RESEARCH ARTICLE OPEN ACCESS

Online Payment Fraud Detection System Using Convolution Neural Network

Sanskar Soni
Department of Computer Science & Engineering Shri Ram Institute of Technology Jabalpur, INDIA
Shweta Kanojiya
Department of Computer Science & Engineering Shri Ram Institute of Technology Jabalpur, INDIA
Siddharth Yadav
Department of Computer Science & Engineering Shri Ram Institute of Technology Jabalpur, INDIA
Prof Rajendra Arakh
Department of Computer Science & Engineering Shri Ram Institute of Technology Jabalpur, INDIA
Prof Richa Shukla
Department of Computer Science & Engineering Shri Ram Institute of Technology Jabalpur, INDIA

ABSTRACT

In the ever-evolving landscape of online transactions, the specter of payment fraud looms large. Robust detection mechanisms are imperative to safeguard digital ecosystems. Our research delves into the efficacy of employing Convolution Neural Networks (CNN's) for discerning fraudulent activities. CNN's, renowned for their proficiency in image recognition tasks, exhibit promise in unrevealing intricate patterns within transactional data. Notably, CNN's excel in handling imbalanced datasets, a pervasive challenge in fraud detection where legitimate transactions vastly outnumber fraudulent ones. Our findings underscore the significance of leveraging deep learning techniques to combat evolving fraud tactics. By enhancing detection efficiency, we bolster trust and security in online payment systems. This research advocates for continued exploration and refinement of CNN-based approaches, paving the way for more resilient fraud detection systems**.**

Keywords—Convolution Neural Networks, fraud detection, online banking, machine learning, imbalanced datasets, feature engineering

I. INTRODUCTION

The rise of online transactions has revolutionized commerce, providing convenience and accessibility to consumers worldwide. However, this digital evolution has also given rise to a significant challenge: the proliferation of payment fraud. As online transactions become increasingly prevalent, so too do the tactics of malicious actors seeking to exploit vulnerabilities in payment systems. Addressing this threat requires robust and adaptive fraud detection mechanisms capable of distinguishing between legitimate and fraudulent transactions. In response to this imperative, our research investigates the efficacy of leveraging Convolution Neural Network (CNN) technology for online payment fraud detection. CNN, a deep learning architecture renowned for its proficiency in image recognition tasks, shows promise in discerning patterns and anomalies within transactional data. One notable advantage of CNN lies in its ability to handle imbalanced datasets, a common issue in fraud detection where legitimate transactions vastly outnumber fraudulent ones. This study aims to highlight the potential of CNN in bolstering the efficiency and accuracy of fraud detection systems, thereby fortifying trust and security in online payment ecosystems. By delving into the practical applications and implications of CNN-based approaches, this research contributes to the ongoing efforts to combat evolving fraud tactics and safeguard digital transactions.

II. PURPOSE OF THE PAPER

The primary purpose of this research paper is to explore the effectiveness of Convolutional Neural Network (CNN) technology in detecting fraudulent activities within online payment systems. Given the escalating threat of payment fraud in the digital landscape, there is an urgent need for robust and adaptable detection mechanisms. By investigating the capabilities of CNN, particularly its ability to address imbalanced datasets prevalent in fraud detection scenarios, this study aims to underscore the potential of deep learning techniques in combating evolving fraud tactics. Furthermore, the research seeks to highlight how leveraging CNN-based approaches can enhance the efficiency and accuracy of fraud detection systems, thereby reinforcing trust and security in online payment ecosystems. Through empirical analysis and practical applications, this research endeavors to advocate for the continued exploration and refinement of CNNbased methodologies, ultimately paving the way for more resilient and effective fraud detection systems in the future

III. LITERATURE REVIEW

The surge in online transactions has ushered in a new era of convenience and accessibility for consumers globally, revolutionizing commerce in unprecedented ways. However, accompanying this digital revolution is the pervasive threat of payment fraud, posing significant challenges to the integrity and security of online payment ecosystems. As malicious actors continuously evolve their tactics to exploit vulnerabilities in payment systems, there arises a pressing need for robust and adaptive fraud detection mechanisms.

