# Distributed Knowledge Distillation with Spatial-Temporal Graph Attention for Financial Fraud detection

*B. S.* Suresh[1*], *M.* Suresh[2], and *Karunakara* Rai. B[3]

[1]Department of Management Studies, St. Peter's Institute of Higher Education and Research, Chennai, India
[2]Department of Commerce, St. Peter's Institute of Higher Education and Research, Chennai, India
[3]Department of Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, Nitte (Deemed to be University), Bengaluru, India

**Abstract.** With current technological advancements, financial fraud cases that harm the interests of investors are becoming increasingly common. However, existing models face challenges due to the difficulty of adapting to evolving fraud patterns, inconsistent financial data and unbalanced data distribution. To overcome this problem, this paper proposed Distributed Knowledge Distillation with Spatial-Temporal Graph Attention (DKD-STGA) for financial fraud detection allows for highly effective and efficient identification of fraudulent activities-based Transformer. First, spatial-temporal graph attention used to intricate relationship across space and time by using graph and used to capture both spatial as well as temporal patterns. Next, neural networks are used to classify financial fraud and distributed knowledge distillation algorithm is used to overcome this complex structures, deep model depths and slow inference speeds. This algorithm gathers detection knowledge from multi-teacher network and transfers, it to a student network, which is used this shared knowledge to identify fraud in financial data from various industries. The experiments results show that the proposed DKD-STGA outperforms in terms of accuracy (98.97%), precision (99.89%), recall (97.99%) and F1-Score (98.90%) compared to existing method like transformer.

## 1 Introduction

In recent years, as technology has improved, most of the people using credit card for buying their needs, because of that fraud associated rising gradually [1]. Online fraud has become a major and difficult problem in the digital age, affecting all age peoples, businesses of different sizes and governments [2]. The increase in the number of vehicles and corresponding insurance policies has been paralleled by a rise in fraudulent claims, resulting in significant economic burden [3]. Cloud services are increasingly maintaining of their user's private information, such as personal details, browsing habits, and payment records, to provide a smooth user experience [4]. Financial fraud refers using fake information, altering records and manipulating markets by individuals or organizations to gain illegal financial advantages. These actions not only cause major financial losses but also threaten the stability of financial markets [5].

The rapid growth of cryptocurrency transactions has further increased the need for advanced fraud detection systems [6]. However, detecting fraud through financial statements often depends on human expertise and judgment, rather than a structured and scientific approach [7]. As fraud methods become more complex and diverse, identifying them is becoming harder. If fraud is not detected and stopped quickly, it can seriously harm the stakeholders of the affected companies and negatively impact capital markets [8]. The tricks used by fraudsters are now more advanced than traditional fraud detection methods, which mostly depend on fixed rules and manual checks [9]. To address this, deep learning models help to detect fraud by processing data through multiple nonlinear layers, though several other models are also available [10].

This research aims to overcome existing problems, by implementing Distributed Knowledge Distillation with Spatial-Temporal Graph Attention (DKD-STGA) for financial fraud detection allows for highly effective and efficient identification of fraudulent activities based on Transformer.

The main contribution of this paper as given as follows,

- Firstly, the spatial-temporal graph attention used to intricate relationship across space and time by using graph and used to capture both spatial as well as temporal patterns.
- Next, neural networks are used to classify financial fraud and distributed knowledge distillation algorithm is used to overcome this complex structures, deep model depths and slow inference speeds.
- This algorithm gathers detection knowledge from multi-teacher network and transfers, it to a student network, which is used this shared knowledge to identify fraud in financial data from various industries

* Corresponding author: bssuresh29@gmail.com

- The proposed DKD-STGA used two publicly available datasets such as IEEE-CIS fraud detection dataset and credit card fraud detection dataset, which help to detect fraud in finance.

The remaining of this research is organized as follows: Section 2 represent literature review, proposed methodology present in section 3, section 4 discuss about the result, section 5 deals with discussion and section 6 presents the conclusion.

## 2 Related work

The related work section examines existing Machine Learning (ML) and Deep Learning (DL) approaches for fraud detection in finance.

