

PREDICCIÓN PRECIO DE VENTA DE VIVIENDA

IA para una evaluación precisa y eficiente



DAVIVIENDA



19 Agosto 2024



Indice

1. Estrategia:

Análisis del problema, enfoque metodológico e impacto esperado.

2. Proceso:

Limpieza de datos, EDA, feature engineer y desarrollo modelo.

3. Resultados:

Rendimiento del modelo, ejecución y costos y análisis de resultados.

4. Funcionalidad:

Valor para el negocio, adopción, usabilidad y escalabilidad.

5. Conclusiones:

Impacto, hallazgos claves, recomendaciones y próximos pasos.

6. Implementación:

Mejoras potenciales, viabilidad técnica/económica y consideraciones futuras.



ESTRATEGIA

EL DESAFÍO: Evaluación **Deficiente** de Viviendas

Procesos lentos, subjetivos y con riesgo a pérdidas

1. La evaluación manual es un proceso *lento y costoso*.
2. La subjetividad puede generar *inconsistencias e inexactitudes*.
3. Las evaluaciones inexactas impactan en la gestión de riesgos y la *rentabilidad*.



LA SOLUCIÓN: IA para la Evaluación Automatizada

Modelo K-Nearest Neighbors (KNN) para una predicción Precisa.

Se implementó un modelo KNN regresor, seleccionado por su superior desempeño en métricas como el MAPE, utilizando la librería Pycaret.

1. Se exploraron 18 modelos de ML (árboles, lineales, no lineales).
2. KNN Regresor obtuvo el MAPE más bajo (4.7543) en *validación cruzada* (10 folds).
3. Se evaluó con curvas de aprendizaje, residuos y feature importance (anexos al final).
4. Se exploraron modelos de ensemble sin mejoras significativas.

EL IMPACTO: Eficiencia, Precisión y Mejora en la Experiencia del Cliente

Beneficios tangibles para el negocio y nuestros clientes.

- Reducción del tiempo de procesamiento de créditos hipotecarios.
- Minimización del riesgo de pérdidas por evaluación inexacta.
- Experiencia del cliente más eficiente y transparente.



PROCESO

1. Preparación de Datos

Limpieza y Preparación para el Modelo

- Estandarización de formatos de texto y fechas.
- Eliminación de datos duplicados y outliers.
- Imputación de valores faltantes utilizando la media.

Entre otros...

2. Exploración de Datos (EDA)

Limpieza y Preparación para el Modelo

Se analizaron las **correlaciones** entre las variables y el precio de venta. Se identificaron **37** variables categóricas y **59** variables numéricas relevantes para el modelo.

3. Desarrollo del Modelo

Se entrenó el modelo KNN utilizando el 80% de los datos para entrenamiento y **validación cruzada** (10 folds), y el 20% restante para **Test**. Se optimizaron los hiperparámetros del modelo para minimizar el MAPE.

Limpieza:





- Se eliminaron columnas con más del 50% de datos vacíos y registros únicos con menos de 100 repeticiones.
- Se realizó one-hot encoding para variables categóricas.
- Se identificaron inconsistencias en la base de datos que afectan el rendimiento del modelo en ciertos folds.

EDA:

- La correlación entre variables numéricas y el precio de venta no fue significativa.
- El 34% de las columnas tenían valores nulos.
- Se visualizaron patrones y tendencias utilizando [Especificar herramientas de visualización].

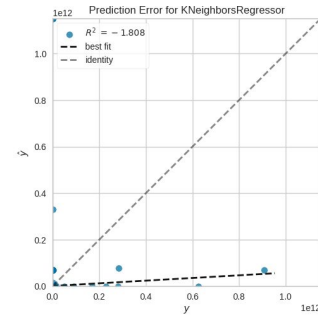
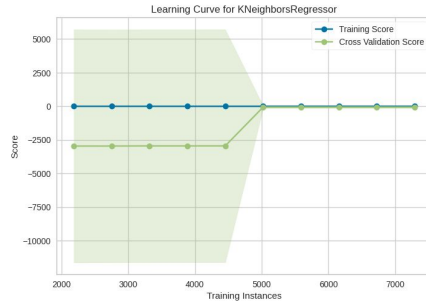


RESULTADOS

METRICA	TRAIN	TEST
MAPE 	4.75	3.5
MAE 	3.5 E9	1.7E9
R2 	-62*	-1.8
RMSE 	67 E9	-

El desarrollo del modelo se realizó utilizando Google Colab y Drive, sin incurrir en costos de infraestructura, pero se debe evaluar la viabilidad de escalar el modelo a una infraestructura más robusta para producción.

El modelo demuestra un alto potencial para automatizar la evaluación de viviendas, pero se requiere una revisión de las inconsistencias en los datos para mejorar aún más su precisión.



Comparación modelos:

- El valor del R2 se debe a dos grandes inconsistencias en 2 folds con outliers, que interfirieron mucho en el promedio.
- Se deben investigar las razones del buen rendimiento de KNN en comparación con otros modelos.



FUNCIONALIDAD

VALOR PARA EL NEGOCIO:

- ▶ **Reducción del Tiempo de Procesamiento:**
Agilización del proceso de créditos hipotecarios.
- ▶ **Disminución del Riesgo:**
Minimización de pérdidas por evaluaciones inexactas.
- ▶ **Experiencia del Cliente Mejorada:**
Proceso más eficiente y transparente para los clientes.

Adopción y Usabilidad:

El modelo se integrará en el sistema de gestión de créditos hipotecarios, con una interfaz intuitiva y fácil de usar para los analistas de crédito.

Escalabilidad:

El modelo se puede escalar para aplicarlo a otros tipos de bienes inmuebles y a diferentes mercados, adaptándose a las necesidades específicas de cada región.



CONCLUSIONES

IMPACTO EN LA TOMA DE DECISIONES:

- ▶ Se ha demostrado la viabilidad de utilizar la IA para automatizar la valuación de viviendas con alta precisión.
- ▶ Se identificaron áreas de mejora en la calidad de los datos para optimizar aún más el rendimiento del modelo.

Recomendaciones y Próximos Pasos:

- Implementar el modelo en un entorno de prueba para evaluar su desempeño en un escenario real.
- Continuar con la limpieza y mejora de la base de datos.
- Desarrollar un plan de implementación que incluya la capacitación del personal y la integración con los sistemas existentes.
- Se estima un retorno de la inversión debido a la reducción de costos operativos y la mejora en la gestión de riesgos.

Comparación modelos:

- Se pueden incorporar datos de web scraping y información georeferenciada para enriquecer la base de datos.
- Se pueden explorar técnicas de deep learning para mejorar la precisión del modelo.



IMPLEMENTACION

Viabilidad Técnica y Económica:

- ▶ Viabilidad Técnica: El modelo es técnicamente viable y se puede implementar con la infraestructura actual.
- ▶ Viabilidad Económica: El proyecto ofrece un alto potencial de retorno de la inversión, con una reducción de costos y una mejora en la eficiencia.

Consideraciones Futuras:

- Monitorear el rendimiento del modelo de forma continua y reentrenarlo periódicamente para mantener su precisión.
- Considerar la incorporación de nuevas variables y la adaptación del modelo a las cambiantes condiciones del mercado inmobiliario.