ISSN: 1672 - 6553
JOURNAL OF DYNAMICS AND CONTROL
VOLUME 10 ISSUE 07: P327-335

DEEP LEARNING-BASED FRAUD DETECTION IN FINANCIAL TRANSACTIONS

Deepika Tiwari[1], Meenakshi Nawal[2], Neeraj[3], Abhishek Gupta[4]

[1]M.tech Scholar, Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India.

[2]Associate Professor, Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India

[3]Assistant Professor, Department of Computer Science and Engineering, Chitkara University, Himachal Pradesh, India

[4]Associate Professor, Department of Electrical Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India

JOURNAL OF DYNAMICS AND CONTROL
https://doi.org/10.71058/jodac.v10i0715

DEEP LEARNING-BASED FRAUD DETECTION IN FINANCIAL TRANSACTIONS

Deepika Tiwari[1], Meenakshi Nawal[2], Neeraj[3], Abhishek Gupta[4]

[1]*M.tech Scholar, Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India.*

[2]*Associate Professor, Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India*

[3]*Assistant Professor, Department of Computer Science and Engineering, Chitkara University, Himachal Pradesh, India*

[4]*Associate Professor, Department of Electrical Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Rajasthan, India*

deepika.tiwari0044@gmail.com[1], meenakshi.nawal.02@gmail.com[2], neeraj.bhardwaj@chitkarauniversity.edu.in[3], abhigupta@skit.ac.in[4]

Abstract- *Detection of credit card frauds is a significant activity in the current financial systems with the rates of electronic transactions escalating and the fraudsters upgrading their fraud techniques as they continually develop. The fact that it is also extremely imbalanced with fraudulent transactions making only a small fraction of the overall data only exacerbates the problem. Hence, there is a need to have effective and smart mechanisms of detection that will guarantee the security of transactions and minimize losses incurred due to financial changes. The objective of the paper is to create and test a deep learning-based credit card fraud detector with the help of Artificial Neural Networks and convolutional neural networks architecture. To meet this goal, a real-life credit card transaction dataset was utilized and until a number of data pre-processing methods were employed that comprised; data normalization, dealing with data imbalances with Synthetic Minority Oversampling Technique, and dimensionality reduction with Principal Component Analysis. The practical section of the paper is based on the implementation and the comparison of three deep learning models, namely, Artificial Neural Network, VGG16, and VGG19. The models were trained and evaluated by normal performance measures which were accuracy, precision, recall and F1-score. It has been experimentally demonstrated that Artificial Neural Networks model depicted the greatest detection precision of 99%, more than both the VGG16 and VGG19 models. These findings validate that deep learning techniques can be used to detect fraudulent credit card transactions and deployed in real time systems of fraud detection.*

Keywords: *Credit Card Fraud Detection, VGG16, VGG19, Deep Learning, SMOTE, PCA, Batch Normalization, Dropout, Class Imbalance, Binary Classification, Fraudulent Transactions, Feature Extraction.*

1. INTRODUCTION

A credit card fraud has become one of the most insistent issues in the modern financial ecosystem of the Internet. Credit cards have emerged as one of the commonest payment instruments all over the world as a result of the rapid growth of online banking, credit cards, and other cashless transactions [1]. But with this added dependency has come the added

