

#SEHACKATHON 2022

Luis García Blanco, Jorge Fernández
Simancas y Mateo Bermann Albalat

Preparación de los datos

CSV

Descargar desde la página del challenge

Unir los dos datasets

+

=

Unir el dataframe y eliminar las columnas con un único valor

JSON

Importar desde URL

Columnas de más: códigos y nombres de la misma información

Modelización

Proceso iterativo:
validación cruzada

Regresión logística, random forest y máquinas de soporte de vectores rechazados por tener un alto nivel de error

Redes neuronales como solución óptima

Predicción

Mediante el software Orange, se ha entrenado una red neuronal y con los datos train_x se ha calculado la predicción final