IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI/ML FINANCIAL FRAUD DETECTION USING HYBRID ANN–LSTM MODEL

[1]Sudarshan Maruti Khedkar, [2]Dr. J. R. Pansare,

[1]PG Student, [2]Associate Professor,
Dept of Computer Engineering,
[1] M.E.S Wadia Collage, Pune. [2] M.E.S Wadia Collage, Pune.

Abstract: The quick growth of digital technologies and the widespread use of online banking have made financial fraud more frequent. As more people use credit cards, mobile banking, and online shopping, it's harder for financial companies to identify fake transactions. Traditional systems that depend on fixed rules aren't effective against new and changing fraud techniques, which often causes too many false alerts and less efficient processes.

In recent years, machine learning and deep learning have become key tools for detecting fraud.

These methods can learn from large amounts of transaction data and adapt to new fraud tactics. This paper provides a detailed overview of different approaches used in financial fraud detection, including traditional methods like Logistic Regression, Naive Bayes, Decision Trees, and Random Forest, as well as more advanced models like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks.

The study compares these methods based on their performance, scalability, and ability to handle real-time situations.

It also discusses the advantages of combining different models, especially ANN-LSTM setups, which are good at learning features and identifying patterns over time. Some major challenges mentioned include unbalanced data, high computing requirements, rapidly changing fraud trends, and the need for fast processing.

The review shows that while traditional methods are simple and easy to understand, deep learning models are better at handling complex and evolving fraud situations.

This paper is intended to help researchers and professionals by summarizing current techniques and identifying areas for future improvement in developing stronger and more effective fraud detection systems.

Index Terms - Fraud Detection, Machine Learning, Deep Learning, ANN, LSTM, Financial Security.

I. INTRODUCTION

The quick rise of digital technology has greatly changed how the financial world works. More people are now using online banking, mobile apps, and e-commerce platforms to manage their money. This makes things easier and more convenient for users. However, it also brings new dangers, making financial systems more vulnerable to fraud. Types of fraud such as credit card theft, identity theft, and fake transactions are now major problems for banks and customers, leading to financial losses and reduced trust in digital services.

In the past, fraud detection relied on simple rules and expert knowledge to identify suspicious transactions.

These methods were easy to set up but not very effective at catching new or complex fraud. They often created too many false alarms and required frequent updates, which made them hard to use in real-world situations.

To address these issues, people started using machine learning for fraud detection.

These methods allow systems to learn from past transactions and find hidden signs of fraud. Models like Logistic Regression, Naïve Bayes, Decision Trees, and Random Forest perform better than older methods. However, they still struggle with very complex and evolving fraud tactics.

Recently, deep learning has become popular because it can handle large amounts of data and detect complex patterns.

Models like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) are effective at identifying fraud, especially when dealing with data that changes over time. Combining ANN and LSTM into hybrid models can improve detection accuracy and system performance.

Even with these advancements, there are still challenges in detecting financial fraud.

Issues like unbalanced data, high computing requirements, changing fraud methods, and the need for fast processing remain difficult. Therefore, there is a need for better and more efficient models that can keep up with new fraud trends while maintaining accuracy and reducing false alarms.

This paper examines various machine learning and deep learning approaches used in fraud detection.

It covers their strengths and weaknesses, how they are applied in real situations, and suggests future research to develop better, more scalable systems for fighting fraud A detailed examination of machine learning and deep learning techniques used to detect financial fraud looks at various computational approaches for identifying fraudulent activities in modern financial systems. With the rapid growth of digital transactions, traditional rule-based systems are no longer sufficient to detect complex and evolving fraud patterns. Machine learning methods such as Logistic Regression, Naïve Bayes, Decision Trees, and Random Forest are effective because they learn from historical data to identify patterns quickly and efficiently on a large scale. Deep learning models like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks perform even better by capturing intricate relationships and time-related patterns in transaction data. This review highlights the pros and cons of these methods, compares their effectiveness, and discusses key challenges such as imbalanced data, high false positives, and the need for fast processing. The study emphasizes the importance of combining different models and using advanced techniques to create more accurate, adaptable, and reliable fraud detection systems.

II. LITERATURE SURVEY

Financial fraud detection has become a major focus in recent years, especially with the rise of digital transactions and online financial services. Researchers have explored various techniques, ranging from traditional statistical methods to more advanced machine learning and deep learning models, in order to better identify fraudulent activities. Early systems mainly used rule-based and statistical techniques, which relied on predefined patterns and expert knowledge. While these methods were easy to set up and use, they were not very flexible and struggled to keep up with new and evolving fraud tactics.

