Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A Hybrid Approach Adopted for Credit Card Fraud Detection Based on Deep Neural Networks and Attention Mechanism

Vikash Chander Maheshwari[1,*], Nurul Aida Osman[2], Norshakirah Aziz[1]

[1] Faculty of Science and IT, Department of Computer and Information Science, Universiti Teknologi PETRONAS, 32610, Iskandar, Perak, Malaysia
[2] Centre for Research in Data Science (CeRDaS), Computer Information Science Department, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia

ARTICLE INFO

Article history:
Received 5 May 2023
Received in revised form 12 August 2023
Accepted 19 August 2023
Available online 4 September 2023

Keywords:
CCFD; Deep Learning RNN-LSTM

ABSTRACT

Over the past few years, credit card fraud has become a serious problem as more individuals rely on credit cards for purchases. The significant increase in fraudulent activities can be attributed to advancements in technology and the prevalence of online transactions, leading to significant financial losses. To address this issue, an effective fraud detection system needs to be developed and put into practice. Machine learning techniques are commonly used to automatically detect credit card fraud, but they do not consider deceptive behaviour or behavioural issues that could lead to false alarms. The objective of this research is to determine how to identify instances of credit card fraud. This paper aimed to create a model using deep learning and SMOTE oversampling technique to anticipate credit card fraud. A Recurrent Neural Network with Long Short-Term Memory (RNN-LSTM) and an attention mechanism is suggested for detecting fraud. This model is known to be effective for processing sequential data with complex relationships between vectors. The performance of RNN-LSTM is compared to XGBoost, Random Forest, Naive Bayes, SVM, and ANN classifiers, and the experiments indicate that our proposed model achieves high accuracy of 99.4% and produces strong results. The suggested model has the potential to decrease financial losses worldwide by identifying instances of credit card scams or frauds.

1. Introduction

The prevalent use of technology and the rise of innovative payment methods like e-commerce and mobile payments have led to a spread of credit card transactions in recent times. The extensive occurrence of cashless transactions has made fraudsters frequently commit fraudulent activities and constantly modify their methods to avoid being caught. In modern fraud, the criminal does not have to be present at the crime scene, however they have a variety of methods to hide their identity and may carry out their malicious activities in the comfort of their own homes [1].

As per the latest Nilson's Report, credit card fraud is projected to result in losses of $43 billion within the next five years, and $408.5 billion worldwide within the next decade. Therefore, detecting

* *Corresponding author.*
E-mail address: vikash_22005211@utp.edu.my

https://doi.org/10.37934/araset.32.1.315331

credit card fraud is now more important than ever [2]. Fraudulent activities are prevalent in various sectors such as government, banking, stock market, and healthcare centres, among others. It is crucial to detect and identify fraudulent activities in order to reduce financial losses. To accomplish this, it is possible to employ deep learning (DL) and machine learning (ML) algorithms for distinguishing between fraudulent and non-fraudulent behaviours. Various DL and ML techniques, such as XGBoost, CatBoost, Naive Bayes, Random Forest, Logistic Regression, Neural Networks (NN), and bio-inspired algorithms such as, Genetic Algorithm (GA) have been applied to detect credit card fraud [3]. Figure 1 illustrates the report of fraudulent activities all over the world.

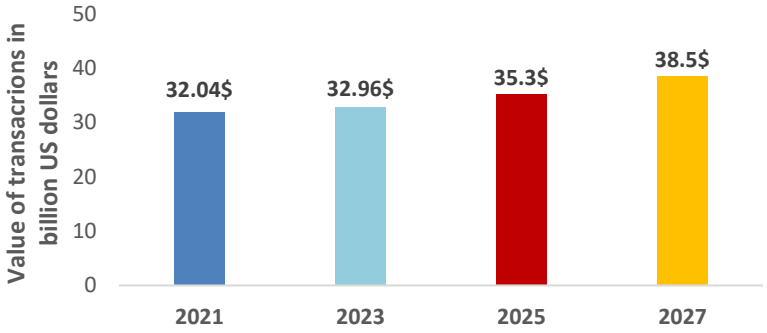

Fig. 1. Report of fraudulent transactions all over the world [2]

While we were trying to address the fraud problem, we encountered some additional problems. Specifically, when one class has more instances than the other, it leads to an imbalanced dataset. An imbalanced dataset means that there are significantly more data samples from one or some of the classes compared to the rest of the classes in the dataset. SMOTE (Synthetic Minority Oversampling Technique) is a popular pre-processing method used to handle imbalanced datasets. It involves oversampling the minority class by generating artificial examples in the feature vector space rather than the data space [4]. Fraudulent activities tend to vary in each attempt, which is referred to as "Concept drift." Detecting credit card fraud is challenging and can result in financial losses. Deep learning can be an effective approach to identify fraud, but concept drift poses a significant problem as fraudsters alter their techniques over time. As a result, the model's ability to predict new patterns reduces when the patterns change [5].

The research presented in this study proposes a novel technique to identify credit card fraud. This method involves utilizing (RNN-LSTM) networks, along with an attention mechanism. This mechanism enables the neural network to automatically prioritize the most relevant data items for the classification task through a data-based weighting system. This results in better detection accuracy by considering local information from each sequence term.

The main contribution of our paper is summarized below:

i. To enhance the learning process of classifiers, it is possible to utilize techniques such as feature selection and dimension reduction, such as Principal Component Analysis (PCA), to optimize the input data.
ii. To address the issue of imbalanced datasets and enhance the learning process, the Synthetic Minority Oversampling Technique (SMOTE) is being utilized.
iii. The combination of the attention mechanism and LSTM recurrent neural networks can significantly aid the classifier in identifying the relevant areas to focus its attention on within the input sequence, leading to an accurate determination of global fraud. This approach has demonstrated impressive outcomes.

iv. We chose accuracy as an assessment metric to evaluate our model's performance and obtained the highest result of 0.99% for LSTM-RNN with Attention Mechanism.

The study is structured into multiple parts. Section II presents an overview of prior research pertaining to approaches for identifying credit card fraud. Section III covers the methodology employed for utilizing deep learning classifiers to detect fraud. Section IV outlines the experimental outcomes and performance evaluation. Lastly, Sections V and VI provide an overview of the conclusion and future prospects of the research. Figure 2 depicts various categories of financial fraud, among which credit card fraud is identified as one of the most prevalent forms.

