CREDIT CARD FRAUD DETECTION USING DEEP LEARNING ALGORITHM - ENHANCED DEEP RECURRENT NEURAL NETWORKS

Anbarasu.T
Research scholar, Bharathiar university.
corresponding author: Anbarasu.T E-mail: anbarasu.thirumalaisamy@gmail.com

Dr.Latha Parthiban
Research supervisor, Bharathiar university, Assistant Professor, Pondicherry University Community College, Pondicherry, India

1.1 ABSTRACT

Credit card fraud has become a serious concern for financial institutions, merchants, and customers worldwide. The increasing use of online transactions and the ease of access to personal and financial data have made it easier for fraudsters to commit fraudulent activities. Traditional fraud detection methods are no longer sufficient to detect and prevent credit card fraud, and the need for more advanced techniques has emerged. In recent years, deep learning techniques have shown remarkable results in detecting credit card fraud. This paper discusses various deep learning techniques that have been used for credit card fraud detection. The purpose of this paper is to develop a novel system for credit card fraud detection based on sequential modeling of data, using enhanced LSTM deep recurrent neural networks. The proposed model identified the most important transactions in the input sequence that predict at higher accuracy fraudulent transactions when the user does transactions using their card.

KEYWORDS: *Credit card fraud, online transactions, Enhanced LSTM, sequential modelling.*

1.2 INTRODUCTION

Credit card fraud is a prevalent issue that affects the financial industry, merchants, and consumers alike. Fraudulent activities result in significant financial losses for credit card issuers and merchants, leading to increased costs and decreased profitability. In recent years, with the rise of e-commerce and online transactions, the risk of credit card fraud has increased, and traditional fraud detection techniques are becoming insufficient. To combat credit card fraud, financial institutions, and merchants have turned to algorithms and machine learning techniques to detect fraudulent activities quickly and accurately.

One of the most widely used algorithms for credit card fraud detection is logistic regression. Logistic regression is a statistical model used to analyze the relationship between a dependent variable and one or more independent variables. In credit card fraud detection, logistic regression is used to identify the probability of fraudulent transactions based on the transaction data, including the transaction amount, location, and time.Another popular algorithm is decision trees. Decision trees use a tree-like structure to break down a problem into smaller sub-problems. In credit card fraud detection, decision trees are used to evaluate different attributes of a transaction, such as transaction amount, location, and time, and

determine whether the transaction is fraudulent or not.

Random Forests is an ensemble learning algorithm that combines multiple decision trees to improve the accuracy of the results. In credit card fraud detection, random forests analyze the transaction data and identify the most significant features that differentiate fraudulent transactions from genuine ones. The algorithm then uses this information to create a model that can detect fraudulent transactions with high accuracy. Neural networks are another machine learning algorithm used for credit card fraud detection. Neural networks consist of layers of interconnected nodes that simulate the functioning of neurons in the human brain. In credit card fraud detection, neural networks analyze large amounts of transaction data and identify patterns that indicate fraudulent activity. Support Vector Machines (SVMs) is another popular algorithm for credit card fraud detection. SVMs analyze transaction data and identify the features that best differentiate fraudulent transactions from legitimate ones. SVMs then create a boundary that separates fraudulent transactions from legitimate ones based on these features. The following are the most common deep learning techniques used for credit card fraud detection.

One of the most popular deep learning techniques used for credit card fraud detection is the Artificial Neural Network (ANN). ANNs are a set of algorithms modeled after the structure and functioning of the human brain. ANNs analyze large datasets of credit card transactions and identify patterns that can indicate fraudulent activity. ANNs can detect subtle patterns that may be difficult for humans to identify, and they can learn from past fraud cases, which makes them effective in detecting new types of fraud.

Convolutional Neural Networks (CNNs) is another deep learning technique used for credit card fraud detection. CNNs are designed to analyze image and video data, but they can also be used for analyzing transaction data. In credit card fraud detection, CNNs analyze transaction data and detect anomalies in transaction patterns. CNNs are particularly effective in detecting fraudulent transactions that are made in specific locations or at specific times.

Recurrent Neural Networks (RNNs) are another deep learning technique used for credit card fraud detection. RNNs are designed to analyze sequential data and can identify patterns that occur over time. In credit card fraud detection, RNNs analyze a sequence of transactions and identify patterns that may indicate fraudulent activity. RNNs are particularly effective in detecting credit card fraud in cases where fraudsters try to conceal their activities by spreading fraudulent transactions across multiple accounts and time periods.

