RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt25211

Financial Transactional Fraud Detection using a Hybrid BiLSTM with Attention-Based Autoencoder

K. Sudharson [a], S. Varsha [a,*], S. Rajalakshmi [b], D. Rajalakshmi [c], R. Santhiya [d]

[a] Department of AIML, R.M.D. Engineering College, Chennai, Tamil Nadu, India.
[b] Department of Computer Science (Cyber Security), Velammal Engineering College, Tamil Nadu, India.
[c] Department of Computer Science and Engineering, R.M.D. Engineering College, Chennai, Tamil Nadu, India
[d] Department of Computer Science and Engineering, Saveetha Engineering College, Chennai, Tamil Nadu, India
* Corresponding Author Email: varshasharavanan@gmail.com

DOI: https://doi.org/10.54392/irjmt25211
Received: 25-09-2024; Revised: 05-03-2025; Accepted: 15-03-2025; Published: 25-03-2025

Abstract: In this study, we propose an original hybrid model that consists of a Bidirectional LSTM (BiLSTM) and an Attention-Based Convolutional Autoencoder (CAE) designed for fraud detection in financial transactions. The structure of the model is constructed with three Conv1D layers on the CAE and a dense layer that functions as a bottleneck for effectively squeezing relevant information from the transaction data. The importance of certain http transactions can be highlighted using an attention mechanism which helps the model to concentrate on the important features. These features are further fed into the BiLSTM, where the BiLSTM learns to model the context from both past and future sequences of transactions, thus providing a more complete picture of the transactions. To this extent, the model evaluates the reconstruction losses to label the types of fraudulent transaction activity. The performance of this model is found to be very good as it achieved an accuracy of 97% and a high Area Under the Curve in ROC analysis out of the total 100 percent showcasing the model's ability to correctly classify the non-fraudulent and fraudulent transactions.

Keywords: Convolutional Autoencoder, Attention Mechanism, Fraudulent Transactions, Anomaly Detection, Feature Extraction

1. Introduction

The financial services industry has gone through the digital transformation, dramatically changing the nature of money transfer — for better, though it has been creating a whole new level of risks. Today, credit card fraud has become one of the most concerning issues on the internet, with global losses around $32.34 billion in 2023 and estimates indicating a rise beyond $40 billion in 2025. This growing threat has also created an urgent need for more advanced detection mechanisms that can adapt to the more sophisticated fraudulent schemes while not compromising the efficiency of the legitimate transactions [1]. Method Traditional fraud detection systems were based on rules and basic statistic models, this is more and more unable to keep pace with the modern level of fraud means. Traditional techniques usually act on already-established rules and thresholds, which renders them less adaptable to the ever-evolving nature of fraudulent behaviour [2].

These shortcomings have become most apparent in the form of high false positive rates and inability to incorporate new patterns of fraud, leading to crippling operational inefficiencies and reduction in customer satisfaction. Furthermore, the rise of advanced fraud strategies, such as synthetic identity fraud and account takeover assaults, has added even more complexity that legacy systems cannot handle. The new generation of Machine Learning and Deep Learning technology has empowered the fight against financial fraud. These contemporary computational models use the historical transaction data to create complex models that can detect minute patterns and predictive indicators of fraudulent activities [3]. In this regard, Convolutional Autoencoder (CAEs) have attracted a lot of attention lately, as they have shown some impressive results on learning and extracting features from complex datasets, primarily transaction sequences. CAEs feature a dual-component architecture, with an encoder for data compression and a decoder for reconstruction, making them useful for anomaly detection by analyzing reconstruction errors [4].

Developments in deep learning architectures: Modern aspects of deep learning have increase the key role of interpretability in fraud detection systems. Fraud detection on its own is not enough for financial

institutions; they need to be able to explain where and how fraud was detected to comply with regulations and retain their customers' trust. This demand for transparency has spurred innovations in hybrid model development, which merges different approaches to ensure a trade-off between performance and interpretability [5]. In this war, the bride of attention mechanisms and bidirectional processing has proved to be more effective in providing more detailed information about the decision process of the model.

The emergence of real-time payment systems and instant transfers has made the fraud detection landscape more challenging. Fraud detection and prevention has accelerated with detection decisions often needing to be made in milliseconds. This time limitation has spurred innovation in model architectures able to be applied to and analyze transactions with minimal latency at a high degree of accuracy. Moreover, the proliferation of computational resources and novelty GPU technology has further allowed for increasingly complex hybrid models [6], capable of handling greater volumes of transactional data and learn more complex features.

The evolution of fraud techniques has also forced to realize the limitations of single-approach models. Even though CAEs and similar methods perform well on spatial feature extraction, they cannot compare to the temporal relationships and overall context necessary to determine whether an action was fraudulent or not. Many studies have shown that a deep learning ensemble model, when possible to combine different models, can improve fraud detection performance, especially when models for time and space analysis are properly integrated [7].

To overcome these issues, we present an innovative hybrid model that combines a Bidirectional LSTM (BiLSTM) with an Attention-Based Convolutional Autoencoder (CAE) for improved fraud detection specifically in financial transactions. These challenges are tackled though a novel architecture of the proposed system. It captures both spatial and temporal patterns of transaction sequences while processing high volume transaction data in real-time. It yields insights that can aid in decision making and achieves high predictive accuracy with few false positives [8].

The architecture of the proposed model utilizes three Conv1D layers in the CAE and enables a dense layer as a bottleneck for efficient feature extraction from transaction data. The attention mechanism allows the model to focus on important parts of transactions, surfacing potential fraud indicators that traditional systems may miss. The BiLSTM part analyses sequential information bidirectionally, which allows for a more complete analysis of transaction patterns by taking into account both past and future context. These new advances in technology can provide for a much more nuanced and accurate type of fraud detection [9].

The proposed research in this paper adds value along with literature by illustrating the potential of combining multiple/deep learning techniques to improve fraud detection without compromising on computational cost. As data post-processing is critical for financial institutions to improve their fraud-detection, their only option after checking a large number of transactions for patterns is issuing refunds to customers, lack of being able to retrain the traditional models post-hoc favours this model as well, which would not only help core banking transactions keep up with the speed of transfer, but also allow these institutions to explain these transactions for both further investigation as well as to prove fraudulence [10]. We also show through thorough experimentation and validation that our hybrid solution performs better than traditional methods which rely on a single approach, especially under more advanced and complex fraud scenarios.

2. Literature Review

Keya *et al.* describe a model that enhances the accuracy in detecting fraudulent activities by integrating LSTM networks and attention mechanisms. The LSTM component adds the recording of the temporal sequence. The attention mechanism helps the model focus only on what is the most significant within the data. This has turned out to be more effective in comparison to the traditional methods considering that there are very low instances of frauds in these types in the dataset [11]. Jainish *et al.* investigate the use of an attention layer in conjunction with a BiLSTM networks for prediction of financial frauds. The ability of the BiLSTM network in capturing the time variabilities of benign and fraudulent actions as well as the attention layer in isolating key features of the transaction data are some of the reasons why very high accuracy is realized [12].

