

Predicción de precios de vivienda

Prueba técnica para la selección de Científico(a) de Datos

Nicolás Sebastián Buitrago Vargas
Economista y magister en estadística

Lunes 20 de marzo de 2023

Objetivo: Predecir el precio de las viviendas de acuerdo con sus características

Se propone el desarrollo de las siguientes tareas

Insumo

Base de datos con:

- **11.571** inmuebles
- **222** variables

- Exploración preliminar de la información

Hay una gran cantidad de datos con lo cual se deben elegir las variables a tener en cuenta. Se considera:

- Su potencial de explicar el precio en un sentido económico.
- Si la variable tiene una composición tal le permita explicar la variabilidad de la variable objetivo.

- Preprocesamiento de los datos

- En los campos categóricos se hacen homologaciones y deja información sin inconsistente como “NIN”, no información.
- En los datos numéricos se excluyen datos atípicos y se eliminan datos inconsistentes, como áreas privadas en cero.

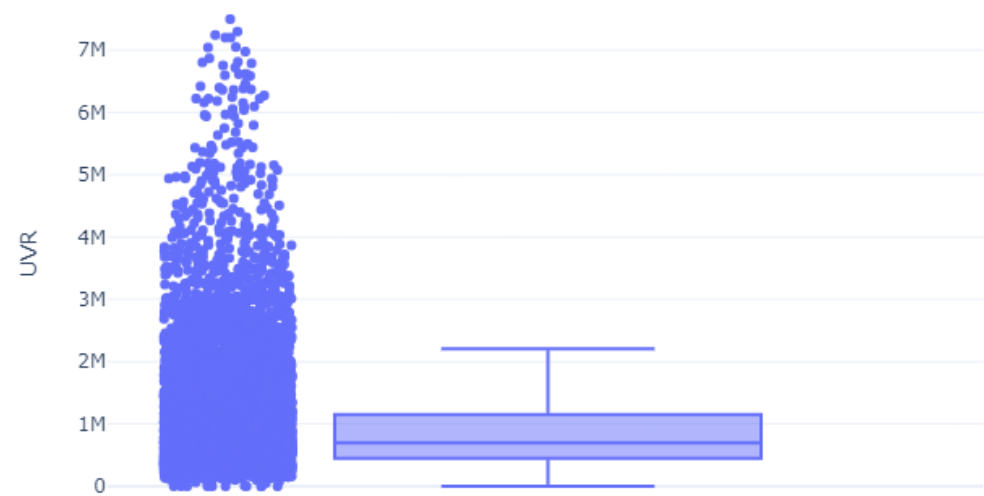
- Propuesta de modelamiento

- Se tienen variables predictoras cualitativas y cuantitativas, un método como Random Forest (Breiman, 2001) se adapta a esto al usar en conjunto ambos tipos.
- Buitrago (2021) encuentra que métodos que usan árboles de decisión son propicios para modelar los precios de la vivienda

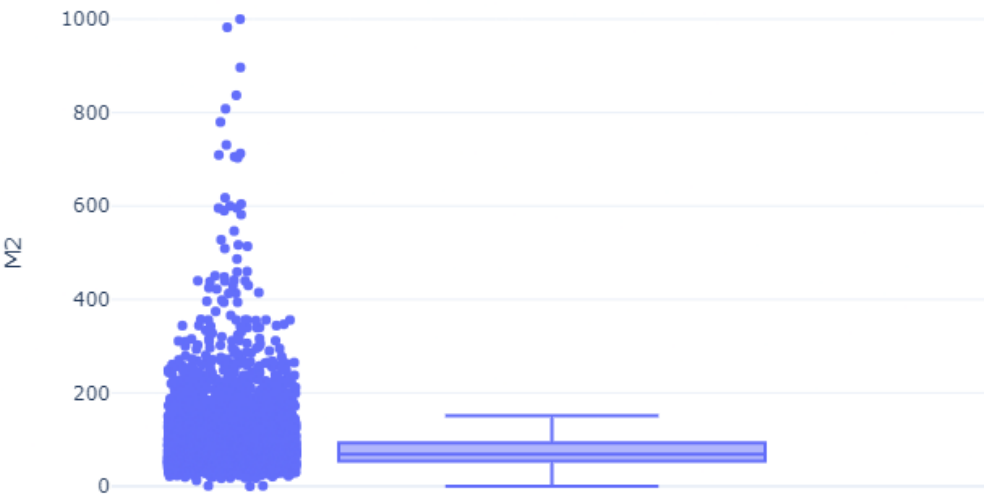
Distribución del precio del inmueble en UVR y del área privada del mismo

De acuerdo, con el valor de la UVR la mediana de precio de la vivienda es de cerca de 237 millones de pesos

