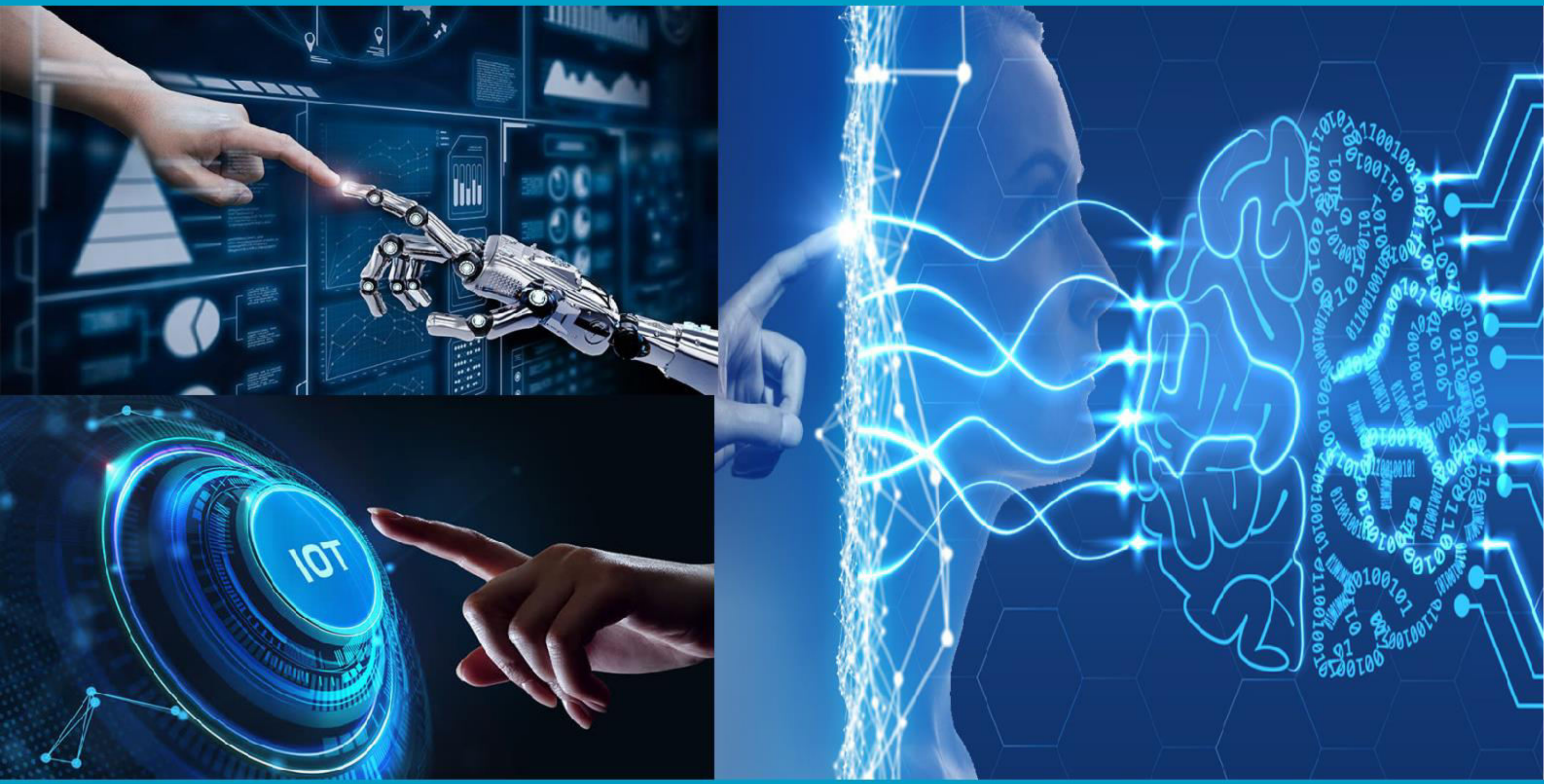
IJMRSETM

e-ISSN: 2395 - 7639



# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 11, Issue 7, July 2024

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.802

+91 99405 72462 +9163819 07438 ijmrsetm@gmail.com www.ijmrsetm.com

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.802| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 11, Issue 7, July 2024 |

# Advancements in Machine Learning for Financial Fraud Risk Assessment: A Financial Perspective

ARUSHI MEHTA

RESEARCH SCHOLAR, DEPARTMENT OF MANAGEMENT STUDIES, JAMIA MILLIA ISLAMIA, NEW DELHI, INDIA

**ABSTRACT:** The field of financial risk management is undergoing a significant transformation due to the advancements in artificial intelligence (AI) and the underlying machine learning (ML) techniques that provide the foundation of AI. These developments hold the potential to revolutionize the way the user's approach and address financial risk. The expansion of AI-driven solutions has opened up various opportunities for comprehending and managing risk. These opportunities encompass a wide range of activities, such as determining appropriate lending amounts for customers in banking, issuing warning signals to financial market traders regarding position risk, identifying instances of customer and insider fraud, enhancing compliance efforts, and mitigating model risk. The prime objective of this study is to investigate the application of AI and ML in the Financial Services industry, with a specific focus on Risk Management and Fraud Detection. This study proposes an intelligent and distributed approach for detecting Internet financial fraud using Big Data.

**KEYWORDS**-Financial fraud,Fraud detection, Risk assessment, Machine learning

## INTRODUCTION

Financial fraud is the act of gaining financial benefits by using illegal and fraudulent methods. Financial fraud can be committed in different areas, such as insurance, banking, taxation, and corporate sectors. Recently, financial transaction fraud, money laundering, and other types of financial fraud have become an increasing challenge among companies and industries. Despite several efforts to reduce fraudulent activities, its persistence affects the economy and society adversely, as large amounts of money are lost to fraud every day. Several fraud detection approaches were introduced many years ago. Most traditional methods are manual, which is not only time consuming, costly, and imprecise but also impractical. More studies are conducted to reduce losses resulting from fraudulent activities, but they are not efficient. With the advancement of the artificial intelligence (AI) approach, machine learning and data mining have been utilized to detect fraudulent activities in the financial sector. Both unsupervised and supervised methods were employed to predict fraud activities. Classification methods have been the most popular method for detecting financial fraudulent transactions. In this scenario, the first stage of model training uses a dataset with class labels and feature vectors. The trained model is then used to classify test samples in the next step.

This study attempts to identify machine-learning-based techniques employed for financial transaction fraud and to analyse gaps to discover research trends in this area. Recently, some reviews have been conducted to detect fraudulent financial activities. Delamaire et al.[11] conducted a review on different categories of fraudulent activities on credit cards, which include bankruptcy and counterfeit frauds, and suggested proper approaches to address them. Similarly, Zhu et al. [12] investigated ML methods for fraud transactions, which include the stock market and other fraud detection processes in financial sectors.Jenipher et al. [13] explored several ML approaches used for credit card fraud detection. Hashedi et al. [14] conducted a comprehensive survey to explore data mining and machine learning techniques to detect frauds in various aspects, including credit card fraud, insurance fraud, and telecoms subscription fraud.

Abdallah et al. [15]introduced a review to investigate different approaches for uncovering fraudulent activities in the health care domain based on statistical approaches. Popat and Chaudhary [16] presented an extensive review work on credit card fraud detection. The authors provide a detailed analysis of various ML classification methods with their methodology and challenges. Ryman-Tubb et al. [17] reviewed several state-of-the-art methods for detecting payment card fraudulent activities using transactional volumes. The study showed that only eight approaches have a practical implication to be used in the industry. A study by Albashrawi[18]analyzed several studies for one decade covering fraud detection in financial sectors using data mining techniques. However, this was not exhaustive and comprehensive enough as they ignored the method of evaluations and the pros and cons of data mining techniques, among others.

Despite several existing reviews in the field, however, most studies particularly focused on specific areas of finance, such as detecting credit card fraudulent activities, fraud in online banking, fraud in bank credit administration, and fraud in payment cards. This paper aims to identify financial fraud transactions based on machine learning methods and to discover datasets applied in the ML-based financial fraud detection. The study reviews existing machine learning (ML)-based methods applied for financial transaction fraud detection.

Fraud in financial statements involves forging financial reports to claim that a company is more profitable than usual, avoiding the payment of taxes, increasing stock prices, or obtaining a bank loan. It can also be regarded as the confidential records generated by organizations that contain their financial recordsthat comprise their expenses, profits made, income loans, etc. These statements also comprise some write-ups made by management for discussing business performances and predicted future tendencies. Different financial records provide the financial reality of the organization, which indicates how successful the organization is and assists in checking if the organization is bankable. In addition, financial statement fraudsters deceive the users of financial statements by correcting misstatements to make the organizations appear beneficial. The main purpose of the financial fraudulent statements is to enhance share prices, minimize tax liabilities, attract more investors as much as possible, and access personal bank loans among others.