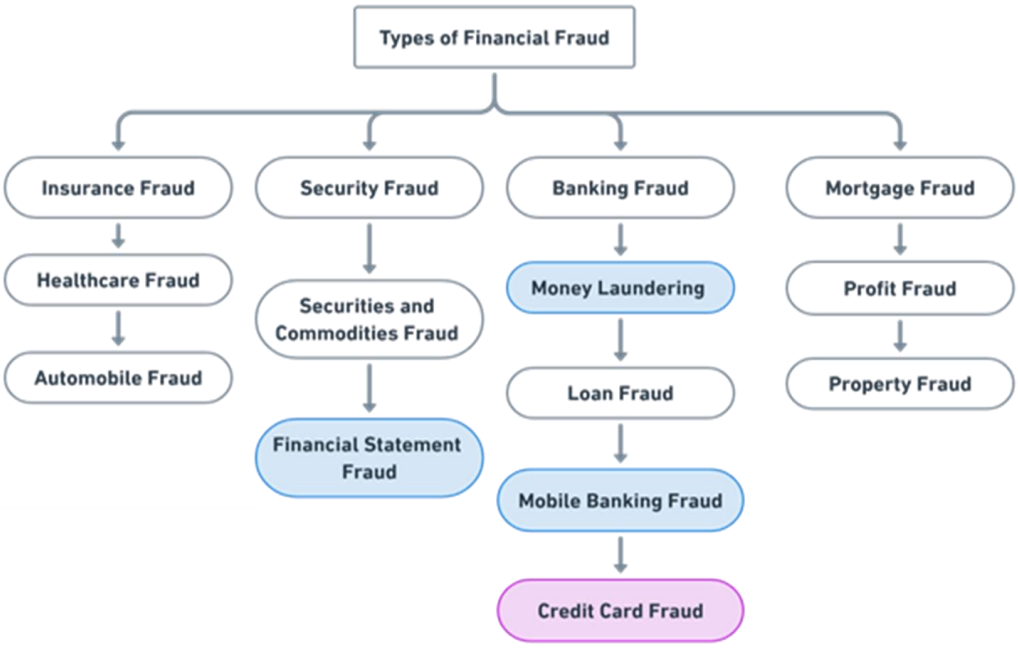

Fig. 2. Types of financial frauds

2. Related Work

Various deep learning and machine learning techniques, including supervised and unsupervised learning, deep neural networks, and ensemble learning, have been applied to detect credit card fraud [1]. Research conducted by G. K. Arun and P. Rajesh indicates that there is a higher possibility of credit card fraud incidents due to the rise in popularity of digital transactions and the demand for new payment methods. They propose a feature selection technique called BEPO-OGRU for credit card fraud detection, which uses binary emperor penguin optimization and optimal gated recurrent unit. This technique improves classification accuracy, achieving 94.78% and 94.16% on two different datasets [6]. Mahbuba *et al*., suggested a technique for detecting credit card fraud using AdaBoost, CNN, and CNN with GRU models, which addressed the imbalance dataset using SMOTE. Among the three models, CNN achieved the highest performance in AUC-ROC, accuracy, precision, and recall [7].

Gupta *et al.,* [8] stated that while plastic money is popular worldwide, new technology has its own vulnerabilities that can harm users financially. The wide adoption of credit cards has led to an increase in fraudulent activities. To detect credit card fraud, a model using Naive Bayes, Random Forest, Logistic Regression, and SVM algorithms was applied to a large dataset. Naive Bayes performed the best, achieving 80.4% accuracy and an AUC of 96.3%. In addition, Ajeet Singh, and Anurag Jain presented in their work that Credit cards are widely used online but this popularity has

also led to more fraud. This poses a challenge for security measures. The paper proposes using a method called HBRF to identify fraud in banking. This method is a combination of feature selection, a bio-inspired algorithm called firefly, and an ensemble classifier called random forest. HBRF overcomes the issue of imbalanced datasets and achieves high accuracy of 96.23% with a low error rate of 3.77% [9].

According to Makki *et al*., the existence of imbalanced datasets presents a major difficulty in identifying fraudulent credit card transactions. To address the imbalance in the dataset, they employed different machine learning algorithms and achieved the highest accuracy, sensitivity, and AUCPR scores with C5.0, LR, (DT), SVM, and ANN [10]. Moreover, Taha *et al.,* have proposed an intelligent approach for detecting fraudulent credit card transactions using an optimized light gradient boosting machine (OLightGBM). They have integrated a Bayesian-based hyperparameter optimization algorithm to adjust the LightGBM parameters intelligently. The method that was suggested outperformed other methods with regards to accuracy, area under the receiver operating characteristic curve (AUC), precision, and F1-score. Specifically, it achieved a 98.40% accuracy, a 92.88% AUC, a 97.34% precision, and a 56.95% F1-score, which were higher than those of other methods [11].

Ileberi *et al.,* developed a machine learning framework to detect credit card fraud using imbalanced datasets from European cardholders. We used SMOTE to address class imbalance and evaluated the effectiveness of several machine learning algorithms, including SVM, LR, RF, XGBoost, DT, and ET, by combining them with AdaBoost to improve their classification performance. The evaluation metrics used were accuracy, recall, precision, MCC, and AUC [12]. Credit cards have transformed into a means of electronically transferring funds as digital payment systems advance to facilitate business transactions worldwide. However, credit card fraud remains a significant global threat to financial institutions. Muhal *et al.,* [13] suggested three different ensemble models for applying a champion-challenger approach and determined the most effective model through evaluation. The model's effectiveness was evaluated by calculating its Accuracy, Precision, Recall, and F1-Score, which were found to be 98.86, 99.73, 99.99, and 99.86, respectively.

In their research, Lebichot *et al.,* [14] noted that financial institutions process thousands of credit and debit card transactions every second, making fraud detection a complex task due to the large volume of data and its sequential nature. They also highlighted the challenge of concept drift in fraud detection. As a result, they suggested that assessing incremental learning methods would be useful. The study developed and tested incremental learning approaches for real-world fraud detection systems. Esenogh *et al.,* expressed in their work that the rise of electronic commerce and credit card use has increased and is challenging to detect due to imbalanced datasets. The article puts forward an effective strategy for identifying fraudulent activities through the implementation of resampling and a neural network ensemble. The LSTM ensemble approach presented in the paper outperformed other algorithms, with a sensitivity of 0.996% and specificity of 0.998%. [15].

Darwish *et al.,* [16], has expressed concerns about the growing utilization of internet credit cards in e-banking systems due to their vulnerability to credit card fraud. Additionally, the problem of data imbalance makes it challenging to detect fraud. The paper proposes a method to address the difficulties associated with identifying credit card fraud. This approach comprises of a two-tier model that combines the k-means clustering algorithm with the artificial bee colony algorithm. The proposed model addresses the issue of imbalanced datasets and utilizes semantic fusion to enhance accuracy. According to the experiments conducted, the model performs better than traditional techniques in identifying fraudulent transactions. Forough [17] developed an ensemble model that uses deep recurrent neural networks to sequentially analyse data and a unique artificial neural network-based voting mechanism to detect fraudulent actions. The performance of the proposed

model was evaluated using two real-world datasets, and the outcomes demonstrated that it outperforms existing models in all assessment metrics.

Naoufal Rtayli and Nourddine Enneya [18] developed a new method called CCFD to detect credit card fraudulent transactions. The proposed approach shows a significant ability to identify fraudulent activity compared to previous studies. The most favourable results in terms of efficiency and effectiveness were achieved by applying their modelling approach to multiple datasets. Online transactions are increasingly common, and Credit Cards are widely used for these transactions. Credit Card Fraud (CCF) is a growing problem for both customers and financial institutions, resulting in significant financial losses. To address this issue, a Convolutional Neural Network (CNN) model using the Adaptive Synthetic (ADASYN) sampling technique has been proposed for detecting fraudulent transactions from normal transactions. The model has achieved high accuracy 0.9982, compared to other studies [19]. In [20] examines the performance of ensemble classifiers for detecting credit card fraud (CCF) using regression and voting methods. The study compares the ensemble classifiers with effective single classifiers including RBF, NB, MLP, DT, KNN, and SVM. Three different datasets were evaluated in the study, which were pre-processed using the SMOTE technique.

