A Convolutional Neural Network Model for Credit Card Fraud Detection

Muhammad Liman Gambo
School of Computing,
Faculty of Engineering
Universiti Teknologi Malaysia
Johor, Malaysia
gambo@graduate.utm.my

Anazida Zainal
School of Computing,
Faculty of Engineering
Universiti Teknologi Malaysia
Johor, Malaysia
anazida@utm.my

Mohamad Nizam Kassim
Department of Strategic Research
National Anti-Financial Crime Center
Putrajaya, Malaysia
mnizam.kassim@jpm.gov.my

2022 International Conference on Data Science and its Applications (ICoDSA) | 978-1-6654-8665-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICoDSA55874.2022.9862930

***Abstract*— Nowadays, online transactions through various ecommerce platforms are becoming more prevalent, and Credit Card (CC) is significantly used in various online transactions. However, Credit Card Fraud (CCF) strategies continue to evolve with the business transformation, causing customers as well as the financial institutions to lose billions of dollars annually. Hence, effective detection of fraudulent transactions initiated by fraudsters from the voluminous array of normal transactions is ever necessary. Hence, a Convolutional Neural Network (CNN) model for credit card fraud detection is proposed in this study using Adaptive Synthetic (ADASYN) sampling technique to address the imbalance dataset. The proposed model has achieved 0.9982, 0.9965, and 0.9999, accuracy, precision, and recall, respectively compared to other existing studies.**

Keywords—credit card fraud detection, convolutional neural network, imbalanced dataset, adaptive synthetic sampling technique, online transactions

I. Introduction

In the financial ecosystem, the quest to make life easier led to the continuous and rapid transformation of the world through digital technology [1]. People have moved from exchanging goods for goods in the barter system of trade to transacting using cash [2], and recently, through credit card (CC) and online payment services [3]. However, various kinds of frauds, for example, credit card fraud (CCF) have continued to engulf this transformation which results in loss of a huge amount of money [4]. A report by LexisNexis indicated that the cost of every single dollar lost to fraud by US retail and ecommerce merchants has increased from $3.13 to $3.60 from 2019 to 2021 [5].

Credit Card Fraud (CCF) as a form of financial fraud occurs when an individual other than a cardholder performs illegitimate transactions using a CC or card details [6]. These are called offline and online CCF, respectively. Because of the growth and proliferation of ecommerce platforms, the use of CC to complete transactions is increasing tremendously [7]. On other hand, with increasing sophistication due to advancements in technology [8], fraudsters see it as an opportunity to defraud customers of their hard-earned finances [9] by employing various means such as hacking, skimming, or the use of lost or stolen CC [10]. Therefore, detecting CCF is a crucial necessity that should be addressed through the implementation of automated and effective fraud detection systems to boost customer confidence and reduce financial losses [11], [12].

Financial institutions traditionally rely on the intuition, experience, and domain knowledge of some fraud experts to write rules for the detection of fraudulent transactions [13]. However, with the breakthrough in technology, a large volume of data generated through various transactions, and increasing fraudsters' intelligence, such a detection approach is difficult to build and maintain. As such, machine learning algorithms are used to develop models for effective detection of credit card fraud transactions.

Moreover, despite the considerable amount of work done by researchers in the development of CC fraud detection models using different machine learning algorithms such as Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), K Nearest Neighbor (KNN) etc., and deep learning algorithms for example Convolutional Neural Network (ConvNets), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM) etc., there is still need for more effective models with better predictive performance and accuracy to stand against the ravaging trend of CCF and successfully detect it. This can also be justified by the dynamicity of the fraud actors and their increasing sophistication, and the inherent imbalance class distribution of the CC dataset. In this study, a ConvNets model is developed, whose performance result is benchmarked against recently developed Credit Card Fraud Detection (CCFD) models developed by other researchers and has shown better predictive performance.

