



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL
IMT2220 – INTRODUCCIÓN A CIENCIA DE DATOS

Propuesta Proyecto

Tasación automática de propiedades

Grupo 23 — Javier Cuitiño, Simón Saravia, Alejandro Orellana

11 de septiembre de 2025

Contexto y motivación

Pregunta: Discuta la motivación y razones para su elección del proyecto, incluyendo el contexto y alguna investigación que haya incluido en su decisión. Comente también sobre la audiencia objetivo de su análisis (quién será el tomador de decisiones que se apoyará en su trabajo).

Respuesta: La compra-venta de viviendas y departamentos exige que ambas partes (comprador y vendedor) cuenten con un valor de referencia objetivo. Una metodología ampliamente utilizada es la valoración de mercado basada en propiedades publicadas en portales inmobiliarios. El proyecto busca automatizar esa tasación aprovechando la disponibilidad de información pública.

Objetivos

Pregunta: Describa los objetivos científicos y de inferencia de su proyecto. ¿Qué pregunta espera resolver y para qué? ¿Cuáles son los beneficios o accionables? ¿Cuál es la audiencia objetivo?

Respuesta (científico): Construir una herramienta de acceso simple que estime el precio de mercado de una propiedad utilizando datos comparables.

Respuesta (inferencia): Predecir el valor esperado de propiedades a partir de sus atributos

Datos

Pregunta: Describa los datos a utilizar, sus características (variables, tipo, formato, volumen), origen y forma de recolección.

Respuesta:

- Variables esenciales (para inferir el valor): Superficie, programa, dirección, año.
- Atributos para definir comparables: ROL o dirección (para ubicación precisa y servicios cerca), cercanía/distancia, programa (dormitorios, baños), estacionamientos, bodegas.
- Tipo y formato: inicialmente CSV. Los datos de portales se obtendrán vía *web scraping* (HTML/JSON) y se transformarán a tablas.
- Origen: portales inmobiliarios públicos y registros abiertos compatibles con uso académico.
- Recolección: Web scraping, APIs cuando existan, geocodificación de direcciones/ROL para obtener coordenadas y métricas de cercanía.

Preguntas de investigación

Pregunta: Proponga al menos 5 preguntas específicas a responder con base en los datos.

Respuesta:

1. ¿Cuáles son los bienes más similares (comparables) a la propiedad en estudio?
2. ¿Es posible construir una muestra representativa de comparables?
3. ¿Cuál es el nivel de dispersión de los precios en la muestra?
4. ¿Es posible aproximar/identificar ventas reales frente a simples publicaciones?
5. ¿Cuál ha sido la evolución histórica de los precios en el área?

Diseño tentativo (métodos computacionales y estadísticos)

Pregunta: Explique brevemente los métodos que espera usar en su análisis.

Respuesta:

- Paso 1: Adquisición y geocodificación: scraping/API de portales; uso de dirección o ROL para geolocalizar (coordenadas y métricas de cercanía).
- Paso 2: Limpieza y normalización: arreglo de valores faltantes, detección de valores atípicos, etc.
- Paso 3: Selección de propiedades comparables: métrica de similitud ponderada (distancia geográfica + diferencia en superficie y programa); seleccionar 10 comparables más cercanos.

Paso 4: A partir de todos los datos obtenidos utilizar un modelo de aprendizaje automático para hacer el cálculo del valor.

Paso 5: Entrega: prototipo web simple que, dada una dirección/ROL y datos importantes, devuelve comparables, mapa, precio estimado e intervalo de confianza.

Resumen ejecutivo

Herramienta de tasación automática basada en comparables cercanos, con ajustes hedónicos y validación cuantitativa; destinada a tasadores, banca y usuarios finales para decisiones informadas de compra-venta e inversión.

Link repositorio: <https://github.com/Saaay20/PROYECTOIMT2200>