



Credit Card Fraud

Did you know that about four percent of the global GDP is lost to fraud annually, with digital payment fraud increasing by 25% over the last three years.

F by Francisco Caupers



What Is Payment Fraud?

Unauthorized Transactions

Payment fraud involves unauthorized transactions using false or stolen payment information. This type of fraud is prevalent across various payment methods and often occurs due to data breaches or exposure on the dark web.





Prevalence of Credit Card Fraud in the U.S.

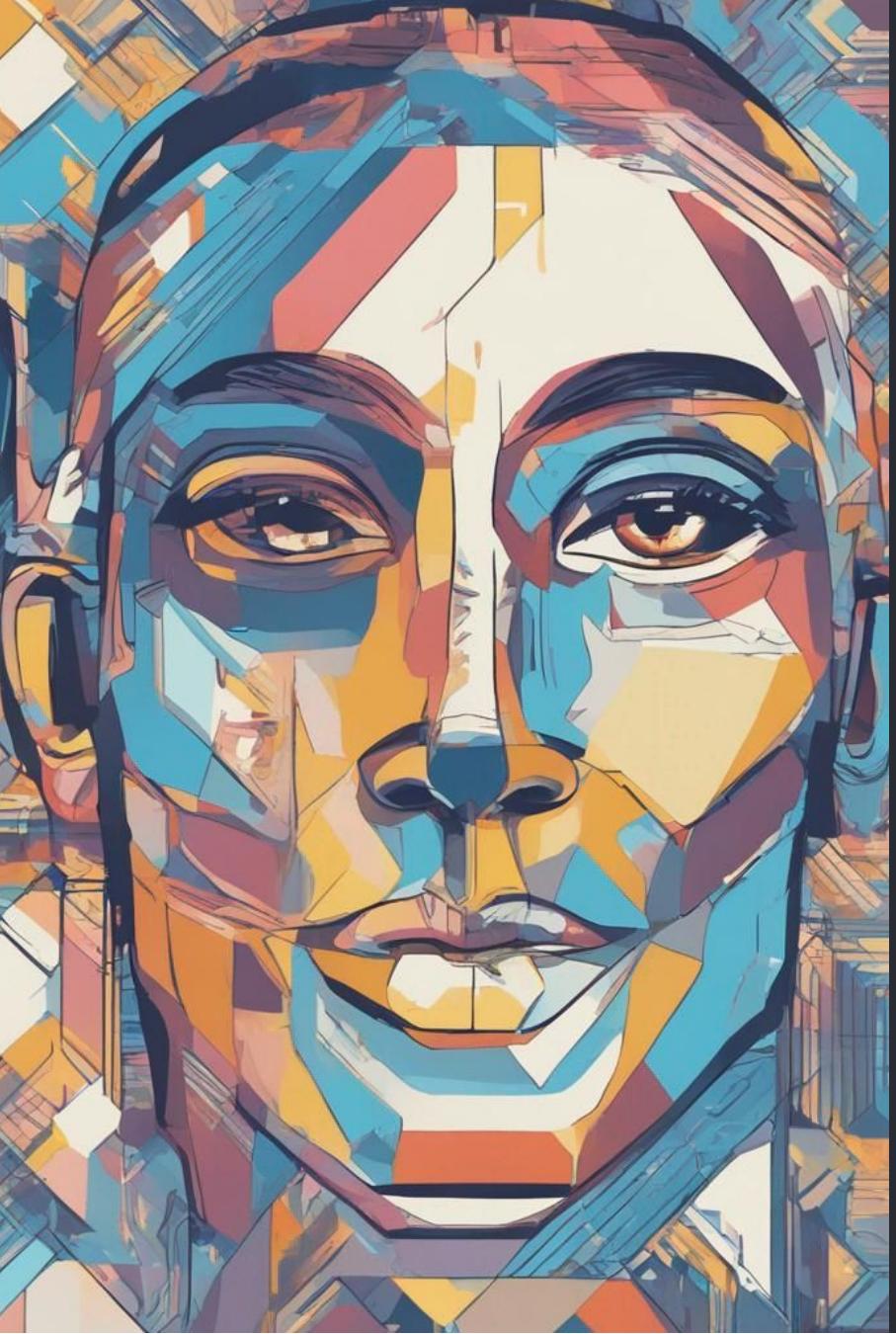
1 38.6% of Global Fraud Losses

The U.S. leads globally in credit card fraud cases, constituting 38.6% of all reported card fraud losses in 2018, totaling \$9.47 billion.

2 Exploitation During Pandemics

Fraudulent transactions increased by 35% in April 2020 during the pandemic, highlighting how fraudsters capitalize on such crises.





Affordability of Stolen Credit Card Data

Low Cost on the Dark Web

Credit and debit card information, along with additional personal data, sells for as low as \$5 on the dark web.





4 Types of Payment Fraud



Chargeback Fraud

Chargeback fraud, also known as friendly fraud, is when a cardholder makes a valid online purchase and then, even though there was no actual problem with the transaction, they fraudulently ask their bank for a refund.



Stolen Card Fraud

The use of unlawfully obtained card information for fraudulent transactions is known as stolen card fraud, and it costs consumers and institutions money.



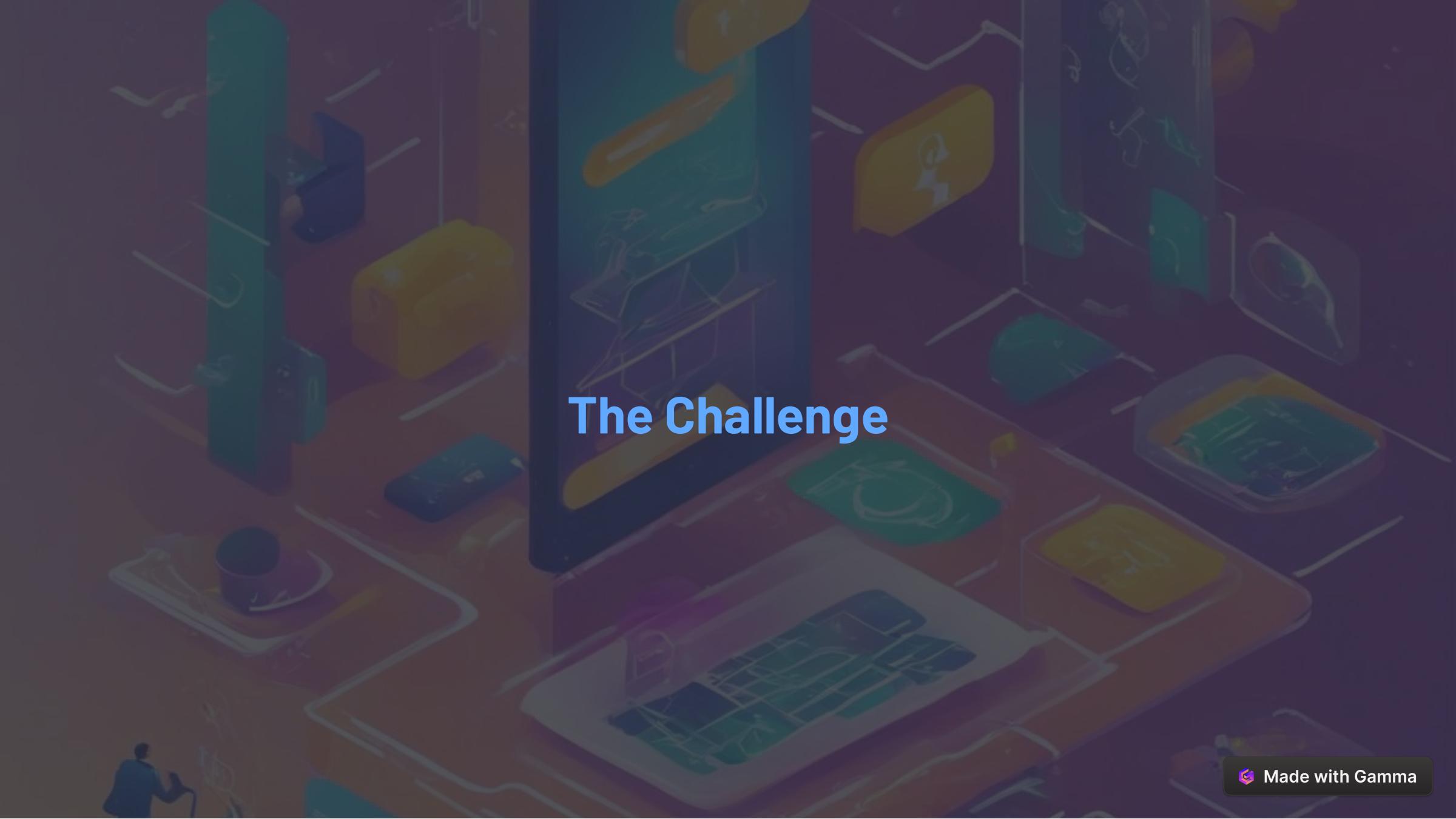
Identity Theft

Identity theft, the most common payment fraud, involves the misuse of personal information for fraudulent transactions, causing financial losses and credit damage to the victim.



Triangulation Fraud

A con artist creates a fake online store, buys items from a legitimate retailer using stolen credit card info, and stores the payment details for future fraud, leaving the customer unaware of the scam.



The Challenge

Tackling the Challenge

Values

Our Dataset has around 285,000 non-fraudulent entries, and just shy of 500 instances of fraud.

Attributes Selected

We've selected 'class,' 'amount,' and 'time' columns for analysis, each playing a crucial role in our study of the data.

