Leveraging LSTM and Attention for High-Accuracy Credit Card Fraud Detection

Ola Imran Obaid[1,*], Ali Yakoob Al-Sultan[2]

[1]College of Science for Women-Computer Science Dept, University of Babylon, Babylon, Iraq

Emails: ola.alghazaly.gsci135@student.uobabylon.edu.iq; ali.alsultan@uobabylon.edu.iq

Abstract

The increasing use of credit cards, especially for online payments, has led to a significant increase in fraud involving credit card payment technologies. Financial companies must enhance fraud detection systems to mitigate significant losses. This study introduces a methodology for developing a credit card fraud detection system that uses the Synthetic Minority Oversampling Technique (SMOTE) to address an imbalanced dataset problem and an attention layer to identify important features in the input sequence, two long short-term memory (LSTM) layers modeling long-run dependencies within a sequence of transactions, a dropout layer that neglects values lower than 0.3, and two dense layers, which allows enhancing the accuracy of prediction of fraudulent transactions. When implemented, the proposed system achieves an accuracy of 0.9434% on the IEEE dataset, 0.9850% on the Banksim dataset, and 0.9757% on the European dataset. This methodology shows improvements in fraud detection, emphasizing its ability to enhance financial security systems and reduce misclassification in credit card transactions.

Keywords: Fraud Detection; Credit Cards; Deep Learning; SMOTE; Attention Mechanism; Long short-Term Memory

1. Introduction

The exponential advancement of technology and the use of the Internet becoming an integral part of everyday life, people have become accustomed to using their credit cards for web-based purchases and Procedure financial transactions [1]. The study was conducted in nine countries and included approximately 3,700 customers. The results indicate that traditional shopping will change to adopting online purchasing permanently[2]. Scammers innovate tactics and discover new vulnerabilities. According to the European Central Bank, they were quick to adapt new methods to carry out deceptive transactions [3].

Annually, a substantial amount of money is lost due to fraudulent operations. Scammers create fraudulent web pages that closely mimic genuine websites. This web page offers a range of discounts to encourage customers to purchase the available products. Once the purchase is completed, the fraudster collects all relevant information related to the card and then exploits it to carry out fraudulent operations [4].

Fraud detection systems are essential to reduce the possibility of financial losses. Developing a powerful system capable of detecting fraudulent transactions is necessary using deep learning algorithms that deal with more complex patterns, and these are considered effective methods for fraud detection systems[5]. This study presents the development of a credit card fraud detection system that uses LSTM with an attention mechanism. It allows the neural network to automatically identify the most important data features for the classification task, leading to improved detection performance. The following are important contributions to this paper.

Incorporating the attention mechanism into LSTM recurrent networks, the classifier can effectively determine which features of the data sequence to focus on for making accurate fraud decisions. The misclassifications were

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

reduced FP and FN to enhance the detection system and increase accuracy. The proposed system will be evaluated using three types of data: European, Banksim, and IEEE.

This paper was organized in the following way: Section 2 provides significant related works to discover fraud in credit cards. Section 3 discusses methodological applications. Section 4 presents the findings and discussion. Section 5: Evaluating the proposed system in comparison to other relevant studies. Finally, conclusions and future work.

2. Related Work

Given the considerable influence of stolen credit cards on the financial market, numerous financial companies have dedicated plenty of money to developing fraud detection systems.

Najadat et al. (2020) [6]. The researchers employed a bidirectional long-short-term memory (BLSTM) with a max-pooling layer oversampling balancing technique. It achieved the highest performance use IEEE with an AUC of 0.9137%, Precision of 0.9114%, and f1-score of 0.9281%.Vengatsan et al. (2020) [7]. The objective of this study is to identify instances of fraud by using machine learning algorithms, such as logistic regression (LR) and k-nearest neighbors (KNN), without using balancing techniques to process the data set. The European dataset has a precision rate of 95%, a recall rate of 72%, and an f1-score of 82%. Benchaji et al. (2021) [8]. The system was developed to identify fraudulent transactions using LSTM with unified manifold approximation and projection (UMAP) for feature selection, attention mechanism, and SMOTE. The results showed that the European datasets had a recall of 0.911%, a precision of 0.967%, a Banksim precision of 0.974%, and a precision of 0.972%. Amusan et al. (2021) [9]. An under-sample was used to balance the data and using several machine-learning algorithms. The study's results revealed that the Random Forest (RF) algorithm achieved a high accuracy of 0.9519% on European data. The use of the under sampling technique leads to the removal of some data in the majority category. Al-Faqir et al. (2022) [10]. This paper introduces a model that aims to decrease the occurrence of fraudulent actions by integrating the outcomes of three advanced deep learning models: Convolutional Neural Network (CNN), Auto encoder (AE), and Recurrent Neural Network (RNN). The European credit card dataset was used, reaching an accuracy of 0.949%without selection of the best features. Kanika et al. (2022) [11]. The decision threshold is adjusted by implementing a new loss function known as Weighted Hard Reduction Focal Loss (WH-RFL) while addressing the class imbalance problem to enhance fraud detection. Results obtained using the European dataset show an accuracy of 0.9125%, and the Banksim dataset shows an accuracy of 0.9311%. Mniai et al. (2023) [12]. This research introduces a model for fraud detection that uses under-sampling, feature selection, and support vector data description (SVDD). The findings were showcased using the European Dataset, revealing that the SVDD achieved an accuracy rate of 93%. The under-sampling algorithm's low accuracy can be attributed to removing specific data points. Jiang et al. (2023) [13]. Building a fraud detection network, UAAD-FDNet, which uses unsupervised intentional anomaly detection. One of the techniques used in this study is to apply an auto encoder with attention to features. UAAD-FD Net achieved an AUC of 0.839% and a recall of 0.602% when tested on the IEEE dataset, and the European dataset achieved an AUC of 0.943% and a recall of 0.751%. Peneti et al. (2024) [14]. They proposed LightGBM machine learning algorithm implementation on an IEEE dataset that involved handling unbalanced data. The results indicate a recall rate of 83% and an F1 score of 45%. Prabha et al. (2024) [15]. Researchers focus on addressing the class imbalance problem by determining the best threshold with the use of machine learning algorithms. The best findings of the XGBoost algorithm on the IEEE dataset obtained a recall of 0.6258% and an f1-score of 0.7325%. Based on previous works above, using balanced data effectively identifies fraud cases. This paper attempts to improve the accuracy by developing a model using LSTM with an attention layer compared to previous works.

