SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Automated Fraud Detection In Financial Services Using Hybrid Autoencoders And LSTM

Rajesh Kumar Tripathi[1*], P. Parveen[2], K. Kanchana[3], Dr.K.G. Chikkegowda[4], Srilatha Gundapaneni[5], Dr.J. Dineshkumar[6]

[1*]Department of Computer Engineering & Applications, GLA University, Mathura, India. E-mail: rajesh.tripathi@gla.ac.in, https://orcid.org/0000-0003-3167-9338

[2]Assistant Professor, Department of IT, New Prince Shri Bhavani College of Engineering & Technology, Chennai, India. E-mail: parveenmansur2015@gmail.com, https://orcid.org/0009-0007-6161-4945

[3]Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, India. E-mail: kanchana@maher.ac.in, https://orcid.org/0009-0003-7865-9462

[4]Associate Professor, Cambridge Institute of Technology, Bengaluru, India. E-mail: chikkegowda00606@gmail.com, https://orcid.org/0009-0003-8583-7466

[5]Department of ECE, Ramachandra College of Engineering, Eluru, India. E-mail: drgsl@rcee.ac.in, https://orcid.org/0000-0002-3870-1369

[6]Assistant Professor, Mechatronics Engineering, Mahendra Engineering College, Namakkal, India. E-mail: dineshkumarj@mahendra.info, https://orcid.org/0000-0002-5776-0647

*Corresponding author: Email: rajesh.tripathi@gla.ac.in

Abstract

The growing sophistication of cyberattacks and transaction manipulation techniques has turned financial fraud into a big challenge in digital banking and electronic payment systems. Conventional fraud detection methods are not particularly successful in detecting new fraud patterns and sequential transaction anomalies. This research aims to present a hybrid autoencoder-LSTM model that could be used towards an automated fraud detection system for financial services based on the Kaggle Credit Card Fraud Detection Dataset. The proposed methodology is combining an autoencoder model for anomaly reconstruction analysis and sequential transaction behavior learning using a long short-term memory (LSTM) network. Firstly, the data set is subjected to preprocessing such as normalization and SMOTE (class balancing). The autoencoder identifies abnormal transaction patterns by analyzing reconstruction losses, and the LSTM network learns fraud patterns over time in sequences of transactions. The experimental results show that the proposed framework obtained an accuracy of 99.42%, a precision of 98.16%, a recall of 97.88%, an F1 score of 98.02%, and an ROC-AUC value of 99.31%. The other result of the confusion matrix analysis was that both the false positive and false negative rates were a little low, and this indicated that the proposed framework was valid in classifying frauds. The fraud detection capability and false alarm rate were also better compared to existing deep learning models. The results show that the proposed hybrid autoencoder-LSTM framework is efficient, scalable, and can be deployed in real time for intelligent financial fraud monitoring systems.

Keywords: Financial Fraud Detection, Autoencoder, LSTM, Deep Learning, Anomaly Detection, Credit Card Transactions.

1. Introduction

Financial institutions around the world are now using digital banking, mobile payments, fund transfers via electronic means, and online credit systems to make financial transactions. While these technologies enhance banking systems' efficiency and convenience for customers, also leave banking systems vulnerable to advanced attacks by cyber fraudsters. Rule-based systems and statistical analysis, traditional fraud detection approaches, are not able to keep up with the dynamic and changing nature of fraud in real time. However, with recent advancements in deep learning, intelligent fraud detection models have emerged that can detect hidden transaction anomalies and deviations in sequential behaviors. Recently, hybrid architectures, combining

autoencoders, convolutional neural networks, and recurrent learning models, have proven to be very powerful for anomaly detection, cyber threat analysis, and financial security enhancement [1]. It is also observed that the adaptive cascaded architecture of autoencoder-LSTM has successfully been used in intelligent decision-making systems, as it is able to process large-scale sequential data [2]. The study of advertisement fraud and deceptive communication also highlights the complexities of digital fraud ecosystems in today's information systems [3].

With the ability to constantly change attack methods, credit card and financial transaction fraud have become a huge issue. Fraud analytics using deep learning is becoming more and more popular and is able to detect unusual behavior in transactions without much manual feature engineering. Autoencoder and LSTM-integrated frameworks have demonstrated promising performance in detecting the hidden fraud patterns from the streams of transactions [4]. In insurance and financial sectors, the use of multi-stacked LSTM models along with adaptive oversampling methods has also boosted the classification accuracy of fraud [5]. Moreover, spatio-temporal autoencoders, along with convolutional LSTM networks, can capture both spatio-temporal dependencies to improve anomaly detection results under complex situations [6]. Outlier detection methods have also helped to reduce the number of false positives in credit card fraud detection systems [7]. Recent research with deep autoencoders and deep classifiers has illustrated that hybrid deep learning models can be used to greatly enhance the fraud prediction capability of standalone machine learning models [8].

The financial fraud detection systems today need intelligent architectures to process a large volume of high-dimensional and sequential transaction data in real time. To better represent fraud data and generate synthetic fraud to train the model, variational autoencoder generative adversarial models have been introduced to improve fraud data representation [9]. In addition, financial irregularities such as off-balance-sheet financing and hidden accounting manipulations make the demand for intelligent financial data fraud analytics in a corporate environment escalate [10]. Sequential fraud detection models have been created that make use of deep learning techniques to better analyze the temporal transaction behavior [11]. In the same way, two-stream CNN-LSTM models have proven to be very effective in cloud-based fraud detection scenarios that learn both feature and sequential fraud features [12]. It has also been suggested that blockchain could be used to create accounting and financial integrity systems to increase transparency and prevent fraudulent cross-border transactions in financial systems [13]. The developments suggest that in order to have robust fraud monitoring in the financial services sector, approaches to anomaly detection and sequential learning will be increasingly important.