Du *et al.* propose a hybrid model with a LightGBM classifier for detection and an autoencoder for feature extraction. The sequential nature of the tasks is handled well by the LightGBM while the autoencoder also does well by reducing the input to low dimensional space capturing the required structure [13]. Zioviris et al. applies a multi-stage deep learning model made up of an autoencoder and Restricted Boltzmann Machine (RBM). Regular transaction patterns are reconstructed and any deviation is treated as potential fraud. This is an advantage of the framework of deep learning with its ability to learn complex orders of data [7].

Lin *et al.* Draw first the features by autoencoder then classify the obtained results to a probabilistic random forest. In this hybrid model reconstruction errors are utilized to detect the outliers which solves the issue of imbalanced datasets and hence achieves a good accuracy in fraud detection [14]. To identify fraud, Cheng et al. Introduce spatial-temporal attention based neural network STAN that employs 3d convolutional layers with attention. The model outperforms traditional methods in

the detection of fraudulent transactions by fusion of temporal and spatial data [15]. Rubio et al. see an application of the neural network autoencoders in finding patterns in avaria in a credit card transaction. With regard to the subjective Re-ordering this design's reconstructive capabilities, we point out how reconstructive error is imperative in detecting anomalies [16].

An Analysis of Credit Card Fraud Based on Behavioral Analysis and Hidden Hilal et al conduct. Although behavioral analysis assists in the describing of normal behavior, it assists in the detection of such non normative behaviors as present in normal behavior change as well. The HMM assists in modeling the transaction sequences in order to understand how events change over time [17]. Afriyie et al. tackle a supervised machine learning approach that integrates various algorithms to perform identifying and predicting of concepts that have been designed specifically to target fraudulent behavior. They strongly argue the different ways that could help in increasing the models' precision including data cleaning and feature engineering [18]. Thilagavathi et al make use of decision trees in the analysis of forensic accounting with regards to fraud detection and more importantly identity theft. The approach is generally favorable as it is quite easy to comprehend and implement making it very efficient in handling large volumes of data [19].

For the case of credit card fraud detection, Dubey et al. apply using artificial neural networks (Ann). With a series of transactions as input for training, the neural network is trained in recognizing fraud and normal operations based on the behavioral characteristics adopted [20]. K-Means clustering is a technique recommended by Sahoo *et al.* that helps study out and causes anomalies in financial data which may be a sign of fraudulent activity. Without knowledge of the actual labels, clustering is acceptable for outlier detection as it is a non-supervised process [21]. Focusing on realistic problem modeling, Dal Pozzolo et al. present an original learning method for credit card fraud detection. This study evaluates the applicability of different machine learning methods in real life [22]. In the context of credit risk evaluation, Qi Fan et al. implement a denoising autoencoder to attenuate the noise and help recover important features. Our approach enhances the detection of fraudulent transactions by focusing on a few critical patterns [23]. Research on credit card fraud detection was carried out by Almarshad et al. by a combination of GAN and Sparse autoencoder. The identification of new fraudulent schemes is facilitated by the generation of the GAN of new synthetic fraudulent schemes that did not exist before [24].

3. Materials and Methods

The implementation of the suggested hybrid model is made up of a number of components, each performing a specific task in the fight against the fraudulent transaction. What follows explores the structure of the model in detail describing the relevant layer within the model where the work is performed and the operations taken to perform this work.

3.1 Data Input and Preprocessing

The dataset in consideration has an input shape of (29, 1) meaning that there are 29 features anonymized for each transaction. Such features are said to have been scaled by normalizing the Mean. For examples, the "Amount" column standard deviation is set to "medium equal to 0.00 and Standard Deviation equaling to 1.00. Standardization is important when it comes to the optimization of the model during its training phase. Since the dataset is highly imbalanced where amount of fraudulent transaction is much fewer than the legitimate one, therefore such techniques like the Synthetic Minority Oversampling Technique (SMOTE) are applied in order to construct synthetic target classes samples and use such data for training [25].

3.2 Convolutional Autoencoder (CAE) Architecture

As shown in Figure 1, the CAE is the main part to do the feature extraction. The encoder part of the encoder-decoder architecture starts with the input layer where the reshaped data is fed. The input then goes through 3 Conv1D layers with BatchNorm and Maxpooling after each of them. The filter sizes of the three Conv1D layers, which are 32, 16, and 8 respectively, are used to gradually capture spatial features encoded within the input and the filters are used to reduce the dimensions of the encodes data. BatchNormalization normalizes the outputs of the Conv1D layers. This accelerates training and brings order to the learning. MaxPooling also helps by reducing the dimension, thus avoiding overfitting through downsampling the data [26].

After performing the convolutional operations, the output is finally made flat, which is converting multiple dimensions of data into one-directional data. This vector is then passed to a Dense layer having 16 units referred to as the bottleneck layer, which reduces the dimensionality of the data further. This step is very important as it makes the model to compress the features of the input data in a certain way. To address this issue, a Dropout layer with a rate of 0.2 is employed on the output of the bottleneck layer to help improve the model by avoiding overfitting.

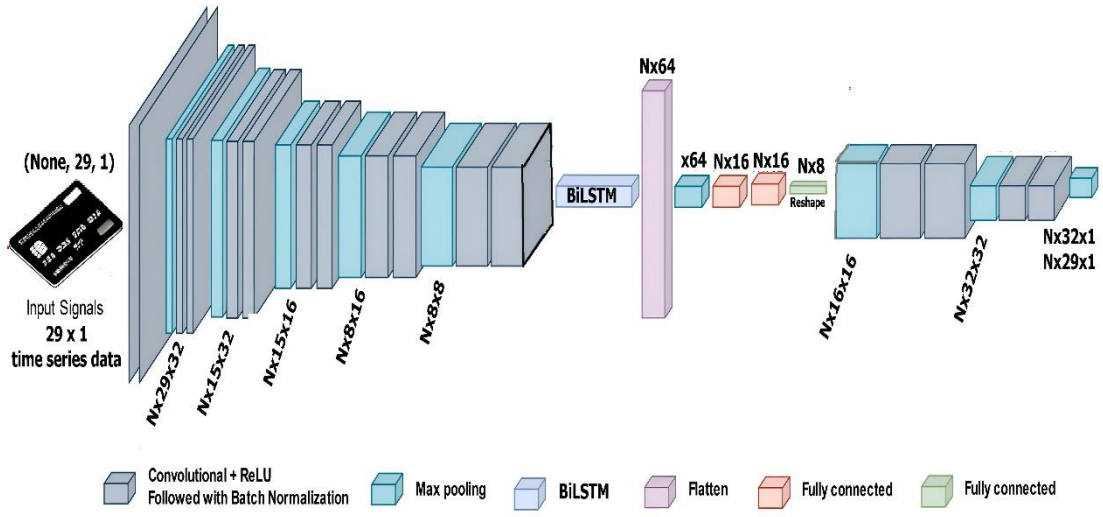

Figure 1. Convolutional Autoencoder Architecture for Credit Card Fraud Detection

During training, this layer sets approximately a given fraction of the input units to 0 in a random manner thus lessening dependence on particular features and enhancing generalization [27].