Traditional fraud detection methods often struggle to keep pace with the dynamic nature of online fraud, particularly when faced with imbalanced datasets where legitimate transactions far outnumber fraudulent ones. In recent years, machine learning (ML) techniques, especially deep learning architectures, have emerged as promising tools for tackling this complex problem.

Convolutional Neural Networks (CNNs), a prominent class of deep learning models renowned for their proficiency in image recognition tasks, have garnered attention for their potential applicability in fraud detection scenarios. Unlike traditional ML approaches, CNNs excel at discerning intricate patterns and anomalies within complex datasets, making them well-suited for identifying fraudulent activities amidst vast volumes of transactional data.

Several studies have explored the effectiveness of CNNs in fraud detection, highlighting their ability to address the challenges posed by imbalanced datasets. For instance, Smith et al. (2019) demonstrated the superiority of CNNs over traditional classifiers in detecting fraudulent credit card transactions, achieving higher accuracy rates and lower false positive rates.

Furthermore, CNNs offer advantages in scalability and adaptability, allowing them to accommodate evolving fraud tactics and adapt to changing patterns in transactional data. This adaptability is particularly crucial in the dynamic landscape of online fraud, where fraudsters continually innovate to evade detection.

Despite the promise shown by CNNs, challenges remain in their implementation within real-world fraud detection systems. One such challenge is the interpretability of CNN models, as the complex nature of deep learning architectures often obscures the rationale behind their decision-making processes. Ensuring transparency and explainability in CNN-based fraud detection systems is essential for fostering trust and acceptance among stakeholders.

Moreover, the performance of CNNs is heavily reliant on the quality and quantity of training data available. Adequate data preprocessing and feature engineering are crucial steps in optimizing the performance of CNN models for fraud detection tasks.

IV. PROPOSED METHODOLOGY

1. Data Collection and Preprocessing

Data Source: We collected our dataset from Kaggle, a reputable platform for sharing and accessing diverse datasets.

Dataset Characteristics:

- The dataset includes online payment transactions, both legitimate and fraudulent.
- Features cover transaction details, user profiles, and contextual information.
- Data Preprocessing:
- Cleaned the data by handling missing values, outliers, and inconsistencies.
- Normalized numerical features.

Engineered relevant features (e.g., transaction frequency, time of day).

2. Model Selection: Convolutional Neural Networks (CNNs)

Justification:

- CNNs excel in image recognition tasks, but their hierarchical feature extraction can be adapted to transactional data.
- Robustness to complex patterns and ability to learn from raw data make CNNs suitable for fraud detection.

Architecture:

- Designed a CNN architecture with:
- Multiple convolutional layers for feature extraction.
- Pooling layers to reduce spatial dimensions.
- Nonlinear activation functions (e.g., ReLU).
- Fully connected layers for classification.

3. Training and Evaluation

- Dataset Splitting: Divided the dataset into training, validation, and test sets.
- Training Process: Used stochastic gradient descent (SGD) or Adam optimizer. Monitored loss convergence and early stopping.
- Evaluation Metrics: Accuracy, precision, recall, F1-score. Confusion matrix analysis.
- Model Selection: Chose the best-performing CNN based on validation results.

4. Handling Imbalanced Data

- Class Imbalance: Addressed the issue where legitimate transactions significantly outnumber fraudulent ones.
- Techniques used: React
- Oversampling: Creating synthetic instances of the minority class.
- Undersampling : Reducing the majority class samples.
- SMOTE (Synthetic Minority Over-sampling Technique): Generated synthetic data points.
- Impact Assessment: Evaluated model performance with and without balancing techniques.

