

## **Mini Proyecto 1**

### **Ciencia de Datos**

#### **Análisis de riesgos en banca y servicios financieros.**

#### **Introducción**

Este caso práctico pretende la aplicación de EDA en un escenario empresarial real. En este estudio de caso, además de aplicar las técnicas aprendidas, también debe desarrollarse una comprensión básica del análisis de riesgos en los servicios bancarios y financieros y entender cómo se utilizan los datos para minimizar el riesgo de perder dinero al prestar a los clientes.

#### **Comprensión empresarial**

A las empresas proveedoras de préstamos les resulta difícil conceder préstamos a las personas debido a su historial crediticio insuficiente o inexistente. Por eso, algunos consumidores lo utilizan como ventaja convirtiéndose en morosos. Supongamos que trabajas para una empresa de financiación al consumo especializada en conceder diversos tipos de préstamos a clientes urbanos. Tiene que utilizar EDA para analizar los patrones presentes en los datos. Así se asegurará de que no se rechaza a los solicitantes capaces de devolver el préstamo.

Cuando la empresa recibe una solicitud de préstamo, tiene que decidir su aprobación en función del perfil del solicitante. La decisión del banco conlleva dos tipos de riesgos:

- Si es probable que el solicitante devuelva el préstamo, la no aprobación del préstamo supone una pérdida de negocio para la empresa.
- Si no es probable que el solicitante devuelva el préstamo, es decir, si es probable que incurra en impago, entonces aprobar el préstamo puede suponer una pérdida financiera para la empresa.

Los datos que se ofrecen a continuación contienen la información sobre la solicitud de préstamo en el momento de solicitarlo. Contiene dos tipos de escenarios:

- El cliente con dificultades de pago: tuvo retrasos de pago de más de X días en al menos una de las primeras Y cuotas del préstamo de nuestra muestra,
- Todos los demás casos: Todos los demás casos en los que se paga puntualmente.

Cuando un cliente solicita un préstamo, hay cuatro tipos de decisiones que podría tomar el cliente/la empresa):

*Aprobado:*

La empresa ha aprobado la solicitud de préstamo

*Cancelada:*

El cliente canceló la solicitud en algún momento durante la aprobación. O bien el cliente cambió de opinión sobre el préstamo o, en algunos casos, debido a un mayor riesgo del cliente recibió un peor precio que no deseaba.

*Rechazado:*

La empresa ha rechazado el préstamo (porque el cliente no cumple sus requisitos, etc.).

*Oferta no utilizada:*

El préstamo ha sido cancelado por el cliente, pero en diferentes fases del proceso.

## **Objetivos empresariales**

El estudio de caso pretende identificar patrones que indiquen si un cliente tiene dificultades para pagar sus cuotas, lo que puede servir para tomar medidas como denegar el préstamo, reducir el importe del préstamo, prestar (a solicitantes de riesgo) a un tipo de interés más alto, etc. Esto garantizará que no se rechace a los consumidores capaces de devolver el préstamo. La identificación de estos solicitantes mediante EDA es el objetivo de este estudio de caso.

En otras palabras, la empresa quiere entender los factores impulsores (o variables impulsoras) del impago de préstamos, es decir, las variables que son fuertes indicadores de impago. La empresa puede utilizar estos conocimientos para evaluar su cartera y sus riesgos.

## **Comprensión de los datos**

### **1. application\_data.csv**

Contiene toda la información del cliente en el momento de la solicitud. Los datos se refieren a si el cliente tiene dificultades de pago.

### **2. previous\_application.csv**

Contiene información sobre los datos del préstamo anterior del cliente. Contiene los datos sobre si la solicitud anterior había sido Aprobada, Cancelada, Rechazada u Oferta no utilizada.

### **3. columns\_description.csv**

Es un diccionario de datos que describe el significado de las variables.

## **Rúbrica de evaluación**

Se tendrán en cuenta dos ítems a evaluar con un porcentaje diferente:

1. Limpieza y transformación de datos 30%
2. Análisis exploratorio de datos. 70%

Se debe presentar un informe en el cual se detallen los dos ítems a evaluar:

1. La primera parte del informe debe detallar el tratamiento que se ha dado a cada una de las variables. En caso de no aplicar algún tratamiento explicar la razón.
2. La segunda parte del informe debe contener un listado de 10 inferencias sobre las variables impulsoras más determinantes para el impago. Preferiblemente, cada inferencia debe estar soportada en estadísticos o gráficos extraídos del EDA.
3. Deben entregarse 3 documentos: 1) Notebook con los procesos desarrollados (preferiblemente documentado). 2) informe. 3) Diapositivas para presentación de sustentación.