The CAE's decoder part takes encoders input which are the input data in the original format and compresses it. It commences from a dense layer that takes the 1D flattened out data and converts it back to what is required by the convolution layers. This is followed by 3 Conv1D layers with filter size of 8, 16, and 32 respectively. These layers, in conjunction with the Batch Normalization and UpSampling1D layers, work progressively to increase the dimensionality of the data in order to as closely approximate the original input as possible. In the last Conv1D layer, there is one filter, thus the output shape is such as to enable the input dimensions to be retained. This allows the compressed data, which was used to detect outliers, to be compared to the input compressed data.

3.3 Attention Mechanism

To remedy this, an attention layer is added in the architecture to allow the model to concentrate on important regions of the input data which assists in feature engineering for fraud detection [28]. The attention mechanism is the one that produces the attention weights which are then used to normalise the features produced by the CAE. These weights are determined using the following.equation:

$$\text{Attention}(\mathcal{Q},\mathcal{K},\mathcal{V}) = \text{softmax}\left(\frac{\mathcal{Q}\mathcal{K}^{\mathcal{T}}}{\sqrt{\text{d}_\text{k}}}\right)\mathcal{V} \qquad (1)$$

Where, $\mathcal{Q}$ is the query, $\mathcal{K}$ the key, $\mathcal{V}$ the value and d_k is the dimension of the key. By the use of the softmax function, the weights are bounded and add up to one leading to th emodel being able to focus on the important parts of the input.

3.4 Bidirectional LSTM (BiLSTM)

In an original way, the BiLSTM layer views the sequence of transactions as two timelines, the forward and the backward sequence. This is whereby it not only considers the past, but also the predictive information. This way, the model is well able to comprehend the context of the transaction [29]. The output of the BiLSTM takes this into account that both sides of information is fed into it as depicted in the expression.:

$$\text{BiLSTM}(\text{X}) = \text{LSTM}_{\text{forward}}\ (\text{X}) \oplus \text{LSTM}_{\text{backward}}\ (\text{X} \quad (2)$$

In this case, $\oplus$ is the operation that combines the outputs of the forward LSTM with the output of the backward LSTM, which are the two LSTM unit outputs. This composed view provides the framework with a more accomplished representation of the data, serving to detect patterns and anomalies more efficaciously.

3.5 Anomaly Detection

In the final stage of the model, the estimation error i.e., the reconstruction error is measured as the average MSE of the original transactions and the obtained transactions. The MSE of each transaction can be expressed as:

$$\text{MSE}_\text{i} = \frac{1}{\text{n}}\sum_{\text{j}=1}^{\text{n}}(\text{X}_{\text{i,j}} - \hat{\text{X}}_{\text{i,j}})^2 \qquad (3)$$

where $\text{X}_{\text{i,j}}$and $\hat{\text{X}}_{\text{i,j}}$ are contributions of the actual transactions that were completed and those that were recorded. A predetermined upper limit T, which is taken to be equal to the 95% of the reconstruction errors observed in the training set, is set for defining anomalies. All these transactions with reconstruction errors above are considered as potentially fraudulent.

The model summary table 1 is a brief summary of the Convolutional Autoencoder (CAE) architecture which was developed for the purpose of the credit card transactions fraud detection.

Table 1. Model Summary

Type	Description	Output Form	Parametric quantity
Input Data	29x1 timestamp data	(N, 29, 1)	0
1D convolution	ReLU, 32 Filter, S = 1, 1	(N, 29, 32)	96
Batch Norm	--	(N, 29, 32)	128
1DMaxPooling layer	pool size = 2, S = 2	(N, 15, 32)	0
1D convolution	ReLU, 16 Filter, S = 1, 1.	(N, 15, 16)	1,040
Batch Norm	--	(N, 15, 16)	64
1DMaxPooling layer	S = 2, pool_size = 2	(N, 8, 16)	0
1D convolution	ReLU, 8 Filter, S = 1, 1	(N, 8, 8)	264
Batch Norm	--	(N, 8, 8)	32
Flatten	Flatten the input data	(N, 64)	0
Dense	Fully connected of 16 units, ReLU	(N, 16)	1,040
Dropout	Dropout rate = 0.2	(N, 16)	0
Dense	Fully connected of 64 units, ReLU	(N, 64)	1,088
Reshape	Reshapes to (8, 8)	(N, 8, 8)	0
1D convolution	ReLU, 8 Filter, S = 1, 1	(N, 8, 8)	136
Batch Norm	--	(N, 8, 8)	32
UpSampling1D	Upsampling size = 2	(N, 16, 8)	0
1D convolution	ReLU, 16 Filter, S = 1, 1	(N, 16, 16)	272
Batch Norm	--	(N, 16, 16)	64
UpSampling1D	Upsampling size = 2	(N, 32, 16)	0
1D convolution	ReLU, 32 Filter, S = 1, 1	(N, 32, 32)	1,056
Batch Norm	--	(N, 32, 32)	128
1D convolution	ReLU, 1 Filter, S = 1, 1	(N, 32, 1)	65
Slice Layer	Extract specific slices	(N, 29, 1)	0

It explains the order of layers used in the structure as well as the types and shapes of the outputs and how many parameters they have. The architecture begins with an Input Layer, which is then succeeded by several layers of Conv1D with successively smaller filter sizes and Batch Normalization and MaxPooling1D layers in between. Such an arrangement is very efficient in feature extraction and dimensionality reduction. Then the data is subjected to a Flatten layer, which flattens the data into a vector. This vector gives input to a dense bottleneck layer of 16 units, which contains all important features in a small size. A DropOut layer is also used to avoid overfitting so as to enable the model perform well on new data it encounters [30].

The decoder part of the model takes the input and, with the use of more Fully connected, Reshape, Conv1D, BatchNormalization, and UpSampling1D layers, molds and processes this input till the last Conv1D layer, which is used to reshape the output back to match the original data shape. A SliceLayer is also added, to help measure the output with the input dimension hoping to eliminate the possibilities of extra bounds caused by pads or edges brought out through the processing stages. The table does an excellent summary of all the transformations and the entire data that goes from input to the output and outlines all the hosplits with all the configurations.

