

Efficiency-Driven Deep Learning Models for Credit Card Fraud Detection: A Comparative Study and Optimization Approach

1st Sumit Haldar
Department of CSE
Brac University, Dhaka
sumit.haldar@g.bracu.ac.bd

2nd Maisha Shabnam Chowdhury
Department of CSE
Brac University, Dhaka
maisha.shabnam.chowdhury@g.bracu.ac.bd

3rd Nazifa Bushra
Department of CSE
Brac University, Dhaka
nazifa.bushra@g.bracu.ac.bd

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

Abstract—Since credit card fraud is a serious danger to both consumers and financial institutions, reliable and effective fraud detection systems must be developed. This paper offers detailed research on efficiency-driven deep learning models for credit card fraud detection, with the goal of exploring and optimizing their performance. To assess the efficacy of several deep learning architectures in identifying fraudulent transactions, we compare convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their hybrid equivalents. Our study uses a broad dataset that includes both legitimate and fraudulent credit card transactions, which makes it easier to evaluate the model's performance realistically. Our goal is to maximize the effectiveness of these models by resolving issues with uneven class distribution, interpretability, and computing capacity. We describe the findings of our comparative analysis and optimization strategy, which highlight the advantages and disadvantages of several deep learning models in situations involving credit card fraud detection. The significance of effective and precise fraud detection systems in the quickly changing field of financial security is emphasized by discussing the consequences of our findings for practical implementations.

Index Terms—credit card, fraud, transaction, optimization, security

I. INTRODUCTION

The evolution of digital transactions brought convenience to human beings worldwide and has revolutionized the financial landscape into a seamlessly connected global network. The central point of this digital network are the credit cards which is enabling the transactions at a surprisingly fast and efficient pace. But this widespread popularity of digital transactions has come with a cost, that is the escalation of credit card fraud. Traditional methods for recognising fraud encounter unparalleled hurdles in keeping up with the growing risks posed by these increasingly sophisticated fraudulent operations.

Credit card fraud is such a worldwide problem which affects all the countries of the world including the Republic of Korea. The Federal Trade Commission (FTC) reports that there were about 1579 data breaches involving 179 million data points, with fraudulent use of credit cards being the most common type of attack. The percentage of US credit and credit card holders who have experienced fraud at some time in their life rose to 65% in 2022 from the 58% recorded in 2021, according to the 2023 Credit Card Fraud Report.

Fraud prevention is crucial in the world of digital transactions, and two primary techniques are used: misuse detection and anomaly detection. Misuse detection uses machine learning to differentiate between normal and possibly fraudulent transactions based on identified patterns, much like a cunning investigator. It's like when a watchful security dog knows the difference between known and unknown activity. However, anomaly detection functions as a watchful defender, identifying the typical characteristics of transactions to build a unique profile and sounding an alarm when deviations take place. It is similar to a security system that is tuned in to normal patterns and prepared to warn users to any unusual events.

Our study attempts to open up new avenues for credit card fraud detection in response to the crucial intersection of financial vulnerability and technology. We provided a detailed analysis of sophisticated deep learning architectures, such as transformer-based models, recurrent neural networks (RNNs), and convolutional neural networks (CNNs). Rethinking the fundamentals of credit card fraud detection in terms of effectiveness, flexibility, and practicality goes beyond just fortifying security mechanisms. In the face of growing fraud complication in financial environments, our work provides direction where

innovation and necessity converge. In the current era, where digital transaction security is critical, this study aims to rethink robustness along with effectiveness criteria. Our work seeks to significantly contribute to the continuing discussion on financial system protection by addressing the difficulties presented by contemporary fraud strategies and offering creative solutions. Come along on this voyage into the future of credit card fraud detection, where cybersecurity imperatives and innovation meet as we work towards an even more adaptable and safeguarded financial landscape.

