Problem Set 2 – Ganancias con ML

Julio 26,2022

Andrea Margarita Beleño, María Valeria Gaona Guevara

*Abstract*—El precio de las viviendas cuenta con diferentes factores que se tienen en consideración para establecer un valor. Por lo tanto, contar con un modelo en el que se identifiquen dichas características principales es fundamental para poder generar una predicción lo más óptima posible. Por lo tanto, en el siguiente documento se realizará el modelo de predicción para los precios de las viviendas por medio del algoritmo Super Learner. El link al Github del presente taller, se encuentra en el siguiente enlace: https://github.com/mvgaona/Problem-Set-3

# Introducción

El precio de una vivienda está dado por diferentes factores como el área total, el número de baños, número de habitaciones, la distancia a puntos importantes de la ciudad, entre otros elementos que explican dicho valor. Además, cada ciudad cuenta con diferentes características que aportan a que los vendedores decidan establecer un precio. Sin embargo, se puede generar un modelo en donde se pueda observar el efecto de cada una de esas características esenciales que da el valor de un inmueble y así, tanto el vendedor como el comprador conoce cada uno de estos impactos para realizar futuras compras y/o ventas de inmuebles.

Por otra parte, Generar una predicción de estos valores es una herramienta óptima para conocer el comportamiento económico de una sociedad y además, cada individuo puede obtener información valiosa para la toma de decisiones a mediano o largo plazo. De acuerdo con lo anterior, en el presente documento se presenta el proceso de la limpieza de datos pertinente para continuar con la caracterización de las variables fundamentales y así, realizar el modelo de predicción de los precios de las viviendas en la localidad de Chapinero en la ciudad de Bogotá y en el área del poblado en Medellín. Ambas ciudades Colombianas

# Datos

El precio de una vivienda puede estar dados por diferentes factores, tanto económicos como sociales. En el caso Colombiano, estos precios pueden estar influenciados por elementos propios de la vivienda y a su vez, por elementos geoespaciales, como la distancia a diversos lugares públicos (Bares, transporte público, etc). Para realizar un modelo de predicción de precios de la vivienda, es necesario contar con las variables determinantes y relevantes del precio, para que este modelo sea robusto, pero no se incurran en gastos que entorpezcan la investigación.

El número de habitaciones de la vivienda es determinante en el precio de la misma, ya que se puede contar con un aproximado del espacio y de cuántos individuos puedes vivir con la mayor calidad de vida posible, es decir, entre más habitaciones, el precio del inmueble tiende a incrementarse. Por lo tanto, dentro del análisis descriptivo de este predictor se encuentra que es una variable numérica, en donde el número mínimo de habitaciones es 1, el máximo son 11 habitaciones por vivienda. Además, el número promedio de habitaciones por vivienda es 3 y de igual manera, son 3 habitaciones el número más común que cada inmueble cuenta.

Otro factor de decisión importante en el precio del inmueble es la cantidad de baños, ya que sin baños, los individuos no pueden satisfacer las necesidades básicas de aseo. Por lo tanto es necesario contar con al menos un baño y con ello, se puede identificar que a medida que aumentan la cantidad de baños, el precio de igual manera se verá afectado e incrementará. De acuerdo con lo anterior, se observa que en la variable numérica, el número mínimo de baños es 1 y así mismo, el máximo es 43 baños, el cuál este puede ser una observación atípica. Por otro lado, en promedio, los inmuebles cuentan con 2.3, es decir 2 baños y finalmente, el número de baños que se repite con mayor frecuencia es 2.

Si el inmueble cuenta con ascensor o no es determinante en el caso de los apartamentos de pisos altos, es por eso que se considera un predictor importante del modelo, ya que quienes viven en pisos más altos, buscan apartamentos donde tenga ascensor y a su vez, el precio aumenta si el inmueble cuenta con esta herramienta de desplazamiento dentro del edificio. Dentro del análisis descriptivo de la variable, es posible identificar que es una variable categórica de 2 niveles: 1 y 0, donde 1 corresponde a si el inmueble cuenta con ascensor y 0 si no. Esta variable cuenta con 85405 viviendas que tienen ascensor y 22162 que no lo tienen; esto puede estar dado porque aquellas viviendas son casas y por lo tanto, no necesitan ascensor.

En la actualidad, gran parte de los individuos cuentan con uno o más carros, por lo tanto, es necesario que el inmueble cuente con al menos un parqueadero para que el individuo no incurra en gastos adicionales en encontrar un lugar seguro para su(s) carro(s), por lo tanto, si el inmueble incluye al menos un garaje, el precio de la vivienda tenderá a aumentar su valor. Por otra parte, de acuerdo con el análisis descriptivo, se identifica que esta es una variable categórica de dos niveles, en donde 1 hace referencia a que la casa o el apartamento cuenta con al menos un parqueadero y 0 que no lo hace. La cantidad de inmuebles que tienen un parqueadero por lo menos, es de 37866 y 69701. Esto puede estar dado porque algunos vendedores no ofrecieron dicha información, por ejemplo.

Para la variable tipo de propiedad XXXX

# Modelo y resultados

XXXX

# Conclusiones y recomendaciones

# REFERENCIAS

XXXXXXX

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**