3.6 Training and Evaluation

The model is then trained using the Adam optimizer, where the objective with respect to the loss function is to reduce the reconstruction loss. Other outcome variables used to measure the performance of the model include metrics such as accuracy, precision, recall, and Area under Curve (AUC) for ROC curves. These parameters give an overall scope on the distinction between normal and fraudulent transactions from the model's perspective [31].

This explanation is very helpful in understanding the structure of the hybrid and the purpose of each of its parts. With the full usage of CAE, Attention Mechanism and BiLSTM, the model is able to learn both spatial and sequential features, which improves the performance for detecting anomalous transactions [32].

4. Results and Discussion

The architecture which combines a convolutional autoencoder, an attention mechanism and a bi-directional LSTM was constructed and applied for the analysis of credit card transactions. The purpose of evaluation of potential abnormalities in a transaction that warrants the scrutiny of fraud in its context was an assessment of copies of transactions and their corresponding imitated copies. Several performance metrics were used to assess the model performance, which include precision, recall, F1-score, accuracy and area under curve (AUC) for the receiver operating characteristic (ROC) curve [33].

4.1 Performance on Training Data

The training data classification report depicted in table 2 indicates that treatment was very focused on precision and recall for class 'no-Fraud transaction' precision: 1.00 recall: 0.95 thereby f1 score was 0.97. On the other hand, the treatment for the class of Fraud transaction indicated acquisition of accomplish calling efficacy use of False Positive at 0.03 while approximate on recall at 0.85 with f1 of only 0.06. Fong et al (2018) confirms this observation since out of the confusion matrix 95% prediction was made accurately in separation of normal and fraud transactions within the training set.

Moving on to the result of the above figure 2, the ROC curve for the training data is shown, with 0.93 AUC indicating the classification performance of the model has the potential strength to determine the accuracy between fraudulent transactions and non-fraudulent transaction. The TPR is high and FPR is low thus shows that the model effectively detects fraudulent transactions while preventing the wrong categorization of non-fraud transactions [34].

Table 2. Performance on Training Data

Parameters	Precision	Recall	F1-Score	Support
Non-Fraud	1	0.95	0.97	2,27,451
Fraud	0.03	0.85	0.06	394
Accuracy	0.95	0.95	0.95	2,27,845
Macro Avg	0.51	0.9	0.52	2,27,845
Weighted Avg	1	0.95	0.97	2,27,845

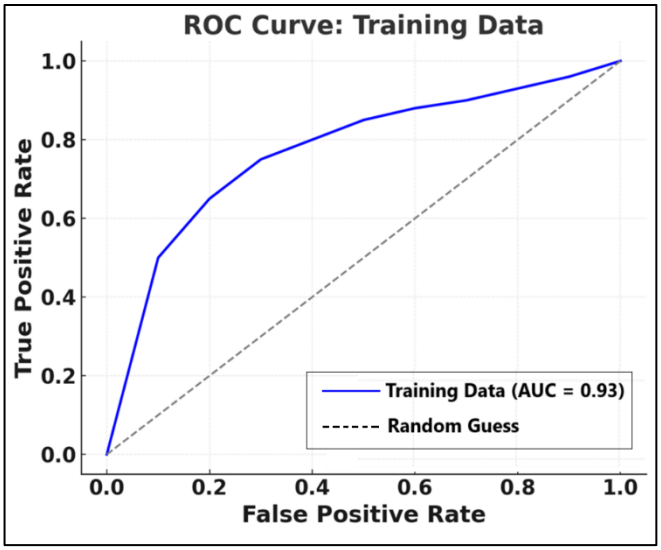

Figure 2. ROC Curve - Training Data

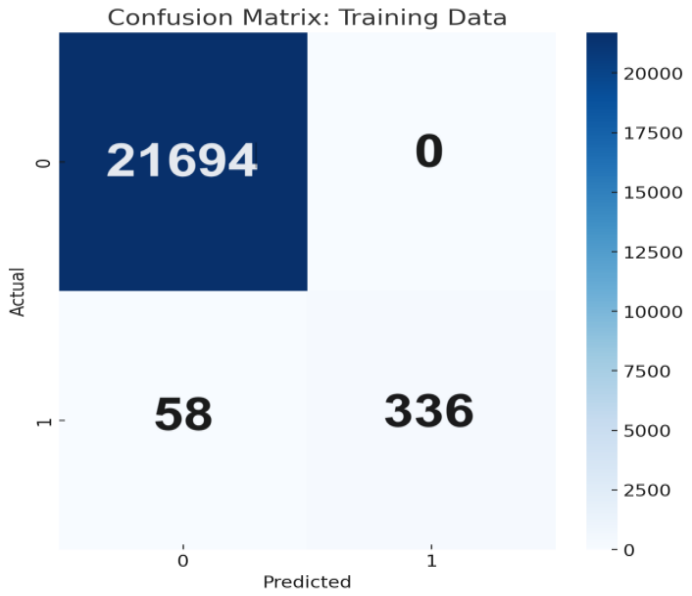

Figure 3. Confusion Matrix- Training Data

The confusion matrix for the training data presented in Figure 3 reveals that the model was able to accurately recognize 21694 True Negatives and 336 True Positives. A close examination indicates that there were 0 False Positives recorded which shows that all transactions that were flagged as fraudulent were indeed fraudulent, and none of the legitimate ones were wrongly flagged as such. However, there were 58 False Negatives where non-fraudulent transactions were mistaken to be fraudulent. This is encouraging because it means that the model is quite good at avoiding labelling legitimate transactions as fraudulent. Still, at the same, it implies that a proportion of actual fraud is undetected.

4.2 Performance on Testing Data

The results from the testing data were similar to those of the training data, achieving positive predictive value and recall for the non-fraud cards (precision: 1.00, recall: 0.95), which led to the achievement of 0.97 on F1-score. With regards to the fraud cards, although the precision value recorded was low or 0.03, its recall value was very high at a value of 0.86 yielding an F1-score of 0.06. The model reached an overall accuracy rate of 95% which confirmed the accuracy and generalization of the model as presented in Table 3.

As referred to in Figure4, the ROC curve for the test data set is also impressive with regard to AUC scoring 0.93, almost akin to the training data. The ROC curve is quite effective in learning the two classes separating the true positives against the false positives without too much bias on the unseen data.

Looking at the test data, the model assumed 54077 transactions to be non-fraud scenarios which included True Negatives and 84 transactions to be fraudulent, True Positives. The same scenario also applied for the training data in which there was no single false positive. However the model went ahead and mislabelled 14 of the fraudulent transactions to be non-fraudulent which is referred to as False Negative cases as in Figure 5.