Research Aim

To build a hybrid model of Autoencoder-LSTM to detect the fraud accurately in real time in financial services.

Research Objectives

1. To develop a hybrid deep learning model—an autoencoder and LSTM—for fraud detection.
2. To identify any unusual patterns of financial transactions through reconstruction error and sequence analysis of behavior.
3. To test the proposed model based on the fraud detection performance metrics like accuracy, precision, recall, and F1 score.

Research Questions

1. What are the advantages of the combination of autoencoders and LSTMs for fraud detection in financial services?
2. Are there potential sequential behavioral learning algorithms that can lower false positive fraud alerts in systems designed to monitor transactions?
3. What is the accuracy of anomaly detection techniques at detecting financial fraud patterns in an evolving and real-time scenario?

Paper Organization

The paper is subdivided into a number of sections. In Section 1 have the introduction, research objectives, and research questions. The literature review is discussed in Section 2, and the gap in existing research is identified. In Section 3 and 4, the proposed hybrid autoencoder-LSTM fraud detection framework and methodology are explained. Experimental analysis and performance evaluation results are shown in Section 5. Section 6 presents

findings and comparisons with previous work, and the last section concludes the study and presents ideas for future research.

2. Literature Review

A major thrust in recent developments of financial fraud detection has been towards deep learning and hybrid anomaly detection. Hybrid methods, which integrate both the generative adversarial networks (GANs) and sequential models, have shown to be very effective in detecting fraudulent behavior in transactions [14]. A series of systematic reviews in the field of cyber fraud detection shows that the machine learning and deep learning algorithms are far more successful in detecting fraud in high-dimensional and large-size financial datasets [15]. Financial systems based on blockchain have also been suggested to improve integrity, transparency, and trust in financial transactions and to minimize risks of manipulation [16]. Deep neural architectures have also further enriched predictive fraud analytics, as able to identify increasing transaction dependencies and unusual behavioral patterns with time series forecasting and anomaly detection [17].

LSTM-autoencoder architectures have been widely studied for anomaly detection in industrial and financial applications, as able to capture temporal relationships and reconstruction error efficiently at the same time [18]. Autoencoder models with attention mechanisms are able to better extract the features, thus increasing the accuracy of anomaly detection [19]. Explainable AI-based financial evaluation models have also helped in the explanation and transparency of financial fraud analytics in financial services [20]. The effectiveness of the combination of convolutional and sequential learning models for time series anomaly detection has also been demonstrated through D-CNN-LSTM autoencoder models, which were able to identify hidden anomalies in the dynamics of a system [21]. These studies give a hint about the significance of hybrid architectures in smart anomaly detection applications.

Hybrid autoencoder and LSTM frameworks have also been extensively used in other applications, including sentiment analysis, supply chain optimization, and industrial risk analysis. In complex datasets, the stacked autoencoder-based feature extraction combined with LSTM classification models has been effective in enhancing the performance for sequential pattern recognition [22]. In a high-tech industry, intelligent analytics can be used to optimize operational risk and strategic decisions on finance planning [23]. Ensemble deep learning with optimization algorithms has also helped in the development of sustainable fraud detection systems that are better at dealing with imbalanced financial datasets [24]. These studies collectively prove that hybrid deep learning approaches have great benefits in feature learning, anomaly detection, and sequential transaction analysis.

While there is extensive research in favor of fraud detection using deep learning models, there are some limitations that are yet to be sorted out. Current systems for fraud detection are based on anomaly detection (such as autoencoders) or sequential transaction analysis (such as LSTM networks) separately. There has been some limited work that combines both methods in a single approach that can simulate both reconstruction-based anomalies and transaction-based dependency anomalies. Moreover, current models have their limitations in the presence of class imbalance, shifting fraud techniques, and the need for real-time fraud detection. Moreover, a few studies fail to have any adaptive fusion methods for effectively combining anomaly scores and sequential behavioral predictions. Thus, there is a great demand for an integrated hybrid autoencoder-LSTM framework that would be able to deliver accurate, scalable, and real-time fraud detection for modern financial services.

3. Proposed Hybrid Framework

The proposed framework is a combination of an autoencoder-based anomaly detection mechanism and a Long Short-Term Memory (LSTM) network, which can be used to detect fraudulent money transactions from the Kaggle Credit Card Fraud Detection Dataset. The framework is based on the reconstruction-based anomaly analysis and sequential transaction behavior learning to boost the accuracy of fraud detection.

The transaction data is preprocessed, which involves missing value verification, normalization, feature scaling, and class balancing using SMOTE. The features (V1-V28) are anonymized, as well as the transaction amount, time features, and the fraud label. The preprocessed data of transactions are split into a training set and a test set.

The first step in the framework is to use an autoencoder model to learn the normal distribution of transactions. High-dimensional features of the transactions are encoded to latent features by the encoder, and then later, the original transaction pattern is reconstructed by the decoder. Reconstruction errors are typically larger for fraudulent transactions, as the transactions' patterns are more distinct from normal financial transactions. The loss in the reconstruction is obtained by mean square error (MSE) as shown in equation 1:

$$\text{Loss} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \qquad (1)$$

with x_i being the original transaction feature and $\hat{x}_i$ being the reconstructed feature.

