Luis Buceta Ojeda: 100303573 | Lucia Alvargonzalez Azqueta: 100315935 | Lauro Bravar Abril-Martorell: 100303876 | Adrián Rodríguez Grillo: 100316457

IBEXPredictor

PRACTICA 3

Inteligencia Artificial en las Organizaciones

# Introducción y objetivos

# Estado del arte

# Obtención y procesamiento de los datos

## Obtención de los datos

Para la obtención de datos de las diferentes bolsas y sus históricos utilizamos la API en python de Quandl, que es una empresa que se dedica a la recopilación de diferentes datos económicos. Esta API nos permitió acceder a los datos históricos de las plazas mundiales en formato *csv*, pudiendo modificar la forma de obtenerlos. Los parámetros que podíamos usar eran los siguientes:

* **Valores o conjuntos de valores:** Este parámetro nos permitía seleccionar que mercado o conjunto de mercados queríamos datos.
* **Datos deseados:** Este parámetro por defecto devolvía todos los posibles datos disponibles, como el precio de apertura, cierre, máximo, mínimo y volumen. Si se deseaba se podía acotar a los deseados.
* **Fecha de inicio y final:** El rango de fechas que se devolverían en el csv
* **Tipo de datos:** Este parámetro, que fue de los más útiles, nos permitió acceder a diferentes representaciones de datos como fueron los valores numéricos normales, los valores de cambio con respecto al día anterior, los porcentajes de subida y bajada con respecto al día anterior y los datos normalizados.
* **Frecuencia de recogida:** Este parámetro permitía ver cada cuanto se recogían los datos, siendo las opciones diario, mensual, semanal. Visto que queríamos una predicción diaria usamos la opción diaria.

Para la selección de plazas mundiales a elegir, observamos las principales mostradas en diferentes páginas de mercado como Yahoo Finance[[1]](#endnote-1), Expansión[[2]](#endnote-2) o Bolsamanía[[3]](#endnote-3), eligiendo finalmente seis bolsas:

* El IBEX35, para tener los datos del día.
* El Dow Jones, que es la bolsa estadounidense y, por tanto, la que más influencia en el resto.
* El Nikkei 225, que es la bolsa de Japón y la principal de la parte oriental del mundo.
* El Euro Stoxx50, que contiene los 100 principales valores de las bolsas europeas y, por tanto, es un generalizador de Europa.
* El Dax, que es la bolsa alemana y unas de las principales europeas.
* El Cac40, que es la bolsa francesa y otra de las principales europeas.

Una de las bolsas que es de las más importantes en el mundo es la de Reino Unido, pero debido a que utilizamos la versión gratuita de la API no estaba disponible.

## Decidicisiones

Para predecir la tendencia del día siguiente, consideramos que los datos más útiles serían los de cierre, debido a que los demás, como el máximo o mínimo del día, no serían muy influyentes sin conocer el resultado final y otros, como el precio de apertura, directamente no resultaban útiles.

Por tanto, para el proceso de obtención de datos lo que hicimos fue recoger, para las variables de entrada, los valores de cierre de las seis bolsas desde el 2000 hasta el día anterior a la recogida y, para la salida, los datos de cierre del IBEX35 para las mismas fechas. Una vez obtenidos estos datos, el siguiente paso era procesarlos para obtener los ficheros deseados para ejecutar las pruebas.

Este problema, que es la predicción de la tendencia del día siguiente de un mercado dados los datos históricos del mismo mercado y otros más, lo planteamos como un problema de clasificación donde la salida serían únicamente dos clases, sube y baja, que representaríamos como 1 y 0 respectivamente y las variables de entrada serían los datos de las diferentes plazas.

Para las pruebas planteamos dos formas diferentes de solucionar el problema, mediante la predicción a un día, donde con los datos de cierre de las bolsas de un día se predeciría el día siguiente, y con la predicción a cinco días, donde se utilizarían los datos de cinco días de las bolsas. El motivo de esta elección fue debido a dos razones, por una parte, diversos artículos de especialistas en la materia indicaban que la relación temporal entre los datos no era duradera en el tiempo y, por otra, queríamos evitar los patrones estacionales y temporales que se pueden producir en ciertas épocas del año.

## Procesamiento

El primer paso para la creación de los ficheros de datos fue obtener la salida, para ello, como hemos indicado anteriormente, tomamos los datos de cierre del IBEX35, en concreto los de porcentaje de cambio y tomamos los valores superiores al 0% como días de subida y el resto como día de bajada. Posteriormente, subimos la columna de resultados un día para convertirlo en la salida del día anterior. Es decir, si un día sube un 3% este pasaría a ser día de subida y posteriormente pasaría a ser la salida del día anterior.

El segundo paso fue juntar todas las columnas con los diferentes datos de entrada y la salida mediante la combinación de las tablas tomando como índice las fechas. Uno de los problemas que encontramos fue que no todas las bolsas abren los mismos días debido a los días festivos, por lo que, para solucionar esto, tuvimos que tomar dos medidas.

Primero, eliminar los días en los que el IBEX no operaba, ya que no tiene sentido predecir un día en el que la plaza no abrirá. En estos casos, la salida lo que predecirá será el siguiente día que se abra. Y segundo, en el caso del resto de bolsas, si en un determinado día una plaza no abre lo que se hace es arrastrar el valor del último día abierto hasta encontrarse otro día en el que se encuentren datos.



Tras estos dos paso, ya tendríamos los datos preparados en un nuevo fichero

# Resolución del problema

# Conclusiones

# Trabajos futuros

# Referencias

1. Yahoo Finanzas: <https://es.finance.yahoo.com/> [↑](#endnote-ref-1)
2. Expansión: <http://www.expansion.com/> [↑](#endnote-ref-2)
3. Bolsamanía: <http://www.bolsamania.com/> [↑](#endnote-ref-3)