

Estimación Automatizada de Precios Inmobiliarios

Solución de Machine Learning para HabitAlpes

Francisco Alejandro Santamaría

Informe Ejecutivo | Noviembre 2025

Desafío de negocio



Problema

Avalúos tradicionales con alto costo operativo:

- **600.000 COP** por inmueble
- Proceso manual que consume días
- Alta variabilidad en criterios de valoración
- Dificultad para estandarizar procesos



Solución

Modelo predictivo automatizado:

- Machine Learning con datos históricos
- Estimación rápida y consistente
- Alta precisión demostrable
- Explicabilidad mediante SHAP/LIME

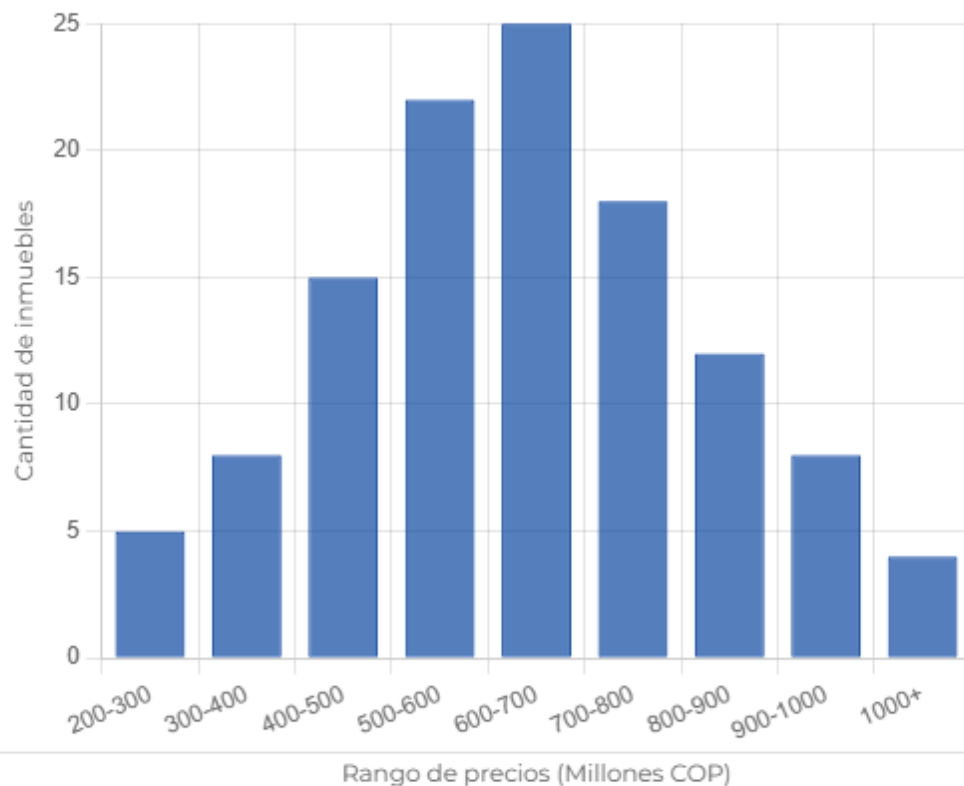


Alcance

Segmento inmobiliario objetivo:

- Apartamentos en Bogotá
- Foco en estratos 4-6
- Usaquén, Chapinero, Suba (>80% muestra)
- Áreas entre 70-180 m²

Distribución de precios de venta (M COP)

**Precio de venta**

Mediana cercana a 600 M COP, con la mayoría de propiedades entre 400–900 M COP.

**Áreas típicas**

Entre 70 y 180 m², con valores extremos poco frecuentes.
Concentración en segmento medio-alto.

**Precio por m²**

Concentrado alrededor de 6–7 M COP/m², con valores mayores en estratos más altos.

**Concentración geográfica**

Usaquén, Chapinero y Suba agrupan más del 80% de las observaciones analizadas.



75M COP

Error medio absoluto (MAE)
≈ 12,5% del precio



92%

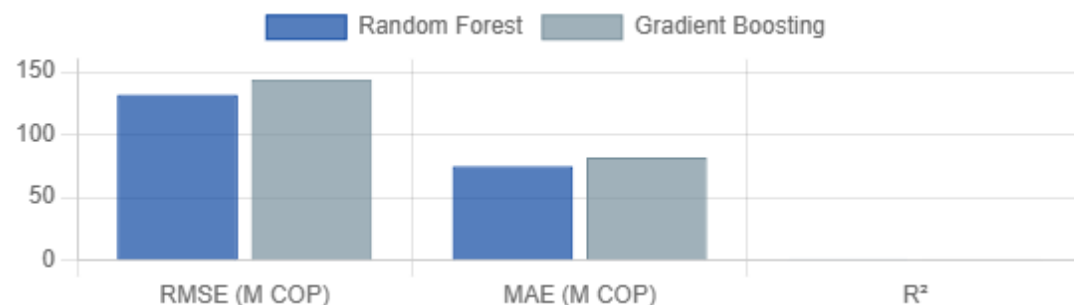
Precisión (R^2)
Varianza explicada



+8-9%

Mejora vs. alternativas
En MAE y RMSE

Comparativa de modelos



Random Forest supera a **Gradient Boosting** en todas las métricas: RMSE ≈ 132M vs 144M COP, MAE ≈ 75M vs 82M COP.

