



Equipos Demoday



Al.rbike

Equipo Turquesa •

Mentor:



Edwin Jiménez



Raúl García



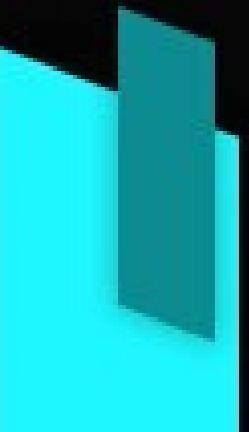
Jacobo
Domínguez



Norma Tapia



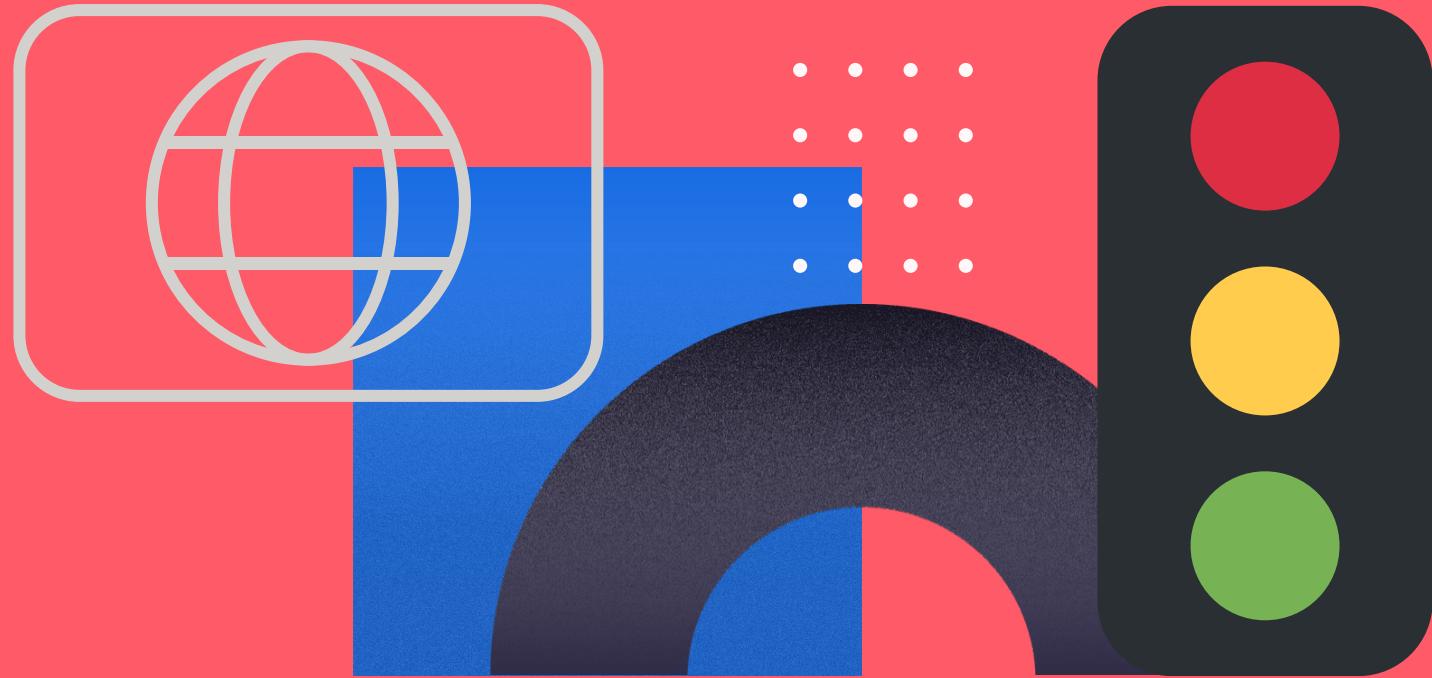
Kimberly Celeste



MiBiciudad Limpia

Al.rbike

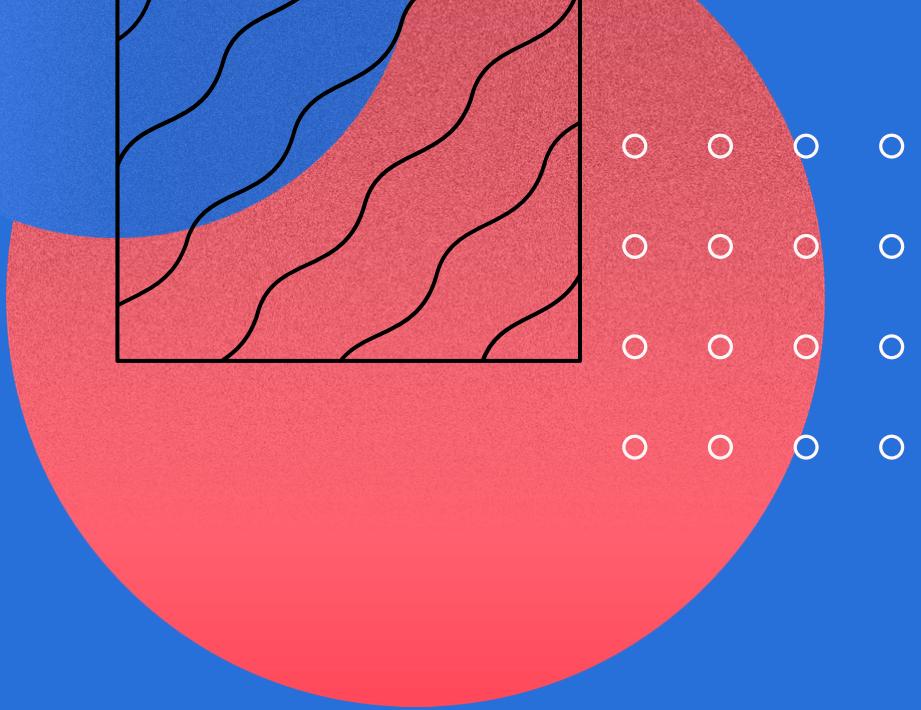




Medio Ambiente

Utilizando datos abiertos del sistema MiBici + calidad de aire de la ZMG, se hace un análisis de la relación entre la calidad de aire y el uso de bicicletas. Con esto se pretende crear un modelo de Machine Learning con la finalidad de demostrar que a un mayor uso de bicicleta como medio de transporte se tiene como consecuencia una mejora en la calidad de aire de la zona.

