

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE OCCIDENTE

DEPARTAMENTO DE PROCESOS TECNOLÓGICOS E INDUSTRIALES

PPMD



ITESO

Universidad Jesuita
de Guadalajara

Examen 3:

PRESENTA:

**Jesus Vargas Pacheco
Jorge Ponce
Angel Aceves**

23 de noviembre del 2025. ITESO, Tlaquepaque, Jalisco

Trabajamos con el conjunto de datos Saratoga Houses, que contiene información real de viviendas vendidas (tamaño del terreno, área habitable, edad, número de cuartos, valor del terreno, servicios, etc.) y el precio de venta como variable objetivo. El objetivo fue construir un modelo de regresión que predijera el precio de la casa a partir de sus características.

Después del preprocesamiento (codificación de variables categóricas, estandarización y aplicación de PCA para reducir dimensionalidad y controlar la colinealidad), el modelo seleccionado como ganador fue una regresión lineal con regularización Elastic Net.

Métricas en entrenamiento y prueba

El desempeño del modelo se resume así:

Entrenamiento

MAE_train ≈ 42,829.04 USD

RMSE_train ≈ 60,749.77 USD

R²_train ≈ 0.6397

Prueba / validación

MAE_test ≈ 40,053.03 USD

RMSE_test ≈ 53,607.81 USD

R²_test ≈ 0.6402

En general, los errores en prueba son del mismo orden que en entrenamiento, lo que indica que el modelo generaliza razonablemente bien y no está sobre ajustando de forma grave.

Evaluación Final (Datos externos)

MAE \approx 50,271.28 USD

RMSE \approx 60,444.01 USD

MAPE = 21.96%

Conclusiones

Estas cifras indican que el modelo comete, en promedio, errores de unas 5×10^4 USD por vivienda y se equivoca alrededor de una quinta parte del valor real. Para una casa típica del conjunto de prueba, cuyo precio promedio ronda los $\approx 247,000$ USD (según los datos de Saratoga_Precios_Reales.txt), esto implica un error esperado cercano a 54,000 USD, lo cual es consistente con el MAE y el MAPE reportados.

Los errores son grandes en términos absolutos (porque los precios de vivienda también lo son), pero son razonables para un modelo lineal aplicado a un mercado con alta variabilidad. Además, el hecho de que el MAE y el RMSE estén relativamente alineados sugiere que, aunque hay algunas casas con errores más altos, el modelo no está totalmente dominado por outliers extremos.