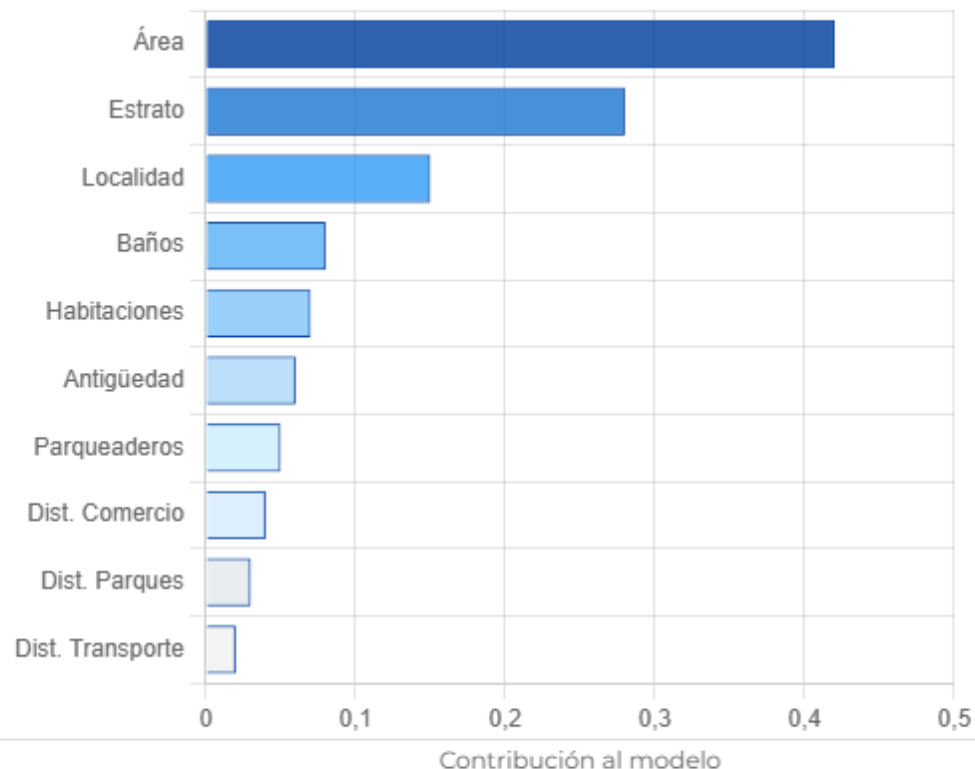


Validación cruzada muestra **estabilidad entre conjuntos** TEST y VALIDATION, sin señales de sobreajuste.



La **distribución de errores** se concentra alrededor de 0, confirmando la precisión del modelo en diferentes segmentos de precio.

Importancia de variables (valores SHAP)



Factores clave de valoración

Área, estrato y localidad son los principales determinantes del precio, seguidos por amenidades específicas.



Análisis SHAP

El área del apartamento tiene la mayor contribución positiva al precio. El impacto se intensifica a partir de 120m².



Efecto localidad

Usaquén y Chapinero muestran valores SHAP positivos, mientras que otras localidades presentan impacto neutral o negativo.



Explicaciones LIME

Cada estimación individual puede explicarse con factores específicos que aumentan o disminuyen la valoración del inmueble.

Propuesta de valor para HabitAlpes



Método Tradicional

Avalúo con perito externo

600.000 COP

- 🕒 **Tiempo de proceso:** Días de ciclo completo desde solicitud hasta entrega final.
- ⚡ **Consistencia:** Variabilidad alta dependiendo del perito y criterios subjetivos aplicados.
- 📈 **Escalabilidad:** Limitada por disponibilidad de peritos calificados y tiempos de respuesta.



Modelo Automatizado

Machine Learning + supervisión

50.000 COP

- ⚡ **Tiempo de proceso:** Minutos para generar estimación completa con explicabilidad incluida.
- ✓ **Consistencia:** Estimaciones estandarizadas bajo criterios objetivos y cuantificables.
- 🔧 **Escalabilidad:** Capacidad para procesar cientos de estimaciones sin incremento significativo de costos.



