



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

MODELAMIENTO DE CAMBIOS EN LA DEMANDA DE TRANSPORTE PÚBLICO EN
SANTIAGO DE CHILE USANDO MODELOS DE DECISIÓN DISCRETA Y REDES
NEURONALES DE GRAFOS

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN COMPUTACIÓN

SEBASTIÁN ALEJANDRO MONTEIRO PARADA

PROFESOR GUÍA:
EDUARDO GRAELLS-GARRIDO

SANTIAGO DE CHILE

2025

Resumen

Estimar el efecto de cambios de demanda en el transporte público de Santiago de Chile debido a cambios de oferta es un desafío complejo, pues estos cambios provocan efectos en cadena en toda la red. Es por ello que se propone una solución en base a modelos de decisión discreta y redes neuronales de grafos para simular estos datos de demanda sintéticos, y así tomar decisiones con mejores datos.

Para ello, se utilizaron los datos de viajes proporcionados por Red Metropolitana de Movilidad, enriquecidos por ADATRAP para modelar la elección de ruta pasando por tres ejes (o componentes) fundamentales del trabajo, el modelado del mundo, en este caso, un grafo bipartito que simula a la red completa, un motor de utilidades, en este caso, el GNN/MNL y experimentos para comprobar y poner a prueba estos modelos.

El modelo MNL resultó ser interpretable y preciso, con un 89% de precisión no trivial. El hallazgo principal en cuanto a las características fue la preferencia de los usuarios a viajes sin transbordo aunque eso signifique un tiempo de viaje mayor.

El modelo GNN logró una mejoría al tener un 91 % de precisión no trivial. Se comparan GNN con capas de decisión discreta activadas y desactivadas. Debido a la leve mejoría del GNN, se opta por hacer experimentos con el MNL gracias a su velocidad e interpretabilidad.

Finalmente, se realizan experimentos para poner a prueba el modelo del MNL, en este caso, se cambian frecuencias de servicios y se suspenden otros, analizando localmente la redistribución de la demanda, observando una redistribución de la demanda en servicios varios. Para terminar, se agrega la Línea 7 del Metro de Santiago pronta a inaugurarse, para analizar los cambios de la demanda. Se observa un aumento de carga general en el metro, especialmente en la L1; un aumento de la cantidad de transbordos y un aumento en la demanda de servicios alimentadores a la línea en cuestión.

Se concluye que los experimentos evidencian fortalezas de la solución, como su capacidad de generar datos sintéticos, pero también limitaciones debido a los datos disponibles.

Tabla de Contenido

1	Introducción	1
2	Revisión de la literatura	3
2.1	Contexto	3
2.1.1	Objeto de la predicción	3
2.1.2	Tipo de datos	4
2.1.3	Factores	4
2.1.4	Actualmente en Chile	5
2.2	Origen y representación de los datos	6
2.2.1	Red Metropolitana de Movilidad	6
2.2.2	Datos de RED y ADATRAP	7
2.2.3	Grafo	9
2.2.4	Algoritmos de ruteo	13
2.3	Modelos de predicción	14
2.3.1	MNL	14
2.3.2	GNN	15
3	Metodología	17
3.1	Solución propuesta	17
3.2	Representación de los datos	19
3.2.1	Grafo Agrupado	19
3.2.2	Grafo Bipartito	21

3.2.3	Frecuencias de los servicios	22
3.2.4	Velocidades promedio de los servicios	22
3.2.5	Nodos	22
3.2.6	Aristas	23
3.2.7	Algoritmo para la creación del grafo bipartito	24
3.3	MNL	25
3.3.1	Métricas de entrenamiento	25
3.3.2	<i>Dataset</i>	26
3.3.3	Entrenamiento	29
3.4	GNN	31
3.4.1	Arquitectura de la solución	31
3.4.2	Datos	32
3.4.3	Embeddings iniciales	32
3.4.4	GNN Bipartito	32
3.4.5	Normalización y contexto	33
3.4.6	<i>Features</i> y concatenación	33
3.4.7	MLP (Perceptron Multi Capa) Scorer	33
3.5	Experimentos	34
3.5.1	Disminución de la oferta de un servicio	34
3.5.2	Suspensión de un servicio	34
3.5.3	Agregar línea 7	34
4	Resultados y discusiones	37
4.1	Representación de los datos	37
4.1.1	Grafo agrupado	37
4.1.2	Grafo bipartito	38
4.2	MNL	40
4.2.1	Resultados y coeficientes	40

4.2.2	Primeros transbordos y la penalización inicial	41
4.2.3	Entrenamiento diario	42
4.2.4	Entrenamiento semanal completo	42
4.3	GNN	46
4.3.1	Resultados	46
4.4	Experimentos	50
4.4.1	Disminución de oferta de un servicio	50
4.4.2	Suspensión de un servicio	55
4.4.3	Agregar Línea 7	57
4.5	Discusión final de la solución	60
5	Conclusión	63

Índice de Tablas

3.1	Descripción de las columnas del dataset de decisiones para el modelo MNL.	28
3.2	Estaciones y coordenadas. El nombre de las estaciones no es el oficial. Están en orden partiendo desde Brasil hasta Estoril. Entre paréntesis aparecen las combinaciones. "esq" abrevia "esquina"	35
4.1	Coeficientes obtenidos del modelo MNL (todos nulos).	40
4.2	Métricas de entrenamiento y validación del modelo MNL.	40
4.3	Coeficientes del modelo MNL entrenado para distintos días de la semana. La última fila muestra el promedio de cada columna. Este entrenamiento fue por cada día.	42
4.4	Resumen de datos y particionado	43
4.5	Coeficientes del modelo MNL	45
4.6	Métricas de la mejor época de la MNL Semanal	45
4.7	Resumen del modelo	46
4.8	Resumen del modelo	48
4.9	Mejor modelo y evaluación en test	48
4.10	Coeficientes del modelo MNL usado para los experimentos	50
4.11	Itinerarios de las alternativas desde T-11-64-P0-30 hasta T-20-177-P0-20	54
4.12	Uso de Servicios (Etapas Reales)	57
4.13	Uso de Servicios (Etapas Sintéticas)	58
4.14	Servicios Alimentadores Principales de L7 (Transbordos)	59
4.15	Comparación de Carga de Metro (Total Estaciones-Pasajero Recorridas)	59

Índice de Ilustraciones

2.1	Zonas tarifarias 777 en Santiago	9
2.2	Esquema resumen del Grafo Bipartito	12
3.1	Diagrama general de la solución en donde se subdivide esta en 4 secciones principales, las fuentes de datos, la representación de estos, el modelo de elección discreto y el GNN	17
3.2	Diagrama del entrenamiento del MNL	30
3.3	Arquitectura de la GNN	31
3.4	Esquema direccional propuesto para la Línea 7	35
4.1	Mapa en Plotly con zoom a un barrio de Cerro Navia	38
4.2	Histograma de Pesos. El eje x representa el tiempo y el eje y la frecuencia	41
4.3	Resultados del modelo MNL. En todos los gráficos, el eje x es la época y el eje y la métrica correspondiente.	44
4.4	Resultados de entrenamiento de la GNN con características del MNL activadas	47
4.5	Resultados de entrenamiento de la GNN sin características del MNL	48
4.6	Costes para ir desde PJ394 a PA300. En azul se muestran los costes reales y en naranja los cambios de oferta. Notar como el tiempo de espera del servicio 517 aumenta (la barra naranja es mas alta que la azul).	51
4.7	Distribución de probabilidades para alternativas de viaje antes y después del cambio de oferta	51
4.8	Costes para ir desde PJ394 a PA433	52
4.9	Distribución de probabilidades para alternativas de viaje antes y después del cambio de oferta (Experimento 2)	52

4.10	Distribución de probabilidades para alternativas de viaje con un aumento del 1500% en el tiempo de espera del servicio 507	53
4.11	Distribución de probabilidades para alternativas de viaje con suspensión de la Línea 1	55
4.12	Costes para ir de San Pablo a Baquedano con suspensión de la Línea 1, antes y después de colocar un tiempo de espera en L1 muy grande.	55
4.13	Costos del paradero E-20-53-PO-115 a Tobalaba con suspensión de la Línea 1 . . .	56
4.14	Distribución de probabilidades para alternativas con suspensión de la Línea 1 . . .	56

Capítulo 1

Introducción

El sistema de transporte público en Santiago de Chile es un componente esencial para el funcionamiento de la ciudad. Cambios en su oferta, sean planificados o inesperados, pueden generar impactos significativos en la movilidad en los barrios cercanos, tanto a corto como a largo plazo.

La planificación o intervenciones a la red de transporte enfrentan el desafío de predecir efectos colaterales en la red existente. Crear un modelo de decisión en base a alternativas ofrecidas al usuario para llegar a un destino dado, un origen y hora, para luego con ese mismo modelo, generar una demanda sintética de una red de transporte en base a cambios nuevos (una redistribución de la demanda) es el objetivo de esta memoria.

Históricamente se han usado soluciones discretas como las MNL (Multinomial Logit) y recientemente modelos como GNN (Redes Neuronales de Grafos) para la predicción de la demanda. Por el lado de la MNL, esta solución se enfoca en una ingeniería de características la cual propone variables de interés en el proceso de decisión de un usuario, tales como el tiempo de espera, la velocidad o que tanto acerca al destino el servicio de transporte público. Por otro lado, es importante notar la correlación espacial de la red, sobre todo si ocurren cambios en ella. Debido a la naturaleza de Red o **set de puntos conectados**, usar Redes Neuronales de Grafos es un área de investigación llamativa, gracias a su versatilidad en la forma de los inputs, a diferencia de los Perceptrón Multi Capa (MLP) o las Redes Neuronales Convolucionales (CNN) las cuales reciben vectores o grillas (como las imágenes).

Actualmente, algunos de los estudios que abordan esta problemática desde Chile lo hacen desde enfoques estadísticos y/o a nivel macro. Estos suelen analizar el antes y el después de una intervención, sin capacidad real de abstracción. Otros modelos tienen una orientación más predictiva, pero se encuentran desactualizados y no reflejan adecuadamente las dinámicas actuales del transporte urbano. También existen enfoques centrados en el transporte privado, que estudian cómo factores como la infraestructura, las tarifas o las políticas públicas afectan la movilidad general. Sin embargo, estos trabajos no se enfocan en cambios estructurales de la red de transporte público, sino que operan sobre la oferta ya existente. Por otro lado, existe el sistema ADATRAP [12], desarrollado por la Universidad de Chile y el Instituto Sistemas Complejos de Ingeniería. ADATRAP es un software que analiza datos y permite planificar y crear estrategias para la priorización en la asignación de servicios públicos de transporte. El software toma en

cuenta la distribución de la oferta para los usuarios del servicio en la Región Metropolitana. ADATRAP será una fuente de datos importante para la predicción de la demanda.

La solución propuesta en este proyecto se basa en el uso de técnicas de aprendizaje automático, modelando el sistema de transporte como un grafo en el que se representen recorridos, paradas y transbordos. Este modelo (sea uno discreto como una MNL o una GNN) tendrá que predecir el proceso de decisión de los usuarios en función de múltiples factores, como la duración del viaje, el número de transbordos y el tiempo de espera. Estos modelos y sus resultados se compararán con datos reales de uso, para afinar el modelo y su precisión.

Objetivo general

Es por ello, que el objetivo de esta memoria es: *Diseñar e implementar un modelo que prediga demanda de transporte dado un escenario (definido como una configuración de red y su respectiva infraestructura urbana); y usar este modelo para predecir demanda en distintos escenarios para medir el impacto de intervenciones en el escenario actual.*

Objetivos específicos a cumplir

1. (OE1) Construir una representación topológica de la red de transporte público de Santiago en un grafo, integrando datos de operación como velocidades, frecuencias y trazados.
2. (OE2) Diseñar, entrenar y evaluar modelos de elección discreta y de aprendizaje profundo, determinando la arquitectura más adecuada para simular la toma de decisiones de los usuarios, usando como datos la demanda histórica de viajes en transporte público de ADATRAP.
3. (OE3) Evaluar el impacto eventos de reconfiguración de la red de transporte público mediante la generación de datos de demanda de viajes sintéticos.

El cuerpo de la memoria comienza con el capítulo 2, en el que se hace una revisión de la literatura, datos, modelado de la red y modelos de predicción. En el capítulo 3 se presenta una metodología para construir una representación de los datos e implementación y entrenamiento del MNL/GNN para luego terminar con los experimentos. El capítulo 4 presenta los resultados y discusión inmediata de ellos. El capítulo 5 presenta la conclusión de la memoria.

