Corprise
Technologies

FraudDefenderAI

PROJECT

1st development stage for a credit card fraud detection AI system using ML (Classification)

- Introduction 01
- Business Understanding 02
- Data understanding 03
- Objectives 04
- Modeling & Evaluation 05
- Recommendations 06
- Next steps 07
- Thankyou 08

Introduction

In the dynamic landscape of digital finance, the increase in fraudulent credit card transactions poses a significant challenge for Mastercard, a leading global credit card provider in Europe. The Fraud Detection Department, led by Chief Risk Officer Sam Miller, faces the critical task of minimizing fraud losses while maintaining customer trust and operational efficiency.

Business Understanding

Stakeholder

At the helm of addressing this issue is the Fraud Detection Department at Mastercard, led by Chief Risk Officer (CRO) Sam Miller. The department shoulders the responsibility of minimizing financial losses from fraud while ensuring customer satisfaction through accurate identification and prevention of fraudulent credit card transactions.

Business Problem

Historically, the department has relied on rule-based systems to flag potentially fraudulent transactions. However, this approach has proven insufficient due to its high rate of false positives (legitimate transactions incorrectly flagged as fraudulent) and its inability to detect more sophisticated fraud tactics, resulting in false negatives (undetected fraud instances).

Data Understanding

Our analysis is based on a dataset containing credit card transactions from European cardholders over two days in September 2013. This dataset, obtained from Kaggle, comprises 284,807 transactions, of which 492 are identified as fraudulent.

PCA feature Transformations

Class Imbalance

Data Quality

Objectives

Minimize False Positives

Reducing the occurrence of legitimate transactions being incorrectly flagged as fraudulent is paramount. This objective aims to enhance customer satisfaction and minimize operational costs associated with addressing false fraud alerts

Maximize Fraud Detection (Minimize False Negatives)

Increasing the detection rate of actual fraudulent transactions is critical to prevent financial losses and safeguard the company's and customers' assets

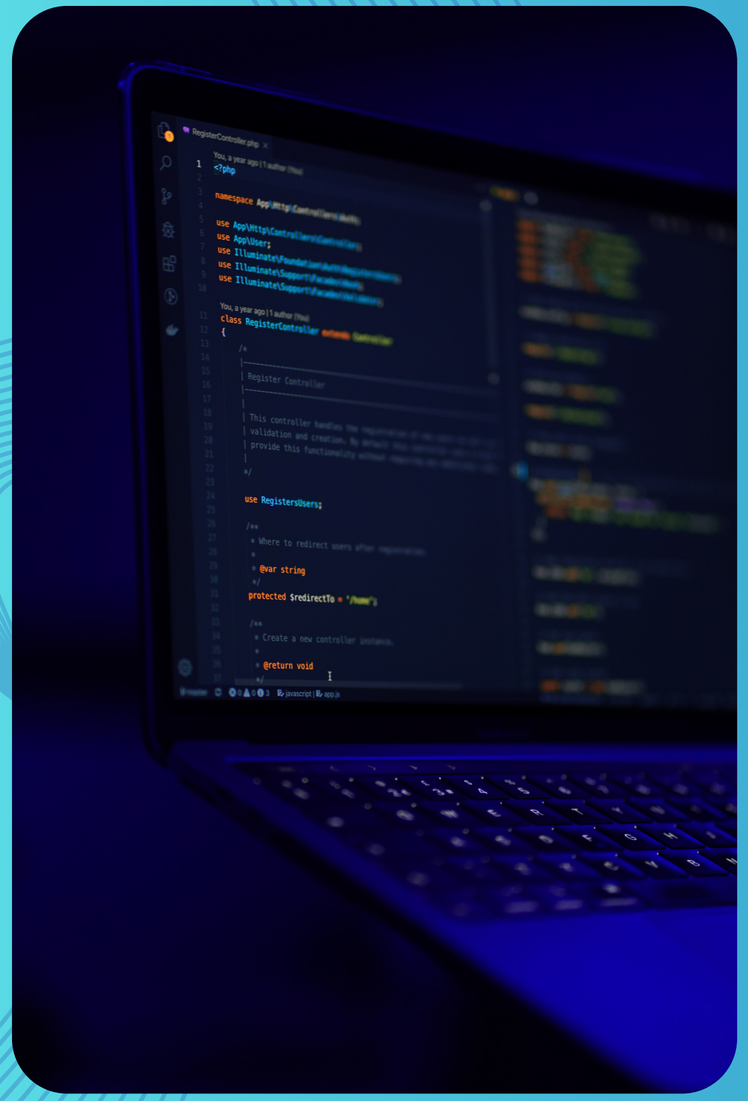
Achieve a Balance between Precision and Recall

Striking an optimal balance between precision (minimizing false positives) and recall (minimizing false negatives) within the highly imbalanced dataset is essential.