3. Methodology

This section comprehensively presents the data description, preprocessing, and application of LSTM with an attention mechanism for data classification, as depicted in Figure 1.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

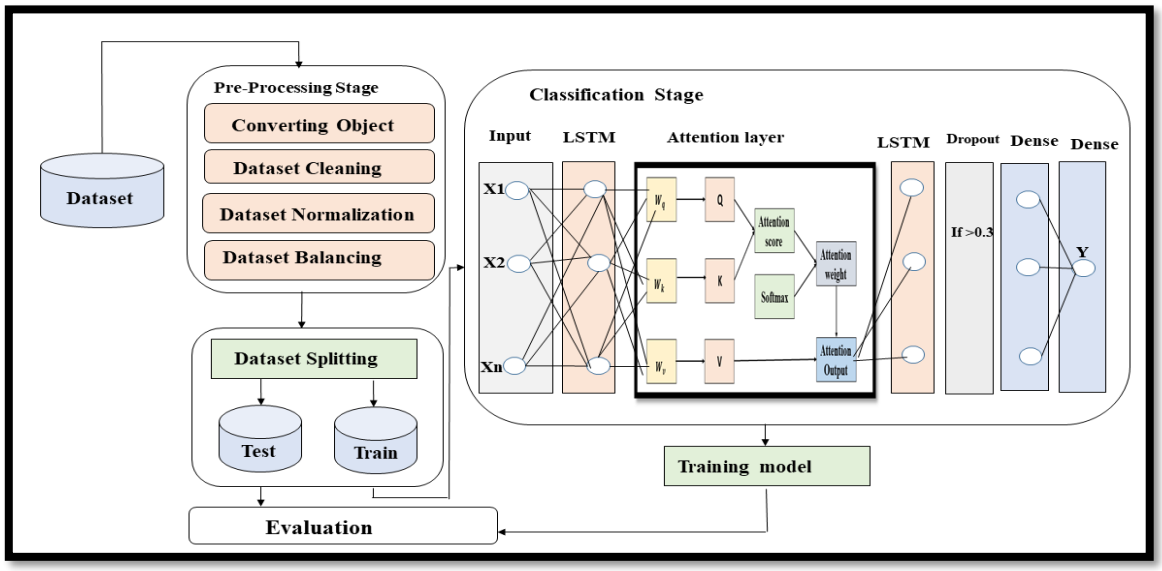

Figure.1 Diagram of the proposed system

3.1 Dataset Description

The proposed system is implemented on three datasets, European, Banksim, and IEEE.

3.1.1 European credit cards dataset

European credit cards dataset acquired from Kaggle. The dataset includes 284,807 transactions, of which only 492 are identified as fraudulent, and 284,315 are normal. The dataset exhibits a significant imbalance. The dataset consists of 31 features, including 'Amount' as transaction value, 'Time' as transaction time, and 28 more factors labelled V1 to V28. Class is target establishes each transaction's attribute, with a binary value of 1 indicating an illegal transaction while 0 indicating a legal transaction [16].

3.1.2 IEEE credit cards dataset

IEEE-CIS credit card datasets obtained from Kaggle. The dataset consists of 531,486 transactions, of which 18,450 are fraudulent and 513,036 are non-fraudulent transactions considered imbalanced. In addition, the dataset contains 269 features. Is fraud a binary target categorizing transactions into legal (0) and illegal (1)[6].

3.1.3 BankSim dataset

BankSim data is sourced from Kaggle and includes 594,643 transactions executed over 180 simulated days; 7,200 transactions were classified as “fraudulent,” while 587,443 transactions were classified as “genuine.” This synthetic dataset was created using BankSim, a fraud data simulator. BankSim uses a multi-agent simulation method using the example of aggregated real transaction data provided by a Spanish bank. BankSim uses agents from three distinct categories, merchants, customers, and fraudsters, to copy the original bank statements. These agents communicate with each other over a series of simulated days, creating a purchase transaction record that resembles the original bank data and consists of 10 features. Fraud is a binary target [17].

3.2 Pre-processing Stage

The proposed system uses a set of basic operations known as pre-processing, which constitute the initial stage. Thus, at this point, the precision of the classification model is assured. The following steps in data pre-processing are described:

3.2.1 Converting object dataset to numbers

Banksim data containing three data types: numeric, float, and object preprocessed by converting the object's column type to a second type, categorical, using astype from the pandas' library and converting it to numeric.

3.2.2 Dataset Cleaning

Dataset cleaning is an important step in pre-processing. It involves identifying missing values, either empty cells or cells containing the word "null" in the dataset, as they can lead to inaccurate outcomes, which are addressed through their deletion [18].