### FinancialFrauds

Financial cyber fraud is a new term capturing the umbrella of crime committed over cyberspace for the sole purpose of illegal economic gain. Financial cybercrime perpetrators are difficult to identify. The fraudsters purposely mask their activities to blend their actions with the normal behavior of any other customer or user of a website or financial service; however, when grouped together, the activity is more obvious in terms of its abnormality. As technical skills and advancements in technology are increasingly available to criminals, their tactics for committing criminal offenses become more difficult to combat. This symbiosis of financial crime and cybersecurity is leading financial institutions to use their in-house developed methods to protect their assets using tools such as real-time analytics and interdiction to prevent financial loss. However, as models are showing signs of an inability to prevent and address these attacks, new methods must be developed and deployed across organizations to prevent further loss to their business, customer data, and their own reputation. The new methods deployed in the research community and industry include machine learning and deep learning models.

Apart from the above fraudulent activities committed in the financial sectors, other frauds are met in the financial domain, which includes commodities and securities fraud, mortgage fraud, corporate fraud, and money laundering. Securities and commodities fraud is a dishonest practice that occurs when a person invests in a company based on given fake information. A mortgage is a material misstatement made by a debtor at any stage of the application procedure when an underwriter relies on those facts to obtain a loan or credit. It intentionally targets documents associated with a mortgage by modifying information during the mortgage loan application processes. Another popular fraud is corporate fraud, which involves the falsification of financial documents by insiders to cover up any fraud or criminal activity. Money laundering is another type of financial fraud in which fraudsters try to change the source of illegal money by convincing criminals to turn their dirty money into legitimate money. Money laundering has a major influence on society because it is the primary method in which other crimes, such as funding terrorism and trade-in weapons, are accomplished. Another popular financial crimeis cryptocurrency fraud. This type of fraud systematically provides fake investments to naïve users in order to defraud them. The main idea of this is to entice innocent individuals with the promise of significant gains from their investments.

Machine learning operates in fraud detection by analyzing historical transactional data to identify patterns indicative of fraudulent activities. Supervised learning techniques, like logistic regression or decision trees, process labelled datasets, enabling models to discern features and behaviours associated with fraud. Additionally, unsupervised methods, such as anomaly detection, scrutinize outliers or irregularities in data, aiding in the detection of novel fraud instances. This amalgamation of algorithms enables adaptive, real-time identification of potentially fraudulent behaviour within financial transactions.

### Feature Engineering and Data Preprocessing

Data preprocessing is a critical stage in preparing data for effective machine learning models in fraud detection. Feature engineering involves selecting, transforming, or creating relevant features from raw data that capture the nuances of fraudulent behaviour. These features might include transaction amounts, frequency, timestamps, geographic locations, and user behavioural patterns.

Data preprocessing encompasses several steps, including normalization, outlier removal, and handling missing values. Normalization ensures that all features are on a consistent scale, preventing certain attributes from dominating the model. Outlier removal helps eliminate extreme values that might distort the learning process. Additionally, handling missing values involves imputation techniques to fill in or estimate missing data points, ensuring the integrity of the dataset.

### Anomaly Detection and Unsupervised Learning

Anomaly detection, a facet of unsupervised learning, is pivotal in uncovering irregularities or outliers within data without the need for labelled examples. In fraud detection, this technique scrutinizes deviations from expected behavior, flagging transactions or patterns that significantly differ from the norm.

Unsupervised learning algorithms like clustering, isolation forests, or autoencoders excel in this domain. Clustering methods group similar data points together, allowing for the identification of anomalies lying outside these clusters. Isolation forests isolate anomalies by constructing binary trees that efficiently distinguish normal from abnormal instances.

### Role of Neural Networks and Deep Learning

Neural networks, a foundational concept in deep learning, have revolutionized fraud detection in finance. Their ability to process complex, unstructured data makes them a potent tool for identifying subtle and intricate fraudulent patterns that evade traditional methods.

In fraud detection, neural networks, particularly deep architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), excel at processing diverse data types such as text, images, sequences, and time-series data.

CNNs, known for their prowess in image recognition, can extract intricate features from transactional data or metadata, aiding in anomaly detection by recognizing complex patterns.

RNNs, with their sequential learning abilities, prove invaluable in analyzing time-series data, and capturing temporal dependencies in transaction sequences or user behaviours. This capability allows for the detection of fraudulent activities that occur over time, enhancing the model’s predictive power.

### Real-Time Fraud Detection and Adaptive Models

Real-time fraud detection, enabled by adaptive machine learning models, revolutionizes the proactive identification of fraudulent activities in financial transactions. These models continuously analyze incoming data, instantly assessing its risk potential, and flagging suspicious behaviour in real-time.

Adaptive models leverage the concept of online learning, updating and evolving with each new data point. They dynamically adjust their detection strategies, learning from the most recent transactions to adapt to evolving fraud tactics.

This capability is crucial in the fast-paced financial landscape, where fraudulent activities constantly evolve. By swiftly adapting to new patterns and behaviors, these models can effectively stay ahead of emerging threats, reducing the window of vulnerability for financial institutions and enhancing their ability to prevent fraudulent transactions.