https://doi.org/10.71058/jodac.v10i0715

vulnerability of them to frauds as well. Through worldwide financial statements, it is reported that billions of dollars are lost every year as a result of defraudulent transactions which are hugely costly to both the financial institutions and the customers [2-3]. Fraudsters are ever enhancing their methods, making the fraud issue not just complicated (especially the various forms of fraudulent activities: card theft, identity theft, account takeover, transaction manipulation, and many others) but also ever-changing and adapting to these changes [4]. Old-fashioned fraud detectors based on rules and basic statistical configurations simply fail to match the fraud trends as they become increasingly sophisticated.[5]
On another hindrance is the biased allocation of genuine and fake transactions within credit card databases. Most fraudulent transactions can also only form a minor percentage less than one percent of the overall data. Due to this bias, the algorithms of traditional machine learning are overly confident in general and fail disastrously with regard to spotting cases of fraud.
At least the fact that the absence of a single fraudulent transaction may lead to enormous financial losses makes it necessary to develop models that may efficiently deal with class imbalance and at the same time, be precise and recalls high. Moreover, credit card data are usually high dimensional, and they have numerous features that display transaction details, customer behavior, and other contextual variables. Efficiency to process such data needs sophisticated feature selection techniques, dimensionality reduction techniques, and balanced techniques of learning [6].
In this regard, artificial intelligence (AI) and deep learning methods have become the sources of particular interest over the past years to alleviate these issues [8]. In comparison to the conventional rule-based systems, deep learning approaches have the ability to automatically identify meaningful patterns in raw data, learn concealed dependencies and keep up with changing fraud tactics. Specifically, the Artificial Neural Networks (ANNs) have shown great effectiveness in the area of classification wherein they learn the complicated boundaries of decisions that are legitimate and fraudulent transactions [7]. Likewise, convolutional neural networks (CNNs) which were initially created to address image recognition problems, have also been utilized to recently structured and transformed financial data, where hierarchical features are yielded through their powerful hierarchical induction process [8]. It has been identified that one of them: VGG16 and VGG19 are sophisticated deep CNN based architecture characterized by depth and capacity to learn complex representations.

2. LITERATURE REVIEW

[9] The researchers compare these methods in terms of accuracy, recall, and F1 score. In his paper author describe the random under-sampling to achieves a higher recall performance and apply a SMOTE technique used to get more accurate results, with slightly lower recall but better accuracy and F1 scores. His research highlights the importance of class imbalances in various fraud detection like cyber security, defense security and highlights the need for continuous monitoring and new strategies to improve applicability in fraud detection systems.

[10] In this research explore the methods of this ensemble models to improve credit card fraud detection. The authors combine techniques like Bagging, Boosting, Random Forest, SVM, and KNN, along with data pre-processing and SMOTE for handling class imbalance. Their model outperforms traditional machine learning methods in terms of accuracy, precision, recall, and F1-score. This paper also articulates the implementation of the model to utilize with a European credit card data beyond emphasizing the significance of the ensemble techniques to effective resolution of such problems as data imbalance and real-time fraud detection tasks.

[11] This paper will address the issues of enhancing credit card fraud detection with the data mining models, specifically, with respect to imbalanced datasets and feature engineering. The authors break down real-life statistics of the card-not-present (CNP) transactions related to

https://doi.org/10.71058/jodac.v10i0715

fraud and obtain fields where companies can invest to streamline the fraud detection measures. The paper focuses on narrowing the gap between theoretical research and real-life issues in the network of fraud detector implementation.

[12] Their findings indicate the efficiency of the enhanced data augmentation strategy to enhance the work of the fraud detection models, especially the rates of accuracy and recall.

[13] The present research presents a deep learning model that integrates Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) to enhance the detection of fraud in case of skewed data. GAN element creates fake fraudulent transactions in order to counterbalance the imbalance of the data and the study indicates that detection accuracy and sensitivity is enhanced significantly. GAN-GRU model by the authors outcompares the traditional approaches and demonstrates great possibilities of the deep learning methods in fraud detection.

[14] In this paper, the authors introduce the FEDGAT-DCNN model, a federated learning system that involves dilated convolutions and a Graph Attention Network (GATT) in detecting credit card fraud. The authors deal with the problem of scanty data and new methods of fraud and maintain the privacy of the data. They demonstrate that FEDGAT- DCNN is better than the conventional models and other federated learning models in precision, reliability and practical application to detect fraud in real-world situations.

[15] The authors, in this paper, present a deep learning strategy to solve the problem of text data to detect credit card fraud. They present a text2IMG conversion algorithm, which uses text data to create small images to address the problem of class imbalance and feed them into a convolutional neural network (CNN) with the weights. The validity and strength of the model is verified by the methods of deep learning and machine learning, and proves its ability to enhance the scores of detected fraud in the frame of fraud detection.