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

3.2.3 Dataset Normalization

Dataset normalization is an essential procedure in data management [19]. The purpose of min-max normalization is to enhance the efficiency and precision of the classification model so that the feature values are standardized within a consistent range between 0 and 1, as shown in equation (1)[20].

$$normalize(X) = \frac{X - X_{min}}{X_{\max} - X_{\min}} \quad (1)$$

X is a variable that represents features, X_{min} is the lowest value of the feature, and X_{max} is the highest value of the feature.

3.2.4 Dataset Balancing

The dataset used in this paper comprises more legitimate transactions, which represent the majority class, compared to fraudulent transactions, which represent the minority class [21].Training our model on this dataset will exhibit a bias towards legitimate transactions. This problem is addressed by employing the Synthetic Minority Over-Sampling Technique (SMOTE), which involves iteratively increasing the number of instances in the minority class to match the majority class[22]. As shown in the algorithm 1 [23].

Algorithm1 : SMOTE

Input: Dataset
Output: Dataset Balancing

Begin
Step 1: Determine the minority cases in dataset.
Step 2: Perform K-Nearest Neighbors calculation.
For each cases in the minority class
Compute the Euclidean distances for every entry in the dataset. Find the K nearest neighbors.
Step 3: Create Synthetic Examples
For each cases in the minority class
Select one of its K nearest neighbors at random.
Produce a synthetic sample by calculating the difference of the chosen adjacent sample and the initial sample
Step 4: Merge the initial dataset with the synthetic samples.
End For
End For
Step 5: Return Dataset Balancing
End

3.3 Classification Stage

The classification stage is important in data processing and consists of two basic steps: LSTM with attention mechanism and evaluation model.

3.3.1 LSTM with Attention Mechanism Model

The methodology presented in this paper utilizes a classifier that employs LSTM with attention mechanisms. Attention mechanisms have become essential in various deep learning models, especially natural language processing (NLP), time series prediction, and computer vision [24]. Each feature is assigned a different significance in the models that employ attention mechanisms. The attention mechanism allocates varying weights to various features based on their significance for the given task, enhancing the model's ability to comprehend the data more efficiently. Moreover, the attention mechanism enhances the neural model by ordering features based on their significance[25]. The key idea of this mechanism is to compute a weight distribution across the input features, giving greater values to those features considered of greater significance. The attention layer consists of attention scores, attention weights, and an attention output[26].

Every feature in the input sequence will be transformed into Q, K, and V using three separate matrices, as shown in equations [27].

$$Q = X.w_q \quad (2)$$
$$K = X.w_k \quad (3)$$
$$V = X.w_v \quad (4)$$

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

Where X represents the input, W_q represents the weight matrix for the query, W_k represent the weight matrix for the key, W_v represent the weight matrix for the value, and Q is represented as the query matrix comprising vectors that find pertinent information within the key matrix; K is represented as the key matrix comprises the vectors that are used to match the information in the query matrix, and V is represented as the value matrix comprises the vectors that yield the ultimate information[27] , which are displayed in the Figure 2.

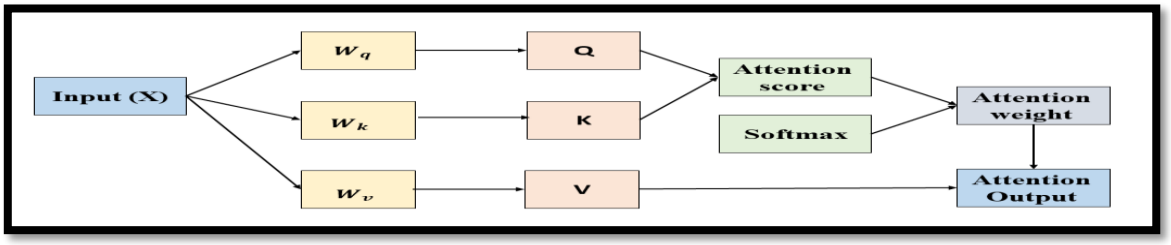

Figure 2. Structure for attention mechanism

Matrix multiplication involves computing the relationship between Q and K. The result that was thus obtained indicates an attention score, as depicted in Equation 5[27].

$$a_t = Q.K \quad (5)$$

Where: a_t represent attention score.

A Scaling matrix is performed to maintain gradients during training, preventing scores from becoming excessively large and avoiding the issue of gradients that decrease. The scaling matrix is passed into the softmax function and converted into attention weights, guaranteeing that the attention weights sum up to 1. These weights represent the significant relevant information, as depicted in Equation 6[28] .

$$w_t = \text{softmax}(\frac{a_t}{\sqrt{k^d}}) \quad (6)$$

Where: w_t represent attention weight, d is the dimensionality of query.

The output of the softmax function is subsequently multiplied by the Value Matrix (V), which represents the final output of the Attention mechanism, as shown in Equation 7 [28].

$$o_t = V.a_t \quad (7)$$

Where: o_t represent attention output.