Yuxuan Tang and Zhanjunliu [11] implemented distributed knowledge distillation architecture based on transformer was implemented, which was used for financial fraud detection. First, multi-attention mechanism was used to highlight the most important features, then feed-forward neural networks were used to extract deeper information from relevant information and finally neural networks were used to classify whether financial fraud occurred. Next, to address the challenges of inconsistent financial data indicators and the unbalanced data distribution across different industries, a distributed knowledge distillation algorithm was proposed. However, existing model face challenges due to difficulty for models to adapt to evolving fraud patterns, inconsistent financial data and unbalanced data distribution.

Mohamed Akouhar et al. [12] introduced DL method that integrates Kolmogorov Arnold Networks (KAN) were integrated with dynamic oversampling and ensemble feature selection. The dynamic oversampling approach combines SMOTE and Generative Adversarial Networks (GANs) with flexible sampling rates, which helps to prevent overfitting and makes better handling of new unseen data. This approach provides a flexible and data-efficient solution to detect fraud in real-world, which helps to more resilient to data imbalance. However, the model restricts the broader generalizability of fraud detection framework because of sensitive and financial data, making it challenging to validate the model across different institutions and evolving fraud scenarios.

Ghosh [13] presented a hybrid model that combines Explainable Artificial Intelligence with Ridgelet Neural Network and Soft Gated Recurrent Unit (XAI RNN-SGRU). This model was designed for financial data prediction and cybersecurity, using advanced pre-processing techniques to manage missing information and address uneven data, as well as normalizing numerical features and encoding categorical data. For feature extraction and selection, improved fast random opposition-based learning Aphid Ant Optimization (AAO) and Enhanced Principal Component Analysis (EPCA) were applied. However, the XAI-RNN-SGRU model faces limitations owing to heavy data processing requirements and GRU-based sequential latency, which was incompatible with real-time fraud detection in low-resource environments.

Yuncan Tang, Yongquan Liang [14] explored Credit Card Fraud Detection (CCFD) model was developed, which combines Graph Neural Networks (GNN) and federated learning (FL). A framework was used to train CCFD models across different financial institutions, allowing them to jointly train while still keeping their data secure. In addition, A graph algorithm using weighted feature similarity was applied to convert each institution's local dataset into a transaction graph, which was then used for model training. However, the model faces certain limitations due to the virtual nature of the generated vertices, the model's judgment can be biased, potentially misinterpreting synthetic connections as genuine ones.

Rahul Kumar Gupta et al. [15] implemented an innovative ML based fraud detection approach was developed, incorporating advanced preprocessing techniques such as SMOTE-ENN used to address class imbalance, autoencoders for reduction of dimensionality and TOPSIS used for optimal feature selection. To improve prediction accuracy, a stacking ensemble model was built by combining K-Nearest Neighbors (KNN), Extreme Learning Machine (ELM) and Support Vector Machine (SVM). However, the proposed framework, despite its impressive performance in fraud detection, faces limitations in broader real-world applicability. Due to the lack of cross-domain validation and real-time optimization strategies, its effectiveness in other financial sectors such as insurance and healthcare fraud detection remains uncertain.

## 3 Proposed methodology

This research proposed the DKD-STGA for detecting fraud in finance. First, the data were collected from datasets. Next, pre-processing techniques such as normalization, dimensionality reduction and data splitting, were applied. Finally, the proposed method was used to detect fraud in the financial sector.

### 3.1 Data acquisition

All experiments are conducted using two publicly available datasets and a detailed description of the dataset is given below:

#### *3.1.1 IEEE-CIS fraud detection dataset*

This detection dataset contains 590,540 transactions and 20,663 transactions are categorize as fraudulent. Its main purpose is to identify fraudulent online transactions. This IEEE-CIS fraud detection dataset is divided into two main branches such as transaction table and identity table. This imbalance makes it harder for machine learning models to detect fraud effectively, so techniques like resampling or specialized evaluation metrics are needed. [16].

#### *3.1.2 Credit card fraud detection dataset*

This dataset includes 284,807 transactions, which only 492 considered as fraudulent. Each transaction includes

29 pre-processed, privacy-protected features along with a corresponding label. Because fraud cases make up only a very small fraction, the dataset is highly imbalanced, which can cause bias during model training. To address this, sampling is applied so that the ratio of normal to fraudulent transactions is adjusted to 3:1. [17].

## 3.2 Data pre-processing

After data collection, the data cannot be used directly in an detection because it leads to erroneous results. Therefore, the preprocessing steps are discussed below:

### *3.2.1 Handling missing values*

Given dataset had only small number of missing values, mean imputation was used to the numerical features. In this method, missing values are replaced with the column's average value, helping to preserve the dataset's overall distribution with minimal impact.

