

CREDIT CARD FRAUD DETECTION AND CUSTOMER SEGMENTATION

PRESENTATION

Submitted on
14th October, 2025

Presented by
CYBERNETIC CRUSADERS

MEMBERS

MEMBERS



SONALI MAHATO

B.Tech (CSE – Data Science and AI)
ARKA **JAIN** UNIVERSITY



NIKITA MISHRA

B.Tech (CSE – Data Science and AI)
ARKA **JAIN** UNIVERSITY



SOHAIL KHAN

B.Tech (CSE – Data Science and AI)
ARKA **JAIN** UNIVERSITY

PROJECT BRIEF

This project involves two data-driven tasks using IBM SPSS Modeler:

Credit Card Fraud Detection – to identify fraudulent transactions using machine learning models.

Customer Segmentation – to group customers based on their banking behavior using clustering.

By combining classification and clustering in a single project, it demonstrates both supervised and unsupervised learning applications in financial data analysis.



INTRODUCTION

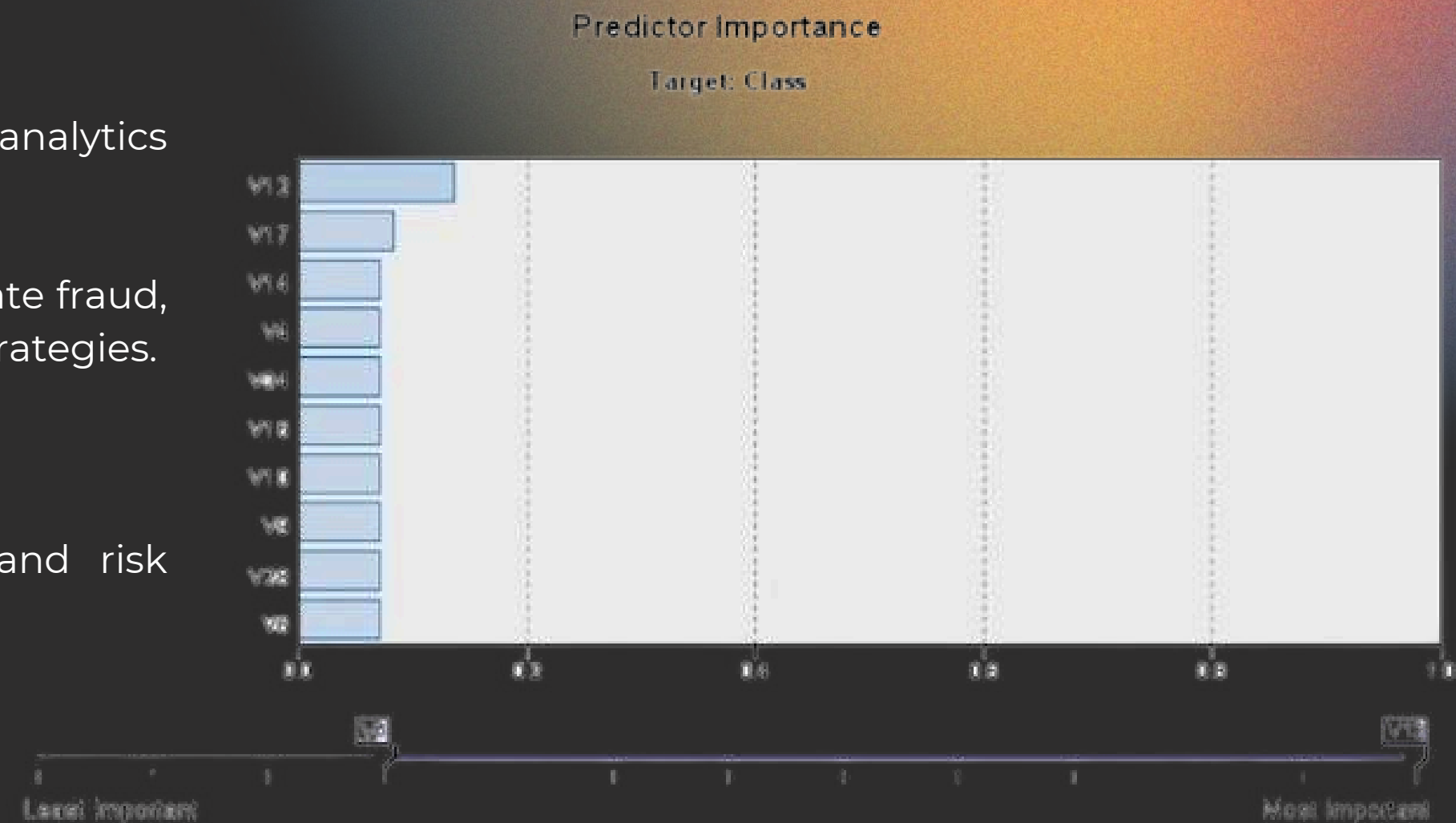
INTRODUCTION

With the increase in online transactions, fraud detection and customer analytics have become critical for banks.

Machine Learning models can identify complex, hidden patterns that indicate fraud, while clustering helps in understanding customer behavior and retention strategies.

This project uses IBM SPSS Modeler to:

- Detect anomalies that could indicate fraud.
- Segment customers into distinct groups for targeted marketing and risk assessment.



FEASIBILITY STUDY

Technical

IBM SPSS Modeler supports classification and clustering through visual drag-and-drop modeling.

Operational

Both datasets are publicly available and easy to process.

Economic

The project uses free datasets and educational access to SPSS Modeler, ensuring cost-effectiveness.

Time Feasibility

The complete workflow can be executed and tested within 1–2 weeks.

