CREDIT CARD FRAUD DETECTION

Student - Parth Bramhecha(33115)

Guide-Dr. Anant M. Bagade

Reviewer-Mrs. Prajakta S. Shinde

OBJECTIVES

Accurate Fraud Detection

distinguishing fraudulent transactions from legitimate ones with high precision and recall.

Real-Time Detection

analyzing transactions in real-time to prevent fraudulent charges from being processed

Low False Positive Rate

Minimize the number of legitimate transactions incorrectly flagged as fraudulent, reducing customer inconvenience.

Comparison

Compare ML models using accuracy, sensitivity, specificity, F1-score, and ROC-AUC to identify the best fraud detection model.

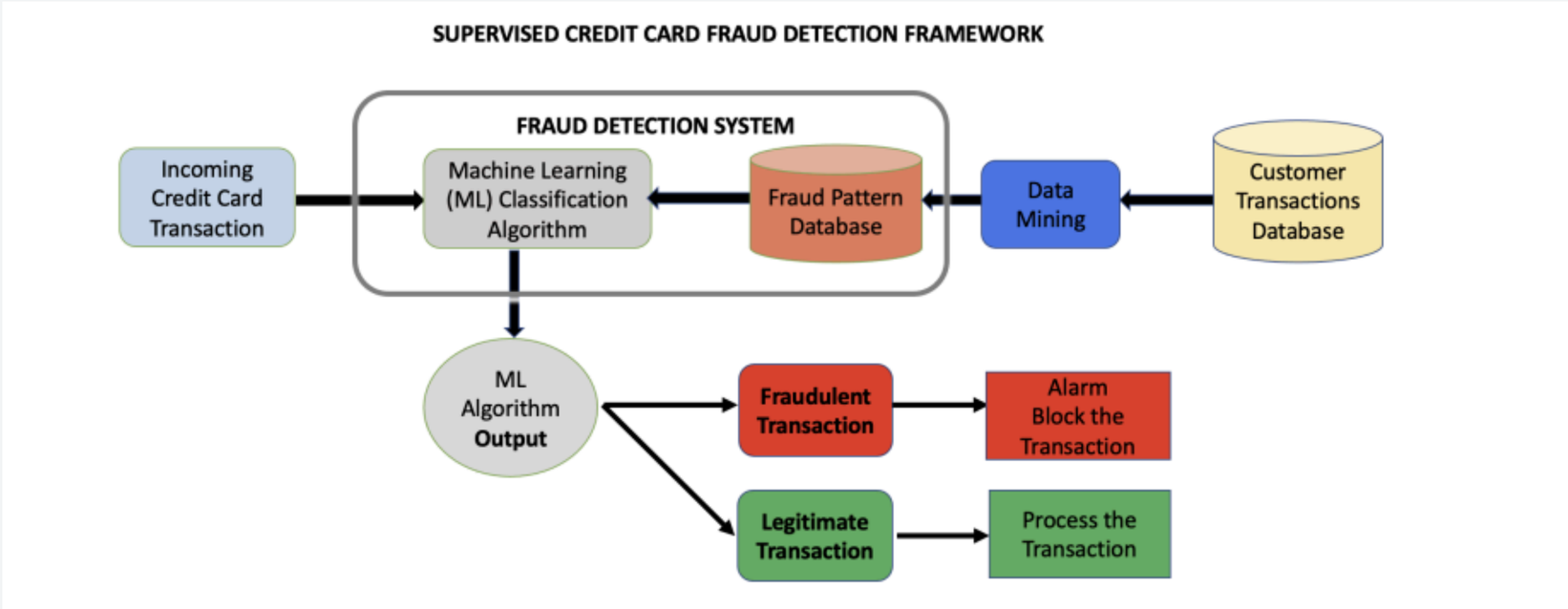
Adaptability

Create a system that can adapt to evolving fraud patterns and new types of fraudulent activities.