Problemática



¿Cómo hacer visible la importancia del uso de bicicleta y su repercusión en la calidad del aire incluso para aquellos ciudadanos que no son usuarios de éstas?

Datos



Infraestructura Pública de Bicicletas

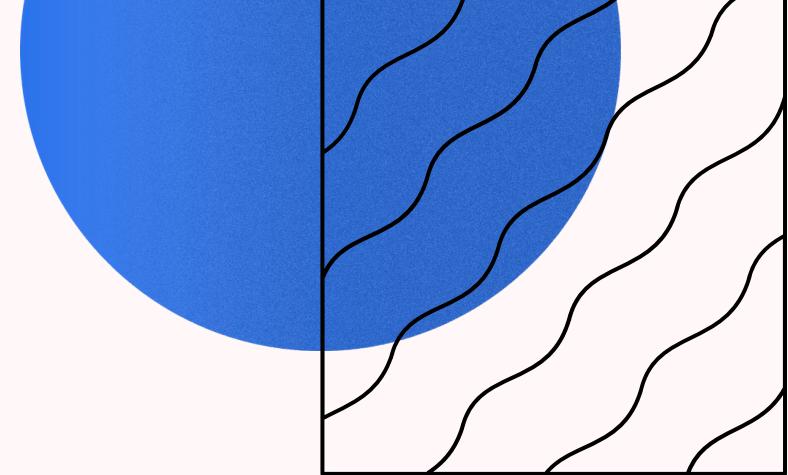
- MiBici (Guadalajara)
- Divvy (Chicago)
- Bluebikes (Boston)
- Citibike (Nueva York)



Calidad del Aire

- Air Quality Historical Data Platform
- SEMADET
- Openaq





Hipótesis

Uno de los principales factores que afectan la calidad de aire en las ciudades es el uso de vehículos de combustibles fósiles.