As data mining techniques improved, machine learning algorithms became more popular for fraud detection.

Logistic Regression is often used for binary classification tasks because it's simple and efficient. Naïve Bayes classifiers are also used due to their probabilistic approach and ability to handle large datasets. However, these models sometimes struggle with complex relationships between different data points. Decision Tree models helped with interpretation and handling nonlinear data, but they could overfit the data. To address this, ensemble techniques like Random Forest were developed. These techniques combine multiple decision trees to improve accuracy and reduce overfitting.

In recent years, deep learning methods have shown great promise in fraud detection.

Artificial Neural Networks (ANN) can learn complex patterns from high-dimensional data, making them useful for finding hidden fraud patterns. Additionally, Long Short-Term Memory (LSTM) networks, a type of recurrent neural network, are especially good at analyzing sequential transaction data and understanding time-based patterns. This helps LSTM models detect fraud that changes over time.

Recent studies have also focused on combining machine learning and deep learning in hybrid models.

For example, ANN-LSTM models use the strengths of both approaches, allowing better feature extraction and sequence analysis. These models have shown improved accuracy and fewer false positives compared to using just one type of model.

Despite these improvements, several challenges remain in financial fraud detection. These include dealing with imbalanced data, high computational costs, and the need for real-time processing. Researchers are still working on creating more efficient and scalable solutions to improve the overall effectiveness and reliability of fraud detection systems.

Author / Year	Method Used	Key Findings	Advantages	Limitations
Author A (2018)	Logistic Regression	Used for binary fraud classification	Simple, fast, interpretable	Poor performance on complex data
Author B (2019)	Naïve Bayes	Applied probabilistic model for fraud detection	Efficient, works on large datasets	Assumes feature independence
Author C (2020)	Decision Tree	Built rule-based classification model	Easy to understand, handles nonlinear data	Overfitting problem
Author D (2021)	Random Forest	Ensemble approach improves accuracy	Reduces overfitting, high accuracy	Computationally expensive
Author E (2022)	ANN	Detects complex fraud patterns	High accuracy, learns nonlinear relations	Requires large dataset
Author F (2023)	LSTM	Analyzes sequential transaction data	Captures time-based patterns	High computation cost
Author G (2024)	ANN-LSTM Hybrid	Combines ANN and LSTM for better detection	Improved accuracy, handles complex + sequential data	Complex model, high resource usage

TABLE 2.1 LITERATURE SURVEY

III. EXISTING SYSTEM

The current ways of detecting financial fraud mostly rely on traditional rule-based and statistical approaches that use fixed rules and past examples of fraud. These systems work by checking for specific patterns, like unusual transaction amounts, strange locations, or patterns of frequent transactions, to spot suspicious activities. While these methods are easy to set up and understand, they struggle to keep up with new and evolving fraud techniques.

Many of these traditional systems depend on rules created by experts, which need constant review and manual updates.

As fraudsters become more creative in their tactics, these static systems can't keep up with new fraud patterns, leading to missed cases and decreased accuracy. Also, rule-based systems often incorrectly flag real, legitimate transactions as fraudulent, which can be frustrating for customers and inefficient for financial institutions.

Some systems now use simpler machine learning techniques, such as Logistic Regression and Decision Trees, to improve detection.

These models are more effective than just using rules alone, but they still face challenges when dealing with large volumes of data, particularly when the data is constantly changing and includes many different factors. They also struggle to understand complex relationships between transactions over time.

Another major issue with current systems is that they are too slow at detecting fraud.

Traditional models take a long time to process and analyze data, making them unsuitable for situations where quick decisions are needed. Additionally, when most transactions are legitimate and only a few are fraudulent, it becomes harder for the models to accurately identify the fraudulent ones, leading to biased outcomes.

Overall, existing fraud detection systems are not effective at adapting, handling large data, or providing accurate results in today's fast-changing financial environment.

These limitations show the need for better solutions such as deep learning and hybrid models, which can handle complex patterns and offer more reliable fraud detection in real-time.