II. LITERATURE REVIEW

In the context of e commerce and electronic payment systems, authors Emmanuel Ileberi, Yanxia Sun & Zenghui Wang [1] introduced a credit card fraud detection engine utilizing Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), and Naive Bayes (NB) of Machine Learning classifiers. It underscores the necessity for robust ML-based solutions for the challenges faced in the evolving nature of fraudulent transactions and the skewed distribution in fraud datasets. To mitigate the challenge, the paper introduced a notable strategy which optimizes the selection of features incorporating the RF method in its fitness function which is a part of feature selection algorithm based on Genetic Algorithm (GA). Using a dataset created from European credit cardholders, an accuracy of 99.98% for GA-RF is attained which demonstrates the superiority of the GA-based feature selection technique. Still due to the highly confidential nature of credit card transaction datasets, experiments can't be fully replicated because ML models use anonymized attributes.

The authors of the article, Shanshan Jiang, Ruiting Dong, Jie Wang, and Min Xia [2], talked about how the widespread use of internet technology has made credit card fraud a growing problem. It draws attention to the shortcomings of conventional machine learning techniques in identifying unfamiliar attack patterns and presents UAAD-FDNet, a novel framework that makes use of autoencoders, feature attention, and generative adversarial networks (GANs) in an Unsupervised Attentional Anomaly Detection Network. Based on tests using the IEEE-CIS Fraud Detection Dataset and the Kaggle Credit Card Fraud Detection Dataset, the suggested strategy asserts better performance in fraud detection. It highlights how technical developments have contributed to the surge in sophisticated credit card fraud, which includes techniques like payment fraud and credit card cashing. Fraud prevention and detection are critical study subjects because of the substantial economic consequences of the problems presented by these developing approaches. The significance of machine learning in fraud detection is highlighted, as it surpasses the constraints of conventional techniques. The authors [2] presented a brand-new unsupervised attentional anomaly detection paradigm and conducted a thorough analysis of the state of credit card fraud detection techniques. The suggested UAAD-FDNet uses a generator

and discriminator to function according to GAN principles. To improve model training, the generator uses a channel-wise feature attention mechanism and self-supervised learning; for supervision, a hybrid weighted loss function is recommended. The suggested method's superiority is demonstrated by experimental findings on pertinent datasets, which frame credit card fraud detection as anomaly detection and provide unique features in UAAD-FDN et.

Jiwon Chung and Kyungho Lee [3] discussed in the context of the growing commerce landscape and IoT devices, how the issue of fraudulent credit card transactions has emerged. By combining three machine learning models: K-nearest neighbor (KNN), linear discriminant analysis (LDA), and linear regression (LR), a significant progress has been represented in credit card fraud detection and mostly to improve the sensitivity and accuracy, the incorporation of operators and conditional statements made the method unique from others. The substantial financial loss in the global surge due to internet fraud, notably in the US, due to the alarming rise of credit card fraud incidents (65% of credit card holders falling victim to fraud in 2022, up from 58% in the previous year) the urgency of the matter is highlighted. The study's credit card fraud real-world datasets are subjected to the integrated technique, showcasing recall scores of 1.0000, 0.9701, 1.0000, and 0.9362 for each of four dataset. An automated machine learning framework, PyCaret was used to validate the methodology and show that it outperforms other individual models. But alongside this, the omission of techniques like regularization and sampling methods to address skewed dataset created uncertainties in benchmarking against state-of-the-art fraud detection and thus for developing strategies and enhanced precision ,future research is suggested.

The paper of Noor Saleh Alfaiz and Suliman Mohamed Fati [4] talks about how the COVID-19 epidemic has made people more reliant on internet services, which has increased credit card theft. Using a dataset of European cardholders, the study investigates 66 machine learning models for credit card fraud detection. Nine algorithms are tested in the first assessment stage, and the top three algorithms are tested in the second stage using 19 resampling strategies. The AllKNN-CatBoost model performs better than previous models with a 97.94% AUC, 95.91% Recall, and 87.40% F1-Score. The research addresses the difficulties associated with imbalanced datasets in fraud detection and suggests a sophisticated method that combines the most efficient machine learning algorithms with practical resampling strategies. The suggested method, which assesses resampling methods and machine learning algorithms for credit card fraud detection, is broken down into two phases. A total of 66 models are obtained by evaluating nine methods and 19 resampling strategies. The best model is found to be AllKNN-CatBoost, which outperforms the others in terms of F1-Score, AUC, and recall. In order to highlight the significance of correcting class imbalance in credit card fraud detection, the research highlights how uncommon