Modeling

In our quest to enhance fraud detection capabilities at Mastercard, we ventured into the realm of machine learning models. These models are like detectives, learning from historical credit card transactions to distinguish between legitimate and fraudulent activities.

Each model brought unique strengths and weaknesses to the table. While some excelled in precision, others shone in recall. Balancing these traits is crucial to strike a fine equilibrium between detecting fraudulent activities and minimizing false alarms



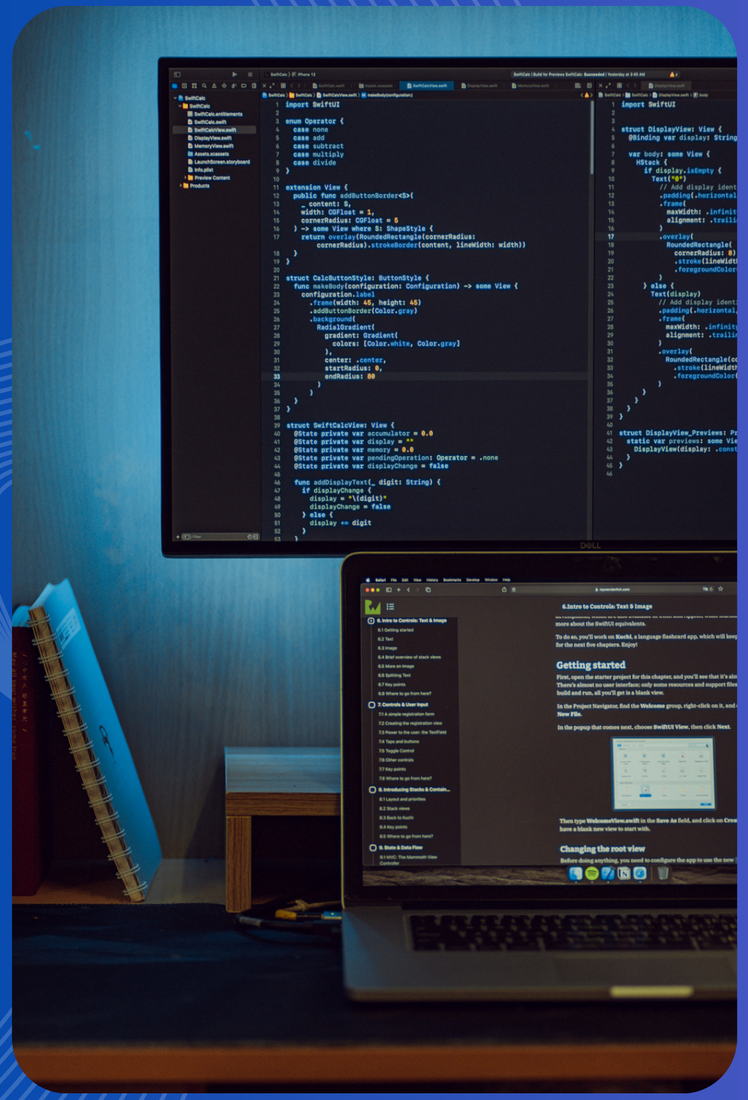
Baseline Model

Logistic Regression Classifier

AUPRC 79.5%

- Accuracy: 97.32%
- Precision: 5.23%
- Recall: 92.65%
- F1 Score: 9.91%

- Logistic Regression showcased a keen eye for spotting fraudulent transactions. However, while it excelled in detecting fraud, it occasionally mistook legitimate transactions as fraudulent, making it prone to false alarms.