3. Proposed Methodology for Credit Card Fraud Detection

This section outlines the implementation process of our proposed model, providing an architectural diagram for clear comprehension. Furthermore, it explains the data collection procedure and the various utilized applications such as the SMOTE technique, RNN, LSTM, and Attention Mechanism algorithms.

3.1 Architectural Diagram of Proposed Model

The proposed model has a well-defined architecture that consists of several steps shown in Figure 3. The first step involves pre-processing the data by resizing, normalizing, and selecting relevant features. To address the imbalanced nature of the data, the SMOTE technique will be employed. The data will then be split into two phases for training and testing. During the training phase, the input data will be subjected to processing using algorithms such as RNN, LSTM, and Attention Mechanism, with the accuracy and loss being calculated at the conclusion of each epoch. After the model has been trained, it will be assessed using testing data and performance metrics like accuracy, precision, recall, and F1-score will be utilized to distinguish between lawful and fraudulent transactions.

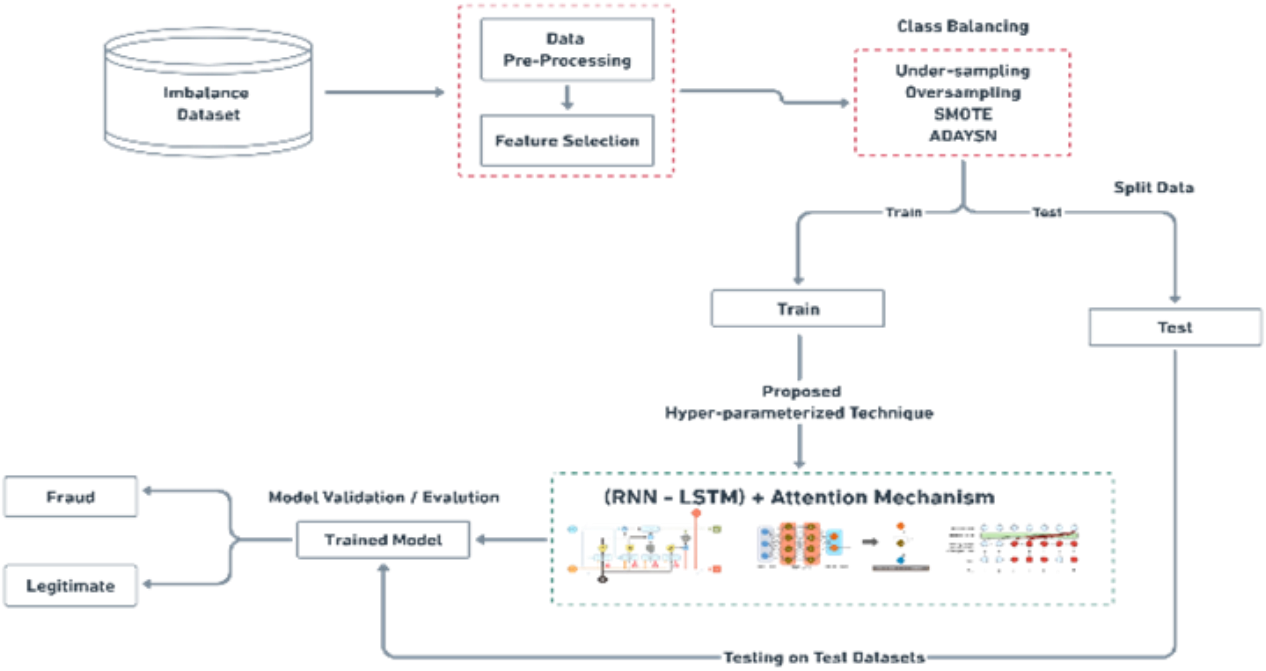

Fig. 3. Architecture of proposed model

3.2 Dataset Collection

The dataset used in the proposed model was obtained from the website www.kaggle.com, and it consists of credit card transactions performed by European cardholders in September 2013. The dataset spans a two-day duration and encompasses a total of 284,807 transactions, including 492 fraudulent transactions. The dataset is imbalanced, with only 0.172% of transactions being fraudulent. The dataset comprises of numerical input variables resulting from a Principal Component Analysis (PCA) transformation. It contains 31 columns, with 28 of them (V1, V2,..., V28) being PCA-derived, while 'Time', 'Class', and 'Amount' are not PCA-derived. The 'Class' column denotes fraud (1) or non-fraud (0).

The imbalance ratio of fraudulent transactions in the dataset is 0.172%, indicating a highly skewed distribution. The dataset includes 31 columns, out of which 28 (V1, V2,..., V28) are the result of Principal Component Analysis (PCA) transformation of numerical input variables. The remaining three columns, namely 'Time', 'Class', and 'Amount', are not derived from PCA. The 'Class' column represents the fraud status of the transaction, with 1 indicating fraudulent and 0 indicating non-fraudulent transactions. Figure 4 shows the nature of the dataset.

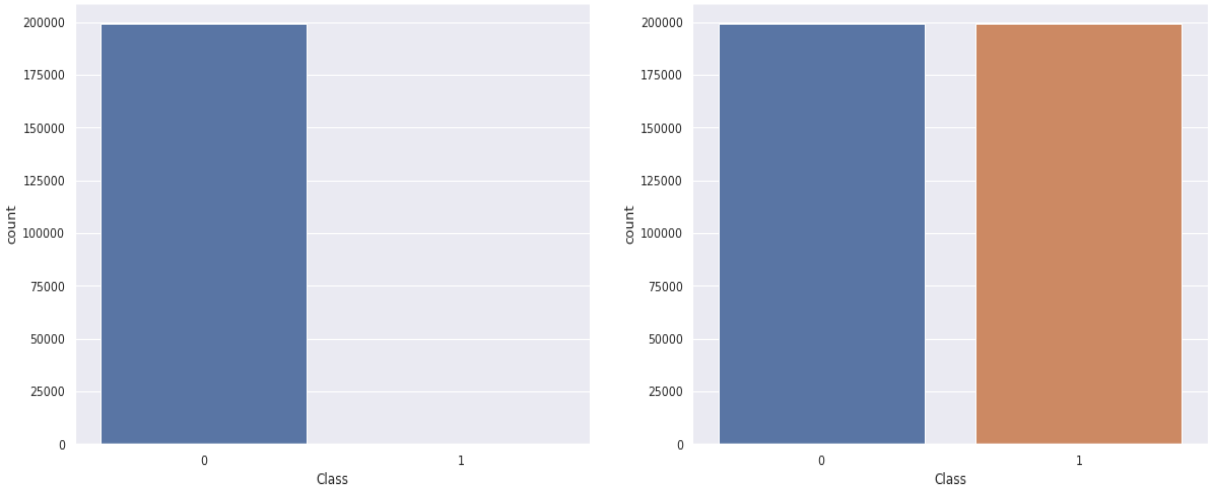

Fig. 4. Balanced and imbalance nature of dataset

Figure 5 illustrates the duration between the initial transaction and its corresponding transaction, which has undergone normalization to decrease processing time.

