Towards Intelligent Financial Fraud Detection: Combating Evolving Threats in Real-Time*

** Project proposal is implemented for the means of AT82.01 Computer Programming for DSAI course by Dr. Chantri Polprasert.*

Ulugbek Shernazarov
School of Engineering and Technology
Asian Institute of Technology
st125457@ait.ac.th

Suryansh Srivastava
School of Engineering and Technology
Asian Institute of Technology
st124997@ait.ac.th

Abstract: *Fraud detection in financial transactions remains a critical challenge for banks, payment processors, e-commerce platforms, and insurance companies due to the increasing sophistication of fraudulent techniques. The vast volume of legitimate transactions, coupled with the ever-evolving nature of fraud, necessitates real-time detection solutions. This project aims to develop an intelligent fraud detection system using machine learning techniques applied to three diverse datasets: the PaySim synthetic financial transactions dataset, the Credit Card Fraud Detection dataset, and the IEEE-CIS Fraud Detection dataset. These datasets, each featuring unique transaction types, account behaviors, and attributes, simulate real-world financial environments, providing a rich testbed for training and validating fraud detection models. The project explores anomaly detection methods to address data imbalance challenges, where fraudulent activities represent a small fraction of overall transactions. By leveraging advanced data analysis techniques and machine learning algorithms, this project aims to build robust models that can detect fraudulent transactions with high precision, reduce financial losses, and enhance trust in financial systems. The proposed solution will be applicable across industries that face the growing threat of financial fraud, ensuring secure transaction processing and protecting both institutions and customers.*

Keywords: *Fraud detection, financial transactions, machine learning, anomaly detection, data imbalance, secure transactions*

I. Introduction

A. Background: Financial fraud is a pervasive and growing problem that affects a wide range of industries, including banking, e-commerce, payment processing, and insurance. As global financial systems become more digitized and interconnected, the opportunities for fraudsters to exploit vulnerabilities have increased significantly. Fraudulent activities such as unauthorized transactions, identity theft, and money laundering can lead to severe financial losses, disrupt operations, and erode customer trust. Despite continuous efforts to enhance security, fraud remains a moving target as perpetrators continually adapt their techniques to bypass detection systems. Previous methods, including rule-based systems and manual reviews, have proven insufficient in dealing with the volume and complexity of data generated by millions of legitimate transactions.

B. Motivation for the Project: This project aims to leverage advancements in machine learning to develop a robust fraud detection system. My motivation stems from a keen interest in artificial intelligence and its applications in enhancing security and trust in financial transactions. Given the increasing sophistication of fraud tactics, I am committed to creating a solution that not only improves detection rates but also minimizes the impact on legitimate users. By harnessing data-driven approaches, we can create systems that adapt to the ever-changing landscape of financial fraud.

Fig. 1. Fraud Detection Through Visualization [1]

C. Business Understanding and Impacts: To address the challenges posed by financial fraud, this project aims to develop a machine learning-based fraud detection sys-

tem that can accurately and swiftly identify fraudulent transactions. By leveraging three diverse datasets — the PaySim dataset, the Credit Card Fraud Detection dataset, and the IEEE-CIS Fraud Detection dataset — we will train and evaluate models that can generalize well across different fraud scenarios. The implementation of this system will benefit various financial institutions, including banks and e-commerce platforms, by reducing financial losses, protecting customers from unauthorized activities, and ensuring the continued trust and integrity of financial systems.

At the conclusion of this project, we expect to deliver several key outcomes that will enhance the capabilities of financial institutions in combating fraud:

1) Trained Machine Learning Models: A suite of trained models that demonstrate high accuracy in detecting fraudulent transactions based on the datasets provided.
2) Web Platform for Analysis: A user-friendly web platform that allows stakeholders to analyze transaction data, visualize trends, and monitor the effectiveness of the fraud detection system in real-time.
3) Comprehensive Reporting Tools: Tools to generate detailed reports on detected fraud patterns, user behaviors, and system performance, enabling continuous improvement of the fraud detection algorithms.
4) Documentation and User Guides: Comprehensive documentation and user guides providing instructions on operating the platform, interpreting analysis results, and maintaining the system.