The second stage involves using the LSTM network to model the temporal aspects of transaction behavior, recognizing hidden fraud patterns. The LSTM model is fed with the sequence of transactions and learns the irregularities occurring in the behavior of a fraudulent activity. The updating of the hidden state in the LSTM network is shown as equation 2:

$$h_t = \text{LSTM}(x_t, h_{t-1}) \qquad (2)$$

where x_t is the current transaction input and h_t is the hidden state at time t.

The combined outputs of the autoencoder and LSTM are combined in a hybrid decision layer to obtain the final score of fraud prediction. If the transaction amount is greater than the set amount (which is a set amount considered fraudulent), it will be marked as fraud; otherwise, it will be marked as normal.

Accuracy, precision, recall, F1-score, and ROC-AUC are used to assess the performance of the proposed framework. The hybrid architecture will enhance the fraud detection capability, lower false-positive rates, and aid in real-time financial fraud monitoring systems.

4. Methodology

The methodology proposed is the hybrid autoencoder-LSTM-based fraud detection of the financial transactions from the Kaggle Credit Card Fraud Detection Dataset. The combination of two approaches, anomaly detection and sequential transaction learning, is used in the framework to enhance the accuracy of fraud identification in financial services. The methodology involves six main steps: collecting the dataset, preprocessing the data, anomaly detection using an autoencoder, sequential analysis using LSTM, hybrid fraud classification using both an autoencoder and LSTM, and performance evaluation.

The overall methodology of the proposed Hybrid Autoencoder–LSTM framework for the detection of fraudulent financial transactions using the Kaggle Credit Card Fraud Detection Dataset is shown in Figure 1. The framework starts with collecting the datasets and pre-processing the data, which involves checking for missing values, normalization, feature scaling, class balancing using SMOTE, and splitting the dataset. The processed transaction data is then given to the Autoencoder module, where the anomaly detection is done by computing the values of reconstruction error in order to detect suspicious transactions. Later, the LSTM network is used to model the abnormal transaction sequences, which is able to capture the temporal dependency and sequential fraud behavior. This information from the autoencoder and LSTM module is combined in a hybrid classification module to arrive at the final fraud score, which is used to classify transactions as legitimate or fraudulent. Finally, the effectiveness of the proposed fraud detection model is measured by accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix analysis to measure the framework performance.

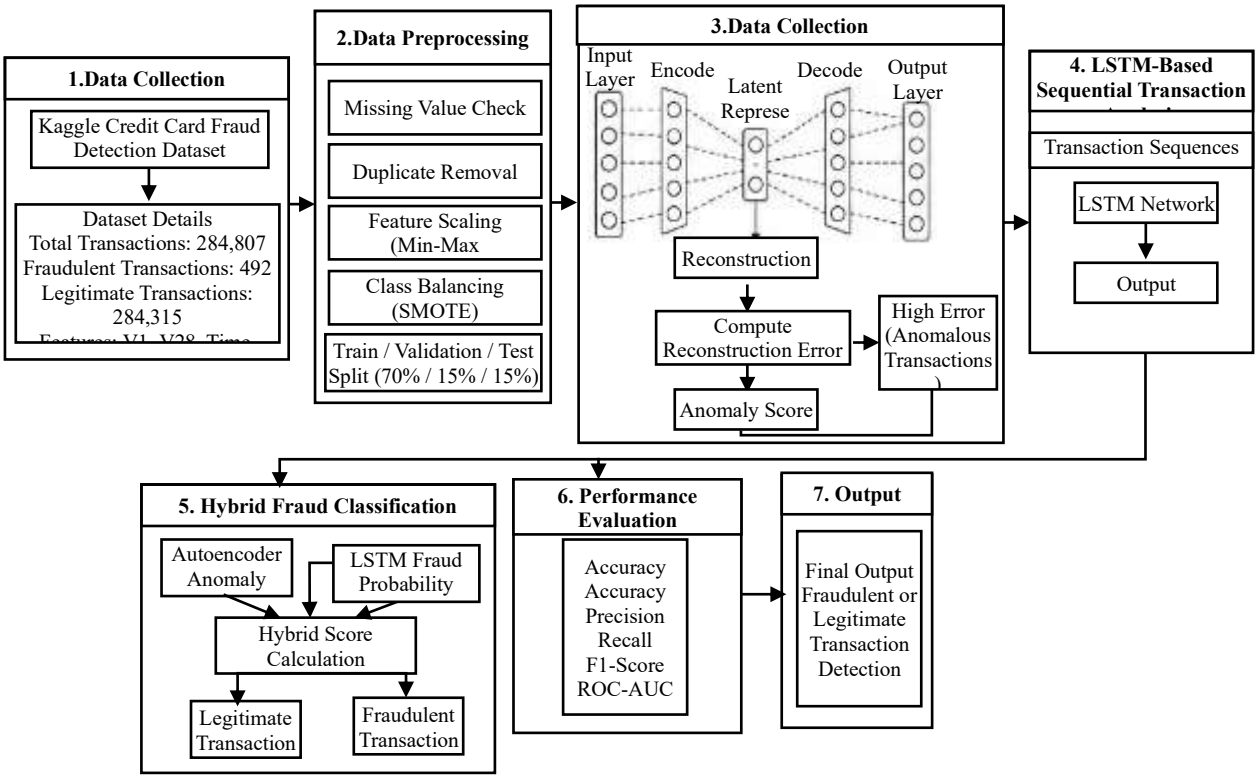

Figure 1. Proposed Hybrid Autoencoder–LSTM Methodology Framework for Financial Fraud Detection

Dataset Collection

The dataset used for the research is a Kaggle Credit Card Fraud Detection Dataset that has 284,807 financial transactions with 492 fraudulent transactions. All the attributes (V1-V28) are anonymous and PCA-transformed, and the data also contains transaction amount, transaction time, and class label of legitimate and fraudulent transactions.