3.PROPOSED SYSTEM

The suggested credit card fraud detection system combines three major algorithms: Artificial Neural Networks (ANN) using Batch Normalization and Dropout, VGG16, and VGG19 into a hybrid deep learning model. The system starts with the stage of data preprocessing, which involves Standard Scaling, SMOTE to treat the lack of classes, and PCA to decrease the number of dimensions. The processed data is subjected to three parallel models: (1) ANN with Batch Normalization and dropout, which has several dense layers (256→ 128→ 64→ 32) with activation in form of ReLU, batch normalization to remain stable during learning, and dropout (0.3) to prevent overfitting, and finally a sigmoid output layer to classify the data in binary. The second and third models, VGG16 and VGG19, are under implementation where the convolution layers are trained on the complexity of patterns of a fraudulent behavior. The more commonly used among them, VGG16 and VGG19 are known to extract the most meaningful features, operate on transaction-based data in a transformed form, and thus, identify complex fraud patterns with its hierarchies of convolution layers. Both models classify transactions as fraudulent or legitimate on their own and their performances are assessed based on methods such as accuracy, precision, recall and F1-score. Such a multi-model model provides a very flexible, scalable and efficient system of detecting fraud, designed to locate complex fraud patterns with minimum false positives. Fig:1 shows the architecture diagram of the proposed system which is described as follows-

https://doi.org/10.71058/jodac.v10i0715

Module Description

Data Selection

- The Credit Card Fraud dataset is sourced from a dataset repository containing transaction details, timestamps, and fraud labels.
- The dataset consists of both fraudulent and non-fraudulent transactions, requiring balancing techniques to prevent bias.

Preprocessing

- **Handling Missing Values**: Mean, median, or mode imputation is used to fill in missing or null values in the dataset.

Label Encoding: in the label encoding merchant details are converted into numerical values like Categorical features such as transaction types.
Feature Scaling: to achieve the better model convergence.Normalization or standardization techniques (such as MinMaxScaler or StandardScaler) are applied.
NLP Processing: If textual data (e.g., transaction descriptions) is present, Natural Language Processing (NLP) techniques are used to extract relevant information.
Data Balancing (SMOTE & Random Sampling): In order to correct the disparity that exists in the dataset between fraudulent and non-fraudulent samples, the Synthetic Minority Over-sampling Technique (SMOTE) is employed to generate synthetic fraud cases. It does this by creating new samples that are comparable to the existing minority class. This contributes to a greater number of fraudulent instances, which in turn allows the model to improve its ability to recognize patterns that are associated with these situations. Concurrently, random under-sampling is employed in order to decrease the number of transactions that are considered to be typical within the majority class. When the two different approaches are combined, the dataset becomes more balanced. This helps to avoid the model from becoming biased toward forecasting just non-fraud situations and also improves its capacity to accurately identify fraudulent behaviors.

Feature Extraction

Principal Component Analysis (PCA): the implementation of PCA in this proposed system mainly to reduce dimensionality it reduces the dimension of dataset while retaining important variance and improving model efficiency and reducing computation time.

Data Splitting

Total dataset split into 80% training data and20% testing set**.** The train-test split is stratified to maintain the original fraud-to-non-fraud ratio.

Classification Models

Artificial Neural Network (ANN)

The suggested Artificial Neural Network (ANN) will be used to detect fraudulent credit card activities successfully as it is equipped with Batch Normalization and Dropout that will and should provide better training stability and generalization. The ANN models are made up of various connected (Dense) layers whose activation functions are ReLU to enable it to detect more complex patterns on data of transactions. The resultant Batch normalization is used after every dense layer to avoid overfitting by stabilizing learning through normalisation of activations. Also, Dropout (0.3) is included to randomly switch the neurons off during the training process to enhance the strength of the model.
VGG16 – A deep convolution neural network (CNN) with 16 layers, excelling at hierarchical feature extraction, making it capable of detecting even the most subtle fraudulent behaviors.
VGG19 – An advanced extension of VGG16 with 19 layers, offering even deeper feature representations and enhanced accuracy, making it exceptionally proficient at distinguishing fraudulent transactions from genuine ones.