PROJECT DETAILS

PROJECT DETAILS

1

PART 1

Credit Card Fraud Detection

2

PART 2

Customer Segmentation

DATASET DETAILS

1

DATASET NAME
creditcard.csv

2

ROWS
284,807

3

COLUMNS
31

FEATURES

- **V1-V28** : PCA-transformed components representing anonymized transaction features.
- **Time** : Time elapsed (in seconds) between each transaction.
- **Amount** : Transaction amount.
- **Class** : Target variable (0 = Non-Fraud, 1 = Fraud)

PART 1: CREDIT CARD FRAUD DETECTION

Logistic Regression

High precision, interpretable results.
Good for baseline.



C5.0 Decision Tree

Easy rule interpretation.
Visual insights.



Random Forest

Highest accuracy & recall.
Robust to imbalance



MODEL

MODEL RESULT

DATASET DETAILS

1

DATASET NAME
bankchurners.csv

2

ROWS
10,127

3

COLUMNS
23

CLUSTER INSIGHTS

- **Cluster 1** : Young customers with low credit limit and moderate usage.
- **Cluster 2** : Middle-aged customers with high income and high transactions.
- **Cluster 3** : Older customers with low activity or nearing attrition.

PART 2: CUSTOMER SEGMENTATION

CONCLUSION

The project successfully demonstrates both supervised and unsupervised learning in IBM SPSS Modeler.

Key Outcomes:

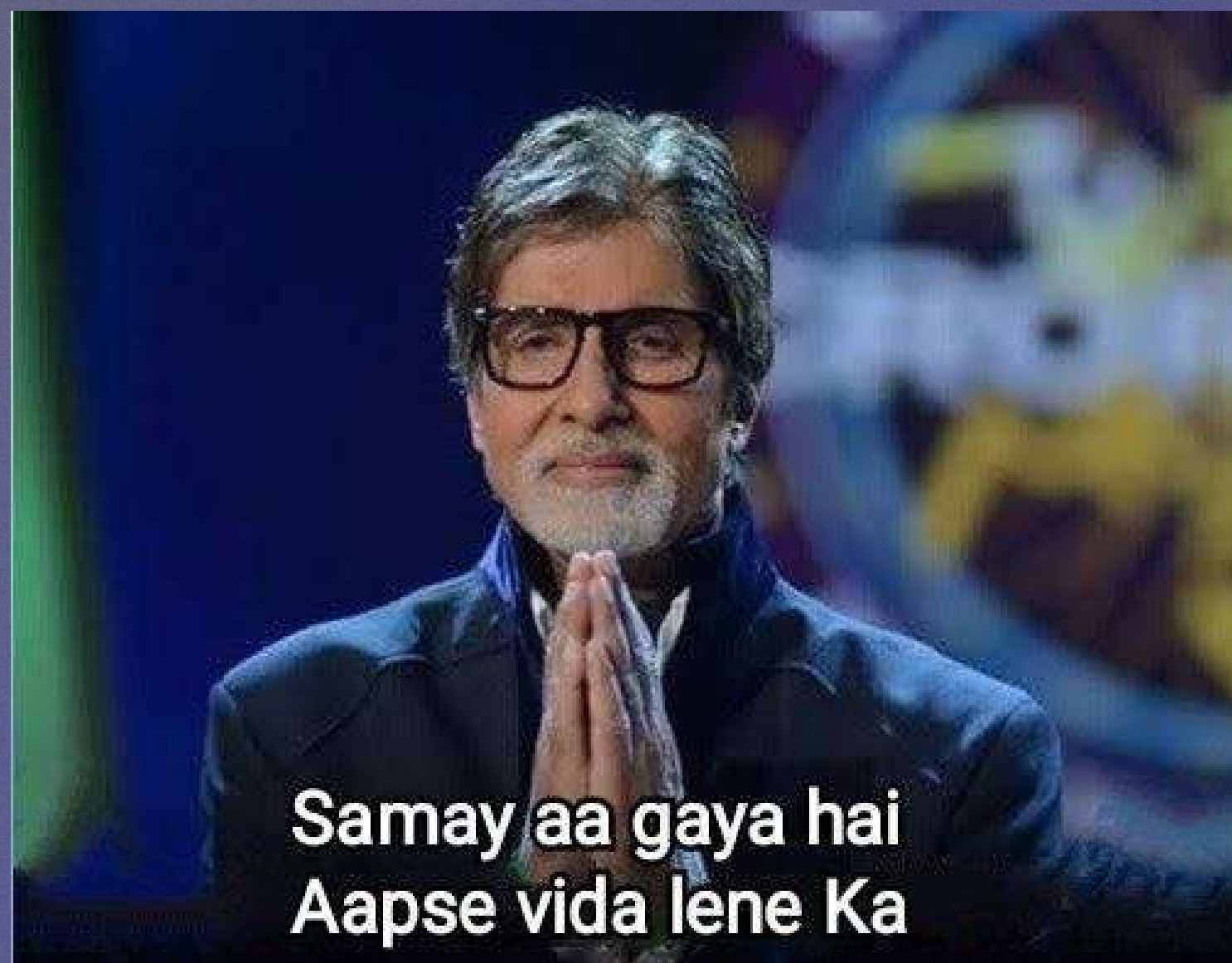
- Fraud detection models (Logistic, C5.0, Random Forest) achieved strong performance on PCA-transformed data.
- Random Forest showed the best overall accuracy.
- Customer segmentation using K-Means revealed meaningful behavioral patterns among clients.
- PCA features ensured privacy while maintaining accuracy.

**YOU HAVE ANY
QUESTIONS?**



**IF YOU HAVE ANY QUESTION IT MEANT
THAT YOU WERE NOT LISTENING SO WE ARE
NOT GOING TO ANSWER ANY QUESTIONS**

THANKS



Presented by
CYBERNETIC CRUSADERS