5]. Ali, A, et al., Financial fraud detection based on machine learning: a systematic literature review. Applied Science, 2022.

[6]. Albrecht, W., et al, Fraud Examination. Cengage, 6th, 2019.

[7]. Barletta, R., An introduction to case-based reasoning. AI Expert, 1991.

[8]. Brown, C. E. and Gupta, U., Applying case-based reasoning to the accounting domain. Intelligent Systems in Accounting, Finance, and Management, 1994, Vol 3.

[9]. Cho, J., The Process of Internal Corruption Detection using the Criminological Approach based on Case-based Reasoning, European Journal of Business and Management Research, 2024, 9(6).

[10]. Dantas, R., et al., Systematic acquired critique of credit card deception exposure through machine learning. Journal of Open Innovation Technology Market and Complexity, 2022, 8(4).

[11]. Hernandez Aros, L., et al., Financial fraud detection through the application of machine learning techniques: a literature review, Humanities and Social Sciences Communications, 2024.

[12]. Kumar, S., et al., Exploitation of machine learning algorithms for detecting financial crimes based on customers' behavior. Sustainability, 2022, 14(21).

[13]. Meservy, R., et al., Internal control evaluation: A computational model of the review process. Auditing: A Journal of Practice & Theory, 1986, Vol. 6.

[14]. Morris, B. W., SCAN: A case-based reasoning model for generating information system control recommendations. International Journal of Intelligent Systems in Accounting, Finance and Management, 1994, Vol 3.

[15]. Nicholls, J., et al., Financial cybercrime: a comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape. IEEE Access 9:163965–163986, 2021.

[16]. Reurink, A., Financial fraud: a literature review. Journal of Economic Surveys. 2018, 32(5).

[17]. Singh, A., et al, Financial fraud detection approach based on the firefly optimization algorithm and the support vector machine. Applied Computational Intelligence and Soft Computing, 2022.

[18]. West, J. and Bhattacharya, M., Intelligent financial fraud detection: a comprehensive review. Computers & Security, 2016, Vol. 57.