

Fraud Detection in Transactions Using Machine Learning

Tackling Fraud in a Real Banking Scenario

Presented by Santiago

07. March 2025

The Challenge

Fraud: A Growing Problem 🕵️

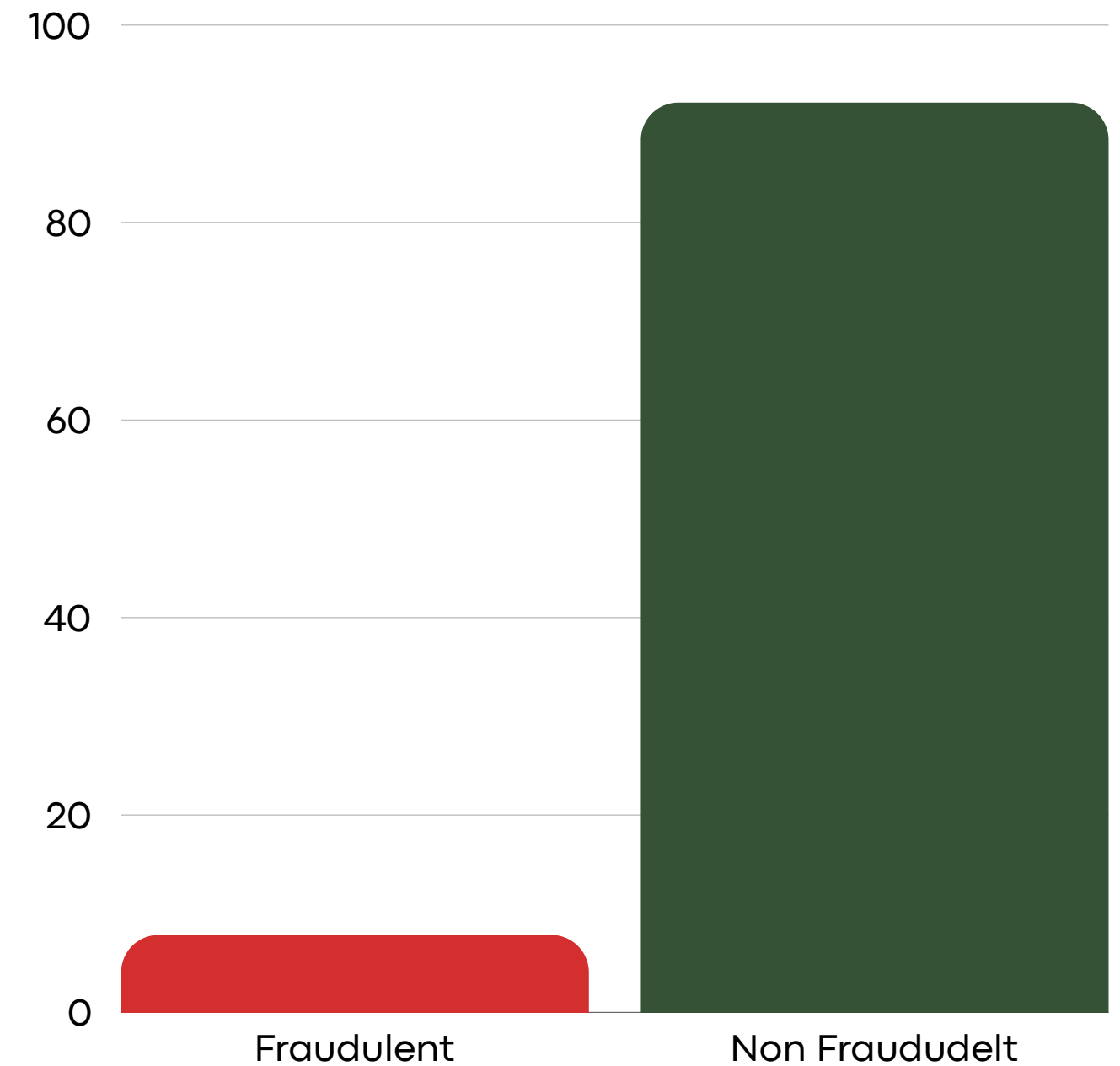
- Fraud in banks is hard to spot and causes big losses
- €4.3 billion lost to fraud in Europe in 2022

This Company's Challenge 📊

- 7,85% of transactions are fraudulent, a big risk

Why It Matters 💡

- Banks lose money 💸
- Customers lose trust 😞
- Systems struggle to keep up ⚙️



The Fraud Detection Solution

A Product to Catch Fraud with Machine Learning

- Designed to spot fraudulent transactions in real time
- Easy for banks to use
- Keeps customers safe

