

Artificial Intelligence Applications in Fraud Risk Management: A Systematic Literature Review

Fiqih Nur Rokhman

Faculty of Economics and Business, Universitas Indonesia

fiqih.nur@ui.ac.id

Abstract

Purpose: This study aims to identify the roles, common techniques, and industry applications of Artificial Intelligence (AI) in Fraud Risk Management (FRM).

Methodology/approach: A Systematic Literature Review (SLR) was conducted following the PRISMA protocol. Data were collected from the Scopus database and managed using Zotero and SciSpace, while Microsoft Excel was used for data processing and analysis. The review covered 95 studies published between 2015 and 2025.

Results/findings: The findings show that most studies focus on fraud detection (80%), followed by risk assessment and prevention. Machine Learning (51%) and Deep Learning (29%) are the most commonly used AI techniques. The banking sector (39%) leads in implementation, followed by the insurance and public sectors. Explainable AI (XAI) is identified as an emerging trend for enhancing transparency.

Conclusions: AI greatly improves FRM by increasing detection accuracy, supporting proactive fraud prediction, and strengthening governance transparency. It transforms traditional systems into adaptive, data-driven frameworks.

Limitations: The study only used the Scopus database and found high variation among study metrics, preventing a meta-analysis.

Contribution: This study provides a comprehensive overview for researchers, practitioners, and policymakers in finance and IT. It maps the AI FRM landscape and emphasizes the importance of balancing technological performance with ethical and transparent governance.

Keywords: *Artificial Intelligence, Fraud Risk Management, Fraud Detection, Systematic Literature Review*

1. Introduction and Literature Review

Fraud Risk Management (FRM) has become one of the most pressing concerns for organizations, regulators, and governments worldwide. The rapid digitization of financial services and the proliferation of online transactions have created both opportunities for economic efficiency and challenges in maintaining data integrity and trust. According to the FRM Guide developed by the Association of Certified Fraud Examiners (ACFE) and the Committee of Sponsoring Organizations of the Treadway Commission (COSO), fraud remains an ever-present threat that can cause severe financial, reputational, and governance damage if not properly managed (ACFE & COSO, 2023).

Within this framework, the FRM Guide outlines five essential components of an effective FRM system: governance and culture, risk assessment, control activities, investigation and corrective action, and continuous monitoring. Among these, fraud detection serves as the primary mechanism for identifying early warning signals and mitigating risks before significant losses occur.

At the same time, Artificial Intelligence (AI) has become a transformative enabler in modern FRM practices. Through machine learning (ML) and deep learning (DL), AI systems can process extensive financial datasets, detect anomalies, and predict fraudulent behaviors that traditional rule-based systems often fail to capture (Dhieb et al., 2020; Carminati et al., 2015). In the financial sector, AI-driven techniques such as ensemble learning and anomaly detection have proven effective in analyzing large

volumes of transaction data, improving both accuracy and responsiveness in fraud detection (Raval et al., 2023). Consequently, integrating AI technologies into the FRM framework aligns with the COSO and ACFE emphasis on analytics-based continuous monitoring to enhance fraud prevention and detection outcomes (ACFE & COSO, 2023).

Recent years have witnessed a notable surge of research exploring AI applications in financial and insurance risk management. For instance, Carminati et al. (2015) introduced Bank Sealer, a decision-support system for online-banking fraud analysis that achieved detection rates above 95 percent using user-behavioral profiling. Likewise, Dhieb et al. (2020) proposed an AI-driven insurance-fraud architecture that integrates extreme-gradient-boosting algorithms with blockchain technology to identify suspicious claims in real time. These studies collectively demonstrate the capacity of AI to automate complex analytical tasks, enable real-time monitoring, and strengthen the predictive accuracy of FRM systems.

Parallel to these technical advances, researchers have emphasized the need for ethical governance and explainability in AI-enabled FRM. Opaque model architectures especially deep neural networks can hinder auditors' and regulators' ability to interpret algorithmic reasoning, posing risks to accountability and compliance. In response, the emerging field of Explainable AI seeks to make AI systems transparent and interpretable (Černevičienė & Kabašinskas, 2024). In the context of financial services, XAI frameworks such as SHapley Additive ExPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) are increasingly employed to justify automated fraud-risk decisions and maintain stakeholder trust (Kieslich et al., 2022).

In addition, recent deep-learning architectures, including recurrent neural networks (RNNs) and long short-term memory (LSTM) models, have enhanced the ability to capture temporal patterns in sequential financial data (Raval et al., 2023). The integration of these techniques within governance-oriented frameworks ensures that AI not only improves predictive performance but also complies with regulatory expectations for transparency and fairness.

Despite this progress, evidence remains fragmented regarding how AI is embedded within broader FRM structures. Many studies emphasize algorithmic optimization while overlooking the alignment of AI adoption with control activities, ethics, and organizational governance (ACFE & COSO, 2023). Addressing this gap requires a systematic and integrative approach that evaluates AI's technical contributions together with its governance and ethical implications.

This study applies a Systematic Literature Review (SLR) using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology (Pati & Lorusso, 2018) to synthesize existing knowledge on AI applications in FRM. It aims to identify prevailing trends, methodological developments, and governance considerations related to AI integration into FRM systems. Accordingly, this study addresses the following research questions (RQs):

- RQ1. What functions or roles of AI have been proven beneficial in FRM?
- RQ2. What AI techniques and models are most commonly applied for fraud detection, prevention, and risk assessment?
- RQ3. In which industries or organizational sectors have AI-based Fraud-Risk-Management systems been most frequently studied or applied?

Therefore, this study conducts a Systematic Literature Review (SLR) to consolidate existing empirical and conceptual findings, aiming to provide a comprehensive understanding of AI's roles, techniques, and implementations in FRM, along with their governance implications. This review specifically addresses three research questions concerning AI's beneficial functions, commonly applied techniques, and predominant industry applications in FRM systems.

2. Methodology

This study employed a Systematic Literature Review (SLR) approach to examine prior research concerning the utilization of AI in the domain of FRM. The primary goal of this review was to explore

prevailing research directions, methodological progress, and the evolution of scholarly insights on how AI technologies are integrated into systems for fraud detection, prevention, and overall risk management. The research design adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol (Pati & Lorusso, 2018). Consistent with the PRISMA procedure, the review was conducted through four sequential stages identification, screening, eligibility, and inclusion which are elaborated in the following subsections describing the data selection and analysis process.

2.1 The Method of Data Selection Criteria

The data collection process was conducted exclusively through the **Scopus** database, which is widely recognized for its extensive coverage of high-quality, peer-reviewed international journals. The **Scopus** database was selected as the primary source because of its rigorous indexing standards and strong reputation within the global academic community (Galvão et al., 2018). A Boolean search strategy was applied using the following keywords: "artificial intelligence" AND ("FRM" OR "fraud detection"). This query was designed to capture studies addressing AI-driven fraud detection and FRM frameworks across different industries. The search covered journal articles published between January 2015 and February 2025, ensuring the inclusion of contemporary works reflecting recent technological and methodological advancements.