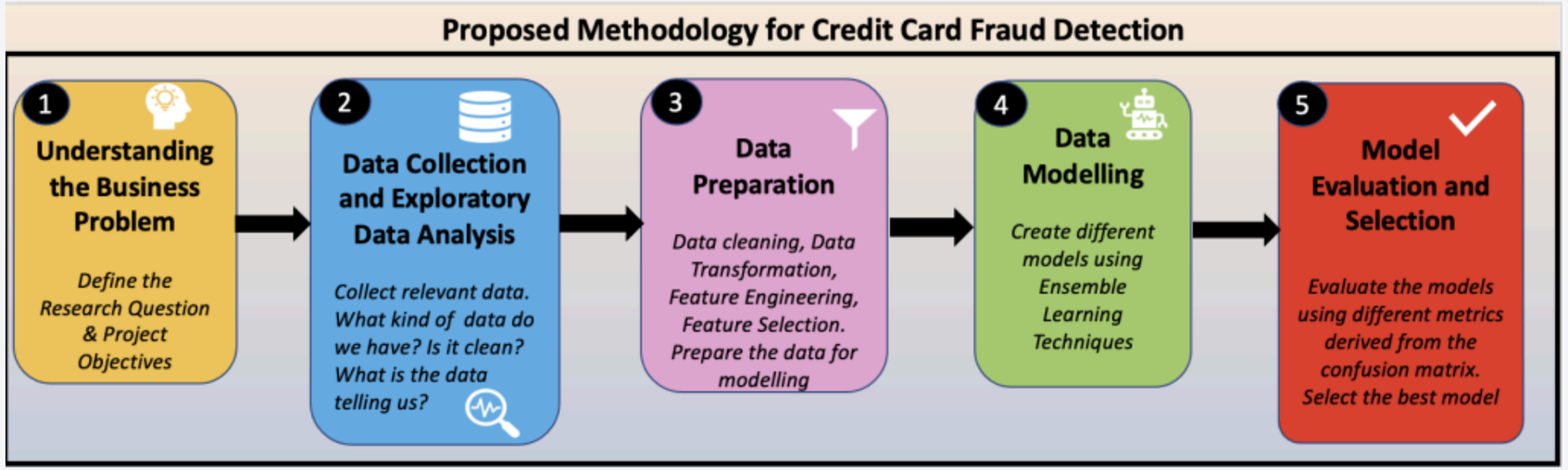
ABSTARCT

This seminar introduces an advanced credit card fraud detection system that combines supervised and unsupervised machine learning techniques to accurately identify anomalous transaction patterns, mitigating significant financial risks to users and the global financial system. The solution meticulously preprocesses and analyzes credit card transaction data, incorporating temporal patterns and distribution analysis, to distinguish fraudulent from legitimate transactions. A comparative analysis of XGBoost, Logistic Regression, Decision Tree, and Random Forest algorithms is conducted, targeting common fraud types such as card-not-present fraud (unauthorized transactions conducted over the phone), card-present fraud (transactions using cloned or stolen physical cards), and account takeover fraud (where fraudsters gain unauthorized access to an account to make transactions). These algorithms are evaluated based on accuracy, sensitivity,
specificity, F1-score, and ROC-AUC, aiming to enhance fraud detection and provide a more secure banking environment at scale.

PROPOSED ARCHITECTURE



PROPOSED ARCHITECTURE



BASIC TERMINOLOGY

- Fraudulent Transactions: Unauthorized or deceptive transactions intended to steal money or personal information.
- Legitimate Transactions: Authorized transactions performed by the rightful owner of the credit card.

TRANSACTION

Illegal use of a credit card to make unauthorized transactions or access an account.
TYPES:
Card Not Present(Online transaction)
Card present Fraud(Cloning of card)
Account Takeover fraud

FRAUD TYPES

XGBOOST

LOGISTIC REGRESSION

DECISION TREE

RANDOM FOREST

ML ALGORITHMS

Accuracy
Recall
Specificity
F1-Score
ROC-AUC

EVALUATION METRICS

A common issue in fraud detection where the number of legitimate transactions greatly exceeds the number of fraudulent transactions, which can skew model performance.

DATA IMBALANCE

UNDERSTANDING CONCEPTS

Data Preprocessing

- Cleanse data: Handle missing values, outliers, inconsistencies.
- Create features: Extract meaningful patterns (e.g., time, amount).
- Normalize data: Scale features for consistent range.

Class Imbalance

- Unbalanced data: Fewer fraudulent cases.
- Balancing techniques:
    - Over-sampling: Increase rare cases.
    - Under-sampling: Decrease common cases.
    - Synthetic data: Generate new minority class examples.

Machine Learning Algorithms

- Logistic Regression: Simple, interpretable, good baseline.
- Decision Tree: Understandable rules, prone to overfitting.
- Random Forest: Multiple trees for better accuracy, reduced overfitting.
- XGBoost: Powerful, often top performer, requires tuning.

Model Evaluation Metrics

- Beyond accuracy: Consider imbalanced data.
- Precision vs. recall: Focus on catching fraud (recall).
- Confusion matrix: Detailed performance breakdown.
- ROC-AUC curve: Overall model performance.

Cross-Validation

- Evaluate model performance: Split data into folds, train on most, test on one.
- Robust evaluation: Average performance across folds for reliable results.

Feature Importance

- Feature importance: Identify key factors driving predictions.
- SHAP values: Explain individual predictions based on features.

Hyperparameter Tuning

- Grid/Random search: Exhaustive/random search for best hyperparameters.
- Bayesian optimization: Intelligent search using probabilistic models.