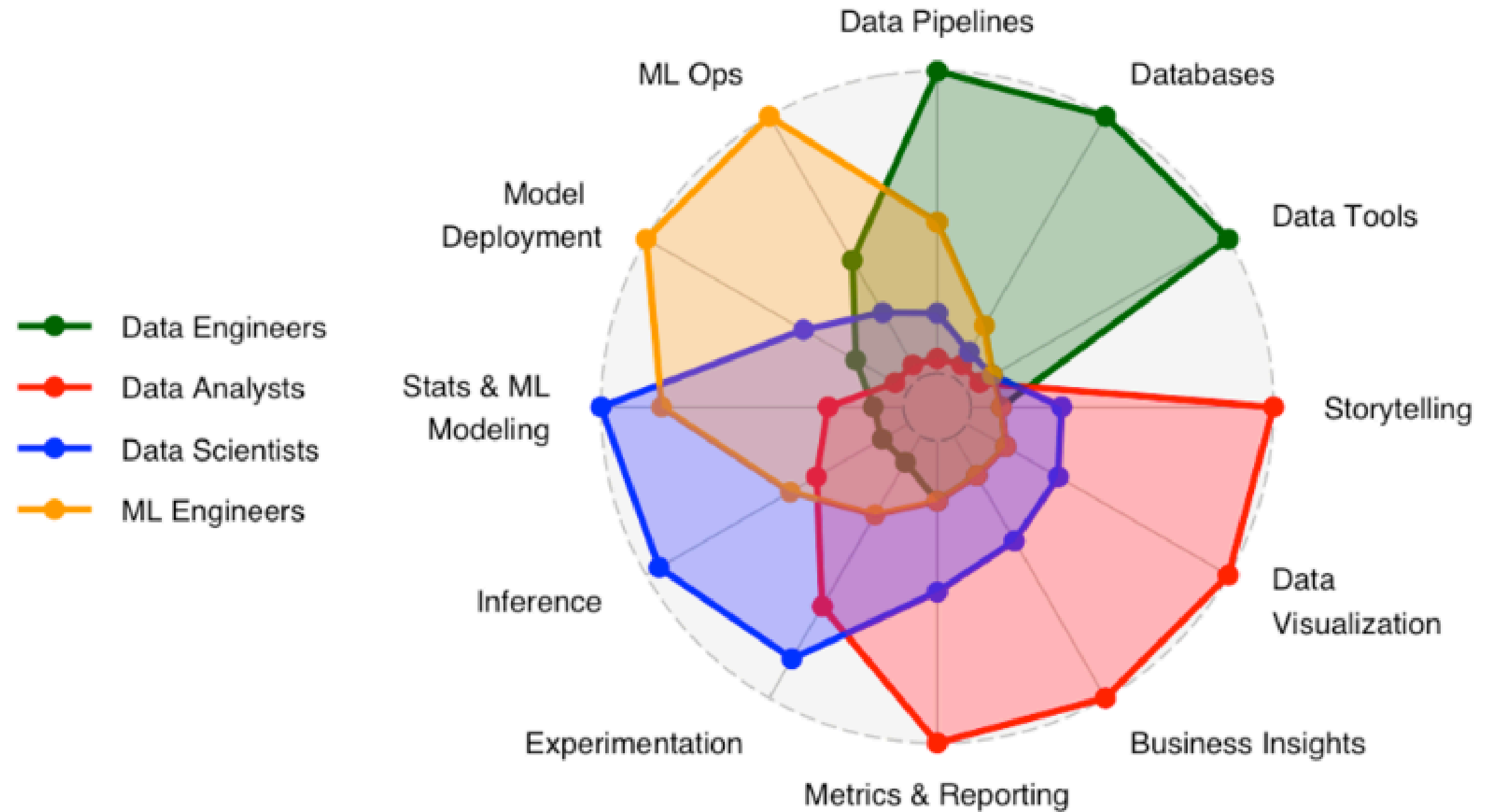
Approach

- Used smart technology to analyze transaction data
- Built a system to clean data and flag risks

Benefits for the Company

- Saves time for account managers
- Protects transactions, builds customer trust

Data Skills across the Project



Understanding and Preparing the Data

The Data Used

- Over 144,000 transactions analyzed
- 434 pieces of information per transaction
- Key details: transaction amounts, card information, email addresses, devices used

Preparation Process

- Data organized with a visual dashboard
- Categories like payments and devices sorted
- Most important details kept, new ways to spot fraud patterns added
- Missing information handled to keep data useful
- Data made ready for a smart fraud detection system