The initial search identified 1,392 articles. During the first exclusion phase, 1,205 records were removed for the following reasons:

- a. The publication was not a peer-reviewed journal article (e.g., conference paper, editorial, or book chapter) (n = 1,033).
- b. The article was not written in English (n = 16).

After this exclusion, 343 journal articles were retained for the screening phase. In the second stage, the titles and abstracts of the remaining articles were carefully reviewed. Studies were included only if they explicitly examined AI, machine learning (ML), deep learning (DL), or explainable AI (XAI) within the context of fraud detection, or risk management. Articles that lacked direct relevance to FRM or discussed AI in unrelated fields (such as general cybersecurity or data mining) were excluded. Additionally, 156 articles were removed at this stage because their full texts were not accessible. This left 187 articles for detailed evaluation in the eligibility phase. During eligibility screening, the full texts were examined based on the following inclusion and exclusion criteria:

Inclusion Criteria:

1. The article is a peer-reviewed journal publication written in English.
2. The study specifically discusses AI, ML, DL, or XAI in the context of FRM or detection.
3. The paper provides either empirical results, conceptual models, or framework discussions relevant to FRM.

Exclusion Criteria:

1. The article's full text was unavailable or contained incomplete bibliographic data.
2. The study did not explicitly link AI methods to FRM objectives (e.g., focused solely on generic data analytics).

After applying these criteria, 92 studies were excluded, leaving a final sample of 95 articles that met all eligibility standards. These papers were then included for comprehensive analysis and synthesis. The overall selection process following the PRISMA methodology is presented in Figure 1, which illustrates each stage of article identification, screening, eligibility, and inclusion.

2.2 Method for Data Analysis

The selected articles were carefully reviewed, and key information was extracted and compiled into a summary matrix for systematic comparison. The extraction and organization of bibliographic data were supported by Zotero, which facilitated reference management and metadata verification, and SciSpace, which was used to assist in article screening, keyword extraction, and thematic coding. Quantitative

data related to publication trends, journal rankings, and methodological classifications were further processed and analyzed using Microsoft Excel, allowing for tabulation, visualization, and descriptive statistics. The extracted data included the following attributes:

1. Authors and year of publication
2. AI techniques or models applied, such as ML, DL, or XAI
3. Application domain or industry sector (e.g., banking, insurance, energy, or public sector)
4. Research objectives and main findings
5. Dataset characteristics and performance indicators, including accuracy, precision, recall, and AUC
6. Relevance to the COSO/ACFE FRM framework, specifically in the areas of detection, prevention, or governance

A descriptive analysis was performed to identify the most frequently applied AI methods, publication trends, and industrial sectors represented in the literature. Additionally, a qualitative synthesis was conducted to interpret how AI technologies contribute to enhancing fraud-risk-management practices particularly in improving detection efficiency, predictive accuracy, and ethical explainability.

The major theme of this review is to map the relationship between AI-driven technologies and the five components of the COSO/ACFE FRM Framework (2023) governance, risk assessment, control activities, investigation and corrective action, and continuous monitoring. By aligning AI techniques with these FRM dimensions, the study provides a conceptual foundation for future research and organizational practice in developing ethical, transparent, and data-driven fraud-prevention systems.

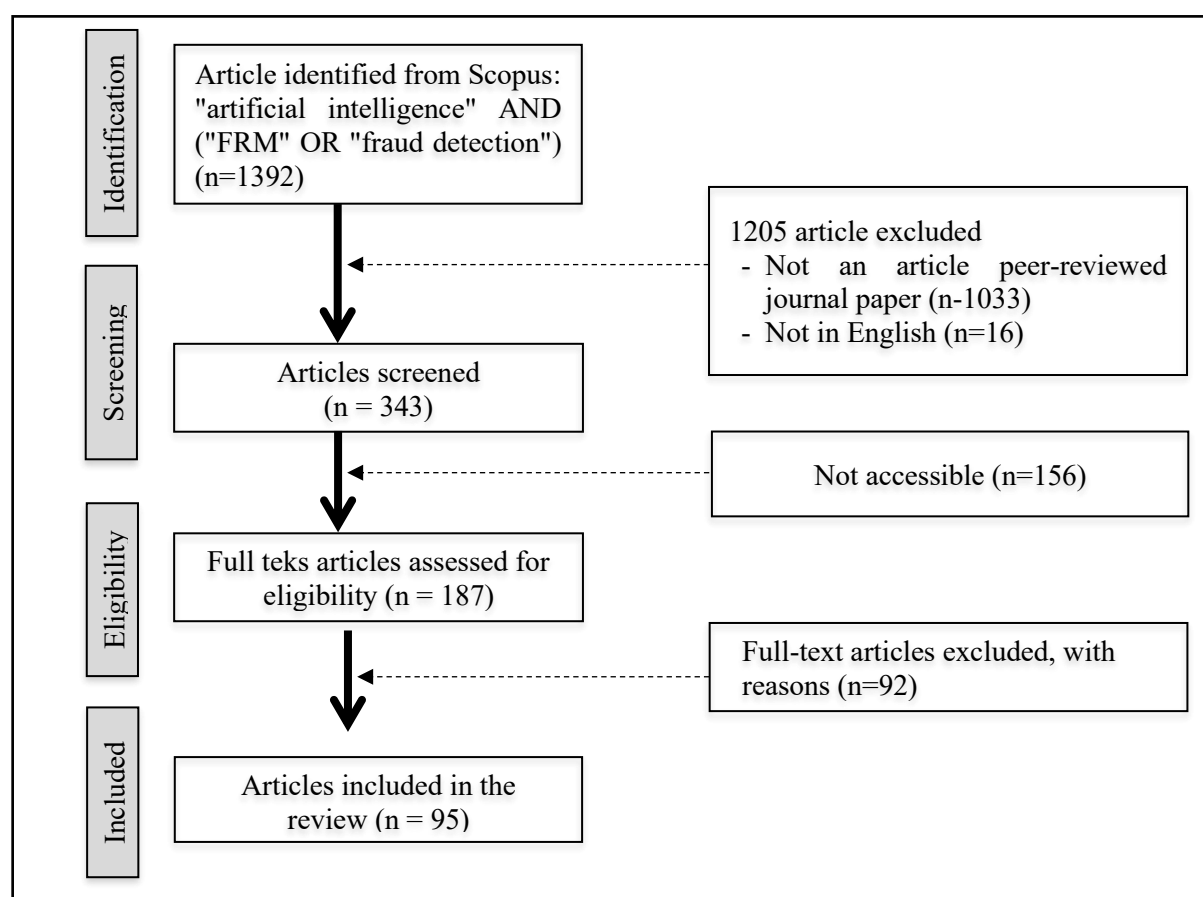


Figure 1. Article selection process based on PRISMA flow diagram

3. Results and Discussion

A few outcomes of the descriptive statistical analysis taken from Scopus included: 1) the trend of articles each year; 2) Distribution of Journal Quality by ScimagoJr; and 3) Keyword Co-occurrence Network.

3.1 Trend of article per year

Figure 2 shows the number of publications on AI in FRM from 2015 to 2025.

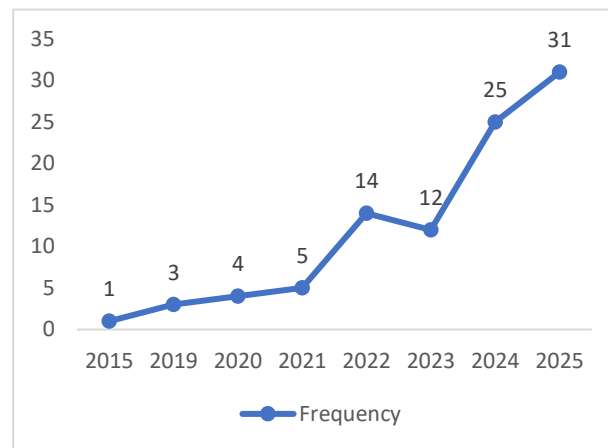


Figure 2. Trend of article per year (processed data by Scopus)

The trend indicates a steady and significant increase in research over the last decade. Only one article was published in 2015, followed by gradual growth through 2019–2021, reaching five studies in 2021. A sharp rise occurred in 2022 with 14 publications, reflecting growing interest in using machine learning and deep learning for fraud detection. Although there was a slight decrease in 2023 (12 articles), the number rose sharply again to 25 in 2024 and 31 in 2025. This upward trend shows that AI-based FRM has become an increasingly important and mature research area, driven by advances in explainable AI and the growing focus on ethical and transparent risk management practices.

3.2 Distribution of Journal Quality (Scimago Journal Rank)

Figure 3 presents the distribution of the reviewed articles according to their Scimago Journal Rank (SJR) classification.

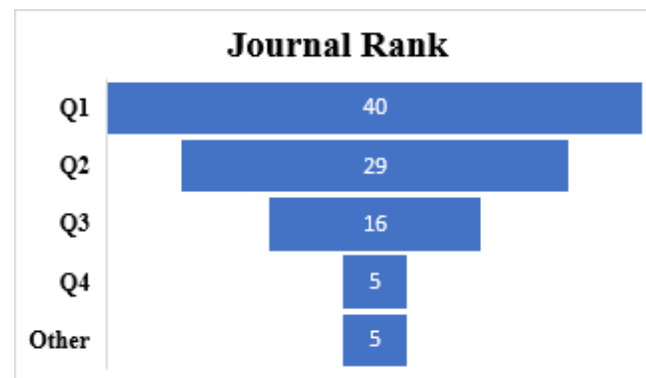


Figure 3. Distribution of Journal Quality

The majority of the studies were published in Q1 journals (40 articles, 42%), indicating that most research on AI-based FRM appears in high-impact and reputable outlets. This demonstrates strong scholarly attention and methodological rigor within the field. Following this, Q2 journals accounted for 29 publications (31%), reflecting a substantial contribution from well-recognized, mid-to-high-tier

journals. Meanwhile, Q3 journals represented 16 studies (17%), and Q4 journals contributed 5 studies (5%). An additional 5 papers (5%) were published in other sources not ranked in Scimago, such as emerging or open-access journals. Overall, this distribution shows that the majority of AI-FRM research has been disseminated through high-quality international journals, reinforcing the credibility and academic maturity of this research domain.

3.3 Keyword Co-occurrence Network

Figure 4 presents the keyword co-occurrence network generated using VOSviewer, which visualizes how frequently and strongly different keywords appear together across the reviewed literature. The size of each node represents the frequency of a keyword, while the color and proximity indicate thematic clustering and relational strength. The network reveals three major clusters representing the main research themes in AI-based FRM:

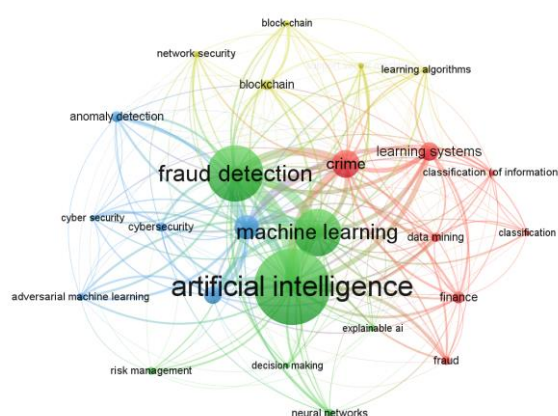


Figure 4. Keyword Co-occurrence Network

- a. **Green Cluster – Core AI and Risk Management Applications**
The largest cluster centers around the keywords “artificial intelligence,” “machine learning,” “fraud detection,” and “risk management.” This cluster reflects the core of the research field, emphasizing the use of AI algorithms particularly machine learning and neural networks for detecting fraudulent transactions, assessing risks, and supporting decision-making processes. The strong linkages among AI, ML, and fraud detection indicate that these terms form the conceptual backbone of most studies.
- b. **Blue Cluster – Cybersecurity and Anomaly Detection Themes**
The second cluster includes terms such as “cybersecurity,” “network security,” “adversarial machine learning,” and “anomaly detection.” This area focuses on fraud prevention within digital infrastructures, highlighting research that applies AI techniques to protect financial networks, detect intrusions, and manage vulnerabilities in online transaction systems. The prominence of cybersecurity and network security suggests a growing overlap between FRM and information-security research.
- c. **Red Cluster – Data Analytics, Learning Systems, and Financial Applications**
The red cluster connects “learning systems,” “data mining,” “classification,” “finance,” and “fraud.” This theme reflects analytical and technical research focused on algorithm development, including supervised and unsupervised learning methods, classification models, and data-driven fraud analysis in financial environments. The presence of finance and data mining terms underscores the quantitative nature of FRM studies.

Across clusters, bridging terms such as “explainable AI,” “blockchain,” and “decision making” indicate emerging interdisciplinary trends, combining AI transparency, fintech innovation, and strategic governance. The strong interconnections among clusters also reveal that AI-based FRM research is highly integrated, with overlapping interests between technical model development, cybersecurity protection, and ethical governance.