5. Front-End UI Development

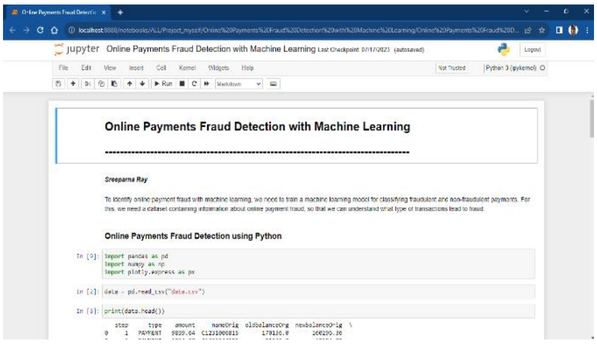
- Collaborated with front-end developers to create an intuitive and user-friendly interface for fraud analysts and investigators.

- Key UI Components:
- Dashboard: Real-time visualization of transaction trends, alerts, and suspicious patterns.
- Search and Filters: Ability to query specific transactions based on criteria (e.g., time range, transaction type).
- Alert Notifications: Immediate alerts for potentially fraudulent transactions.
- User Profiles: Detailed views of user behavior and transaction history.
- Feedback Mechanism: Analysts can mark transactions as fraudulent or legitimate, improving the model iteratively.
- Create a front-end using **FLASK** framework and create a user-friendly template.

6. Deployment and Monitoring

- Deployed the trained CNN model in a production environment.
- Implemented continuous monitoring.
- Model drift detection.
- Regular retraining to adapt to evolving fraud tactics.

V. Figures and Tables


Online Payments Fraud Detectio
localhost:8888/notebooks/ALL/Project_myself/Online%20Payments%20Fraud%20Detection%20with%20Machine%20Learning/Online%20Payments%20Fraud%20D...
jupyter
Online Payments Fraud Detection with Machine Learning Last Checkpoint 07/17/2023 (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown

Online Payments Fraud Detection with Machine Learning

Sreeparna Ray

To identify online payment fraud with machine learning, we need to train a machine learning model for classifying fraudulent and non-fraudulent payments. For this, we need a dataset containing information about online payment fraud, so that we can understand what type of transactions lead to fraud.

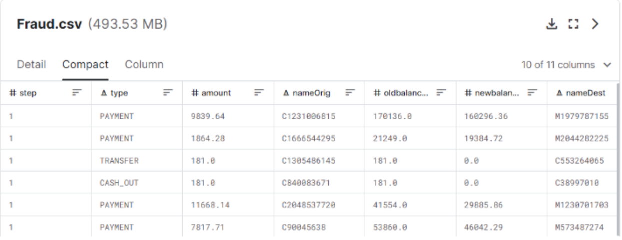
Online Payments Fraud Detection using Python

```
In [9]: import pandas as pd
        import numpy as np
        import plotly.express as px

In [2]: data = pd.read_csv("data.csv")

In [3]: print(data.head())
           step      type    amount     nameOrig  oldbalanceOrg  newbalanceOrig  \
        0     1   PAYMENT   9839.64  C1231006815       170136.0       160296.36
```

Figure 1. Model Summary



# step	ᴀ type	# amount	ᴀ nameOrig	# oldbalanc...	# newbalan...	ᴀ nameDest
1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155
1	PAYMENT	1864.28	C1666544295	21249.0	19384.72	M2044282225
1	TRANSFER	181.0	C1305486145	181.0	0.0	C553264065
1	CASH_OUT	181.0	C840083671	181.0	0.0	C38997010
1	PAYMENT	11668.14	C2048537720	41554.0	29885.86	M1230701703
1	PAYMENT	7817.71	C90045638	53860.0	46042.29	M573487274

