

Aprendizaje Automático

Departamento de Informática – UC3M

PRACTICA 1 – PREDICCIÓN DE LA PRODUCCIÓN DE
ENERGÍA EÓLICA

Práctica 1

Objetivo

- Practicar con métodos de aprendizaje automático
- Practicar con las diferentes técnicas de ajuste de hiperparámetros (HPO)
- Realizar todo el proceso: selección del modelo (seleccionar el mejor modelo, incluyendo HPO), estimar el rendimiento futuro (evaluación del modelo), construir el modelo final y utilizarlo para hacer predicciones.
- Pasar un problema de predicción a clasificación.

Práctica 1

Introducción

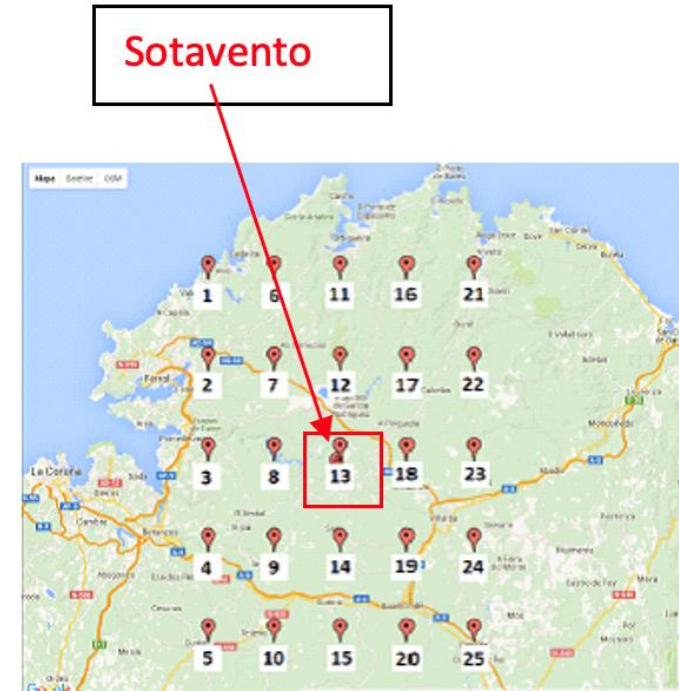
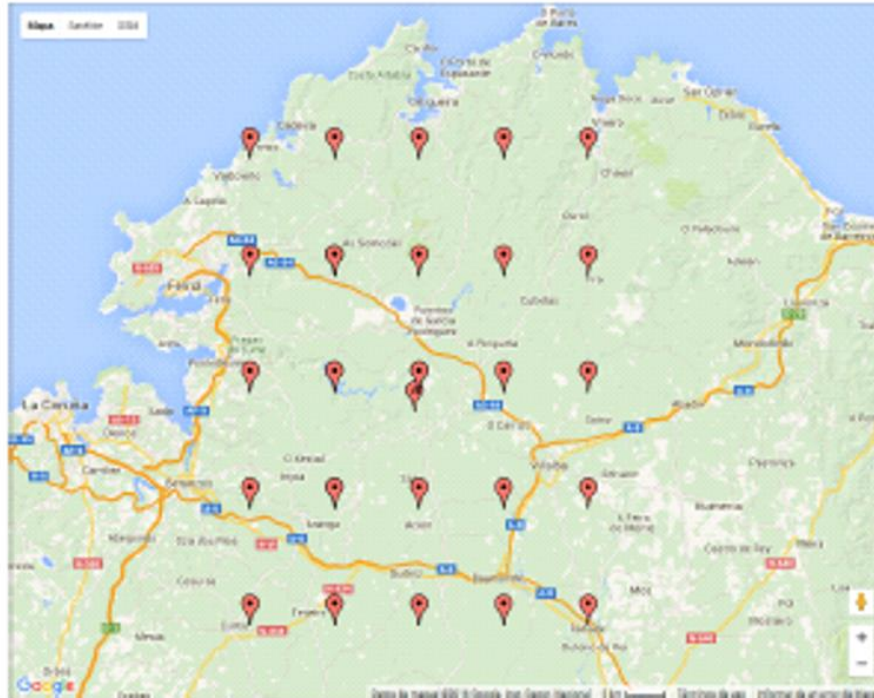
- Uno de los principales problemas de la energía eólica es la variabilidad y la incertidumbre. Los productores de energía necesitan prever un calendario de producción de energía con 24 horas de antelación. Por eso es importante disponer de previsiones eólicas precisas.
- Nuestro principal objetivo será predecir la energía eólica diaria en el parque eólico de Sotavento (Lugo), utilizando las previsiones meteorológicas como atributos (predictores).



Parque eólico de Sotavento (Lugo).

Práctica 1

Introducción



ECMWF (European Centre for Medium-Range Weather Forecasts)
<https://www.ecmwf.int/>

Práctica 1

Descripción del problema: Entradas (predictores)

t2m: 2 metre temperature	lai_lv: Leaf area index, low vegetation
u10: 10 metre U wind component	u10n: Neutral wind at 10 m u-component
v10: 10 metre V wind component	v10n: Neutral wind at 10 m v-component
u100: 100 metre U wind component	stl1: Soil temperature level 1
v100: 100 metre V wind component	stl2: Soil temperature level 2
cape: Convective available potential energy	stl3: Soil temperature level 3
flsr: Forecast logarithm of surface roughness for heat	stl4: Soil temperature level 4
fsr: Forecast surface roughness	sp: Surface pressure
iews: Instantaneous eastward turbulent surface stress	p54.162: Vertical integral of temperature
inss: Instantaneous northward turbulent surface	p59.162: Vertical integral of divergence of kinetic energy
lai_hv: Leaf area index, high vegetation	p55.162: Vertical integral of water vapour

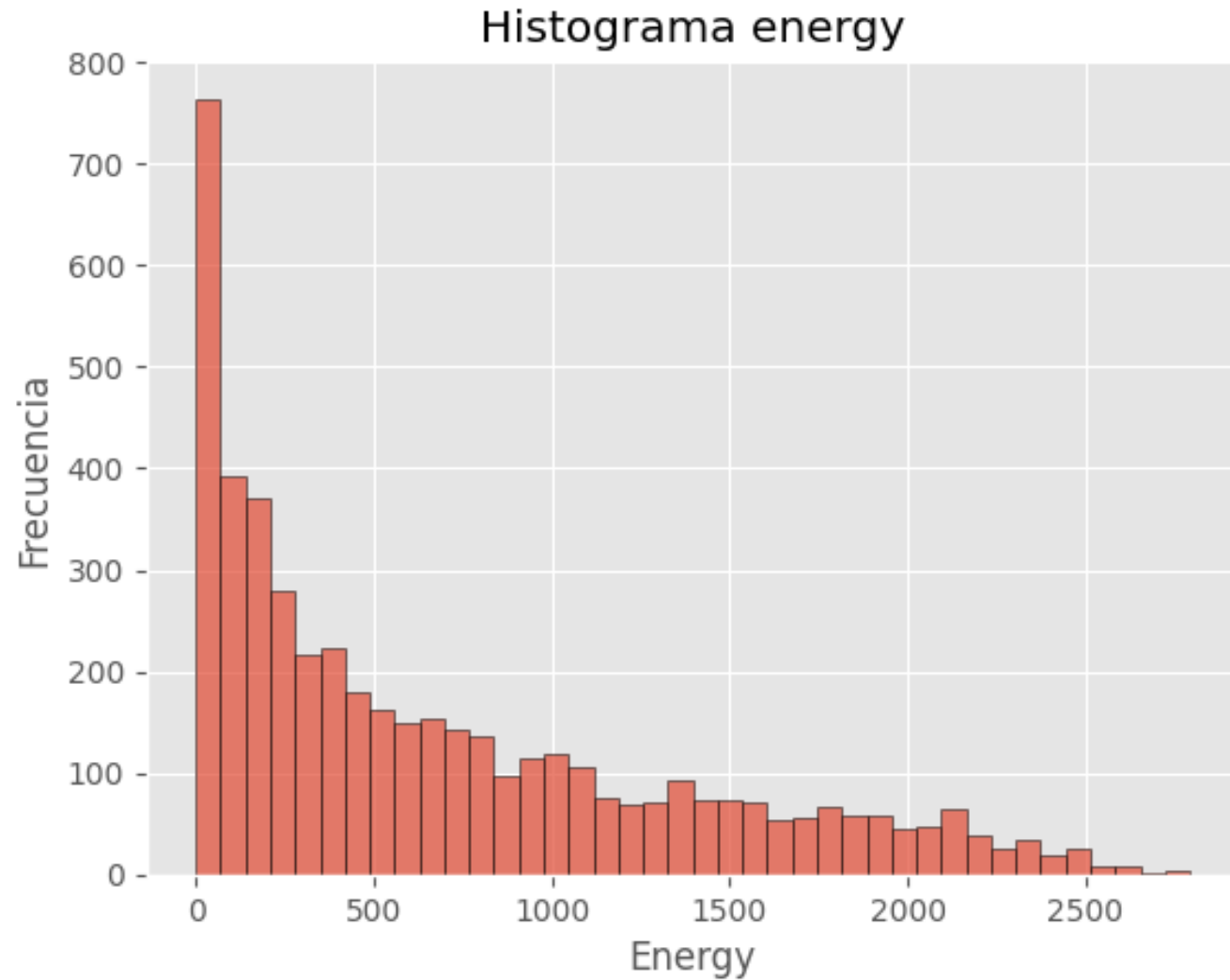
Práctica 1

Descripción del problema

- Variable dependiente "energy". Problema de regresión:
 - EDA (Exploratory Data Analysis) simplificado.
 - Métodos de escalado (KNN).
 - KNN, árboles de regresión, regresión lineal y SVM.
- Variantes:
 - Regresión para bajas y altas energías.
 - Clasificador bajas/altas energías.
- ChatGPT
- GitHub

Práctica 1

Variable de salida



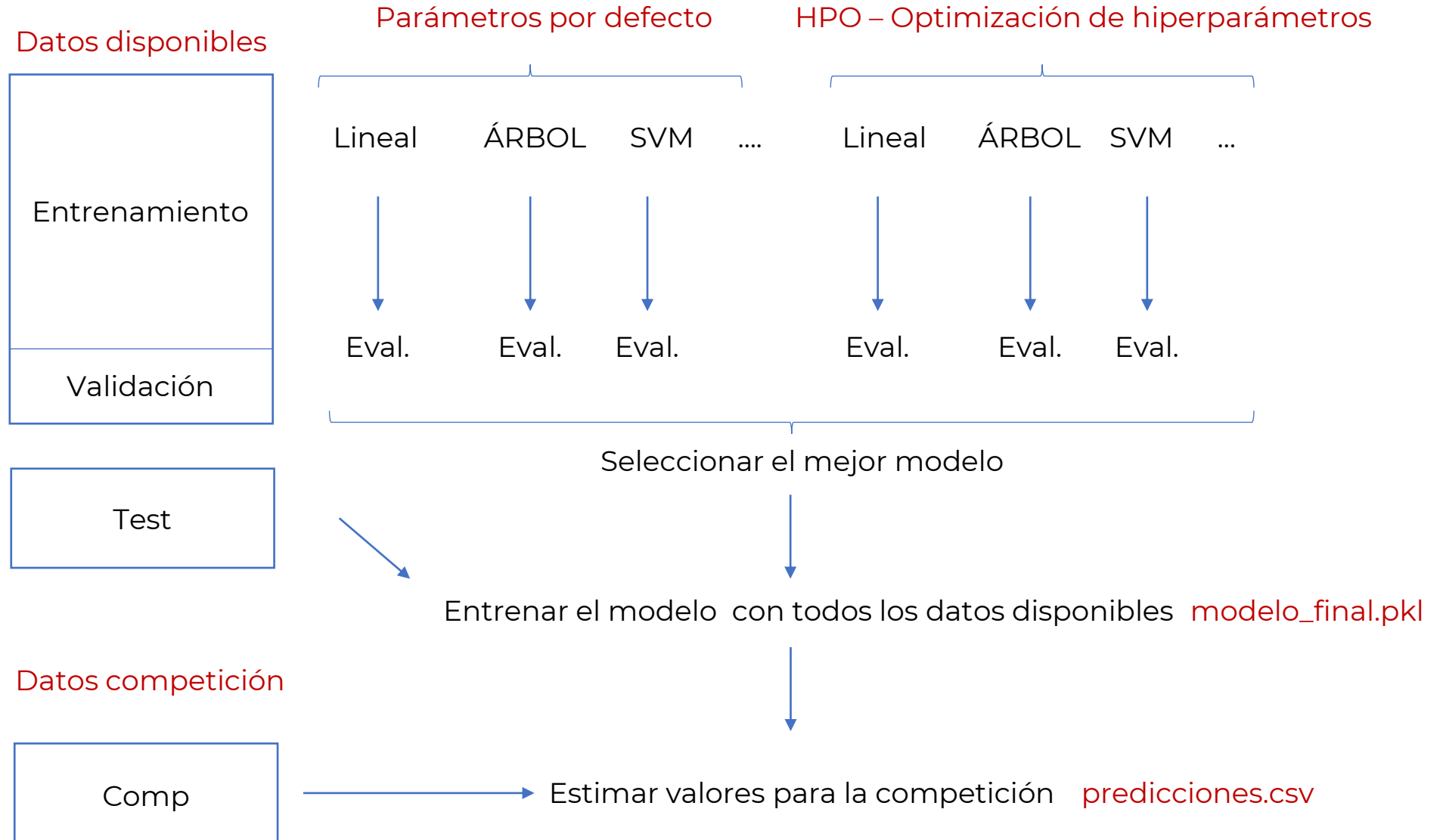
Práctica 1

Ficheros

- Se proporcionan dos archivos de conjuntos de datos:
 - **wind_available.csv.gzip**: datos disponibles para HPO, evaluación del modelo y entrenamiento del modelo final.
 - **wind_competition.csv.gzip**: datos de la competición sin variable de respuesta. Este conjunto de datos sirve para utilizar el modelo final para hacer predicciones sobre nuevos datos.

```
wind_ava = pd.read_csv('wind_available.csv.gzip', compression="gzip")
```


Práctica 1



Práctica 1



- Commit semanal.
- Sólo un repositorio por grupo de prácticas
- **Repositorio privado**
- Nombre repositorio: "GrupoXX_Practica1"
- Invitar al profesor como "Colaborador"
"mpatrici UC3M"
- Enviar el enlace del repositorio:
mpatrici@inf.uc3m.es