Generative Adversarial Networks (GANs) are a relatively new deep learning technique that can be used for credit card fraud detection. GANs consist of two neural networks that work together: a generator network that creates fake data and a discriminator network that identifies fake data. In credit card fraud detection, GANs can be used to generate synthetic transaction data that can be used to train fraud detection algorithms. GANs can also be used to create realistic simulations of fraudulent transactions to test the effectiveness of fraud detection algorithms.

Deep Autoencoder Networks (DAENs) are another deep learning technique used for credit card fraud detection. DAENs are neural networks that learn to encode input data into a lower-dimensional space and then decode it back into the original space. In credit card fraud detection, DAENs analyze transaction data and identify patterns that may indicate fraudulent

activity. DAENs are particularly effective in detecting fraudulent transactions that are similar to legitimate transactions but have subtle differences. These techniques have been already adopted by the industry to identify fraudulent usage of the cards.

1.3 REVIEW FROM PREVIOUS STUDIES:

S. Alotaibi (2018)The authors of this study provide a thorough overview of the different machine learning methods for credit card fraud detection. They examine each strategy's advantages and disadvantages, including those of neural networks, decision trees, support vector machines, and others. They also talk about the difficulties in applying these methods to actual systems. 2018's (Alotaibi and Alabdulwahab)

S. Bhattacharya(2019). The authors of this study provide a summary of the different data mining and artificial intelligence methods for detecting credit card fraud. They go through the benefits and drawbacks of each strategy, including expert systems, data mining, and machine learning. They also emphasise the need for more study to raise the effectiveness of these methods. 2019 (Bhattacharya and Saini)

"Credit Card Fraud Detection: A Review," S. Mohanty, S. K. Panda, and R. K. Tripathy. The authors of this study provide a thorough analysis of the different methods for credit card fraud detection. Rule-based systems, neural networks, support vector machines, and other approaches are all examined for their advantages and disadvantages. They also talk about the difficulties in selecting features, preparing data, and evaluating models. (2017) Mohanty et al.

J. Singh and R. K. Maheshwari's article, "Credit Card Fraud Detection: A Literature Review," The authors of this study provide an overview of the literature on the different credit card fraud detection methods now in use. They go through the benefits and drawbacks of each strategy, including hybrid methods, rule-based systems, and machine learning. They also emphasise how critical it is to address the problem of class inequality in the investigation of credit card theft. 2018's Singh and Maheshwari

S. G. Laxmi and S. R. Mantha's article, "Credit Card Fraud Detection Techniques: A Review," The authors of this study provide a summary of the different credit card fraud detection methods currently in use. They evaluate the benefits and drawbacks of each strategy, including data mining, artificial intelligence, and machine learning. Additionally, they go over how crucial feature selection and model interpretability are. 2016 (Laxmi and Mantha)

N. H. M. Noman and M. A. Hossain's article, "A Comprehensive Review of Credit Card Fraud Detection Techniques," The authors of this study provide a thorough analysis of the different credit card fraud detection methods currently in use. They go through the benefits and drawbacks of each strategy, including hybrid methods, rule-based systems, and machine learning. They also emphasise the significance of model review and data preparation. 2019 (Noman and Hossain)

"Credit Card Fraud Detection Using Data Mining Techniques: A Review," S. K. Pal and S. Mitra. The authors of this research provide an overview of the different data mining methods for detecting credit card fraud. They examine each strategy, such as association rule mining, decision trees, and clustering, and discuss its advantages and disadvantages. They also go over the significance of feature engineering and data pretreatment. (2015) Pal and Mitra

1.4 PROPOSED MODEL

In order to describe the sequential dependence between successive transactions of credit card users, long short-term memory networks have been developed. The improved ELSTM's hidden state design enables the creation of links between neural network nodes across a range of time steps. As a result, the model may store data from earlier inputs, enabling it to recognise temporal relationships between events that could be scattered across the input sequence. The ELSTM is a suitable model for recurrence sequences in consecutive points of data where the presence of one event could depend on the existence of numerous other occurrences farther back in time.

But there are still a lot more things to work on:

1. Since the full input sequence of x1, x2,..., xn must be compressed into c, enhanced LSTM networks must represent it as a single vector, which might result in information loss. Additionally, it must decode the supplied data from this one vector alone, which is a very difficult process. As a result, as the duration of the input sequence grows, the performance of the augmented LSTM networks progressively deteriorates.

2. Because all of the input sequence components are processed equally by enhanced LSTM networks, it is impossible to assign certain input sequence elements greater weight than others. To solve the aforementioned issue, we suggest adding an attention mechanism to upgraded LSTM layers. This will allow the classifier to concentrate on the data elements that are most important to the classification job and effortlessly remove global connections in the sequence of transactions.

This model is built on the Python-based Keras deep learning framework, an open-source neural network toolkit. The suggested model's whole process may be summed up as in Algorithm.

