

Oportunidades de importación en el mercado europeo de coches de segunda mano para marcas premium.

Objetivo

El objetivo de este proyecto es buscar oportunidades de importación en el mercado europeo de coches de segunda mano para marcas premium. (Mercedes-Benz, Audi, BMW, Porsche y Land Rover)

Resumen

Comienza con la adquisición de los datos, ofertas de coches de segunda mano, tanto del mercado español como del europeo. A continuación, se manipulan los datos para poder trabajar con ellos.

Una vez que los datos están listos para ser utilizados se cogen las ofertas del mercado español y se entrena un algoritmo de Machine Learning para que sea capaz de predecir el precio de un coche de segunda mano en el mercado español. Por último, el algoritmo de Machine Learning predice el precio de ofertas europeas de coches de segunda mano en el mercado español.

Aquellas ofertas que presenten una mayor diferencia entre su precio real en Europa y su precio predicho para España serán las mejores oportunidades de importación.

Etapas

Adquisición de datos

Las ofertas de coches de segunda mano se han obtenido haciendo Web Scraping del portal <https://www.autoscout24.es/>. Es uno de los que más ofertas tiene tanto a nivel nacional como a nivel europeo. Se pueden encontrar los notebooks que realizan esta tarea en la carpeta (Web_Scraping). Los datos se almacenan en archivos CSV en la carpeta Raw_Data.

Manipulación de datos

Las ofertas contenidas en 94 ficheros CSV se manipulan y transforman para poder explotar los datos de forma eficiente en algoritmos de Machine Learning. Cada marca de coches es tratada por separado con el objetivo de recategorizar los modelos de forma correcta. Se pueden encontrar los notebooks que realizan esta tarea en la carpeta (Data Manipulation). Los datos se almacenan en archivos CSV en la carpeta Clean_Data.

Exploración y visualización de los datos

Esta etapa sirve para familiarizarse con los datos y ver qué relación tienen las distintas variables extraídas para cada oferta con la variable objetivo a predecir, el precio. Se puede encontrar el notebook que realiza esta tarea en la carpeta (Data Exploration and Visualization).

Modelos de Machine Learning

Se busca el mejor algoritmo de Machine Learning para predecir los precios de los coches de segunda mano en el mercado español. También se prueban varios tratamientos previos a las variables. El algoritmo elegido es el GradientBoostingRegressor de Sklearn. Se pueden encontrar los notebooks que realizan esta tarea en la carpeta (Machine_Learning_Models).

Oportunidades de importación

Se utiliza el GradientBoostingRegressor de Sklearn para predecir el precio de las ofertas europeas en el mercado español. Estas son ordenadas en función de la diferencia entre su valor real europeo y su predicción en el mercado español. Además, estas son segmentadas en función de su precio real para poder ajustarse a la inversión deseada. Se puede encontrar el notebook que realiza esta tarea en la carpeta (Import_Oportunities).

Resultados

Se revisan cada una de las ofertas a mano analizando si se trata de ofertas reales o si se tratase de algún tipo de estafa o similar. Por último, se entregan cuatro ficheros CSV con las mejores ofertas ya revisadas.

Conclusiones

Existen actualmente oportunidades de importación en el mercado de coches de segunda mano europeo. Las mejores oportunidades de inversión se encuentran en un conjunto de coches caracterizado por ser coches de gran tamaño, pocos kilómetros, gran potencia y una antigüedad de entre 2 y 6 años. Esto es debido a que en Europa sufren una depreciación más rápida.