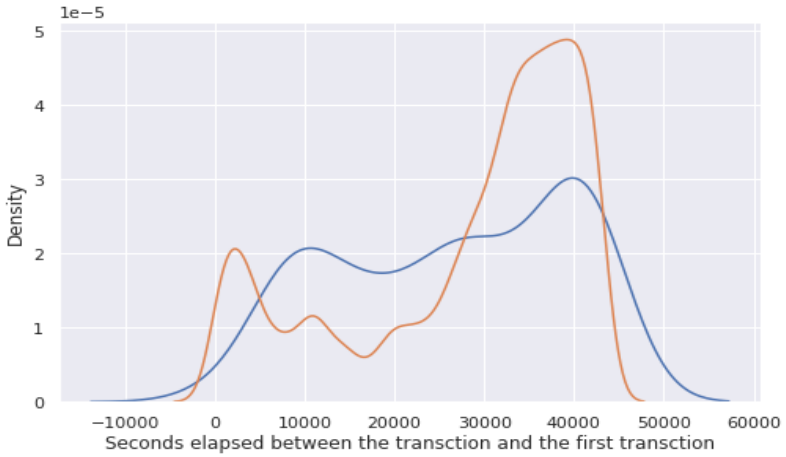

Fig. 5. Visualization of time elapsed between transactions and first transaction

Figure 6 displays a visualization of the transaction amount value, which has been adjusted to a normalized format in order to improve computational efficiency.

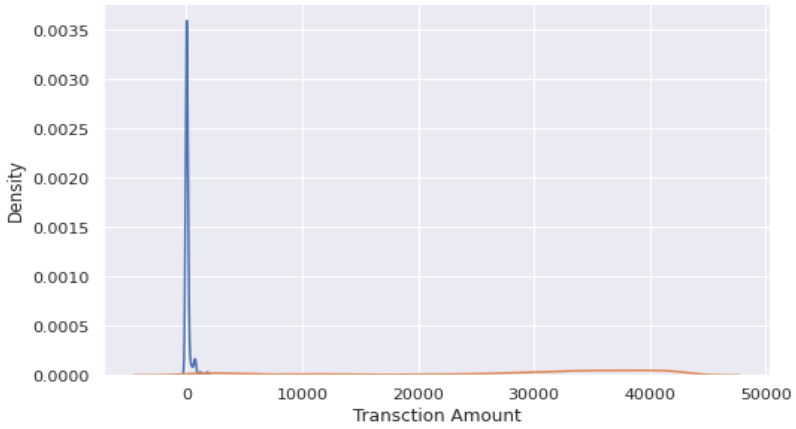

Fig. 6. Visualization of transaction amount

Figure 7 displays the visualization characteristics, which consist of features labelled as V1, V2, V3, up to V28.

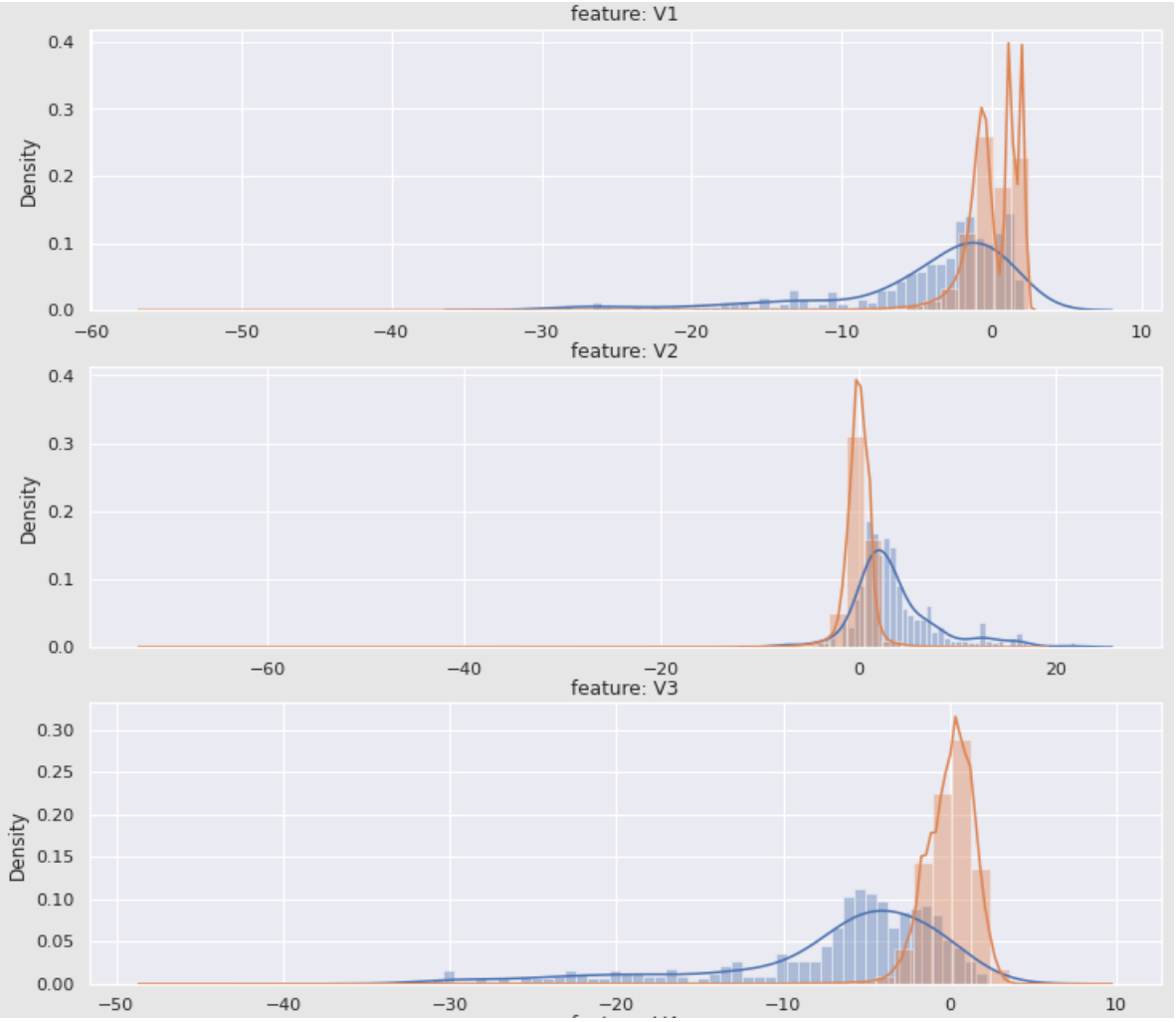

Fig. 7. Visualization of features V1, V2, and V3

3.3 SMOTE Technique

The term "SMOTE" is an acronym that stands for "Synthetic Minority Oversampling Technique". SMOTE is a commonly used approach for addressing the issue of imbalanced classes in datasets, especially for tasks such as developing models for detecting credit card fraud. [21]. It creates synthetic data points to address the scarcity of minority class instances, without replicating them directly. This approach generates new synthetic data points by utilizing the K-nearest neighbour algorithm to connect instances in the minority class, as demonstrated in Figure 8. This algorithm addresses the issue of overfitting that arises from random oversampling. It centres on the feature space and utilizes interpolation among closely located positive instances to produce new instances [22].