The remaining sections of the paper are organized as follows: Section II presents the related works. A brief background on the ConvNets algorithm and discussion on the detailed experimental procedure adopted in this study is given in Section III. Section IV presents the results and discussions, and finally, the conclusion of the paper as well as recommendation is given in Section V

978-1-6654-8665-1/22/$31.00 ©2022 IEEE

II. Related Work

Financial fraud has a long history and will continue to transform and intensify as technology advances [14]. Various studies have been proposed by researchers in the subject area, in particular, credit card fraud. [15] proposed a CCFD model using Convolutional Neural Network (CNN) as the classifier and Synthetic Minority Oversampling Technique (SMOTE) to balance the dataset. Using three evaluation metrics, the performance of the model was compared with that of a Multi-Layer Perceptron model which was trained using both balanced and unbalanced datasets. [16] proposed an anomaly detection model based on CNN and LSTM using the KDD99 dataset and European cardholder dataset. The study uses CNN to encode the time series data, and subsequently, LSTM to decode the extracted features from the sequence. Also, the study evaluated the developed model and compared the performance results with other models in terms of various performance metrics. [12] indicated the criticality of developing CCFD models which give a fair prediction of both the normal and fraudulent transactions. The study proposes three CCFD models based on Naïve Bayes (NB), Generative Adversarial Network (GAN), and Neural Network (NN) while using GAN to resample the minority class of the dataset. Results of the study indicated that the NN model demonstrated better performance than the other developed models. Similarly, [17] compare the predictive performance of four machine learning algorithms which include SVM, Naïve Bayes (NB), LR, and KNN. Without information on how the imbalance dataset was handled, the study shows that the model developed using the NB algorithm has the highest accuracy compared to the other three developed. Furthermore, [18] considers four approaches to handle the imbalance distributions of three different datasets. These include no sampling, Synthetic Minority Oversampling Technique (SMOTE), random undersampling, and Near Miss Sampling (NMS). The study implemented three deep learning algorithms which are: LSTM, Artificial Neural Network (ANN), and CNN (Conv1D and Conv2D), and two machine learning algorithms which are: SVM and RF algorithms for CCFD. The authors have shown that, while all the models demonstrated good results, the LSTM model produced better results compared to the other models considered. CCFD is an important research area that continues to receive a contribution from various researchers to create a secure space for online transactions. Also, while deep learning has demonstrated promising results in the detection of CC fraud and other applications, the intrinsic imbalance distributions of the CC dataset remain a critical issue in the development of effective models for CCFD. This study proposes an effective CCFD model using the CNN algorithm and ADASYN sampling technique and has shown better performance.

III. Experimental Procedure

This Section presents the experimental procedure followed in this study. It begins by giving a brief description of the deep learning algorithm and the dataset used in the study. Subsequently, how the dataset is preprocessed is described, before finally a discussion on the model development process.

A. Convolutional Neural Network

Convolutional Neural Network (CNN) is a commonly used deep learning algorithm that has provided good results in a wide range of applications [19]. ConvNets can uncover latent features of illegitimate transactions and avoid model overfitting [20]. ConvNets algorithm has three main layers which are: convolution layer, pooling layer, and fully connected layer. Generally, the function of the convolution and pooling layers is to perform feature extraction, while the third layer, a fully connected layer, performs the function of mapping the extracted features into the final output, for example, classification [21].

The three layers can be further described as follows:

Convolutional layer: The principal building block of ConvNets, and where the majority of computation (feature extraction) occurs which includes both linear and nonlinear processes, i.e., convolution operation and activation function.

Pooling Layer: This provides a down-sampling operation which decreases the in-plane dimensionality of the feature maps which introduce a translation invariance to small shifts and distortions, and reduce the size of successive learnable parameters.

Fully Connected Layer: The result of the feature maps from the final convolution or the pooling layer is changed into a one-dimensional (1D) array of numbers (or vector), and linked to one or more fully connected layers which are also called dense layers. At this layer, every input is connected to each output through a learnable weight. After the features extracted by the convolution layers and which the pooling layer down-sampled subsequently are generated, they are then mapped by fully connected layers to the final output of the network, for example, the probabilities for each class in the classification problem.