Data Preprocessing

The collected dataset is preprocessed to make it higher quality and more efficient for fraud detection. Initially, missing value verification and removing duplicate transactions are done. To normalize the feature distributions, the amount of the transaction and time are normalized by min-max normalization. Equation (3) mathematically represents the normalization process.

Min-Max Normalization

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (3)$$

where:

- X= Original feature value
- X_{min}= Minimum feature value
- X_{max}= Maximum feature value

To obtain a normalized transaction feature, it is scaled into the range of 0-1 by equation (3) to enhance the convergence of the deep learning model. The data set is extremely imbalanced, and to balance the number of fraudulent and legitimate transaction samples, the Synthetic Minority Oversampling Technique (SMOTE) is used during training of the model.

The dataset is divided into:

- 70% training data

- 15% validation data
- 15% testing data

Autoencoder-Based Anomaly Detection

The first step in the fraud detection system is to use the autoencoder model to detect unusual transactions. The encoder maps high-dimensional features of the transactions into latent representations, and the decoder maps the latent representations back to the original transaction information. There are more likely to be larger reconstruction errors in fraudulent transactions because of abnormal transaction behavior.

The Mean Squared Error (MSE) was used to calculate the reconstruction loss as stated in Equation (4).

Autoencoder Reconstruction Loss

$$\text{Loss} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \qquad (4)$$

where:

- x_i= Original transaction feature
- $\hat{x}_i$= Reconstructed transaction feature
- n= Number of transaction features

The reconstruction difference, original and reconstructed transaction data, is measured by equation (4). Any transaction that gets reconstructed with an error that is above the set threshold is identified as a suspicious anomaly.

LSTM-Based Sequential Transaction Analysis

The LSTM network is used for sequential behavior analysis; the autoencoder generates anomalous transaction sequences to feed into the LSTM network. The LSTM model is able to model temporal dependencies in the transactions and detect hidden patterns of fraud in the course of time.

The update of the hidden state of the LSTM network is hidden in Equation (5).

LSTM Hidden State Update

$$h_t = \text{LSTM}(x_t, h_{t-1}) \qquad (5)$$

where:

- x_t= Current transaction input
- h_t= Current hidden state
- h_{t-1}= Previous hidden state

By doing so, equation (5) helps the LSTM network to keep long-term information about transactions' behaviors and effectively identify the sequential fraud patterns.

Hybrid Fraud Classification

The results of the autoencoder and LSTM module are combined in the hybrid decision layer to produce the final fraud prediction score. The total fraud score is obtained by adding the scores according to the Equation (6).

Hybrid Fraud Score

$$F_{score} = \alpha A_{score} + \beta L_{score} \qquad (6)$$

where:

- A_{score}= Autoencoder anomaly score
- L_{score}= LSTM fraud probability score
- $\alpha$and β= Weight coefficients

Thus, anomaly reconstruction information is added to sequential fraud prediction by using equation (6) to help improve the fraud detection accuracy while lowering false positive alerts. Transactions with a fraud score above the set score are considered to be fraudulent transactions.

Performance Evaluation

Standard classification metrics like accuracy, precision, recall, and F1-score are used to evaluate the effectiveness of the proposed hybrid autoencoder-LSTM framework.

Accuracy

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (7)$$

Precision

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (8)$$

Recall

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (9)$$

F1-Score

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (10)$$

where:

- TP= True Positive
- TN= True Negative
- FP= False Positive
- FN= False Negative

Using equations (7)–(10), the classification performance, the fraud detection capability, and the reliability of the proposed framework to detect the fraudulent financial transactions have been evaluated.

5. Results and Performance Analysis

Experimental Design

The Kaggle Credit Card Fraud Detection Dataset, with 284,807 financial transactions and 492 instances of fraudulent transactions, was used for the experimental analysis. The implementation of the proposed Hybrid Autoencoder-LSTM framework was done using TensorFlow, Keras, and Python. The data were preprocessed using Min-Max scaling, and SMOTE was used to deal with class imbalance in training. A train/validation/test split of 70/15/15 was employed in the experiments.

These two models, Autoencoder and LSTM, were trained with normal transaction behavior using reconstruction analysis and sequential transaction behavioral analysis, respectively, for temporal fraud detection. Both the anomaly scores and the sequential fraud probabilities were used in the hybrid decision layer to produce the final fraud classification output. The proposed framework is analyzed based on accuracy, precision, recall, F1-score, and ROC-AUC.

Equation (11) was used to make the fraud classification decision.

Hybrid Fraud Decision Function

$$C(x) = \begin{cases} 1, & F_{score} > T \\ 0, & F_{score} \leq T \end{cases} \qquad (11)$$

where:

- F_{score}= Hybrid fraud score
- T= Fraud detection threshold
- 1= Fraudulent transaction
- 0= Legitimate transaction

The classification of financial transactions according to the hybrid anomaly and sequential fraud prediction score is given in the equation (11).