The LSTM with attention mechanisms model structure. Below is a detailed analysis of the layers:

- LSTM_1 Layer: This layer receives the input data provided to the network. In this case, the input is a one-dimensional vector. It consists of 269 features in the IEEE dataset, 30 in the European dataset, and the Banksim dataset, which contains 8 features and extracts the best features based on the best weights. It consists of 100 units and uses an activation function (tanh).
- Attention layer: A dedicated layer uses attention mechanisms to determine the most important parts of the sequence by working to adjust the weights of the LSTM outputs.
- LSTM_2 layer: The second LSTM layer processes the output from the attention layer, including 32 units, and uses the activation function tanh.
- Dropout Layer: A dropout layer is added with a rate of 0.3 to prevent overfitting. During the training process, any individual neuron that exhibits undesirable features with a probability of less than 0.3 will be removed.
- Dense Layer: This layer consists of 50 neurons that are fully connected. Every neuron in this layer obtains input from each of the neurons in the prior layer. This layer works as an intermediate layer that aids the transfer of educated features to the end output.
- Dense layer: represents the last layer. It is specifically designed to perform binary classification tasks. It applies a sigmoid activation function to classify transactions as normal or fraudulent.
- The model employs an Adam optimizer with a learning rate of 0.0001 and employs binary cross-entropy as the loss function.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

- The model employs the model employs a Checkpoint to save the best weights during training when the lower validation loss and early stopping are implemented when the model's performance plateaus with consistent accuracy, leading to the ending of the training process.

3.3.2 Evaluation methodology

This paper evaluates the performance using a confusion matrix consisting of TP, FN, FP, and TN, through which the measures can be calculated: accuracy, precision, recall, and F1 score. In addition, the AUC measure can be calculated, as shown in the equations.

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \quad (8)$$

$$Precision = TP/(TP + FP) \quad (9)$$

$$\text{Sensitivity (Recall)} = \text{TP}/(\text{TP} + \text{FN}) \quad (10)$$

$$\text{F1 score} = 2 * ((\text{precision} * \text{recall})/(\text{precision} + \text{recall})) \quad (11)$$

$$\text{Specificity} = TN/(TN + FP \quad (12)$$

$$AUC = (Sensitivity + Specificity)/2 \quad (13)$$

"True Positive" (TP) refers to the number of cases correctly identified as positive. A "false negative" (FN) refers to many positive cases mistakenly recorded as negative. False positive (FP) refers to negative cases mistakenly identified as positive. "True negative" (TN) refers to correctly identified cases of negativity [29].

4. Result and Discussion

Preprocessing results and LSTM with attention mechanism results will be discussed for three data types.

4.1 The Results of Pre-processing

The pre-processing results consist of the conversion results and the balancing results.

4.1.1 Conversion Results

This type of processing on Banksim data. Two columns (Zipcodeori, Zipmerchant) were deleted because each column contains the same value. It includes column values of object type, which converts them to categorical and then to numeric.

4.1.2 Balancing Results

Table 1 displays the data processing procedure before and after applying SMOTE so that fraudulent transactions become equal to legitimate transactions.

Table 1: Data Balancing Result

Dataset	Before Balancing	Fraud	Normal	After Balancing
IEEE	531,486	18,450	513,036	1,026,072
European	284,807	492	284,315	568,630
Banksim	594,643	7,200	587,443	1,174,886

4.2 The Results of LSTM without attention layer

The first method was applied, which is the standard LSTM method, which comprises two layers: an LSTM layer with 200 units, a layer with 100 units, a dropout layer that disregards values less than 0.3, and 2 dense layers responsible for the classification output. Where it appeared the Total parameters for Banksim and dataset European are 287101, non-trainable are 0, and the Total parameters for the IEEE dataset are 31,013, non-trainable are 0. We applied the model to three types of data, as illustrated in Table 2.

Table 2: LSTM without attention layer results

Dataset	Accuracy	Precision	Recall	F1 score	AUC
IEEE	0.9036	0.92	0.90	0.90	0.96
European	0.9254	0.93	0.93	0.93	0.97
Banksim	0.9810	0.98	0.98	0.98	0.99

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

Observing the results shown in Table 2, we can see that since important financial transactions require accuracy because decision-making gives fraudulent or natural results, they are accurate and sensitive. Therefore, we turned to a model that increases accuracy by adding an attention layer.

4.3 The Results of LSTM with attention

The proposed system is evaluated using three types of data: Banksim, IEEE, European, and the dataset was partitioned into two groups: training data at 80% and testing at 20%.

4.3.1 IEEE Dataset Results

Figure 3(a) shows the accuracy values for training and testing over 70 epochs using the IEEE dataset to achieve a training accuracy of 0.9417% and a test accuracy of 0.9434%, precision, and F1 score of 0.9434%, recall of 0.9438%, AUC of 0.9437.

Figure 3 (b) shows that the training process through 70 epochs in the IEEE dataset started with a loss function for training at epoch 1 of (0.5587), arrived at a low of (0.1589), and for the testing of (0.4854), arrived at a low of (0.1517) because the best weight in the epoch 70 was obtained, was stored for the model based on the lowest val loss.