Capítulo 2

Revisión de la literatura

Una gran motivación para este proyecto es la optimización en el uso de los recursos públicos. Modificar dinámicamente la frecuencia de los buses, crear nuevos recorridos o eliminar aquellos que han quedado obsoletos son decisiones que pueden tener un impacto significativo en la calidad del servicio y en la satisfacción de los usuarios. Sin embargo, estas decisiones deben basarse en datos precisos y actualizados sobre el uso del transporte público, así como en una comprensión profunda de cómo los cambios en la red afectan la demanda.

La siguiente sección se dividirá en dos, la primera, una revisión de literatura para entender el contexto y el estado del arte en el ámbito de la predicción de uso de transporte público, y la segunda, un marco teórico que explora las técnicas de predicción y las herramientas a utilizar en la memoria.

2.1 Contexto

Siguiendo el trabajo de Torrepadula más a fondo [14], un paper review sobre este objeto de estudio, se abren muchas soluciones y consideraciones: La primera, el objeto de la predicción.

2.1.1 Objeto de la predicción

Diversos trabajos se enfocan tanto en:

1. Cantidad de personas en una parada en la ruta. Trabajos como el de Wei et al. [16] usan enfoques no lineales para estimar la demanda en algunas estaciones de metro.
2. Cantidad de personas en la ruta. El trabajo de Zhao [19] utiliza Prophet para estimar las personas en la ruta 320 de Zhengzhou, China
3. Cantidad de personas en un vehículo. Algunos trabajos lo predicen, como el de Wang et al.[6] con un Support Vector Machine más un filtro de Kalman.

4. Cantidad de personas en un área. El trabajo de Wang et al. [15] explora predicciones espacio temporales con un modelo llamado GALLAT (GrAphic preddiction with ALL ATtention), que modela la red como un grafo

Notar que cada enfoque u objeto requiere un set de datos distintos, por ejemplo, para saber cuanta gente hay en un momento dado en un vehículo, se deben usar cámaras o sensores, en cambio, para saber un estimado de gente en la ruta, se usan los datos de las validaciones de la tarjeta bip! en la ruta.

2.1.2 Tipo de datos

Algunos ejemplos son:

1. Datos de validación de la tarjeta Bip! (que se puede usar para saber cuanta gente hay en una ruta, o en un área).
2. Datos de sensores (que se pueden usar para saber cuanta gente hay en un vehículo).
3. Datos de cámaras (que se pueden usar para saber cuanta gente hay en un vehículo, o en un área o un paradero).
4. GPS para el flujo de personas en un área.

Analizando el trabajo de Torrepadula [14], los datos de validación de la tarjeta , como la Bip! o sus equivalentes en otros países son los más utilizados, ya que son fáciles de obtener y tienen una buena cobertura geográfica. Sin embargo, también tienen limitaciones, como la falta de información sobre el origen y destino de los viajes. Los datos de sensores y cámaras son más precisos, pero son más difíciles de obtener y tienen una cobertura geográfica limitada. Los datos de GPS son muy precisos, pero también son difíciles de obtener y tienen una cobertura geográfica limitada.

Trabajos como los de Ye [18], Jian [4] y Li [5] utilizan datasets provenientes de tarjetas de validación con tecnología similar o idéntica a la de la tarjeta Bip!.

2.1.3 Factores

Los factores que afectan la demanda son diversos y pueden variar según el contexto. Algunos de los más relevantes son:

1. **Tarifas:** El costo del transporte público puede influir en la demanda, especialmente en áreas donde existen alternativas de transporte privado.
2. **Frecuencia:** La cantidad de buses o trenes disponibles en una ruta puede afectar la demanda, ya que una mayor frecuencia puede atraer a más usuarios.

3. **Tiempo de viaje:** La duración del trayecto es un factor clave en la decisión de utilizar el transporte público. Un tiempo de viaje más corto puede aumentar la demanda.
4. **Comodidad:** La calidad del servicio, como la limpieza, el confort y la seguridad, puede influir en la decisión de utilizar el transporte público.
5. **Accesibilidad:** La facilidad de acceso a las paradas o estaciones, así como la disponibilidad de servicios complementarios (como estacionamientos o bicicletas compartidas), puede afectar la demanda.
6. **Condiciones climáticas:** Factores como la lluvia, el frío o el calor extremo pueden influir en la decisión de utilizar el transporte público.
7. **Eventos especiales:** La realización de eventos masivos, como conciertos o ferias, puede generar picos de demanda en ciertas rutas.
8. **Fiestas y feriados:** La demanda de transporte público puede variar significativamente durante días festivos o feriados, lo que puede afectar la planificación de la oferta.
9. **Búsqueda Web** Los turistas, generalmente, se informan de las rutas y horarios de los buses en la web, por lo que el tráfico web puede ser un buen indicador de la demanda. También aplicaciones móviles de mapas pueden sugerir rutas de transporte público, lo que puede influir en la demanda.

2.1.4 Actualmente en Chile

Hoy en día, la red está enfrentando transformaciones importantes. La construcción e implementación de nuevas líneas de metro, como la Línea 7 y la futura Línea 8, tendrán un efecto profundo sobre el uso de ciertos recorridos de buses. Algunos servicios podrían volverse redundantes, mientras que otros —como los recorridos locales tipo [LETRA]-XX— podrían experimentar un aumento significativo en la demanda, al convertirse en alimentadores hacia las nuevas estaciones. Esta situación presenta una oportunidad para replantear frecuencias, redistribuir flotas y mejorar la eficiencia general del sistema.

El más destacado es ADATRAP, desarrollado por la Universidad de Chile y el Instituto Sistemas Complejos de Ingeniería. Este software permite analizar datos y planificar estrategias para la priorización en la asignación de servicios públicos de transporte. ADATRAP toma en cuenta la distribución de la oferta para los usuarios del servicio en la Región Metropolitana.

ADATRAP [12] es un software que utiliza la información geotemporal referenciada (GPS) en buses de Transantiago, en conjunto con la información que entrega la tarjeta Bip!, con el objetivo de estimar desempeño de transporte público, velocidades de traslado, hacinamiento, perfiles de carga, etc. Logra crear perfiles de velocidad por servicio y por tramo de ruta, perfiles de carga por servicio, matrices origen-destino, indicadores de calidad de servicio. El software está registrado a nombre de la Universidad de Chile y transferido mediante acuerdo de licencia a la Subsecretaría de Transportes. Se utiliza diariamente para tomar decisiones tales como la definición semanal de programas de operación, modificación de servicios y decisiones de infraestructura

Estos fenómenos (el cambio de demanda) han sido objeto de análisis en trabajos previos. Un ejemplo representativo es el de Ramírez [11], quien estudia el cambio espacial en la demanda de

transporte público tras la apertura de una nueva línea de metro, empleando un enfoque estadístico. Si bien su análisis es útil para evaluar efectos pasados y prever algunos eventos futuros, su falta de generalización da una ventana de oportunidad.

Por otra parte, el trabajo de Camus [1] propone una simulación basada en agentes dentro de la red de transporte público. Sin embargo, dicho modelo considera la oferta como un elemento estático y no contempla escenarios en los que esta pueda ser modificada.

También existe el modelo desarrollado para el Directorio de Transporte Público Metropolitano (DTPM) [2], mediante el software EMME de Bentley. Algunas características del modelo de demanda generado se basan en entradas como el diseño, la demanda y los datos operacionales. Luego, puede predecir y simular el impacto de cambios en la infraestructura para planear cambios. Este modelo de demanda fue creado con el plan de operación de 2020 (Marzo) y la demanda de 2019 (Agosto). Las franjas horarias (o períodos de análisis) son 3, Punta Mañana, Fuera de Punta Mañana y Punta Tarde. En el documento expuesto por la DTPM, el proceso de ajuste de matrices de viaje no contó con los aforos de la zona oriente, pero pudieron ser subsanados con datos anteriores, aunque se reconoce una posible subestimación de la demanda en esa zona. Al no ser de código abierto el software, no mucha más información se puede recabar.

Asimismo, existen modelos de demanda agregada, como el desarrollado por Méndez [13], que se apoyan en técnicas econométricas y estudian elasticidades en función de variables como tarifas o cantidad de servicios disponibles. Aunque valiosos, estos trabajos no abordan cambios estructurales en la red, sino que se enfocan en la oferta existente.

2.2 Origen y representación de los datos

Transformar los datos en una *estructura de datos* es un paso importante para que los algoritmos de ruteo entreguen características de cada alternativa del viaje. GNNs requieren preprocesar los datos en matrices o grafos. Trabajos como los de Liu et al.[7] utilizan grafos representados por matrices del tipo (o,d), donde o es el origen y d es el destino de la persona. Otros enfoques, como el de Massobrio[8] modelan una red con nodos que representan las paradas de las rutas.

2.2.1 Red Metropolitana de Movilidad

Red Metropolitana de Movilidad es el sistema de transporte público de Santiago, el cual opera mediante empresas como Metro S.A, el tren urbano Estación Central-Nos y el sistema de Buses anteriormente llamado Transantiago .

Su operación se basa en distintas unidades de negocio (UN) que trabajan en conjunto con un set de servicios acordados mediante licitaciones. Los datos necesarios para crear el grafo y enriquecerlo con costes para entrenar están disponibles de manera pública en los planes de operaciones y en las matrices de viaje.

Recaltar algunas características que impone el sistema de transporte a la solución. El usuario solo valida su tarjeta en el paradero de subida, esto implica que no se sabe con inequívocidad donde

se baja el usuario. Además, el coste de un pasaje depende de la hora, siendo estas franjas tarifarias valle y peak. Otro caso borde incluye a las zonas paga, en las cuales se valida en un tótem fuera del servicio, por lo que se pierde información sobre el servicio a bordo real del usuario una vez se sube al vehículo. En estaciones de metro de combinación un usuario puede validar su tarjeta en el tótem de una de las líneas, pero combinar inmediatamente y subirse a la otra.

2.2.2 Datos de RED y ADATRAP

ADATRAP entrega datos de viajes y etapas. Los datos están públicos en el siguiente enlace: <https://www.dtpm.cl/index.php/documentos/matrices-de-viaje>. Cada viaje tiene n etapas, hasta 4 como máximo.

Cada viaje tiene un origen y un destino. El sistema de transportes capitalino no posee validación de la Bip! o sus derivados al término de la etapa, por lo que la estimación de este parámetro fue realizada por el software ADATRAP. ADATRAP analiza los patrones de viaje de usuarios para detectar donde se sube y baja. Tomar el siguiente caso:

Un usuario sube a las 7:00 AM en el servicio X en el paradero P, y aborda a las 19:00 en el servicio Y en el paradero P', esto con cierta regularidad, entonces se concluye que en la mañana el usuario se bajó cerca del paradero P' usando el servicio X, y que en la tarde el usuario se bajó cerca del paradero P en el servicio Y.

Tabla de viajes y etapas

Las tablas de viajes y etapas serán las demandas históricas.

La tabla de viajes contiene la información de los viajes del usuario, registrando hasta 4 etapas o 3 combinaciones. Combinaciones en metro no cuentan, pues no se valida la tarjeta al cambiar de línea. Cada tabla de viajes o de etapas corresponde a un solo día de análisis. Las tablas de viajes y de etapas vienen generalmente en packs de una semana completa.

Código TS y código usuario

Los servicios y paraderos se encuentran codificados en formato TS, esto es, un código interno usado por DTPM para identificar a los recorridos. La mayoría de los recorridos tiene un código TS que coincide con el de usuario. Por ejemplo, el servicio **T507 OOI** codifica al servicio 507 de ida (servicio en sentido ENEA- AV GRECIA). En algunas ocasiones no coincide, esto ocurre mayoritariamente en servicios locales con prefijo alfabético, casos como el servicio con código de usuario **J01** en código TS es en **T521**. Esta es la razón por la cual algunos recorridos nuevos tienen códigos de usuario que no siguen el numerado del usuario, ya que si lo siguieran, habrían colisiones de nombres.

Por otro lado, los códigos de paradero también poseen esta distinción. Ningún código de paradero de usuario coincide con su versión en TS. En el set de datos de tabla de viajes y de etapas

ambos códigos, tanto el de paraderos como el de servicios vienen en código TS.

Paraderos subida y bajada

Ambas en código TS, denotan, para las 4 posibles etapas, las subidas y bajadas del usuario. Máximo 8 (2 por cada etapa).

Horas de subida y bajada

Estimados con la velocidad promedio de los buses y los itinerarios, cada etapa tiene un horario de subida y bajada. Máximo 8 (2 por cada etapa). Estos se pueden separar en bins de 30 minutos cada uno. En resumen, 48 bins de tiempo. Se denominan en el lenguaje de ADATRAP como mediahora.

Servicios de las 4 etapas

En formato TS. Servicio de cada etapa. Máximo 4 (1 por cada etapa).

Hay más columnas, pero para el análisis posterior no son de relevancia. La tabla de etapas contiene la misma información pero de manera disgregada, es decir, cada fila es una etapa.