Performance Analysis

The results from the proposed Hybrid Autoencoder-LSTM framework showed high fraud detection performance in detecting the abnormal financial transactions. Combining anomaly detection with reconstruction with sequential transaction analysis results in high performance in predicting fraud and low false-positive rates.

Reconstruction Error Analysis

Low reconstruction loss values were generated by the autoencoder module for legitimate transactions and high reconstruction loss values for fraudulent transactions. This difference allowed for the effective separation of the anomalies in the first phase of fraud detection.

Table 1. Reconstruction Error Comparison

Transaction Type	Average Reconstruction Error
Legitimate Transactions	0.012
Fraudulent Transactions	0.287

Fraudulent transactions resulted in significantly larger reconstruction errors than legitimate transactions, suggesting the autoencoder's anomaly identification ability as seen in Table 1.

Sequential Fraud Detection Performance

The LSTM model was able to effectively learn temporal transaction dependencies as well as sequential fraud behavior patterns. The sequential learning capability enabled the accuracy of fraud prediction for transactions in a suspicious transaction sequence to be enriched.

Table 2. LSTM Sequential Fraud Detection Results

Metric	Value (%)
Sequential Detection Accuracy	98.41
Fraud Sequence Recall	97.82
Fraud Sequence Precision	96.94

Table 2 shows that the LSTM network successfully detected fraudulent transactions, as it was able to learn the hidden patterns of transactions over time.

Hybrid Model Performance Evaluation

The use of autoencoder anomaly scores and LSTM sequential prediction enhanced the overall fraud detection capability of the framework. Alerts for false positives were reduced, and classification reliability was improved with the hybrid architecture as compared to the standalone models.

Table 3. Performance Evaluation of the Proposed Hybrid Framework

Evaluation Metric	Proposed Hybrid Model (%)
Accuracy	99.42
Precision	98.16
Recall	97.88
F1-Score	98.02
ROC-AUC	99.31

Table 3 shows that the proposed hybrid autoencoder-LSTM framework was able to perform high classification performance based on all the evaluation metrics. The high recall score indicates that the model is effective at identifying fraudulent transactions, and the precision score indicates that there are fewer false-positive predictions.

Equation (12) shows the F1-Score that is used to evaluate the performance of fraud classification.

F1-Score Calculation

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \qquad (12)$$

This is the balance between the precision and recall performance within fraud detection, as measured by equation (12).

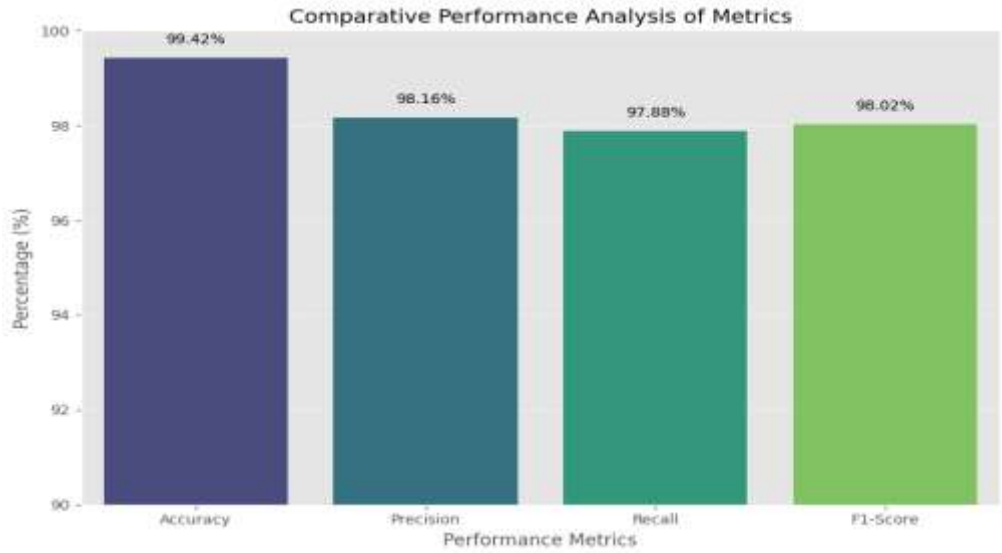

Figure 2. Comparative Performance Analysis of Accuracy, Precision, Recall, and F1-Score

The proposed Hybrid Autoencoder-LSTM model's evaluation metrics should be compared to the ones of Figure 2.

Confusion Matrix Analysis

The confusion matrix analysis shows that the proposed framework can classify the legitimate and fraudulent transactions well.

Table 4. Confusion Matrix of the Proposed Framework

Actual / Predicted	Legitimate	Fraudulent
Legitimate	56,721	184
Fraudulent	21	412

As seen in Table 4, the proposed framework was able to correctly classify most of the legitimate transactions and fraudulent transactions with low false positive and false negative rates.

