**J**ournal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**


# CREDIT CARD FRAUD DETECTION SYSTEM

**Swayansyu sahoo** 4th Year, Department of CSE, Gandhi Institute for Technology, BPUT,
swayanshu2021@gift.edu.in
**Jyotirmyee sethy** 4th Year, Department of CSE, Gandhi Institute for Technology, BPUT, India
jyotirmayee2021@gift.edu.in
[3] Assistant Professor, Department of CSE, Gandhi Institute for Technology, BPUT, India

- Introduction and Problem Statement
- Title: Credit Card Fraud Detection System
-
- 1. Introduction
-
- Credit card fraud is a growing concern in the digital age, as more transactions are conducted online and through electronic payment systems. With the rapid growth in e-commerce, there is an urgent need to develop intelligent systems that can automatically detect fraudulent transactions and prevent financial losses. This project presents a machine learning-based system for detecting credit card fraud by analyzing transaction data for suspicious behavior.
-
- 2. Problem Statement
-
- The primary objective of this project is to design and implement a model that accurately classifies credit card transactions as legitimate or fraudulent. Fraudulent activities are rare compared to legitimate transactions, leading to imbalanced datasets and increased difficulty in detection. Hence, the system must handle the class imbalance and ensure high accuracy with minimal false positives and false negatives.
-
- Page 2: Methodology and Implementation
-
- 3. Methodology
- The project uses supervised machine learning techniques for binary classification. The major steps involved are:
-
- Data Collection: A publicly available dataset from Kaggle containing anonymized credit card transaction records is used.
-
- Data Preprocessing: Includes handling missing values, normalizing data, and addressing class imbalance using techniques like SMOTE (Synthetic Minority Over-sampling Technique).
-
- Feature Selection: Dimensionality reduction techniques like PCA (Principal Component Analysis) are applied to retain important features and improve model efficiency.
-
- Model Building: Several models are trained, including Logistic Regression, Random Forest, Decision Tree, and XGBoost.

- Evaluation Metrics: Accuracy, Precision, Recall, F1-Score, and ROC-AUC are used to evaluate the models.

4. Implementation Details

Language/Tools: Python with libraries such as Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn.

Model Training: Models are trained using an 80-20 train-test split, and cross-validation is used to fine-tune hyperparameters.

Performance: Among all models, Random Forest and XGBoost show the best performance with an F1-Score above 0.90.

---

Page 3: Results, Conclusion, and Future Work

5. Results

The final models show high accuracy in identifying fraudulent transactions. Notably:

Random Forest: F1-Score = 0.93, AUC = 0.98

XGBoost: F1-Score = 0.95, AUC = 0.99

These models significantly reduce the number of false positives and false negatives, making them suitable for real-time deployment.

**6. Conclusion**

This project successfully demonstrates how machine learning can be effectively used to detect credit card fraud. By using ensemble models and proper data preprocessing, the system achieves high precision and recall. Handling class imbalance and optimizing hyperparameters were crucial to improving detection performance.

**7. Future Work**

- Integrate the model into a real-time fraud detection system.
- Enhance the dataset with real-world transactional metadata (like location, device, merchant info).
- Use deep learning techniques such as autoencoders or recurrent neural networks for advanced anomaly detection.

**8. References**

- Kaggle Credit Card Fraud Detection Dataset
- Scikit-learn Documentation
- Research papers on fraud detection and imbalanced data classification