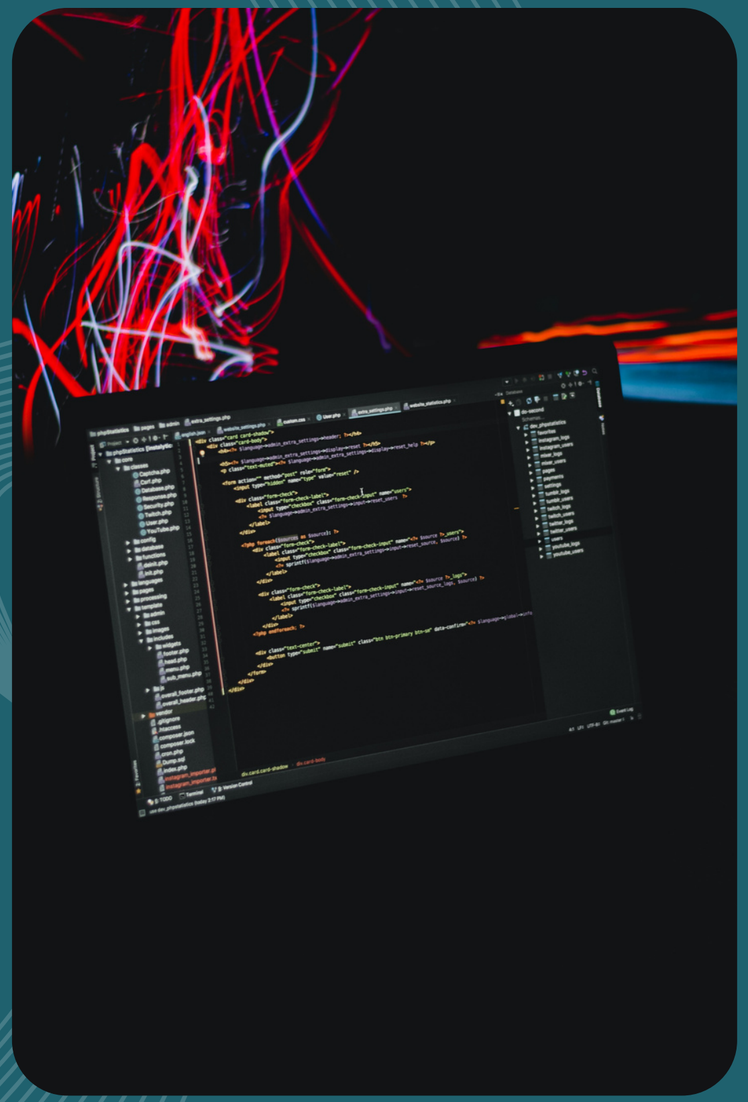
Advanced Model 1

Random Forest Classifier

AUPRC 88.35%

- Accuracy: 99.95%
- Precision: 83.22%
- Recall: 87.50%
- F1 Score: 85.30%

Random Forest emerged as a robust choice, demonstrating a balance between precision and recall. It identified fraudulent transactions with commendable accuracy, maintaining a low false positive rate while capturing a significant portion of actual frauds.



