

INFORME EJECUTIVO

1. Contexto de negocio

El mercado inmobiliario colombiano representa más del 20% del PIB nacional y es un sector intensivo en capital y empleo. En este contexto, HabitAlpes busca ofrecer a propietarios y a entidades públicas herramientas analíticas para estimar de forma ágil y consistente el precio de venta de apartamentos en Bogotá, reduciendo la dependencia en avalúos totalmente manuales y mejorando la experiencia del cliente.

Actualmente, el proceso de avalúo presencial requiere en promedio 6 horas de trabajo de un perito (a 9.500 COP por hora) por cada apartamento. HabitAlpes desea apoyarse en un modelo de Machine Learning que entregue una estimación inicial de precio, de modo que el experto solo valide y ajuste el resultado, reduciendo su dedicación a 1 hora por caso. La capacidad operativa de la empresa es de hasta 500 apartamentos al mes.

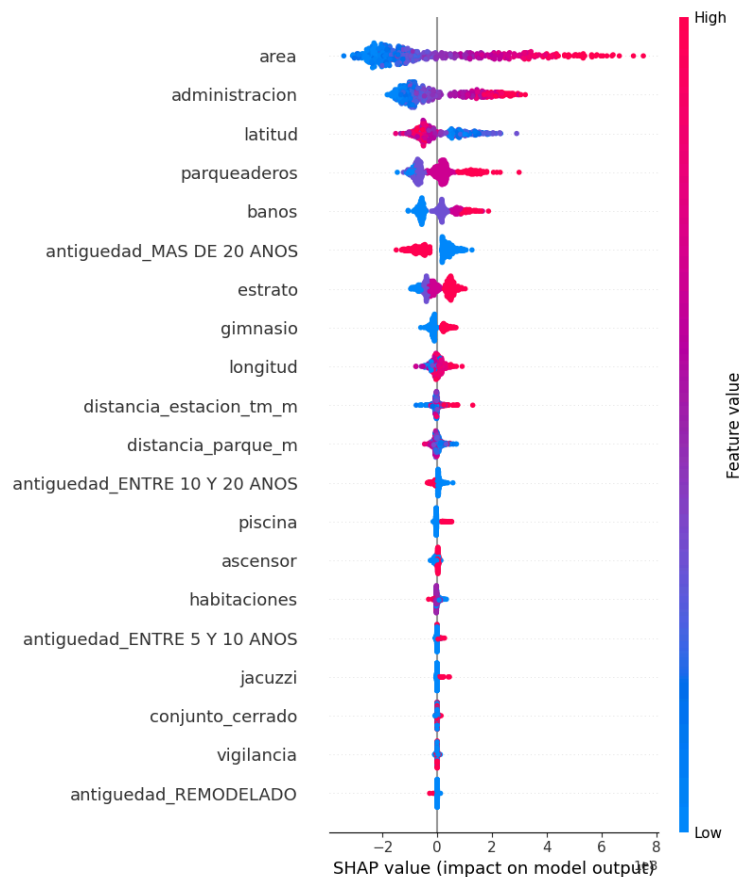
2. Datos y preparación - Insights Iniciales

El dataset original contiene 46 variables para anuncios de apartamentos publicados en los últimos dos meses. Se eliminaron registros cuyo tipo de inmueble no corresponde a apartamentos (por ejemplo, casas en conjunto), dado que el servicio objetivo es exclusivo para apartamentos. Se descartaron registros de tipo de operación "ARRIENDO" para concentrarse únicamente en precio de venta. Se observaron valores negativos y montos extremos sin sentido para el mercado bogotano, por esta razón, se fijó un límite inferior operativo de 150.000.000 COP, eliminando valores por debajo de este umbral. Se logró identificar áreas negativas y superficies irrealmente grandes; se retuvieron únicamente apartamentos con área mayor o igual a 50 m². Se priorizaron las variables de distancia (por ejemplo, distancia_parque_m, distancia_estacion_tm_m) frente a sus equivalentes categóricas binarias (is_cerca_parque, is_cerca_estacion_tm), al aportar mayor granularidad. Dada la alta correlación entre localidad y (latitud, longitud), se decidió conservar coordenadas geográficas y eliminar la redundancia en variables de ubicación, manteniendo aquellas de mayor valor explicativo (barrio, sector).

3. Resultados cuantitativos mejor modelo

Se probó un conjunto de cuatro modelos de regresión lineales y no lineales. El mejor modelo por sus métricas en los datos de test fue Random Forest. En promedio, el modelo se equivoca en 100,290,598 millones de pesos al estimar el precio de un apartamento. Dado que el precio promedio del dataset es de 775,800,559 millones de pesos, este error corresponde aproximadamente al 12.04% sobre el valor del inmueble. El modelo explica el 91,4% de la variabilidad de los precios de los apartamentos. Esto implica que las características incluidas capturan muy bien los factores que afectan el precio real.

