

Ciencia de Datos Aplicada a la Predicción de los Periodos de Mantenimiento de la Red del Metro, Metrobús y RTP

Trabajo Terminal No. — — — — —

*Alumnos: Carrillo Oviedo Ian, Gutiérrez González Diego, *Serrano Covarrubias Marco Antonio.*

Directores: Martínez Acosta Lilian, Corona Bermúdez Uriel

**e-mail: marcoaserrano2001@gmail.com*

Resumen –

La movilidad de la población dentro de la ciudad es una actividad de gran importancia, tanto en ámbitos sociales como económicos, es por ello, que se crea la necesidad de un funcionamiento óptimo del transporte público. Sin embargo, actualmente existen múltiples problemáticas dentro de esta red, debido a la falta de mantenimiento. Por lo que, a lo largo de este trabajo se analizarán datos de afluencia de los sistemas Metro, Metrobús y RTP, además de datos de calidad del aire, que, en ocasiones, intervienen. Esto, para predecir la concurrencia y determinar temporadas viables para su mantenimiento, sin afectaciones a usuarios.

Palabras clave – Ciencia de datos, predicción, transporte público y mantenimiento.

1. Introducción

El transporte público de la Ciudad de México ha sido, a lo largo de los años, una parte fundamental para la movilidad dentro de la zona urbana. Actualmente, los medios de transporte público más populares son: el Metro, el cual ha funcionado desde su inauguración en 1969, que a día de hoy consta de 195 estaciones divididas en 12 líneas, con una longitud total de 226.48 km [1]; el Metrobús inicio sus operaciones en 2005, consta de 283 estaciones divididas en 7 líneas, con longitud de 174.6 km [2] y el RTP creada en el año 2000, un sistema de rutas de camiones que consiste en 103 rutas divididas en 7 módulos y 8 diferentes tipos de servicio. [3]

Los sistemas del Metro y Metrobús laboran normalmente jornadas de 19 horas con horario de 5:00 a 0:00 horas y en el sistema de RTP, algunas de las rutas son de jornadas de 24 horas debido a cuentan con el servicio de “Nochebús”, el cual, tiene un costo más alto que en horarios normales. La red de transporte ha logrado mejorar la movilidad de las personas dentro de la Ciudad de México, por ejemplo, cuando se activa el protocolo “Hoy no circula” debido a la mala calidad del aire, muchos suelen optar por algún servicio de transporte público en lugar de un taxi o un Uber por el alto costo que estos servicios de transporte representan. [1,2,3]

Hoy en día, dentro de la red de transporte, existen múltiples problemas debido a la falta de mantenimiento o al largo tiempo en el que han estado funcionando de manera continua. [4] Dichos problemas pueden ir desde falta de iluminación en las instalaciones, problemas en los vehículos, fallos en las vías de transporte. Todos estos problemas pueden tener consecuencias, ya sean riesgos menores o un riesgo mayor, por ejemplo, un accidente grave que incluya muertos o heridos. [5]

Lamentablemente, han sucedido desgracias en el Metro por falta de mantenimiento. Tal es el caso del choque que se dio en la línea 1 en la estación Tacubaya en el año 2020, en el percance resultaron 43 heridos y un muerto. De acuerdo con el análisis de la caja negra de los trenes involucrados, el accidente se presentó por poca presión de aire en los frenos, lo que originó que el tren se deslizara en sentido contrario impactando con un tren estacionado [6].

Como consecuencia de múltiples incidentes en el Metro, el gobierno de la CDMX está realizando el programa de mantenimiento profundo a un tramo de la línea 1 del metro, esta sería la primera vez después de 53 años que recibirá intervención. La nula atención a las instalaciones de servicios de transporte público a lo largo de los años ha empeorado su estado gravemente. El tiempo que tomará renovar las instalaciones de la línea 1 es de un año, y repercute en la movilidad de los usuarios aunado a la saturación de rutas alternativas con tal de satisfacer las necesidades de los pasajeros [7].