Machine Learning Behind the Scenes

A Supervised Learning Model 🤖

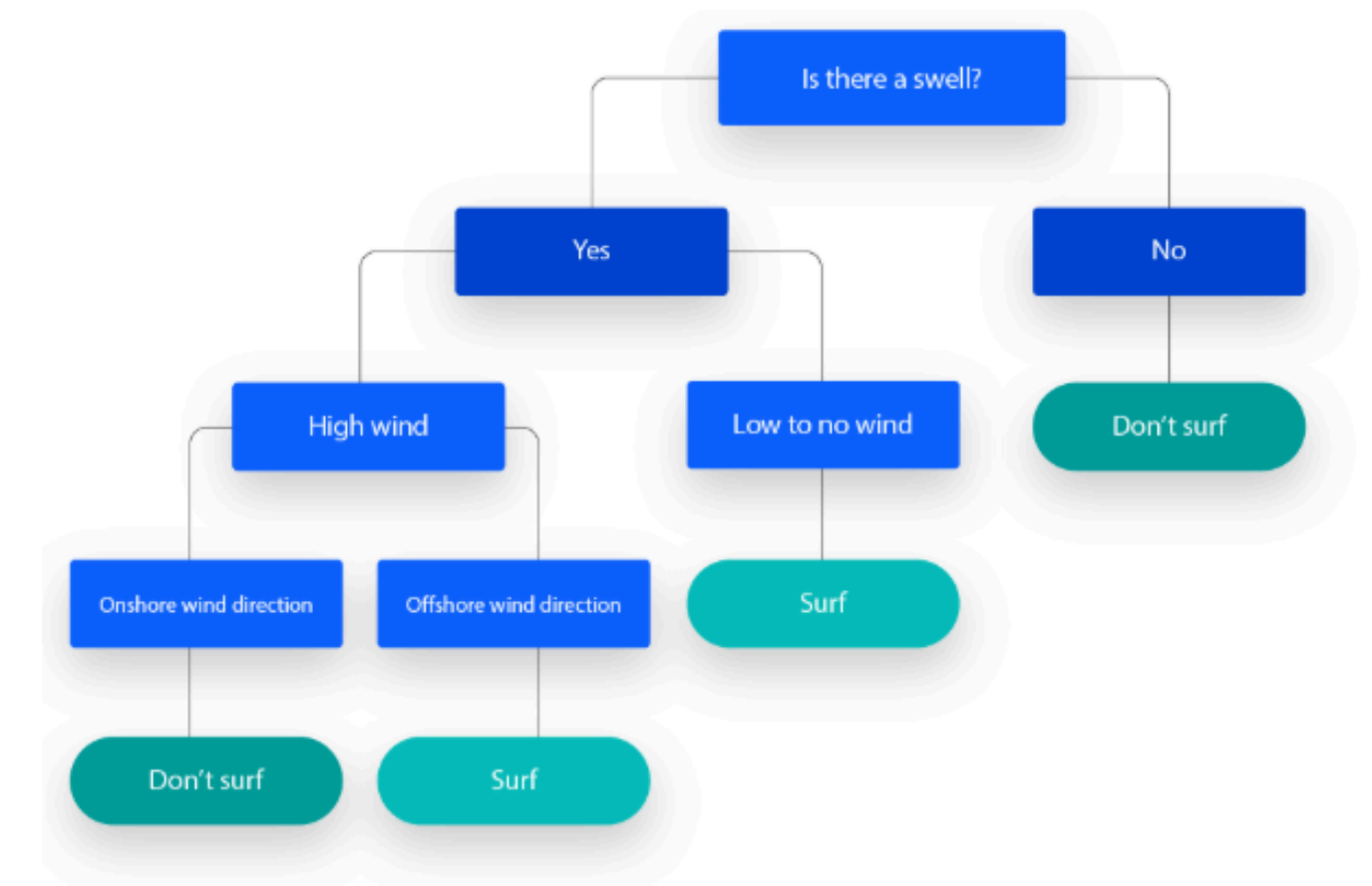
- Learns from past data with known fraud cases to spot new ones

Training and Testing 🌳

- Trained on examples like a student studying, then tested to check it works
- Uses transaction details to decide what's risky

LightGBM: Predicting the Future ⚡

- Combines many decision trees to predict fraud from past patterns
- Chosen for speed and power with lots of data



Making the Model Smarter

Model:

LightGBM using
Cross
Validation

Fine-Tuning for Better Results 🛠️

- The system was carefully adjusted to catch more fraud

Focus on Catching Fraud 🔍

- Prioritized finding as many fraud cases as possible
- Balanced speed and accuracy for real-time use