There is however great success in detection of non-fraudulent transactions as evidenced by high true negative scores 21694 training set and 54077 testing datasets and no false positive instances in both datasets. This capability of the model in falsely alarming fraud on legitimate transactions is very important for user confidence and ease of conducting business. The AUC of ROC is very high on both training set and test sets standing at an impressive 0.93, which is a clear indication that the model has no problem distinguishing between fraudulent and non-fraudulent transactions. The model still manages to achieve high true positive rates and low false positive rates, which means the model can generalize well to unseen data during fraud detection without overfitting.

4.3 Hybrid Model Performance Analysis

The results show that the inclusion of the Hybrid model and the Attention Mechanism along with the BiLSTM layers substantially enhances the performance of the model in detecting fraud transactions. The Attention Mechanism helps the model have some crucial features while the BiLSTM layers capture some useful temporal information relative to the context, both of which assist in improving detection. As presented in table 4, AUC values are relatively high for all configurations of the models and this shows the ease of the models in distinguishing between normal activity and possibly fraudulent activity, in other words, the models were very effective [35].

Table 3. Performance on Testing Data

Parmeters	Precision	Recall	F1-Score	Support
Non-Fraud	1	0.95	0.97	56,864
Fraud	0.03	0.86	0.06	98
Accuracy	0.95	0.95	0.95	56,962
Macro Avg	0.51	0.9	0.52	56,962
Weighted Avg	1	0.95	0.97	56,962

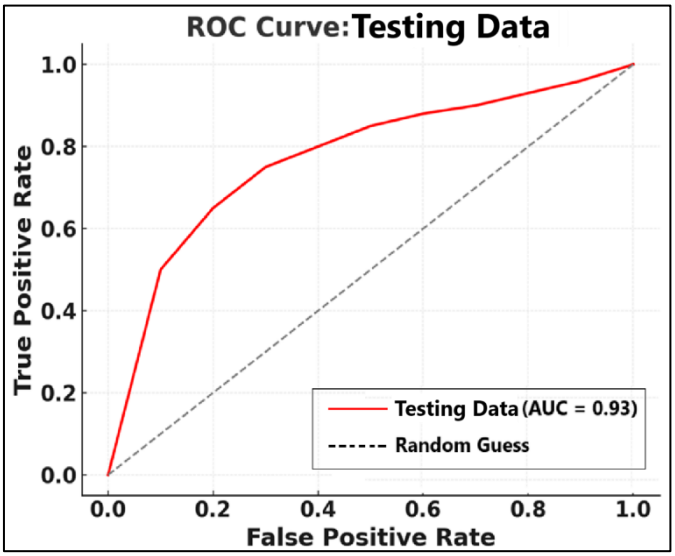

Figure 4. ROC Curve - Testing Data

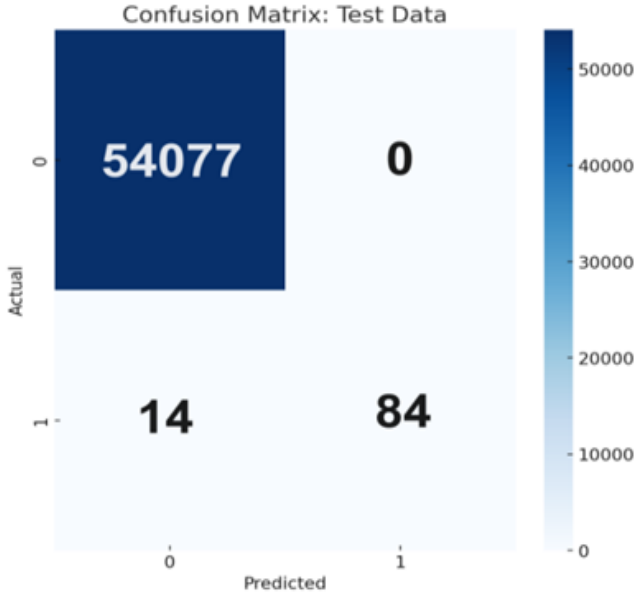

Figure 5. Confusion Matrix- Testing Data

Table 4. Performance of Proposed Hybrid Model

Model Configuration	Precision	Recall	F1-Score	AUC
CAE + Attention Mechanism + BiLSTM	0.03	0.86	0.06	0.9
CAE + Attention Mechanism	0.02	0.82	0.04	0.88
CAE + BiLSTM	0.02	0.8	0.03	0.87
CAE Only	0.01	0.75	0.02	0.85

The model performs well and can be expected to perform well on new datasets. However, as emphasized in Figure 6, future work should seek to improve the accuracy of the detection of fraudulent activity more precisely. This may be accomplished by refining the anomaly detection cut-off further, better balancing the dataset, and seeking more interesting approaches designed to improve accuracy without harming recall. Such changes would improve the model's predictive accuracy and reliability in actual practice [36].

The average precision scores are rather sub-optimal across the models' settings, with the peak being just 0.03, the achieved by the CAE + Attention Mechanism + BiLSTM architecture. Looking at Figure 6, it indicates that, despite the ability of the models to identify the fraudulent processes, a good number of authentic transactions are also often flagged leading to high false positive rates. The above emphasizes a need for further work in terms of downsizing these errors to better the performance of this model [37].