In this study, we employed the SMOTE technique to balance the number of legitimate and fraudulent credit card transactions by oversampling the minority class. We utilized SMOTE to augment the sample size of the minority class in order to achieve an equal number of instances in both classes.

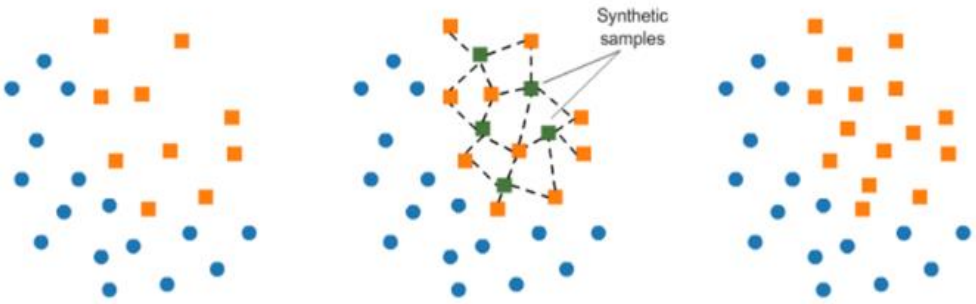

Fig. 8. Synthetic minority oversampling technique (SMOTE)

3.4 Recurrent Neural Network

Recurrent neural networks (RNNs) are a type of artificial neural network that are specifically designed to handle sequential data [23]. The primary characteristic that sets recurrent neural networks (RNNs) apart from other types of neural networks is their recurrent architecture, where the interconnections between nodes form a directed graph that facilitates the retention and retrieval of previous input sequences in a consecutive series [24]. RNNs have demonstrated remarkable flexibility in learning from experience, particularly when it comes to processing, categorizing, and predicting sequences based on their ability to retain information [24]. Artificial neural networks have limitations in modelling large sequential datasets. Recurrent neural networks solve this problem by creating connections between neurons within the same layer, resulting in cycles in the network's structure. These networks allow neurons to share weights across multiple time steps, enabling them to consider past information. The performance of Recurrent Neural Networks (RNNs) can be significantly affected by several essential factors, such as the activation function, dropout rate, and loss function [25]. Figure 9 demonstrates the architecture of RNNs.

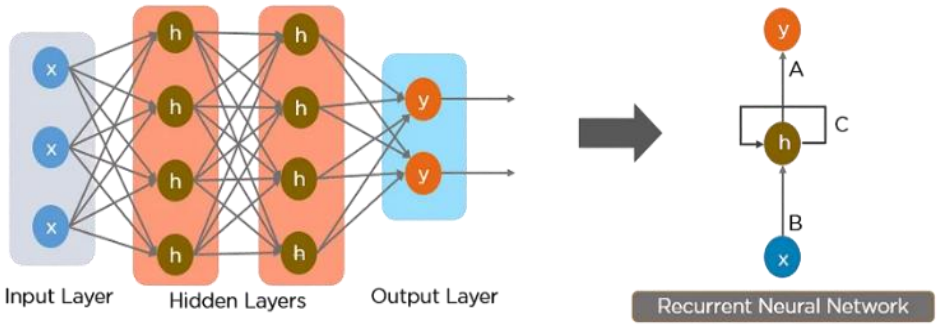

Fig. 9. Architecture of recurrent Neural Networks

3.5 Long Short-Term Memory

The Long Short-Term Memory (LSTM) technique represents an expansion of Recurrent Neural Networks (RNNs), which are deep learning neural networks that possess the capability to store memories [17]. The Long Short-Term Memory (LSTM) neural network is a suitable method for detection tasks due to its aptitude to learn from past experiences and its capability to manage long-term dependencies in sequential data. Its exceptional design permits it to retain information for a prolonged duration, thereby making it a useful tool for detecting and predicting long-term patterns. Therefore, LSTM is a useful tool for tasks that require prediction and detection capabilities [26]. The LSTM neural network is a type of recurrent neural network that is specifically designed to effectively

learn relationships and patterns over extended periods of time. It is capable of overcoming the problem of gradient disappearance that is commonly associated with traditional RNN [27,28]. The LSTM model consists of three gates that control the processing and utilization of historical data, along with a memory cell called $c_t$that stores the previously processed information [15]. These three gates are known as forget gate f_t, input gate i_t, and output gate o_t. Mathematical equations are used to update the LSTM layers.

$$i_t = \sigma\ (V_i x_t\ +\ W_i h_{t-1}\ +\ b_i) \tag{1}$$

$$f_t\ =\ \sigma\ (V_f\ x_t\ +\ W_f h_{t-1}\ +\ b_f) \tag{2}$$

$$c\sim_t\ =\ tan_h\ (V_c\ x_t\ +\ W_c\ h_{t-1}\ b_c) \tag{3}$$

$$c_t\ =\ f_t\ \otimes\ c_{t-1}\ +\ i_t\ \otimes\ \sim c_t \tag{4}$$

$$o_t\ =\ \sigma\ (V_o\ x_t\ W_o\ h_{t-1}\ +\ b_o) \tag{5}$$

$$h_t\ =\ o_t \otimes\ tan_h(c_t) \tag{6}$$

The symbol "*" can stand for different things depending on the context, such as f, i, o or c. These letters indicate specific gates or the memory cell. V_* and W_* are weight matrices, while h_*represents the hidden state, and b_* is the bias. At each time t step, h_t is the resulting output vector. The functions σ and tanh are activation functions, where σ represents the sigmoid function and $tanh$ represents the hyperbolic tangent function. Figure 10 presents the architecture of the LSTM network.

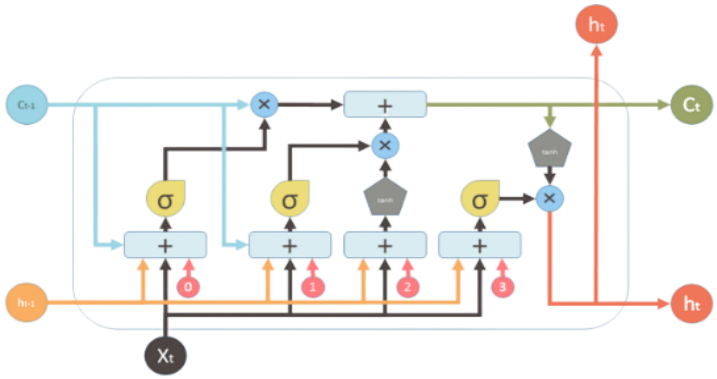

Fig. 10. Architecture of long short-term memory

3.5 Memory Attention Mechanism

The attention mechanism was initially proposed by Bahdanau *et al.,* [29] for the field of natural language processing (NLP) as a way to enhance encoder-decoder based neural machine translation systems. Since then, this mechanism or its variations have been utilized in other domains such as fraud detection, computer vision, and speech processing [30]. The attention mechanism enables neural networks to selectively concentrate on crucial data elements for the classification task by computing a weighted average of the local information contained in each sequence term. This results in better detection performance.