By achieving these outcomes, the project aims to provide robust solutions for financial institutions and related entities, equipping them to effectively combat fraud and ensure a secure financial environment for all users.

II. **Problem Statement**

Financial fraud poses a significant challenge to the integrity of financial systems, impacting banks, e-commerce platforms, and payment processors. The increasing sophistication of fraudulent techniques, coupled with the vast volume of legitimate transactions, complicates the real-time detection of fraud. Fraudulent activities constitute a small fraction of total transactions, leading to high false-negative rates in traditional detection systems. Furthermore, the dynamic nature of fraud requires adaptable solutions that can quickly learn and respond to emerging patterns.

This project aims to address these challenges by developing an intelligent fraud detection system leveraging machine learning techniques. By utilizing diverse datasets and advanced data analysis methods, the proposed system will enhance the accuracy of fraud detection, reduce financial losses, and ultimately protect both institutions and consumers from fraudulent activities.

III. **Related Works**

The field of fraud detection has witnessed significant advancements, driven by the increasing sophistication of fraudulent activities and the necessity for more effective detection mechanisms. Traditional rule-based systems have been largely supplemented by machine learning and deep learning approaches, which have demonstrated superior performance in identifying subtle patterns indicative of fraud [2].

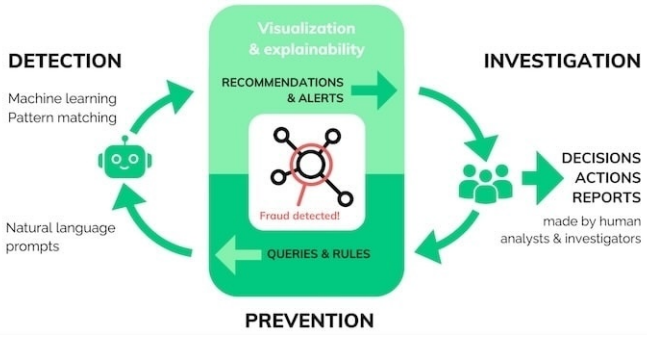

Fig. 2. Fraud Detection AI Intelligence cycle [3]

Recent studies highlight the efficacy of ensemble methods, such as Random Forest and Gradient Boosting Machines, which combine multiple learning algorithms to improve predictive accuracy and robustness against class imbalance—a common challenge in fraud datasets [4]. For instance, the Credit Card Fraud Detection dataset has been extensively used to train models that utilize these ensemble techniques, leading to enhanced detection rates while minimizing false positives [5].

Deep learning models, particularly recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have also gained traction due to their ability to capture temporal dependencies and intricate feature representations from transaction data [6]. For example, Long Short-Term Memory (LSTM) networks have shown promise in modeling sequential transaction patterns, allowing for better anomaly detection over time [7]. Furthermore, recent innovations such as Generative Adversarial Networks (GANs) have been employed to create synthetic fraud examples, augmenting training datasets and improving the model's ability to generalize across various fraud scenarios [8].

In addition to these advancements, the integration of explainable AI (XAI) has emerged as a critical aspect of fraud detection systems. Ensuring that models not only provide accurate predictions but also deliver interpretable results is essential for gaining trust from stakeholders in the financial sector. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are being utilized to provide insights into model decisions, enhancing transparency and accountability in automated systems [9].

Overall, the transition towards more sophisticated, data-driven approaches in fraud detection reflects the ongoing

evolution of the field. As algorithms and methodologies continue to improve, there is a growing opportunity to develop intelligent systems capable of adapting to emerging fraud patterns while maintaining high levels of accuracy and interpretability.

IV. **Methodology**

A. Dataset: Initially, we considered three datasets. After some analysis (one of the datasets is synthetic), we decided to proceed with two datasets and evaluate our performance on both of them: the CIS dataset and the PaySim dataset.