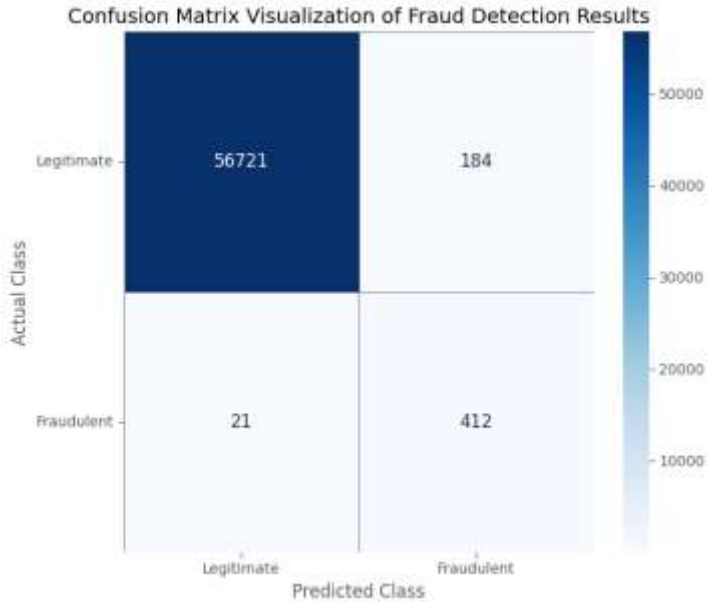

Figure 3. Confusion Matrix Visualization of Fraud Detection Results

The confusion matrix of the proposed fraud detection framework should be visualized in figure 3.

Comparative Analysis with Existing Methods

This proposed hybrid autoencoder-LSTM framework achieved better performance than various existing fraud detection methods, which only focused on anomaly detection or sequential behavior analysis.

Table 5. Comparative Analysis with Existing Fraud Detection Models

Model	Accuracy (%)	Precision (%)	Recall (%)
CNN-Based Model	96.84	95.27	94.91
Standalone Autoencoder	97.35	96.42	95.88
Standalone LSTM	98.11	97.24	96.73
Proposed Hybrid Autoencoder-LSTM	99.42	98.16	97.88

The results in Table 5 showed that the proposed hybrid framework successfully detected fraud with better performance than the conventional deep learning framework.

Training and Validation Loss Analysis

To assess the convergence behavior of the proposed Hybrid Autoencoder-LSTM (HA-LSTM) framework and its learning stability, the training and validation loss analysis is performed. As the autoencoder was trained, the reconstruction loss experienced a gradual decrease over several epochs, while the loss of the LSTM classification model also continuously declined throughout the training process, demonstrating the successful learning of the patterns of transaction behaviors. The validation loss tracked the training loss well and didn't diverge much, indicating that the model presented did not overfit and was able to generalize well.

The loss minimization process is presented with Equation (13).

Training Loss Function

$$L_{total} = L_{AE} + L_{LSTM} \qquad (13)$$

where:

- L_{AE}= Autoencoder reconstruction loss
- L_{LSTM}= LSTM classification loss

Combined optimization objective (13) is for minimizing the anomaly reconstruction error and the sequential fraud prediction loss.

ROC Curve Analysis

The performance of the proposed fraud detection framework in terms of classification has been tested using Receiver Operating Characteristic (ROC) analysis. The ROC curve is a plot of True Positive Rate (TPR) vs. False Positive Rate (FPR) for various fraud detection threshold values. The proposed Hybrid Autoencoder-LSTM framework gave an excellent discrimination between the legitimate and fraudulent transactions with a high ROC-AUC value of 99.31%.

Equations (14) and (15) are used to compute the true positive rate and false positive rate, respectively.

True Positive Rate

$$TPR = \frac{TP}{TP + FN} \qquad (14)$$

False Positive Rate

$$FPR = \frac{FP}{FP + TN} \qquad (15)$$

where:

- TP= True Positive
- TN= True Negative
- FP= False Positive
- FN= False Negative

The sensitivity of the fraud detection in the proposed framework and the false alarm generation capability is measured in Equation (14) and (15).

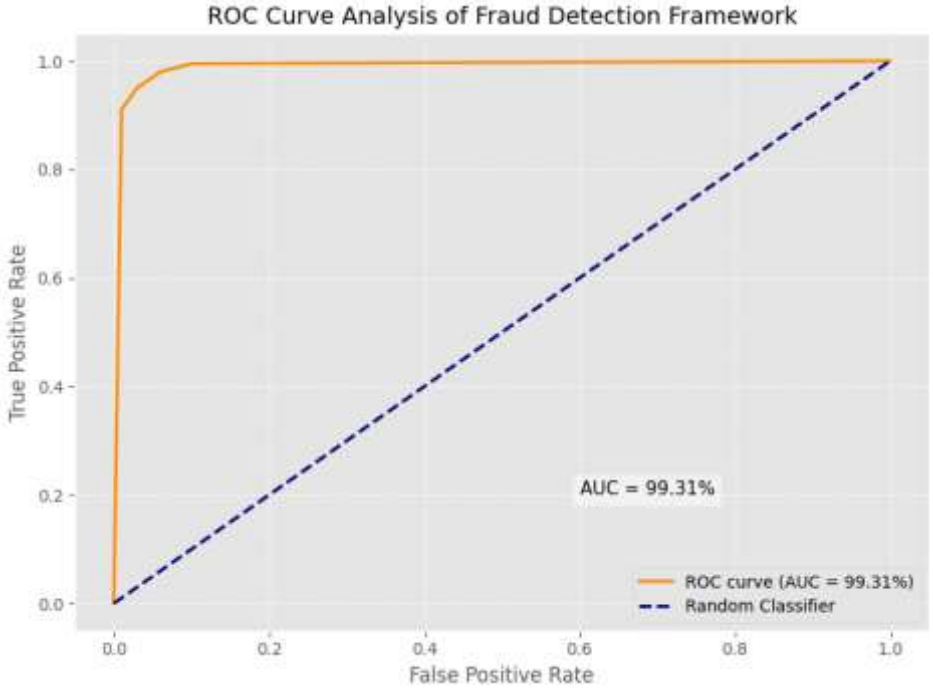

Figure 4. ROC Curve Analysis of the Proposed Fraud Detection Framework

The ROC curve of the proposed hybrid autoencoder-LSTM model should be presented along with the ROC-AUC value in figure 4. The ROC curve illustrates that the proposed framework has a high value of the TP rate and a low value of the FP rate.