4. Interpretabilidad del modelo

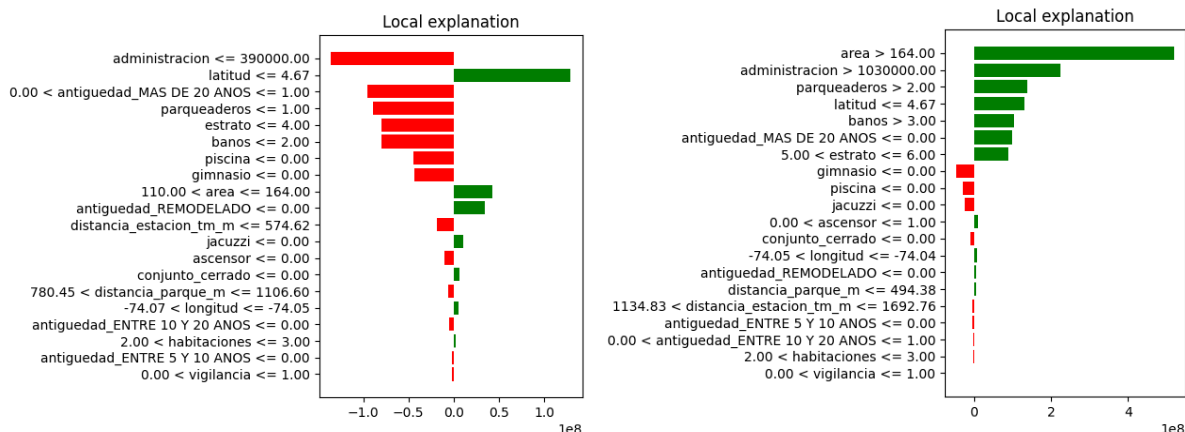


Análisis global (SHAP):

- Las variables con mayor impacto global en las predicciones son:
 - Área del apartamento.
 - Cuota de administración.
 - Latitud
 - Número de baños y habitaciones.
 - Número de parqueaderos.

Mayor área, estrato alto, más parqueaderos y más baños se asocian con incrementos en el precio estimado. Apartamentos con antigüedad baja o recientemente remodelados tienden a recibir valores más altos. Distancias menores a estaciones de transporte masivo y a parques se relacionan con aumentos en el valor, coherente con la lógica de valorización urbana en Bogotá.

Análisis local (LIME):



En ejemplos individuales analizados, el modelo combina de forma coherente los efectos de administración, área, estrato, parqueaderos y antigüedad. Por ejemplo, la ausencia de conjunto cerrado, distancias intermedias a transporte y niveles de antigüedad desfavorables aparecen como factores que reducen el precio estimado respecto a un apartamento de referencia, mientras que área y estrato altos suelen ejercer efectos positivos significativos. Esto respalda que el modelo no solo tiene buen desempeño agregado, sino que sus justificaciones caso a caso son razonables para un analista inmobiliario.

5. Generación de valor económico

El ahorro por estimación es de 28,117 COP, teniendo en cuenta que son 500 apartamentos al mes, el ahorro esperado es alrededor de 14,1 millones de COP. Además, se calcula una precisión mínima requerida de aproximadamente 0.16 (16%) para que el modelo, aun considerando subestimaciones costosas, no destruya valor. El modelo actual logra $I \approx 0.66$ (66%), es decir, más de cuatro veces la precisión mínima necesaria, lo que indica que el esquema propuesto sí genera valor económico. Asumiendo un costo de proyecto (desarrollo + despliegue + integración) de 20.000.000 COP, el ROI para HabitAlpes es de 1,42 meses. Es decir, en alrededor de mes y medio de operación, el proyecto recuperaría la inversión inicial bajo los supuestos de volumen (500 estimaciones/mes) y desempeño del modelo observados en la validación.

6. Recomendaciones para HabitAlpes y conclusiones

- Recomendaciones

1) Implementar el Random Forest como primera versión del motor de precios

El modelo presenta buen desempeño y un comportamiento interpretable con SHAP y LIME.

Mejora de forma consistente frente a alternativas lineales y otros ensambles probados.

Inclusive económicamente tiene un ahorro con un impacto significativo

2) Gobernanza y actualización del modelo

- Actualizar el modelo con datos nuevos al menos cada 6–12 meses para capturar cambios de mercado.

El modelo construido demuestra que HabitAlpes puede automatizar una parte significativa del proceso de valoración inmobiliaria manteniendo precisión analítica y reduciendo costos operativos. Con un MAE equivalente al 12% del valor del inmueble, un R^2 de 0,915 y un nivel de precisión aceptable ($I = 0,66$) muy superior al mínimo requerido para generar valor (0,16), el modelo genera ahorros mensuales estimados de alrededor de 14 millones de pesos.

Esto permite recuperar la inversión inicial en aproximadamente mes y medio, lo que valida la viabilidad económica del proyecto. Implementar el Random Forest como solución inicial, complementada con mecanismos de supervisión, monitoreo continuo y actualizaciones periódicas, permitirá a HabitAlpes ofrecer un servicio más rápido, consistente y competitivo en el mercado inmobiliario. Además, sienta las bases para el desarrollo de nuevos servicios analíticos que pueden fortalecer el portafolio de la empresa y ampliar sus oportunidades de crecimiento.