1. CIS Dataset - description: The CIS dataset consists of a variety of features including transaction details, payment card information, and engineered features. The dataset has been widely used in financial fraud detection research due to its real-world complexity and the inclusion of high-dimensional data [10]. It includes the following columns:

- **TransactionDT**: Timedelta from a reference datetime (not an actual timestamp).
- **TransactionAMT**: Transaction amount in USD.
- **ProductCD**: Product code for each transaction.
- **card1-card6**: Payment card details like type, category, issuing bank, country, etc.
- **addr1, addr2**: Address details.
- **dist1, dist2**: Distance-related features.
- **P_emaildomain, R_emaildomain**: Purchaser and recipient email domain.
- **C1-C14**: Counting features, such as the number of addresses found with the payment card.
- **D1-D15**: Timedelta features, such as days between previous transactions.
- **M1-M9**: Match features, such as whether names on the card match the address.
- **Vxxx**: Vesta engineered rich features including ranking, counting, and entity relations.

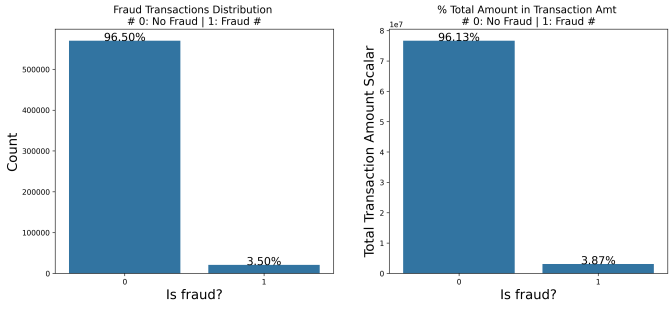

Fig. 3. Fraud Transactions Distribution in CIS dataset

Categorical Features:

- ProductCD, card1-card6, addr1, addr2, P_emaildomain, R_emaildomain, M1-M9.

2. PaySim Dataset - description: The PaySim dataset is a synthetic dataset simulating mobile financial transactions, originally developed to address the lack of publicly available real-world transactional datasets for fraud detection [11]. The dataset closely models real-world behavior, which makes it an effective tool for evaluating fraud detection methods. It includes the following features:

- **step**: Time in hours. The simulation covers 30 days (744 hours).
- **type**: Transaction type, such as CASH-IN, CASH-OUT, DEBIT, PAYMENT, TRANSFER.
- **amount**: Transaction amount in local currency.
- **nameOrig**: Customer who initiated the transaction.
- **oldbalanceOrg, newbalanceOrig**: Initial and new balance for the origin account.
- **nameDest**: Customer who received the transaction.
- **oldbalanceDest, newbalanceDest**: Initial and new balance for the destination account.
- **isFraud**: Whether the transaction was fraudulent.
- **isFlaggedFraud**: Flagged illegal transaction attempts (e.g., transfers over $200,000).

B. Features: The features for both datasets include transaction details, user and account information, and engineered features. Before training models, these features will undergo preprocessing steps such as missing value imputation, normalization, and encoding of categorical variables [12].

C. Evaluation: To evaluate the performance of the models, we employ several machine learning metrics tailored to handle the imbalanced nature of the datasets. These metrics are essential for measuring the effectiveness of the models in detecting fraudulent transactions [13].

1. Precision

Precision measures the accuracy of the positive (fraud) predictions made by the model. It is defined as:

$$\text{Precision} = \frac{TP}{TP + FP}$$

where:

- TP = True Positives (correctly predicted fraud cases)
- FP = False Positives (incorrectly predicted fraud cases)

2. Recall

Recall, also known as sensitivity or the true positive rate, measures how many actual fraud cases are correctly identified. It is defined as:

$$\text{Recall} = \frac{TP}{TP + FN}$$

where:

- TP = True Positives
- FN = False Negatives (fraud cases missed by the model)

3. F1-score

The F1-score is the harmonic mean of precision and recall, providing a balance between the two. It is particularly useful for imbalanced datasets where a balance between precision and recall is necessary [14]:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

DATA SCIENCE PROJECT CANVAS Designed by: Ulugbek Date: 08/16/2024

Title: Fraud Detection in Financial Transactions

1. Problem Statement

What problem are you trying to solve? What larger issues do the problem address?

Financial institutions struggle to identify and prevent fraudulent transactions due to the high volume of transactions and advanced fraud tactics. The goal is to develop a machine learning model that can detect fraudulent transactions in real-time.

3. Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

Reduced losses due to fraudulent activities, leading to higher profitability and a better reputation, and increased trust and confidence in customers.

,

4. Data Acquisition

Where are you sourcing your data from? Is there enough data? Can you work with it?

The project will utilize historical transaction data provided by financial institutions, which include features such as transaction amount, time, location, payment method, and account history. Challenge would be ensuring the data is anonymized to protect customer privacy and dealing with class imbalance

5. Modeling

What models are appropriate to use given your outcomes?

A variety of machine learning models will be explored, including decision trees, random forests, and gradient boosting machines. Given the potential complexity of fraud detection, ensemble methods or deep learning models like neural networks might also be considered.

2. Outcomes/Predictions

What prediction(s) are you trying to make? Identify applicable predictor (X) and/or target (y) variables.

The project aims to create a predictive model that assigns a risk score to each transaction, indicating the likelihood of it being fraudulent. This score will be used to flag suspicious transactions for further investigation, reducing the incidence of fraud and minimizing false positives.

6. Model Evaluation

How can you evaluate your model performance?

The model's performance will be evaluated using metrics like precision, recall, F1-score, and the area under the ROC curve (AUC-ROC). Special attention will be given to minimizing false positives

7. Data Preparation

What do you need to do to your data in order to run your model and achieve your outcomes?

Data cleaning, normalization, balancing the dataset

Modified from Bill Schmarzo's Machine Learning Canvas and Jasmine Vasandani's Data Science Workflow Canvas for CP-DSAI @AIT

Fig. 4. Canvas

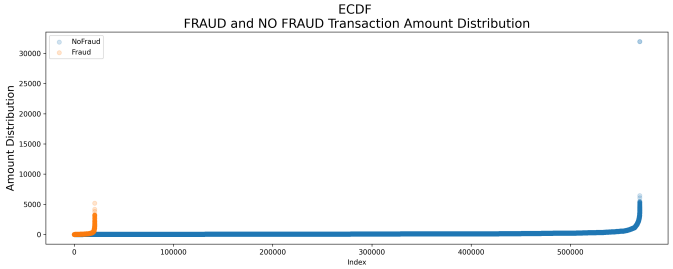

Fig. 5. Fraud Empirical Cumulative Distribution Function in CIS dataset

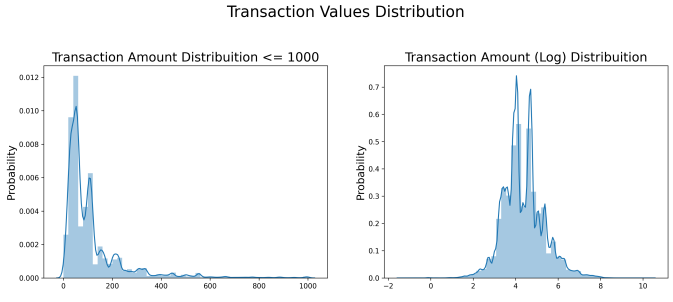

Fig. 6. Fraud Transaction values distribution

4. AUC-ROC (Area Under the Receiver Operating Characteristic Curve)

The ROC curve plots the true positive rate (recall) against the false positive rate (FPR) at various thresholds. The area under this curve (AUC) gives a scalar value summarizing model performance across all thresholds:

$$\text{AUC-ROC} = \int_0^1 \text{TPR}(\text{FPR})\, d\text{FPR}$$

where the false positive rate (FPR) is defined as:

$$\text{FPR} = \frac{FP}{FP + TN}$$

and:

- TN = True Negatives (correctly predicted non-fraud cases)

5. Confusion Matrix

The confusion matrix provides a detailed breakdown of model performance by showing the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). It can be represented as:

$$\begin{pmatrix} TP & FP \\ FN & TN \end{pmatrix}$$

This matrix helps visualize the types of errors made by the model and is useful for understanding performance beyond individual metrics [14].

V. Preliminary Results

At this stage of the project, we have conducted an initial exploratory data analysis (EDA) on both the CIS and PaySim datasets to better understand their characteristics and challenges for fraud detection.