Consolidado de recorridos

Para crear el grafo, lógicamente es necesario el trazado de todos los recorridos de RED. Para ello, se RED tiene en su página web el trazado activo hasta ahora. Este archivo contiene en sus columnas:

1. Los códigos de los servicios y paraderos en TS y en formato usuario.
2. El nombre del paradero.
3. Excepciones del paradero.
4. Las posiciones X e Y del paradero.(UTGSM)

Cada fila contiene una parada de un trazado de un servicio.

Algo importante a notar es la fecha de esta tabla de recorridos. Es válida desde el 31/05/2025 hasta a fin de año (al momento de hacer este informe)

Zonas 777

Red delimita Santiago en zonas, llamadas Zonas777. Estas están accesibles tanto desde la tabla de etapas (es decir, la subida y bajada denota en que zona ocurrieron), como también en el

consolidado en cada paradero. Además, RED entrega archivos SHAPE que pueden analizarse con GEOPANDAS (paquete del lenguaje de programación Python) para visualizar las zonas 777. La Figura 2.1 muestra las zonas 777 de Santiago.

Mapa Interactivo de Zonas 777

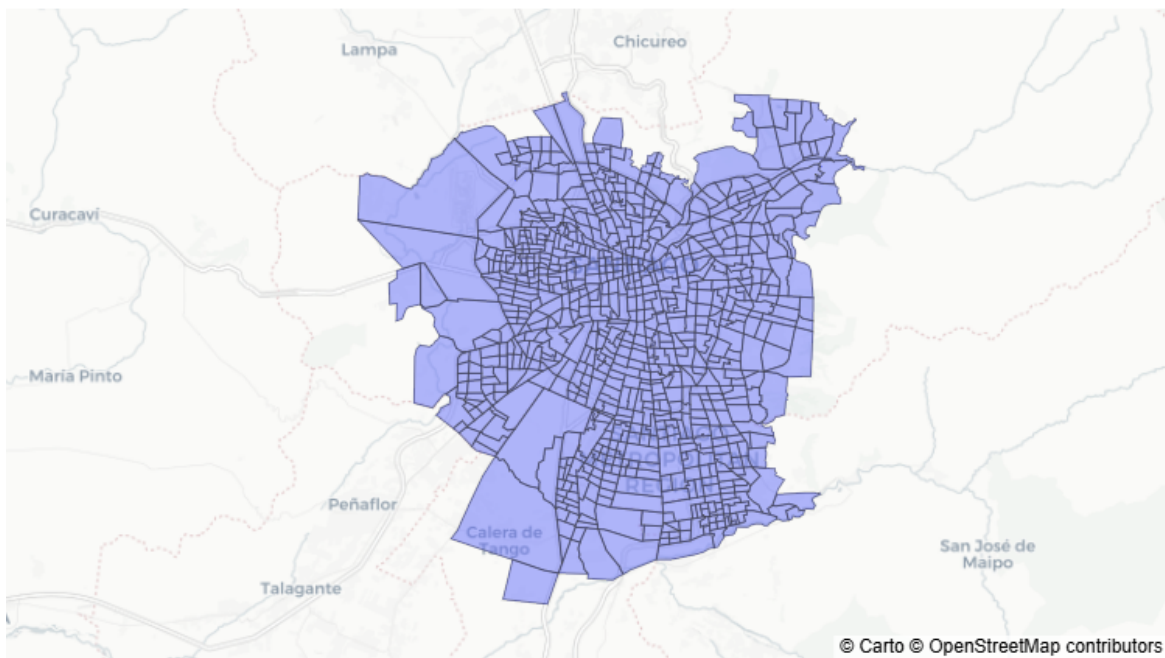


Figura 2.1: Zonas tarifarias 777 en Santiago

Con todo esto presente, se hacen las siguientes afirmaciones:

- La demanda de origen y destino es inamovible en esta solución, es decir, cada usuario tiene que comenzar y terminar su viaje en los paraderos reales históricos. Esto debido a que no se sabe donde vive exactamente cada usuario, ni donde trabaja o estudia.
- La demanda de transición (las bajadas y subidas interetapas) pueden cambiar si es que la oferta cambia, es decir, si es que el usuario decide que es mejor hacer transbordo en un paradero P en la zona777 z1 en vez del paradero Q en la zona777 z2. Esto es perfectamente posible. La idea de la redistribución de la demanda propuesta en esta memoria, es mover el trayecto de una persona con dos puntos fijos, el origen y el destino final.

2.2.3 Grafo

Un grafo $G(E,V)$ es un conjunto de aristas(E) y vertices(V). Estos pueden ser dirigidos (los vértices tienen dirección única) o no (ambas direcciones posibles). Se pueden asignar pesos a las aristas, que representan el coste de ir de un nodo a otro.

Considerar una red de transporte de la siguiente manera:

- Cada paradero P es un nodo V de este grafo. Por ejemplo, Estación Central es un paradero o PA433 es un paradero (el paradero justo fuera de la Facultad de Ciencias Físicas y Matemáticas).
- Cada arista E es una conexión entre dos paraderos, dada por un servicio. Se define un servicio como un recorrido de la red. Por ejemplo, la Línea 1 (L1) o la 506 son servicios.

En este sentido, se definen dos tipos de grafos que se utilizarán en la memoria, uno para exploración visual, el Grafo Agrupado, y otro para el uso de algoritmos de *ruteo* y las técnicas de aprendizaje que se mencionarán en un futuro, el Grafo Bipartito.

Grafo agrupado

El primer grafo, denominado grafo agrupado, es un grafo necesario para la visualización de la información. No es útil para entrenar modelos ni para ejecutar algoritmos de ruteo, pero permite un mayor entendimiento de la estructura de la red.

Consiste en un grafo el cual los nodos son los paraderos y las aristas son las conexiones consecutivas de un servicio al recorrer los paraderos. Un servicio S tiene un set de paraderos P_k con k el índice del paradero en el orden en que los recorre.

Por ejemplo, se tienen dos servicios S_1 y S_2 , ambos tienen un set de paraderos distintos P_k y Q_i , con P denotando los paraderos de S_1 y Q los de S_2 . Notar que un paradero V en el grafo (un nodo V) puede ser el paradero P_3 (el tercer paradero para S_1) y el paradero Q_4 (el cuarto paradero del recorrido S_2). Entonces, si es que el primer servicio recorre en un subset de paraderos en el mismo orden que el segundo servicio, la arista que une ambos paraderos tendrá la información de que es recorrida por ambos servicios. Es directo ver que este grafo no conviene para ejecutar algoritmos de ruteo, pues se pierde la información del peso de la arista, ya que S_1 puede demorarse más en recorrer la arista que S_2 .

Este grafo si conviene para visualizar la red, pues agrupa todos los servicios que recorren los mismos paraderos en el mismo orden en una sola arista. Cada nodo entonces tendrá la información de los servicios que paran en él y las aristas los servicios que conectan los nodos.

Grafo bipartito

Un grafo Bipartito en cambio se puede entender como un grafo de estado que captura las transiciones por las que pasa un usuario al tomar un transporte. Se entienden 4 cambios de fases. Se denomina bipartito pues se tienen dos tipos de nodos en el grafo y se puede dividir en dos conjuntos.

- *CAMINAR*: Se permite caminar entre paraderos para combinar o hacer transbordo si es necesario.
- *SUBIR*: Subir al servicio.
- *BAJAR*: Bajarse del servicio.

- VIAJAR: Viajar a bordo del servicio.

Para cada cambio de fase, se asocia un coste, estos son:

- COSTE DE CAMINAR: Calculado como la distancia entre dos paraderos caminando a una velocidad dada promedio.
- COSTE DE ESPERA : Calculado como el tiempo de espera esperado para subir al servicio.
- COSTE DE BAJAR : Es cero, pues bajar del servicio es un acto puntual.
- COSTE DE VIAJAR : Es el tiempo que tarda en recorrer un servicio entre dos paraderos de una etapa de viaje.

Como los cambios de fase con costes, es directo entender que los cambios de fases definen los tipos de aristas que habrá en el grafo.

Pero, cada cambio de fase o arista debe de mostrar la transición de un estado al otro. Se definen los siguientes dos estados (nodos):

- EN PARADERO: Cuando se espera al servicio.
- EN SERVICIO: Cuando se está en el servicio.

Entonces, entre dos estados EN PARADERO pueden haber aristas CAMINAR y aristas VIAJAR. En cambio, entre un PARADERO y un SERVICIO, pueden haber aristas SUBIR o BAJAR. La direccionalidad de las aristas es importante. Lógicamente la arista BAJAR irá desde el SERVICIO hasta el PARADERO. En cambio, la arista CAMINAR es bidireccional. Una arista VIAJAR es de un solo sentido en el caso de los buses, pero bidireccional en el METRO.

El diagrama de la Figura 2.2 muestra un esquema resumen del grafo Bipartito.

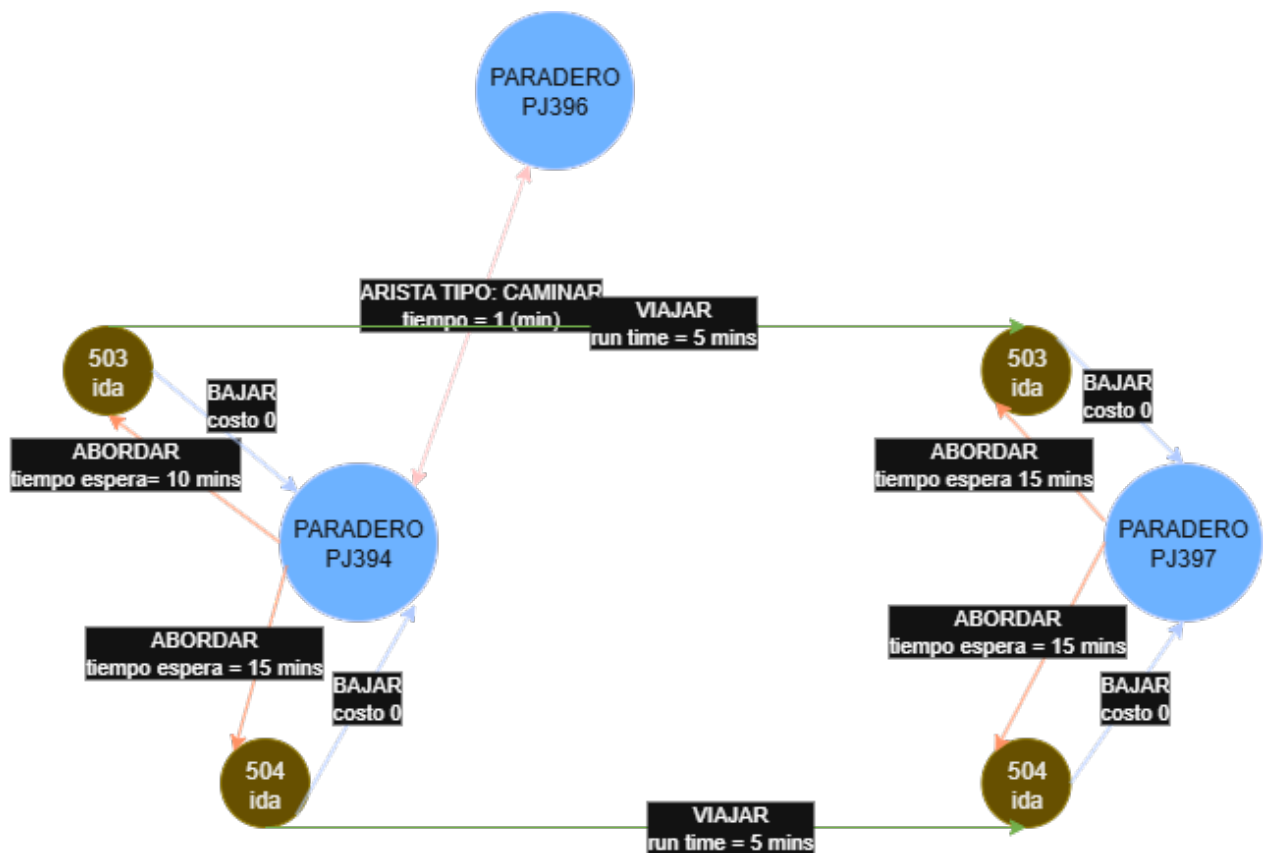


Figura 2.2: Esquema resumen del Grafo Bipartito

Notar como el grafo codifica cambios de estado en un viaje de un usuario. En este grafo, el usuario puede caminar, combinar y viajar.