Nuestra hipótesis es que si se incrementa el número de **viajes en bicicleta**; la calidad de aire mejorará; notando una mejora en la calidad del aire





Descripción de la solución



Viajes en
bicicleta

+

Calidad de aire

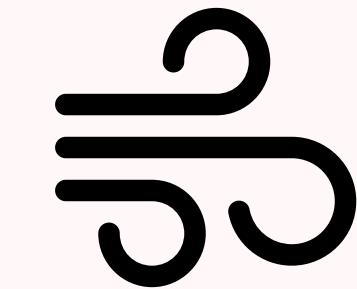
↓

Modelo de
Predicción

Preprocesamiento



Datos estaciones
2014-2020



Mediciones Contaminante
2014-2020

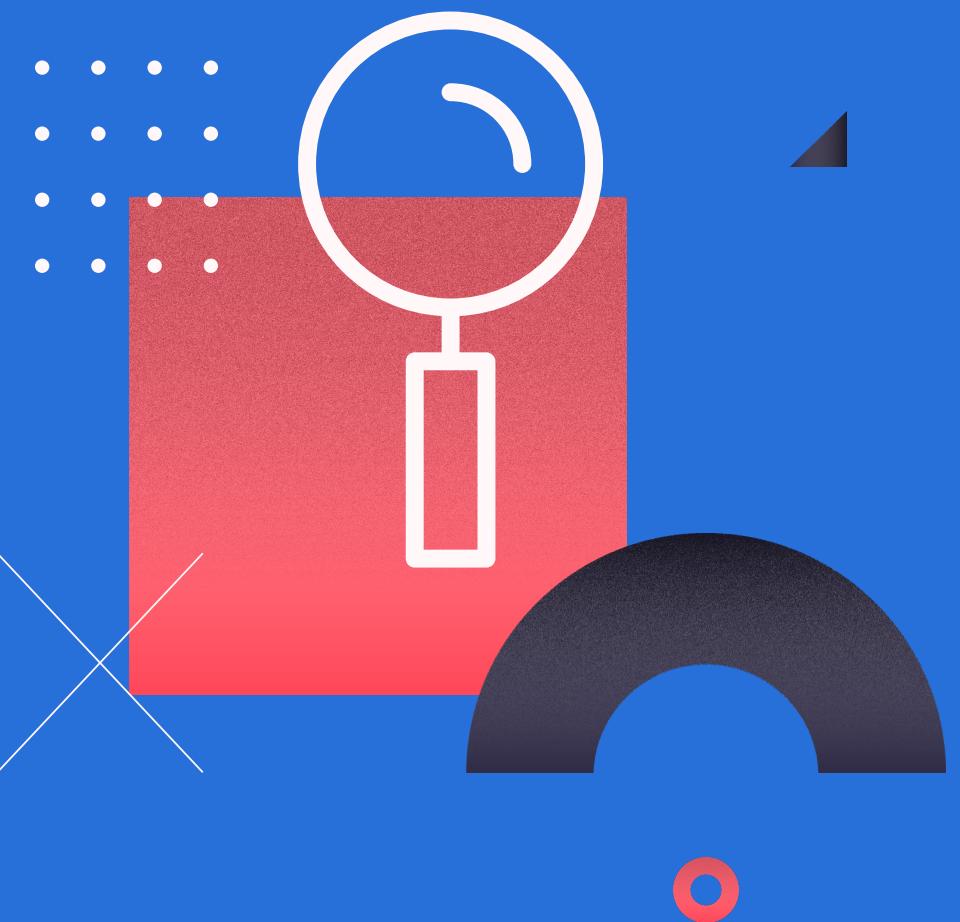


Dataset por ciudad

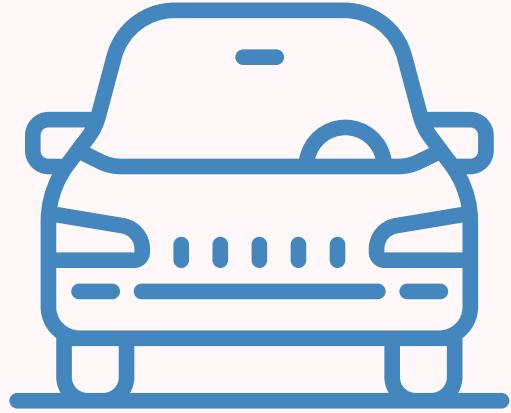
# Viajes	Distancia	Duración	Hora	Estación	Fecha	Contaminante
----------	-----------	----------	------	----------	-------	--------------