Ahorro neto por estimación:

500.000 COP (incluye costo esperado de errores de ≈50.000 COP)

HabitAlpes

Análisis financiero y ROI



Inversión inicial

70M COP



Punto de equilibrio

140 estimaciones



Proyección anual

1.000 estimaciones



Ahorro anual

500M COP

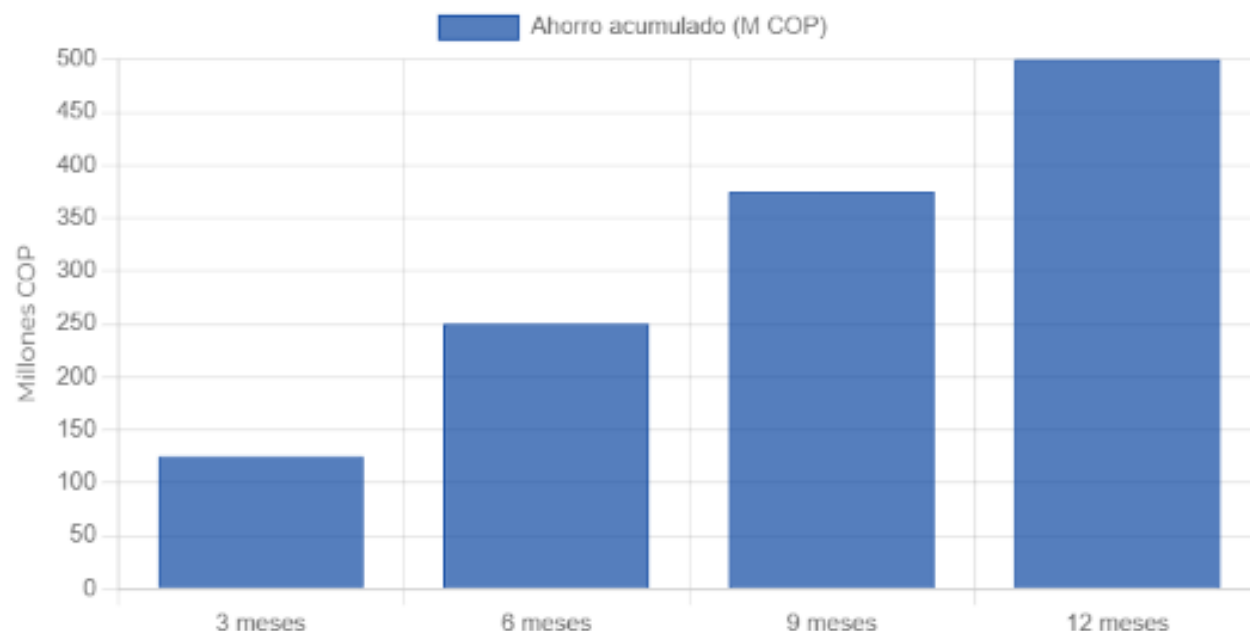


Retorno sobre la inversión (ROI) anual

600%

La inversión inicial se recupera en menos de 2 meses de operación con el modelo automatizado, generando un retorno 6 veces superior al capital invertido en el primer año.

Proyección de ahorro acumulado en M COP





Modelo y Operación

Estrategia predictiva y flujo operativo

3 recomendaciones

- ✓ **Adoptar Random Forest** como modelo base de estimación de precios, integrado al flujo de trabajo de HabitAlpes.
- ↔ **Esquema híbrido:** El modelo genera estimaciones y el perito revisa y ajusta sólo los casos con mayor error esperado o menor confianza.
- 📊 **Transparencia:** Utilizar SHAP y LIME para explicar estimaciones críticas a stakeholders (comerciales, clientes, aliados financieros).



Datos y Monitoreo

Fortalecimiento y supervisión continua

2 recomendaciones

- 🔧 **Reforzar datos:** Reinvertir parte del ahorro en mejorar la cobertura de datos (más zonas y estratos bajos) para ampliar alcance del modelo.
- 🔄 **Monitoreo continuo:** Implementar seguimiento de drift, recalibración periódica y retraining trimestral para mantener desempeño del modelo en el tiempo.



0-4 semanas

Corto plazo

- ✓ Integración MVP con sistemas actuales
- ✓ Piloto con 50 propiedades seleccionadas
- ✓ Definición de métricas de control
- ✓ Desarrollo de tablero SHAP para explicabilidad



1-3 meses

Mediano plazo

- ✓ Despliegue a equipos comerciales
- ✓ Capacitación para agentes comerciales
- ✓ Definición de umbrales de confianza para revisión manual
- ✓ Implementación de pipeline de datos automatizado



3-6+ meses

Largo plazo

- ✓ Expansión a otras ciudades y tipos de propiedad
- ✓ Optimización de features y desarrollo de modelos específicos
- ✓ A/B testing de modelos alternativos en producción
- ✓ Integración con plataformas externas de valoración