Overall, the visualization suggests that fraud detection using machine learning and artificial intelligence remains the dominant research focus, while newer areas such as explainable AI and blockchain-based

risk management are rapidly gaining importance. This confirms that the field is evolving from pure algorithmic studies toward integrated, transparent, and governance-driven approaches to FRM.

Discussion

RQ 1. What functions or roles of AI have been proven beneficial in FRM?

Based on the analysis of 95 peer-reviewed journal articles published between 2015 and 2025, five major functional themes of AI in FRM were identified: (1) fraud detection, (2) risk assessment and prediction, (3) fraud prevention and early warning systems, (4) governance, auditing, and compliance, and (5) explainability, trust, and human-AI collaboration. These categories collectively illustrate how AI has enhanced the effectiveness, efficiency, and transparency of fraud management processes across diverse sectors, including banking, insurance, energy, and the public sector.

Table 1. Quantitative distribution of AI functions

AI Function / Role	Number of Studies	%	Key Representative Studies
Fraud Detection	76	80%	Machine/deep learning and ensemble models for anomaly and transaction detection. (Dhieb et al., 2020; Carminati et al., 2015; Khan et al., 2022; Almazroi & Ayub, 2023; Ismail & Haq, 2024)
Risk Assessment & Prediction	44	46%	AI used for risk scoring, probability modeling, and credit risk prediction. (Mhlanga, 2020; Buyuktepe et al., 2023; Fritz-Morgenthal et al., 2022; Hajj & Hammoud, 2023; COSO FRM Guide, 2023)
Fraud Prevention / Early Warning Systems	29	31%	Predictive monitoring, alert-based systems, and behavioral analytics for proactive fraud control. (Narsimha et al., 2022; Mill et al., 2023; Taher et al., 2023)
Governance, Auditing, & Compliance	17	18%	Automation of internal control, algorithmic auditing, and governance support. (A Full-Population Auditing Method, 2022; Artificial Intelligence for Financial Accountability in the Public Sector, 2025; COSO FRM Guide, 2023)
Explainability / Human-AI Trust	12	13%	XAI frameworks, ethical design, and transparency for trust and accountability. (Capuano et al., 2022; Chaquet-Ulldemolins et al., 2022; Kieslich et al., 2022; Fritz-Morgenthal et al., 2022)

a. Fraud Detection

Fraud detection emerged as the most prevalent function of AI within the reviewed literature, appearing in approximately 80% of the studies. AI-based detection systems rely on advanced algorithms that identify anomalous or suspicious behaviors in real time, substantially improving both speed and precision compared with traditional rule-based systems. Machine learning (ML) and deep learning (DL) models have been widely applied to detect credit card fraud, financial statement manipulation, and online transaction irregularities (Carminati et al., 2015; Dhieb et al., 2020; Ismail & Haq, 2024). Hybrid and ensemble approaches, such as XG Boost, Light GBM, and federated learning models, have achieved high detection accuracies some exceeding 97% (Taher et al., 2023; Reddy et al., 2024).

AI also plays an instrumental role in identifying anomalies within decentralized platforms like blockchain and cryptocurrency exchanges, thereby improving transparency and integrity (Taher et al., 2023). Studies in sectors beyond finance, such as food safety (Vinothkanna et al., 2024) and energy distribution (Coma-Puig & Carmona, 2019), further demonstrate AI's cross-sector adaptability in detecting diverse types of fraud. Overall, AI-driven detection systems enable organizations to automate monitoring processes, reduce false positives, and ensure more reliable identification of fraudulent activities in complex data environments.

b. Risk Assessment and Prediction

The second dominant theme concerns the use of AI for risk assessment and predictive analytics. Approximately 46% of the reviewed studies employed AI models to quantify fraud risk, predict the likelihood of occurrence, and support data-driven decision-making. Predictive models leveraging regression, reinforcement learning, and neural networks have improved the accuracy of forecasting fraudulent events (Raval et al., 2023; Buyuktepe et al., 2023). For instance, Buyuktepe et al. (2023) applied explainable AI (XAI) to assess fraud risk in food supply chains, demonstrating how interpretability enhances decision quality. Similarly, Hajj and Hammoud (2023) found that AI-based systems strengthen risk measurement in financial institutions by addressing information asymmetry and improving customer risk profiling.

In financial contexts, AI-powered credit scoring and risk assessment frameworks have replaced manual evaluations, enhancing fairness, efficiency, and consistency in risk judgments (Mhlanga, 2020). These systems leverage massive datasets, integrating both structured and unstructured data, to forecast potential fraud risks with remarkable precision. Consequently, AI facilitates a proactive approach to fraud management by identifying vulnerabilities before they escalate into significant losses.

c. Fraud Prevention and Early Warning Systems

Beyond detection, AI has been increasingly utilized for fraud prevention and early warning mechanisms. Approximately one-third of the studies analyzed emphasize this role. AI systems continuously monitor behaviors, enabling early identification of potential fraudulent activities through pattern recognition and behavioral analytics. Several studies proposed AI-driven frameworks that combine predictive modeling and real-time monitoring to generate automatic alerts when abnormal patterns emerge (Narsimha et al., 2022; Mill et al., 2023). These systems allow financial institutions to intervene before fraudulent transactions are completed, significantly reducing financial and reputational damage.

Other works emphasize AI's preventive capacity through behavioral analytics and sentiment modeling, detecting emerging fraud trends or cyberattacks in dynamic environments such as mobile banking or decentralized finance (Yang, 2021; Oleiniks, 2022). Overall, AI-enabled early warning systems enhance organizational resilience by offering predictive intelligence, improving responsiveness, and enabling adaptive fraud prevention strategies in rapidly evolving digital ecosystems.

d. Governance, Auditing, and Compliance

Another recurring theme in the reviewed literature relates to the role of AI in governance, auditing, and regulatory compliance. Around 18% of the studies highlighted AI's potential to modernize internal control and audit processes. AI-driven auditing systems automate the detection of irregularities across entire datasets, eliminating sampling bias and improving internal control quality (A Full-Population Auditing Method, 2022; Hajj & Hammoud, 2023). In the public sector, AI has been implemented to enhance accountability in government spending by identifying anomalies in procurement and budget allocations (AI Financial Accountability and Governance, 2025).

Moreover, explainable AI (XAI) has strengthened the transparency of decision-making processes, which is essential for regulatory compliance and ethical governance (Mhlanga, 2020). The integration of AI into audit and compliance frameworks not only minimizes human error but also supports continuous monitoring, enabling real-time assurance and governance improvements. Thus, AI acts as both a technological and ethical enabler of robust FRM governance.

e. Explainability, Trust, and Human–AI Collaboration

The final theme involves explainability and human–AI collaboration, reflecting a growing emphasis on ethical and interpretable AI in fraud management. Although only 13% of studies directly address this issue, it represents an emerging frontier in FRM research. Explainable AI (XAI) techniques such as SHAP and LIME have been used to interpret complex model outputs, providing transparency and interpretability to auditors and regulators (Capuano et al., 2022; Chaquet-Ulldemolins et al., 2022). Raval et al. (2023) demonstrated that integrating LSTM with XAI improves trust and accountability in

fraud classification models, while S41597-025-04664-Y (2025) revealed that human-AI collaboration in reviewing alerts enhances both accuracy and efficiency.