### *3.2.2 Normalization*

To standardize all numerical features to the same scale and min-max normalization is applied. This technique used to adjust feature values within the range of 0 to 1, Equation (1) give the explanation.

$$X_{normalized} = \frac{X_{standardized} - min}{max - min} \quad (1)$$

### *3.2.3 Dimensionality reduction with autoencoder*

An autoencoder is used to reduce the dimensionality, while preserving most important information. In Equation (2) represent, the encoding layer compresses input into a smaller and more compact representation:

$$Encoded\ representation = f_{relu}(W_e X + b_e) \quad (2)$$

### *3.2.4 Data splitting*

The dataset underwent various partitioning to ensure effective model training and evaluation. For IEEE-CIS fraud detection dataset, data divided into training for 70% and testing for 30%. Similarly, second dataset is split into three parts such as training for 70%, validation for 10% and testing for 20%.

## 3.3 Proposed method architecture

The architecture of the proposed DKD-STGA for financial fraud detection presented in Figure 1. The model consists of three modules. The first module is a multi-industry data processing module. The second module is a Transformer Encode Block module, which includes a spatial temporal graph attention module and feedforward neural network with fully connected. The feedforward network consists of a linear transformation followed by a ReLU activation function, together with a residual connection and layer normalization. The third module is the output neural network module, containing a linear neural network for output and a SoftMax function for result normalization.

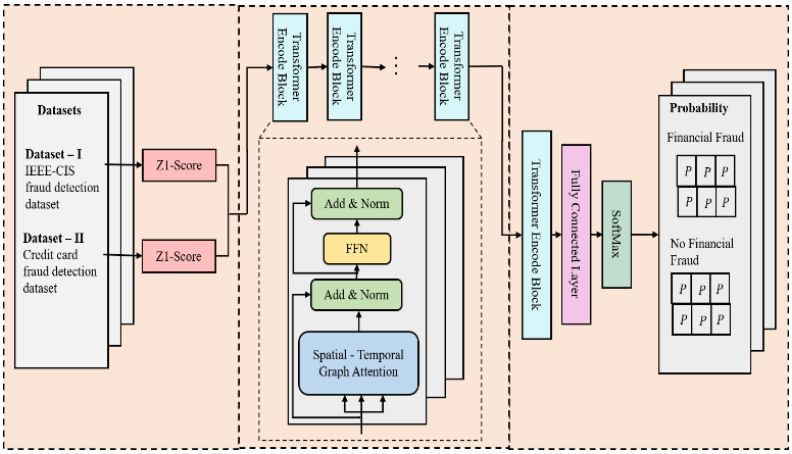


**Fig. 1.** Architecture of the proposed distributed knowledge distillation with spatial-temporal graph attention.

The flow of this architecture diagram is given below:

- First the data is collected from datasets such as IEEE-CIS fraud detection dataset and credit card fraud detection dataset. Next, pre-processing is applied.
- After finishing pre-processing, the output is passed through a sequence of Transformer Encoder Blocks. Each encoder block integrates a Spatial-Temporal Graph Attention mechanism, which effectively captures relationships between transactions and temporal dependencies over time.
- Next, the attention mechanism is used followed by Add & Norm layers and a Feed Forward Network (FFN) to refine feature representations.
- The output from the previous block is passed through to a fully connected layer, which condenses the learned features before passing them through a SoftMax classifier.
- Finally, the model uses a probability distribution distinguishing between Financial Fraud and No Financial Fraud.

### *3.3.1 Spatial and temporal graph attention*

The major advantages of attention mechanism is to calculate how nodes in the adjacency matrix depend on each other. This dependency is characterized by the dynamic interactions between nodes across both temporal and spatial domains.

An attention-based Transformer handles input data that comes in a sequence like $X \in R^{(N\times C)}$ , which $N$ represent element and $C$ denotes Channel. Time information an important role in understanding the relationship between space and time. The element integrated with time information, which used to remove difference between space and time. The sequential data as follows: $X \in R^{(N\times C)}$ , where $N = N \times T$.