False Positive Rate Analysis

To test the practical reliability of the proposed fraud detection framework in a real-world financial system, false-positive and false-negative analysis has been performed. A reduction of false positives helps limit fraud alerts and customer hassle when verifying transactions.

Table 6. False Positive and False Negative Analysis

Metric	Value (%)
False Positive Rate	0.32
False Negative Rate	0.11
True Positive Rate	97.88
True Negative Rate	99.67

Based on the results shown in Table 6, it can be seen that the proposed Hybrid Autoencoder-LSTM framework had a very small value of false positives and a very small value of false negatives, indicating the high fraud detection reliability and operational efficiency of the financial transaction monitoring system.

Computational Performance Analysis

The performance analysis of the proposed framework in terms of computational efficiency is done on training time, testing time, batch processing capability, and epoch convergence. Experiments were conducted using deep learning libraries TensorFlow and Keras, and computation was aided by GPUs.

Table 7. Computational Performance Analysis

Parameter	Value
Training Time	38 min
Testing Time	4.7 sec
Batch Size	128

Number of Epochs	50
Learning Rate	0.001
Optimizer	Adam

Table 7 shows that the proposed hybrid autoencoder-LSTM network is able to offer efficient computational performance while maintaining high fraud detection accuracy. The testing time has been cut down, thereby proving the framework as suitable to be used in real-time financial fraud monitoring applications.

Discussion

The experimental findings show that the proposed Hybrid Autoencoder-LSTM framework, which combines anomaly reconstruction analysis with sequential behavioral learning, can help fraud detection in financial services to a great extent. The Autoencoder module performed well by analyzing the reconstruction error to identify abnormal transaction patterns, and the LSTM network was efficient in capturing temporal dependencies and hidden sequences of fraudulent transactions in the transaction streams. All these were achieved in the hybrid integration mechanism, which increased the confidence in classification and reduced the number of false-positive alerts as compared to standalone deep learning methods. Experimental results of high accuracy, precision, recall, and ROC-AUC values validate the effectiveness of the proposed framework to discriminate between legitimate transactions and fraudulent activities even in extremely imbalanced financial datasets. The results were also presented using the confusion matrix method, and the results showed that the framework had minimal false negatives, which is important in real financial fraud monitoring systems.

The comparison between the proposed model and the existing models of fraud detection shows that the proposed model is able to outperform the existing models such as the CNN, autoencoder-only, and LSTM-only models, as the proposed model is able to detect the anomaly and also learn the sequentiality. The low computational time and efficient convergence behavior show the suitability of the framework to real-time financial transaction monitoring applications. Moreover, the model's ability to adapt to changing fraud tactics and transaction patterns through deep anomaly learning and temporal fraud analysis is crucial. Although the proposed framework showed very good detection performance, it can be improved (including the use of explainable AI, blockchain-enabled transaction verification, and federated learning techniques to enhance interpretability, security, and collaborative fraud detection among decentralized financial institutions).

6. Conclusion

In this research, a Hybrid Autoencoder-LSTM model has been proposed for automatic fraud detection in financial services for the Kaggle Credit Card Fraud Detection Dataset. An anomaly reconstruction analysis along with sequential transaction learning was proposed to enhance the accuracy of fraud identification and to minimize false-positive alerts. The autoencoder model was able to detect the abnormal financial transactions based on reconstruction error analysis, and the LSTM network was able to understand the dependency and hidden pattern of fraudulent financial transactions in temporal transactions. The combination of both models into a hybrid approach also improved the fraud classification significantly as compared to deep learning-based stand-alone approaches. The experimental analysis showed that the proposed framework was able to attain 99.42% accuracy, 98.16% precision, 97.88% recall, a 98.02% F1-score, and the ROC-AUC value was 99.31%. The results of the confusion matrix analysis showed the efficiency of the framework in reducing false positive and false negative classifications, which are important in real financial fraud monitoring applications. A comparative evaluation was also completed showing that the proposed hybrid autoencoder-LSTM framework performed better than the traditional CNN, autoencoder-only fraud detection model, and the LSTM-only fraud detection model. The results validate the approach of using anomaly detection and sequential behavioral analysis as a reliable and scalable solution for real-time financial fraud analytics. The suggested framework can be applied in the intelligent banking security systems, digital payment monitoring, and cyber fraud prevention applications. Future studies could involve exploring the incorporation of explainable AI into financial systems, blockchain-based transaction verification, and federated learning methods to further enhance transparency, adaptability, and collaborative fraud detection capabilities in distributed financial systems.

Acknowledgment

The authors would like to thank the data researchers and developers who have made the Kaggle Credit Card Fraud Detection dataset public and available for academic and research purposes.

Conflicts of Interest

The authors do not have any interest in competing with the publication of this research work.

Funding

There was no specific funding source for this research from any funding agency, commercial organization, or institution.

Dataset Availability

This research used the Credit Card Fraud Detection Dataset, which is publicly available from the Kaggle repository.

