

Informe del proyecto final

- **Carrera:** Licenciatura en Ciencias de Datos
 - **Materia:** Introducción a la Ciencia de Datos
 - **Profesor:** Claudio Pavón
 - **Integrantes:**
 - Abraham, Milena Paula
 - Arese, Juan Cruz
 - Ceschini, Francisco
 - Cogorno Díaz, Lautaro Javier
-

1. Objetivo del proyecto:

Analizando un dataset y utilizando modelos de Machine Learning (Random Forest y Decision Tree), predecimos si una persona incumplirá o no el pago de una tarjeta de crédito.

2. Descripción de variables del dataset:

En la pagina del dataset en Kaggle.com, el autor del mismo explica el significado de cada columna:

COLUMNA	DESCRIPCIÓN
ID	ID de cada cliente
LIMIT_BAL	Monto del crédito otorgado en dólares taiwaneses (incluye crédito individual y familiar/complementario)
SEX	Género (1=masculino, 2=femenino)
EDUCATION	1=posgrado, 2=universitario, 3=preparatoria, 4=otros, 5=desconocido, 6=desconocido
MARRIAGE	Estado civil (1=casado, 2=soltero, 3=otros)
AGE	Edad en años
PAY_X	Estado de pago en un determinado mes de 2005 (-1=pago al día, 1=un mes de retraso en el pago, 2=dos meses de retraso en el pago,...9=nueve meses o más de retraso en el pago)

BILL_AMTX	Importe del extracto de factura en un determinado mes de 2005 (dólares taiwaneses)
PAY_AMTX	Importe del pago anterior en un determinado mes de 2005 (dólares taiwaneses)
default.payment.next.month	Pago default (1=sí, 0=no)

3. Limpieza y preparación de los datos