The initial feature value of a node be $f_{(t,p)}$ presented in Equation (3),

$$Q^n_{(t,p)} = W^n_Q f_{(t,p)}, K^n_{(t,p)} = W^n_K f_{(t,p)}, V^n_{(t,p)} = W^n_V f_{(t,p)} \quad (3)$$

$K$ denotes key, $Q$ denotes query and $V$ denotes value. In this process, the attention mechanism first calculates the dot product between the query and key vectors within the same time frame, and then normalizes the result using a SoftMax activation function, as shown in Equations (4) and (5),

$$u^m_{(t,p)\in(t,q)} = \frac{(Q^m_{(t,p)}, K^m_{(t,p)})}{\sqrt{d}} \quad (4)$$

$$\alpha^m_{(t,p)\in(t,q)} = \frac{\exp(u^m_{(t,p)\in(t,q)})}{\sum_{n=1}^{N}\exp(u^m_{(t,p)\in(t,q)})} \quad (5)$$

Here, $<.>$ denotes dot product between two nodes and α denotes attention weights.

At this stage, masking operation was applied, which help to decision like attention model works in the temporal or spatial domain. For the spatial domain, all time-related information was blocked by assigning a value of 0 to the temporal edges during the masking process. Then, the attention head calculation is presented in Equations (6) and (7),

$$\bar{f}^m_{t,p} = \sum_{q=1}^{N}\left(\alpha^m_{(t,p)\in(t,q)}.V^m_{t,q}\right) \quad (6)$$

$$\bar{f}_{t,p} = Concat\left[\bar{f}^1_{t,p}, \bar{f}^2_{t,p}, \bar{f}^3_{t,p}, \dots, \bar{f}^H_{t,p}\right] \quad (7)$$

### *3.3.2 Distributed knowledge distillation detection framework*

On the one hand, due to the presence of various challenges related to financial data manipulation in different industries, there exist distinct characteristics and internal correlations in the financial data of different industries. Moreover, there are significant differences in the financial data indicators that different industries focus on. Therefore, it is challenging to use a universal model to detect financial data with such substantial variations. On the other hand, traditional models face challenges such as complex structures, deep model depths and slow inference speeds, making it difficult to deploy them in practical application scenarios. Based on the above problem considerations, this paper uses a distributed architecture to train multiple teacher detection models for multiple industries. On the one hand, the detection model is compressed to adapt to practical application scenarios, and on the other hand, the model in the case of unbalanced data distribution is improved. The distributed knowledge distillation detection frame work, as shown in Figure 1, is illustrated as follows. Firstly, datasets from various industries are prepared, and these datasets are utilized to train teacher models. Subsequently, untrained student models with simpler structures than the teacher models are prepared. A knowledge distillation algorithm is employed to transfer knowledge from multiple teacher models to the student network, which is then used to test financial data from different industries. The knowledge distillation algorithm is a model compression technique. It involves transferring knowledge from the teacher to the student, with the goal of preserving the teacher's performance while making the student model more lightweight and efficient.

For a detection task, the final output is the probabilities for each class, which are referred to as soft targets. The true labels for each sample are called hard targets. The difference with hard targets is that soft targets not only inform us about the most likely class for a sample but also provide probabilities for other classes, indicating that soft targets contain more information than hard targets. Therefore, when training the Multi-teacher model, we use hard targets. The predictions generated by the teacher network on a sample provide richer information to the student network. The distillation process is represented in Equations (8)-(11),

$$Softmax(Z/T) = \frac{\exp(z_i/T)}{\sum_j \exp(z_j/T)} \quad (8)$$

$$p^{(t)} = softmax(Y^{pre}_T/T) \quad (9)$$

$$p^{(s)} = softmax(Y^{pre}_S/T) \quad (10)$$

$$L_{KD} = \frac{T^2}{B}\sum_{i=1}^{B}\sum_{j=1}^{C} log\left(\frac{p^{(s)}_{i,j}}{p^{(t)}_{i,j}}\right) \quad (11)$$

The second step is to compute the student loss. This involves using a temperature softmax distiller (with T=1) on the output $Z^{(s)}$ to calculate soft targets $P$, and then calculating cross-entropy loss between $P^{(t=1)}$and the hard targets $Y_n$from the financial data using Equation (12),

$$L_{cls} = \frac{1}{B}\sum_{i=1}^{B}\sum_{j=1}^{C} -\left[Y_{i,j}.\log(i,j) + \left(1 - Y_{i,j}\right).\log\left(1 - P_{i,j}\right)\right] \quad (12)$$

The final knowledge distillation loss is obtained by taking the weighted sum of both distillation loss and student loss presented in Equation (13). The model utilizes data samples for training to update model parameters $W$ presented in Equation (14),

$$L_{tr} = \alpha.L_{cls} + \beta.L_{KD} \quad (13)$$

$$W = W + \eta . \frac{\partial L_{tr}}{\partial W} \quad (14)$$

Where, $\eta$ represents the learning rate and the proposed method results are discussed in section 4.

# 4 Result

This section describes the experiments conducted using multiple ML and DL models applied to detect financial fraud. Following the results are discussed in detail.

## 4.1 Experiments

All experiments are conducted on the hardware platform of 13th Gen Intel(R) Core (TM) i9-13900KF 3.00 GHz and NVIDIA GeForce RTX3060 Ti. The primary configuration environment for the experiment includes Python 3.9.1, torch 2.0.1, NumPy 1.22.4, and pandas 2.1.1. The Transformer-based financial fraud detection model was constructed using the PyTorch deep learning framework.

## 4.2 Evaluation metrics

To evaluate the performance of the detection of fraud in financial, the following parameters are considered:

### *4.2.1 Accuracy*

Accuracy is used to prediction of accurate number is divided by the total number of pr4edictions is used to measure the model's accuracy. The mathematical expression given below Equation (15).

$$Accuracy = \frac{TrueNegative + TruePositive}{TrueNegative + FalsePositive + TruePositive + FalseNegative} \quad (15)$$

### *4.2.2 Precision*

Precision used to assess a model's ability to accurately identify as positive class and calculated as ratio of true positives to both true and false positives. The mathematical formula is given below as Equation (16):

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (16)$$

### *4.2.3 Recall*

Recall used to measure the proportion of actual fraudulent transactions that a model correctly identifies out of all samples. The Equation (17) is given below:

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (17)$$

### *4.2.4 F1-Score*

F1-score is a harmonic mean of precision and recall, used to find balance between precision and recall and provides a single value that considers both false positives and false negatives. The mathematical Equation (18) given below:

$$FI - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (18)$$

## 4.3 Performance analysis

This section evaluates the performance of proposed method by comparing against existing models such CNN, GNN, RNN and LSTM. This comparison aims to demonstrates performance of the proposed method in terms of accuracy, precision, recall and F1-score. In first experiment, the first dataset compared with existing methods, which results are present in Table 1 and second dataset results are presented in Table 2.

**Table 1.** Performance analysis for proposed method using IEEE-CIS fraud detection dataset.

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| CNN | 90.45 | 93.56 | 97.54 | 90.75 |
| GNN | 92.76 | 91.99 | 94.35 | 92.46 |
| RNN | 91.57 | 94.67 | 95.87 | 96.34 |
| LSTM | 89.99 | 97.56 | 96,45 | 93.67 |
| Proposed DKD-STGA | 98.97 | 99.89 | 97.99 | 98.90 |

**Table 2.** Performance analysis for proposed method using credit card fraud detection dataset.

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| CNN | 90.34 | 97.27 | 95.74 | 94.74 |
| GNN | 93.76 | 96.83 | 93.89 | 98.34 |
| RNN | 95.65 | 98.34 | 98.45 | 99.37 |
| LSTM | 98.45 | 95.73 | 94.65 | 92.67 |
| Proposed DKD-STGA | 99.96 | 99.95 | 99.98 | 99.96 |

## 4.4 Computational complexity

This section used to evaluate the complexity of DKD-STGA, which higher compared to existing methods such as CNN, GNN, RNN and LSTM. The proposed DKD-STGA demonstrated 734 seconds, memory size also low and parameters. Table 3 represent the comparison of parameters, memory size and inference time across existing methods.

**Table 3.** Comparison of inference time and computational efficiency for financial fraud detection.

| Methods | Parameters (in millions) | Memory (in MB) | Inference Time (in seconds) |
|---|---|---|---|
| CNN | 13.7 | 0.29 | 12.56 |
| GNN | 9.6 | 0.39 | 23.83 |
| RNN | 10.3 | 0.56 | 16.54 |
| LSTM | 8.4 | 0.43 | 10.45 |
| Proposed DKD-STGA | 7.4 | 0.22 | 7.34 |

### 4.5 Statistical analysis

This section validates statistical, using Wilcoxon signed-rank based on p-value (<0.05) of 0.00150 indicates a statistically significant difference at the 0.01 significance level. The Wilcoxon signed-rank test demonstrated a statistically significant advantage of DKD-STGA over existing model (p = 0.00195).