These studies underscore the importance of integrating human oversight with machine intelligence to ensure responsible AI deployment. As regulatory frameworks such as the EU AI Act (2021) emphasize ethical AI governance, research in this area is likely to expand. The focus is gradually shifting from maximizing predictive power to ensuring trustworthy, transparent, and human-centered AI systems in FRM.

Overall, the analysis reveals that AI plays multifaceted roles across all phases of the FRM cycle. Fraud detection remains the most mature and widely studied application, while risk assessment and early prevention demonstrate AI's predictive and proactive potential. Meanwhile, governance and explainability reflect the ongoing evolution toward responsible, transparent, and human-aligned AI systems. The reviewed evidence collectively demonstrates that AI has not only improved the accuracy and efficiency of fraud management but also transformed traditional FRM practices into more data-driven, adaptive, and ethically governed systems. This transformation underscores AI's critical role as a catalyst for innovation and integrity in modern FRM frameworks.

RQ 2. What AI techniques and models are most commonly applied for fraud detection, prevention, and risk assessment?

Analysis of the 95 reviewed studies revealed a diverse range of AI techniques applied to detect, prevent, and assess fraud risks across banking, insurance, public, and digital-finance sectors. The dominant categories include machine learning (ML), deep learning (DL), hybrid and ensemble models, explainable AI (XAI), and graph-based or big-data approaches. Each class of technique addresses a specific stage of the fraud-risk management (FRM) process ranging from early detection to ongoing governance.

Tabel 2. Quantitative summary of AI technique families

AI Technique Family	Examples of Models	%	Main Application Focus
Machine Learning (ML)	Logistic Regression, Decision Tree, Random Forest, SVM, Naïve Bayes, KNN	51%	Core fraud detection, supervised classification on structured data.
Deep Learning (DL)	ANN, CNN, RNN, LSTM, Autoencoder, GAN, Deep Reinforcement Learning	29%	Detecting complex sequential or image-based fraud patterns.
Ensemble / Hybrid Optimization	XG Boost, Light GBM, CatBoost, Firefly, Genetic Algorithm	23%	Improving predictive stability and addressing class imbalance.
Explainable AI (XAI)	SHAP, LIME, Decision Rules, White-Box ML	15%	Enhancing interpretability and governance in FRM systems.
Graph-Based & Big Data	GNN, Node2Vec, Federated Learning, Hadoop, Spark	12%	Detecting network or collusion fraud; large-scale risk management.

a. Machine Learning

Machine learning models remain the most widely implemented techniques, dominating over half of all FRM studies. They rely on historical, structured data (e.g., transaction logs, claims records) to train classifiers that differentiate between legitimate and fraudulent patterns. Random Forest (RF) and Support Vector Machine (SVM) are particularly dominant because of their high detection accuracy and robustness to noise (Dhieb et al., 2020; Carminati et al., 2015).

Logistic Regression (LR) remains popular as a baseline benchmark due to its interpretability and ease of implementation (Ismail & Haq, 2024). Naïve Bayes and K-Nearest Neighbors (KNN) are applied in lightweight fraud-screening and e-commerce scenarios due to their low computational cost (Tuning Machine Learning Models Using a Group Search Firefly Algorithm, 2022). These models are valued for their speed, interpretability, and cost-efficiency, making them ideal for operational FRM systems where transparency is critical.

b. Deep Learning

Deep-learning techniques have expanded rapidly since 2018, supported by larger datasets, cloud computing, and GPU acceleration (Mosavi et al., 2020). They dominate high-volume, nonlinear, and temporal fraud analysis, such as credit-card and online-payment fraud detection. Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) detect complex feature patterns in transactional data, particularly when integrated with autoencoders for unsupervised anomaly detection (Dhieb et al., 2020).

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models capture time-series dependencies, identifying evolving patterns and concept drift over time (Raval et al., 2023). Generative Adversarial Networks (GANs) augment training data in highly imbalanced datasets, while Deep Reinforcement Learning (DRL) provides adaptive and continuous decision support for fraud prevention (Mosavi et al., 2020). This family of models emphasizes autonomous feature learning and self-improving risk analysis, making them suitable for dynamic fraud environments such as blockchain, fintech, and mobile payment systems.

c. Ensemble and Hybrid Models

Hybrid and ensemble models have gained popularity because they combine the predictive strengths of multiple algorithms, mitigating data imbalance and model overfitting. Gradient-boosted algorithms (e.g., XG Boost, Light GBM, Cat Boost) consistently achieve state of the art accuracy for financial-fraud datasets (Tuning Machine Learning Models Using a Group Search Firefly Algorithm, 2022). Metaheuristic optimization methods such as the Firefly Algorithm, Genetic Algorithm, and Particle Swarm Optimization are often applied to fine-tune hyperparameters and accelerate convergence (Dhieb et al., 2020). Hybrid deep-ensemble systems, such as combinations of Autoencoder, ResNet or RF, ANN, fuse shallow and deep representations to balance interpretability and performance (Ismail & Haq, 2024). These approaches are pivotal in real-time fraud monitoring, where speed, reliability, and false-positive reduction are essential.

d. Explainable AI

With regulators such as COSO (2023) and the EU Artificial Intelligence Act (2021) emphasizing auditability, Explainable AI (XAI) has become a fast-emerging area in FRM research. Techniques like SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) provide feature-level transparency for complex “black-box” models, allowing financial institutions to justify algorithmic fraud decisions (Kieslich et al., 2022).

Similarly, white-box ML models, such as interpretable decision trees and rule-based algorithms, support governance compliance by ensuring traceability of AI predictions (Černevičienė & Kabašinskas, 2024). The integration of XAI ensures that AI-based FRM systems remain ethically aligned and maintain stakeholder trust through interpretability and explainability.

e. Graph and Big Data Methods

A smaller yet strategically significant segment of FRM research leverages graph-based analytics and distributed computing to detect organized or hidden fraud networks. Graph Neural Networks (GNNs) and Node2Vec embeddings represent entities (e.g., accounts, users, or transactions) as nodes and edges, enabling relational modeling to uncover collusive fraud (Zhou et al., 2021).

