



Este es uno de los retos clasificatorios que forman parte del hackathon online de CaixaBank Tech.

Background

Overview: the dataset and challenge

Se emplearán tres datasets: el dataset de training, el dataset de tweets, dataset de test.

Reto:

Desarrollar un modelo predictivo que permita predecir la variable target (si el precio de cierre del IBEX35 será superior o inferior al precio de cierre actual).

La variable objetivo:

Target: Esta es la variable a predecir. Es una variable binaria.

- 1: Indica que el precio de cierre tres días adelante será más alto que el precio de cierre actual.
- 0: Indica que el precio de cierre tres días adelante será igual o menor al precio actual.

Librerías Necesarias:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report
```

Desarrollo:

He preparado los datos *datetime* como una variable numérica a pesar de que tuve la tentación de eliminarlo, porque para el conjunto de datos aprecié que los datos de entrada, salida, máximo y mínimo son suficientes. Tanto más cuanto que se trata de valores históricos desde 1994, con un margen de +/- 3 años.

Para averiguar la precisión, la sensibilidad del algoritmo utilizamos la librería metrics de scikit-learn e importamos `precision_score`, y la implementamos junto con los datos reales y los que hemos obtenidos utilizando el modelo.