Tested for Reliability ✅

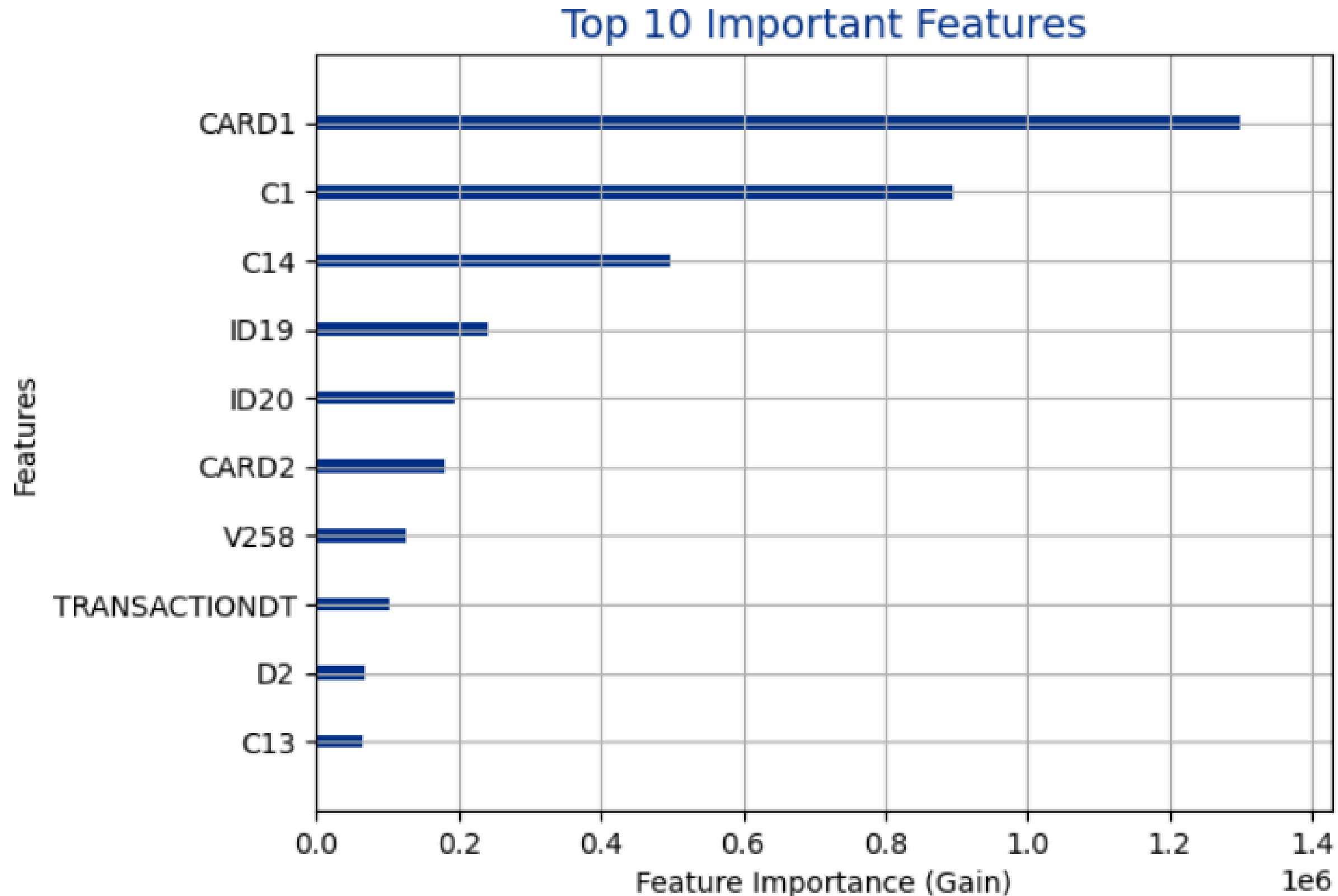
- Checked with past data to ensure it works well
- Ready to protect transactions effectively

Key Findings & Insights



- Recall: 93.6% of frauds detected
- Precision: 43.7% of predicted frauds are true
- False Positive Rate (FPR): 10.27% false alarms