Despite its effectiveness, deploying ML in fraud detection poses challenges. Model interpretability, bias in algorithms, and the balance between false positives and false negatives are critical considerations. Moreover, ensuring data privacy and complying with regulations while handling sensitive financial data remains a priority.

### Future Trends and Innovations

Future trends in fraud detection within the financial sector are poised to be driven by advancements in machine learning and innovative technologies.

Explainable AI (XAI): The push for more transparent and interpretable AI models will continue. Explainable AI techniques aim to elucidate the decision-making process of complex models, fostering trust and understanding among stakeholders.

Federated Learning: Collaborative model training without sharing sensitive data will gain traction. Federated learning enables multiple institutions to jointly train models while keeping their data decentralized, enhancing privacy and security.

Blockchain Technology: The integration of blockchain offers immutable and transparent transaction records. Its implementation in financial systems can enhance security and trust by preventing tampering with transaction histories.

Machine learning has significantly transformed the landscape of fraud detection in the financial sector. Its ability to analyze vast amounts of data, detect intricate patterns, and adapt in real-time has made it an indispensable tool for financial institutions. As the financial landscape continues to evolve, the integration of machine learning with advanced analytics will play a pivotal role in staying ahead of fraudulent activities, ensuring the security and trust of financial systems.

In today's rapidly evolving financial landscape, the rise of digital transactions and technological advancements has brought exciting opportunities and unprecedented challenges.As the world becomes increasingly interconnected, the threat of fraud looms large, demanding robust and innovative solutions. This is where Artificial Intelligence (AI) steps into the limelight, promising to revolutionize transaction fraud detection and bolster security measures. Mastercard and Fintech Nexus surveyed financial institutions (FIs) to learn their opinions about AI, how they use it, the latest digital payment rails, and their technology needs and challenges.

## CONCLUSION

It is important to note that the remarkable advancements in machine learning are made possible by, and otherwise depend on, the emergence of big data. The ability of a computer algorithm to generate useful solutions from the data relies on the existence of a lot of data. More data means more opportunity for a computer algorithm to find associations. As more associations are found, the greater the accuracy of predictions. Just like with humans, the more experience a computer has, the better the results will be. The trial-and-error approach to computer learning requires an immense amount of computer processing power. It also requires specialized processing power, designed specifically to enhance the performance of machine learning algorithms.

## REFERENCES

[1] The Securities and Exchange Commission.
[2] SEC Speech, Has Big Data Made us Lazy?, Midwest Region Meeting of the American Accounting Association, October 2016. https://www.sec.gov/news/speech/bauguess-american-accounting-association-102116.html.
[3] http://cfe.columbia.edu/files/seasieor/center-financial-engineering/presentations/MachineLearningSECRiskAssessment030615public.pdf.
[4] Arthur Samuel, 1959, Some Studies in Machine Learning Using the Game of Checkers. IBM Journal 3, (3): 210-229.
[5] Gideon Lewis- Kraus, The New York Times, December 14, 2016, The Great A.I. Awakening.
[6] http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf.
[7] G. Hoberg and C. Lewis, 2017, Do Fraudulent Firms Produce Abnormal Disclosure? Journal of Corporate Finance, Vol. 43, pp. 58-85.
[8] Loughran, Tim, and McDonald, Bill, 2011. When is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. Journal of Finance 66: 35–65.
[9] https://www.sec.gov/divisions/marketreg/rule613-info.htm.
[10] Securities and Exchange Commission Strategic Plan Fiscal years 2014-2018, https://www.sec.gov/about/sec-strategic-plan-2014-2018.pdf.
[11]Delamaire, L., Abdou, H. A. H., & Pointon, J. (2009). Credit card fraud and detection techniques: a review. Banks and Bank systems, 4(2).
[12] Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in post-pandemic era. The Innovation, 2(4).
[13] Jenipher, V. N., Rose, J. D., Sabharam, M., & Nithin, M. (2021, November). Learning Algorithms with Data Balancing in Credit Card Fraud Detection Application. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 1-6). IEEE.
[14]Al-Hashedi, K. G., &Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 40, 100402.[15]Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. Journal of Network and Computer Applications, 68, 90-113.
[16]Popat, R. R., & Chaudhary, J. (2018, May). A survey on credit card fraud detection using machine learning. In 2018 2nd international conference on trends in electronics and informatics (ICOEI) (pp. 1120-1125). IEEE.

[17]Ryman-Tubb, N. F., Krause, P., & Garn, W. (2018). How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. Engineering Applications of Artificial Intelligence, 76, 130-157.

[18] Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: A decade review from 2004 to 2015. Journal of Data Science, 14(3), 553-569.

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com

www.ijmrsetm.com