In this article, we employ sequence models based on RNN-LSTM and attention models to identify connections over time between events that may be distant from each other in the input sequence. This leads to improved performance in the classification task and enhances the ability to detect fraudulent transactions compared to conventional models. Figure 11 illustrate the architecture of attention mechanism.

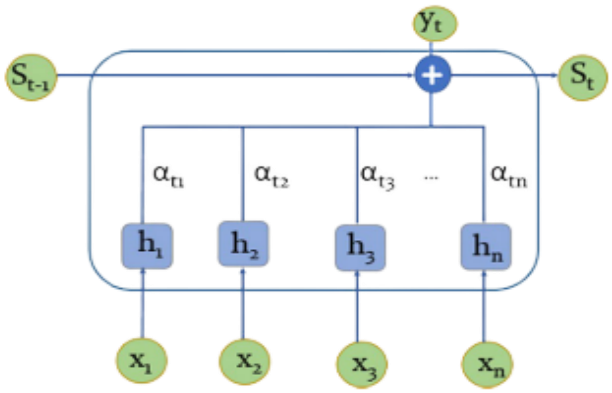

Fig. 11. Architecture of long short-term memory

4. Experimental Setup

The experiment involved training on 80% of the dataset and testing on the remaining 20%. The training data was prepared using the SMOTE method. Various parameters such as optimizers, activation functions, and modules like RNN and LSTM with attention mechanism were used to obtain the best results. The system employed σ and tanh as activation functions, Adam as optimizer, while batch size, epochs, and maximum number of epochs were used as parameters. The implementation was done using Python programming language, and Keras and TensorFlow libraries in Jupyter Notebook on a computer with an Intel(R) Core(TM) i5-2.80 GHz processor and 16 GB RAM. Python libraries like Numpy, Pandas, Scikitlearn, and Matplotlib were also used.

4.1 Performance Measurement Metrics

In this research, a dataset containing both legitimate and fraudulent credit card transactions have been used. The dataset is labelled with 1s for fraudulent transactions and 0s for legitimate transactions, and the problem is framed as a binary classification task. To evaluate the performance of the classification model, we use metrics such as accuracy, recall, precision, and AUC-ROC. These metrics are calculated using mathematical formulas. The most commonly used metric for measuring performance, which is accuracy, is presented in Figure 12 depicts a visual representation of the performance measurement metrics that occur most frequently.

Accuracy can be defined as the frequency of correct estimations, which can be calculated using the following mathematical expression:

$$Accuracy = \frac{Count\ of\ accurate\ predictions.}{The\ total\ count\ of\ predictions\ performed.} \quad (7)$$

Precision, which is also referred to as positive predictive value, represents the model's ability to accurately predict positive values in relation to all the positive values predicted by the model. The concept of precision is defined as follows:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive'} \quad (8)$$

Recall is a metric that can be utilized to evaluate the model's ability to detect true positives. A high recall score indicates that the model has successfully identified many true positives, whereas a low recall value implies that the model has encountered a large number of false negatives. The term "recall" denotes the following concept:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative'} \quad (9)$$

The F1-score is the harmonic mean of precision and recall, providing a more accurate estimate of the incorrectly classified cases compared to the accuracy metric.

$$F1 - score\ = 2 * \frac{Recall*Precision}{(Recall+Precision)} \quad (10)$$

F1-score is necessary to achieve a balance between precision and recall. As we previously observed, True Negatives have a significant impact on accuracy. However, when dealing with an imbalanced class distribution with a substantial number of Actual Negatives, the F1-score might be a more appropriate evaluation metric for achieving balance between precision and recall.

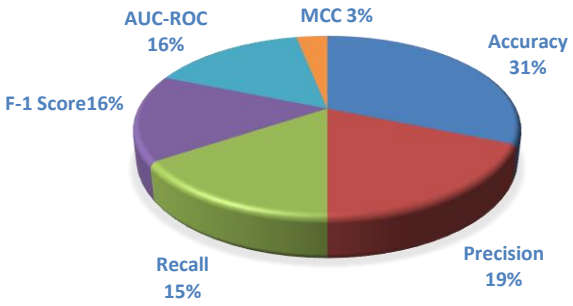

Fig. 12. Most occurring performance measurement metrics

The Receiver Operating Characteristic Curve (ROC) and Area Under the ROC Curve (AUC) are utilized to evaluate the performance of algorithms for classification tasks. The ROC is a probability curve, and the AUC represents the level of separability. The AUC reflects the model's ability to accurately distinguish between classes, indicating its efficacy in predicting the 0 and 1 classes. The ROC curve visualizes the correlation between True Positive Rate and False Positive Rate for different classification thresholds.

4.2 Results

The paper investigates the effectiveness of deep learning models, specifically RNN-LSTM with attention mechanism, in detecting financial fraud in transactions. The results show that the RNN-LSTM with attention mechanism model achieved the best performance with an accuracy of 99.94%.

Moreover, our study aimed to improve the accuracy and efficiency of fraud detection, which was successfully demonstrated. Our findings indicate that the RNN-LSTM model with attention

mechanism outperformed previously published results by a significant margin. Figure 13 presents the accuracy and loss of the RNN-LSTM model with attention mechanism technique.

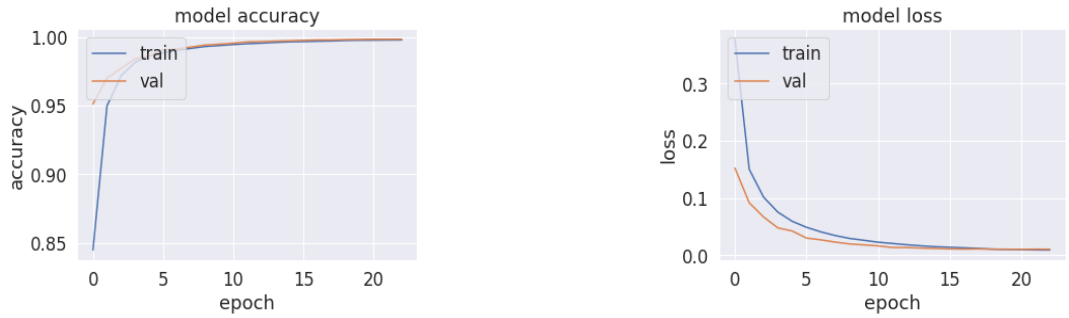

Fig. 13. Accuracy and loss of RNN-LSTM with attention mechanism model

This study includes a comparison of our results with the best previous outcome, revealing that the RNN-LSTM model with attention mechanism significantly outperformed prior published results. Figure 14 displays the ROC (Receiver operating characteristic) curve. Based on the experimental results, our proposed model outperformed various classification methods such as GRU, LSTM, SVM, XGBoost, Naïve Byes, Random Forest, Logistics Regression, RNN, CNN, and ANN, thus demonstrating its effectiveness in detecting credit card fraud.