Etapas de experimentación

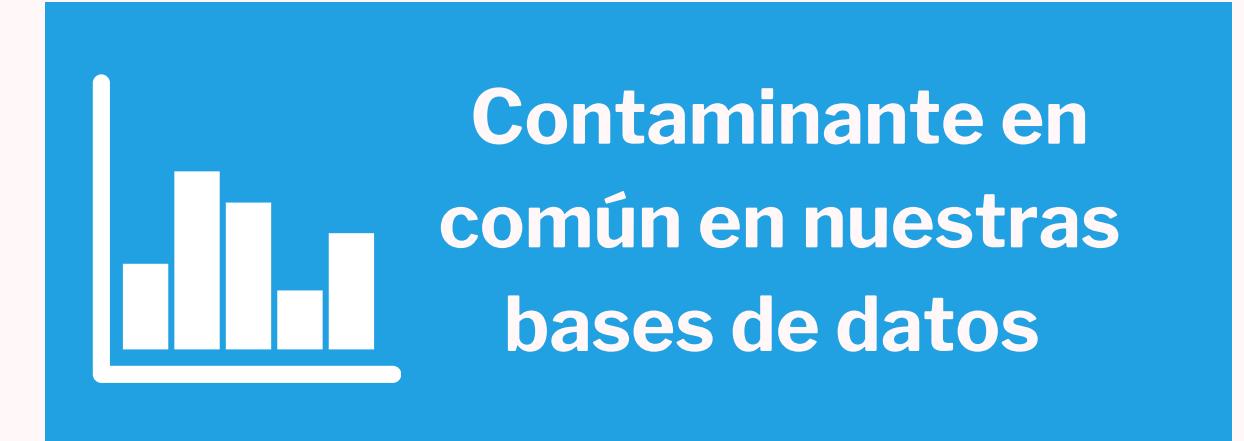
- Modelos específicos por estación
- Generalizar para la ciudad



Experimentación y Resultados



Óxidos nitrosos (NOx)
Monóxido de carbono (CO)
Dióxido de carbono (CO₂)



Sin embargo, para NY se tomó el ozono debido a que era el contaminante con más datos disponibles

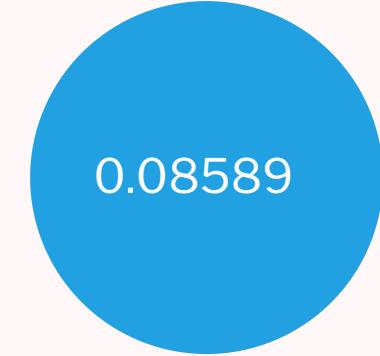
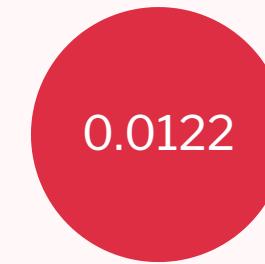
Estos modelos no fueron tan exitosos

Nueva York : Ciudad

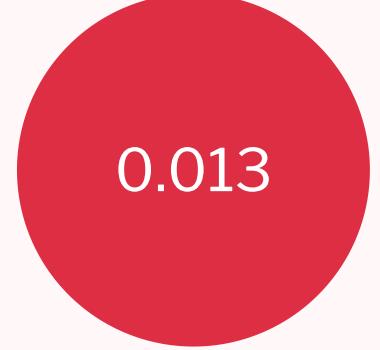
1'728,988 viajes en 2019

- Root Mean Square Error
- Determination coefficiente

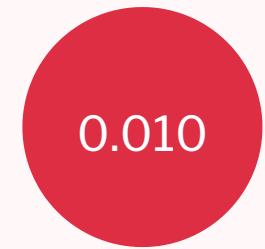
Linear Regression



Decision Tree Regressor

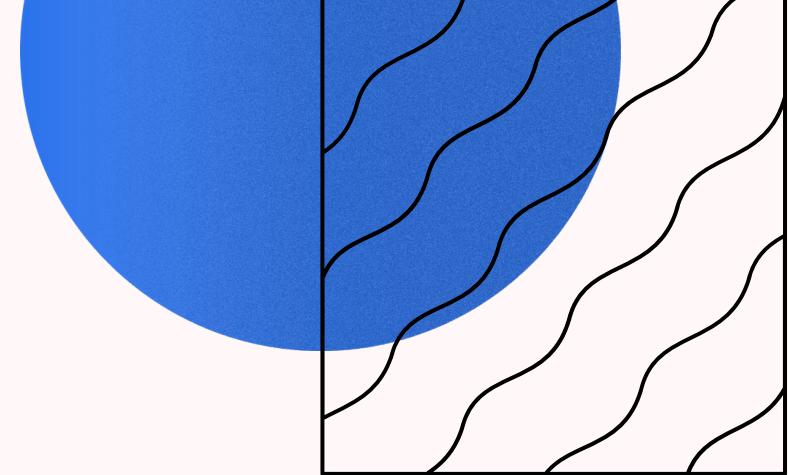


MLP Regressor



Había muy pocos registros de otros contaminantes, por lo que aquel que se consideró para hacer el modelo fue el ozono.





¿Por qué el modelo se comporta mejor con ozono?

- Los contaminantes emitidos por el sector transporte contribuyen a la formación de ozono en la atmósfera
- En los últimos 5 años ha habido 22 contingencias en la Ciudad de México y 17 de ellas han sido causadas por ozono

¿Será que estamos analizando el contaminante menos adecuado?



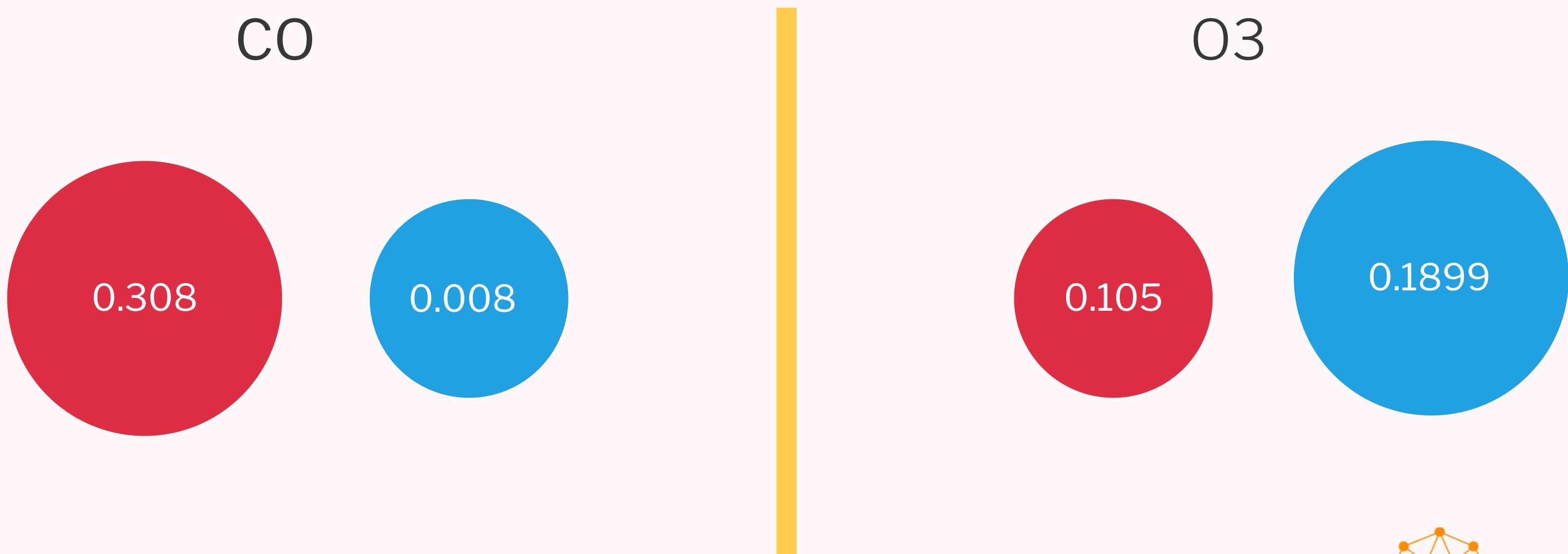
Boston : Station 39

Estación mayor número de viajes

5395 viajes en 2019

Linear Regression

- Root Mean Square Error
- Determination coefficient



El modelo de la ciudad puede no ser tan óptimo como el de esta estación debido a la lejanía de las estaciones a la locación de medición del contaminante