Federated Learning (FL) is employed in multi-institutional contexts, allowing collaborative model training across organizations without direct data exchange (Capuano et al., 2022). Big-data ecosystems like Hadoop and Apache Spark support scalable, high-velocity analysis, making these methods particularly relevant for anti-money laundering (AML), insurance-claim analysis, and public-sector

risk monitoring (Mhlanga, 2020). These innovations enhance cross-organizational fraud detection while maintaining data privacy and computational efficiency.

In summary, among the 95 studies analyzed, Machine Learning (51%) dominates as the most practical and interpretable family of AI methods in FRM. Deep Learning (29%) and Ensemble Models (23%) drive innovation in accuracy and adaptability, whereas Explainable AI (15%) and Graph-based Methods (12%) represent growing research directions emphasizing ethical oversight and interconnected fraud risk. Together, these techniques reflect a shift from static, rule-based detection toward proactive, intelligent, and transparent FRM systems.

RQ3. In Which Industries or Organizational Sectors Have AI-Based Fraud-Risk-Management Systems Been Most Frequently Studied or Applied?

The synthesis of 95 peer-reviewed studies reveals that the implementation of AI in FRM spans a wide range of industries. The majority of research focuses on sectors characterized by high financial transaction volumes and data-intensive risk exposure, particularly banking and financial services, insurance, and public-sector governance. Other studies extend to energy and utilities, healthcare, corporate auditing, and emerging digital commerce applications.

This distribution underscores the cross-sectoral importance of AI in detecting, preventing, and managing fraud across economic system.

Tabel 3. Quantitative distribution of sectors

Industry / Sector	Number of Studies	Percentage (%)	Representative Studies
Banking & Financial Services	37	39%	Carminati et al. (2015); Raval et al. (2023); Ismail & Haq (2024)
Insurance	13	14%	Dhieb et al. (2020); Buyuktepe et al. (2023)
Public Sector / Government	11	12%	Van Bekkum & Zuiderveen Borgesius (2021); Kieslich et al. (2022); COSO (2023)
Energy & Utilities	7	7%	Coma-Puig & Carmona (2019); Pérez-Lara et al. (2020)
Healthcare	5	5%	Capuano et al. (2022); Chaquet-Ulldemolins et al. (2022)
Corporate Auditing / Governance	6	6%	Hajj & Hammoud (2023); COSO (2023)
E-Commerce / Digital Trade	9	9%	Taher et al. (2023); Zhou et al. (2021)
Other or Cross-Sectoral Studies	7	7%	Mosavi et al. (2020); Mhlanga (2020)

a. Banking and Financial Services

The banking and financial sector is the most extensively explored domain for AI-driven FRM, accounting for nearly two-fifths of all reviewed studies. The volume of digital transactions and susceptibility to credit-card, online-payment, and loan-application fraud make this sector a primary target for AI innovation. Machine-learning models such as Random Forest, Support Vector Machine, and XG Boost dominate detection systems for credit-card and digital-payment transactions (Carminati et al., 2015; Ismail & Haq, 2024).

Deep-learning architectures including LSTM and Autoencoder networks enable sequential anomaly detection in transaction streams (Raval et al., 2023). The widespread availability of transactional big

data has encouraged the adoption of ensemble and hybrid models that balance accuracy and interpretability. Financial institutions increasingly combine predictive analytics with Explainable AI (XAI) frameworks such as SHAP to justify automated decisions to regulators, thereby aligning operational efficiency with compliance obligations (Černevičienė & Kabašinskas, 2024)

b. Insurance

Insurance firms adopt AI primarily for claims-fraud detection and risk-underwriting optimization. Studies demonstrate that supervised ML models including logistic regression and gradient-boosting ensembles achieve high accuracy in identifying anomalous claim patterns (Dhieb et al., 2020). In addition, explainable and hybrid frameworks integrate textual and numerical claim data to enhance transparency for both regulators and customers (Buyuktepe et al., 2023). AI applications in insurance thus combine fraud prevention with risk scoring, creating early-warning mechanisms that enable dynamic premium adjustments and targeted audits.

c. Public Sector and Government

AI deployment within the public sector focuses on improving accountability and reducing misuse of public funds. Research highlights applications in social-welfare distribution, procurement oversight, and tax-compliance monitoring. For example, the Dutch *SyRI* project used predictive analytics for welfare-fraud detection, prompting debates over data ethics and algorithmic transparency (Van Bekkum & Zuiderveen Borgesius, 2021). Further, ethics-by-design frameworks and public-trust analyses (Kieslich et al., 2022) have underscored the necessity for explainability in governmental AI. The COSO–ACFE FRM Guide (2023) positions AI as a governance enabler, enhancing continuous monitoring and internal-control efficiency. Collectively, public-sector studies emphasize accountability, transparency, and compliance, making XAI principles indispensable in this domain.

d. Energy and Utilities

Within energy and utilities, AI research focuses on detecting non-technical losses (NTLs) such as electricity theft and irregular energy consumption. Techniques like entropy-based anomaly detection and graph-embedding models (Node2Vec) analyze meter readings and network topologies to identify abnormal patterns (Coma-Puig & Carmona, 2019). Hybrid models that integrate big-data platforms (Hadoop, Spark) enable scalable, near-real-time monitoring of energy distribution systems. These applications highlight AI's role in safeguarding infrastructure, promoting sustainability, and minimizing operational losses.

e. Healthcare

Although smaller in volume, the healthcare sector leverages AI for medical-billing and insurance-claim fraud detection. Convolutional and autoencoder-based models identify discrepancies in diagnosis and treatment codes (Capuano et al., 2022; Chaquet-Ulledemolins et al., 2022). Here, privacy-preserving AI is critical, ensuring compliance with health-data protection laws while maintaining detection performance.

f. Corporate Auditing and Governance

AI in corporate auditing supports continuous auditing, anomaly detection in ledgers, and financial statement integrity. Studies (Hajj & Hammoud, 2023; COSO, 2023) show that AI-assisted auditing enhances both accuracy and speed in internal-control systems. Explainable algorithms, often coupled with ensemble methods, strengthen accountability and transparency in financial reporting. This sector demonstrates AI's potential to transform audit functions from periodic to real-time assurance processes.

4.8 E-Commerce and Digital Trade

AI applications in e-commerce target payment-gateway protection, account-takeover prevention, and identity verification. Graph and deep-learning models analyze user-behavior patterns and device fingerprints to detect fraudulent interactions (Zhou et al., 2021; Taher et al., 2023). Moreover, ensemble and reinforcement-learning systems facilitate adaptive fraud detection, adjusting thresholds dynamically as market conditions evolve. This domain benefits from the scalability of AI to process millions of transactions in real time while maintaining a low false-positive rate.

g. Cross-Sectoral and Emerging Studies

Several studies adopt a cross-sectoral perspective, investigating AI's integration across industries and its potential for macro-level risk prediction. For instance, deep-reinforcement and meta-learning approaches have been applied to simulate risk behaviors in macroeconomic and fintech contexts (Mosavi et al., 2020). Mhlanga (2020) further emphasizes AI's role in promoting financial inclusion and ethical automation under the Industry 4.0 framework. Such studies demonstrate AI's versatility beyond domain-specific use, paving the way for enterprise-wide FRM ecosystems that combine analytics, governance, and human oversight.

