Desarrollo de una herramienta para **pymes que ofrecen financiación directa a clientes**

Proyecto aplicado a pymes, pues es un nicho en el que se podría "vender" como productos.

Desarrollar una herramienta sencilla para **pymes que ofrecen financiación directa a clientes**, como agencias inmobiliarias, concesionarios o comercios locales. El sistema calculará de forma automática el **tipo de interés mínimo recomendado** para cada solicitud, basándose en el perfil de riesgo del cliente. Utilizando unos pocos datos básicos (edad, ingresos, importe y duración del préstamo), permitirá a la pyme **proteger su rentabilidad y minimizar riesgos**. La solución está diseñada para ser rápida, práctica y fácilmente integrable, sin necesidad de grandes infraestructuras ni conocimientos técnicos avanzados.

Para construir esta solución, entrenaría un **modelo de clasificación ligero** (como una regresión logística o un árbol de decisión) usando variables básicas del cliente. A partir de la **probabilidad de impago (PD)** estimada, aplicaría una fórmula sencilla de cálculo de interés mínimo:  
Interés mínimo = tipo base + (coeficiente de riesgo \* PD predicha).  
La validación del modelo se realizaría mediante el **AUC** (Área Bajo la Curva ROC) para garantizar su fiabilidad. Finalmente, implementaría un **formulario web interactivo** donde la pyme pueda introducir los datos del cliente y recibir una recomendación inmediata.

Para crear el modelo de **probabilidad de impago (PD)** usando el dataset prestamos.csv, primero se seleccionarían solo las variables más relevantes y fácilmente disponibles en una pyme: edad, ingresos, importe solicitado, duración del préstamo y tipo de empleo. La variable objetivo sería el estado de impago (0 = no impago, 1 = impago).

Se preprocesarían los datos corrigiendo valores nulos y codificando variables categóricas. Luego, se entrenaría un **modelo simple** de regresión logística o árbol de decisión, dividiendo el dataset en entrenamiento y test para validar su desempeño. El modelo se evaluaría usando la **métrica AUC**, asegurando buena capacidad de distinguir clientes de alto y bajo riesgo. Una vez entrenado, el modelo predeciría la PD de nuevas solicitudes. Finalmente, se aplicaría una fórmula de **pricing ajustado al riesgo**, recomendando un interés mínimo basado en la probabilidad de default estimada, ayudando así a la pyme a proteger su rentabilidad en operaciones de financiación.

• Ingresos mensuales 🡪 ingresos 🡪 float64

• Importe del préstamo solicitado 🡪 principal 🡪 float64

• Duración del préstamo 🡪 num\_cuotas 🡪 object

• Tipo de contrato o situación laboral (si está disponible)

• Variable objetivo: si el préstamo fue impagado o no

**PD – Probability of Default (Probabilidad de impago)**  
Predice si un cliente va a impagar o no (evento binario: sí o no).

**Algoritmo: Regresión Logística**

* Es ideal para resolver problemas de **clasificación binaria**.
* Estima directamente la **probabilidad** de que ocurra un evento (impago).
* Fácil de interpretar y muy utilizado en banca por su robustez y transparencia regulatoria.

1. Regresión logística como baseline

2. Validar con árbol de decisión simple para explorar mejoras