Finalmente, un usuario si quiere ir de un paradero P al paradero Q, tendrá que:

- ESPERAR al servicio deseado (ABORDAR)
- VIAJAR hasta el destino o al paradero de bajada para el primer transbordo si lo amerita (N ARISTAS VIAJAR)
- BAJAR
- CAMINAR si es necesario o ESPERAR si es que el servicio deseado para en el mismo paradero de bajada
- Volver a VIAJAR ...
- BAJAR en el destino final.

Cada acción requiere pagar un coste.

Notar como el grafo bipartito es ideal para solucionar el problema, pues permite capturar los distintos costes que un usuario tiene que pagar al hacer un viaje, y además permite ejecutar algoritmos de ruteo fácilmente, los cuales denotan tanto el camino físico, como los cambios de estado que experimenta el usuario. Esto no se puede hacer con un grafo común no bipartito, pues se pierden los costes de subir (coste de espera) ya que no se modelan los estados de subir y bajar, estados claves para capturar los costes reales de un viaje.

2.2.4 Algoritmos de ruteo

Para saber el camino óptimo de un usuario dado un paradero, se usará el Algoritmo de Dijkstra, debido a su simplicidad, y porque en muchas librerías ya está implementado y muy optimizado. Dado dos nodos, el algoritmo de Dijkstra buscará el camino con menos coste entre ambos nodos.

No se usará A^* pues este algoritmo es útil para cálculos punto a punto, pero Dijkstra será usado para la generación de atributos en un espacio de estados completo, el grafo bipartito. Mas adelante se verá como el algoritmo de Dijkstra hace un barrido completo de toda la red por cada destino, generando paradas de bajada óptimas para cada paradero de origen y servicio.

Gracias a que el grafo indirectamente penaliza los transbordos, el algoritmo de Dijkstra podrá encontrar caminos con sentido lógico, en vez de cambiar de recorrido en cada paradero, hecho que pasaría en el grafo agrupado, en el cual no se penaliza hacer transbordos.

2.3 Modelos de predicción

Para predecir una decisión de un usuario, se exploraron distintos métodos, algunos son el MNL, el Logit Anidado, modelos de Machine Learning como Random Forests, XGBoost y Redes Neuronales como la GNN.

Se ahondará en dos modelos, el MNL y la GNN, pues son los que mejores resultados arrojaron en las pruebas preliminares.

2.3.1 MNL

El modelo MNL (Multinomial Logit Model) es un modelo de elección discreta que se utiliza para predecir la probabilidad de que un individuo elija una alternativa dentro de un set de ellas.

Por ejemplo, si un usuario tiene N alternativas de servicio en un paradero, sean S_1, S_2, \dots, S_n en un paradero de origen P y un destino Q , el modelo MNL permite predecir la probabilidad de que el usuario elija cada una de las alternativas en base a variables propuestas como *determinantes* por el investigador, las cuales considera importantes para la toma de decisiones, pero no se ponderan aún. El modelo será encargado de decir qué variable pondera más que otra en el proceso de entrenamiento. Notar que todas estas variables son *atributos* de la alternativa.

Tal como lo definen Ortuzar & Willumsen [10], una utilidad U_n se define como la suma de la parte determinística (los atributos) y una parte aleatoria ϵ_n que captura la incertidumbre del modelo.

$$U_n = V_n + \epsilon_n$$

V_n se construye como una función predictora lineal de los atributos ponderados por coeficientes β_i que representan la importancia de cada atributo en la decisión del usuario. Es decir:

$$V_n = \beta_1 x_{1n} + \beta_2 x_{2n} + \dots + \beta_k x_{kn} = \sum_{i=1}^k \beta_i x_{in}$$

Cada beta es un peso que pondera la importancia del atributo x_i en la decisión del usuario. Por ejemplo, si β_1 es muy grande, el atributo x_1 es muy importante en la decisión del usuario. Si β_2 es negativo, el atributo x_2 tiene un efecto negativo en la decisión del usuario.

Ejemplos de atributos x_i pueden ser el tiempo de viaje, el tiempo de caminata, el coste que queda después de tomar el servicio y bajarse en el paradero óptimo, etc.

Estos atributos son los mencionados anteriormente, los cuales definen un vector el cual pondera con un producto lineal los atributos con los coeficientes β_i . Esto genera una probabilidad de que la alternativa sea elegida, dada la fórmula:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

En el ámbito de predicción de demanda en transporte público, el modelo MNL se es conveniente pues permite incorporar múltiples factores que influyen la decisión final.

Ventajas

- Fácil de interpretar, pues se obtienen valores para cada característica de los costes, permitiendo interpretar que considera más valioso el usuario en términos de utilidad.
- Más rápida de entrenar, pues requiere menos datos.
- Resistente a outliers. Al ser una función global la obtenida, servicios que tengan desvíos o datos corruptos no se verán afectados por esto.
- Fácil de crear escenarios ficticios de nuevos servicios pues no es necesario reentrenar.

Desventajas

- No captura correlación espacial .
- Depende fuertemente de las variables propuestas.

Una evolución del MNL es el Logit Anidado, el cual flexibiliza el supuesto de la independencia de alternativas irrelevantes. Este modelo permite agrupar alternativas similares para correlacionar entre los términos de error.

Este modelo requiere que el investigador defina a priori la estructura jerárquica. Esta estructura jerárquica estática puede ser Modo Público/Privado > Tipo de Servicio (Metro, Bus, etc)

Además el logit anidado asume una estructura fija, el arbol de decisión es el mismo para todos los usuarios. Dado que el conjunto de alternativas es dinámico entre usuarios, definir una estructura de nidos universal que capture competencia entre miles de servicios resulta inviable.

Por otro lado, modelos de Machine Learning como Random Forest o XGBoost han demostrado una buena capacidad predictiva, sin embargo, solo operan sobre datos tabulares, perdiendo la topología y las relaciones entre nodos. Ademas, los modelos de clasificación estándar asumen un número fijo de clases. En el contexto de elección de ruta, el conjunto de alternativas tiene cardinalidad variable según el origen, el destino, e incluso la hora.

2.3.2 GNN

Una red neuronal de grafos (GNN) son redes neuronales especializadas para recibir como inputs grafos. A diferencia de redes neuronales como las LSTM o las convolucionales, las cuales reciben

datos con una estructura más rígida (una secuencia o una grilla), las GNN reciben datos en grafos abstractos. Entonces, se puede pensar a las GNN como una abstracción general de un set de datos relacionados entre sí. Wu et al. [17] hacen una revisión exhaustiva de las GNN y sus aplicaciones.

Una GNN aplicada al transporte público es una red neuronal capaz de aprender características espaciales. Se explora la solución de una GNN Heterogénea que aproveche la riqueza de los tipos del grafo bipartito entre los nodos y las aristas.

Una GNN tiene *embeddings* o representaciones vectoriales, tanto en los nodos paradero como en los nodos servicio, que permite añadir riqueza y similitud entre paraderos. Entre estos paraderos y servicios, se propagan mensajes que permiten capturar la correlación espacial. GraphConv, construido sobre el trabajo de Morris et al.[9], es una capa de convolución de grafos que permite capturar la información local de los nodos y sus vecinos. Existen otros tipos de capas, como SageConv construido por Morris Hamilton et al.[3], el cual aplanar los mensajes con promedios.

Ventajas

- Captura correlación espacial
- más riqueza local gracias a los *embeddings*

Desventajas

- Más tiempo de entrenamiento y más datos necesarios.
- Menos resistente a outliers (*embeddings* de recorridos corruptos o desviados afectan localmente)
- Menos interpretabilidad debido a lo abstracto que son los vectores de *embeddings*.
- Requiere reentrenar para escenarios ficticios de servicios que no existen (ya que hay que obtener los *embeddings*)

Capítulo 3

Metodología

En la presente sección, primero se dará a conocer la solución propuesta y finalmente la metodología para cumplir cada paso de la solución.

3.1 Solución propuesta

La solución se basa en dos pilares fundamentales, el primero, la *representación de los datos* en una estructura topológica, en donde se creó el grafo agrupado y bipartito usando el programa de operaciones de red y el segundo, el *modelado del proceso de decisión de los usuarios*, en donde se entrenaron dos modelos, el MNL y el GNN, usando datos históricos de demanda de ADATRAP. Esta solución cumple los objetivos específicos (OE1, OE2). La Figura 3.1 muestra un diagrama general de la solución propuesta.

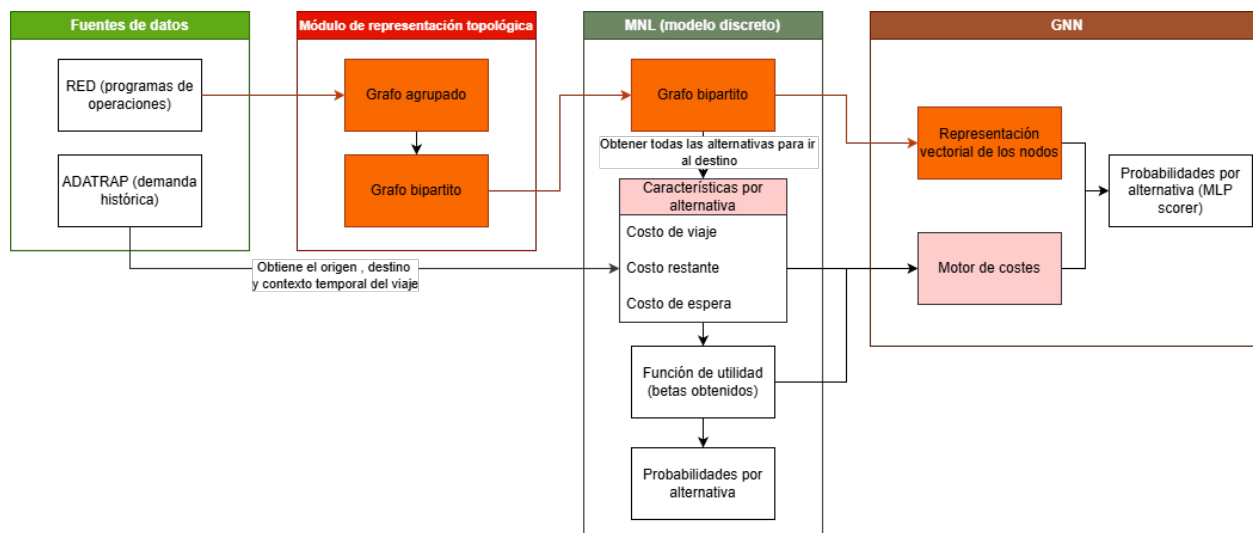


Figura 3.1: Diagrama general de la solución en donde se subdivide esta en 4 secciones principales, las fuentes de datos, la representación de estos, el modelo de elección discreto y el GNN

El objetivo del MNL es modelar el proceso de decisión de un usuario al elegir entre N alternativas de viaje en la primera etapa del viaje, entendiendo las alternativas como los servicios que paran en el paradero de origen del usuario a la hora y día que se sube. Cada alternativa tendrá un set de atributos, los cuales definen el coste total del viaje si es que el usuario elige esa alternativa. Este set de atributos o características de cada alternativa son:

- **Tiempo de espera** del servicio: Obtenido desde el programa de operaciones de RED, es el tiempo promedio que un usuario espera para abordar el servicio, asumiendo que el usuario puede llegar en cualquier momento al paradero. Se toma una distribución uniforme de llegada, por lo que el tiempo de espera promedio es la mitad del *headway* (frecuencia) del servicio en ese paradero.
- **Tiempo restante** de viaje desde el primer transbordo hasta el destino final: Obtenido usando un algoritmo de ruteo en el grafo bipartito. Este algoritmo debe encontrar el paradero óptimo en el que el usuario debe bajarse dado que quiera ir al destino dado.
- **Tiempo de viaje** en el servicio hasta el primer transbordo en el paradero óptimo.

Para los tres costes, es necesario tener una representación del programa de operaciones mas robusta y flexible que una tabla. Es por ello que se desarrolló un grafo bipartito. El grafo bipartito provee la información del tiempo de espera, tiempo de viaje y tiempo restante como los pesos de las aristas, en los que cada cambio de estado denota un coste que el usuario debe de pagar. Lógicamente el usuario no penaliza de igual manera pagar cinco minutos a bordo de un servicio que esperando el bus. Es por ello, que el MNL debe de ponderar estas características usando datos históricos, entrenando los coeficientes β_i de cada característica de la alternativa. Notar que un algoritmo enrutador que encuentre la ruta mas corta en el grafo bipartito entregará el camino con menor coste total de tiempo, pero no pondera la importancia de cada coste en la decisión del usuario. Es decir, **la ruta mas corta en tiempo no es la ruta mas corta en utilidad**.