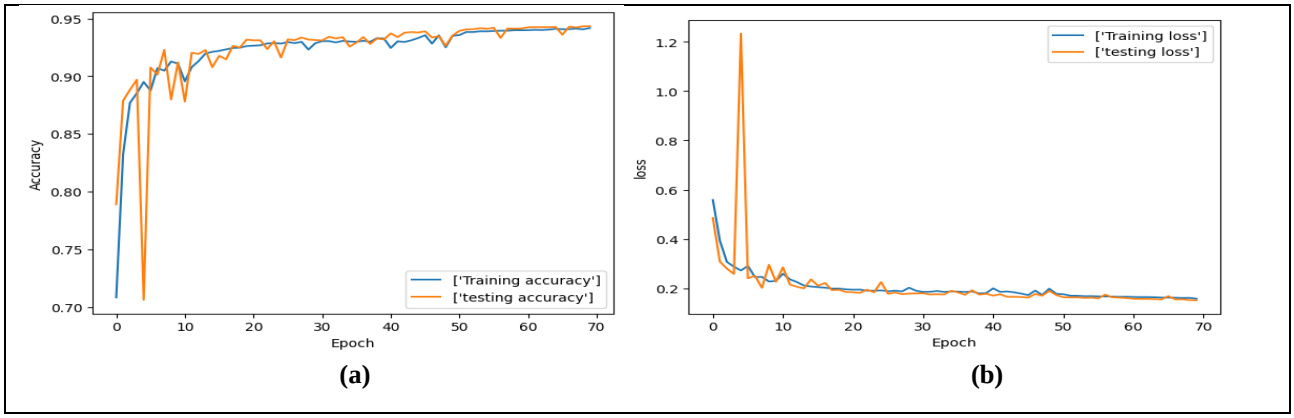

Figure 3. (a) The accuracy for the IEEE dataset; (b) The loss function for the IEEE dataset

The confusion matrix depicted in Figure 4 utilizes the IEEE dataset. The number of false classifications was determined to be (FP=4254) and (FN=7368), resulting in an accuracy of 0.9437%. However, despite the large size of the data and the number of features 269, the total number of correctly identified fraudulent and natural (TP=98395) (TN=95198) transactions was high.

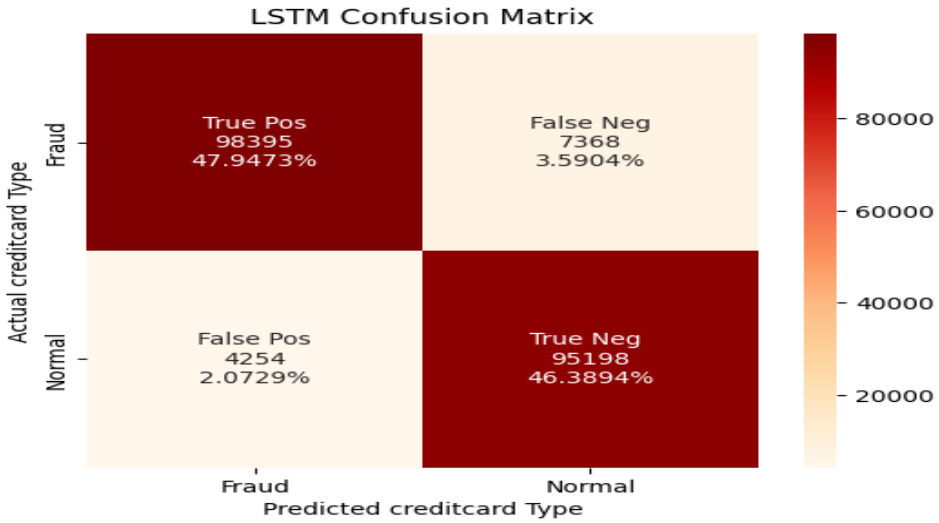

Figure 4. The confusion matrix for the IEEE dataset

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

4.3.2. European Dataset Results

The European dataset was trained over 70 epochs and obtained an accuracy for training of 0.9742% and an accuracy for testing of 0.9757%, as shown in Figure 5(a). The small difference between the training and testing accuracy values indicates that the model effectively avoids overfitting.

The European dataset was used over 70 epochs, starting with a loss function for training at epoch 1 of (0.3887), which arrived at a low of (0.0771), and for the testing of (0.2709), arrived at a low of (0.0736) because the best weight in the epoch 69 was obtained, was stored for the model based on the lowest val loss that busted by monitor indicator, as depicted in Figure 5 (b). Relatively low loss values indicate that the model performs well in training and testing.

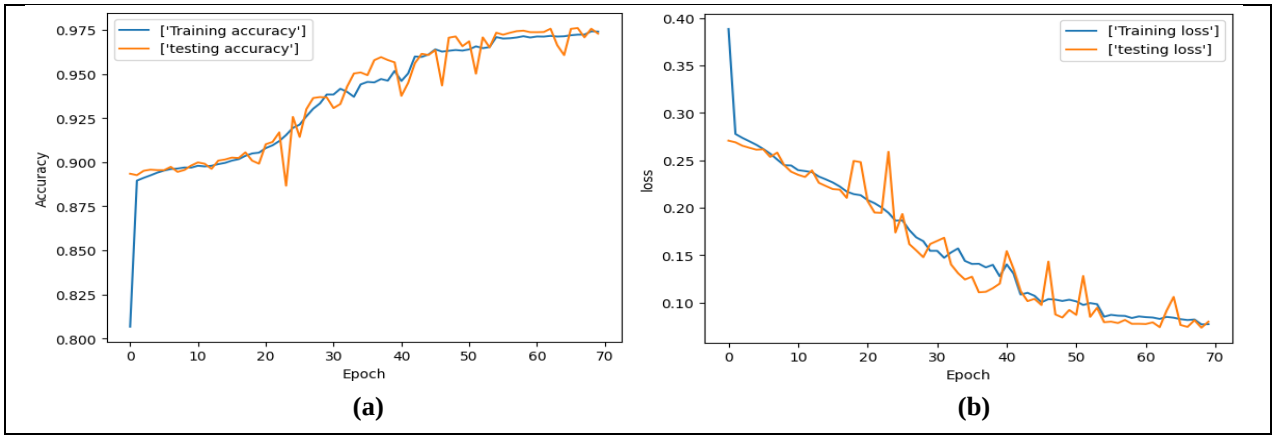

Figure 5. (a) The accuracy for the European dataset: (b) The loss function for the European dataset

The improvement in the model's performance is attributed to the attention mechanism, which effectively extracts the most pertinent information for the classification task.