B. Dataset

The dataset used in this study is the European Cardholders' Dataset which is hosted in the Kaggle dataset repository. The European Cardholders Dataset which was collected and analyzed during collaborative research between Worldline and the Machine Learning Group of Universitas Libre de Bruxelles ´ (ULB) on big data mining and fraud detection consists of transactions made by European Credit Cardholders within the span of two days in September 2013. The dataset contains numerical values as a result of PCA (Principal Component Analysis) Transformation which was done to preserve the confidentiality of the cardholders, where 492 transactions are fraudulent out of 284,807 total transactions which translates to 0.172% of all transactions, and this indicates the high imbalance distribution of the dataset. The transactions are described by 30 features (V1, V2, . . ., V28, Time, and Amount) and then the class label which denotes a fraudulent transaction as ”1” and a normal transaction as ”0”

C. Data Preprocessing

Data used for training machine learning models have to be in a form suitable for application to the machine learning algorithms. Some preprocessing steps are usually done which include addressing null or missing values, encoding categorical attributes, data normalization and standardization, feature extraction, balancing the data, etc. European Cardholders' dataset contains all numerical values and no feature contains missing or null values. However, the imbalanced class distribution is highly critical and was addressed using Adaptive Synthetic (ADASYN) sampling technique in this study.

ADASYN adaptively generates different number of samples based on an estimate of the local distribution of the class to be oversampled; in this case, the minority class. Using default values for the parameters of the ADASYN sampling technique, the fraudulent class was oversampled to a size (284,298 samples) comparable to the normal class. Thereafter, the complete dataset which consists of a total of 568,613 data points was partitioned into three parts; 60% training data, 20% validation data, and 20% testing data as it was done by [22]. Using the train test split method, the dataset was first divided into random training (80%) and testing (20%) datasets, and subsequently, 25% of the training dataset was taken as the validation dataset which is equivalent to 20% of the total dataset. Table 1 illustrates the distribution of the dataset, before and after balancing. Also, standardization was performed using the StandardScaler method which is another preprocessing to address the high range of feature values of a dataset. StandardScaler works by making the mean of a distribution to be zero and scaling to unit variance.

TABLE I. DATASET DISTRIBUTION

Dataset	Legitimate samples	Fraud Samples	Total
Before Balancing	284,315	492	284,807
After Balancing	284,315	284,298	568,613

D. Model Development

The implementation of this study consists of two parts. It starts by training the classifier using the imbalance dataset which gives rise to Model 1 and then subsequently using the balanced dataset from which Model 2 is developed. In all the two parts, the dataset is split up into three parts used for training, validation, and testing the machine learning model as described in Section III C. Moreover, the ConvNets algorithm (Conv1D) is used as the classifier to develop the CCFD model based on the layer orientation and hyperparameter values presented in Table 2 and 3. Also, an effective model is a result of a comprehensive evaluation using standard metrics. Therefore, the two models developed in this study were evaluated using four performance metrics. These are confusion matrix, accuracy, precision, and recall. Additionally, the performance of the second model which was developed using the balanced dataset was compared with some recently proposed models from other studies.

TABLE II. CONVNETS MODEL LAYER ORIENTATION

Layer	Description
1	Conv1D Layer (filters = 32, kernel_size = 2, activation = relu)
2	BatchNormalization
3	MaxPool1D (pool_size = 2)
4	Dropout (rate = 0.2)
5	Conv1D Layer (filters = 64, kernel_size = 2, activation = relu)
6	BatchNormalization
7	MaxPool1D (pool_size = 2)
8	Dropout (rate = 0.5)
9	Flatten
10	Dense (units = 64, activation = relu)
11	Dropout (rate = 0.5)
12	Dense (units = 64, activation = relu)

TABLE III. CONVNETS MODEL LAYER ORIENTATION HYPERPARAMETER VALUES

Hyperparameter	Value
epochs	46
optimizer	Adam
learning_rate	0.0001
loss	binary_crossentropy
metrics	accuracy

IV. RESULTS AND DISCUSSIONS

This section presents the performance results obtained after training the two models developed in this study using the two approaches described. As a binary classification problem, Table 4 describes the results for the two predicted classes using the confusion matrix.

