

Entrega 1



ASIGNATURA

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL

Profesor

RAUL RAMOS POLLAN

RESPONSABLES

Julian Andres Castaño Jimenez

**UNIVERSIDAD DE ANTIOQUIA
MEDELLÍN
2023**

Descripción del problema: Dadas las características de los alquileres de una casa disponible en Airbnb en Madrid. La idea es predecir el precio a cobrar de un alquiler a corto plazo.

Para este análisis se utilizara el dataset proporcionado por kaggle [Madrid Airbnb Data](#) el cual cuenta con 18794 muestras de las cuales se utilizaran entre 5000 a 10000 de las primeras filas y con las siguientes columnas de información:

- 1) **id:** Forma de identificar el Airbnb de manera única.
- 2) **name:** Anunciador.
- 3) **descripción:** Descripción del anuncio.
- 4) **neighborhood_overview:** Descripción del barrio.
- 5) **host_id:** Identificador único del Airbnb para el usuario.
- 6) **host_since:** Fecha de creación del usuario.
- 7) **host_location:** Ubicación del anfitrión.
- 8) **host_response_time:** Tiempo en que tarda en responder el anfitrión por medio de Airbnb.
- 9) **host_response_rate:** Velocidad con la que el anfitrión responde los mensajes por medio de Airbnb.
- 10) **host_acceptance_rate:** Proporción de solicitudes de reserva del anfitrión que han sido aceptadas.
- 11) **host_is_superhost:** Variable utilizada por Airbnb para anfitriones que cumplen ciertos criterios (Estándares de calidad).
- 12) **property_type:** Tipo de propiedad auto seleccionada. Los hoteles, camas y desayunos son descritos como tales por sus anfitriones en este campo.
- 13) **room_type:** Tipo de casa (Casa entera - apartamento - Habitación privada - Habitación compartida - Hotel).
- 14) **accommodates:** La capacidad máxima del listado.
- 15) **bathrooms_text:** El número de baños del anuncio. En el sitio web de Airbnb, el campo de los baños ha pasado de ser un número a.
- 16) **bedrooms:** Número de habitaciones.
- 17) **beds:** Número de camas.
- 18) **amenities:** Lista de servicios incluidos en el alquiler.
- 19) **price:** Precio diario en moneda local.
- 20) **minimum_nights:** Número mínimo de noches de estancia para el listado.
- 21) **maximum_nights:** Número máximo de noches de estancia para el anuncio.
- 22) **has_availability:** Tiene disponibilidad [t=true; f=false].
- 23) **number_of_reviews:** El número de reseñas que tiene el anuncio.
- 24) **first_review:** La fecha de la primera/antigua revisión.
- 25) **review_scores_rating:** Puntuación global de las calificaciones.
- 26) **review_scores_location:** Puntuación de la ubicación de la propiedad.
- 27) **review_scores_accuracy:** Forma en que se califica la precisión de la descripción proporcionada por el anfitrión.

- 28) **review_scores_cleanliness**: Puntuación de la higiene de la propiedad.
- 29) **review_scores_checkin**: Puntuación de la experiencia del usuario en el proceso de registro en la propiedad.
- 30) **instant_bookable**: Reserva instantánea [t=true; f=false].

Métricas de Machine Learning: Se utilizara el RMSE (Root Mean Squared Error) es particularmente relevante en este contexto, ya que su valor refleja cuán cercanas están las predicciones del modelo a los precios reales de alquiler. Un RMSE bajo indica que el modelo tiene la capacidad de realizar predicciones altamente precisas.

Supongamos que el RMSE tiene un valor de 5, de un rango de precios entre 100 y 200, este valor refleja el error de la predicción al ser bajo significa que el modelo es efectivo para predecir los precios del alquiler.

Métricas del negocio:

- Se tomará en cuenta el aumento de los alquileres del 10% debido a las condiciones legales de España que son del 2% antes de esto no tiene sentido la ejecución del modelo.
- Aumento de puntuaciones en el rango del alquiler debe de ser del 10%