



National University
of computer and emerging sciences

Data Science Tools and Technologies

Semester Project

Title

Predicting Credit Card Default Risk Using Machine Learning: A Comparative Analysis of Classification Model

Supervised By
Dr. Kashif Zafar

Members

| | |
|----------------|----------|
| Aousaf Sulaman | 25L-8013 |
| Ali Hussan | 25L-8007 |
| Sana Imtiaz | 25L-8037 |

Overview:

This project focuses on analyzing the **Default of Credit Card Clients dataset**, sourced from Kaggle and the UCI Machine Learning Repository. The objective is to build a classification model that predicts whether a credit card client will default on their next payment. By leveraging past payment records, demographic details, and financial behavior, the project applies machine learning techniques to classify clients into **default** or **non-default** categories. This predictive modelling not only helps in understanding risk factors but also provides insights useful for financial institutions in credit risk management and decision-making.

Problem Statement:

Credit card defaults pose a major financial risk to banks and institutions. Identifying potential defaulters in advance enables better decision-making and risk management.

Expected Outcomes:

- A trained classification model to predict client defaults.
- Performance comparison across multiple ML algorithms.
- Insights into key factors influencing defaults.

Tools & Technologies

Python, Pandas, Scikit-learn, Matplotlib/Seaborn, Jupyter Notebook, WEKA

Data Sources:

- Kaggle
- UC Irvine Machine Learning Repository
- Hugging Face

Steps:

- Data Visualisation using Weka
- Data Preprocessing
- Feature Selection
- Splitting Dataset into Training and Testing Sets
- Model Training
- Evaluating Model Performance by different evaluation metrics
- Visualisation
- Story Telling

The process will continue in a tight loop, with different models, feature selection and Pre Processing to get the best accuracy of the model

Project Objective:

To predict credit card client defaults using classification models for improved credit risk assessment.