balanced datasets are in real-world circumstances. One month is allotted to the comprehensive evaluation procedure, which illustrates how comprehensive the methodology is. In order to prevent more sophisticated fraud, the paper's conclusion underlines the necessity for proactive solutions that make use of AI and machine learning. All things considered, the work presents a novel strategy to address class disparity in credit card fraud detection, exhibiting encouraging outcomes and providing directions for further investigation.

In their paper, the authors J Femila Roseline, GBSR Naidu, Dr. V. Samuthira Pandi, S Alamelu alias Rajasree and Dr.N. Mageswari [5] discuss the growing issue of credit card fraud, which is a result of both increased card use and technology improvements. It promotes the use of a Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN) with an attention mechanism as an innovative approach to an efficient fraud detection system. When compared to alternative classifiers, the suggested model exhibits strong performance and high accuracy. The need of protecting digital transactions is emphasized, but the difficulties in spotting fraudulent activity are acknowledged. Although automated systems are capable of identifying unusual behavior, the analysis that follows needs human interaction, which presents a financial hurdle. The study emphasizes how well machine learning—particularly LSTM-RNN—performs in pattern detection. The topic of fraud detection in real-time is covered, using human rules and machine learning algorithms to quickly identify transactions that seem suspect. The main objective is to identify suspect credit card transactions by using machine learning algorithms (Naive Bayes, SVM, ANN, and LSTM) for pattern identification. The creation of a novel fraud detection system based on LSTM-RNN, effective behavior prediction of illicit service charge behavior, and competitive performance in comparison to current techniques are among the contributions. The study concludes by highlighting the need of adjusting to big data difficulties and presenting an LSTM-based model that is supported by a variety of datasets and performance measures for the identification of credit card fraud.

In the study, Fahdah A. Almarshad, Ghada Abdalaziz Gashgari, and Abdullah I. A. Alzahrani [6] explores the growing issue of credit card fraud in the age of digital transactions and emphasizes the demand for sophisticated fraud detection techniques. For the traditional methods, the complexity of fraudulent activities is a matter of struggle to cope up, which is why innovation in the industry is essential. So for this reason, the authors came up with a unique method for detecting fraud using Generative Adversarial Networks (GANs) which eventually solve the imbalance in datasets connected to credit cards and also overcome the difficulties presented by developing fraud approaches. The study emphasizes the model's efficacy against a number of measures, highlighting its resilience, decreased false positives, and increased efficiency while highlighting the significance of

developing a substantial dataset for fraud identification. The development of a new GANs-based identification method, dataset generation, resilience, a drop in false positives, a reduction in processing needs, and an increase in accuracy are among the contributions. The study does, however, admit its shortcomings, including the lack of available datasets, privacy issues, and the need for more investigation. Along with this, the study makes clear that more investigation is required to handle new fraud strategies and enhance the suggested methodology because privacy issues arise whenever there is transaction of data using a credit card.

Authors, Zainab Firdous, Sushma, Aftab Pasha, M Shahista Banu, and Najmusher [7], in their paper have provided a thorough comparison between the different classification models. Such as Logistic regression, support vector machine and the random forest classification model. These models are compared on the account of recall precision and accuracy. For their research they have used UCI machine learning repository . Their data contained users payment history, bill amount, amount paid by the user and also the payment static from april to september. To Visualize the data they have used MATPLOTLIB and seaborn libraries. They have used the sklearn library to scale and preprocess the data. Moreover this paper also shows different works of researchers and the overview of their existing methods and approaches. The research work briefs about all the algorithms and their respective performance on the dataset. Provided a confusion matrix upon their results. To achieve the goal to evaluate the most accurate and effective model for Credit Card fraud Detection, this paper also offers their own experimental investigation of classification algorithms. From their analysis they have concluded that both super vector machine and random forest have a good accuracy compared to logistic regression. Even by tuning the threshold slightly, best recall, precision and accuracy can be achieved.