Se estima, que para darle un mantenimiento profundo al sistema del metro se necesitan más de 3,500 millones de pesos. Sin embargo, esta labor no se puede llevar a cabo simultáneamente, lo que crearía un caos en la movilidad, por lo que es necesario identificar temporadas y tramos pequeños para no afectar de forma significativa el traslado de los ciudadanos [8].

Al considerar la afluencia de gente que usa el transporte público día con día y la previsión de la calidad del aire, se puede predecir en que periodos la afluencia será baja y así planificar el mantenimiento de unidades y de pequeños tramos de las rutas de los sistemas principales del transporte público, es decir, Metro, Metrobús y RTP. Con ello, se podrá reducir las afectaciones a la movilidad de la ciudad.

A continuación, se muestran algunas aplicaciones o productos similares a lo planteado.

SOFTWARE	CARACTERÍSTICAS	PRECIO EN EL MERCADO
TT 2013-A033	Base de datos para registro y control de personal, actividades y almacén del Metro. [9]	No aplica
Computerized Maintenance Management Software eMaint CMMS	Software de planeación de mantenimiento, organización de actividades colaborativas, manejo de inventario y revisión de rendimiento de producción. [10]	Costo de su servicio popular de 85 dólares al mes
SafetyCulture	Aplicación donde se registran las actividades de una empresa, ya sean mantenimiento, auditorías, gestión de actividades, entre otros. Después del registro de cada actividad todo el equipo será notificado cuando están actividades ocurran para estar al tanto. [11]	Costo de 19 dólares al mes por usuario.

Tabla 1. Productos Similares

Como se muestra en la tabla 1, estos productos solamente sirven para administrar el tiempo y actividades, avisando al resto del equipo sobre los programas que se realizaran, o funcionan para llevar un mejor control del material y personal para ser más productivos. Sin embargo, estos no sirven para saber cuándo es más conveniente hacer los mantenimientos basándose en múltiples factores a menos que sea referente a una cadena de producción en el caso de eMaint [8,9,10].

2. Objetivo

Analizar y predecir los datos de la afluencia de los usuarios del Metro, Metrobús y RTP, en conjunto con la calidad del aire para establecer temporadas óptimas de mantenimiento a estos servicios de movilidad de la Ciudad de México.

3. Justificación

El sistema de transporte urbano de la Ciudad de México ya sea Sistema de Transporte Colectivo (STC), mejor conocido como Metro, Metrobús o Red de Transporte de Pasajeros (RTP), en junio de 2021 de acuerdo con cifras proporcionadas por el Instituto Nacional de Estadística y Geografía (INEGI) [12], el sistema de transporte colectivo metro transporto a 63 millones de pasajeros, recorriendo un aproximado de 3 millones de kilómetros, mientras que el Metrobús, siendo el método de transporte más popular en la ciudad, traslado a 22 millones de pasajeros, en tercer lugar, está el RTP que tuvo una afluencia de 8 millones de pasajeros.

Cada servicio de transporte labora por jornadas de 19 horas, por lo que, el mantenimiento de las unidades es esencial para que se provea de un servicio de calidad y no se presenten fallas que retrasen la llegada de los usuarios. De acuerdo con los técnicos del STC [13], el metro o material rodante, recibe cuatro tipos de mantenimiento: fallas o averías, mantenimiento sistemático, mantenimiento cíclico y la Revisión General programada (RG-pr), donde los trenes son desarmados completamente, para sustituir las piezas desgastadas o los equipos que ya no funcionen. Este mantenimiento se debe realizar cuando los trenes han recorrido 500 mil kilómetros.

A través de una solicitud de información realizada por medio de la Plataforma Nacional de Transparencia, se sabe que hasta el pasado 06 de junio de 2022, en los talleres del metro, había 96 trenes de 384 en la espera de mantenimiento y refacciones [13].

Nuestro proyecto pretende analizar la afluencia de usuarios en Metro, Metrobús y RTP para encontrar las mejores fechas o temporadas para que se les de mantenimiento a las unidades. Como apoyo a nuestro análisis, incluiremos la contaminación del aire como variable de análisis para determinar si esta tiene relación con la afluencia a los servicios de transporte.