Guadalajara : Ciudad

Estación mayor número de viajes

5395 viajes en 2019

MLP Regressor



Root Mean Square Error

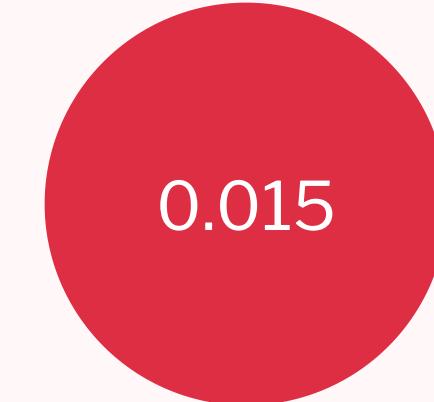


Determination coefficient

CO



O3

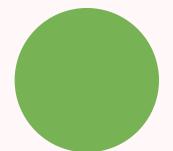


Red Neuronal por cantidad de datos.

Estación medición con más viajes a 2km, la estación más lejos 4km



Guadalajara : MLP Regressor



0.04

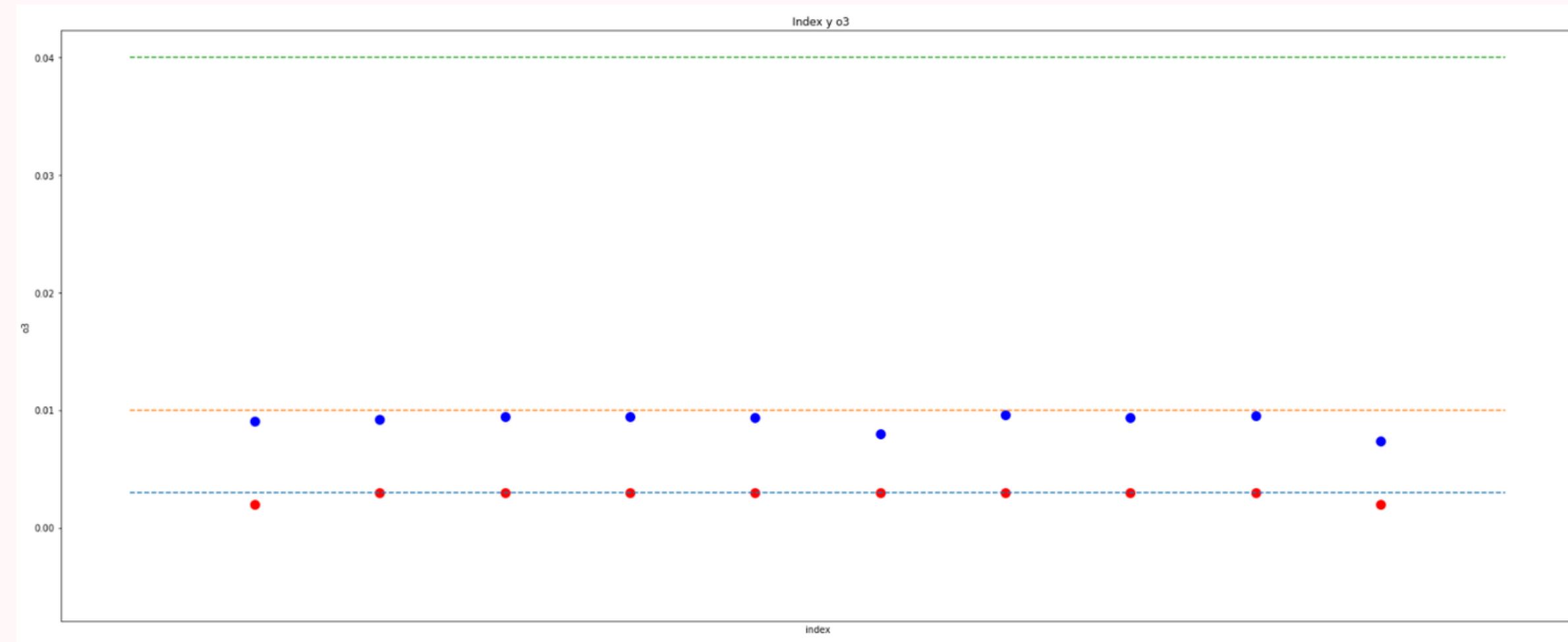


0.01

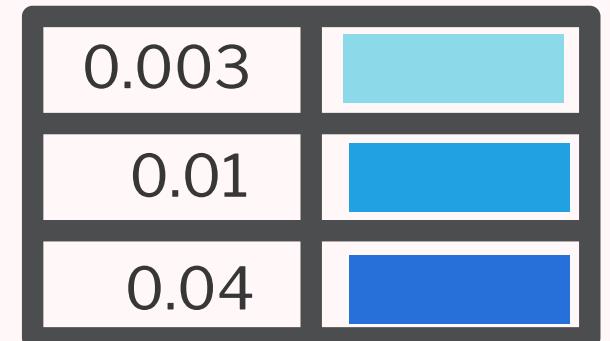


0.003

Concentración PPM Ozono



Rango de concentración ozono



Las distancias entre los puntos de los datos reales y del modelo de predicción se consideran mínimas de acuerdo al contexto. Es decir, en términos de la escala del ozono, estas distancias no repercuten entre una medición aceptable a una nociva.

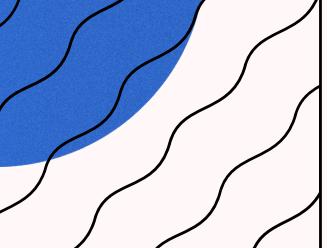
Conclusiones

Comportamiento entre modelos de ciudad & estación