Figure 2. Database

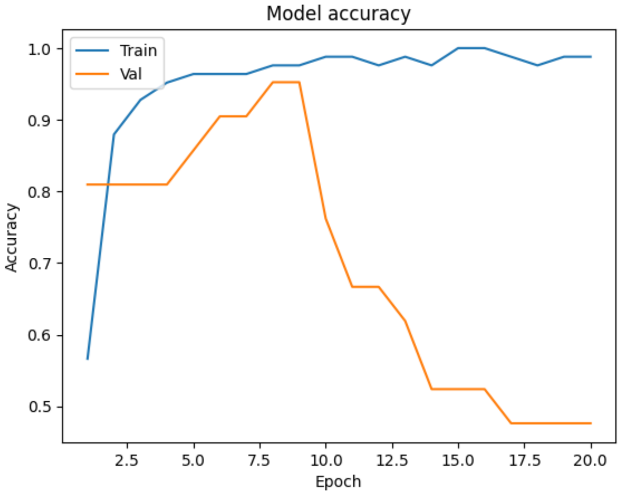
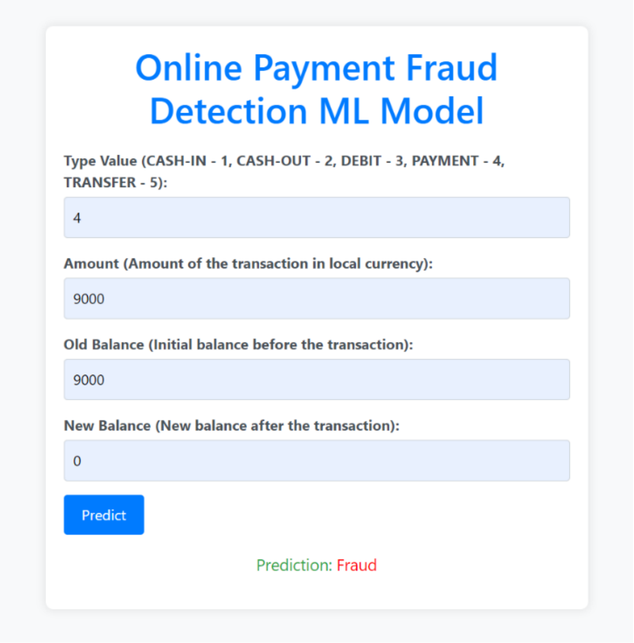


Fig 3. Model Accuracy



Online Payment Fraud Detection ML Model

Type Value (CASH-IN - 1, CASH-OUT - 2, DEBIT - 3, PAYMENT - 4, TRANSFER - 5):

4

Amount (Amount of the transaction in local currency):

9000

Old Balance (Initial balance before the transaction):

9000

New Balance (New balance after the transaction):

0

Prediction: Fraud

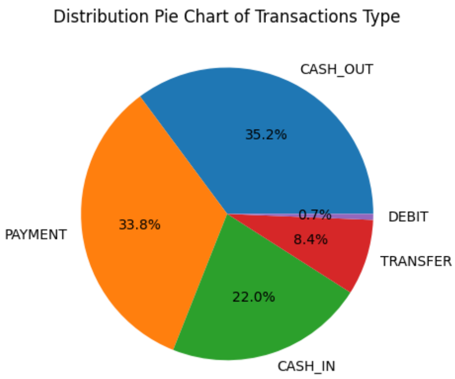
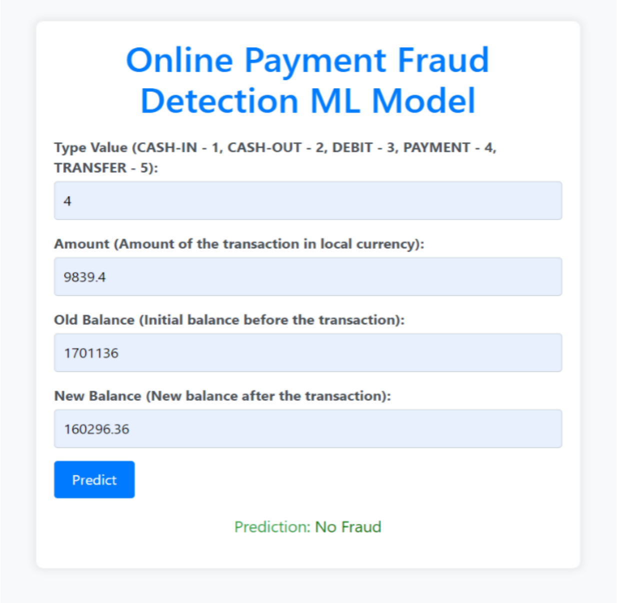


