

CREDIT CARD FRAUD DETECTION

Present by Soumita Patra





INTRODUCTION

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

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

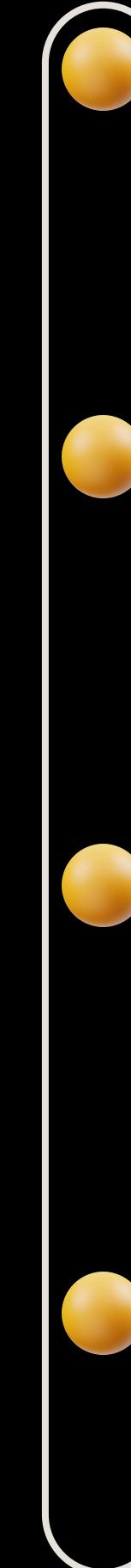
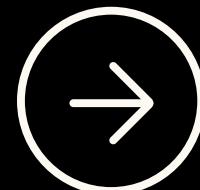
Customer Segmentation – to group customers based on their banking behavior using clustering.

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





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 costeffectiveness.

Time Feasibility

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

DATASET DETAILS



DATASET NAME

creditcard.csv

ROWS

284,807

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)



MODEL RESULT



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

THANK YOU

for your time and attention