Para entrenar al modelo, por cada intención de etapa de un viaje (es decir, un origen, destino, hora y día), se obtuvieron las N alternativas posibles en el paradero usando el grafo bipartito y se calcularon los atributos de cada una de ellas usando el algoritmo enrutador en el mismo grafo. Una de esas alternativas fue la elegida realmente (la demanda histórica), mientras que las otras no. Con ello, se entrenó al MNL, tomando como acierto del modelo elegir a la alternativa real tomada como la que tiene menor utilidad según la función lineal y se obtuvieron los coeficientes β_i que ponderan la importancia de cada atributo en la decisión del usuario. Esto es, un modelo de elección discreto.

La fase dos consistió en enriquecer el proceso de elección sumando la contextualización espacial, para ello, una GNN que utilice los *embeddings* de los paraderos y servicios fue implementada. Esta GNN aprovecha el motor de costes obtenido por el MNL para crear un set de atributos enriquecido para cada alternativa.

A continuación se detallan los pasos seguidos para cumplir cada fase de la solución propuesta.

3.2 Representación de los datos

Se usó el programa de operaciones de RED para crear los grafos necesarios para la *visualización* y el *entrenamiento* de los modelos. Estos son los grafos **agrupado** y **bipartito** respectivamente. El grafo agrupado tiene dos utilidades, primero, permite visualizar la red de transporte y segundo, permite ser una excelente base para desagrupar las aristas y crear el grafo bipartito.

3.2.1 Grafo Agrupado

Recordando que el grafo agrupado es un grafo dirigido donde los nodos son los paraderos y las aristas son las conexiones entre paraderos consecutivos en un servicio, se definieron los siguientes elementos:

Aristas

En este caso, las aristas E fueron las conexiones entre dos paraderos en un recorrido. Por ejemplo, una arista conecta la estación Los Héroes con Moneda. Una arista, por lo tanto, deben guardar los servicios que la recorren. Otro caso, son las aristas que unen paradas de servicios en superficie. Una arista va a representar la conexión entre dos paraderos consecutivos mediante un servicio.

Si varios servicios paran en las mismas paradas consecutivas, se unieron todos los recorridos en la misma arista. Es más simple computacionalmente, pero datos como la distancia o tiempo que toma al servicio recorrer la arista (el peso de la arista) se pierden.

Vértices

Los vértices V son las paradas. Cada parada tiene un par coordenado (latitud, longitud) que la posiciona en el grafo. Una parada se identifica con el código TS del paradero. Una parada contiene 1 o más servicios.

Algoritmo para crear el grafo agrupado

El primer paso, consistió en agrupar a todas las conexiones de dos paraderos consecutivos en una arista en común. Es decir:

1. El servicio X tiene una secuencia de paraderos P_k , con k el número de paradero en el recorrido. P_0 es el paradero inicial y P_N es el paradero final del recorrido.
2. Los paraderos se configuran en nodos V . Cada nodo V tiene como llave su código de usuario C , una lista de servicios $S[]$ y un par coordenado (lat, lon) para ubicarlo geográficamente.

3. Cada servicio tiene una secuencia de nodos que visita en orden. Por ejemplo, la secuencia de paraderos que visita un recorrido X es $P[\cdot]$. Si el set de nodos es $V[\cdot]$, se puede hacer una biyección entre P_k y V_i . Siendo k el k -ésimo paradero en orden e i el i -ésimo paradero de toda la red. Obviamente i no tiene por qué ser igual a k .
4. Si hay dos servicios, X e Y , que tienen secuencias de paraderos P_k y Q_i y tienen dos paraderos consecutivos que coinciden, es decir, $P_k = Q_i$ y $P_{k+1} = Q_{i+1}$, luego se puede decir que desde $P_k = Q_i = V_l$ a $P_{k+1} = Q_{i+1} = V_m$ habrá una arista en esa dirección, con m y l no necesariamente consecutivos.
5. Esta arista direccionada desde V_l a V_m tendrá como información que los servicios X e Y pasan por ella.

Siguiendo estas reglas, se creó el grafo con el siguiente pseudocódigo:

1. Se obtuvieron todos los servicios únicos en el dataframe del programa de operaciones de red.
2. Se creó un diccionario con la información Código Usuario, Variante (PM o Normal), Sentido Servicio (Ida o Regreso).
3. Por cada servicio, se filtraron del dataframe todas las filas que corresponden al servicio.
4. Se ordenó el dataframe viendo la columna “orden_circ”. Esta es la columna que denota el orden de circulación del servicio por los paraderos.
5. Por cada fila (paradero) del dataframe, se creó o actualizó un diccionario que corresponde al paradero, con llave código paradero, con los siguientes datos:
 - llave(código paradero)
 - lat
 - lon
 - servicios
 - nombre (Por ejemplo, José Joaquín Pérez esq Las Lomas)
 - nombre completo (código del paradero + nombre del paradero)
 - tipo (BUS o Metro)
6. Por cada fila del dataframe, se revisó el parámetro “siguiente_parada” que contiene la siguiente parada desde la que se está revisando (básicamente un puntero). Se creó una arista E_l en un diccionario que une ambas paradas con la siguiente información:
 - conexion_id (llave formada por el par código_paradero_origen, código_paradero_siguiente)
 - servicios
 - nodo_origen
 - nodo_destino
 - tipo (Bus o Metro)

Notar que al hacer esto por todos los servicios, se agregaron a cada arista los servicios que recorren ambos nodos en el mismo orden.

7. Se realizó el mismo procedimiento para el Metro, pero las aristas fueron bidireccionales (es decir, por cada conexión, se hizo una simétrica pero en sentido inverso).
8. Se unieron los nodos con las aristas.

Es inmediato notar que las aristas de este grafo están degeneradas, es decir, tienen información de múltiples servicios, pero no tienen un peso definido, esto hace que este grafo sea inviable para obtener las características de un servicio (el tiempo de viaje).

3.2.2 Grafo Bipartito

Antes de definir el grafo bipartito, es necesario definir la notación y los datos base que se usarán en la definición del grafo.

Notación y datos base

De ahora en adelante, se usará la siguiente notación para definir los elementos del grafo bipartito:

- **Paraderos:** P, Q, \dots
- **Servicios:** S (ej: 507), con **sentido** $d \in \{\text{Ida}, \text{Ret}\}$.
- **Tipo de día:** $D \in \{\text{LAB}, \text{SAB}, \text{DOM}\}$.
- **Tiempo discreto:** 48 bins de media hora $b \in \{0, \dots, 47\}$.
- **Frecuencia** (buses/h): $f_{S,d}(D, b)$.
- **Headway** (min entre buses): $H_{S,d}(D, b) = \frac{60}{f_{S,d}(D, b)}$.
- **Aristas VIAJAR:** tramo $(u \rightarrow v, S, d)$ con:
 - **distancia** L_e (m),
 - **velocidad** $v_e(D, b)$ (km/h),
 - **tiempo a bordo:**

$$\tau_e(D, b) = \frac{L_e/1000}{v_e(D, b)} \cdot 60 \quad [\text{min}]$$

Esto permite definir lo siguiente:

3.2.3 Frecuencias de los servicios

Para cada tupla servicio, sentido, variante es necesario definir la frecuencia de esta tupla para cada bin, es decir, la función:

$$f(S, V, d, b)$$

Esta función f retorna la frecuencia de la tupla (S, V, d) en el bin temporal b . Esta información la retorna el programa de operaciones de RED. RED provee de tablas de frecuencias para todos los servicios TS (es decir, los buses). Esta frecuencia es por hora. Es decir, buses/hora. Si para una tupla (S, V, d) la frecuencia es 6 buses/hora, esto significa que cada 10 minutos pasa un bus (si asumimos equipartición, que es lo que se asumirá desde ahora hasta que se diga lo contrario). El tiempo está dividido en bins de 30 minutos de ahora en adelante, por lo tanto, tendremos 48 bins en un día, desde el 0 hasta el 47.

No se encontraron datos de frecuencias del metro. Se decidió asumir una frecuencia fija de 3 minutos para todo el día durante su funcionamiento entre las 6:30 AM y las 11:00 PM. Esto equivale a una frecuencia de 20 vehículos/hora.

3.2.4 Velocidades promedio de los servicios

Para cada tupla servicio, sentido, variante es necesario definir la velocidad promedio de esta tupla para cada bin b , es decir, la función: $v_e(S, V, d, D, b)$

Esta función g retorna la velocidad promedio de la tupla (S, V, d) en el bin temporal b . Notar que al ser promedio, para una tupla, es la misma por todo el recorrido, es decir, no influye la arista por la que circula el servicio.

Para el metro, se decidió una velocidad promedio de 30km/h, debido a que no se encontraban en el programa de operaciones de RED. Esta estimación se basó en los tiempos que tarda el metro en ir de una estación a otra en Google Maps.

Ambas tablas (de frecuencias y velocidades) las provee RED en su plan de operaciones. Revisar acá: <https://www.dtpm.cl/index.php/programa-de-operacion>

3.2.5 Nodos

Los tipos de nodos que tuvo el grafo son:

Paraderos

Cada paradero es un nodo. Cada nodo tiene la siguiente información:

- Código de paradero (TS y Usuario)
- Latitud y longitud
- Nombre del paradero

- Tipo (BUS o Metro)
- Servicios que pasan por el paradero (lista) en cualquier bin b y día D .
- Zona 777

Servicios

Cada paradero tiene un conjunto de servicios que pasan por él. Por lo tanto, se define un nodo servicio por cada paradero y servicio que pasa por él. Estos nodos permiten cerrar la transición entre estar en un paradero y subirse a un servicio.

Por ejemplo, si un paradero P tiene los servicios 507 y 512 que pasan por él, se crean dos nodos servicio: Servicio 507 en paradero P y Servicio 512 en paradero P . Cada nodo servicio tiene la siguiente información:

- Nodo paradero asociado (Paradero P)
- Servicio
- Sentido
- Variante (PM o Normal)
- Tipo (BUS o Metro)

3.2.6 Aristas

Los tipos de aristas que tuvo el grafo son:

VIAJAR

Aristas que corren entre nodos *Servicio*. Representan la conexión dirigida entre dos paradas consecutivas de un servicio. Estas aristas tienen la siguiente información:

- Nodo origen (Servicio en paradero P)
- Nodo destino (Servicio en paradero Q)
- Servicio
- Sentido
- Variante (PM o Normal)
- Distancia (metros)
- Velocidad promedio (kilómetros/hora) por bin y tipo de día (en un diccionario)
- Tiempo a bordo (min) por bin y tipo de día (en un diccionario)
- Tipo (BUS o Metro)

Notar que las aristas VIAJAR tienen peso, el cual es el tiempo a bordo. Estas aristas son temporalmente dependientes.

CAMINAR

Aristas que corren entre nodos *Paradero*. Representan la conexión no dirigida entre dos paraderos cercanos. Estas aristas tienen la siguiente información:

- Nodo origen (Paradero P)
- Nodo destino (Paradero Q)
- Distancia (metros)
- Tiempo estimado (min). Son iguales para ambas direcciones y son atemporales. Dependen de la velocidad promedio del usuario.

SUBIR

Aristas que corren entre nodos *Paradero* y *Servicio*. Representan la acción de subirse a un servicio en un paradero. Estas aristas tienen la siguiente información:

- Nodo origen (Paradero P)
- Nodo destino (Servicio en paradero P)
- Servicio
- Sentido
- Variante (PM o Normal)
- Tiempo de espera (min) por bin y tipo de día (en un diccionario)
- Tipo (BUS o Metro)

BAJAR

Aristas que corren entre nodos *Servicio* y *Paradero*. Representan la acción de bajarse de un servicio en un paradero. Estas aristas tienen la siguiente información:

- Nodo origen (Servicio en paradero P)
- Nodo destino (Paradero P)
- Servicio
- Sentido
- Variante (PM o Normal)
- Tipo (BUS o Metro)

Estas aristas no tienen coste alguno. Bajarse es inmediato.

Todas las distancias son proyectando líneas rectas entre ambos nodos. No tiene en cuenta la topología de la urbe.

3.2.7 Algoritmo para la creación del grafo bipartito

Hacer el grafo de estado es directo teniendo el grafo agrupado. Un algoritmo recorre cada nodo de tipo paradero y crea :

1. Aristas de tipo CAMINAR entre los k nodos PARADERO más cercanos. Se definió 10 paraderos a menos de 200 metros. Si un paradero no tiene ningún paradero a menos de 200 metros no tendrá aristas CAMINAR entrantes ni salientes.

2. Nodos de tipo servicio por cada servicio que pase.
3. Aristas de tipo SUBIR a cada nodo SERVICIO con los pesos en un diccionario.
4. Aristas de tipo BAJAR desde cada nodo SERVICIO al PARADERO.

Luego, se conectaron todos los nodos de tipo servicio con aristas VIAJAR según el recorrido .