### 4.6 Ablation study

Ablation experiments are essential for assessing the individual features that affect a model's performance. Based on the proposed DKD-STGA, generated three cases for the conduction of ablation study. The first case considers baseline STGA, next STGA with MHA and last STGA with CBAM. The output of this ablation study is presented in Table 4 and Table 5.

**Table 4.** Ablation study for proposed method for IEEE-CIS fraud detection dataset.

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| Baseline STGA | 93.76 | 96.36 | 95.87 | 94.67 |
| STGA-MHA | 92.82 | 94.67 | 96.46 | 96.57 |
| STGA-CBAM | 97.47 | 98.56 | 93.56 | 97.54 |
| Proposed DKD-STGA | 98.97 | 99.89 | 97.99 | 98.90 |

**Table 5.** Ablation study for proposed method for credit card fraud detection dataset

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| Baseline STGA | 93.6 | 95.73 | 94.54 | 94.76 |
| STGA-MHA | 96.76 | 98.52 | 97.45 | 93.56 |
| STGA-CBAM | 98.45 | 99.36 | 98.99 | 98.97 |
| Proposed DKD-STGA | 99.96 | 99.95 | 99.98 | 99.96 |

### 4.7 Comparative analysis

Tables 6 and 7 present a comparison of the proposed DKD-STGA with existing models such as KAN, FMC, XAI-RNN-SGRU and stacking ensemble. The proposed DKD-STGA demonstrates superior performance across publicly available datasets. In first IEEE-CIS fraud detection dataset, the proposed DKD-STGA achieved a high 98.97% for accuracy, 99.89% for precision, 97.99% for recall, and 98.90% for F1-score. Similarly, for the second dataset, the proposed method achieved an 99.96% of accuracy, 99.95% of precision, 99.98% of recall and 99.96% of F1-score.

**Table 6.** Comparative analysis of proposed method using IEEE-CIS fraud detection dataset

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| KAN [ 12] | 97.47 | 87.54 | 35.49 | 50.50 |
| FMC [ 14] | 81.60 | 69.23 | 61.02 | 64.86 |
| Proposed DKD-STGA | 98.97 | 99.89 | 97.99 | 98.90 |

**Table 7.** Comparative analysis of proposed method using credit card fraud detection dataset.

| Models | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) |
|---|---|---|---|---|
| XAI-RNN-SGRU [13] | 99.95 | 97.53 | 98.21 | 96.81 |
| FMC [14] | 95.99 | 93.81 | 91.10 | 92.44 |
| Stacking ensemble [15] | 99.95 | 99.93 | 99.97 | 99.95 |
| Proposed DKD-ATGA | 99.96 | 99.95 | 99.98 | 99.96 |

## 5 Discussion

The proposed DKD-STGA model exhibits high performance on publicly available datasets. The proposed DKD-STGA achieved an high accuracy (98.97%), precision (99.89%), recall (97.99%) and F1-score (98.90%) when the IEEE-CIS fraud detection dataset was used. Similarly, credit card fraud detection dataset is applied, the proposed method achieved high results such as accuracy (99.96%), precision (99.95%), recall (99.98%) and F1-score (99.96%), these improvements are attributed, knowledge distillation with spatial-temporal graph attention for financial fraud detection allows for highly effective and efficient identification of fraudulent activities based on Transformer. The proposed DKD-STGA was used to overcome this problem, including models to adapt to

evolving fraud patterns, inconsistent financial data, and unbalanced data distributions. The experiments were only tested using the Wilcoxon signed-rank test, which demonstrated a statistically significant difference between the proposed and the existing methods.

## 6 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 DKD-STGA based on Transformer for detecting fraudulent financial data in listed companies. Experimental validation was conducted using the dataset from IEEE-CIS fraud detection and credit card fraud detection dataset. The performance of the proposed DKD-STGA is evaluated by comparing with state-of-arts. The experiments results show that the proposed DKD-STGA outperforms compared to other deep learning algorithms, achieving superior performance in terms of accuracy, precision, recall and F1-score. In future, the DKD-STGA model use Adaptive knowledge distillation and integration with blockchain or federated learning can ensure secure, privacy-preserving detection. It extended for real-time fraud detection and improved cross-domain adaptability to handle diverse financial sectors.