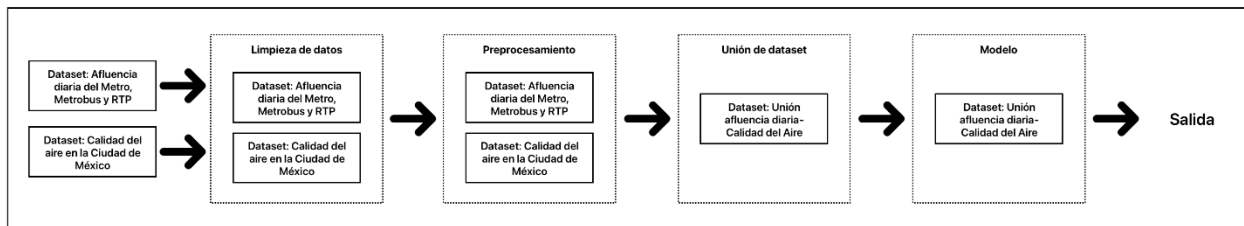
Los resultados de esta investigación aportaran elementos para que las personas encargadas de planear y realizar el mantenimiento en estas redes de transporte puedan establecer las fechas adecuadas para llevarlo a cabo. De tal forma, que se puedan reducir los fallos, así como las interrupciones del servicio.

Los resultados de esta investigación aportaran elementos para que las personas encargadas de planear y realizar el mantenimiento en estas redes de transporte puedan establecer las fechas adecuadas para llevarlo a cabo. De tal forma, que se puedan reducir los fallos, así como las interrupciones del servicio.

Para alcanzar nuestro objetivo, se usarán datos públicos que ofrece el portal de datos abiertos de la Ciudad de México y el sistema de monitoreo atmosférico, con los que tendremos acceso a la afluencia diaria por estación del Metro, Metrobús y RTP, de igual forma se usarán las métricas de los componentes contaminantes del índice IMECA, las cuales son ozono (O3), partículas menores a diez micrómetros (PM10), dióxido de azufre (SO2), dióxido de nitrógeno (NO2) y monóxido de carbono (CO).

4. Productos o Resultados esperados

Derivado de las predicciones y el análisis de los datos, se espera obtener el desglose de las temporadas en las que es factible hacer el mantenimiento a cada uno de los servicios, de movilidad, así como las estaciones, el tipo de trabajo que se realizará, e incluso, algunas sugerencias para neutralizar las intermitencias en el servicio que se puedan generar.



5. Metodología

La metodología con la que se desarrollará nuestro proyecto es SEMMA, su acrónimo es *Sample, Explore, Modify, Model* y *Access*. Esta metodología está enfocada en proyectos de Ciencia de Datos debido a que el flujo de trabajo abarca la obtención de datos, exploración de datos, selección de variables, creación y entrenamiento de un modelo y, por último, la evaluación de dicho modelo. La metodología SEMMA está estructurada para ser cíclica, es decir, en caso de que se realice un cambio en alguna etapa, se tiene que retomar desde esa etapa para modificar todas las partes posteriores, por lo que se pueden realizar todas las modificaciones convenientes. [14]

Ahora que conocemos el flujo de trabajo que se llevara a cabo, a continuación, se explicaran las fases aplicadas en nuestro proyecto.

En el apartado de Sample, dentro del portal de datos de la Ciudad de México, se obtendrán los datos acerca de la afluencia del Metro, Metrobús y RTP, mientras que los datos relacionados a la calidad del aire se conseguirán en el portal que ofrece la “Dirección de Monitoreo Atmosférico”. El conjunto de datos del Metro almacena registros sobre la afluencia diaria desde el 2010, los datos históricos del Metrobús comienzan en el año 2005 y el *dataset* sobre el RTP reúne datos a partir del 2022. En cuanto al *dataset* de calidad del aire, se tienen registros desde 1986 hasta la actualidad.

En las secciones posteriores, la herramienta que será utilizada es el lenguaje de programación Python con sus respectivas bibliotecas para facilitar el desarrollo del proyecto.

