

# Proposal for Predicting Term Deposit Subscriptions Using Machine Learning

## Introduction

This proposal outlines a machine learning project aimed at solving a business problem for a banking organization using supervised machine learning techniques. The dataset, sourced from Kaggle, contains information about customers and their responses to marketing campaigns. The goal is to use machine learning models to predict whether a customer will subscribe to a term deposit (binary classification problem).

## Business Problem

**Objective:** To improve the effectiveness of marketing campaigns by predicting customer responses, enabling targeted marketing strategies that can increase subscription rates for term deposits.

**Problem Statement:** The bank wants to identify which customers are likely to subscribe to a term deposit based on their demographic and transactional data. By leveraging machine learning, the bank can focus its efforts on the most promising leads, improving conversion rates and reducing marketing costs.

## Dataset Description

The dataset includes the following features:

- **age:** Age of the customer
- **job:** Job title
- **marital:** Marital status
- **education:** Education level
- **default:** Whether the customer has credit in default
- **balance:** Average yearly balance in euros
- **housing:** Whether the customer has a housing loan
- **loan:** Whether the customer has a personal loan
- **contact:** Contact communication type
- **day:** Last contact day of the month
- **month:** Last contact month of the year
- **duration:** Last contact duration in seconds
- **campaign:** Number of contacts performed during this campaign
- **pdays:** Number of days since the client was last contacted from a previous campaign
- **previous:** Number of contacts performed before this campaign
- **poutcome:** Outcome of the previous marketing campaign
- **y:** Whether the client subscribed to a term deposit (target variable)

## Machine Learning Techniques

To address the problem, we propose to implement three different machine learning techniques using Orange Data Mining Software:

1. Support Vector Machine (SVM)
2. K-Nearest Neighbors (KNN)
3. Neural Network

#### **Implementation Plan:**

**Dataset Link :** <https://www.kaggle.com/datasets/prakharrathi25/banking-dataset-marketing-targets>

#### **Step 1: Data Preparation**

1. **Load the Dataset:** Import the dataset into Orange.
2. **Preprocess the Data:**
  - Handle missing values.
  - Encode categorical variables (e.g., job, marital, education) using one-hot encoding or label encoding.
  - Normalize numerical features to ensure all features contribute equally to the model.

#### **Step 2: Model Implementation**

Implement the following models in Orange Software:

1. Support Vector Machine (SVM)
2. K-Nearest Neighbors (KNN)
3. Neural Network

#### **Deployment Considerations**

1. **Model Integration:** Integrate the best-performing model into the bank's existing marketing platform.
2. **Scalability:** Ensure the model can handle real-time predictions for a large number of customers.
3. **Monitoring and Maintenance:** Set up processes for continuous monitoring of model performance and periodic retraining with new data to maintain accuracy over time.

#### **Benefits to the Organization**

1. **Increased Subscription Rates:** By targeting the right customers, the bank can improve the success rate of marketing campaigns.
2. **Cost Efficiency:** Reduce marketing costs by focusing efforts on customers who are more likely to subscribe.
3. **Data-Driven Decision Making:** Enhance the bank's ability to make informed decisions based on predictive insights.