The confusion matrix in Figure.6 displays the model's performance on European data. It shows that the correct classification values of (TP = 55820) and (TN = 55140) are high compared to faille values (FP=1054) and (FN=1712), The efficacy of this model can be ascribed to its capacity to efficiently acquire knowledge from training data and reduce misclassification rates due to fraudulent transformations that closely resemble natural transformations, making them indistinguishable.

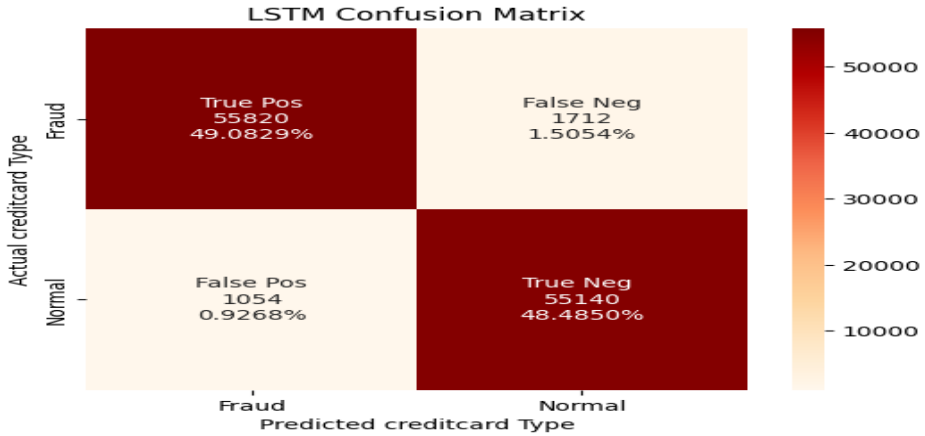

Figure 6. The confusion matrix for the European dataset

4.3.3 Banksim Dataset Result

The Banksim dataset was used and executed over 70 epochs, achieving 98.34% training accuracy and 98.50% testing accuracy, recall of 98.51%, and F1 score of 98.50%, as shown in Figure 7(a).

The Banksim dataset is shown in Figure 7(b) for over 70 epochs. It started with a loss function for training at epoch 1 of (0.6712), which decreased to (0.0497), and for the testing of (0.6486), that decrease to (0.0447), due to best weight in the epoch 69 was obtained.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

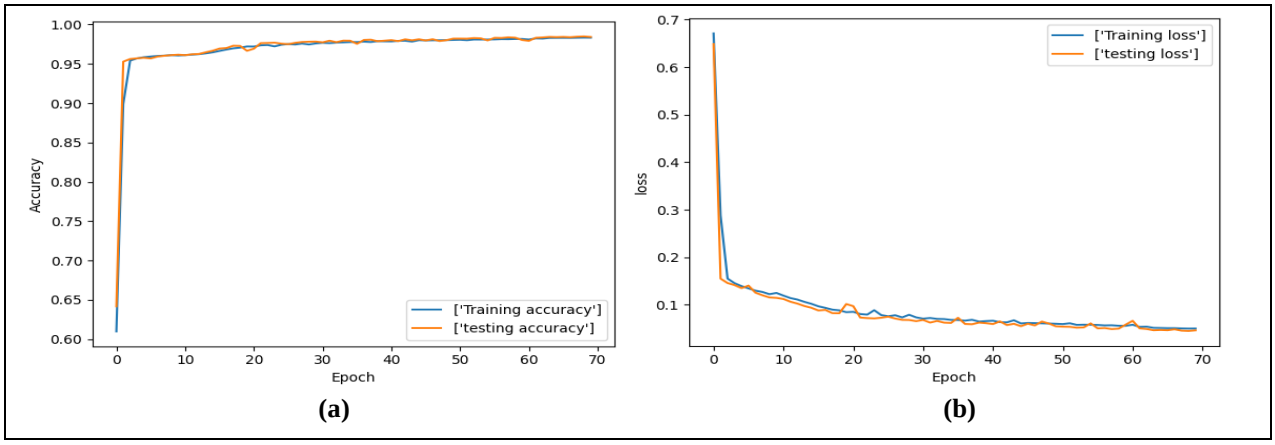

Figure 7. (a) The accuracy for the Banksim dataset; (b) The loss function for the Banksim dataset

The confusion matrix in Figure 8 displays the Banksim data. The number of correctly detected normal transactions (TN=117125) and the (FN=759) indicate the number of false classifications fraudulent transaction is classified as normal.

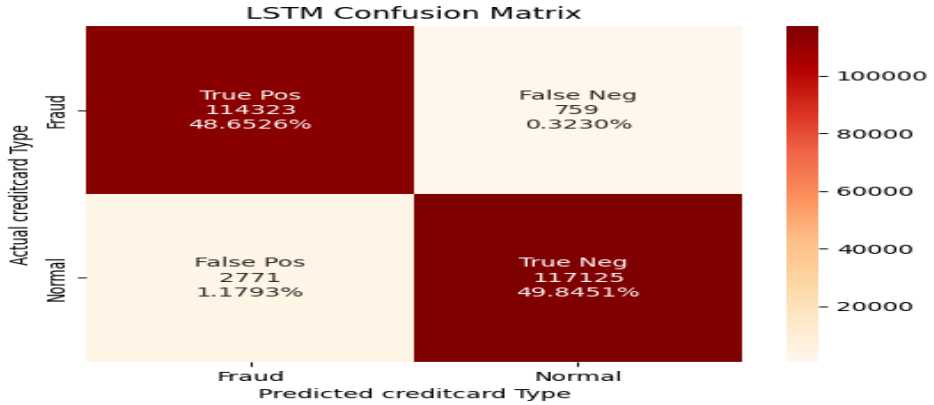

Figure 8. The confusion matrix for the Banksim dataset

5. Comparison of the Proposed System with Related Works

Table 3 compares the suggested model with prior studies that employ machine learning and deep learning algorithms. The proposed system exhibited superior performance compared to previous researchers who utilized the same dataset.