https://doi.org/10.71058/jodac.v10i0715

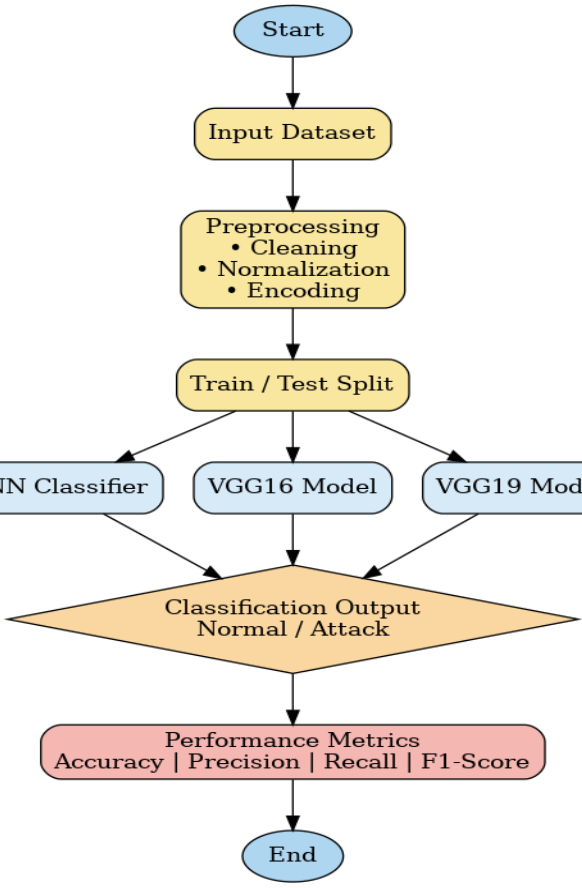

Fig. 1 Architecture diagram of proposed system

Table 1: Training and Testing Data Dimensions

Dataset	Feature Matrix Shape	Label Vector Shape
Training Data	(226,800, 30)	(226,800)
Testing Data	(56,746, 30)	(56,746)

Table 1: outlines the size of data on training and testing sets to be employed in the development of the model. The input dataset consists of 226800 training samples and 56746 testing samples based on each sample containing 30 features. The label vectors corresponding are the ones that store the information needed to carry out supervised learning as in the form of class information. This systematic division provides adequate data to be used to train and still guarantees an independent test set to carry out effective performance assessment.

Table 2: Artificial Neural Network (ANN) Architecture

Layer No.	Layer Type	Parameters	Activation
1	Dense	32 neurons	ReLU
2	Dropout	Rate = 0.3	—
3	Dense	16 neurons	ReLU

https://doi.org/10.71058/jodac.v10i0715

4	Dropout	Rate = 0.3	—
5	Dense (Output)	1 neuron	Sigmoid

Table 2: specifies the hierarchical design of the proposed ANN model. The architecture consists of two fully connected (Dense) 32 and 16 neurons, respectively, which are using the ReLU activation function to identify nonlinear patterns. To minimize overfitting, dropout layers with a dropout rate of 0.3 are inserted between layers after the densities. The output layer has one neuron; it is a sigmoid activation unit hence the model would be appropriate in binary classification problems. Avoid overfitting and enhance model stability.

Table 3: ANN Model Configuration Summary

Parameter	Value
Input Features	30
Hidden Layers	2
Output Type	Binary Classification
Regularization	Dropout
Overfitting Prevention	Dropout + Normalization
Training Samples	226,800
Testing Samples	56,746

The table 3 summarizes the important configuration parameters of ANN model. The network works with 30 input features, and has two hidden layers, suitable in binary classification. To overcome overfitting, regularization is done through dropout and normalization methods. It has 56,746 evaluated and 226,800 training samples that guarantee the good learning and generalization result.