Boxplot del valor del avaluo en UVR



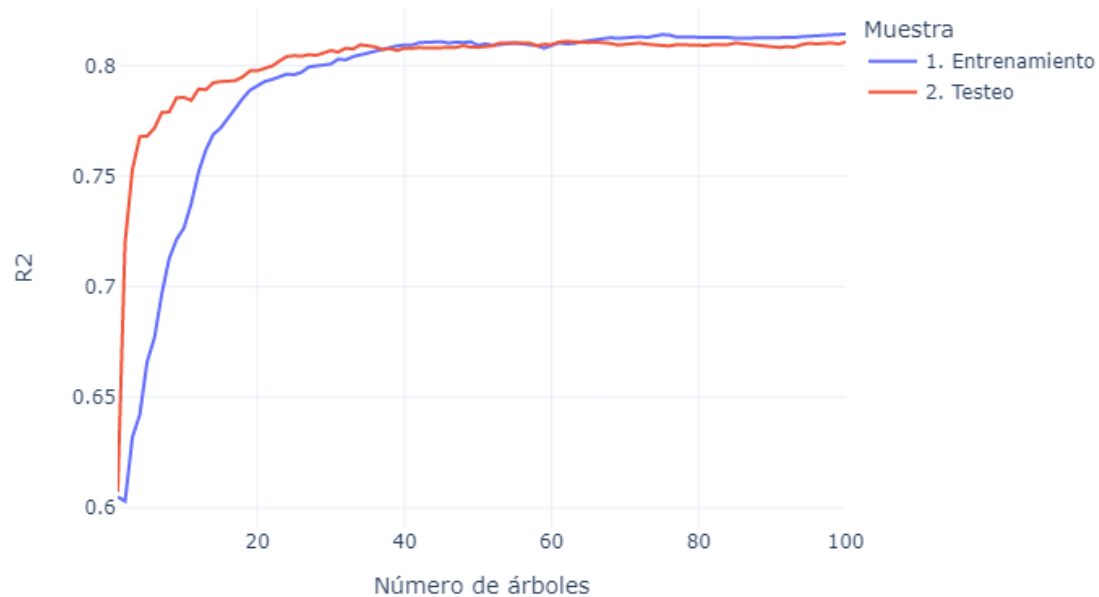
Boxplot del área privada de las viviendas



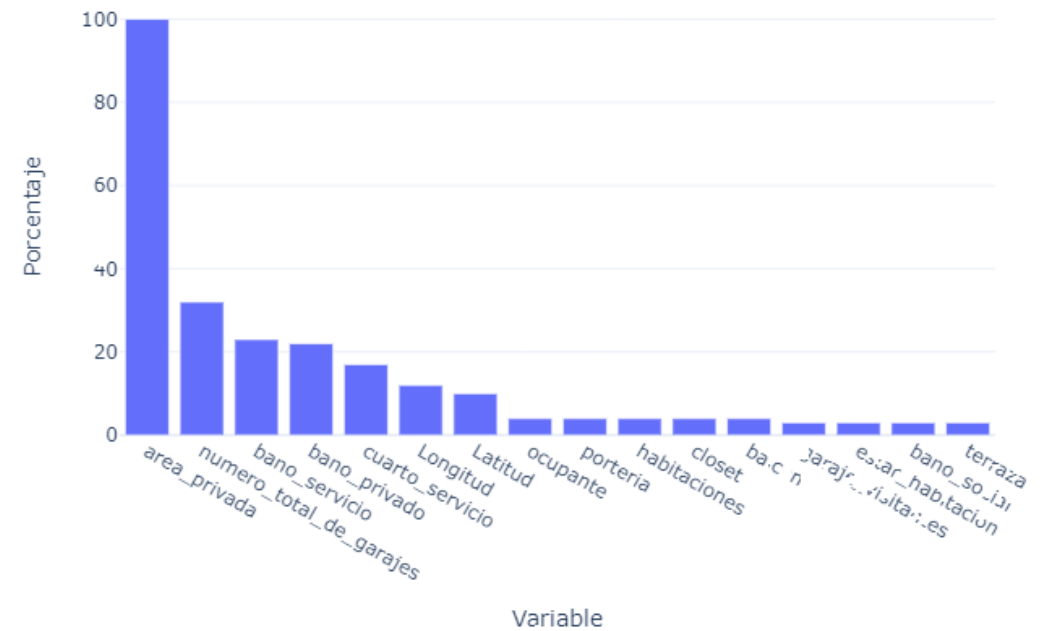
Resultados de la modelación

El modelo logra explicar un poco más del **80%** de la varianza del precio sin sobre ajustarse.
El **área de la vivienda** es la variable más importante en explicar el precio.

Importancia relativa de las variables en la predicción del precio



Importancia relativa de las variables en la predicción del precio



Trabajo futuro

● Involucrar información del espacio circundante

Agregar a la vivienda distancias a puntos de interés y calidad de esos puntos de interés

● Ejercicios del agrupamiento

Además de predecir el precio es interesante entender la configuración del mercado inmobiliario. ¿Qué patrón hay en los inmuebles? ¿se pueden resumir en grupos?

● Construir índices de características

Hay información de características de la propiedad horizontal, del interior del inmueble y de los acabados del mismo. Esto se puede resumir en índices mediante ACM

● Tableros de consulta

Consultar la predicción en tableros de consulta que permitan establecer escenarios. Pero así mismo que esos tableros permitan identificar patrones identificados en los ejercicios de agrupamiento

● Involucrar PLN

En la información se encontraba descripciones de los inmuebles. Estos también pueden ser insumo para identificar información adicional, así como para pronosticar el precio usando redes neuronales

Predicción de precios de vivienda

Prueba técnica para la selección de Científico(a) de Datos

Nicolás Sebastián Buitrago Vargas
Economista y magister en estadística

Lunes 20 de marzo de 2023