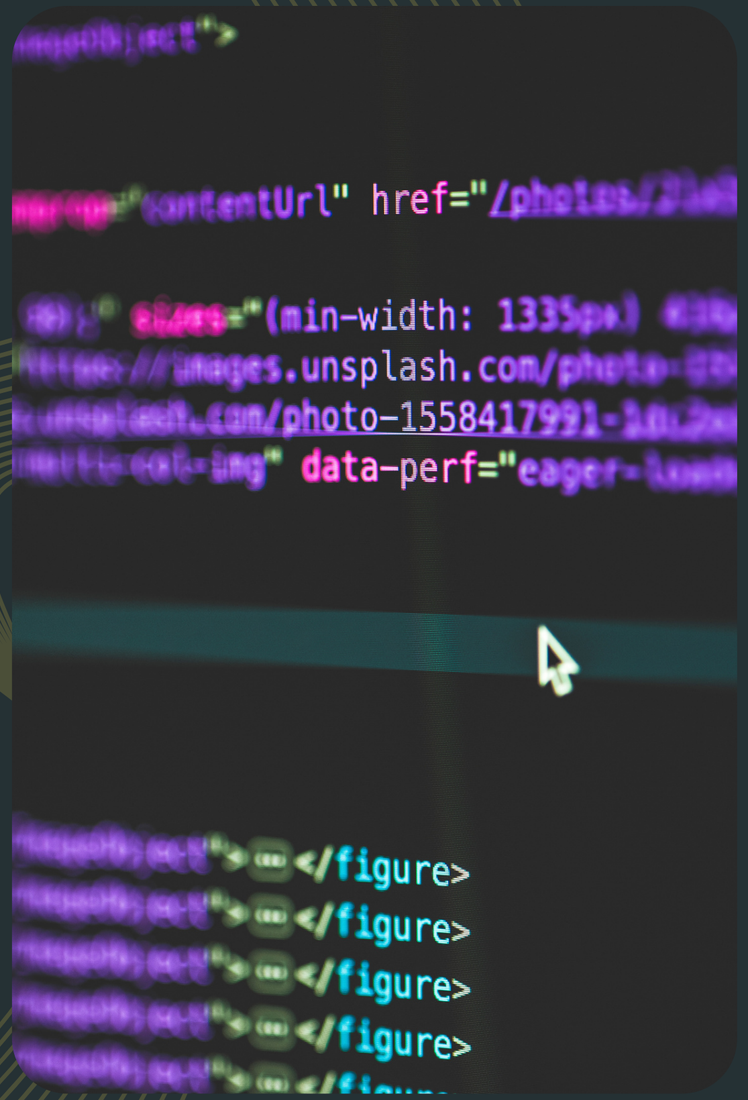
Advanced Model 2

Gradient Boosting Classifier

AUPRC 78.14%

- Accuracy: 99.35%
- Precision: 18.56%
- Recall: 91.18%
- F1 Score: 30.85%

Although skilled at detecting fraudulent transactions, Gradient Boosting exhibited a higher false positive rate, potentially raising unnecessary alerts for legitimate transactions.



Finetuned Model (Iterated)

Logistic Regression Classifier

AUPRC 79.5%

- **Accuracy**: 97.00%
- **Precision**: 5.00%
- **Recall**: 93.00%
- **F1 Score**: 10.00%
- **AUPRC**: 79.37%
-

The tuned Logistic Regression improved its recall while maintaining a similar precision. While it continued to excel in identifying fraudulent transactions, it still struggled with false alarms for legitimate ones.