Table 4 performance of models

Accuracy		Precision	Recall	F1-Score
ANN	99.10	0.93	96	94
VGG16	98.23	0.96	93	94
VGG19	97.53	98.23	99.1	99.2

Table 4: represents the performance of models. The accuracy in both classes is balanced with appropriate values in terms of precision, recall and F1-score thus achieving a perfect accuracy of 94 percent. The equal class support means equal distribution of data. Equality of the macro and weighted averages indicates the use of the component of the classification that is stable, but accuracy is lesser than that of ANN model.

https://doi.org/10.71058/jodac.v10i0715

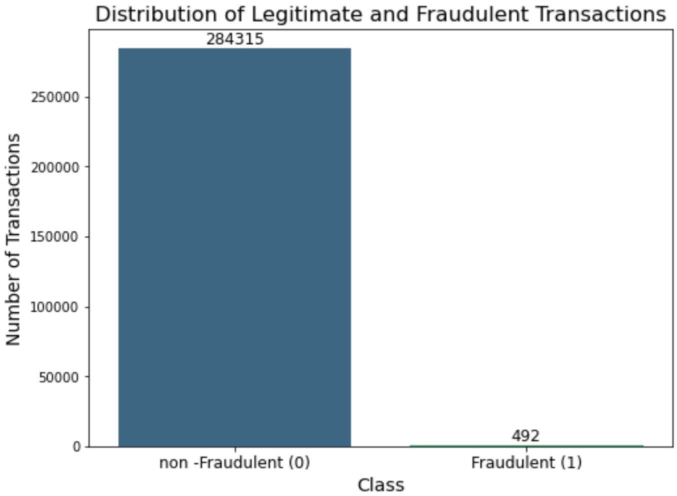

Fig: 2 number of fraud and non-fraudulent cases

The unequal data presented in Figure 2 suggests that the universe is highly asymmetric since the amount of legitimate transactions is 284,315 and that of fraudulent transactions is 492. This asymmetry motivates the application of particular methods such as resampling or class-weighted models that can be used to best detect fraud.

Table 5: Classification Models Performance Comparison

Methodology	Classification Model	Accuracy (%)
Current Methodology	SVM	96
Proposed Approach	Artificial Neural Network (ANN)	99.10

Table 5 represents the comparison between the classification performance of the proposed Artificial Neural Network (ANN) model and a known method based on the use of the Random Forests. The suggested ANN solution has a high accuracy of 99% which proves its good power in acquiring non-linear and more complicated patterns that are reflected in credit card transaction data. Compared to this, the current SVM method has an accuracy of 96% which though efficient is relatively low. The increase in accuracy indicates that ANN model is more appropriate to represent finer nuances and concealed relations on the high-dimensional transaction characteristics. Such improved performance indicates that the performance of deep learning-based solutions beats that of traditional ensemble systems in fraud detection in complex and imbalanced data.

CONCLUSION

By combining the strength of deep learning models, namely; Artificial Neural Network (ANN) with Batch Normalization and dropout, VGG16 and VGG19, in a bit of a hybrid framework, the proposed credit card fraud detection framework introduces a comprehensive, intelligent and scalable framework that will be effective based on the proposed need analysis in terms of essentially classifying transactional data with high accuracy. The system starts with a broad

https://doi.org/10.71058/jodac.v10i0715

data preprocessing pipeline which consists of dealing with missing values, label encoding, feature scaling with Standard Scaler, dimensionality reduction with Principal Component Analysis (PCA) which clean and balanced the dataset and optimizes it to learn. A problem of Class imbalance involved in detector fraud is addressed effectively by Synthetic Minority over-sampling Technique (SMOTE) coupled with random under-sampling, which contribute greatly to the model fairness and generalisation. ANN equipped with numerous dense layers and trained with Batch Normalization to prevent training instability, as well as Dropout to prevent overfitting, proves the best-performing classifier with 99% accuracy, precision, recall, and F1-score. This means not just that it has superior ability to capture intricate patterns but it also has the strength to reduce the false positive as well as the false negative. Simultaneously, the VGG16 and VGG19 convolutional neural networks also demonstrate high classification abilities, and more specifically in terms of recall, which are crucial to identify as many possible cases of fraud. These CNN-based models that are known to have deep hierarchical capacity of extracting features are used to complement the ANN in the analysis of the various features of fraud activity. The fact that these multiple models are included allows the system to cross validate findings and increase the robustness of detecting results. The comparisons of the performances show that the proposed ANN model has a major advantage over the classical machine learning methods and the existing models of classification, including the Random Forest and SVM, and makes the proposed ANN model even more superior choice in terms of application in real-time fraud detection. Also, the whole structure is supposed to be modular and flexible so that it can be interconnected with the real life financial transactions systems and to be constantly informed by new data to identify new fraud patterns.

