**Abiola Ijaola**

**Predicting Default of Credit Card Clients**

**Project Description**

This project aims to use 23 customers case default payments in Taiwan to predict fraud. The problem statement is based on default payments and compares the predictive accuracy of probability of default among data mining methods.

**Methodology**

The dataset is a binarily classed multivariate data with 23 features and default payment (Yes = 1, No = 0), as the response variable. From the perspective of risk management, the result of predictive accuracy of the estimated probability of default will be more valuable than the binary result of classification.

**Target Question**

Compares the predictive accuracy of probability of default among data mining methods.

**Source of Dataset**

The dataset is collected from Department of Information Management, Chung Hua University, Taiwan and Department of Civil Engineering, Tamkang University, Taiwan. The source URLs for the data is

<https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>.

Accessing this data does not violate any laws. This data does not appear to have been previously analyzed based on a Google search.

A preliminary survey of the data indicates an aggregate of over 30,000 rows and has 24 columns and the file is about 5.28MB.

**Brief Project Process Summary**

* Data Cleaning
* Feature extraction and engineering
* Data Visualization
* Model Training.
* Model Evaluation and Testing