

EDA →

Variables a Quitar	Var Necesit Transform
Var Colineales	

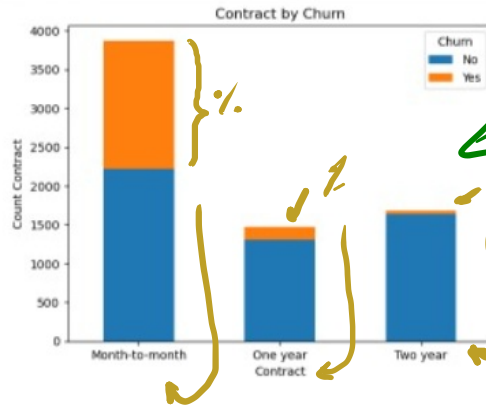
TELCO

- Quitar: Customer Id,
- Imputar: Total Charge
- Balanceo

BANK

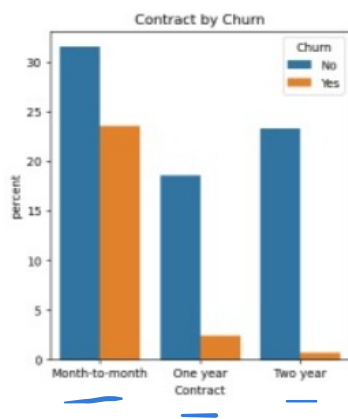
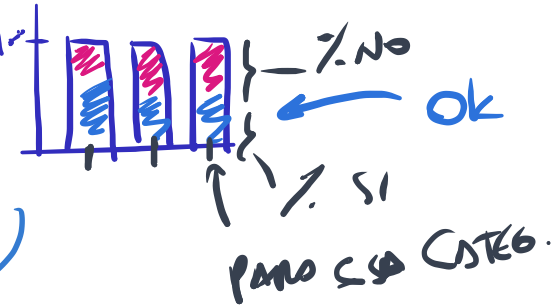
- Idea: ① Clusterizar primero → VARIABLE cluster
- ② Modelo Classif $y = f(x)$ X : + cluster

- Quitar: Rowname, Surname, Customer Id
- Imputar:
- Balanceo

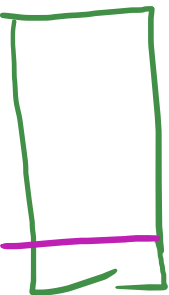


No Recom

fill = 100%



Complementary
ABSOLUTES



TRAIN
TEST
Preprocessor
+
Ing Variables

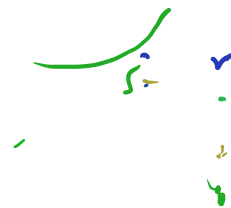
Modelamiento con hiperparam
Remuestreo | Métricas |
Modelo → Meta Búsqueda

Afinar
hiperparam

¿Cuáles métricas serán
las más importantes?
- JUSTIFICAR

Modelo final

Evaluar
¿Sobreajuste?



TARE

- 2 Modelos

- Random Forest
- esloger ← AUISAR

- Reg log Elastic Net
- GAM
- MARS
- Bagging MARS
- Bagging GAM
- Boosting $\begin{cases} \text{light} \\ \text{XGB} \\ \text{GAM} \end{cases}$

(ALL) ← GUID
- R - no repetir
- Py - no repetir

Fecha

2-MARZO-2024

GitHub
- c group un
 archivo diferente
- Modelo Room 01 $\begin{cases} R \\ RMD \\ IPYNB \\ QMD \end{cases}$

		SI	↓ Real	No	
Pred →	SI	100	20	120	
	No	150	730		
		250	750		

$$Acc = \frac{100 + 730}{1000} = 83$$

$$Prec\ Pos = \frac{100}{120} = 83$$

$$Recall\ Pos = \frac{100}{250} = 0,4$$

Problema
Dados No
Balanceados

		SI	↓ Real	No	
Pred →	SI	200	120		
	No	50	630		
		250	750		

$$Precis\ Pos = \frac{200}{320} = 0,62$$

$$Recall\ Pos = \frac{200}{250} = 0,80$$