Credit Card Default Risk Analysis Report

1. Project Overview

This project uses the UCI Credit Card Default dataset to predict whether a customer will default next month. Fr

2. Data Preprocessing

The dataset has 30,000 samples and 25 features. No missing data was found. Class imbalance was addresse

3. Modeling and Evaluation

We used logistic regression.
Original performance:
- Accuracy: 0.81
- Recall (default): 0.25
- AUC: 0.71

After SMOTE:
- Accuracy: 0.66
- Recall (default): 0.60
- AUC: 0.69

The recall improvement highlights the model's stronger ability to identify potential defaulters.

4. Skills Demonstrated

- Logistic regression
- Handling imbalanced data (SMOTE)
- Metrics: Confusion Matrix, Precision, Recall, F1-score, ROC AUC
- Practical use of sklearn
- Business interpretation for risk mitigation