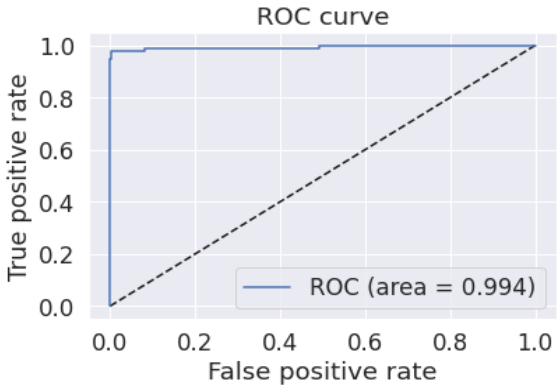

Fig. 14. ROC curve of RNN-LSTM with attention mechanism model

5. Discussion

This section discusses the methods and algorithms for detecting credit card fraud transactions, with a particular emphasis on the use of machine learning and deep learning techniques in the detection process. The issue of credit card fraud detection is typically addressed using both Deep Learning and Machine Learning approaches. The most commonly used technique for detecting credit card fraud transactions is SVM, which was utilized four times. Other methods employed include recurrent neural networks (RNN), artificial neural networks (ANN), long short-term memory (LSTM), decision trees (DT), random forests (RF), gradient boosting (LightGBM, XGBoost), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Logistic Regression (LR), Naive Bayes (NB), ensemble techniques (ET), deep reinforcement learning (DRL), and Convolutional Neural Networks (CNN). To enhance their performance and address the problem of imbalanced classes, these techniques have been combined with different methods of feature engineering and data resampling. Figure 15 presents the graphical presentation of credit card fraud detection techniques.

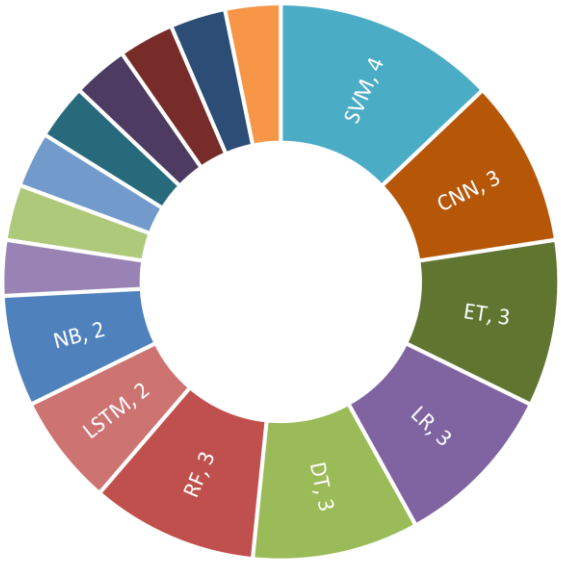

Fig. 15. Graphical presentation of credit card fraud detection techniques

After analysing the results, it was discovered that a SMOTE-based RNN-LSTM model with attention mechanism had significant success in detecting credit card fraud. In contrast to existing methods, this model demonstrated a higher accuracy in differentiating between valid and fraudulent transactions. Table III presents a comparison between the proposed model and other models based on various performance metrics such as accuracy. From the table, it can be observed that some existing models achieved lower accuracy ranging from 80.48% to 96.23%. However, a few models such as [13,19,26] and attained medium to high accuracy ranging from 98.86% to 99%. On the other hand, our proposed SMOTE-based RNN-LSTM model with attention mechanism achieved the highest accuracy of 99.4%. This indicates that our proposed model outperformed the existing models in terms of accuracy. In summary, our proposed model provides superior results compared to other current models, as shown in Table 1.

Table 1
Comparative analysis between the proposed and existing techniques

Study Ref.	Techniques	Accuracy (%)
[6]	BEPO-OGRU	94.78%
[8]	NB	80.4%
	RF	
	LR	
	SVM	
[9]	HBRF	96.23%
[11]	OLightGBM	98.40%
[26]	LSTM	98.95%
[13]	Ensemble Techniques	98.86%
[19]	CNN	99%
Proposed Mode RNN-LSTM with Attention Mechanism		99.4%

5. Conclusions and Future Recommendations

The objective of our research was to enhance the prediction efficiency for identifying fraudulent transactions, by integrating various Machine Learning techniques such as Swarm intelligence for feature selection, SMOTE for resolving imbalanced data, LSTM networks for modelling long-term dependencies within transaction sequences, and attention mechanism for identifying the most relevant data items for classification. Our proposed model is adept at identifying useful patterns in consumer behaviour, which facilitates the effective detection of fraudulent transactions. To validate our findings, we applied our model to a dataset of European credit card holders, and it demonstrated high accuracy in detecting fraudulent instances, which is crucial in this domain. Our model also outperforms recent works in this area.

In future research will explore a novel approach to credit card fraud detection using attention and transformers architecture as the sole means of processing sequences without relying on recurrent networks and improve the accuracy of fraud detection. Our model will also tackle the problem of concept drift and overlapping issues between different classes. We believe that this novel technique will provide a more efficient and effective means of detecting credit card fraud, benefiting both consumers and financial institutions. This research has the potential to make significant contributions to the field of fraud detection and could pave the way for future advancements in this area. Table 2. presents the full forms of the techniques used in this research paper.

Table 2
Comparative analysis between the proposed and existing techniques

Abbreviations	Full Form
DL	Deep Learning
ML	Machine
NB	Naïve Bayes
RF	Random Forest
DT	Decision Tree
LR	Logistic Regression
SVM	Support Vector Machine
OLightGBM	Optimized Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
MLP	Multilayer Perception
ANN	Artificial Neural Network
SMOTE	Synthetic Minority Oversampling Technique
KNN	K-Nearest Neighbour

Acknowledgement
This work was supported by the STIRF Research Grant project, Cost Centre: 015LA0-035, Universiti Teknologi PETRONAS.

References
[1] Alkhatib, Khalid I., Ahmad I. Al-Aiad, Mothanna H. Almahmoud, and Omar N. Elayan. "Credit card fraud detection based on deep neural network approach." In *2021 12th International Conference on Information and Communication Systems (ICICS)*, pp. 153-156. IEEE, 2021. https://doi.org/10.1109/ICICS52457.2021.9464555
[2] T. N. Report. "Card Fraud Worldwide – Nelson." vol. 1187, no. 1068, (2020).