Table 3: Comparison with related works

Reference	Technique	Dataset	Accuracy	Precision	Recall	F1 score	AUC
[6]	BLSTM with a max-pooling layer	IEEE	-	0.9114	-	0.9281	0.9137
[7]	KNN	European	-	0.95	0.72	0.82	-
[8]	UMAP and LSTM with attention	European	0.9670	-	0.911	-	-
		Banksim	0.974	-	0.972	-	-
[9]	RF	European	0.951	-	0.922	-	-
[10]	Ensemble (CNN, AE, RNN)	European	0.949	0.971	0.838	-	-
[11]		European	0.9311	-	-	-	-

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

	Threshold with WH-RFL loss function	Banksim	0.9125	-	-	-	-
[12]	Support Vector Data Description SVDD	European	0.93	0.90	0.97	0.93	-
[13]	(UAAD-FDNet)	IEEE	-	-	0.602	0.7349	0.839
		European	-	-	0.751	0.848	0.943
[14]	LightGBM	IEEE	-	-	0.83	0.45	
[15]	threshold with XGBoost	IEEE	-	0.883	0.625	0.732	0.76
Proposed system	LSTM with attention	IEEE	0.9434	0.9434	0.9438	0.9434	0.9437
		European	0.9757	0.9757	0.9757	0.9757	0.9757
		Banksim	0.9850	0.9849	0.9851	0.9850	0.9851

The main reasons for the proposed model obtaining high accuracy are: Firstly, the SMOTE was used to balance the dataset. Secondly, adding an attention layer focusing on important features by correcting the weights that are used best to handle large datasets in credit card fraud detection. Third, the use of early stopping during model training and Adam's optimizer algorithm.

6. Conclusions and Future Works

The study introduces a methodology for detecting illegal transactions in credit cards through the used model of LSTM with an attention mechanism. The experimental results indicate that the detection suggested model has high efficiency in terms of accuracy and reduces classification error when compared with related studies because of the preprocessing process for the dataset that includes removing missing values, normalization, and the Synthetic Minority Oversampling Technique (SMOTE) for addressing imbalanced data problem and add an attention mechanism that automatically focuses on the most important features for the classification task by assigning and adjusting weights to the features with LSTM networks to model long-term dependencies and accurately detect fraudulent transactions. This study contributes to the use of deep learning algorithms in credit card security and provides valuable perspectives for continued improvements in fraud detection systems. Ensuring the long-term continuity and progress of these systems in the face of future challenges. In future work, we aspire to improve accuracy by adding feature selection algorithms to our model features and using the proposed system in real time.

Reference

[1] H. John and S. Naaz, "Credit Card Fraud Detection using Local Outlier Factor and Isolation Forest International Journal of Computer Sciences and Engineering Open Access Credit Card Fraud Detection using Local Outlier Factor and Isolation," no. April, 2019, doi: 10.26438/ijcse/v7i4.10601064.

[2] H. Ahmad, B. Kasasbeh, B. Aldabaybah, and E. Rawashdeh, "Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS)," Int. J. Inf. Technol., vol. 15, no. 1, pp. 325–333, 2023, doi: 10.1007/s41870-022-00987-w.

[3] H. Wang, P. Zhu, X. Zou, and S. Qin, "An ensemble learning framework for credit card fraud detection based on training set partitioning and clustering," Proc. - 2018 IEEE SmartWorld, Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Cloud Big Data Comput. Internet People Smart City Innov. SmartWorld/UIC/ATC/ScalCom/CBDCo, pp. 94–98, 2018, doi: 10.1109/SmartWorld.2018.00051.

[4] Y. Jain, N. Tiwari, S. Dubey, and S. Jain, "A comparative analysis of various credit card fraud detection techniques," Int. J. Recent Technol. Eng., vol. 7, no. 5, pp. 402–407, 2019.

[5] T. Y. Wu and Y. T. Wang, "Locally Interpretable One-Class Anomaly Detection for Credit Card Fraud Detection," Proc. - 2021 Int. Conf. Technol. Appl. Artif. Intell. TAAI 2021, pp. 25–30, 2021, doi: 10.1109/TAAI54685.2021.00014.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

[6] H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, "Credit Card Fraud Detection Based on Machine and Deep Learning," 2020 11th Int. Conf. Inf. Commun. Syst. ICICS 2020, no. May, pp. 204–208, 2020, doi: 10.1109/ICICS49469.2020.239524.

[7] K. Vengatesan, A. Kumar, S. Yuvraj, V. D. Ambeth Kumar, and S. S. Sabnis, "Credit card fraud detection using data analytic techniques," Adv. Math. Sci. J., vol. 9, no. 3, pp. 1185–1196, 2020, doi: 10.37418/amsj.9.3.43.

[8] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," J. Big Data, 2021, doi: 10.1186/s40537-021-00541-8.

[9] E. Amusan et al., "Credit Card Fraud Detection on Skewed Data using Machine Learning Techniques," LAUTECH J. Comput. Informatics, vol. 2, no. 1, pp. 49–56, 2021, [Online]. Available: https://www.researchgate.net/publication/354780529

[10] S. Al-Faqir and O. Ouda, "Credit Card Frauds Scoring Model Based on Deep Learning Ensemble," J. Theor. Appl. Inf. Technol., vol. 100, no. 14, pp. 5223–5234, 2022.

