



REGRESSION PROJECT

Paula Naranjo y Ana García



OBJETIVO DEL PROYECTO

Debemos construir un modelo que prediga el precio de una casa basándose en las características proporcionadas en el conjunto de datos.

También se quiere explorar las características de las casas utilizando algunas parcelas.

En nuestro caso hemos buscado una propiedad con unas características concretas.

Vamos a utilizar la regresión lineal para construir un modelo.



Proceso de la búsqueda del modelo

Para poder buscar unas condiciones en nuestro dataset de casi 22000 datos y averiguar el precio debemos seguir un proceso:

1. Cargar los datos del dataset
2. Limpiar el dataset, eliminando las columnas con poca correlación o nula y los que tenían muchos outliers.
3. Analizar los datos de cada columna
4. Hacer uso del método de descripción



Proceso de la búsqueda del modelo

1. Realizar una gráfica usando matplotlib y seaborn
2. Con unos datos concretos imprimir el precio de dicha propiedad
3. Comparamos las métricas de error en los conjuntos 'train' y 'test' para asegurarnos de que el modelo no sufre sobreajuste o un ajuste insuficiente.



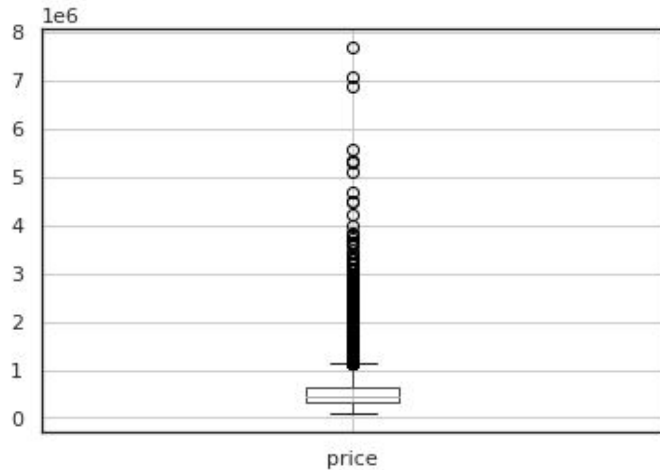
Conclusiones del modelo

Como hemos realizado un modelo lineal, el error es mayor, por lo que no podemos predecir con exactitud

Con un modelo no lineal podríamos predecir mejor el modelo ya que los datos son muy variados y es difícil aproximarlos con una línea recta.

Dependiendo del precio de la casa, el error obtenido de 200000E es mejorable, pero entra dentro del rango establecido puesto que la diferencia entre test y train es muy pequeña.

Desviación Estándar



La desviación estándar es

$3.673681e+05$

Si comparamos este dato con la cantidad de datos obtenidos la solución es óptima