

# 1 Proyecto de ML para predecir clientes potenciales de prestamos Bancaña

## 1.0.1 Introducción

Se trata de una entidad bancaria “Bancaña” que busca identificar clientes potenciales mediante los datos obtenidos de su última campaña de carteras.

## 1.0.2 Objetivos del proyecto

1. ¿Cuáles son los objetivos del negocio?

El objetivo de negocio de Bancaña es aumentar el número de clientes que adquieran productos de cartera utilizando los datos de su última campaña, buscando así incrementar la conversión de clientes potenciales, mediante la optimización de estrategias de marketing.

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

Bancaña busca mediante la implementación de ML identificar los clientes potenciales mediante la segmentación de estos.

3. ¿Se podría solucionar el problema de forma no automatizada?

Sí, existen un sin número de soluciones manuales para la identificación de clientes potenciales; sin embargo, esto implica una gran cantidad de limitaciones y desafíos que pueden ser solucionados mediante ML.

## 1.0.3 Metodología Propuesta

4. ¿Cuál es el algoritmo de Machine Learning más adecuado para resolver este problema? ¿Cómo justifica la elección de ese algoritmo? ¿Qué métricas de evaluación se utilizarán para medir el rendimiento del modelo?

El algoritmo seleccionado es un Random Forest de clasificación, ya que el objetivo es predecir la probabilidad de que un cliente se convierta en cliente potencial y este modelo es eficaz para manejar conjuntos de datos complejos con múltiples características. El motivo por el que usaremos dicho algoritmo es debido a que este maneja relaciones no lineales entre variables y es adecuado para trabajar con datasets con valores faltantes, además permite extraer información sobre la importancia de algunas variables en la toma de decisiones estratégicas. Con respecto a las métricas de evaluación que utilizaremos, serán las siguientes: 1. Accuracy: Proporción de predicciones correctas. 2. Recall: Capacidad del modelo para identificar verdaderos positivos (clientes potenciales reales). 3. Precision: Precisión en las predicciones positivas realizadas por el modelo.

#### 1.0.4 Datos Disponibles

5. ¿Qué datos están disponibles para abordar este problema?

Los datos que tenemos disponibles para abarcar este proceso son los siguiente: 1. 50.000 Entradas correspondiente a clientes que participaron en la campaña. 2. 15 características que nos permitirán conocer características tanto demográficas, como personales y financieras. 3. Contamos con tipos de datos que abarcan variables categóricas (estado civil y ocupación), variables numéricas (ingresos y depósitos) y variables de resultado.

La variable de resultado de nuestro modelo será binaria e indicará si el cliente podría o no aceptar un producto de cartera, arrojándonos los siguientes valores posibles: - 0. El cliente acepta el producto.  
- 1. El cliente no acepta el producto.

#### 1.0.5 Métrica de Éxito

6. ¿Cuál es la métrica de éxito para este proyecto?

La métrica de éxito de Banca es maximizar la conversión de clientes hacia productos de cartera (depósitos) de manera eficiente y rentable.

#### 1.0.6 Responsabilidades Éticas y Sociales

7. ¿Qué responsabilidades éticas y sociales es importante tener en cuenta?

Ya que involucramos datos personales en decisiones automatizada es crucial considerar varias responsabilidades éticas y sociales, tales como: - Privacidad y protección de datos. Que nos obligan a cumplir estrictamente las normativas establecidas como el RGPD - Minimizar los riesgos de exclusión financiera, mitigando el sesgo mediante la revisión y la depuración del dataset para detectar posibles comportamientos erróneos del modelo con respecto a edad, género u otros atributos sensibles. - Transparencia en la forma como se utilizan los datos y como se toman decisiones basadas en el modelo. - Impacto social: Evitar que el modelo cree desigualdad económica y social implementando estrategias que faciliten el acceso equitativo a los recursos financieros.

[ ]: