

Elements of Machine Learning 2025

Primer Proyecto

26.febrero.2025

Datos

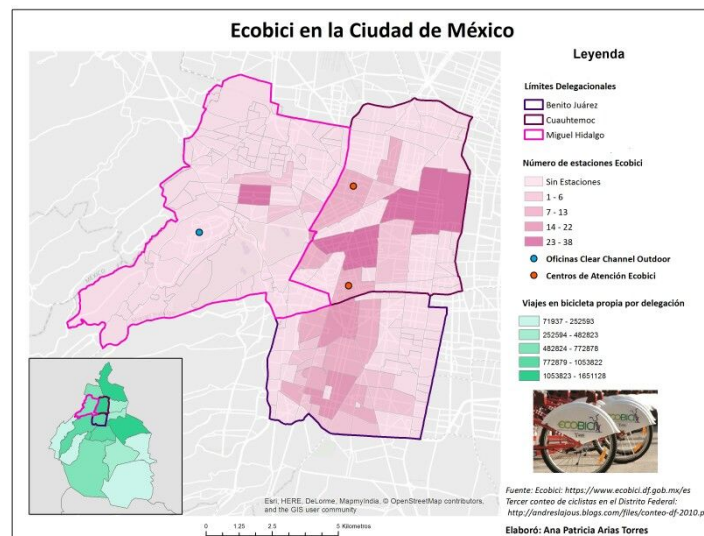
El proyecto consiste en elegir trabajar con los datos de ECOBICI, el cual es un sistema público de bicicletas en la Ciudad de México. Puede encontrar los históricos de ECOBICI en

- <https://ecobici.cdmx.gob.mx/>
- <https://ecobici.cdmx.gob.mx/datos-abiertos/>.

Un página con algunas estadísticas útiles es la siguiente:

- <https://ecobici.cdmx.gob.mx/estadisticas/>.

Además, los archivos `stations.csv` o `stations.json` proporciona las coordenadas geográficas de cada estación.



Proyecto

La idea es utilizar las herramienta vistas en el curso, y algoritmos o métodos que no hemos visto pero que ustedes deseen incorporar, para obtener información valiosa a partir de los datos.

El proyecto consiste en elaborar un análisis de datos. Para ello, deberá

- Redactar un informe detallado de su análisis. Tomar en cuenta lo siguiente:
 - breve descripción de los datos (tamaño, variables, tipos de datos)
 - exploración
 - selección de variables importantes
 - tratamiento de datos faltantes y preprocesamiento
 - técnicas de visualización

- análisis (indicar con detalle la metodología)
 - resultados
 - conclusiones
 - referencias
- Entregar copias del código utilizado (en un archivo comprimido).
 - Realizar una presentación de 20 minuto en la semana de presentaciones.

Observaciones

- Pueden colaborar y compartir ideas entre equipos, pero cada equipo debe hacer su análisis, reporte y presentación.
- Pueden utilizar algoritmos o técnicas no vistas aún en el curso, desde que se indiquen las referencias.
- Pueden usar código o partes de código ajeno (disponible en la web), desde que se indique los créditos y la referencia correspondiente. El uso de herramientas automáticas y de inteligencia generativa está permitido, desde que se incluya la referencia correspondiente.

Fechas de entrega

- Presentaciones: **semana del 24 al 28 de marzo.**
- Entrega de informe, código y presentación: **viernes 28 de marzo.**

Evaluación

El proyecto vale 25 puntos, que se distribuirán de la siguiente forma:

Aspecto a evaluar	Proporción
Análisis de datos, metodología y elaboración del informe	15 puntos
Innovación y creatividad en ML (Sorpréndanme)	05 puntos
Presentación de los resultados, preguntas y participación	05 puntos