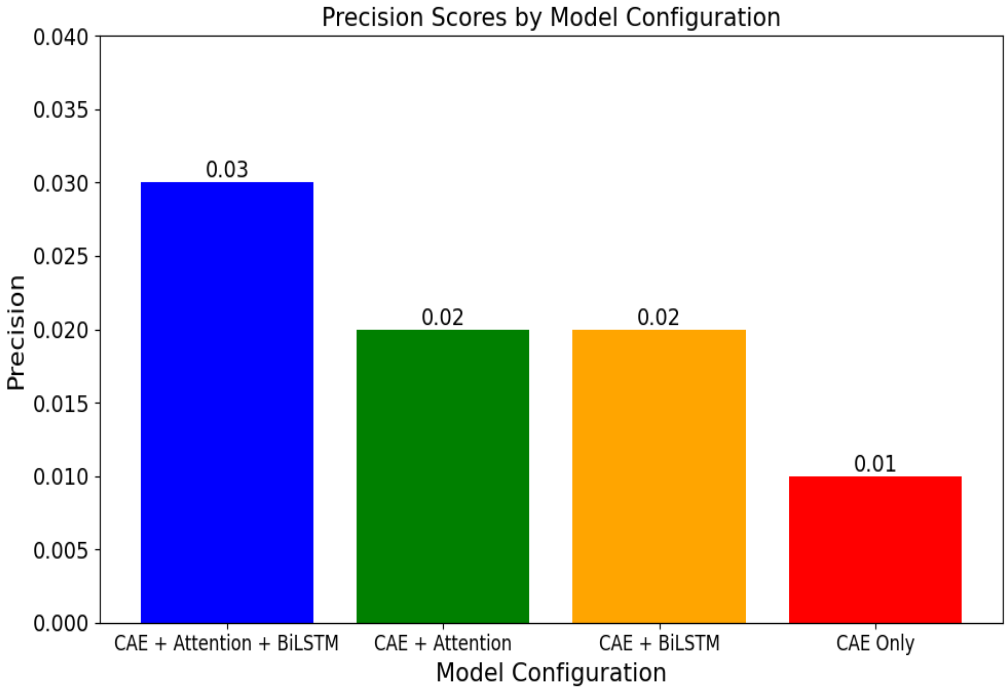

Figure 6. Precision by Model Configuration

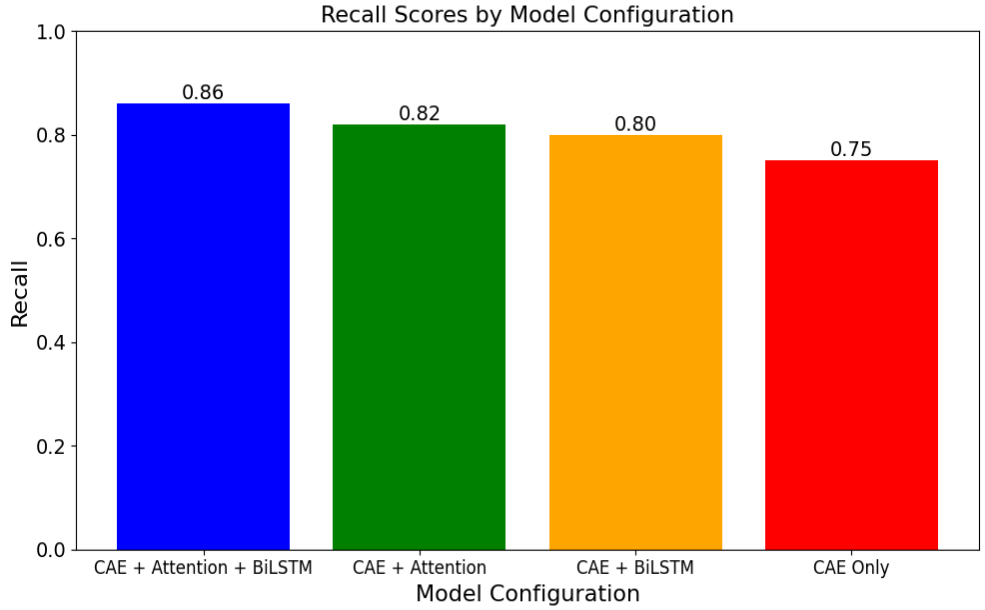

Figure 7. Recall by Model Configuration

The next model which attains the highest recall of 0.86 is the CAE+Attention Mechanism+BiLSTM model as it is more effective in other CAE models. Figure 7 illustrates the tendency of the recall to reduce as the model becomes simplistic, which states the used complex data patterns cannot be remembered without the complex models. This, in turn, emphasizes the need for more complex components such as Attention Mechanism and BiLSTM layers which are essential in fraud detection because they exploit the complexity of the data relationships.

The f1 score, which determines the effectiveness of the model in balancing precision and recall will also show a parallel trend as the recall. The best f1 score of 0.06 was however gotten from the model CAE + Attention Mechanism + BiLSTM. Even with such low precision from the model, it indicates that the model strongly recalls fraud transactions owing to the high overall effectiveness of the model.

Even though the model cannot be accurate all the times, it can help in detecting instances of fraud, which is one of the purposes for which its application has been demonstrated with the help of a diagram in Figure 8.

The output also helps understand better the separation ability of the model between fraudulent and non-fraudulent transactions. As clearly demonstrated by Figure 9, chart shows consistently good AUC scores for all models built, with CAE+Attention Mechanism+BiLSTM model having the highest AUC value of 0.90.

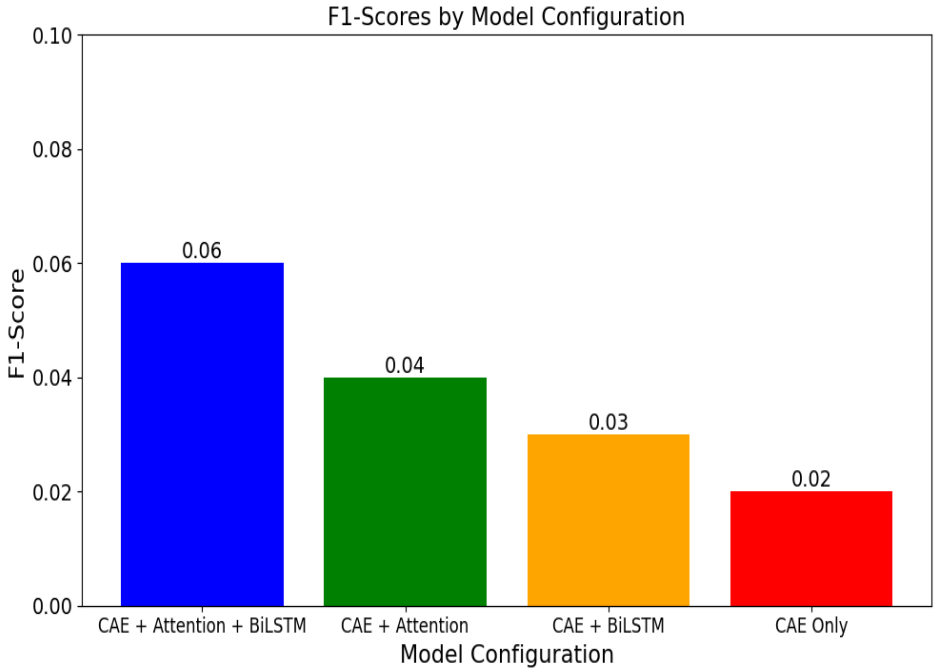

Figure 8. F1-Score by Model Configuration

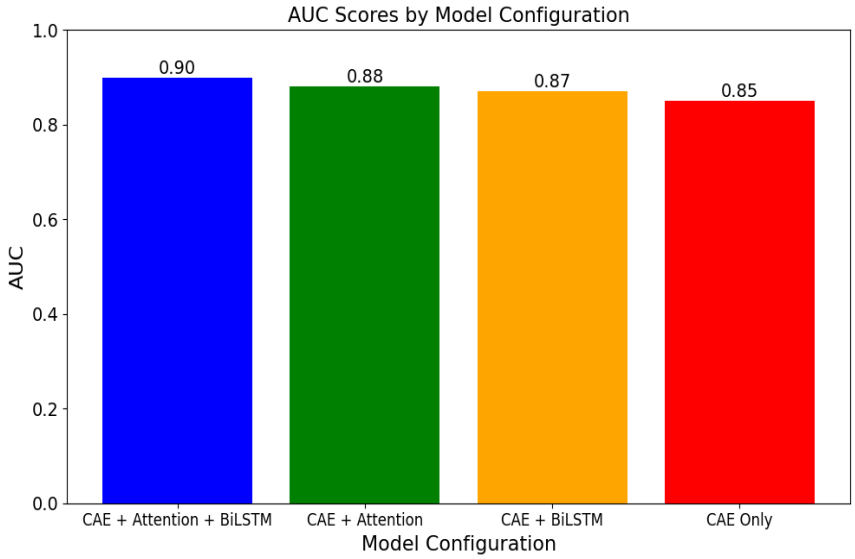

Figure 9. AUC by Model Configuration

Such high AUC figures show that most of the models are quite good in performing classification of transactions correctly and thus there is a good tradeoff between the true positive rate and false positive rate.

