

Modelo de Machine learning para predicción de precios de apartamentos en venta

HabitAlpes

Entregado por: Lina Bejarano y Juan Diego Enriquez

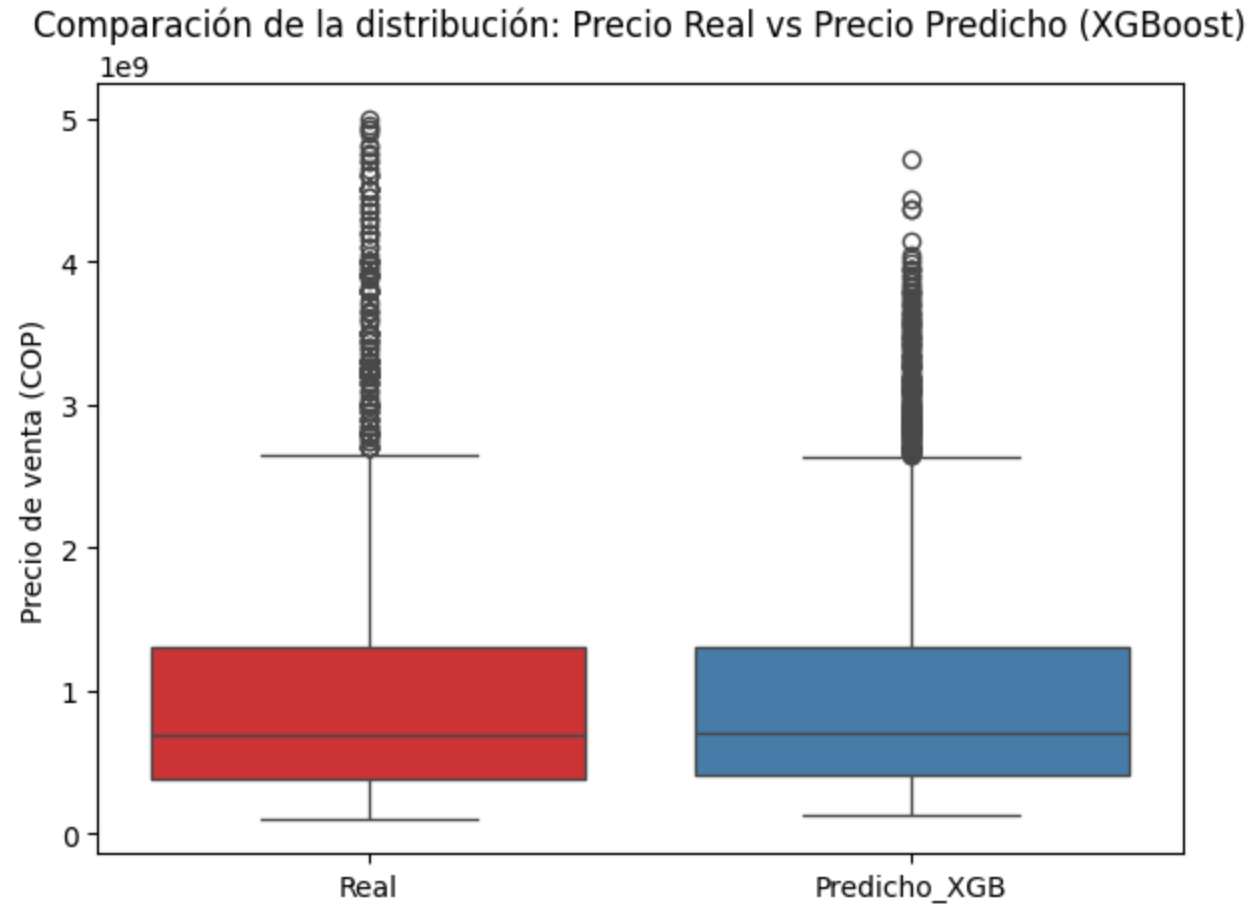
Contexto

- HabitAlpes evalúa un total de 500 apartamentos al mes, para ello, invierte 28.500.000 pesos en gastos de peritaje para determinar el precio de venta de apartamentos. Dado esto, esta interesado en el diseño de un modelo de Machine Learning para predecir el valor de los apartamentos a partir de las características que ingresan los clientes en una plataforma. Lo cual llevará a disminución de costos y tiempo de peritaje.