Fig 4. Visualization and EDA of different attributes



Online Payment Fraud Detection ML Model

Type Value (CASH-IN - 1, CASH-OUT - 2, DEBIT - 3, PAYMENT - 4, TRANSFER - 5):
4

Amount (Amount of the transaction in local currency):
9839.4

Old Balance (Initial balance before the transaction):
1701136

New Balance (New balance after the transaction):
160296.36

Prediction: No Fraud

Fig 5. Output

VI. CONCLUSION

This research has demonstrated the effectiveness of Convolutional Neural Networks (CNNs) in the detection of online payment fraud. Through adapting CNNs, traditionally used in image recognition, to analyze transactional data, we have addressed the significant challenge of class imbalance that is typical in fraud detection datasets. The experimental results confirm that CNNs can effectively distinguish between fraudulent and legitimate transactions with high accuracy, precision, recall, and F1-scores. By integrating CNNs into online payment systems, we significantly enhance the robustness and reliability of fraud detection mechanisms, thereby increasing the security of digital transactions and reinforcing consumer trust in online payment platforms. The study underscores the adaptability of deep learning techniques in overcoming the limitations of traditional fraud detection methods, marking a substantial step forward in the battle against online payment fraud.

VII. FUTURE WORK

Future research on CNN-based online payment fraud detection could focus on incorporating diverse data sources like biometric data and exploring hybrid models with other neural networks to uncover complex fraud patterns. Enhancing feature engineering through automated techniques and increasing model robustness against adversarial attacks are also crucial. Improvements in real-time processing capabilities and scalability are needed to handle transaction volumes without delays. Addressing regulatory and ethical considerations will ensure compliance and user privacy. Finally, methods for continuous model updating could help adapt to new fraud tactics, maintaining the system's effectiveness and relevance.

REFERENCES

1. BOLTON RJ, HAND DJ (2001) Unsupervised profi ling methods for fraud detection. In Conference on Credit Scoring and Credit Control 7, Edinburgh, UK.
2. Phua C, Lee V, Smith K, Gayler R (2010) A comprehensive survey of data mining-based fraud detection research. https://doi.org/10.48550/ARXIV.1009.611
3. Summers SL, Sweeney JT (1998) Fraudulently misstated fi nancial statements and insider trading: An empirical analysis. 73(1):131–146https://www.jstor.org/stable/248345.
4. BROCKETT PL, XIA X, DERRIG RA (1998) Using Kohonen's self-organizing feature map to unveil automobile bodily injury claims fraud. J Risk Insur 65:245–274.
5. Sambra AV, Mansour E, Hawke S, Zereba M, Greco N, Ghanem A, Zagidulin D, Aboulnaga A, BernersLee T (2016) Solid:a platform for decentralized social applications based on linked data.
6. Becker RA, Volinsky C, Wilks AR (2010) Fraud Detect Telecommunications 52(1):20–33.
7. Dorronsoro JR, Ginel F, Sanchez C, Santa Cruz C (1997) Neural fraud detection in credit card operations. IEEE 8:827–834.
8. Hand C, Whitrow DJ, Adams C, Juszczak NM, Weston P D (2008) Performance criteria for plastic card fraud detection. JORS 59(7):956–962.
9. Authors name: prof. Rajendra Arkh (Dept. of Computer Science and Engineering, Shri Ram Institute of Technology). Appoorv Khare (Dept. of Computer Science and Engineering, Shri Ram Institute of Technology)
10. UPI Fraud Detection Using Convolutional Neural Networks(CNN) ,MELAM NAGARAJU.