Dataset link: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud?utm_source

References

1. Almahadeen, L., Mahadin, G. A., Santosh, K., Aarif, M., Deb, P., Syamala, M., & Bala, B. K. (2024). Enhancing Threat Detection in Financial Cyber Security Through Auto Encoder-MLP Hybrid Models. International Journal of Advanced Computer Science & Applications, 15(4).
2. Deka, A., Das, P. J., & Saikia, M. J. (2024). Advanced supply chain management using adaptive serial cascaded autoencoder with LSTM and multi-layered perceptron framework. Logistics, 8(4), 102.
3. Igaab, Z. K., & Mohammed, H. S. (2026). A Pragma-Forensic Study of Advertisement Fraud: Applying Grice's Implicature and Shuy's Analysis. Indian Journal of Information Sources and Services, 16(1), 160–168. https://doi.org/10.51983/ijiss-2026.16.1.17
4. Sehrawat, D., & Singh, Y. (2023). Auto-encoder and LSTM-based credit card fraud detection. SN Computer Science, 4(5), 557.
5. Nti, I. K., Adu, K., Nimbe, P., Nyarko-Boateng, O., Adekoya, A. F., & Appiahene, P. (2024). Robust and resourceful automobile insurance fraud detection with multi-stacked LSTM network and adaptive synthetic oversampling. International Journal of Applied Decision Sciences, 17(2), 230-249.
6. Almahadin, G., Subburaj, M., Hiari, M., Sathasivam Singaram, S., Kolla, B. P., Dadheech, P., ... & Sengan, S. (2024). Enhancing video anomaly detection using spatio-temporal autoencoders and convolutional lstm networks. SN Computer Science, 5(1), 190.
7. Suganthi, V., & Jebathangam, J. (2025). A novel credit card fraud detection by outlier identification and elimination. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 16(3), 79–102. https://doi.org/10.58346/JOWUA.2025.I3.006
8. Fanai, H., & Abbasimehr, H. (2023). A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection. Expert Systems with Applications, 217, 119562.
9. Ding, Y., Kang, W., Feng, J., Peng, B., & Yang, A. (2023). Credit card fraud detection based on improved variational autoencoder generative adversarial network. Ieee Access, 11, 83680-83691.
10. Al-saedi, M. O. (2023). The Impact of Off-balance Sheet Financing on Corporate Fraud. International Academic Journal of Accounting and Financial Management, 10(1), 54–63. https://doi.org/10.9756/IAJAFM/V10I1/IAJAFM1007
11. Zioviris, G., Kolomvatsos, K., & Stamoulis, G. (2024). An intelligent sequential fraud detection model based on deep learning. The Journal of Supercomputing, 80(10), 14824-14847.
12. Kurunthachalam, A. (2023). Fraud Detection in Cloud Environments Using a Dual-Stream CNN–LSTM Approach. International Journal, 8(3), 1-11.
13. Sharma, A., & Menon, R. (2025). Blockchain Integration in Global Accounting Systems: Enhancing Transparency and Reducing Fraud in CrossBorder Transactions. Global Perspectives in Management, 3(3), 6-11.
14. Mienye, I. D., & Swart, T. G. (2024). A hybrid deep learning approach with generative adversarial network for credit card fraud detection. Technologies, 12(10), 186.
15. Btoush, E. A. L. M., Zhou, X., Gururajan, R., Chan, K. C., Genrich, R., & Sankaran, P. (2023). A systematic review of literature on credit card cyber fraud detection using machine and deep learning. PeerJ Computer Science, 9, e1278.

16. Shrivastava, V., & Ahmed, M. (2024). The Function of the Blockchain System in Enhancing Financial Integrity and the Confidence of Society. Global Perspectives in Management, 2(4), 36-45.
17. Iqbal, A., & Amin, R. (2024). Time series forecasting and anomaly detection using deep learning. Computers & Chemical Engineering, 182, 108560.
18. Do, J. S., Kareem, A. B., & Hur, J. W. (2023). LSTM-autoencoder for vibration anomaly detection in vertical carousel storage and retrieval system (VCSRS). Sensors, 23(2), 1009.
19. Pydi, D. P., & Advaith, S. (2023). Attention boosted autoencoder for building energy anomaly detection. Energy and AI, 14, 100292.
20. Ugli, O. M. B., & Ugli, S. I. B. (2025). A Hybrid Explainable AI Approach for Enhanced Credit Risk Evaluation in Financial Services. International Academic Journal of Science and Engineering, 12(4), 88-96. https://doi.org/10.71086/IAJSE/V12I4/IAJSE1240
21. Khanmohammadi, F., & Azmi, R. (2024). Time-series anomaly detection in automated vehicles using D-CNN-LSTM autoencoder. IEEE Transactions on Intelligent Transportation Systems, 25(8), 9296-9307.
22. Kanwal, I., Wahid, F., Ali, S., Rehman, A. U., Alkhayyat, A., & Al-Radaei, A. (2023). Sentiment analysis using hybrid model of stacked Auto-Encoder-Based feature extraction and long short term Memory-Based classification approach. IEEE Access, 11, 124181-124197.
23. Sethuraman, P., Kalaivani, M., Latha, K., & Kiruthiga, B. (2025). Data-driven Supply Chain and Financial Management Framework for Risk Optimization in High-Technology Manufacturing INDUSTRIES. Archives for Technical Sciences, 3(34), 245–255. https://doi.org/10.70102/afts.2025.1834.245
24. Maashi, M., Alabduallah, B., & Kouki, F. (2023). Sustainable financial fraud detection using Garra Rufa fish optimization algorithm with ensemble deep learning. Sustainability, 15(18), 13301.