



DIARI DEL PROJECTE



11/02/2025

Planificació del projecte i tecnologies a útlilitzar.



PLANIFICACIO

La idea es obtindre en webscraping una datases per a entrenar la IA. Poducte: Mobils Dades a recollir: Marca,Modelo,Fecha de lanzamiento,Precio de venta,Procesador,RAM (GB),Almacenamiento (GB),Batería (mAh),Cámaras traseras,Resolución cámara principal (MP),FPS cámara trasera,Resolución cámara frontal (MP),FPS cámara frontal,Tamaño pantalla (pulgadas),Resolución pantalla,Frecuencia actualización (Hz),Precio de lanzamiento Despues de recollir les dades necessitarem fer webscraing per a veure quines dades podem obtindre , també hem de determinar quin resultat volem obtindre esa dir si volem un true o false per a si es una bona resposta o no o volem un percentattje de quan bona oferta es Target: De moment nem ha deixar-ho en True o False per a ssaber si es o no una bona oferta en un futur es valorara fer per percentatge de bona oferta. Despues de tindre el dataset entrenarem la ia, de moment nem a centrarmos en el dataset.

Nem a intentar fer webscraping de amazon , phonearea i si no de mercado libre Hem decidir gastar la api de rapidapi per a poder fer web scraping de amazon

Coses a fer:

- Web scrapping
- Pujar dataset a mongo
- Crear un model per a entrenar
- Crear una api
- Crear el frontend pogent fer peticio a la api



TECNOLOGIES

Python BBDD : MongoDB Interficie: Flet Llibreries IA: (sckitlearn / keras)

Coses a fer:

- Web scrapping
- Pujar dataset a mongo
- Crear un model per a entrenar
- Crear una api
- Crear el frontend pogent fer peticio a la api




GUIA DE FICHEROS

- Scrapping (sacamos ID, convertimos los ID en json)
- Limpieza (limpiamos JSON, lo convertimos a CSV)
- Preprocesamiento y entrenamiento (preparamos los datos, entrenamos el modelo y lo guardamos)
- Interfaz del usuario (el usuario introduce un movil, cargamos el modelo y hacemos la prediccion)

- Res: ids.json - mobile_raw.json - mobile_clean.json - mobile.csv

PD: Subir a mongo: - JSON Raw - JSON Clean - CSV

 12/02/2025

Gerard: Començar la interfície amb flet Andreu: Obtencio de els productes

COLUMNAS CSV

- product_id
- precio_actual
- precio_anterior
- Tipo de pantalla
- Tipo de procesador
- Ano del modelo
- Memoria RAM
- Tecnologia de pantalla
- Sistema operativo
- Capacidad de la memoria flash instalada
- Fabricante
- Resolucion máxima
- Numero de procesadores
- Velocidad del procesador
- Frecuencia de actualización
- Marca
- Capacidad de almacenamiento digital
- Capacidad de la memoria
- Capacidad de la memoria RAM instalada
- Tecnologia GSM
- Dimensiones del producto
- Relacion de aspecto -Tamano de la pantalla

POSIBLES MODELOS IA

SKLEARN

- LinearRegression
- DecisionTreeRegressor
- RandomForestRegressor 🏆
- GradientBoostingRegressor 🏆
- Regresión con Support Vector Machines (SVR)

RESULTADOS ESPERADOS

- R2 Score Proximo a 1 (0,80 ja es bo)
- MAE Lo mas bajo posible
- MSE Lo mas bajo posible

MILLOR MODEL ACTUAL

RANDOM FOREST AMB MODA EN MARCA

- R2 de 0,23
- MAE de 85
- MSE de 23343

TO DO ACTUAL

- `"/training/TrainModel.ipynb"` Ajustar modelo
- Detectar columnas que no interesen para eliminarlas de dataset
- Reorganitzar Repositri
- `"/FlasckServer/app.py"` Crear una api pa el link de amazon