LITERATURE SURVEY

Title	Author	Publication Date	Aim/Objective
Credit Card Fraud Detection Using Machine Learning	Deep Prajapati, Ankit Tripathi, Jeel Mehta, Kirtan Jhaveri, Vishakha Kelkar	04 December 2021	The aim is to evaluate and compare machine learning algorithms like Random Forest, XGBoost, and ANN for their effectiveness in accurately detecting and preventing fraudulent credit card transactions.
Credit Card Fraud Detection using Machine Learning	Naga Ashwini Nayak V J , C. Suchika , N Sandhya , M Lakshmi , Roja J	7, April 2023	The aim is to identify credit card fraud by employing machine learning algorithms like Decision Tree, Random Forest, and Extreme Gradient Boosting, and to evaluate their effectiveness using both public and real-world financial data. The goal is to assess the robustness and accuracy of these models in detecting fraudulent activities.
Credit Card Fraud Detection System based on Operational & Transaction features using SVM and Random Forest Classifiers	C. Sudha; D. Akila	25 February 2021	The aim is to develop a Credit Card Fraud Detection system utilizing Operational and Transaction features, applying Random Forest and Support Vector Machine classifiers. The system aims to classify user features into benign or suspected categories and evaluate its performance using precision, accuracy, recall, and F1-score metrics, demonstrating high detection rates and accuracy for both classifiers.

LITERATURE SURVEY

Title	Author	Publication Date	Aim/Objective
Credit Card Fraud Detection Using Machine Learning	Sanjay Bharadwaj	21 March 2024	The aim is to compare the effectiveness of two low-cost machine learning techniques, Random Forest and K-Nearest Neighbor (K-NN), in predicting credit card fraud using a highly unbalanced dataset. The research evaluates the models based on key machine learning metrics, ease of implementation, and cost-effectiveness to draw conclusions applicable to real-world scenarios.
Credit Card Fraud Detection Using Machine Learning	Neethu Tressa, V Asha, Govindaraj M, Sangamesh Padanoor, Rahila Tabassum, Desai Vatsal Dharmesh, Binju Saju	10 October 2023	The aim is to develop a system that uses machine learning techniques to detect fraudulent credit card transactions, ensuring customers are not charged for unauthorized purchases. The goal is to model historical transaction data, including fraudulent cases, and use this pattern to identify and prevent future fraudulent transactions.

Probable Application

01

SECURITY FOR E-COMMERCE

Deploy fraud detection algorithms in online payment gateways to protect against card-not-present fraud during online purchases, reducing the risk of fraudulent transactions.

02

BEHAVIOR ANALYTICS

Analyze customer transaction data to understand typical spending patterns, allowing the system to detect deviations that may indicate fraudulent behavior.

03

RISK ASSESSMENT

Use machine learning to assess transaction risk by analyzing historical data, customer behavior, and patterns, enabling financial institutions to make informed decisions.

04

ACCOUNT TAKEOVER

Use machine learning to detect account takeovers by monitoring unusual activities like multiple logins, behavioral changes, and irregular transaction patterns.

CONCLUSION

Credit card fraud will continue to be a significant challenge for the financial industry. This seminar will showcase the potential of machine learning to mitigate these risks. By leveraging algorithms like XGBoost, Logistic Regression, Decision Trees, and Random Forest, future fraud detection systems will be capable of more accurately identifying fraudulent transactions.

Data preprocessing and feature engineering will remain crucial for optimizing model performance. As fraudsters evolve their tactics, ongoing research and development will be essential to stay ahead. By integrating emerging technologies and refining existing models, we anticipate a future where credit card transactions are increasingly secure.

REFRENCES

i. G. Rushin, C. Stancil, M. Sun, S. Adams and P. Beling, "Credit card fraud—deep learning, logistic regression, and Gradient Boosted Tree,"2020 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, 2019, pp. 117-121.

ii. A.A.Taha, S.J.Malebary, An intelligent approach to credit card fraud detection using-An optimized light gradient boosting machine, IEEE Access 8(2020) 25579–25587, doi:10.1109/ACCESS.2020.2971354.

iii. Y. Lucas and J. Jurgovsky, ‘‘Credit card fraud detection using machine learning: A survey,’’ 2020 arXiv:2010.06479.

iv. A. Rb and S. K. Kr, ‘‘Credit card fraud detection using artificial neural network,’’ Global Transitions Proc., vol. 2, no. 1, pp. 35–41, Jun. 2021, doi: 10.1016/j.gltp.2021.01.006.

v. Emmanuel Ileberi1*, Yanxia Sun1 and Zenghui Wang2 A machine learning based credit card fraud detection using the GA algorithm for feature selection Proc., vol. 2, no. 1, pp. 35–41, 2022

THANK YOU