

Análisis Precio de la Vivienda en Boston con Regresión

Introducción

La predicción del precio de viviendas es un desafío crucial en el mercado inmobiliario, y para abordarlo utilizaremos varios métodos de regresión aplicados al conocido conjunto de datos de casas en Boston. Emplearemos técnicas como la regresión lineal, Random Forest, aprendizaje profundo (deep learning) y Máquinas de Soporte Vectorial (SVM). Compararemos estos métodos en términos de precisión y capacidad de generalización, evaluando su rendimiento mediante métricas como el error cuadrático medio (MSE) y el coeficiente de determinación (R^2), con el objetivo de identificar el enfoque más eficaz para predecir los precios de las viviendas en esta área.

Proceso del Análisis

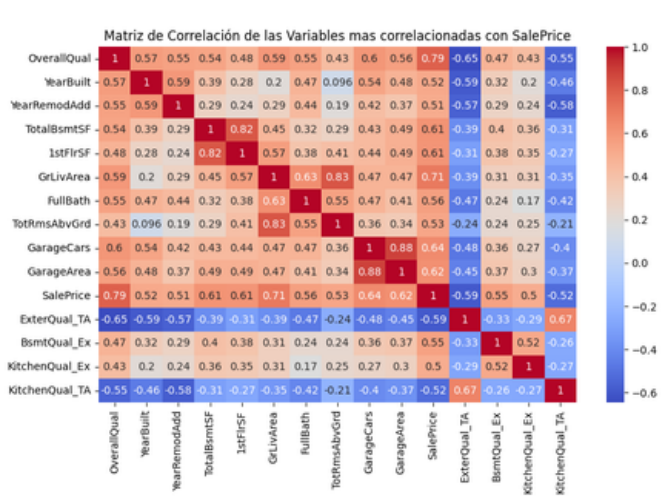
Primero se realizo una limpieza de los datos nulos da base de datos para luego realizar una dumificación de las variables categóricas y finalmente una estandarización de la variables

Luego por medio de la matriz de correlación y una selección de importancia de las variables realizada con un random forest obtuvimos las mejores variables para entrenar el modelo

Finalmente con las variables elegidas con el método se realizo una regresión lineal, un SVM y un deep Learning, con los datos obtenidos de estos entrenamientos se usaron métricas como el r cuadrado, error absoluto medio y el r cuadrado ajustado para ver la eficacia de cada modelo