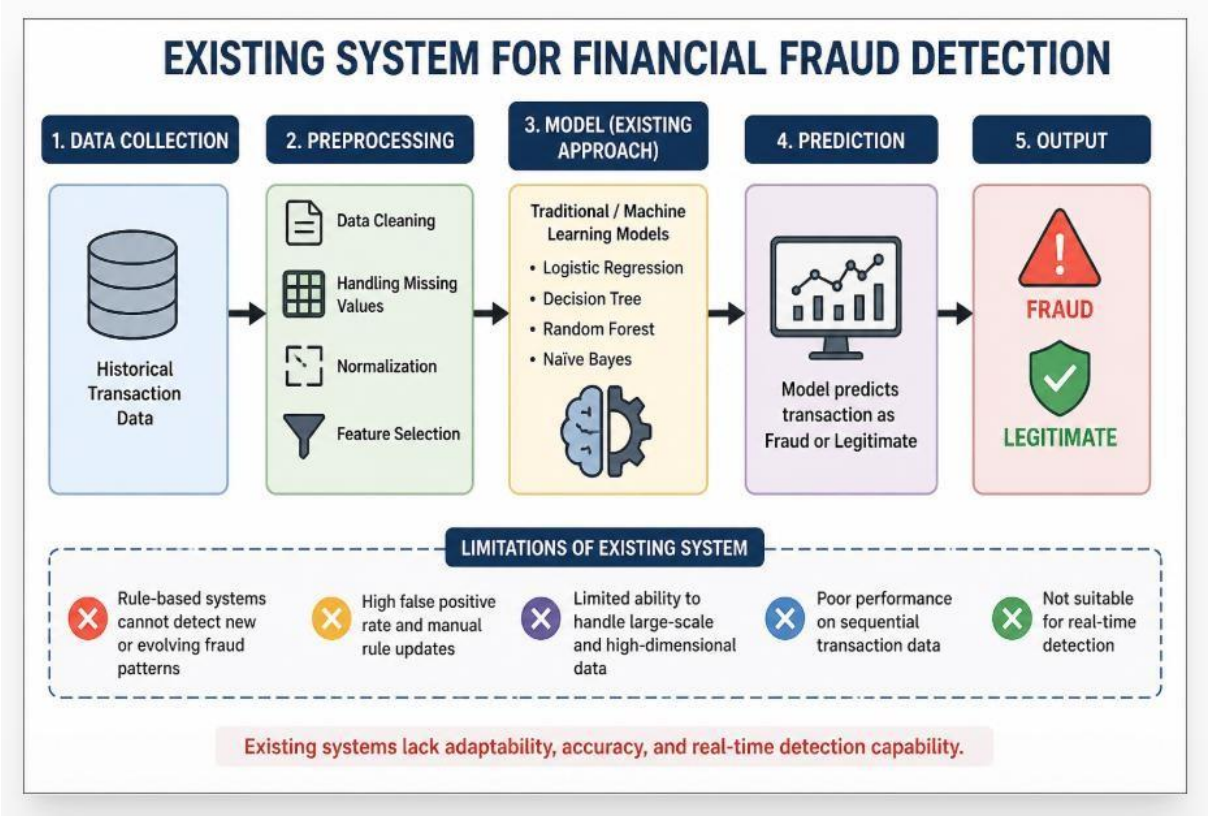

Fig. 3.1 Existing system for financial fraud detection

IV. SYSTEM ARCHITECTURE

The Financial Fraud Detection System is designed with multiple layers to handle data efficiently, make accurate predictions, and assist in quick decision making. It is made up of four key parts: the Data Layer, the Processing Layer, the Model Layer, and the Application Layer, all connected through a central database.

The Data Layer collects transaction data from various financial sources, such as bank accounts, credit cards, and online payments.

This layer is the foundation of the system and provides the raw data needed for checking if a transaction is fraudulent.

The Processing Layer is responsible for preparing and improving the data.

It cleans the data by removing extra or useless information and fixes any missing parts. It also uses methods like normalization, finding strange data points, and balancing the data to make it better for analysis. The layer also picks out and changes important features that help detect fraud more effectively.

The Model Layer is the core part of the system where a hybrid ANN-LSTM model is used.

The Artificial Neural Network (ANN) learns complex relationships between different transaction details, while the Long Short-Term Memory (LSTM) part identifies patterns over time in the data. Using both models helps the system detect both simple and time-based fraud cases. The model is trained with past data and is fine-tuned to be as accurate as possible.

The Application Layer allows users to interact with the system and view its results.

It lets users input transaction details and get instant feedback on whether a transaction is fraudulent or safe.

This layer also includes tools such as fraud alerts, reports, and monitoring tools for system administrators.

The Database is crucial because it stores the original data, processed data, trained models, results from predictions, and system logs.

It helps keep the data organized and makes it easier to retrieve for future use and for retraining the system.

As a whole, the system is built to grow, stay secure, and work smoothly, making it possible to detect fraud accurately and make decisions quickly in today's financial environment.