[11] Kanika, J. Singla, A. K. Bashir, Y. Nam, N. U. I. Hasan, and U. Tariq, "Handling class imbalance in online transaction fraud detection," Comput. Mater. Contin., vol. 70, no. 2, pp. 2861–2877, 2022, doi: 10.32604/cmc.2022.019990.

[12] A. Mniai, M. Tarik, and K. Jebari, "A Novel Framework for Credit Card Fraud Detection," IEEE Access, vol. 11, no. October, pp. 112776–112786, 2023, doi: 10.1109/ACCESS.2023.3323842.

[13] S. Jiang, R. Dong, J. Wang, and M. Xia, "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," Systems, vol. 11, no. 6, pp. 1–14, 2023, doi: 10.3390/systems11060305.

[14] S. Peneti, S. Rama Krishna, A. Kiran, and H. K. Tripathy, "Credit Card Fraud Detection Using Machine Learning," Proc. 2nd Int. Conf. Adv. Smart, Secur. Intell. Comput. ASSIC 2024, no. November, 2024, doi: 10.1109/ASSIC60049.2024.10508010.

[15] D. P. Prabha and C. V. Priscilla, "Estimation of optimal threshold shifting to handle class imbalance in credit card fraud detection using machine learning techniques," AIP Conf. Proc., vol. 2802, no. 1, 2024, doi: 10.1063/5.0182386.

[16] V. N. Dornadula and S. Geetha, "Credit Card Fraud Detection using Machine Learning Algorithms," Procedia Comput. Sci., vol. 165, pp. 631–641, 2019, doi: 10.1016/j.procs.2020.01.057.

[17] I. Benchaji, S. Douzi, and B. El Ouahidi, "Credit card fraud detection model based on LSTM recurrent neural networks," J. Adv. Inf. Technol., vol. 12, no. 2, pp. 113–118, 2021, doi: 10.12720/jait.12.2.113-118.

[18] M. J. Madhurya, H. L. Gururaj, B. C. Soundarya, K. P. Vidyashree, and A. B. Rajendra, "Exploratory analysis of credit card fraud detection using machine learning techniques," vol. 3, no. April, pp. 31–37, 2022, doi: 10.1016/j.gltp.2022.04.006.

[19] S. Rao, P. Poojary, J. Somaiya, and P. Mahajan, "a Comparative Study Between Various Preprocessing Techniques for Machine Learning," Int. J. Eng. Appl. Sci. Technol., vol. 5, no. 3, pp. 431–438, 2020, doi: 10.33564/ijeast.2020.v05i03.069.

[20] O. S. Yee, S. Sagadevan, and N. H. A. H. Malim, "Credit card fraud detection using machine learning as data mining technique," J. Telecommun. Electron. Comput. Eng., vol. 10, no. 1–4, pp. 23–27, 2018.

[21] U. Ependi, A. F. Rochim, and A. Wibowo, "A Hybrid Sampling Approach for Improving the Classification of Imbalanced Data Using ROS and NCL Methods," Int. J. Intell. Eng. Syst., vol. 16, no. 3, pp. 345–361, 2023, doi: 10.22266/ijies2023.0630.28.

[22] Z. Zhao and T. Bai, "Using SMOTE and Machine Learning Algorithms," Entropy, vol. 24, no. 1157, pp. 1–18, 2022, doi: org/10.3390/e24081157.

[23] Y. Sun et al., "Borderline SMOTE Algorithm and Feature Selection-Based Network Anomalies Detection Strategy," Energies, vol. 15, no. 13, 2022, doi: 10.3390/en15134751.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024

[24] M. H. Al-Tai and B. M. Nema, "Detecting Arabic Misinformation Using an Attention Mechanism-Based Model," Iraqi J. Comput. Sci. Math., vol. 5, no. 1, pp. 285–298, 2024, doi: 10.52866/ijcsm.2024.05.01.020.

[25] H. Song, K. Huang, X. Ji, and F. Wan, "Fraud web page detection based on bidirectional LSTM and attention mechanism," vol. 12329, no. Aiea, p. 100, 2022, doi: 10.1117/12.2646903.

[26] M. Fazil, A. K. Sah, and M. Abulaish, "DeepSBD: A Deep Neural Network Model with Attention Mechanism for SocialBot Detection," IEEE Trans. Inf. Forensics Secur., vol. 16, pp. 4211–4223, 2021, doi: 10.1109/TIFS.2021.3102498.

[27] K. Mohiuddin et al., "Retention Is All You Need," Int. Conf. Inf. Knowl. Manag. Proc., no. Nips, pp. 4752–4758, 2023, doi: 10.1145/3583780.3615497.

[28] G. Wang, J. Ma, and G. Chen, "Attentive statement fraud detection: Distinguishing multimodal financial data with fine-grained attention," Decis. Support Syst., vol. 167, no. December 2022, p. 113913, 2023, doi: 10.1016/j.dss.2022.113913.

[29] A. Mohari, J. Dowerah, K. Das, F. Koucher, and D. J. Bora, "Credit Card Fraud Detection Techniques: A Review," no. July, pp. 157–166, 2021, doi: 10.1007/978-981-16-1048-6_12.

DOI: https://doi.org/10.54216/FPA.170115
Received: November 26, 2023 Revised: March 16, 2024 Accepted: July 20, 2024