Insights

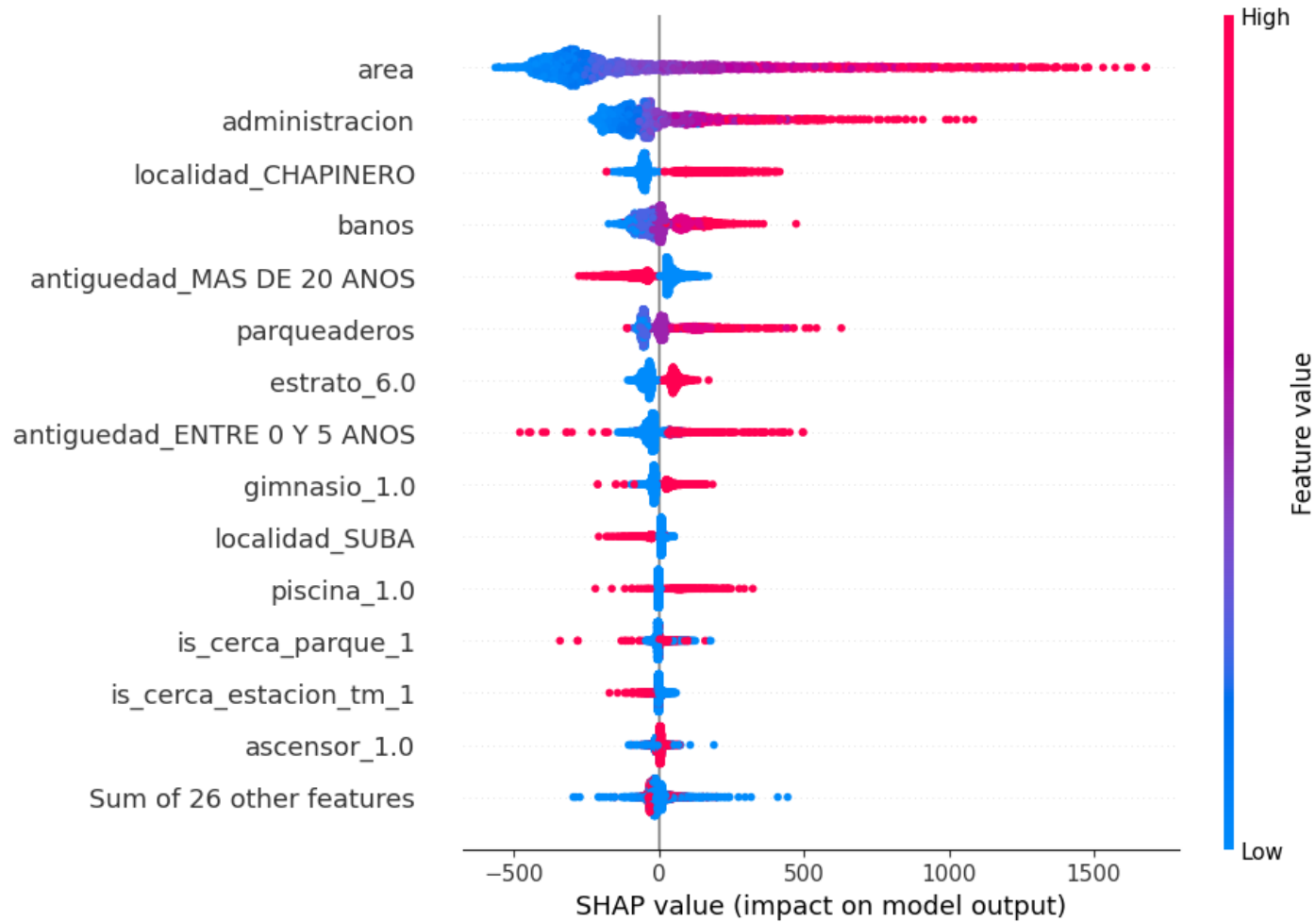
- Se evaluaron 2 modelos de machine learning de regresión: Regresión lineal y XGBoost-regresor para predicción de valor de venta. Entre estos modelos el que mejor métricas arrojó con el conjunto de validación, fue XGBoost. Este modelo presento un R^2 de 0.85, lo que indica que explica el 85% del precio de venta de los apartamentos.

Insights



El modelo de ML (XGBoost) para la predicción de valor de venta de apartamentos, logró presentar predicciones cercanas al valor real como se observa en los boxplot de la imagen. Lo que quiere decir que presenta buenos rendimientos en términos de predicción.

Insights



Las características que más aportan al precio de venta de los apartamentos son: área, valor de administración, # de baños, antigüedad, estrato y localidad.

Insights-Generación de valor

A partir de datos históricos se asume que:

- Si el modelo subestima el precio en más de \$20 millones, el cliente solicitará un avalúo presencial.
- Las estimaciones que sobreestiman el valor no generan reclamo.

Dado esto, en la validación del modelo se encontró que un 35.8% se requiere un avalúo presencial por el desfase presente entre el valor real y el valor predicho. Esto sugiere que el ~64% no requiere de ello y va a presentar un ahorro en peritaje. Además, este valor destaca un buen rendimiento del modelo.

Insights-Generación de valor

- El modelo reduce costos de peritaje en 15.245.404 pesos al mes. Que representa una disminución del 53,4% del valor que se usa sin modelo (\$28.500.000).
- Este modelo, permite una reducción de 30.491 pesos promedio por avalúo. Lo que representa un valor de \$13.254.596 para predicción de valor de apartamentos al mes.
- El modelo en general, requiere una inversión de 52 millones de pesos en infraestructura y remuneración al científico de datos. Con la reducción anual de ~182 millones de pesos, esta inversión se recupera en menos de un año, y en un año, puede retornar valores de ahorro en 3.5 veces la inversión del modelo de acuerdo con el valor ROI.

Recomendaciones

- Mejorar y ampliar la base de datos utilizada para el entrenamiento del modelo, podría incrementar la precisión en la estimación del precio de venta.
- Evaluar la escalabilidad del modelo, en como puede adaptar la incorporación de los datos.
- Con el pasar del tiempo se va a requerir monitoreo del modelo, por lo que se sugiere hacer un estudio al respecto para incluir el valor de esta actividad en la inversión del modelo.