Analytical Approach

To address the significant class imbalance, we will implement a combination of resampling techniques and tailored algorithms.





Machine Learning Model

1

Logistic Regression

OverSampling, UnderSampling, changing the Threshold and Featured Engineering

2

Random Forest

Random Forest is a powerful ensemble learning algorithm that combines multiple decision trees for enhanced predictive accuracy and robustness against overfitting.

3

Decision Tree

A decision tree is a popular machine learning algorithm used for classification and regression tasks. It represents decisions as a tree-like structure, where each node represents a decision, and each branch represents an outcome.

4

LSTM Neural Network

LSTM, or Long Short-Term Memory, is a specialized recurrent neural network ideal for sequential data analysis with a strong focus on capturing long-range dependencies.



Model Selecting

Random Forest

After experimenting with a number of different models that we finally decided to use Random Forest . Due to our efforts, we were able to get 473 positive detections out of 498.



Before

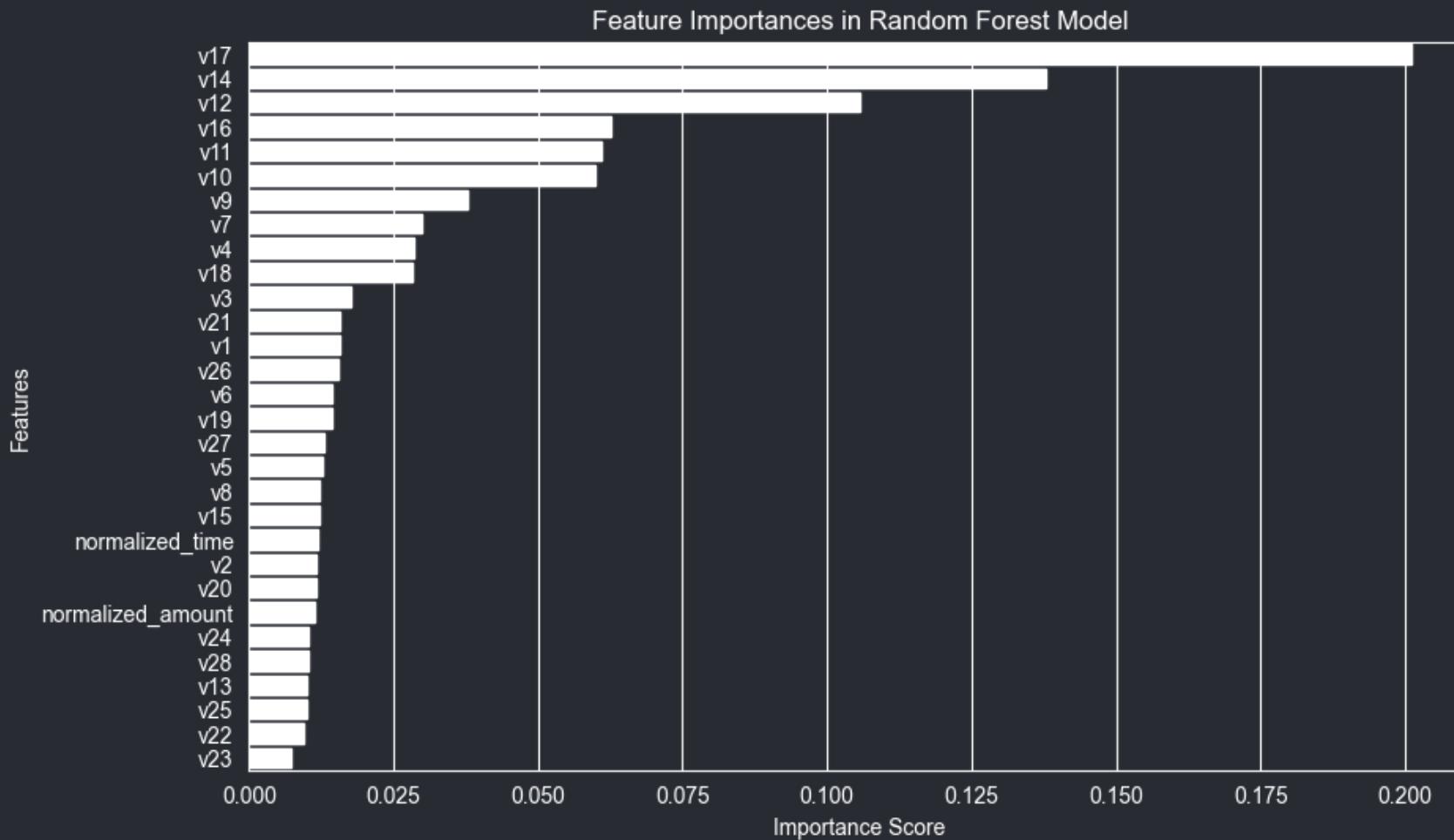


During



After

Feature Importances



Conclusion

IRON
HACK