En la fase de exploración se visualizarán los datos, de tal forma se hallarán relaciones y tendencias entre las variables a analizar. Al igual, esta sección se enfoca en encontrar anomalías dentro de los datos para que puedan ser tratados en el preprocesamiento de datos.

Una vez se entiendan las relaciones entre variables y el comportamiento de los datos, en esta sección, se pretende preparar los datos para la creación de un modelo. Para ello, es necesario usar técnicas de preprocesamiento como la codificación de variables categóricas, normalización, centrado de los datos, ajustes de asimetría, selección de variables, entre otras.

A continuación, es el paso en el que se genera un modelo de Machine Learning de acuerdo con la tarea que se realizara, las más conocidas son regresión, clasificación y *clustering*. El proyecto usará regresión, debido a que se predecirá la afluencia de cada servicio de transporte para el año 2024, y a partir de la predicción se seleccionarán las temporadas en las que se puede hacer mantenimiento.

En la última sección de la metodología SEMMA se centra en evaluar el desempeño de nuestro modelo con datos distintos con los que fue entrenado. A partir de las métricas de error, es posible verificar que tan bien generaliza el modelo.

6. Cronograma

Nombre del alumno: Carrillo Oviedo Ian

Actividad	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY
Obtención de datos.										
Exploración de los datos.										
Preprocesamiento de datos de afluencia de Metrobús y calidad del aire.										
Evaluación TT I.										
Creación del modelo de predicción.										
Pruebas										

Ajuste del modelo.										
Generación de documentación y reporte de análisis.										
Presentación de resultados.										
Evaluación de TT II.										

CRONOGRAMA Nombre del alumno: Gutiérrez González Diego

Actividad	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY
Obtención de datos.										
Exploración de los datos.										
Preprocesamiento de datos de afluencia de calidad del aire y Metro.										
Evaluación TT I.										
Creación del modelo de predicción.										
Pruebas										
Ajuste del modelo.										
Generación de documentación y reporte de análisis.										
Presentación de resultados.										
Evaluación de TT II.										

CRONOGRAMA Nombre del alumno: Marco Antonio Serrano Covarrubias

Actividad	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY
-----------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Obtención de datos.										
Exploración de los datos.										
Preprocesamiento de datos de afluencia de Metro y Metrobús.										
Evaluación TT I.										
Creación del modelo de predicción.										
Pruebas										
Ajuste del modelo.										
Generación de documentación y reporte de análisis.										
Presentación de resultados.										
Evaluación de TT II.										

7. Referencias

- [1] Gobierno de la Ciudad de México, «Sistema de Transporte Colectivo Metro,» 19 Junio 2019. [En línea]. Available: <https://www.metro.cdmx.gob.mx/comunicacion/nota/hace-52-anos-inicio-la-construccion-de-la-red-del-metro>. [Último acceso: 16 Abril 2023].
- [2] Gobierno de la Ciudad de México, «Metrobus,» 23 Mayo 2021. [En línea]. Available: <https://www.metrobus.cdmx.gob.mx/17MB>. [Último acceso: 16 Abril 2023].
- [3] Gobierno de la Ciudad de México, «Red de Transporte de Pasajeros,» 14 Julio 2020. [En línea]. Available: <https://www.rtp.cdmx.gob.mx/red-de-rutas>. [Último acceso: 16 Abril 2023].
- [4] El Economista, «El Economista,» 10 Enero 2023. [En línea]. Available: <https://www.eleconomista.com.mx/politica/Mantenimiento-del-Metro-de-CDMX-se-planea-mas-de-lo-que-se-hace-20230110-0066.html>. [Último acceso: 16 Abril 2023].