TABLE IV. CLASSIFICATION RESULT USING CONFUSION MATRIX

Model		Confusion Matrix			
		TN	FP	FN	TP
Model_1	Validation	56852	11	24	75
	Testing	56850	14	17	81
Model_2	Validation	56650	213	5	56855
	Testing	56661	202	4	56856

Also, Table 5 presents the performance of each of the two models developed in this study to further demonstrate the effect of addressing the imbalance distribution of the dataset using the ADASYN sampling technique.

From the results presented in Table V, it can be seen that the performance of Model 2 which was developed using the ADASYN resampled dataset is better than Model 1 which was developed using the imbalanced dataset. Although the accuracy scores of Model 1 are a bit higher than those of Model 2 in both the validation and testing phase, the scores for the other three metrics have increased substantially. In the validation phase, the precision score has increased from 0.8720 to 0.9963, while the recall score has increased from 0.7576 to 0.9999. Similarly, the precision score in the testing phase has increased from 0.8526 to 0.9965, and the recall score has increased from 0.8265 to 0.9999. In other words, the results indicate that the percentage of credit card transactions flagged as fraudulent that are actually fraudulent to the total number of predicted fraudulent transactions is 99.63% and 99.65% for the validation and testing phases. Likewise, the predictive ability of the model to correctly classify fraudulent transactions is 99.99% for both the validation and testing phases.

TABLE V. PERFORMANCE RESULTS

Model		Evaluation Results			
		Accuracy	Precision	Recall	F1 Score
Model_1	Validation	0.9993	0.8720	0.7576	0.8108
	Testing	0.9994	0.8526	0.8265	0.8393
Model_2	Validation	0.9981	0.9963	0.9999	0.9981
	Testing	0.9982	0.9965	0.9999	0.9982

Furthermore, Table 6 presents the performance comparison results of this study with other recently proposed studies.

TABLE VI. PERFORMANCE COMPARISON

Ref.	Technique	Performance Results			
		Accuracy	Precision	Recall	F1 Score
[15]	CNN + SMOTE	0.9890	0.9700	0.9100	0.9390
[16]	CNN + LSTM	0.9612	0.0382	0.9910	0.0742
[12]	NN + GAN	0.9987	0.9876	0.9871	0.9873
Proposed Model	CNN + ADASYN	0.9982	0.9965	0.9999	0.9982

From the results presented in Table 6, it can be seen that the model proposed in this study outperforms all the other studies in terms of accuracy score, precision score, and recall score.

V. CONCLUSION AND FUTURE WORK

Having a secure and enabling space for online transactions is an important requirement for e-commerce businesses. In this study, a Convolutional Neural Network Model for CCFD is successfully implemented by training the algorithm using training data, and then validating and testing the performance using the validation and testing datasets. The performance results show a significant improvement in the predictive performance of the CCFD model after the dataset is balanced using the ADASYN sampling technique, and hence, Model 2 is proposed. This is in addition to the results of the performance comparison performed with other existing studies using three evaluation parameters. For future study, the performance of the proposed model needs to be further improved to keep pace with the growing number of credit card fraud, and also explore other deep learning algorithms.

REFERENCES

[1] S. Vitaly, B. S. Rejwan, and P. Sant, “Review of Machine Learning Approach on Credit Card Fraud Detection,” *Human-Centric Intell. Syst.*, no. 0123456789, pp. 939–943, 2022.

[2] A. Sahu, H. Gm, and M. K. Gourisaria, “A Dual Approach for Credit Card Fraud Detection using Neural Network and Data Mining Techniques,” *2020 IEEE 17th India Counc. Int. Conf. INDICON 2020*, no. 1997, 2020.

[3] Y. Alghofaili, A. Albattah, and M. A. Rassam, “A Financial Fraud Detection Model Based on LSTM Deep Learning Technique,” *J. Appl. Secur. Res.*, vol. 15, no. 4, pp. 498–516, 2020.

[4] T. Jemima Jebaseeli, R. Venkatesan, and K. Ramalakshmi, “Fraud detection for credit card transactions using random forest algorithm,” *Adv. Intell. Syst. Comput.*, vol. 1167, no. January, pp. 189–197, 2021.