## References

1. R.B. Asha, K.R.S. Kumar, Credit card fraud detection using artificial neural network, Global Transitions Proc. **2**, 35-41 (2021). https://doi.org/10.1016/j.gltp.2021.01.006
2. A. Papasavva, S. Lundrigan, E. Lowther, S. Johnson, E. Mariconti, A. Markovska, N. Tuptuk, Applications of AI-based models for online fraud detection and analysis, Crime Sci. **14**, 7 (2025). https://doi.org/10.1186/s40163-025-00248-8
3. A. Gheysarbeigi, M. Rakhshaninejad, M. Fathian, F. Barzinpour, An ensemble-based auto insurance fraud detection using BQANA hyperparameter tuning, IEEE Access **13**, 42997-43012 (2025). https://doi.org/10.1109/ACCESS.2025.3548529
4. K.P. Joshi, L. Elluri, A. Nagar, An integrated knowledge graph to automate cloud data compliance, IEEE Access **8**, 148541-148555 (2020). https://doi.org/10.1109/ACCESS.2020.3008964
5. J. Qian, G. Tong, Metapath-guided graph neural networks for financial fraud detection, Comput. Electr. Eng. **126**, 110428 (2025). https://doi.org/10.1016/j.compeleceng.2024.110428
6. A. Badjie, G.M. Ntuala, Q. Xia, J. Gao, H. Xia, MGGPT: A multi-graph GPT-enhanced framework for dynamic fraud detection in cryptocurrency networks, Comput. Netw. **270**, 111508 (2025). https://doi.org/10.1016/j.comnet.2025.111508
7. H. Su, I.M. Jiang, D. Liu, Detecting financial fraud risk using machine learning: Evidence based on different categories and matching samples, Finance Res. Lett. **85**, 107858 (2025). https://doi.org/10.1016/j.frl.2025.107858
8. J. Li, Corporate governance, fraud learning cycles, and financial fraud detection: Evidence from Chinese listed firms, Res. Int. Bus. Finance **76**, 102832 (2025). https://doi.org/10.1016/j.ribaf.2025.102832
9. O.I. Odufisan, O.V. Abhulimen, E.O. Ogunti, Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria, J. Econ. Criminol. **7**, 100127 (2025). https://doi.org/10.1016/j.joec.2025.100127
10. V.V.K. Reddy, R.V.K. Reddy, M.S.K. Munaga, B. Karnam, S.K. Maddila, C.S. Kolli, Deep learning-based credit card fraud detection in federated learning, Expert Syst. Appl. **255**, 124493 (2024). https://doi.org/10.1016/j.eswa.2024.124493
11. Y. Tang, Z. Liu, A distributed knowledge distillation framework for financial fraud detection based on transformer, IEEE Access **12**, 62899-62911 (2024). https://doi.org/10.1109/ACCESS.2024.3387841
12. M. Akouhar, M. Ouhssini, M. El Fatini, A. Abarda, E. Agherrabi, Dynamic oversampling-driven Kolmogorov–Arnold networks for credit card fraud detection: An ensemble approach to robust financial security, Egypt. Inform. J. **31**, 100712 (2025). https://doi.org/10.1016/j.eij.2025.100712
13. S. Ghosh, A novel framework for financial cybersecurity and fraud detection using XAI-RNN-SGRU, IEEE Access **13**, 88134-88155 (2025). https://doi.org/10.1109/ACCESS.2025.3570216
14. Y. Tang, Y. Liang, Credit card fraud detection based on federated graph learning, Expert Syst. Appl. **256**, 124979 (2024). https://doi.org/10.1016/j.eswa.2024.124979
15. R.K. Gupta, A. Hassan, S.K. Majhi, N. Parveen, A.T. Zamani, R. Anitha, B. Ojha, A.K. Singh, D. Muduli, Enhanced framework for credit card fraud detection using robust feature selection and a stacking ensemble model approach, Results Eng. **26**, 105084 (2025). https://doi.org/10.1016/j.rineng.2025.105084
16. The IEEE-CIS fraud detection dataset link: https://www.kaggle.com/c/ieee-fraud-detection/
17. The Credit card fraud detection dataset link: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud