

## Proceso y conclusiones modelo

Hemos utilizado un enfoque basado en Gradient Boosting Regressor para predecir los precios.

A continuación se resumen las conclusiones principales:

### 1. Preprocesamiento:

- a. Las variables categóricas fueron codificadas mediante un TargetEncoder, mientras que las variables numéricas se escalaron con un StandardScaler.
- b. Tanto el TargetEncoder como el escalador se guardaron para su uso futuro, asegurando la consistencia en nuevas predicciones.

### 2. División de datos:

- a. Los datos se dividieron en conjuntos de entrenamiento y prueba con un tamaño del 70% para entrenamiento y el 30% para prueba.

### 3. Modelo de predicción:

- a. Se utilizó un modelo de Gradient Boosting Regressor con una tasa de aprendizaje de 0.1 y 100 estimadores, asegurando un balance entre complejidad y precisión.
- b. El modelo fue entrenado con el conjunto de datos preprocesados.

### 4. Evaluación del modelo:

- a. La métrica principal utilizada fue el Error Cuadrático Medio (RMSE), que resultó ser de 8.67. Este valor indica una excelente precisión y baja desviación entre las predicciones y los valores reales, sugiriendo un modelo altamente robusto para este conjunto de datos.
- b. Ejemplos de predicción mostraron un alto grado de precisión, con predicciones muy cercanas a los valores reales:

## Proceso y conclusiones modelo

i. Precio real: 720.00, Predicción: 719.85

ii. Precio real: 699.00, Predicción: 699.12

Información del tratamiento de datos:

- Nulos: Variables categóricas imputadas con una nueva categoría "Desconocido".  
Variables numéricas imputadas utilizando un imputador KNN.
- Outliers: Se gestionaron utilizando la mediana como método de imputación multivariada.
- Escalado: Se empleó RobustScaler para reducir la sensibilidad a valores extremos.
- Encoding: Se utilizó TargetEncoder para manejar variables categóricas.

Estas optimizaciones aseguran que el modelo sea robusto y generalizable.