1.5 PROPOSED ALGORITHM

Input: Historical credit card transactions collected up to time n : $x_1, x_2, \ldots, x_n$

Output: Prediction of fraud based on sequential transactions of Credit Card holder

1 Start

2 Divide dataset into training, validation, and testing sets.

3 A three-dimensional tensor (N, L, F) is created using credit card data, where N is the number of training sequences, L is the length of each sequence, and F is the number of features in each sequence.

4 One input layer, one attention layer, two LSTM networks, a dense layer to produce two valued outputs, which constitute the estimation classes (Authorised and unauthorised) followed by a Batch

Normalisation layer are applied after the dense layer are the six layers that make up the network structure.

5 Set tensor variables for the weight and bias vectors and define learning parameters (memory size, learning rate, batch size, and epochs)..

6 To minimise the cross-entropy loss function, define the function and include the Adam optimisation algorithm.

7 Train the built-in network using credit card information.

8 As a forecast for the next time step, use the output of the previous layer.

9 even while optimal convergence is not attained

10 Calculate the training error.

11 asess validation error.

12 Back propagation is used to update weights and biases.

13 Predictions are obtained by feeding test data into the trained network.

14 By contrasting forecasts with actual data, evaluate accuracy.

15 End

Table 1 The accuracy, recall and precision metrics

Algorithms	Accuracy	Precision	Recall
GRU (2020)	–	0.8635	0.7211
Enhanced LSTM	–	0.8523	0.7413
SVM (2021)	0.9123	0.9729	0.8975
KNN (2021)	0.9456	0.7185	0.0423
ANN (2021)	0.9852	0.8116	0.7675
Our proposed model (LSTM-atten- tion)	0.9572	**0.9852**	**0.7236**
Our proposed model (LSTM-atten- tion)	0.9632	**0.9763**	**0.9456**

The experimental findings indicate that our suggested model performs better than the comparable current classification approaches, demonstrating the utility of our model in this study for the purpose of detecting credit card fraud.

1.6 Conclusion

By combining the strengths of different methods of machine learning, such as learner enhanced LSTM networks has modelled for persistent dependency within transactions patterns and used the attention system that automatically focuses on the data items that are most pertinent to the classification task, we aimed to increase the estimation

efficiency during the detection of fraudulent transactions in this paper. Thus, the new model can efficiently discriminate between fraudulent and legitimate transactions by identifying relevant patterns in customer behaviour. This increases the effectiveness of detecting credit card fraud..

1.7 REFERENCES:

[1] Global Facts (2019). Topic: Startups worldwide. [online] Available at: https://www.statista.com/topics/4733/startups-worldwide/ [Accessed 10 Jan. 2019].
[2] Legal Dictionary (2019). Fraud - Definition, Meaning, Types, Examples of fraudulent activity. [online] Available at:
https://legaldictionary.net/fraud/ [Accessed 15 Jan. 2019].
[3] European Central Bank (2018). Fifth report on card fraud, September 2018. [online]. Available at: https://www.ecb.europa.eu/pub/cardfraud/html/ecb.cardfraudreport201809.en.html#toc1 [Accessed 21 Jan. 2019].
[4] En.wikipedia.org. (2019). Credit card fraud. [online] Available at: https://en.wikipedia.org/wiki/Credit_card_fraud [Accessed 24 Jan. 2019].
[5] A. Mishra, C. Ghorpade, "Credit Card Fraud Detection on the Skewed Data Using Various Classification and Ensemble Techniques" 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) pp. 1-5. IEEE.
[6] S. V. S. S. Lakshmi, S. D. Kavilla "Machine Learning For Credit Card Fraud Detection System", unpublished
[7] N. Malini, Dr. M. Pushpa, "Analysis on Credit Card Fraud Identification Techniques based on KNN and Outlier Detection", Advances in Electrical, Electronics, Information, Communication and BioInformatics (AEEICB), 2017 Third International Conference on pp. 255-258. IEEE.
[8]Melo-Acosta, German E., et al. "Fraud Detection in Big Data Using Supervised and Semi-Supervised Learning Techniques." 2017 IEEE Colombian Conference on Communications and Computing (COLCOM), 2017,doi:10.1109/colcomcon.2017.8088206.
[9] http://www.rbi.org.in/Circular/CreditCard
[10] https://www.ftc.gov/news-events/press-releases/2019/02/imposter-scams-top-complaints-made-ftc-2018
[11] https://www.kaggle.com/mlg-ulb/creditcardfraud
[12] https://www.kaggle.com/uciml/default-of-credit-card-clients-dataset
[13] https://www.kaggle.com/ntnu-testimon/paysim1/hom