Para el grafo bipartito solo se tuvo en cuenta el Metro y los servicios de superficie como los buses. No se consideró el Metro Tren Nos pues no posee plan de operaciones en la página de RED.

3.3 MNL

El MNL se entrenó para predecir la probabilidad de que un usuario elija una alternativa entre N alternativas en un paradero de origen O , dado un destino D , una hora y un día. Para ello utiliza tres atributos principales, el tiempo de espera, el tiempo de viaje en el servicio y el tiempo restante desde el primer transbordo hasta el destino final.

Suponer que para ir a un destino D desde un origen O tienen dos opciones. Un servicio S_1 que deja directamente en el destino, con un coste de viaje asociado Cv_1 y un servicio S_2 que tiene un coste de viaje Cv_2 hasta el primer transbordo, para luego tener un costo de viaje de ese servicio de transbordo Cr_2 .

Si es que el tiempo de viaje de S_1 es menor y además deja directamente en su destino, es lógico que tomar este servicio es la decisión idónea u óptima. Ahora, si el costo de viaje de S_1 es mucho más alto, quizás convenga tomar un transbordo. Un ejemplo clásico de esto sería hacer transbordo al metro usando un bus alimentador para llegar al sistema subterráneo. A priori, dependiendo de la urgencia del usuario, deberá de elegir una de las dos alternativas. No todos los usuarios piensan igual. Algunos prefieren comodidad y no hacer transbordos, sobre todo si están con algo de tiempo de sobra. Otras personas confían más en servicios más rápidos que les obligan a hacer transbordo. Como no todo el mundo piensa igual, el MNL es muy útil para estos casos, ya que entrega una distribución de probabilidad sobre que servicio se va a tomar, sobre todo cuando las utilidades de ambos son parecidas. El objetivo de este modelo es descubrir que prefieren los usuarios, si viajes más directos con menos transbordos -pero más largos- , o viajes más rápidos pero con transbordos. Notar que los transbordos tienen tiempos de viajes más variables. Poca confianza en los headways de los buses de transbordo pueden inflar el tiempo de viaje real, ya que la variable de tiempo de espera suele tener más varianza que el tiempo de viaje. Más transbordos implican más varianza en el tiempo de viaje total y por lo tanto menos confianza en el trayecto, o sea, menos comodidad.

Con esta reflexión, es directo darse cuenta que lo que se busca con este modelo es descubrir como se comparan el tiempo de viaje total v/s qué tanto acerca el servicio inicial al destino.

3.3.1 Métricas de entrenamiento

Esta sección aplicó tanto para la MNL como la próxima GNN. Las métricas de entrenamiento fueron:

- NLL (Loss): Indica que tan bien calibradas las probabilidades. Penaliza fuertemente la sobreconfianza cuando se falla. *Más bajo es mejor.*
- NLL Normalizado: Normaliza NLL dependiendo del tamaño del set de alternativas. *0 es perfecto, 1.0 es uniforme.*
- Acc (Accuracy TOP1): Indica la proporción de predicciones que acertaron en el primer rank de las probabilidades, es decir, si la elegida realmente fue la más alta probabilidad. *Más alto es mejor.*
- AccNT (Accuracy Non Trivial): Precisión de decisiones no triviales (es decir, con set de alternativas mayor a 1). *Más alto es mejor.* Esta métrica es más importante que Acc, debido a la gran proporción de decisiones con solo una posibilidad.
- MRR (Mean Reciprocal Rank): Valora que la elegida esté alta en el ranking a pesar de que no esté top 1. *Más alto es mejor.*
- LL (Log Likelihood): Se usa solo en el MNL. Suma de $\log(p|elegida)$. *Cuanto más negativo y cercano a cero mejor.*
- LL_null: LL del modelo uniforme. Para medir ganancia sobre el azar.
- Pseudo- R^2 de McFadden: $1 - LL_modelo/LL_null$. Análogo a R^2 ; 0 a <1 (mayor es mejor). En elección discreta, ~0.2–0.4 suele considerarse muy bueno.

3.3.2 Dataset

Para el modelo MNL, se utilizó la tabla de etapas, ya que venía disgregada y permitía manejar con mayor facilidad los datos que la tabla de viajes, la cual tenía que expandirse.

Los pasos para generar el *dataset* fueron los siguientes:

1. De la tabla de etapas, se obtuvieron todos los viajes que tienen bajada registrada gracias a ADATRAP.
2. Por cada fila se obtuvo el paradero de origen, el servicio tomado, el bin, el paradero de bajada observado, el tipo de día y el tipo de servicio. Se agregó el destino final para cada etapa gracias a agrupar las etapas con el mismo ID.

Costo Restante

El costo restante es la medida en tiempo que el usuario le queda por pagar al bajarse en el paradero óptimo y los transbordos que le preceden. Tomar el siguiente ejemplo. Una persona que quiere ir desde PJ394 a PA433 (Beauchef) a las 10 de la mañana un día laboral.

- En el paradero PJ394 se tienen las siguientes alternativas a las diez de la mañana un día LABORAL: 503, 504, 507, 517, 518, B38.
- Convenientemente también para el 507 en el paradero destino, así que el costo restante es cero para esa alternativa, pues después de bajarse en la parada óptima, ya se llegó al destino.
- Para los otros servicios, el costo restante es mayor que cero, ya que ninguno deja directamente en PA433. Entonces, se debe calcular el costo restante desde el paradero de bajada óptimo.

Dijkstra Inverso para el coste restante y el paradero óptimo

Para calcular el paradero óptimo y el costo restante al bajarse en ese paradero es importante la noción del Algoritmo de Dijkstra (AD).

- Se parte en un nodo origen y se le asigna un costo 0.
- Se exploran todos los nodos vecinos y se les asigna un costo igual al peso de la arista que los conecta con el nodo origen.
- Se marca el nodo origen como visitado.
- Se selecciona el nodo no visitado con el costo más bajo y se repite el proceso hasta que todos los nodos hayan sido visitados o se haya alcanzado el nodo destino.

No se usó A^* para esta tarea, pues se necesita ejecutar este algoritmo de ruteo para toda la red en una pasada desde un destino. A^* es perfecto para cálculos punto a punto, pero si se hubiera usado este algoritmo, se hubiera tenido que ejecutar A^* N veces, la complejidad del código completo hubiera sido enorme.

Volviendo a Dijkstra, en este caso, el algoritmo se corre en sentido inverso, es decir, se parte del nodo destino y se avanza hacia atrás. De esta forma, se obtiene el costo mínimo para llegar al destino desde cualquier otro nodo. Es esta razón por la que se usó Dijkstra y no A^* , ya que se quiere una tabla completa de costos restantes para todos los paraderos.

Entonces, para un paradero de destino, un bin y un día se obtiene una lista de costos restantes para cada paradero de la red. Notar que es costoso ejecutar este algoritmo en un grafo tan grande, así que hay que ejecutar estrategias para evitar el sobrecoste.

Por ejemplo, se puede ejecutar el AD para PA433 y el costo restante para ir desde cada paradero hasta PA433 es el costo de viajar. Como se separaron los servicios (es decir, el grafo no es agrupado), cada camino es una combinación de aristas. Lo bueno de este enfoque, es que penaliza fuertemente los transbordos, haciendo más realistas los caminos.

Entonces, se obtiene un camino C que tiene de extremos dos nodos PARADERO y una cantidad par de aristas SUBIR + BAJAR y un número arbitrario de aristas VIAJAR visitadas. Dicho de otra manera, el costo restante tiene el costo de los transbordos, esperas, tiempo a bordo y todo lo incluido. El coste restante representa la parte determinista del viaje después de la primera decisión.

Si se ejecuta el AD para el origen y destino, se obtendrán paraderos de bajada óptimos para cada alternativa. Estos son, el destino para el 507, y paraderos de transbordo para los otros.

Volviendo al proceso de la creación del dataset:

3. Se expandió esta tabla de decisiones con las alternativas posibles disponibles para el usuario en ese paradero y bin. Esta expansión se hizo de la siguiente forma:

- Por cada decisión, se obtienen todos los servicios disponibles.

- Por cada alternativa, incluida la elegida, se extraen desde el grafo de estado el tiempo de espera en ese paradero y bin.
- Por cada alternativa que NO sea la decisión, se ejecuta el AD Inverso desde el paradero de destino FINAL . El paradero óptimo es el que minimiza el costo restante dentro del perfil de paraderos del servicio, siendo el perfil de paraderos una lista de paraderos que le siguen al paradero actual .
- Por cada alternativa, se calcula el costo de esperar el servicio más el costo de viajar al paradero óptimo (o el de bajada real en el caso del usado) y el costo restante ya obtenido por Dijkstra.
- Se agregan todos estos atributos a la fila por cada alternativa.

Algunas optimizaciones para acelerar el proceso fueron:

- Se agruparon todos los viajes que iban al mismo destino. Con ello, se calculaba solo una vez el algoritmo de Dijkstra para muchas decisiones a la vez, y estos resultados se cacheaban.
- Se *cacheó* el perfil de cada servicio.

Con ello, una tabla de etapas de un día (300K decisiones) se podía procesar en 2 horas solamente.

La tabla de decisiones resultante tiene las siguientes columnas:

Columna	Descripción
decision_id	ID de la decisión. Común entre alternativas de la misma persona
person_id	ID de la tarjeta
trip_id_real	ID de viaje real: [person_id]_[número_de_viaje]_[número_de_etapa]
servicio_usuario_alternativa	Servicio alternativa en código formato usuario
sentido	Sentido del servicio
wait_time	Tiempo de espera
optimal_alight_stop	Parada óptima de bajada del servicio alternativa. Es la bajada real observada cuando el servicio fue tomado.
cost_to_go	Coste restante entre la parada de bajada y el destino final
viajar_cost	Coste entre el nodo servicio inicial y final del servicio inicial tomado
total_cost	Coste total sumando todos los costes (wait, cost_to_go, viajar_cost)
chosen	Si la alternativa fue elegida o no (solo una fila con 1 por decision_id)
choice_set_size	Tamaño del set de alternativas
day_type	Día (LAB, SAB, DOM)
bin30_k	Bin de tiempo en el que se llegó al paradero
origin_p_k	Paradero de origen en código TS
dest_p_final_obs	Paradero de destino final real en código TS
srv_obs_ts	Servicio real tomado en código TS
srv_obs_usuario	Servicio real tomado en código formato usuario
is_corrupt	Si la fila está corrompida
is_initial_transfer	1 si Dijkstra decide caminar a otro paradero para tomar un servicio adecuado, 0 en caso contrario

Tabla 3.1: Descripción de las columnas del dataset de decisiones para el modelo MNL.

Notar el último atributo, que indica si el usuario tuvo que caminar a otro paradero para tomar un servicio adecuado. Esto es relevante, pues implica un costo extra de caminata. Esto es la causa de muchos problemas que se verán mas adelante.

Nota sobre el coste restante

El coste restante tiene una característica **determinista**, pues asume que después de la primera decisión las personas son deterministas y no eligen con distribución de probabilidad. Esto es intencional. Modelar todo el trayecto como una concatenación de decisiones probabilísticas complica el modelo. Una pequeña intuición que no se desarrollará en este trabajo indica que posiblemente el coste restante sería una distribución o variable aleatoria más que un valor fijo escalar. Esto tiene más sentido real. Una persona sabe que hacer transbordo aumenta su varianza en su tiempo de viaje debido a que debe de esperar otro servicio, que induce una incerteza temporal. Aunque en el coste restante está incluido el tiempo de espera, realmente el tiempo de espera *esperado* debería de ser un tiempo que tenga en cuenta todos los tiempos de espera del paradero que le puedan servir al usuario. Lo mismo con los tiempos de viaje del servicio que pueda tomar el usuario. Es inmediato notar como se complica el problema, pues ahora cada decisión subsecuente tiene una distribución.

3.3.3 Entrenamiento

Algoritmo de entrenamiento

Para el entrenamiento, se consideró lo siguiente:

- Variables base: cost_to_go, wait_time, viajar_cost.
 - Variables derivadas : is_initial_transfer (binaria) y first_walk_time= penalty (5 mins) si is_initial_transfer es True.
 - zero_onboard = 1 si viajar_cost es 0.
 - ASC_METRO si usa el metro.
 - intercepto. Constante de probabilidad.
-
1. Se dividió un 20% del dataset para pruebas (test).
 2. El modelo tiene la utilidad ya mencionada anteriormente $u = X\beta$. La probabilidad por alternativa es un softmax estable por decisión.
 3. Tiene una regularización L2 sobre β .
 4. Optimización. Minimiza NLL con L-BFGS-B.
 5. Gradiente Analítico.
 6. 300 iteraciones máximas por época. tol = 1e-7 y l2_reg= 1e-3.
 7. Métricas las ya anteriormente mencionadas.

La Figura 3.2 resume el proceso de entrenamiento del MNL.

