International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 31 Issue 05 – SEPTEMBER 2025.

Collaborative Learning Approach for Identifying Fraudulent Transactions Using Transformer Models

Prof. Nagamma [#1], Shamanth H R [*2]

[#] *HOD, Department of CSE (cyber security), Akash institute of engineering and technology, Bangalore*
[*] *Student, 4th Semester MCA, Akash Institute Of Engineering And Technology, Devanahalli, Bangalore*

Abstract— **Financial fraud cases causing serious damage to the interests of investors are not uncommon. As a result, a wide range of intelligent detection techniques are put forth to support financial institutions' decision-making. Currently, existing methods have problems such as poor detection accuracy, slow inference speed, and weak generalization ability. Therefore, we suggest a distributed knowledge distillation architecture for financial fraud detection based on Transformer. Firstly, the multi-attention mechanism is used to give weights to the features, followed by feed-forward neural networks to extract high-level features that include relevant information, and finally neural networks are used to categorize financial fraud. Secondly, for the problem of inconsistent financial data indicators and unbalanced data distribution focused on different industries, a distributed knowledge distillation algorithm is proposed. This algorithm combines the detection knowledge of the multi-teacher network and migrates the knowledge to the student network, which detects the financial data of different industries.**

***Index Terms*— keywords: CSRC, Transformer, migrate**

I. INTRODUCTION

The Financial fraud detection has become a significant concern for institutions and investors alike, as fraudulent activities continue to cause substantial financial losses. The rise of sophisticated fraudulent schemes and the increasing volume and complexity of financial data have made it challenging for traditional detection methods to remain effective. Financial fraud detection systems are critical tools that aim to detect, prevent, and mitigate fraud in various financial transactions and services. The traditional methods for detecting financial fraud often rely on statistical analysis, machine learning, or rule-based systems to identify suspicious activities. However, these methods have certain limitations, including poor accuracy, slow processing times, and difficulties in generalizing across diverse datasets. This has led to the development of more advanced methods, including deep learning, which can handle the complexity and high-dimensional nature of financial dataFinancial fraud detection has become a significant concern for institutions and investors alike, as fraudulent activities continue to cause substantial financial losses. The rise of sophisticated fraudulent schemes and the increasing volume and complexity of financial data have made it challenging for traditional detection methods to remain effective. Financial fraud detection systems are critical tools that aim to detect, prevent, and mitigate fraud in various financial transactions and services. The traditional methods for detecting financial fraud often rely on statistical analysis, machine learning, or rule-based systems to identify suspicious activities. However, these methods have certain limitations, including poor accuracy, slow processing times, and difficulties in generalizing across diverse datasets. This has led to the development of more advanced methods, including deep learning, which can handle the complexity and high-dimensional nature of financial data

II. LITERATURE SURVEY

C SVM Model for Financial Fraud Detection.

This paper proposes the use of Support Vector Machine (SVM) models for financial fraud detection. SVM, known for its robustness in classification tasks, is applied to financial data to detect anomalies and fraudulent activities. The paper examines various kernel functions and hyperparameter tuning techniques to optimize SVM's performance in this context. Experimental results show that SVM models outperform traditional methods in detecting financial fraud, especially when combined with feature scaling and cross-validation techniques. The study highlights the potential of SVMs to provide accurate and reliable fraud detection in financial transactions, which can help reduce financial losses and enhance decision-making in financial institutions.

This systematic review explores the advancements in financial statement fraud detection, focusing on the state-of-the-art methods and technologies. The paper categorizes existing approaches, ranging from traditional statistical methods to more recent machine learning and deep

learning techniques. It discusses the strengths and weaknesses of each approach, emphasizing the increasing use of artificial intelligence (AI) in automating fraud detection tasks. The review also identifies gaps in current research and suggests areas for future exploration, such as the integration of hybrid models and explainable AI to improve transparency and accuracy in detecting financial fraud.

This paper investigates the application of deep learning techniques to detect financial statement fraud in Chinese listed companies. By leveraging deep neural networks (DNNs), the study explores how these models can uncover complex, non-linear relationships in financial data that are often missed by traditional methods. The results show that deep learning models significantly outperform conventional algorithms like logistic regression in detecting fraudulent activities, demonstrating their ability to handle the intricacies of large-scale financial data. The study also proposes a framework for integrating deep learning models into the existing regulatory frameworks for more effective fraud detection. [2]

III. Existing System

Financial fraud detection has evolved over the years, from simple rule-based systems to more sophisticated machine learning (ML) and deep learning (DL) algorithms. Traditional financial fraud detection methods rely heavily on statistical and rule-based techniques, which, while effective for certain types of fraud, have significant limitations in addressing the complexities of modern financial systems. These conventional systems primarily focus on detecting fraudulent behavior based on predefined rules or thresholds, often lacking the ability to adapt to new and emerging fraud patterns.The most widely used approach in the existing system is based on supervised machine learning algorithms, such as logistic regression (LR), decision trees, support vector machines (SVM), and ensemble methods like random forests and gradient boosting machines. These techniques are typically applied to historical financial data, where features such as transaction amount, frequency, location, and time are analyzed to predict the likelihood of fraud. For example, logistic regression models classify financial transactions as either fraudulent or legitimate, depending on a set of features extracted from the data.

IV. Proposed System

To overcome the limitations of the existing fraud detection systems, this project proposes an advanced financial fraud detection framework based on Transformer networks and distributed knowledge distillation. The Transformer model, which has revolutionized fields like natural language processing (NLP) and computer vision, is known for its ability to capture long-range dependencies between features, making it well-suited for detecting complex fraud patterns in high-dimensional financial data. The proposed system uses the Transformer's multi-head attention mechanism to weigh the importance of different features in financial data, which helps in identifying relationships between disparate data points that may indicate fraud. This attention mechanism allows the model to dynamically focus on the most relevant features of the data while ignoring irrelevant information, improving the accuracy and efficiency of fraud detection.