The writers Varun Kumar, Vijaya Kumar, Vijaya Shankar, Pratibha K [8] tried to build a model which predicts fraud and non fraud transactions with efficiency using machine learning algorithms. The prediction will be on account of time and the amount of the transaction using classification models and linear algebra in building more complex machine learning models. The authors started by briefing about the need for such an efficient solution for Credit card fraud detection. Moreover they have described each algorithm's working process and their performance upon the DATASET. This research paper worked with a dataset containing the transaction form europe card owners where 492 out of 2,84,807 are fraud transactions. The Dataset was then converted into PCA transformation thus it contains numeric values. To predict the result they have used statistical analysis to visualize the data. Furthermore, they have used a confusion matrix of all the ML models such as RF, NB, ANN, logistic regression for better comparison. Upon their thorough analysis, they have concluded that ANN has more accuracy of 98 percent from the rest of the ML

models.

The authors of this paper, Priti jadhav, Rutuja Ghadge, Utkarsha Halpatrao, Prof. Vilas Jadhav [9] worked on modeling the dataset for credit card fraud detection. They tried to achieve detection of fraudsters and the number of them. They have used machine learning algorithms such as Artificial neural networks, Gradient Boosting Algorithms. Logistic regression and DTM. On analyzing logistic regression, and more of the ML models. The authors conclude in their study that the gradient boosting model has more accuracy than any of the ML models and also it helps in teaching how the credit card detection model can be improved.

III. DATASET

Dataset contains 2,84,807 transactions of Europe card owners. From here only 492 transactions are fraud. Thus, the dataset is not balanced. It has fewer fraud cases compared to the huge number of transactions. Also, the dataset is in PCA transformation only containing the numeric values. Here only time and amount are not converted into PCA value. All the other values from V1- V28 are transformed PCA values. For privacy concerns, many values are already given as PCA values. As for the feature value in this dataset, the feature class contains 1 for fraud and 0 for normal transactions.



Fig. 1. Transaction Dataset.

IV. METHODOLOGY

We are going to follow a few steps to understand the problem and the data. First, visualization and statistical analysis upon the data will be performed, to see how much of the data is imbalanced. Later on, to balance the data, oversampling

and scaling will be used. standardization and normalization will also be used to scale the data. Furthermore, to run the ML models and visualize the dataset, Numpy, Matplotlib and seaborn libraries will be used respectively.

REFERENCES

- [1] Ileberi, E., Sun, Y., & Wang, Z. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data*, 9(1), 1-17.
- [2] Jiang, S., Dong, R., Wang, J., & Xia, M. (2023). Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network. *Systems*, 11(6), 305.
- [3] Chung, J., & Lee, K. (2023). Credit Card Fraud Detection: An Improved Strategy for High Recall Using KNN, LDA, and Linear Regression. *Sensors*, 23(18), 7788.
- [4] Alfaiz, N. S., & Fati, S. M. (2022). Enhanced credit card fraud detection model using machine learning. *Electronics*, 11(4), 662.
- [5] Roseline, J. F., Naidu, G. B. S. R., Pandi, V. S., alias Rajasree, S. A., & Mageswari, N. (2022). Autonomous credit card fraud detection using machine learning approach. *Computers and Electrical Engineering*, 102, 108132.
- [6] Almarshad, F. A., Gashgari, G. A., & Alzahrani, A. I. (2023). Generative Adversarial Networks-Based Novel Approach for Fraud Detection for the European Cardholders 2013 Dataset. *IEEE Access*.
- [7] Journal, I. R. J. E. T. (2023). Credit Card Fraud Detection Using Machine Learning. *IRJET*.
- [8] S. Varun. (2020). Credit Card Fraud Detection using Machine Learning Algorithms. *International Journal of Engineering Research and*. V9, 10.17577/IJERTV9IS070649.
- [9] Journal, I. (n.d.). Credit Card Fraud Detection using Machine Learning. https://www.academia.edu/51587581/Credit_Card_Fraud_Detection_using_Machine_Learning.