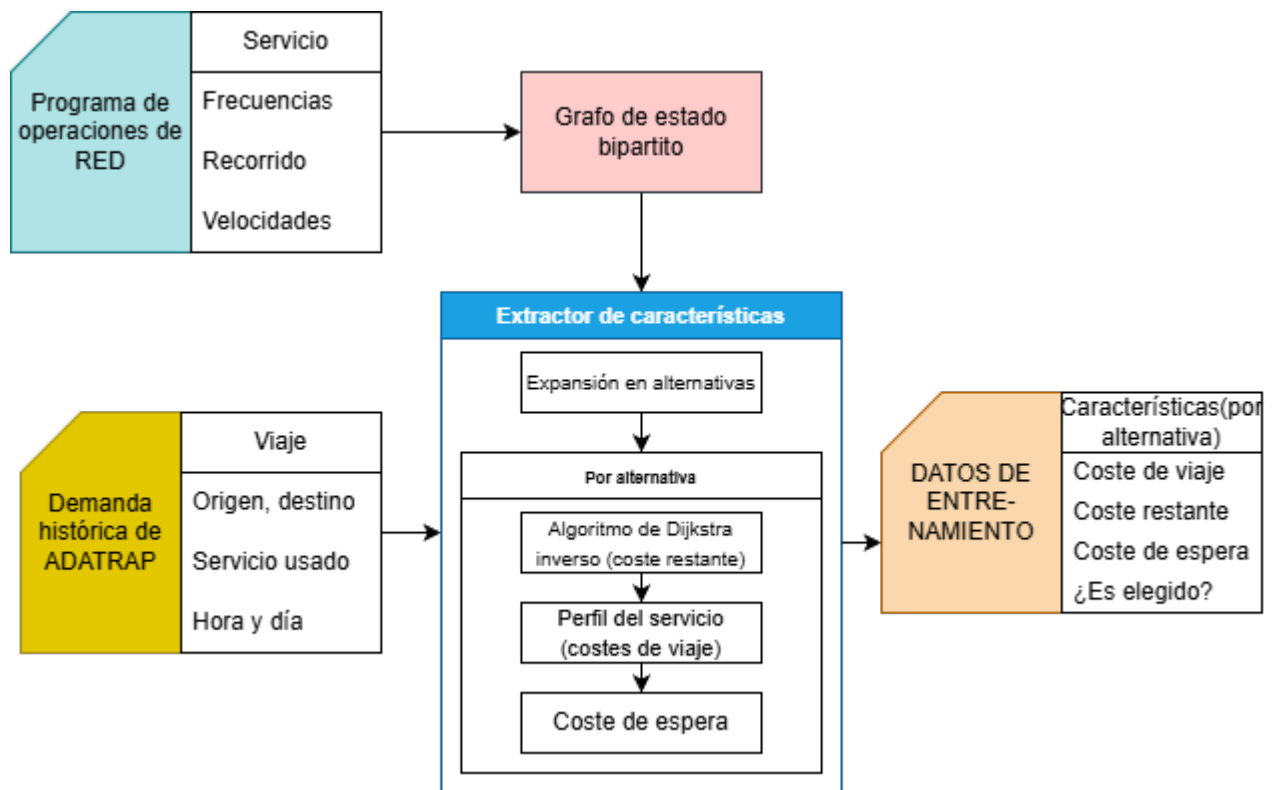


Figura 3.2: Diagrama del entrenamiento del MNL

3.4 GNN

3.4.1 Arquitectura de la solución

Se implementó una GNN con la siguiente arquitectura mostrada en la Figura 3.3.

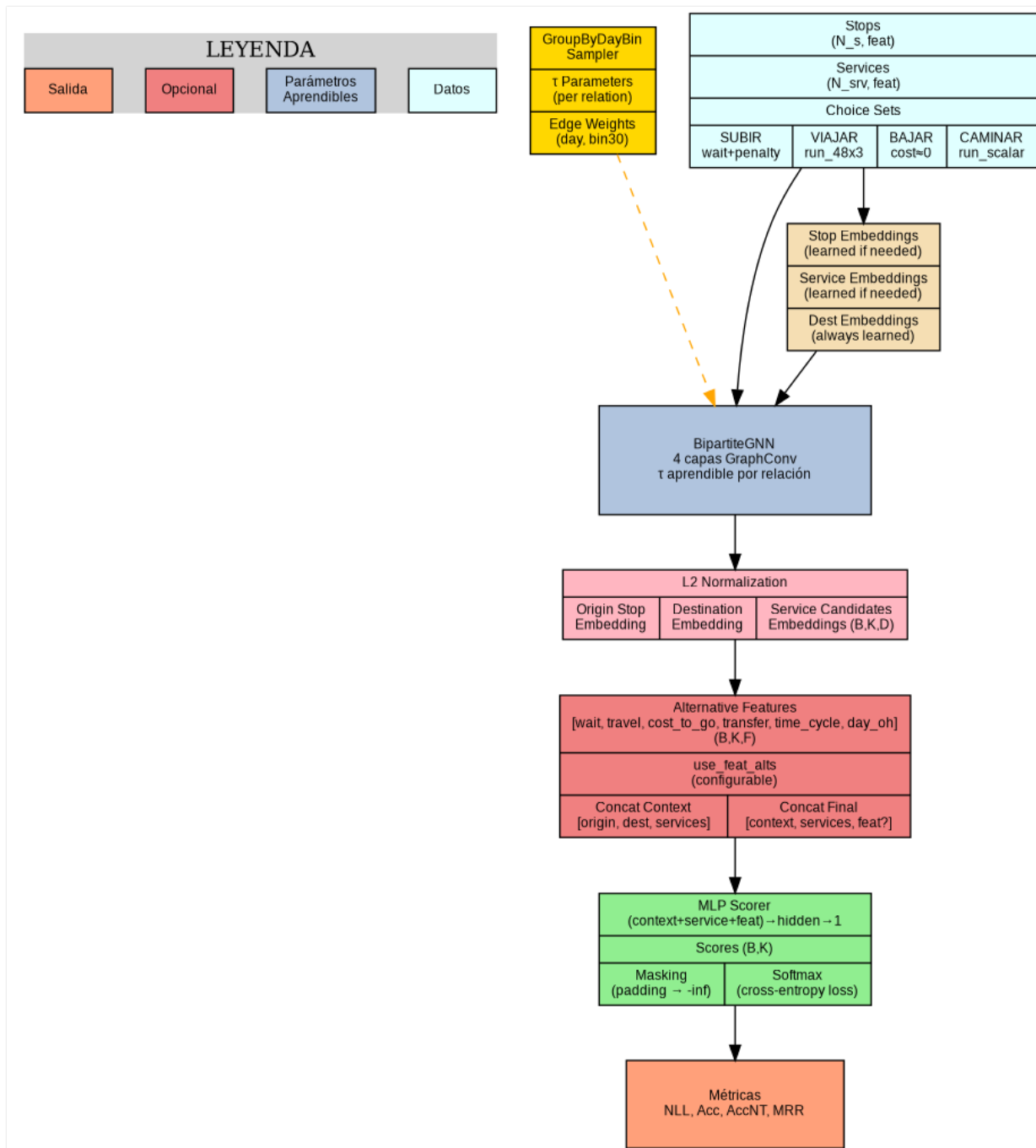


Figura 3.3: Arquitectura de la GNN

3.4.2 Datos

Para los datos, se usó el grafo Bipartito ya mencionado anteriormente. Además, se utilizó la misma tabla de decisiones para entrenar al MNL para poder compararlos justamente. Se añadió la variante a la tabla de decisiones infiriéndola desde el bin, día y paradero en que tomó el servicio el usuario. Esto para homogeneizar los datos con respecto al grafo bipartito.

Las aristas SUBIR, VIAJAR, BAJAR Y CAMINAR tienen tensores relacionados con los costes de transicionar de estado en el grafo bipartito.

- SUBIR: “wait_48x3” un tensor [servicio, tipo dia, bin30].
- VIAJAR: “run_48x3” un tensor [arista, day_type, bin30].
- BAJAR: “cost” un tensor de ceros (no penalizar bajar del servicio).
- CAMINAR: “run_scalar” un tensor de rango 0 (un escalar) que denota el tiempo de caminata calculando distancia euclidiana dividido por la velocidad.

3.4.3 Embeddings iniciales

Los *embeddings* son atributos vectoriales aprendibles por la red locales a cada nodo. Hay *embeddings* para :

- Paraderos
- Servicios
- Destinos

Una representación vectorial de un nodo es útil pues permite reducir la dimensionalidad, establecer similitudes y aprender localmente características de los nodos. Notar que la MNL no tiene esta característica espacial. Esto hace a la GNN más completa. Se podrían hacer *embeddings* para las aristas, pero esto hace al modelo menos interpretable.

3.4.4 GNN Bipartito

Una GNN Bipartito tiene 4 capas de GraphConv (en un inicio se usó SageConv, pero SageConv daba resultados poco convincentes debido a que aplanaba los mensajes de los vecinos). GraphConv aplica una convolución que permite que cada nodo agregue información de sus vecinos, el *message passing*.

Se agrega un parámetro τ_e que es aprendible por relación. Esto para todas las aristas. Este τ_e modula la intensidad del paso de mensajes, lo que implica aristas más importantes que otras (en otras palabras, unos costes pueden ponderar más, básicamente lo que hace la MNL)

3.4.5 Normalización y contexto

Se aplica una Normalización L2 para que algunos *embeddings* dominen por magnitud, mejore la convergencia, facilita las comparaciones entre *embeddings* y previene el overfitting.

3.4.6 Features y concatenación

Las *features* del MNL (o las características de las alternativas) son añadidas opcionalmente como una capa extra. Estos *features* se concatenan con todas las características ya aprendidas mediante los *embeddings*. El vector final entonces tendrá dimensión 64, 64, 64 + DIMENSIONES de características del MNL (*FEATURES*).

3.4.7 MLP (Perceptron Multi Capa) Scorer

Toma el vector concatenado y retorna una probabilidad para cada alternativa del modelo, implementando la lógica de decisión del modelo, muy parecido al MNL. El scorer es dinámico, por lo que respeta las dimensiones del vector concatenado. El MLP tiene las siguientes fases:

- Capa de entrada (nn.Sequential)
- Scores de Utilidad: Retorna los scores de cada servicio.
- Masking de alternativas : Tamaño variable demanda padding.
- Softmax: Transforma Scores a probabilidades de la misma manera que el MNL.
- CrossEntropyLoss: Penaliza predicciones incorrectas.

Las métricas de evaluación serán las mismas que las del MNL para poder compararlos efectivamente.

3.5 Experimentos

Para exponer los modelos a distintos cambios topológicos, y así cumplir el objetivo específico OE3, se realizaron los siguientes experimentos:

3.5.1 Disminución de la oferta de un servicio

Se disminuyó la oferta de un servicio (el servicio 517 y el 507) modificando los tiempos de espera en el grafo. Con ello, se analizó la redistribución de la probabilidad comparando las probabilidades *baseline* (basales) y las *contrafactuales* (el nuevo grafo). En específico, se analizaron dos casos, uno en el que dos servicios compiten (tienen probabilidades comparables, en este caso, el 503 y el 517) y uno de ellos verá su tiempo de espera modificado (el 517), mientras que el otro caso se enfrenta un servicio dominante (probabilidad muy alta, el 507) versus el resto (503, 504, 517, 518, B38 en el paradero PJ394), analizando la redistribución de la demanda en los transbordos. Esto se hizo de manera local en PJ394 y no en todo el grafo.

3.5.2 Suspensión de un servicio

Se suspendió la L1, colocando un indicador booleano en sus aristas para que Dijkstra no permita subir al servicio, y se ejecutó el modelo de predicción. Se realizó la misma comparación mencionada en el caso anterior. Esto se hizo de manera local, es decir, no se ejecutó el algoritmo de predicción sobre todo el día, ya que tardará demasiado.

3.5.3 Agregar línea 7

Los datos de esta nueva línea 7 siguen en la tabla 3.2. El trazado se obtuvo desde Metro y las coordenadas se aproximaron viendo Google Maps.