Key Findings & Insights



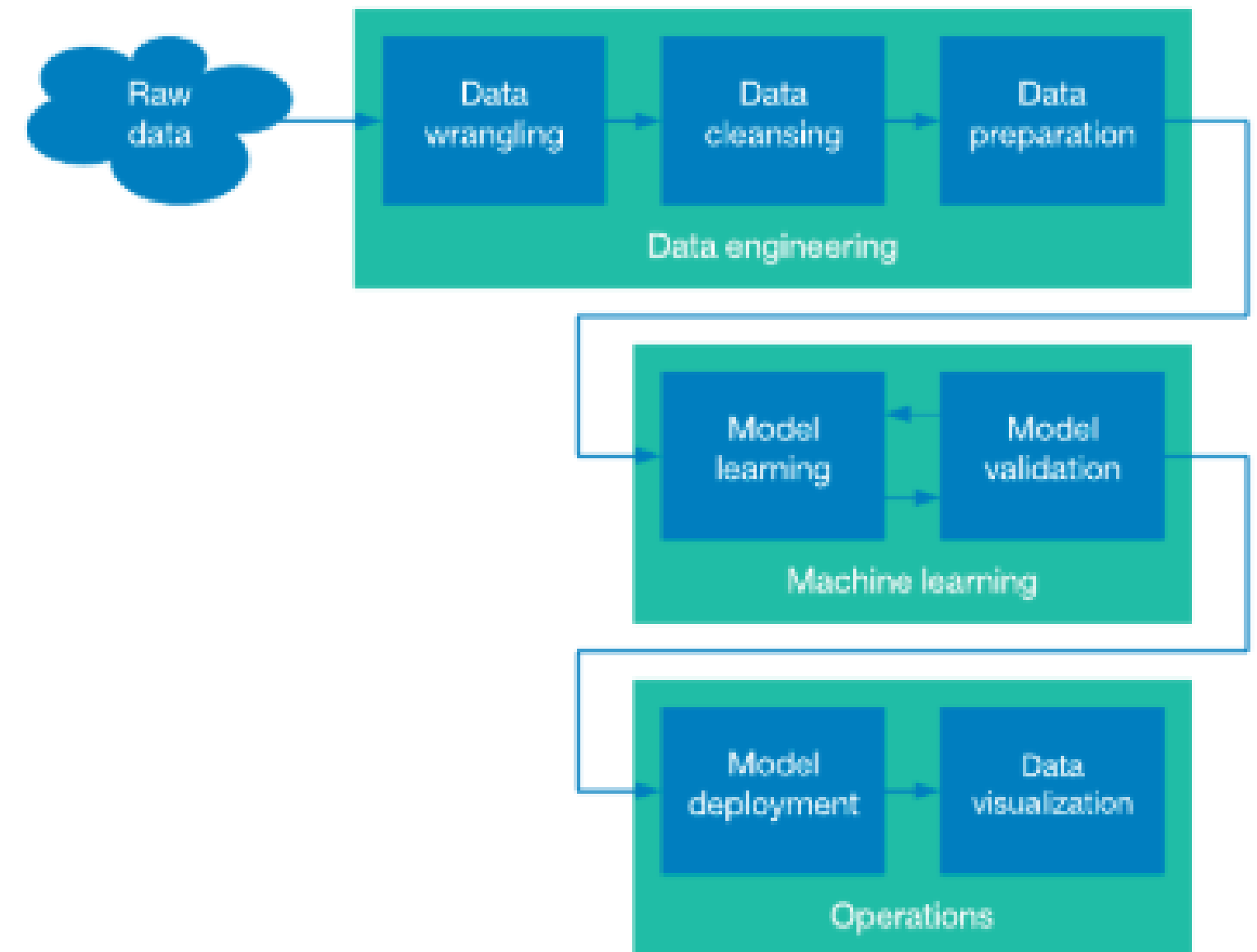
Features:

- TransactionDT: Time delta attributes
- card1 - card2: Card details
- C1-C13-C14: Counting metrics
- id 19 - id 20: Identity attributes
- D2: Transaction intervals
- V258: Engineered features




How the Pipeline Works

An Automatic Process to Spot Fraud 🚀

- A fast, automated system for account managers:
 - Account manager uploads transaction and identity data about a client
 - Data is automatically cleaned and organized to focus on key details
 - Extra clues are added automatically to spot fraud patterns
 - A smart system (LightGBM) analyzes the data to predict fraud
 - Results are delivered instantly, showing which transactions might be fraudulent



Conclusions

- **Effective Fraud Detection Achieved** 
 - Recall of 93.6% detects most frauds
 - False Positive Rate of 10.27% ensure practical use
- **Banking Application Ready** 
 - App integrates model for real-time fraud prediction
 - Key features
- **Future Improvements Identified** 
 - Explore ensemble methods to enhance precision further
 - Model maintenance recommended to adapt to evolving fraud patterns
 - Further feature engineering suggested to boost performance

Thanks

Presented by Santiago

07. March 2025