- [5] El Financiero, «El Financiero,» 12 Enero 2023. [En línea]. Available: <https://www.elfinanciero.com.mx/cdmx/2023/01/12/guardia-nacional-en-el-metro-fallas-son-por-falta-de-mantenimiento-no-por-seguridad-dice-sindicato/>. [Último acceso: 2023 Abril 16].
- [6] L. Gómez, Flores, «La Jornada,» 13 marzo 2020. [En línea]. Available: <https://www.jornada.com.mx/2020/03/13/capital/030n1cap>. [Último acceso: 24 abril 2023].
- [7] Gobierno de la Ciudad de México, «La Nueva Línea 1,» [En línea]. Available: <https://www.lanueval1.cdmx.gob.mx/informacion/>. [Último acceso: 17 Abril 2023].
- [8] El Economista, «El Economista,» 12 Enero 2023. [En línea]. Available: <https://www.eleconomista.com.mx/politica/Necesarios-para-el-mantenimiento-del-Metro-3500-mdp-20230112-0005.html>. [Último acceso: 17 Abril 2023].
- [9] H. B. C. E. V. J. S. I. Chávez Millán Diego Eduardo, «Tesis IPN,» Mayo 2014. [En línea]. Available: <https://tesis.ipn.mx/jspui/bitstream/123456789/22547/1/Sistema%20de%20computo%20para%20el%20mantenimiento%20preventivo%20de%20los%20trenes%20del%20stc%20metro.pdf?fbclid=IwAR2kz9GKak7yqsGvsyhMlik6TNKBN5BSnTx0vOXoPybBOlrPL6HfZfP7hL0#:~:text=En%20a%20actua>. [Último acceso: 18 Abril 2023].
- [10] Fluke, «Fluke,» [En línea]. Available: <https://www.fluke.com/en-us/products/fluke-software/emaint-cmms>. [Último acceso: 18 Abril 2023].
- [11] SafetyCulture, «SafetyCulture,» [En línea]. Available: <https://safetyculture.com/es/>. [Último acceso: 18 Abril 2023].
- [12] INEGI, «ESTADÍSTICA DE TRANSPORTE URBANO,» 13 agosto 2021. [En línea]. Available: https://www.inegi.org.mx/contenidos/saladeprensa/notasinformativas/2021/ETUP/ETUP2021_08.pdf. [Último acceso: 18 abril 2023].
- [13] C. Roberto Barco , «El Sol de México,» 12 octubre 2022. [En línea]. Available: <https://www.elsoldemexico.com.mx/metropoli/cdmx/metro-cdmx-asi-es-el-mantenimiento-para-que-los-trenes-sigan-andando-9026510.html#:~:text=Fallas%20o%20aver%C3%ADas%3A%20los%20trenes%20entran%20a%20taller,a%20los%2010%20o%2012%20mil%20kil%C3%B3metros%20rec>. [Último acceso: 18 abril 2023].
- [14] N. Hotz, «Data Science Process,» 31 enero 2023. [En línea]. Available: <https://www.datascience-pm.com/semma/>. [Último acceso: 20 abril 2023].
- [15] a. Perez, «la poderosa,» 7 enero 2011. [En línea]. Available: <https://www.nqa.com/es-mx/certification/standards/iso-9001>.

8. Alumnos y Directores

Ian Carrillo Oviedo.- Alumno de la carrera de Lic. en Ciencia de Datos en ESCOM, , Boleta: 2021630204, Tel. 5516575424 , email iancaov06.uni@gmail.com

Firma: _____



Diego Gutiérrez González.- Alumno de la carrera de Lic. en Ciencia de Datos en ESCOM, Boleta: 2021630808, Tel. 5581576681 , email diego.gugo@outlook.com

Firma: _____



Marco Antonio Serrano Covarrubias.- Alumno de la carrera de Lic. en Ciencia de Datos en ESCOM, Boleta: 2021630591, Tel. 5536334836 , email marcoaserrano2001@gmail.com

Firma: _____



Martínez Acosta Lilian.- Doctorado sociología por la UAM, línea de investigación en impactos del desarrollo tecnológico. Titular C de tiempo completo en ESCOM-IPN, áreas de interés: desarrollo tecnológico, E-mail: lmartineza@ipn.mx

Firma: _____



CARÁCTER: Confidencial
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.
PARTES CONFIDENCIALES: Número de boleta y teléfono.

Corona Bermudez Uriel.- M. en C. por parte del IPN,
línea de investigación enfocada en Aprendizaje
Automático. Profesor de asignatura en ESCOM. Tel:
5565117721 E-mail ucoronab@ipn.mx

Firma:  _____