[5] LexisNexis, “Explosive Growth of Ecommerce and Retail Fraud,” 2021. [Online]. Available: https://risk.lexisnexis.com/insights-resources/research/us-ca-true-cost-of-fraud-study#ecommerce. [Accessed: 12-Apr-2022].

[6] B. Al Smadi and M. Min, “A Critical review of Credit Card Fraud Detection Techniques,” *2020 11th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2020*, pp. 0732–0736, 2020.

[7] G. M. Suhas Jain, N. Rakesh, K. Pranavi, and L. Bale, “A Novel Approach in Credit Card Fraud Detection System Using Machine Learning Techniques,” no. September 2013, pp. 1–5, 2022.

[8] N. Boutaher, A. Elomri, N. Abghour, K. Moussaid, and M. Rida, "A Review of Credit Card Fraud Detection Using Machine Learning Techniques," *Proc. 2020 5th Int. Conf. Cloud Comput. Artif. Intell. Technol. Appl. CloudTech 2020*, 2020.

[9] J. C. Mathew, D. B. Nithya, V. C. R, P. Shetty, P. H, and K. G, "An Analysis on Fraud Detection in Credit Card Transactions using Machine Learning Techniques," *Proc. Second Int. Conf. Artif. Intell. Smart Energy*, 2022.

[10] C. Shilpa and S. A.H., "A Comparative Analysis of Supervised Classifiers for Detecting Credit Card Frauds," *2022 Int. Conf. Comput. Commun. Informatics*, no. 978, 2022.

[11] A. Alharbi *et al.*, "A Novel text2IMG Mechanism of Credit Card Fraud Detection: A Deep Learning Approach," *Electron.*, vol. 11, no. 5, pp. 1–18, 2022.

[12] I. Ali, K. Aurangzeb, M. Awais, R. J. Ul Hussen Khan, and S. Aslam, "An Efficient Credit Card Fraud Detection System using Deep-learning based Approaches," *Proc. - 2020 23rd IEEE Int. Multi-Topic Conf. INMIC 2020*, 2020.

[13] B. Bart, H. Sebastiaan, and V. Tim, "Data Engineering for Fraud Detection," *Decis. Support Syst.*, vol. 150, no. July 2020, pp. 677–688, 2021.

[14] K. Shing Lim, L. Hong Lee, and Y.-W. Sim, "A Review of Machine Learning Algorithms for Fraud Detection in Credit Card Transaction," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 21, no. 9, 2021.

[15] A. A. El Naby, E. El-Din Hemdan, and A. El-Sayed, "Deep learning approach for credit card fraud detection," *ICEEM 2021 - 2nd IEEE Int. Conf. Electron. Eng.*, pp. 0–4, 2021.

[16] A. Zhang, X. Zhao, and L. Wang, "CNN and LSTM based Encoder-Decoder for Anomaly Detection in Multivariate Time Series," *IEEE Inf. Technol. Networking, Electron. Autom. Control Conf. ITNEC 2021*, pp. 571–575, 2021.

[17] O. Adepoju, J. Wosowei, S. Lawte, and H. Jaiman, "Comparative Evaluation of Credit Card Fraud Detection Using Machine Learning Techniques," *2019 Glob. Conf. Adv. Technol. GCAT 2019*, pp. 1–6, 2019.

[18] T. T. Nguyen, H. Tahir, M. Abdelrazek, and A. Babar, "Deep Learning Methods for Credit Card Fraud Detection," 2020.

[19] M. Krishna and D. Reshma, "Review On Fraud Detection Methods in Credit Card Transactions," *Int. Conf. Intell. Comput. Control*, 2017.

[20] K. Fu, D. Cheng, Y. Tu, and L. Z. B, "Credit Card Fraud Detection Using Convolutional Neural Networks," *Neural Inf. Process. - 23rd Int. Conf. ICONIP*, pp. 483–490, 2016.

[21] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4. pp. 611–629, 2018.

[22] I. Sohony, R. Pratap, and U. Nambiar, "Ensemble learning for credit card fraud detection," *ACM Int. Conf. Proceeding Ser.*, no. January 2018, pp. 289–294, 2018.