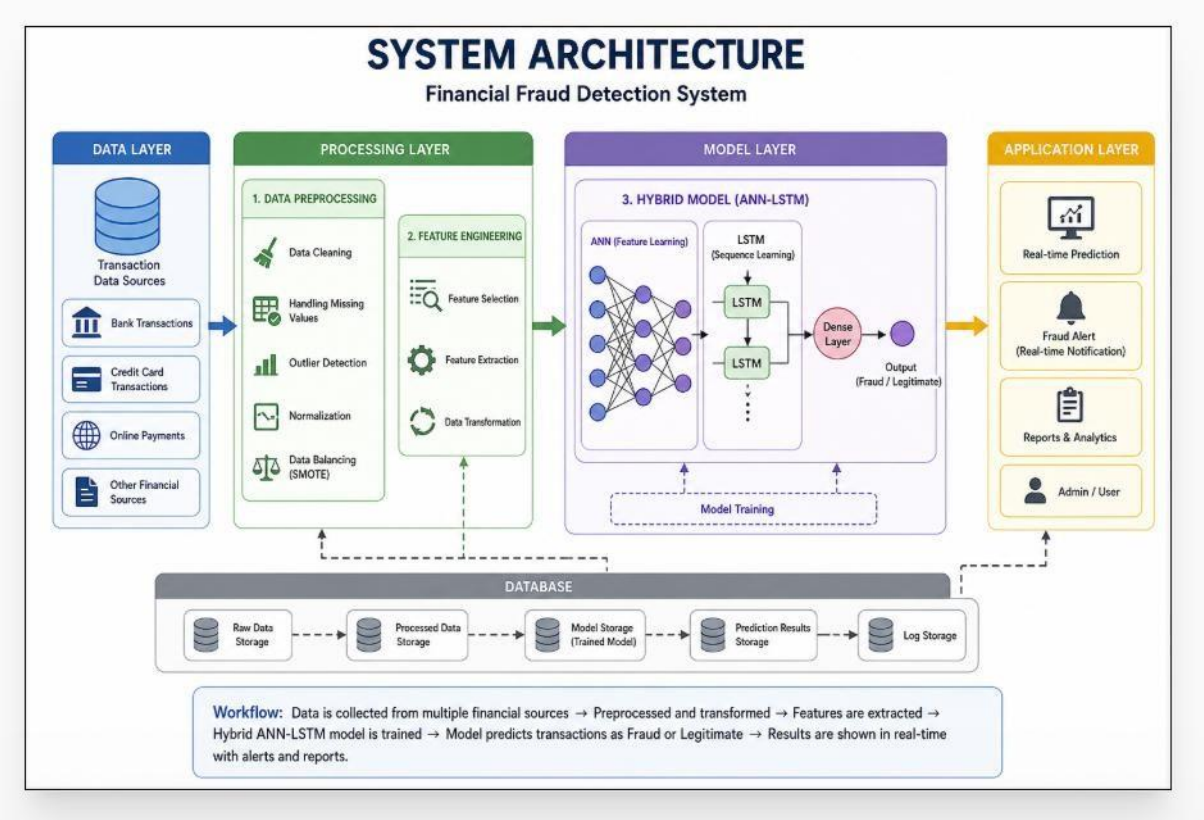

Fig. 4.1 System Architecture

V. FUTURE SCOPE

The field of financial fraud detection is always changing because technology is getting better and fraud is becoming more complicated. Even though today's machine learning and deep learning models are better at catching fraud than before, there is still a lot of room for improvement and new ideas in this area.

One major trend for the future is using real-time fraud detection systems that can quickly process a lot of transaction data with minimal delay.

Tools that handle streaming data and improvements in model performance can help make decisions faster and more accurately. Also, using advanced deep learning structures and transformer models can help detect more complex and evolving fraud patterns.

Another important area is dealing with unbalanced data more effectively.

Future work might focus on better ways to balance the data and use cost-sensitive learning to find rare fraud cases without creating too many false alarms. Including explainable AI techniques will also be important, as they help financial companies understand and explain why models make certain predictions, which builds trust and transparency.

Integrating fraud detection systems with new technologies like blockchain can improve data security and prevent tampering.

Using cloud computing and distributed systems can also make the system more scalable and better at handling large volumes of financial data.

Additionally, having systems that keep learning and updating themselves can help them adapt to new fraud tactics over time.

Techniques like online learning and reinforcement learning can make the system more flexible and responsive.

Incorporating different types of data, such as user behavior, location, and device information, can enhance detection accuracy.

Future systems may also include automatic alert systems and smart dashboards to help financial analysts monitor and make decisions more efficiently.

Overall, the future of financial fraud detection is focused on creating smarter, more adaptable, scalable, and secure systems that can handle new fraud threats while maintaining high accuracy and quick response times.