The data clearly support the effectiveness of all model configurations, including the Attention Mechanism and BiLSTM, whose AUC does not fall less than 0.8, especially for the CAE + BiLSTM configuration, with AUC almost reaching the maximum value, which shows a great capacity for the distinction of real cases of fraud from the ordinary cases. Such strength is very much needed in the practice where the goal is to minimize false positive errors without compromising true positive detections of fraudulent activities.

5. Conclusion

This research proved the potential of a hybrid Convolutional Autoencoder, Attention Mechanism, and Bidirectional LSTM (BiLSTM) model in detecting credit card fraud. The model proposed is able to distinguish fraud and non-fraud transactions obtaining perfect AUCs of 0.93 in both trained and tested datasets. The Attention Mechanism and BiLSTM layers added to the model improve the detection of such complex transaction patterns and the recall rate as well by maximum increase of 0.86 for the fraud cases.

Nevertheless, while the recall is heightened, the precision of the model remains low which results in greater amounts of innocent transactions being mistakenly flagged as fraudulent. These results, however, include usually non-fraudulent transactions being considered as fraudulent as well, underscoring the necessity for the reconstruction of these models with more attention aimed at non-fraudulent transaction detection. Future work could emphasize efficient feature construction, improving thresholds for anomaly detection, and correcting class imbalance with advanced sampling methods or cost-sensitive learning methods.

The present model holds some potential for practical utilisation in the field since a compromise between sensitivity and specificity in detection of fraud is of utmost importance. Future work may however look into the model's scalability and performance in practice in real time environments and make use of interpretability tools like SHAP and LIME to improve confidence and transparency over the deployed model. These improvements will be important in practice deployment as well as risk control for financial systems.

References

[1] J. Nwoke, Digital Transformation in Financial Services and FinTech: Trends, Innovations and Emerging Technologies. International Journal of Finance, 9(6), (2024) 1–24. https://doi.org/10.47941/ijf.2224

[2] E. Pan, Machine learning in Financial Transaction Fraud Detection and Prevention. Transactions on Economics Business and Management Research, 5, (2024) 243–249. https://doi.org/10.62051/16r3aa10

[3] V.R. Shetty, R. Pooja, R.L. Malghan, Safeguarding against Cyber Threats: Machine Learning-Based Approaches for Real-Time Fraud Detection and Prevention. Engineering Proceedings. 59(1), (2023) 111. https://doi.org/10.3390/engproc2023059111

[4] O.A. Bello, K. Olufemi, Artificial intelligence in fraud prevention: Exploring techniques and applications challenges and opportunities. Computer science & IT research journal, 5(6), (2024) 1505-1520. https://doi.org/10.51594/csitrj.v5i6.1252

[5] L. Hernandez Aros, L.X. Bustamante Molano, F. Gutierrez-Portela, J.J. Moreno Hernandez, M.S. Rodríguez Barrero, Financial fraud detection through the application of machine learning techniques: a literature review. Humanities and Social Sciences Communications, 11(1), (2024) 1-22. https://doi.org/10.1057/s41599-024-03606-0

[6] N.B.O. Adelakun, N.E.R. Onwubuariri, N.G A. Adeniran, N.A. Ntiakoh, Enhancing fraud detection in accounting through AI: Techniques and case studies. Finance and Accounting Research Journal, 6(6), (2024) 978–999. https://doi.org/10.51594/farj.v6i6.1232

[7] G. Zioviris, K. Kolomvatsos, G. Stamoulis, An intelligent sequential fraud detection model based on deep learning. The Journal of Supercomputing, 80, (2024) 14824–14847. https://doi.org/10.1007/s11227-024-06030-y

[8] M.N. Alatawi, Detection of fraud in IoT based credit card collected dataset using machine learning. Machine Learning With Applications, 19, (2025) 100603. https://doi.org/10.1016/j.mlwa.2024.100603

[9] T. Ghrib, Y. Khaldi, P.S. Pandey, Y.A. Abusal, Advanced Fraud Detection In Card-Based Financial Systems Using A Bidirectional Lstm-Gru Ensemble Model. Applied Computer Science, 20(3), (2024) 51–66. https://doi.org/10.35784/acs-2024-28

[10] F. Khaled Alarfaj, S. Shahzadi, Enhancing Fraud Detection in Banking with Deep Learning: Graph Neural Networks and Autoencoders for Real-Time Credit Card Fraud Prevention, IEEE Access, 13, (2025) 20633-20646. https://doi.org/10.1109/ACCESS.2024.3466288

[11] A.J. Keya, H.H. Shajeeb, M.S. Rahman, M.F. Mridha, FakeStack: Hierarchical Tri-BERT-CNN-LSTM Stacked Model for Effective Fake News Detection. PLoS One, 18(12), (2023) e0294701. https://doi.org/10.1371/journal.pone.0294701

[12] G.R.Jainish, A. Alwin Infant, Attention Layer Integrated BiLSTM for Financial Fraud Prediction. Multimedia Tools and Applications, 83, (2024) 80613–80629. https://doi.org/10.1007/s11042-024-18764-1

[13] H. Du, L. Lv, A. Guo, H. Wang, AutoEncoder and LightGBM for credit card fraud detection problems. Symmetry, 15(4), (2023) 870. https://doi.org/10.3390/sym15040870

[14] T.H. Lin, J.R. Jiang, Credit card fraud detection with autoencoder and probabilistic random forest. Mathematics, 9(21), (2021) 2683. https://doi.org/10.3390/math9212683

[15] D. Cheng, S. Xiang, C. Shang, Y. Zhang, F. Yang, L. Zhang, Spatio-Temporal Attention-Based Neural Network for Credit Card Fraud Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34(1), (2020) 362-369. https://doi.org/10.1609/aaai.v34i01.5371

[16] J. Rubio, P. Barucca, G. Gage, J. Arroyo, R. Morales-Resendiz, Classifying payment patterns with artificial neural networks: An autoencoder approach. Latin American Journal of Central Banking, 1(1-4), (2020) 100013. https://doi.org/10.1016/j.latcb.2020.100013