Para Guadalajara, los modelos por estación tenía diferencias mínimas entre sí de valor de Root Mean Square Error y el Coeficiente de determinación. La **máxima distancia** entre la medición del contaminante era de **4km**. Para este modelo era conveniente utilizar los **datos de toda la ciudad**

Para Boston, había estaciones con **muy pocos datos**, por lo que esto afectaba el resultado del modelo por ciudad. Por otro lado, había estaciones incluso a **29 km de distancia** de la locación de medición del contaminante.



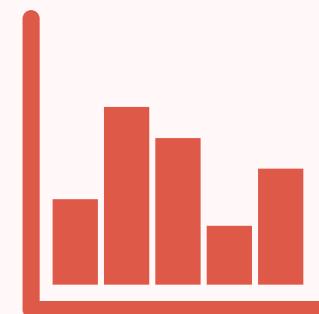


Trabajo futuro



Aplicar los modelos de predicción para visualizar el impacto ambiental de acuerdo al número de bicis por zona.

- Decisiones de infraestructura
- Políticas Pùblicas
- Cambio de percepción de la importancia de la bicicleta

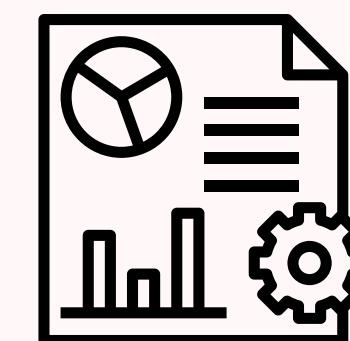
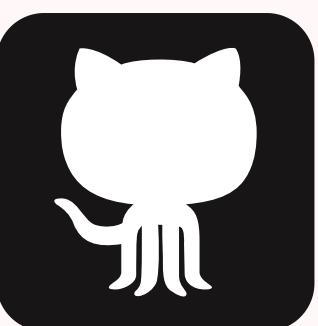


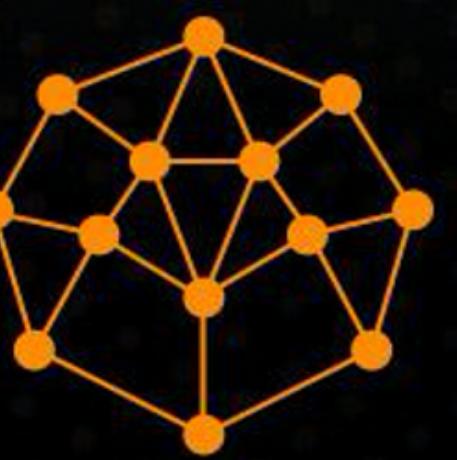
Cantidad y calidad de mediciones de los datos



Tiempo de preprocesamiento de los datos y acceso a las bases

Recursos





Saturdays.AI



Saturdays.AI
Guadalajara