Top Metrics

88%

AUPRC Random forest

85%

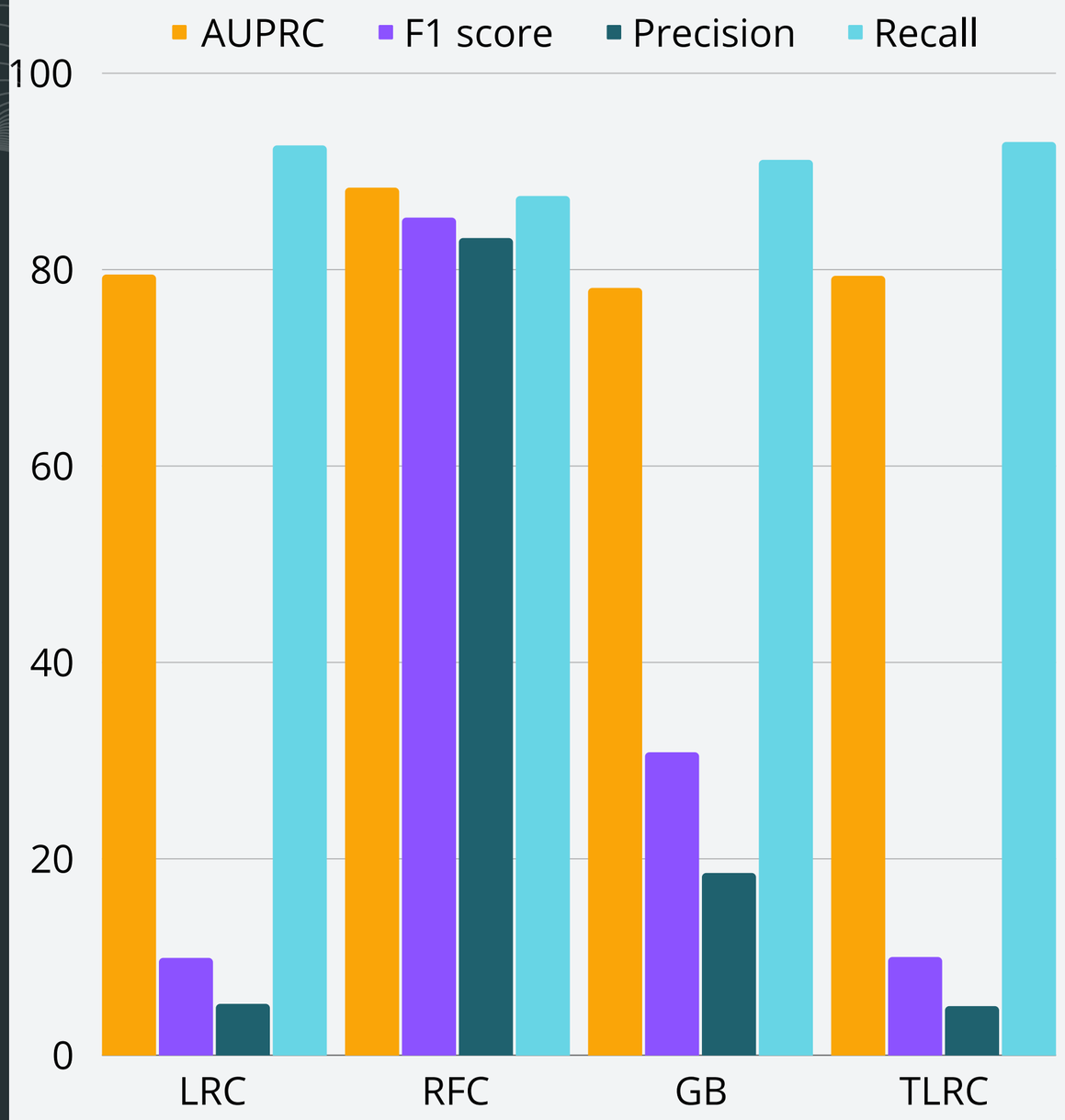
F1 Score Random Forest

83%

Precision Random Forest

93%

Recall Tuned Logistic Regression



Recommendations

At the helm of addressing this issue is the Fraud Detection Department at Mastercard, led by Chief Risk Officer (CRO) Sam Miller. The departm1. Implement a Tiered Fraud Detection Approach:

- Utilize Logistic Regression as Initial Filter:
  - Leverage the high recall of the Logistic Regression model to capture a majority of potential fraud cases in the first pass.
- Deploy Random Forest as a Secondary Layer:
  - Apply Random Forest to achieve a balance between precision and recall, ensuring accurate identification of fraudulent transactions while minimizing false positives.

2. Explore Advanced Techniques:

- Unsupervised Learning Methods:
  - Investigate anomaly detection techniques such as Isolation Forests or Autoencoders to identify new patterns of fraudulent behavior that may not have been captured by existing models.
- Consider Deep Learning Architectures:
  - Explore neural network models like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) to uncover intricate patterns within transactional data for enhanced fraud detection accuracy.

Next steps

1. Continuous Monitoring and Model Updating:
   - Regular Model Evaluation:
     - Periodically assess model performance against new data to ensure adaptability to evolving fraud tactics.
   - Implement Automated Updates:
     - Develop mechanisms for automatic model updates based on emerging fraud patterns, ensuring real-time protection.
2. Collaboration and Knowledge Sharing:
   - Cross-Department Collaboration:
     - Foster collaboration between Fraud Detection, Data Science, and Security teams to share insights and strategies for improved fraud prevention.
   - Industry Knowledge Exchange:
     - Engage in industry conferences or forums to exchange best practices and stay updated on the latest fraud trends and prevention methodologies.
3. Customer-Centric Approach:
   - Balancing Security and User Experience:
     - Maintain a balance between robust fraud detection and seamless customer experience by fine-tuning models to minimize false positives that may inconvenience customers.
   - Customer Education and Communication:
     - Develop clear communication strategies to inform customers about fraud prevention measures without causing alarm or inconvenience.

Corprise
Technologies

Thank You

Contact Us

011-345-9918

mwendajames341@gmail

www.mwendajames.com