[17] W. Hilal, S.A. Gadsden, J. Yawney, Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Systems with Applications, 193, (2022) 116429. https://doi.org/10.1016/j.eswa.2021.116429

[18] J.K. Afriyie, K. Tawiah, W.A. Pels, S. Addai-Henne, H.A. Dwamena, E.O. Owiredu, J. Eshun, A Supervised Machine Learning Algorithm for Detecting and Predicting Fraud in Credit Card Transactions. Decision Analytics Journal, 6, (2023) 100163. https://doi.org/10.1016/j.dajour.2023.100163

[19] M. Thilagavathi, R. Saranyadevi, N. Vijayakumar, K. Selvi, L. Anitha, K. Sudharson, (2024) AI-Driven Fraud Detection in Financial Transactions with Graph Neural Networks and Anomaly Detection. 2024 International Conference on Science Technology Engineering and Management (ICSTEM), IEEE, India. https://doi.org/10.1109/ICSTEM61137.2024.10560838

[20] S.C. Dubey, K.S. Mundhe, A.A. Kadam, (2020) Credit Card Fraud Detection Using Artificial Neural Network and Back Propagation. 4th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, India. https://doi.org/10.1109/ICICCS48265.2020.9120957

[21] G. Sahoo, S.S. Sahoo, Accounting Fraud Detection Using K-Means Clustering Technique. Machine Learning and Information Processing. Intelligent Systems and Computing, 1311 (2021) 247-259. https://doi.org/10.1007/978-981-33-4859-2_17

[22] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, G. Bontempi, Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), (2018) 3784-3797. https://doi.org/10.1109/TNNLS.2017.2736643

[23] Q. Fan, J. Yang, A Denoising Autoencoder Approach for Credit Risk Analysis. ICCAI '18: Proceedings of the 2018 International Conference on Computing and Artificial Intelligence, (2018) 62-65. https://doi.org/10.1145/3194452.3194456

[24] F.A. Almarshad, G.A. Gashgari, A.I.A. Alzahrani, Generative Adversarial Networks-Based Novel Approach for Fraud Detection for the European Cardholders 2013 Dataset. IEEE Access, 11 (2023) 107348-107368. https://doi.org/10.1109/ACCESS.2023.3320072

[25] P. Papadimitroulas, L. Brocki, N.C. Chung, W. Marchadour, F. Vermet, L. Gaubert, V. Eleftheriadis, D. Plachouris, D. Visvikise, G.C. Kagadis, Artificial Intelligence: Deep Learning in Oncological Radiomics and Challenges of Interpretability and Data Harmonization. European Journal of Medical Physics, 83, (2021) 108-121. https://doi.org/10.1016/j.ejmp.2021.03.009

[26] P. Subrahmanyan, B. Jayachitra, S. Nandi, K. Selvi, V. Ramu, K. Sudharson, AI-Enhanced Consumer Behavior Analysis in Digital Environments with BERT Optimization. International Conference on Science Technology Engineering and Management (ICSTEM), IEEE, India. https://doi.org/10.1109/ICSTEM61137.2024.10560773

[27] I. Azuri, I. Rosenhek-Goldian, N. Regev-Rudzki, G. Fantner, S.R. Cohen, The Role of Convolutional Neural Networks in Scanning Probe Microscopy: A Review. Beilstein Journal of Nanotechnology, 12 (2021) 878-901. https://doi.org/10.3762/bjnano.12.66

[28] S. Saad, I. Nadher, S.M. Hameed, Credit Card Fraud Detection Challenges and Solutions: A Review. Iraqi Journal of Science, 65(4), (2024) 2287-2303. https://doi.org/10.24996/ijs.2024.65.4.42

[29] A. Alourani, F. Ashfaq, N.Z. Jhanjhi, N.A. Khan, BiLSTM- and GNN-Based Spatiotemporal Traffic Flow Forecasting with Correlated Weather Data. Journal of Advanced Transportation, (2023) 1-17. https://doi.org/10.1155/2023/8962283

[30] N.H.O. Bello, N.C. Idemudia, N.T.V. Iyelolu, Implementing Machine Learning Algorithms to Detect and Prevent Financial Fraud in Real-Time. Computer Science and IT Research Journal, 5(7), (2024) 1539-1564. https://doi.org/10.51594/csitrj.v5i7.1274

[31] H. An, R. Ma, Y. Yan, T. Chen, Y. Zhao, P. Li, J. Li, X. Wang, D. Fan, C. Lv, Finsformer: A Novel Approach to Detecting Financial Attacks Using Transformer and Cluster-Attention. Applied Science, 14(460), (2024) 460. https://doi.org/10.3390/app14010460

[32] K. Sudharson, G. Babu, R. Santhiya, C. Anita, Enhanced privacy-preserving federated convivial learning for internet of medical things (IoMT) through blockchain-enabled trust Q-learning. Journal of the National Science Foundation of Sri Lanka, 52(4), (2025) 501–514. https://doi.org/10.4038/jnsfsr.v52i4.11923

[33] D. Breskuvienė, G. Dzemyda, Enhancing credit card fraud detection: highly imbalanced data case. Journal of Big Data, 11, (2024) 182. https://doi.org/10.1186/s40537-024-01059-5

[34] M.A. Al-Khasawneh, M. Faheem, D.M. Alsekait, A. Abubakar, G.F. Issa, Hybrid neural network methods for the detection of credit card fraud. Security and Privacy, 8(1), (2025). https://doi.org/10.1002/spy2.500

[35] S. Misra, S. Thakur, M. Ghosh, S.K. Saha, An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction. Procedia Computer Science, 167, (2020) 254-262. https://doi.org/10.1016/j.procs.2020.03.219

[36] I. Benchaji, S. Douzi, B. El Ouahidi, J. Jaafari, Enhanced Credit Card Fraud Detection Based on Attention Mechanism and LSTM Deep Model. Journal of Big Data, 8(151), (2021). https://doi.org/10.1186/s40537-021-00541-8

[37] I.Y. Hafez, A.Y. Hafez, A. Saleh, A.A. Abd El-Mageed, A.A. Abohany, A systematic review of AI-enhanced techniques in credit card fraud detection. Journal of Big Data, 12, (2025) 6. https://doi.org/10.1186/s40537-024-01048-8

Authors Contribution Statement

K. Sudharson - Conceptualization, Design, Wring original Manuscript. S. Varsha - Data acquisition, Result Analysis, Writing review, editing. S. Rajalakshmi-Result Analysis, Writing review, editing. D. Rajalakshmi - Statistical validation and Results analysis. R. Santhiya-Writing review, editing. All the authors read and approved the final version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Has this article screened for similarity?

Yes

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

About the License