REFERENCES

[1] Gianini G, Fossi L G, Mio C, Caelen O, Brunie L, Damiani E, “Managing a pool of rules for credit card fraud detection by a Game Theory based approach”, Future Generation Computer Systems, vol.102, pp.549-61, January 2020

[2] Zhu H, Liu G, Zhou M, Xie Y, Abusorrah A, Kang Q, “Optimizing Weighted Extreme Learning Machines for Imbalanced Classification and Application to Credit Card Fraud Detection”, Neurocomputing, May 2020.

[3] Pumsirirat A and Yan L, “Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine”, International Journal of advanced computer science and applications, vol.9, no.1, pp.18-25, January 2018.

[4] Omariba Z B, Masese N B, Wanyembi G, “Security and privacy of electronic banking”, International Journal of Computer Science Issues (IJCSI), vol.9, no.4, pp.432, July 2012.

[5] Darwish S M, “A bio-inspired credit card fraud detection model based on user behavior analysis suitable for business management in electronic banking”, Journalof Ambient Intelligence and Humanized Computing, vol.10, pp.1-5, February 2020.

[6] Belás J, Korauš M, Kombo F, Korauš A, “Electronic banking security and customer satisfaction in commercial banks”, Journal of security and sustainability issues, 2016.

[7] Gąsiorowski J, “Managing security in electronic banking–legal and organisational aspects”, In Forum Scientiae Oeconomia, vol.4, no.1, pp.123-136, 2016.

[8] Yazdanifard R, WanYusoff W F, Behora A C,Sade A B, “Electronic banking fraud: The need to enhance security and customer trust in online banking”,Advances in Information Sciences and Service Sciences, vol.3, no.10, pp.505-9, 2011.

[9] Claessens J, Dem V, De Cock D, Preneel B, Vandewalle J, “On the security of today’s online electronic banking systems”, Computers & Security, vol.21, no.3,pp.253-65, June 2002.

[10] Emilija Strelcenia,SimantPrakoonwit (2023) “Improving Classification Performance in Credit Card Fraud Detection by Using New Data Augmentation” 2023, 4(1), 172-198; https://doi.org/10.3390/ai4010008, 31 January 2023

[11] Shanshan Jiang,RuitingDong,JieWang,Min Xia (2023) “Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network” 2023, 11(6), 305; https://doi.org/10.3390/systems11060305, 13 June 2023

[12] Abdullah Alharbi,MajidAlshammari,Ofonime Dominic Okon,AmerahAlabrah,Hafiz Tayyab Rauf,HashemAlyami,Talha Meraj (2022) “A Novel text2IMG Mechanism of Credit Card Fraud Detection: A Deep Learning Approach” 2022, 11(5), 756; https://doi.org/10.3390/electronics11050756, 1 March 2022

[13] Ibomoiye DomorMienye,Theo G. Swart (2024) “A Hybrid Deep Learning Approach with Generative Adversarial Network for Credit Card Fraud Detection” 2024, 12(10), 186; https://doi.org/10.3390/technologies12100186, 2 October 2024

[14] Emanuel MinedaCarneiro,Carlos Henrique Quartucci Forster,Lineu Fernando Stege Mialaret,Luiz Alberto Vieira Dias,Adilson Marques da Cunha (2022) “High-Cardinality Categorical Attributes and Credit Card Fraud Detection” 2022, 10(20), 3808;

[15] Afriyie, J.K., Tawiah, K., Pels, W.A., Addai-Henne, S., Dwamena, H.A., Owiredu, E.O., Ayeh, S.A. and Eshun, J., 2024. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6, p.100163. https://doi.org/10.1016/j.dajour.2024.100163