Across the 95 studies analyzed, the financial ecosystem (banking + insurance) accounts for over half of all AI-FRM research ($\approx 53\%$), confirming its position as the technological epicenter of fraud analytics. Public-sector applications (12%) underscore transparency and ethics, while energy, healthcare, and auditing research illustrate AI's extension into infrastructure, social services, and corporate governance. E-commerce, though relatively young, is an emerging growth area due to rapid digitalization and the increasing complexity of cyber-fraud schemes.

Overall, these findings reveal a clear pattern: AI adoption in FRM follows the trajectory of digital transaction intensity and regulatory oversight the more data-driven and compliance-sensitive a sector, the greater its reliance on AI for fraud detection, prevention, and governance.

5. Conclusion

5.1. Conclusion

This study conducted a Systematic Literature Review (SLR) of 95 peer-reviewed articles published between 2015 and 2025 to examine the role of AI in FRM. The research aimed to identify (1) the functions and benefits of AI in FRM, (2) the AI techniques and models most commonly applied, and (3) the industrial sectors where AI-based FRM systems are implemented.

The findings reveal that AI significantly enhances the effectiveness, efficiency, and transparency of FRM across multiple dimensions. First, AI contributes most prominently to fraud detection, followed by risk assessment and prediction, fraud prevention and early warning, governance and compliance, and explainability and human-AI collaboration. This aligns with the COSO-ACFE FRM Framework (2023), where AI plays a key role in improving continuous monitoring, risk assessment, and control activities.

Second, Machine Learning (ML) and Deep Learning (DL) techniques are the most frequently applied, with hybrid and ensemble models (e.g., XGBoost, LightGBM) demonstrating high accuracy in fraud detection. Furthermore, Explainable AI (XAI) methods, such as SHAP and LIME, strengthen auditability and stakeholder trust, ensuring that AI-driven systems remain transparent and ethically governed.

Third, the implementation of AI-based FRM is most prevalent in banking and financial services, followed by insurance, public-sector auditing, e-commerce, and energy utilities. These sectors exhibit high levels of digital transaction intensity and regulatory oversight, both of which drive the adoption of AI for fraud prevention and governance.

Overall, this study concludes that AI not only improves predictive accuracy and detection capability but also transforms traditional FRM systems into adaptive, data-driven, and ethically aligned governance frameworks, enhancing both operational efficiency and accountability.

5.2. Limitation

Despite its comprehensive approach, this study has several limitations. First, the review relied solely on the Scopus database, which may have excluded relevant studies from other sources or non-English publications. Second, the heterogeneity of research methods and evaluation metrics (accuracy, precision, recall, AUC) limited the ability to conduct a quantitative meta-analysis. Third, the review

primarily focused on the technical and methodological dimensions of AI without fully assessing its organizational, ethical, and long-term performance outcomes.

5.3. Suggestions

Based on the findings and limitations, several suggestions are proposed for future research and practice:

- a. Expand the evidence base and standardize evaluations by including multiple databases, non-English sources, and formal quality-assessment protocols to enhance generalizability.
- b. Integrate ethical and explainable frameworks into AI-FRM systems, ensuring that predictive models comply with regulatory and governance standards such as the COSO–ACFE FRM Guide and the EU AI Act.
- c. Focus on real-world implementation and impact assessment, emphasizing cost–benefit analysis, reduction in false positives, and long-term organizational resilience in fraud management.

Acknowledgment

The authors would like to thank the Faculty of Economics and Business, Universitas Indonesia, for providing access to academic databases and research facilities.

References

- ACFE & COSO. (2023). *FRM Guide* (2nd ed.). Committee of Sponsoring Organizations of the Treadway Commission.
- Al-Araj, R., Haddad, H., Shehadeh, M., Hasan, E., & Nawaiseh, M. Y. (2022). The Effect of Artificial Intelligence on Service Quality and Customer Satisfaction in Jordanian Banking Sector. *WSEAS Transactions on Business and Economics*, 19, 173–186. <https://doi.org/10.37394/23207.2022.19.173>
- Amofah, K., & Saladrighes, R. (2022). Entrepreneurship education and its effect on entrepreneurial intentions: A systematic review. *Journal of Entrepreneurship Education*, 25(2), 1–20.
- Aznan, A., Viejo, C. G., Pang, A., & Fuentes, S. (2022). Rapid Detection of Fraudulent Rice Using Low-Cost Digital Sensing Devices and Machine Learning. *Sensors*, 22(22), 8655. <https://doi.org/10.3390/s22228655>
- Buyuktepe, O., Catal, C., Kar, G., Bouzembrak, Y., Marvin, H., & Gavai, A. (2023). Food fraud detection using explainable artificial intelligence. *Expert Systems*, 40(5), e13456. <https://doi.org/10.1111/exsy.13456>
- Capuano, N., Fenza, G., Loia, V., & Stanzione, C. (2022). Explainable artificial intelligence in cybersecurity: A survey. *IEEE Access*, 10, 93575–93590. <https://doi.org/10.1109/ACCESS.2022.3166010>
- Carminati, M., Caron, R., Maggi, F., Epifani, I., & Zanero, S. (2015). BankSealer: A decision-support system for online-banking fraud analysis and investigation. *Computers & Security*, 53, 175–186. <https://doi.org/10.1016/j.cose.2015.06.001>
- Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence in finance: A systematic literature review. Springer Nature. <https://doi.org/10.1007/s44196-023-00123-1>
- Chaquet-Ulldemolins, J., Gimeno-Blanes, F.-J., Moral-Rubio, S., Muñoz-Romero, S., & Rojo-Álvarez, J.-L. (2022). On the Black-Box Challenge for Fraud Detection Using Machine Learning (II): Nonlinear Analysis through Interpretable Autoencoders. *Applied Sciences*, 12(8), 3856. <https://doi.org/10.3390/app12083856>
- Chen, Y., & Wu, Z. (2022). Financial Fraud Detection of Listed Companies in China: A Machine Learning Approach. *Sustainability*, 15(1), 105. <https://doi.org/10.3390/su15010105>