[3] Sharma, Shagun, Anjali Kataria, Jasminder Kaur Sandhu, and K. R. Ramkumar. "Credit Card Fraud Detection using Machine and Deep Learning Techniques." In *2022 3rd International Conference for Emerging Technology (INCET)*, pp. 1-7. IEEE, 2022. https://doi.org/10.1109/INCET54531.2022.9824065

[4] Abd El Naby, Aya, Ezz El-Din Hemdan, and Ayman El-Sayed. "Deep learning approach for credit card fraud detection." In *2021 International Conference on Electronic Engineering (ICEEM)*, pp. 1-5. IEEE, 2021. https://doi.org/10.1109/ICEEM52022.2021.9480639

[5] Shamitha, S. Kotekani, and V. Ilango. "Importance of Self-Learning Algorithms for Fraud Detection Under Concept Drift." In *International Conference on Artificial Intelligence and Sustainable Engineering: Select Proceedings of AISE 2020, Volume 2*, pp. 343-354. Singapore: Springer Singapore, 2022. https://doi.org/10.1007/978-981-16-8546-0_28

[6] Arun, G. K., and P. Rajesh. "Design of Metaheuristic Feature Selection with Deep Learning Based Credit Card Fraud Detection Model." In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, pp. 191-197. IEEE, 2022. https://doi.org/10.1109/ICAIS53314.2022.9742937

[7] Turaba, Mahbuba Yesmin, Mehedi Hasan, Nazrul Islam Khan, and Hafiz Abdur Rahman. "Fraud Detection During Financial Transactions Using Machine Learning and Deep Learning Techniques." In *2022 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, pp. 1-8. IEEE, 2022. https://doi.org/10.1109/CCCI55352.2022.9926503

[8] Gupta, Amit, M. C. Lohani, and Mahesh Manchanda. "Financial fraud detection using naive bayes algorithm in highly imbalance data set." *Journal of Discrete Mathematical Sciences and Cryptography* 24, no. 5 (2021): 1559-1572. https://doi.org/10.1080/09720529.2021.1969733

[9] Singh, Ajeet, and Anurag Jain. "Hybrid bio-inspired model for fraud detection with correlation based feature selection." *Journal of Discrete Mathematical Sciences and Cryptography* 24, no. 5 (2021): 1365-1374. https://doi.org/10.1080/09720529.2021.1932929

[10] Makki, Sara, Zainab Assaghir, Yehia Taher, Rafiqul Haque, Mohand-Said Hacid, and Hassan Zeineddine. "An experimental study with imbalanced classification approaches for credit card fraud detection." *IEEE Access* 7 (2019): 93010-93022. https://doi.org/10.1109/ACCESS.2019.2927266

[11] Taha, Altyeb Altaher, and Sharaf Jameel Malebary. "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine." *IEEE Access* 8 (2020): 25579-25587. https://doi.org/10.1109/ACCESS.2020.2971354

[12] Ileberi, Emmanuel, Yanxia Sun, and Zenghui Wang. "Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost." *IEEE Access* 9 (2021): 165286-165294. https://doi.org/10.1109/ACCESS.2021.3134330

[13] Muhal, Harshit, Gaurav Khatri, Gaurav Kumar Dhama, and Deepika Bansal. "Ensemble Approach for Credit Card Fraud Detection Using Champion-Challenger Analysis." In *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, pp. 1-8. IEEE, 2022. https://doi.org/10.1109/ICSES55317.2022.9914387

[14] Lebichot, Bertrand, Gian Marco Paldino, Gianluca Bontempi, Wissam Siblini, Liyun He-Guelton, and Frédéric Oblé. "Incremental learning strategies for credit cards fraud detection." In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pp. 785-786. IEEE, 2020. https://doi.org/10.1109/DSAA49011.2020.00116

[15] Esenogho, Ebenezer, Ibomoiye Domor Mienye, Theo G. Swart, Kehinde Aruleba, and George Obaido. "A neural network ensemble with feature engineering for improved credit card fraud detection." *IEEE Access* 10 (2022): 16400-16407. https://doi.org/10.1109/ACCESS.2022.3148298

[16] Darwish, Saad M. "A bio-inspired credit card fraud detection model based on user behavior analysis suitable for business management in electronic banking." *Journal of Ambient Intelligence and Humanized Computing* 11, no. 11 (2020): 4873-4887. https://doi.org/10.1007/s12652-020-01759-9

[17] Forough, Javad, and Saeedeh Momtazi. "Ensemble of deep sequential models for credit card fraud detection." *Applied Soft Computing* 99 (2021): 106883. https://doi.org/10.1016/j.asoc.2020.106883

[18] Rtayli, Naoufal, and Nourddine Enneya. "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization." *Journal of Information Security and Applications* 55 (2020): 102596. https://doi.org/10.1016/j.jisa.2020.102596

[19] Gambo, Muhammad Liman, Anazida Zainal, and Mohamad Nizam Kassim. "A Convolutional Neural Network Model for Credit Card Fraud Detection." In *2022 International Conference on Data Science and Its Applications (ICoDSA)*, pp. 198-202. IEEE, 2022. https://doi.org/10.1109/ICoDSA55874.2022.9862930

[20] Kumari, Priyanka, and Smita Prava Mishra. "Analysis of credit card fraud detection using fusion classifiers." In *Computational Intelligence in Data Mining: Proceedings of the International Conference on CIDM 2017*, pp. 111-122. Springer Singapore, 2019. https://doi.org/10.1007/978-981-10-8055-5_11

[21] Elreedy, Dina, and Amir F. Atiya. "A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance." *Information Sciences* 505 (2019): 32-64. https://doi.org/10.1016/j.ins.2019.07.070

[22] Lemaître, Guillaume, Fernando Nogueira, and Christos K. Aridas. "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning." *The Journal of Machine Learning Research* 18, no. 1 (2017): 559-563.

[23] Roseline, J. Femila, G. B. S. R. Naidu, V. Samuthira Pandi, S. Alamelu alias Rajasree, and N. Mageswari. "Autonomous credit card fraud detection using machine learning approach☆." *Computers and Electrical Engineering* 102 (2022): 108132. https://doi.org/10.1016/j.compeleceng.2022.108132

[24] Zhang, Xinwei, Yaoci Han, Wei Xu, and Qili Wang. "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture." *Information Sciences* 557 (2021): 302-316. https://doi.org/10.1016/j.ins.2019.05.023

[25] Roy, Abhimanyu, Jingyi Sun, Robert Mahoney, Loreto Alonzi, Stephen Adams, and Peter Beling. "Deep learning detecting fraud in credit card transactions." In *2018 systems and information engineering design symposium (SIEDS)*, pp. 129-134. IEEE, 2018. https://doi.org/10.1109/SIEDS.2018.8374722

[26] Alghofaili, Yara, Albatul Albattah, and Murad A. Rassam. "A financial fraud detection model based on LSTM deep learning technique." *Journal of Applied Security Research* 15, no. 4 (2020): 498-516. https://doi.org/10.1080/19361610.2020.1815491

[27] Benchaji, Ibtissam, Samira Douzi, Bouabid El Ouahidi, and Jaafar Jaafari. "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model." *Journal of Big Data* 8 (2021): 1-21. https://doi.org/10.1186/s40537-021-00541-8

[28] Xie, Yu, Guanjun Liu, Chungang Yan, Changjun Jiang, MengChu Zhou, and Maozhen Li. "Learning transactional behavioral representations for credit card fraud detection." *IEEE Transactions on Neural Networks and Learning Systems* (2022). https://doi.org/10.1109/TNNLS.2022.3208967

[29] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." *arXiv preprint arXiv:1409.0473* (2014).

[30] Benchaji, Ibtissam, Samira Douzi, and Bouabid El Ouahidi. "Credit card fraud detection model based on LSTM recurrent neural networks." *Journal of Advances in Information Technology* 12, no. 2 (2021). https://doi.org/10.12720/jait.12.2.113-118