

Análisis de Predicción para Tarjetas de Crédito

Carmen Sánchez Servio





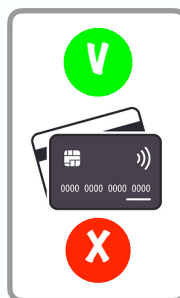
Condiciones laborales



Ausencia historial crediticio



<21 >60



Retrasos pagos previos



Procedencia ingresos



Bajos ingresos



Avales pago

Introducción



Datos → **Kaggle**

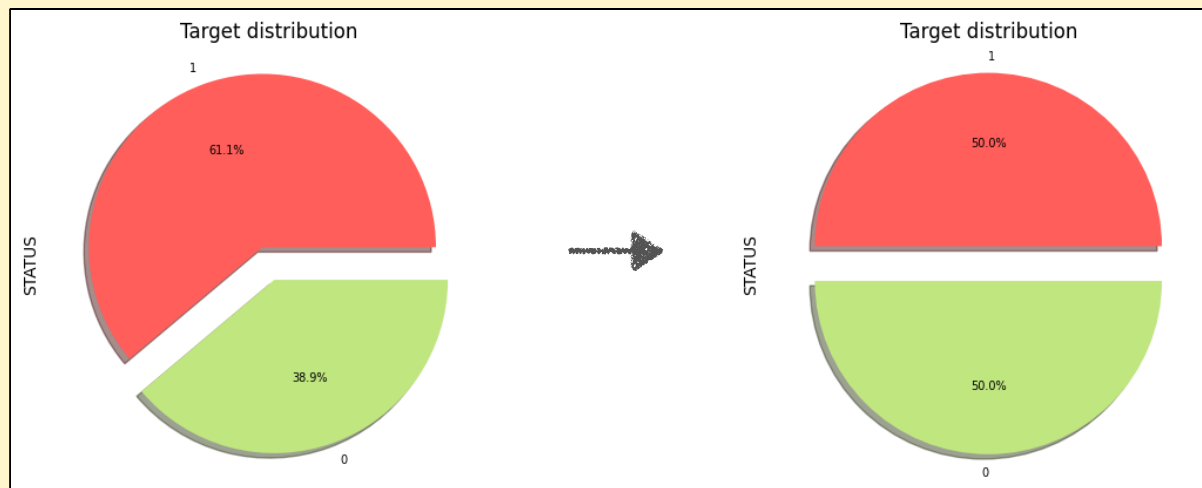
Target → **Status (estado de pago)**

- **Está al corriente**
- **No está al corriente**

Datos



- a. Tamaño elevado datos
- b. Variables categóricas (numéricas)
- c. Target desbalanceado (under sampling)
- d. Missings



Dificultades



1. REGRESIÓN LOGÍSTICA
2. KNN
3. RANDOM FOREST
4. ÁRBOL DE DECISIÓN
5. GRADIENT BOOSTING
6. XG BOOST
7. SVM



Recall

$$\frac{TP}{TP + FN}$$

F1-score

$$2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

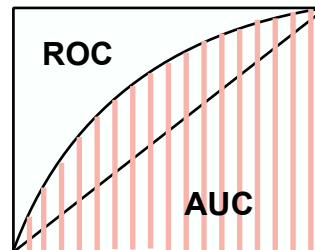
Precision

$$\frac{TP}{TP + FP}$$

Accuracy

$$\frac{TP + FP}{\text{Total}}$$

Curva ROC y AUC



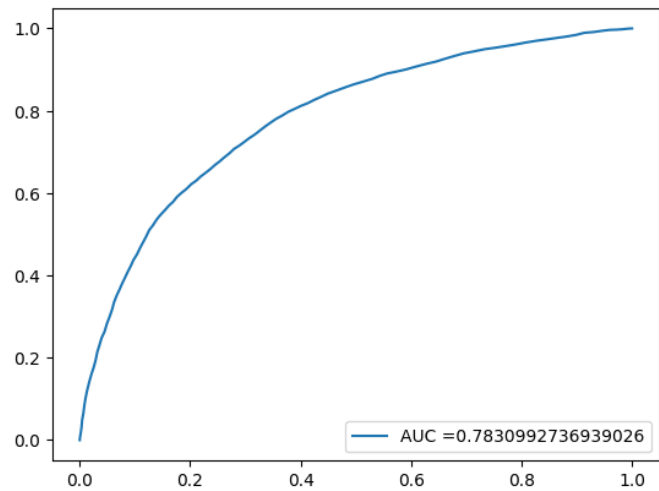
Matriz de confusión

| | |
|----|----|
| TP | FP |
| FN | TN |

Métricas



Random Forest

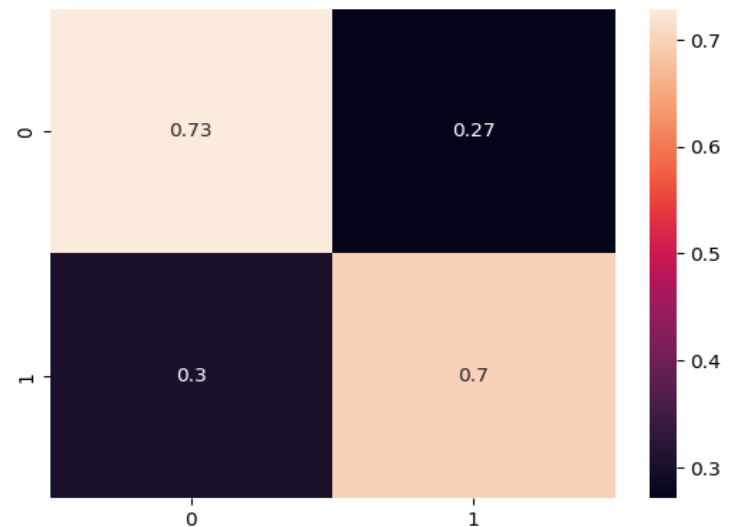


Recall = 0.7001

F1-score = 0.7046

Accuracy = 0.7119

AUC = 0.7830



Resultados



1. Desempeño aceptable del modelo.
2. Precisión relativamente alta del modelo.

Conclusiones

