

# HACKATHON

28 MAYO | ONLINE



## "PLAY THE CHALLENGE" WITH CAIXABANK + TECH



REACT



DATA SCIENCE



+



3.000€ EN PREMIOS

Carlos Pérez Ricardo

Ingeniero  
Aeroespacial en UPC  
de Terrassa

Data Scientist en  
Grupo Hotusa

#TECH4U



CaixaBank  
TECH

x



NLIW

# Solución propuesta

El objetivo es construir un dataset que **permita entender y detectar patrones en el mercado reciente** para poder conocer si el precio de cierre a 3 días vista, será menor o mayor que el precio de cierre de hoy.

**No nos importa tanto el valor exacto del cierre**, sino la evolución en el mercado reciente que permite conocer si el mercado subirá o bajará (por lo que referenciamos los valores a los valores del mercado hace 7 días)

```
train[col] = train[col] - train['Close'].shift(7)
```

Los atributos generados básicamente consisten en shift y rolling window.

Se trata de un problema de **Clasificación** (balanceada).

# Resultados

Dado que tratamos con Series Temporales => El train-test split para validar nuestro modelo es muy importante. Se coge el último año para validar el modelo.

Se probaron diferentes modelos (SVM, RF, LGBM y LSTM (pendiente))

Random Forest	TRAIN	TEST
Accuracy	0.621844	0.537254
Precision	0.58983	0.522448
Recall	0.883399	0.992248
F1-score	0.70737	0.684491
ROC AUC score	0.69872	0.56263

Confusion  
matrix  
TRAIN

**PREDICTED**

**REAL**

	True	False
True	1038	2002
False	380	2879