Matriz de Correlación



Variables elegidas y su importancia

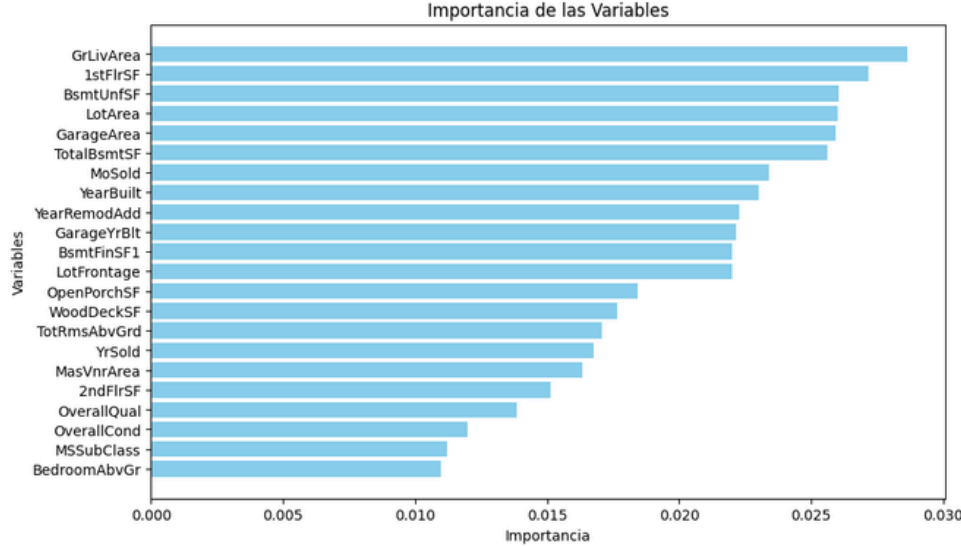
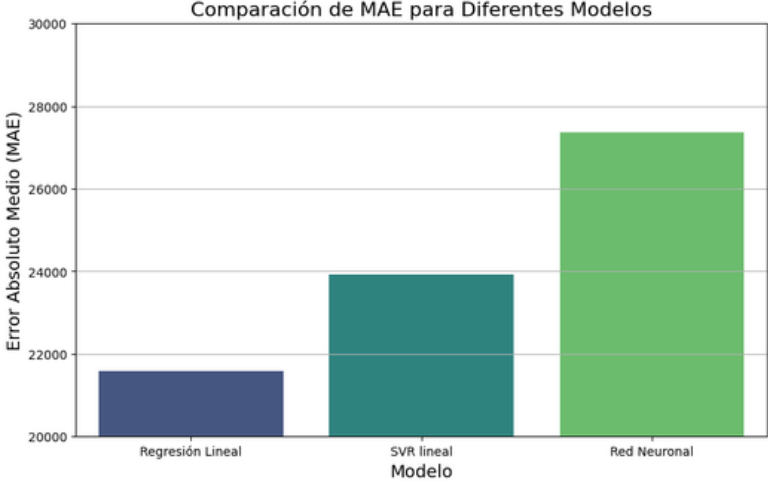
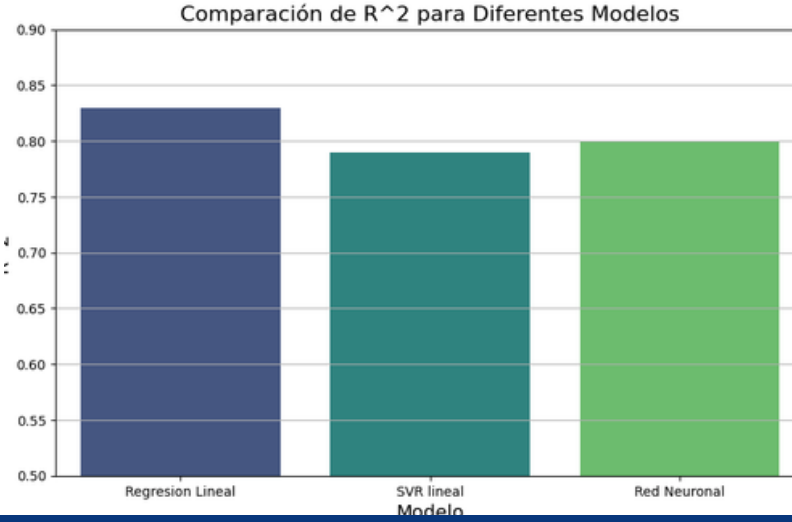


Tabla Comparativa de Métricas

Método	Error Cuadrático Medio (MSE)	Error Absoluto Medio (MAE)	Coefficiente de Determinación (R^2)
Regresión Lineal	1,509,159,116.10	21,577.74	0.8348
SVM	1,863,134,208.74	23,926.07	0.7961
Redes Neuronales	1,976,956,485.01	27,362.01	0.7836

Comparación de Métricas Financieras



Conclusiones

Regresión Lineal es el modelo más eficaz en términos de precisión y capacidad explicativa de la variabilidad en los precios de las viviendas, seguido por SVM y Redes Neuronales.

SVM ofrece un rendimiento decente, pero no tan bueno como la regresión lineal.

Redes Neuronales, a pesar de su capacidad para modelar relaciones complejas, no superan a los otros modelos en este caso específico, mostrando los mayores errores de predicción.

Estos resultados sugieren que, para el conjunto de datos y el problema específico de predicción de precios de viviendas en Boston, la regresión lineal es la opción más recomendable.