## Introduction

In the rapidly evolving field of financial services, predictive analytics has become a cornerstone for decision-making and strategic planning. Financial institutions increasingly rely on data-driven insights to optimize their services and manage risks effectively. One critical area of interest is the ability to predict the likelihood of financial arrangements being terminated prematurely. Such predictions can help institutions better understand client behaviours, manage financial risks, and tailor services to client needs.

This use case focuses on building a predictive model using an anonymized dataset containing detailed records of financial arrangements. The goal is to develop a model that can accurately predict whether a financial arrangement will be terminated, based on various attributes such as demographic data, financial details, and arrangement specifics. These predictions will aid in proactive decision-making, allowing financial advisors to intervene or adjust terms to prevent potential defaults and improve client retention.

## Use Case Overview

Develop a predictive model using an attached dataset to assess the likelihood of a financial arrangement being terminated. This process includes data pre-processing, feature selection, building and tuning a classification model, evaluating its performance, and interpreting the results. The goal is to provide actionable insights that could influence decision-making regarding these financial arrangements.

## Model Building

Utilize machine learning algorithms to predict the "Terminated" status, a binary variable indicating whether an arrangement was discontinued. Suitable models could include Logistic Regression, Decision Trees, or Ensemble Methods like Random Forests and Gradient Boosting Machines. Model selection will depend on the accuracy, interpretability, and computational efficiency required by the stakeholders.

## Presentation Preparation

As part of our start-up environment, all development team members could interact directly with clients and present their work to stakeholders. For this project, you will prepare a presentation tailored to both technical and non-technical audiences. The presentation should outline your analytical methods, model-building process, key findings, and recommendations. Visual aids such as charts, graphs, and tables should be used to enhance understanding and engagement.

| Field Name | Description |
|---|---|
| clientid | Unique identifier for each client. |
| Lower_Super_Output_Area_Code | Generic area code, used if the postcode is standard. |
| Output_Area_Classification_Code | Represents high-level social grades. |
| AB | Percentage of individuals in the highest social grade (AB). |
| C1 | Percentage of individuals in the lower middle social grade (C1). |
| C2 | Percentage of individuals in the skilled working social grade (C2). |
| DE | Percentage of individuals in the semi-skilled and unskilled social grade (DE). |
| DOB_Year | Year of birth of the client. |
| DOB_Month | Month of birth of the client. |
| Gender | Gender of the client, encoded as an integer. |
| Marital_Status | Marital status of the client, encoded as an integer. |
| Employment_Status | Employment status of the client, encoded as an integer. |
| iva_joint | Indicates whether the arrangement is a joint one (boolean). |
| Partner_DOB_Year | Partner's year of birth (if applicable). |
| Partner_DOB_Month | Partner's month of birth (if applicable). |
| Partner_Gender | Partner's gender (if applicable). |
| Partner_Employment_Status | Partner's employment status (if applicable). |
| household_income | Total household income. |
| household_expenses | Total household expenses. |
| household_DI | Disposable income of the household. |
| Household_Total_Debt | Total debt of the household. |
| no_adults | Number of adults in the household. |
| under_16 | Number of children under 16 in the household. |
| under_18 | Number of children aged 16 to 18 in the household. |
| arrangement_status | Status of the financial arrangement, encoded as an integer. |
| Arrangement_startdate_year | Start year of the financial arrangement. |
| Arrangement_Startdate_month | Start month of the financial arrangement. |
| duration_Arrangement_to_Date | Duration of the arrangement in months. |
| Total_Expected_Duration | Expected total duration of the arrangement in months. |
| home_owner_flag | Indicates whether the client owns a home (boolean). |
| owned_property_count | Number of properties owned. |
| vehicle_count | Number of vehicles owned. |
| Closure_Status | Status of the arrangement closure, encoded as an integer. |

| Field Name | Description |
|---|---|
| Terminated | Indicates whether the arrangement was terminated (boolean). |
| Terminated_Year | Year the arrangement was terminated (if applicable). |
| Terminated_Month | Month the arrangement was terminated (if applicable). |
| Completed | Indicates whether the arrangement was completed (boolean). |
| dividend_approved | Dividend approved, represented in pence per pound. |
| payment_frequency | Frequency of payments, encoded as an integer. |
| contributions_expected_to_date | Contributions expected up to the current date. |
| contributions_received_to_date | Contributions received up to the current date. |
| TotalPaymentsDue | Total payments due under the arrangement. |
| Arrears_Amount | Amount of arrears under the arrangement. |
| arrears_months | Number of months the arrangement has been in arrears. |
| Arrears_Category | Category of arrears, encoded as an integer. |
| NoMonths_FirstPayment | Number of months until the first payment was made. |
| agreed_missed_flag | Indicates whether there was an agreed missed payment (boolean). |
| NoMonth_FirstMissedPayment | Number of months until the first payment was missed. |
| Physical_Disability_Vulnerability | Indicates physical disability (boolean). |
| Illness_Vulnerability | Indicates illness vulnerability (boolean). |
| Addiction_Vulnerability | Indicates addiction vulnerability (boolean). |
| Mental_Health_Vulnerability | Indicates mental health vulnerability (boolean). |