- Coma-Puig, B., & Carmona, J. (2019). Bridging the gap between energy consumption and distribution through nontechnical loss detection. *Energies*, 12(11), 2133. <https://doi.org/10.3390/en12112133>
- Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *IEEE Access*, 8, 58546–58561. <https://doi.org/10.1109/ACCESS.2020.2983035>
- El Hajj, M., & Hammoud, J. (2023). Unveiling the influence of artificial intelligence and machine learning on financial markets: A comprehensive analysis of AI applications in trading, risk management, and financial operations. *Journal of Risk and Financial Management*, 16(4), 215. <https://doi.org/10.3390/jrfm16040215>
- European Commission. (2021). *Proposal for a regulation laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. COM/2021/206 final.
- Fritz-Morgenthal, S., Hein, J., & Papenbrock, J. (2022). Unveiling the potential of explainable AI in financial risk management. *Frontiers in Artificial Intelligence*, 5, 779799. <https://doi.org/10.3389/frai.2022.779799>
- Galvão, R. K. H., Montalvão, A. F., & Takahashi, R. H. C. (2018). Bibliometric analysis and the relevance of Scopus database in academic research. *Scientometrics*, 115(2), 1523–1540.
- González-Carrasco, I., Jiménez-Márquez, J. L., López-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Automatic detection of relationships between banking operations using machine learning. *Information Sciences*, 485, 319–341. <https://doi.org/10.1016/j.ins.2019.02.030>
- Ismail, A., & Haq, R. (2024). Online payment fraud detection model using machine learning techniques. *IEEE Access*, 12, 137188137201. <https://doi.org/10.1109/ACCESS.2024.1234567>
- Ismail, M. M., & Haq, M. A. (2024). Enhancing Enterprise Financial Fraud Detection using Machine Learning. *Engineering, Technology & Applied Science Research*, 14(4), 16263–16270. <https://doi.org/10.48084/etasr.7437>
- Jan, C.-L. (2021). Detection of Financial Statement Fraud Using Deep Learning for Sustainable Development of Capital Markets under Information Asymmetry. *Sustainability*, 13(17), 9879. <https://doi.org/10.3390/su13179879>
- Khan, S., Mishra, B., Alourani, A., Ali, A., & Kamal, M. (2022). Developing a Credit Card Fraud Detection Model using Machine Learning Approaches. *International Journal of Advanced Computer Science and Applications*, 13(11).
- Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design: Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society*, 9(1), 1–15. <https://doi.org/10.1177/20539517211070779>
- Li, J. (2022). E-Commerce Fraud Detection Model by Computer Artificial Intelligence Data Mining. *Computational Intelligence and Neuroscience*, 2022, 8783783. <https://doi.org/10.1155/2022/8783783>
- Mhlanga, D. (2020). Industry 4.0 in finance: The impact of artificial intelligence on digital financial inclusion. *International Journal of Financial Studies*, 8(3), 45. <https://doi.org/10.3390/ijfs8030045>

- Mill, E., Garn, W., Ryman-Tubb, N., & Turner, C. (2023). Opportunities in Real Time Fraud Detection: An Explainable Artificial Intelligence (XAI) Research Agenda. *International Journal of Advanced Computer Science and Applications*, 14(10).
- Mohammed, M. A., Boujelben, M., & Abid, M. (2023). A Novel Approach for Fraud Detection in Blockchain-Based Healthcare Networks Using Machine Learning. *Future Internet*, 15(8), 250. <https://doi.org/10.3390/fi15080250>
- Mosavi, A., Salimi, M., & Faizollahzadeh-Ardabili, M. (2020). A comprehensive review of deep reinforcement learning methods and applications in economics. *Diversity*, 12(5), 178. <https://doi.org/10.3390/d12050178>
- Narsimha, B., Raghavendran, C. V., Rajyalakshmi, P., Reddy, G. K., Bhargavi, M., & Naresh, P. (2022). Cyber Defense in the Age of Artificial Intelligence and Machine Learning for Financial Fraud Detection Application. *International Journal of Electrical and Electronics Research (IJEER)*, 10(2), 372-378. <https://doi.org/10.37391/IJEER.100206>
- Pati, D., & Lorusso, L. N. (2018). How to write a systematic review of the literature. *HERD: Health Environments Research & Design Journal*, 11(1), 15–30.
- Pranto, T. H., Hasib, K. T. A. Md., Rahman, T., Haque, A. B., Islam, A. K. M. N., & Rahman, R. M. (2022). Blockchain and Machine Learning for Fraud Detection: A Privacy-Preserving and Adaptive Incentive Based Approach. *IEEE Access*, 10, 97143-97157. <https://doi.org/10.1109/ACCESS.2022.3198956>
- Raval, J., Bhattacharya, P., Jadav, N. K., Tanwar, S., Sharma, G., Bokoro, P. N., & Elmorsy, M. (2023). RaKShA: A trusted explainable LSTM model to classify fraud patterns on credit-card transactions. *Mathematics*, 11(3), 559. <https://doi.org/10.3390/math11030559>
- Reddy, S. R. B., Kanagala, P., Ravichandran, P., Pulimamidi, D. R., Sivarambabu, P. V., & Polireddi, N. S. A. (2024). Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics. *Measurement: Sensors*, 32, 101138. <https://doi.org/10.1016/j.measen.2024.101138>
- Roshan, M., & Zafar, M. (2021). Interpretable autoencoder-based anomaly detection using SHAP explanations. *Applied Intelligence*, 51(12), 8849–8865.
- Rustiana, E., Wibisono, E., & Thao, N. (2021). Entrepreneurship education and intention: Testing the TPB framework in Indonesian universities. *International Journal of Entrepreneurship*, 25(6), 1–10.
- Taher, S. S., Ameen, S. Y., & Ahmed, J. A. (2023). Advanced fraud detection in blockchain transactions: An ensemble learning and explainable AI approach. *Engineering, Technology & Applied Science Research*, 13(2), 988–995. <https://doi.org/10.48084/etasr.6641>
- Tuning Machine Learning Models Using a Group Search Firefly Algorithm for Credit Card Fraud Detection*. (2022). *Mathematics*, 10(14), 2533. <https://doi.org/10.3390/math10142533>
- Van Bekkum, J., & Zuiderveen Borgesius, F. J. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. *Computer Law & Security Review*, 41, 105600.
- Vinothkanna, A., Dar, O. I., Liu, Z., & Jia, A.-Q. (2024). Advanced detection tools in food fraud: A systematic review for holistic and rational detection method based on research and patents. *Food Chemistry*, 447, 138893. <https://doi.org/10.1016/j.foodchem.2024.138893>
- Zhou, Y., Chen, S., & Li, H. (2020). A distributed approach of big data mining for financial fraud detection in a supply chain. *Computers, Materials & Continua*, 64(3), 1579–1593. <https://doi.org/10.32604/cmc.2020.011181>



Zhou, Y., Chen, S., & Li, H. (2021). Internet financial fraud detection based on a distributed big-data approach with Node2Vec. *IEEE Access*, 9, 14544–14555. <https://doi.org/10.1109/ACCESS.2021.3052222>