**ORDEN DE EJECUCIÓN**

Las libretas que se exponen a continuación tratan de estudiar y generar una predicción de las temperaturas para Valencia y Alicante mediante los datos reportados por los distintos dispositivos receptores de temperaturas, precipitaciones, presiones y CO2.

Observamos que las temperaturas medias y mínimas se explican bien en el largo plazo con la variable explicativa “co2”. Pero la variable explicativa “co2” no es capaz de captar la variabilidad de las temperaturas máximas.

Por este motivo, para la predicción de las temperaturas máximas de los meses de verano, en la libreta que busca predecir olas de calor, se realizará de un modo más creativo.

Método de predicción de las temperaturas:

Libreta 1\_1\_PFM:

Esta libreta tiene todos los datos de todos los dispositivos. En esta libreta se realiza una revisión de las variables, gestión del formato de los datos, limpieza de los datos, un análisis visual de las variables y una división por localización.

Esta división derivó en la creación de los ficheros de Valencia y Alicante.

Libreta 1\_2\_CO2 y 1\_3\_co2\_prediction:

* Observamos que las emisiones de “co2” muestran signos de estacionalidad y tendencia, pero no es estacionaria. Pero se diferenció dos veces para hacerla estacionaria, por lo que generamos una predicción mediante el uso de series temporales.
* Estas predicciones las usaremos para predecir las temperaturas futuras debido a la correlación del CO2 con las temperaturas medias en el largo plazo.
* Generamos un DataFrame para Valencia y otro para Alicante. Dos distintos por el distinto tamaño de las muestras de Alicante y Valencia.

Libretas 2\_1\_v\_tmed, 2\_2\_v\_tmin, 2\_3\_v\_tmax, 2\_4\_valencia\_verano\_rgl, 2\_5\_a\_tmed, 2\_6\_a\_tmin, 2\_7\_a\_tmax, 2\_8\_alicante\_verano\_rgl.

Las libretas referentes a “tmed”, “tmin” y “tmax”:

* Ejecutan un código que comparan los resultados de la regresión lineal simple y regresión polinómica de las variables objetivos a las que se refieren con la variable explicativa “co2”. En el modelo de regresión lineal múltiple se seleccionan múltiples variables.
* Al final de cada libreta se observa una pequeña valoración entre datos pasados y futuros que dan visibilidad a un posible futuro.

Comprobamos como la evolución de las temperaturas parecen guardar una relación con la evolución temperaturas pasadas.

Las libretas referentes a “verano”:

* Generaremos un DataFrame anual con la media de las temperaturas en los meses más calurosos (junio, julio y agosto).
* Entrenamos el modelo con la variable explicativa “co2”.
* Guardamos el modelo entrenado y generamos predicciones para la temperatura media con el DataFrame que tenemos con predicciones de “co2”.
* Ahora, entrenamos la variable “tmax”, temperaturas máximas, con la variable “tmed, temperaturas medias.
* El modelo resultante lo guardamos en una variable.
* Utilizamos las predicciones realizadas para “tmed” y generamos predicciones para “tmax” con el modelo que habíamos entrenado.
* Realizamos la misma operación con “tmin”, temperaturas mínimas.
* Generamos un DataFrame para cada predicción y los representamos gráficamente.
* Para finalizar unimos los DataFrames con las predicciones.
* Este nuevo DataFrame con las predicciones lo unimos al DataFrame original, habiendo eliminado previamente las variables que no representan temperaturas o “co2”.

Este fichero, generado en el último punto, representa con temporalidad anual los meses de verano más calurosos de Valencia o Alicante. Depende si se ha ejecutado para Valencia o para Alicante.

Por otro lado, para finalizar, comprobamos como se comporta un modelo de machine learning y comparamos los resultados de las predicciones los resultados de las predicciones de nuestro modelo de regresión lineal simple.

Libretas 3\_1\_v\_ML y 3\_2\_a\_ML:

* Modelo con aprendizaje no automático.
* Modelo con aprendizaje automático.
* Modelo con redes neuronales aplicando método de optimización “GridSearchCV.
* Se valoran los resultados en función de la desviación entre los valores reales y predichos (mape).
* Se comparan las predicciones realizadas con el modelo de redes neuronales con las previsiones del modelo de regresión simple.
* Se genera un nuevo DataFrame en el que se muestra la diferencia entre ambos valores.