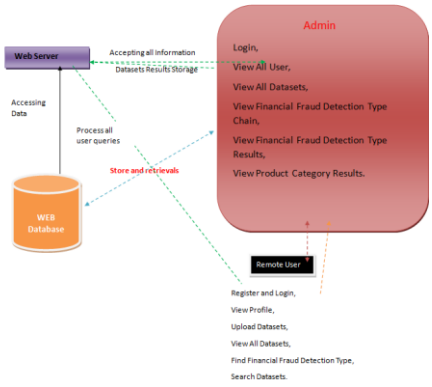

Fig: Architecture Digram

Advantages

A key component of the proposed system is the use of distributed knowledge distillation. In this approach, knowledge is transferred from multiple teacher models, each trained on different subsets of financial data from various industries, to a student model. This allows the student model to learn from a diverse range of fraud detection strategies and generalize better across industries. Knowledge distillation not only improves model performance but also reduces the complexity of the model, making it more suitable for deployment in real-time environments where computational resources are limited.

I. Implementation

The implementation phase of the financial fraud detection system is a crucial stage in the project lifecycle, where the theoretical designs and methodologies are transformed into a working system. This phase involves configuring the environment, developing the necessary software modules, integrating the different components, testing the system functionalities, and deploying it for operational use. The goal of the implementation stage is to ensure that the system can accurately detect financial fraud, provide actionable insights, and operate efficiently under real-world conditions. This chapter provides a detailed explanation of the implementation process, covering all the essential steps and considerations. The first step in the implementation process is setting up the development and production environments. The system requires a combination of software and hardware resources, including servers for data storage, high-performance processors for executing machine learning algorithms, and secure databases for maintaining transaction records. The development environment is typically configured with programming tools, libraries, and frameworks necessary for building machine learning models and integrating them with the application interface. Popular technologies for this system include Python for algorithm implementation, Flask or Django for backend development, MySQL for database management, and supporting libraries such as scikit-learn, TensorFlow, and PyTorch for machine

learning tasks. Setting up a robust and scalable environment ensures that the system can handle large datasets efficiently and perform real-time fraud detection.

II. RESULT

The Data security and privacy are prioritized during implementation. The system processes sensitive financial data that must be protected from unauthorized access, tampering, and breaches. Encryption protocols are implemented for both data at rest and data in transit. Access control mechanisms are established to restrict data visibility based on user roles. Compliance with data protection regulations such as GDPR and PCI DSS is ensured throughout the system's deployment. Security audits and penetration testing are conducted to validate the robustness of the system against potential cyber threats.

.

Fig: Resultant graph

III. CONCLUSION

The detection of fraudulent financial data in listed companies is of significant importance for safeguarding the interests of shareholders and investors. This paper proposes a distributed knowledge distillation framework based on Transformer for detecting fraudulent financial data in listed companies. Experimental validation was conducted using the dataset from the 9th ''TipDM Cup'' Financial Analysis Competition for Listed Companies. The performance of the proposed method was evaluated by comparing it with other advanced machine learning algorithms, including logistic regression, linear support vector machine, decision tree, random forest, XGBoost, and Adaboost. The experimental results demonstrate that the proposed method outperforms other machine learning algorithms, achieving the highest performance in terms of AUC, accuracy, precision, recall, and F1 score.

IV. REFERENCES

[1] C. Defang and L. Baichi, "SVM model for financial fraud detection," Northeastern Univ., Natural Sci., vol. 40, pp. 295–299, Feb. 2019.
[2] T. Shahana, V. Lavanya, and A. R. Bhat, "State of the art in financial statement fraud detection: A systematic review," Technological Forecasting Social Change, vol. 192, Jul. 2023, Art. no. 122527.
[3] W. Xiuguo and D. Shengyong, "An analysis on financial statement fraud detection for Chinese listed companies using deep learning," IEEE Access, vol. 10, pp. 22516–22532, 2022.
[4] M. N. Ashtiani and B. Raahemi, "Intelligent fraud detection in financial statements using machine learning and data mining: A systematic literature review," IEEE Access, vol. 10, pp. 72504–72525, 202. Bassiliades IEEE Trans. Intell. Transp. Syst., vol. 16, no. 4, pp. 1619–1635, August 2015, doi: 10.1109/TITS.2014.2376873.
[5] M. El-Bannany, A. H. Dehghan, and A. M. Khedr, "Prediction of financial statement fraud using machine learning techniques in UAE," in Proc. 18th Int. Multi-Conf. Syst., Signals Devices (SSD), Mar. 2021, pp. 649–654
[6] R. Cao, G. Liu, Y. Xie, and C. Jiang, "Two-level attention model of representation learning for fraud detection," IEEE Trans. Computat. Social Syst., vol. 8, no. 6, pp. 1291–1301, Dec. 2021.
[7] A. Singh, A. Singh, A. Aggarwal, and A. Chauhan, "Design and implementation of different machine learning algorithms for credit card fraud detection," in Proc. Int. Conf. Electr., Comput., Commun. Mechatronics Eng. (ICECCME), Nov. 2022, pp. 1–6.
[8] C. Liu, Y.-C. Chan, S. H. Alam, and H. Fu, "Financial fraud detection model: Based on random forest," in Econometrics: Econometric Model Construction, 201
[9] H. Shivraman, U. Garg, A. Panth, A. Kandpal, and A. Gupta, "A model framework to segregate clusters through K-means method," in Proc. 2nd Int. Conf. Comput. Sci., Eng. Appl. (ICCSEA), Sep. 2022, pp.