

# CAMPAÑA DEPOSITOS A PLAZO

## INTRODUCCION

Este proyecto se centra en una entidad financiera que busca incrementar el número de usuarios que poseen un depósito a plazo fijo. Para lograrlo, se utilizan una serie de datos personales, como la edad, estado civil, tipo de empleo y nivel educativo. Además, se incluyen variables financieras como si el cliente tiene deudas impagas, su balance financiero anual promedio, y si cuenta con préstamos hipotecarios o personales.

También se consideran los datos de contacto y las interacciones previas del cliente en relación con este tipo de producto.

El objetivo principal es que un mayor número de clientes suscriban depósitos a plazo, lo cual aumentaría los recursos económicos de la entidad para futuras inversiones rentables.

## OBJETIVOS DEL PROYECTO

### **1. *¿Cuáles son los objetivos del negocio?***

Aumentar el número de clientes que tienen un depósito a plazo con la entidad financiera y no ser muy agresivo con las comunicaciones, para ello se debe tener en cuenta si ya se le ha ofrecido el servicio anteriormente, hace cuanto se le ha contactado o si tiene o ha tenido el mismo.

### **2. *¿Qué decisiones o procesos específicos desea mejorar o automatizar con ML?***

Evaluar de manera automatizada si un cliente es elegible para acceder al servicio, tomando en cuenta tanto su salud financiera como su historial de contacto previo, de manera que no se produzca un acercamiento excesivamente agresivo.

### **3. *¿Se podría resolver el problema de manera no automatizada?***

Es posible utilizar un algoritmo para evaluar la salud financiera de los clientes, considerando si tienen préstamos activos, su balance económico, y si ya disponen del servicio. Además, otro algoritmo podría analizar las interacciones previas para reducir la agresividad en los contactos.

## METODOLOGÍA PROPUESTA

Como el objetivo es predecir si un cliente suscribirá un depósito a plazo, este problema se aborda como una clasificación binaria que indica si un cliente decidirá contratar o no el producto financiero. Para esto, considero que un modelo de clasificación como el Random Forest es adecuado, ya que este modelo es reconocido por su alta precisión en problemas de clasificación. Al combinar los resultados de muchos árboles de decisión, Random Forest genera predicciones más estables y precisas. Además, cada árbol es entrenado con distintas muestras de los datos, lo cual reduce errores específicos y evita el sobreajuste.

Con este enfoque, el modelo captura patrones generales en los datos sin adaptarse demasiado a casos específicos, lo cual es esencial para que pueda predecir bien cuando se aplican datos de nuevos clientes.

### Métricas de Evaluación

Como estamos tratando un problema de clasificación con datos balanceados (es decir, una distribución relativamente equitativa de clientes con y sin depósito), elegí las siguientes métricas de evaluación:

- **Precisión:** Mide la proporción de verdaderos positivos respecto al total de predicciones positivas. Esta métrica nos ayuda a ver cuántas predicciones de clientes interesados en un depósito fueron correctas, minimizando los falsos positivos.
- **Recall:** Mide la proporción de verdaderos positivos sobre el total de casos positivos reales, ayudando a identificar la mayoría de los clientes interesados en el depósito y evitando los falsos negativos.
- **F1 Score:** Es el balance entre Precisión y Recall. Esta medida resulta útil para equilibrar ambos aspectos y obtener una idea general del rendimiento, especialmente en problemas de clasificación como este.
- **Exactitud:** Dado que los datos están balanceados, la exactitud es una métrica representativa que nos muestra el porcentaje de predicciones correctas sobre el total de casos. Esto nos permite evaluar el rendimiento global del modelo de manera confiable.

Este enfoque asegura que evaluemos al modelo no solo en términos de su precisión general, sino también en su capacidad para identificar correctamente a los clientes interesados en un depósito, lo cual es crucial para tomar decisiones efectivas en el contexto bancario.

## DATOS DISPONIBLES

Los datos incluyen características personales del cliente (edad, estado civil, nivel educativo, entre otros), datos financieros (crédito en mora, saldo promedio anual, préstamos de vivienda y personales), y datos de contacto e interacciones previas (número de contactos durante la campaña, resultado de campañas anteriores, entre otros). Incluyendo también si el cliente ya cuenta con un depósito a plazo o no.

## METRICA DE ÉXITO DEL PROYECTO

La métrica clave será la **tasa de conversión**, que mide el porcentaje de clientes contactados que se suscriben a un depósito a plazo. Un incremento en esta tasa indicará el éxito del modelo.

## RESPONSABILIDADES ETICAS Y SOCIALES

- **Privacidad y Protección de los Datos:** Es fundamental garantizar que los datos de los clientes se manejen de manera segura y se cumplan todas las regulaciones sobre protección de datos. También es muy importante garantizar el almacenamiento seguro de estos datos.
- **Consentimiento Informado:** Los clientes deben ser informados de cómo y por qué se están utilizando sus datos y deben dar su consentimiento explícito para el uso de estos en modelos predictivos.
- **Imparcialidad del Modelo:** El modelo no debe introducir sesgos que puedan discriminar a ciertos grupos de clientes por características como género, edad, nivel educativo, o cualquier otra característica.
- **Derecho a acceder a sus Datos:** Los clientes deben tener derecho a acceder, corregir y eliminar su información personal en cualquier momento. Además, deben poder optar por no participar en los procesos de toma de decisiones exclusivamente automatizadas.