**PROYECTO INTEGRADOR - PREDICCIÓN DE PRECIOS DE AUTOMÓVILES**

**Descripción**

Este proyecto tiene como objetivo desarrollar un sistema de predicción de precios de automóviles utilizando técnicas de Machine Learning. El proyecto se divide en tres fases principales:

1. **Análisis Exploratorio de Datos (EDA)**: Exploración del dataset para entender las relaciones entre las variables y el precio de los automóviles, detectar outliers, valores faltantes y patrones.
2. **Preparación de Datos**: Limpieza, transformación y codificación de las variables, así como el manejo de outliers y la estandarización de datos.
3. **Modelado y Evaluación**: Creación de dos modelos predictivos: uno de clasificación para identificar autos caros y baratos, y otro de regresión para predecir el precio exacto.

**Archivos**

* **cleaned\_dataset.csv**: Dataset limpio utilizado para el análisis y la modelización.
* **ML\_cars.csv**: Archivo adicional con datos para las fases de modelado.
* **Limpieza\_datos copy.ipynb**: Jupyter Notebook donde se realiza la limpieza y preparación de datos.

**Tecnologías Utilizadas**

* **Python**: Lenguaje de programación utilizado para todas las fases del proyecto.
* **Pandas**: Manipulación y análisis de datos.
* **NumPy**: Operaciones matemáticas y de álgebra lineal.
* **Matplotlib y Seaborn**: Visualización de datos durante el EDA.
* **Scikit-learn**: Creación de los modelos de Machine Learning.

**Fases del Proyecto**

**1. Análisis Exploratorio de Datos (EDA)**

Durante esta fase, realicé la exploración de las variables más relevantes del dataset. Algunos de los gráficos clave incluyen:

* Distribuciones de precios y características numéricas.
* Identificación de outliers mediante boxplots.
* Relación entre las variables numéricas y categóricas con el precio.

**2. Preparación de Datos**

En esta fase, se llevaron a cabo las siguientes tareas:

* Manejo de valores faltantes.
* Codificación de variables categóricas.
* Estandarización de las variables numéricas.
* Tratamiento de outliers.

**3. Modelado y Evaluación**

Creé dos modelos:

* **Modelo de Clasificación**: Dividí los autos en "baratos" y "caros" usando la mediana de los precios como punto de corte. Utilicé un modelo de clasificación como Random Forest.
* **Modelo de Regresión**: Para predecir el precio exacto del vehículo, entrené varios modelos de regresión, evaluando su rendimiento con métricas como RMSE y MAE.

**Resultados**

El modelo de clasificación logró una precisión del X% para identificar autos caros y baratos, mientras que el modelo de regresión tuvo un error medio absoluto (MAE) de X.