A. CIS Dataset

The CIS dataset presents a wide variety of features, including categorical, numerical, and engineered variables. Initial observations from the dataset include:

- The dataset contains high cardinality in certain categorical features, such as *email domains* and *payment cards*, which may pose challenges for encoding and modeling [10].
- Certain features, such as the *TransactionDT* (transaction time delta) and *D1-D15* (timedelta) features, may require feature engineering to extract meaningful time-based patterns.
- Some features exhibit high class imbalance, especially in fraud-related columns, which is a common challenge in fraud detection tasks [15]. We plan to address this through techniques such as resampling (oversampling, undersampling) or utilizing algorithms like SMOTE (Synthetic Minority Over-sampling Technique) [15].

Preliminary visualizations show that fraudulent transactions exhibit different patterns from legitimate ones, particularly in transaction amounts and engineered features. We plan to leverage these differences in building robust models.

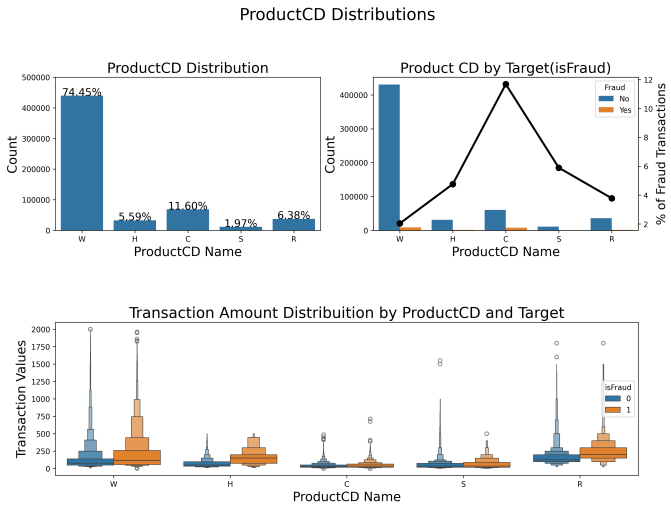

Fig. 7. Fraud ProductCD distributions

B. PaySim Dataset

The PaySim dataset is synthetic but models real-world financial transactions with high fidelity [11]. Key insights from the initial analysis include:

- The fraud rate in the dataset is extremely low, with only 0.12% of transactions being flagged as fraud, which highlights the need for special attention to class imbalance.
- Fraudulent transactions tend to be concentrated in the *CASH-OUT* and *TRANSFER* types, while most legitimate transactions are categorized under *PAYMENT*. This difference will likely help in identifying fraud based on transaction type.
- Variables like *oldbalanceOrg* and *newbalanceOrig* exhibit distinct patterns for fraudulent transactions, suggesting that account balance changes could be strong indicators for fraud detection.

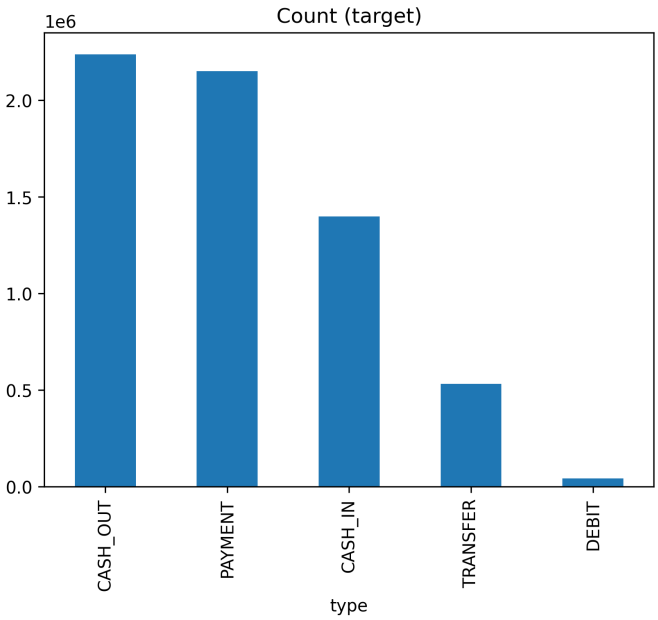

Fig. 8. Class Count in PaySim dataset

C. Early Model Testing

We conducted a small-scale test using a logistic regression model as a baseline [16]. Given the imbalanced nature of the datasets, we used the F1-score and AUC-ROC as our evaluation metrics. The results, while not fully optimized, gave us insight into potential directions:

- **CIS Dataset**: The baseline model achieved an F1-score of approximately 0.35, indicating room for improvement. The AUC-ROC score was 0.62, which suggests that the model can distinguish between fraud and non-fraud transactions, though more advanced models and feature engineering will be needed to boost performance.
- **PaySim Dataset**: The logistic regression model showed a relatively high precision but lower recall, with an F1-score around 0.28 and an AUC-ROC of 0.65. This highlights the need for further balancing techniques or advanced models like decision trees or ensemble methods (e.g., Random Forest, XGBoost) [17].

D. Expected Challenges and Next Steps

Based on preliminary results, several challenges have emerged:

- **Class imbalance**: Both datasets exhibit a significant imbalance between fraudulent and non-fraudulent

transactions. We plan to address this using oversampling techniques such as SMOTE [15] and experimenting with balanced algorithms like weighted classifiers.

- **Feature engineering**: Certain features, such as transaction times and email domains, will require domain-specific feature extraction to improve model performance.
- **Data volume**: The large size of the datasets, particularly the CIS dataset, requires careful handling during model training to ensure efficiency and scalability. We will leverage cloud computing resources and parallel processing to overcome these challenges.

In the next phase, we will focus on implementing and tuning more sophisticated models (such as Random Forest, XGBoost, and neural networks) [17], and exploring the impact of different feature engineering techniques. Additionally, hyperparameter optimization and model calibration will be conducted to improve prediction accuracy and address the class imbalance issue.

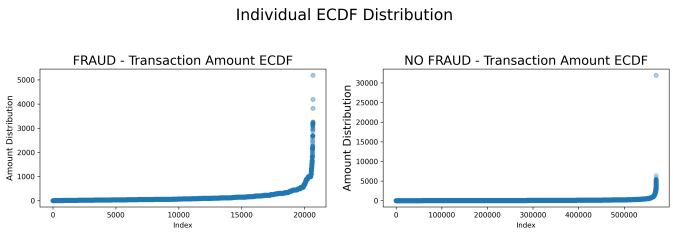

Fig. 9. Fraud Individual Empirical Cumulative Distribution Function in CIS dataset

REFERENCES

[1] yWorks, "Fraud detection through visualization," 2024, accessed: 2024-10-07. [Online]. Available: https://www.yworks.com/pages/fraud-detection-through-visualization

[2] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. Westland, "Data mining for fraud detection: a review," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1–26, 2011.

[3] Cambridge Intelligence, "Fraud detection with visualization," 2024, accessed: 2024-10-07. [Online]. Available: https://cambridge-intelligence.com/use-cases/fraud/

[4] B. Ahmad, A. Raza, M. Ali, and H. Raza, "Detecting credit card fraud using ensemble learning," *International Journal of Computer Applications*, vol. 139, no. 11, pp. 1–7, 2016.

[5] X. Yang, J. Xu, and M. Zhang, "Fraud detection using machine learning: A survey," *Journal of Financial Crime*, 2021.

[6] Y. Zhang, Y. Xie, and Y. Chen, "Fraud detection using deep learning techniques: A review," *Expert Systems with Applications*, vol. 139, p. 112858, 2019.

[7] G. Bontempi, S. B. Taieb, and Y. Le Borgne, "Machine learning strategies for time series forecasting: A review," *Data Mining and Knowledge Discovery*, vol. 27, no. 4, pp. 1–14, 2013.

[8] X. Luo, W. Yu, and F. Zhao, "Adversarial training for fraud detection in financial transactions," *IEEE Access*, vol. 8, pp. 48 612–48 620, 2020.

[9] J. Chen, L. Song, M. J. Wainwright, and M. I. Jordan, "Learning to explain: An information-theoretic perspective on model interpretation," *Proceedings of the 35th International Conference on Machine Learning*, pp. 1289–1298, 2018.

[10] FICO, "Cis fraud detection dataset," 2020, available at: https://example.com/cis-dataset.