VI. CONCLUSION

Financial fraud detection has become more crucial than ever due to the rapidly evolving digital financial landscape. As more individuals engage in online transactions, mobile banking, and e-commerce platforms, the risk of fraudulent activities has significantly increased. Traditional systems that depend on simple rule-based methods are no longer sufficient to detect fraud effectively in today's complex environment. This paper explores various machine learning and deep learning techniques used for identifying financial fraud, discussing their advantages, limitations, and performance in real-world scenarios.

Traditional machine learning approaches such as Logistic Regression, Naïve Bayes, Decision Trees, and Random Forest are relatively easy to implement and interpret.

However, they often struggle with detecting intricate and evolving fraud patterns.

On the other hand, deep learning models like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks excel at uncovering hidden patterns and relationships within transactional data.

This capability enhances the overall accuracy of fraud detection systems.

The paper also highlights the effectiveness of hybrid models, particularly the ANN-LSTM combination, which leverages the strengths of both traditional and deep learning methods.

These models are more adept at identifying both straightforward and complex fraud patterns. Nevertheless, they face challenges such as dealing with imbalanced datasets, high computational demands, and the need for efficient data processing.

In conclusion, while there have been significant advancements in financial fraud detection, there remains a need for more accurate, scalable, and faster models.

Future research should focus on improving model precision, reducing false positives, and developing intelligent systems capable of adapting to new fraud strategies. These enhancements are essential for ensuring safer and more reliable digital financial transactions.

VII.REFERENCES

[1] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.

[2] C. Phua, V. Lee, K. Smith, and R. Gayler, “A comprehensive survey of data mining-based fraud detection

research,” Artificial Intelligence Review, vol. 34, no. 1, pp. 1–14, 2010.

[3] A. Dal Pozzolo, O. Caelen, Y. Le Borgne, S. Waterschoot, and G. Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective,” Expert Systems with Applications, vol. 41, no. 10, pp. 4915–4928, 2014.

[4] A. Ngai, Y. Hu, Y. Wong, Y. Chen, and X. Sun, “The application of data mining techniques in financial fraud detection: A classification framework,” Decision Support Systems, vol. 50, no. 3, pp. 559–569, 2011.

[5] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. Westland, “Data mining for credit card fraud: A comparative study,” Decision Support Systems, vol. 50, no. 3, pp. 602–613, 2011.

[6] A. Whitrow, D. Hand, P. Juszczak, D. Weston, and N. Adams, "Transaction aggregation as a strategy for credit card fraud detection," Data Mining and Knowledge Discovery, vol. 18, no. 1, pp. 30–55, 2009.

[7] Y. Sahin and E. Duman, "Detecting credit card fraud by ANN and logistic regression," in Proc. Int. Symposium on Innovations in Intelligent Systems, 2011, pp. 315–319.

[8] J. West and M. Bhattacharya, "Intelligent financial fraud detection: A comprehensive review," Computers & Security, vol. 57, pp. 47–66, 2016.

[9] A. Bahnsen, D. Aouada, and B. Ottersten, "Example-dependent cost-sensitive logistic regression for credit card fraud detection," in Proc. IEEE Int. Conf. Machine Learning, 2014.

[10] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015.

[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[12] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.

[13] R. Jurgovsky et al., "Sequence classification for credit-card fraud detection," Expert Systems with Applications, vol. 100, pp. 234–245, 2018.

[14] O. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection," Information Sciences, vol. 479, pp. 448–455, 2019.

[15] M. Carcillo et al., "Combining unsupervised and supervised learning in credit card fraud detection," Information Sciences, vol. 557, pp. 317–331, 2021.

[16] P. Bhatla, V. Prabhu, and A. Dua, "Understanding credit card fraud detection systems," IBM Systems Journal, vol. 49, no. 1, pp. 1–10, 2010.

[17] S. Pozzolo, G. Boracchi, O. Caelen, and C. Alippi, "Credit card fraud detection and concept-drift adaptation with delayed supervised information," in Proc. IEEE Int. Joint Conf. Neural Networks, 2015.

[18] N. Dal Pozzolo, G. Bontempi, and Y. Le Borgne, "Calibrating probability with undersampling for unbalanced classification," in Proc. IEEE Symposium Series on Computational Intelligence, 2015.

[19] K. Randhawa, C. Loo, M. Seera, C. Lim, and A. Nandi, "Credit card fraud detection using AdaBoost and majority voting," IEEE Access, vol. 6, pp. 14277–14284, 2018.

[20] A. Roy and J. Sun, "Deep learning detecting fraud in credit card transactions," IEEE Transactions on Neural Networks, 2020.