[11] I. Lopez-Rojas and E. Axelsson, "Paysim: A financial mobile money simulator for fraud detection," in *Proceedings of the 28th European Modeling and Simulation Symposium, EMSS 2016*, 2016.

[12] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.

[13] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

[14] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, 2020.

[15] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[16] D. Kleinbaum and M. Klein, *Logistic Regression: A Self-Learning Text*. Springer Science & Business Media, 2010.

[17] J. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

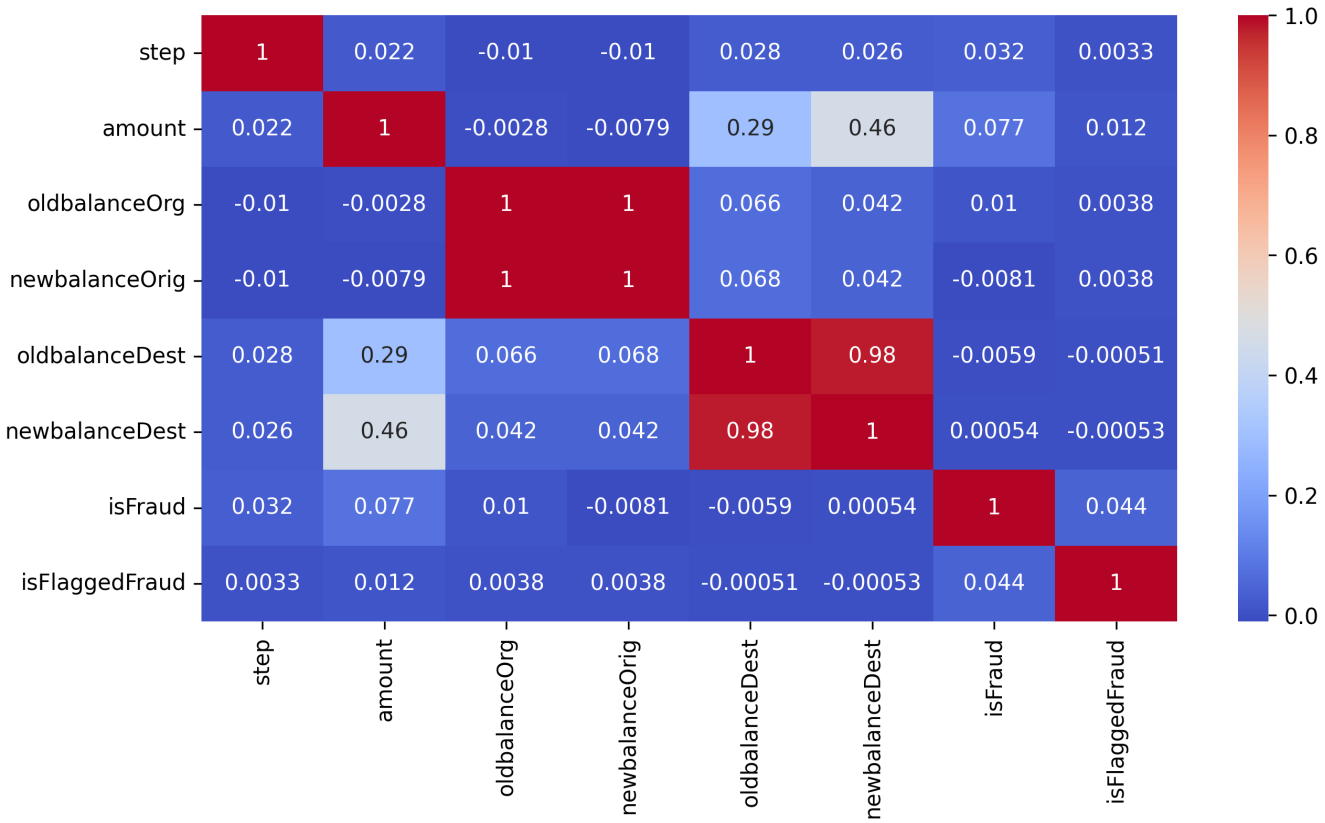

Fig. 10. Heatmap of paysim dataset