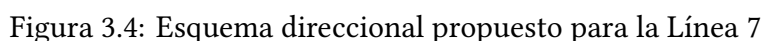


Tabla 3.2: Estaciones y coordenadas. El nombre de las estaciones no es el oficial. Están en orden partiendo desde Brasil hasta Estoril. Entre paréntesis aparecen las combinaciones. "esq" abrevia "esquina"

Posteriormente, se ejecutó el predictor con todos los datos de etapas de un día, específicamente

el lunes 24 de Abril de este año (2025). Para ello, se realizó el siguiente algoritmo:

- Eliminar las etapas intermedias, es decir, solo quedan las intenciones de viaje (paradero inicial, final, bin30, día)
- Por cada intención de viaje, ejecutar el predictor, y tomar el camino con mayor probabilidad de ser elegido . El motor de rutas ahora es el algoritmo A^* , pues es mas eficiente para rutas punto a punto que Dijkstra.
- Expandir el camino en etapas. En el caso del Metro, ADATRAP provee las tablas con viajes en metro sin contar las combinaciones, es decir, un usuario que se suba en San Pablo (L1 o L5) y se baje en Vicuña Mackenna (L4) solo mostrará un viaje en ADATRAP, no dos o tres. Por lo tanto, se expandió el camino en etapas separando estas etapas colapsadas en varios subcaminos del camino global. El camino se obtuvo con un algoritmo de ruteo teniendo en cuenta las velocidades y tiempos de espera del metro.
- Expandir la tabla de etapas original, para poder comparar efectivamente ambas tablas de etapas.

Con esta tabla de demandas sintética, se comparará ésta última con la tabla de etapas original en los siguientes aspectos:

- Servicios alimentadores a la L7. Se entiende como alimentador a servicios utilizados en la etapa $k-1$ del viaje, siendo la etapa k la etapa en que se usó el servicio L7.
- Servicios que ganan o pierden demanda. Se obtuvo comparando ambas tablas de etapas.
- Cantidad de etapas promedio obtenida. Se obtuvo contando las etapas de servicios clave, en ese caso, el metro.
- Carga del servicio. Se comparan la cantidad de veces que se recorren las aristas de un servicio, esto para denotar la carga que los usuarios imprimen en el servicio.

Capítulo 4

Resultados y discusiones

En la siguiente sección se mostrarán resultados e inmediatamente una discusión de ellos para mayor comodidad. Esta parte está estructurada de la misma manera que la metodología para relacionarlas directamente de manera más sencilla.

4.1 Representación de los datos

4.1.1 Grafo agrupado

La creación del grafo agrupado dio como resultados un grafo con las siguientes características:

- 11890 paraderos de bus
- 126 estaciones de metro
- 15465 conexiones de bus
- 272 conexiones de metro
- 15737 conexiones totales

Para visualizar el grafo, se puede descargar el archivo en el link, haciendo *clic* en el ícono de descarga cerca de “Raw” llamado “download raw file” y abrirlo en un navegador. La Figura 4.1 muestra un zoom a un barrio de Cerro Navia en este mismo mapa.

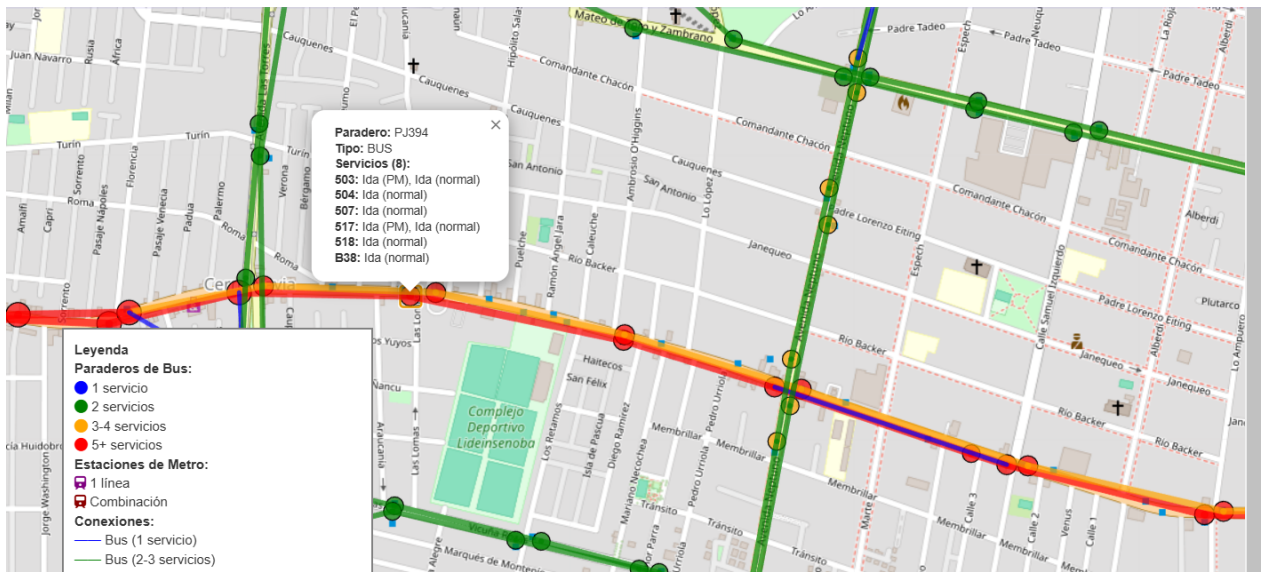


Figura 4.1: Mapa en Plotly con zoom a un barrio de Cerro Navia

Este mapa permite visualizar el grafo completo, pero carece de funcionalidad para el entrenamiento de los modelos.

4.1.2 Grafo bipartito

En la Figura 2.2 se muestra un ejemplo del grafo de estado en la misma zona que se mostró en la Figura 4.1.

Si se audita el grafo para sanear errores, se obtiene lo siguiente:

=== Estado del grafo bipartito ===

Nodos totales : 60679

- Paraderos : 11588
- A bordo (Servicio) : 49091

Aristas totales : 217305

Aristas por tipo :

- CAMINAR : 70826
- SUBIR : 49091
- VIAJAR : 48297
- BAJAR : 49091

Se obtiene un grafo muy útil. Por ejemplo, ya con este grafo con pesos se puede ejecutar un algoritmo de Dijkstra para encontrar la ruta más corta entre dos paraderos. Notar que esta ruta

más corta es teniendo en cuenta que todos los pesos “pesan” lo mismo, es decir, al enrutador le es indiferente recorrer 15 minutos caminando, que en bus o metro, ni que un minuto de espera vale lo mismo que un minuto a bordo. Esto es lo que se tiene que descubrir viendo los parámetros, en este caso, del MNL.

Este grafo tiene toda la información de la OFERTA de transporte. Junto con las tablas de etapas y viajes se tiene la DEMANDA.

Recordar que el objetivo es tener un grafo de estado artificial con OFERTA ARTIFICIAL y obtener, en base a tablas de etapas y viajes reales, DEMANDA ARTIFICIAL al tener modelos de elección y de grafos que las generen .

4.2 MNL

Recordar que las variables base del modelo son:

- Tiempo de espera
- Tiempo de viaje
- Coste restante

4.2.1 Resultados y coeficientes

El entrenamiento dio los siguientes resultados.

	intercept	wait_time	viajar_cost	cost_to_go	zero_onboard	ASC_metro
Coefficiente	0.0	0.0	0.0	0.0	0.0	0.0

Tabla 4.1: Coeficientes obtenidos del modelo MNL (todos nulos).

	loglik	loglik_null	pseudo-R^2 (McFadden)	Top-1 Accuracy	# Decisiones	# Alternativas
Entrenamiento	NaN	-178245.70	NaN	0.435	151279	701136
Validación	NaN	-44693.39	NaN	0.433	37791	175648

Tabla 4.2: Métricas de entrenamiento y validación del modelo MNL.

Como se observa en las tablas, todos los coeficientes resultaron nulos y las métricas de log-likelihood y pseudo- R^2 no son numéricas (NaN), lo que indica que el modelo no logró aprender una relación significativa entre las variables y la elección observada. Sin embargo, la *top-1 accuracy* se mantiene en torno al 43%, lo que sugiere que, pese a la falta de ajuste en los coeficientes, el modelo logra predecir la alternativa elegida en una proporción considerable de los casos, posiblemente debido a la estructura de los datos o a la presencia de alternativas dominantes.

La Figura 4.2 ilustra la razón de este problema. Hay muchas alternativas con *viajar_cost* cero.

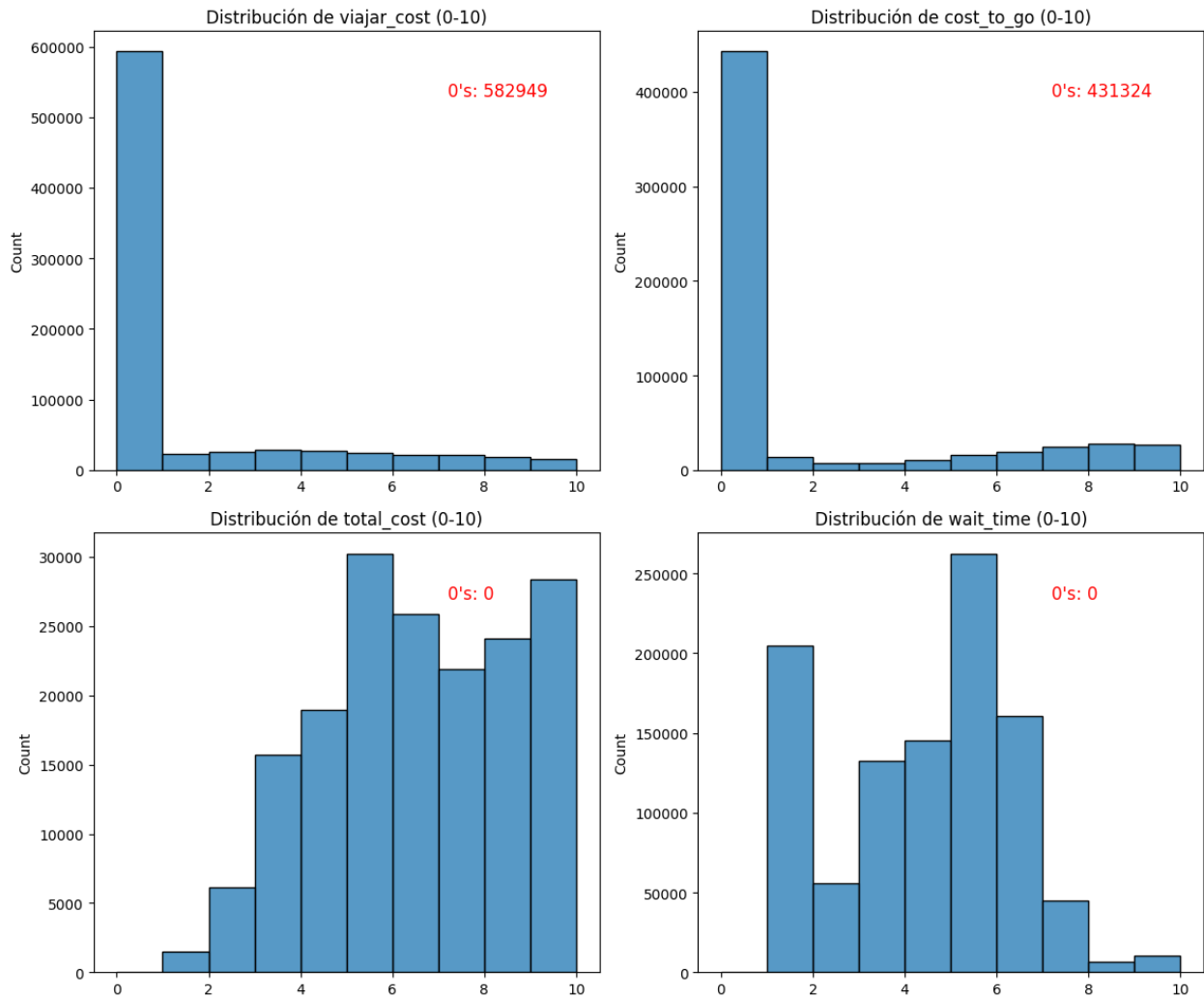


Figura 4.2: Histograma de Pesos. El eje x representa el tiempo y el eje y la frecuencia

La cantidad de 0's en *viajar_cost* y en *cost_to_go* tienen distintas razones.

Para *viajar_cost*, el coste de viajar es 0 cuando el usuario llega a un paradero, y una de las alternativas decide en hacer transbordo a otro paradero, ya que la ruta más corta comienza en ese paradero. Esto es un problema, pues hace que el modelo no converja.

Para *cost_to_go* 0, es cuando es necesario solo una etapa para completar el viaje. Esto no es problema que sea 0.

4.2.2 Primeros transbordos y la penalización inicial

Una solución elegida fue penalizar a las alternativas que requieran hacer transbordo inicial (es decir, cambiarse de paradero sin siquiera tomar el primer servicio), con un tiempo extra que se pueda configurar, esto por dos razones:

- Es más simple y directo.

- No es necesario que Dijkstra retorne el camino completo para verificar hacia donde fue.

Una idea interesante podría haber sido guardar el camino completo hecho por Dijkstra (y no solo el peso) y cuando hayan primeros transbordos, guardar el coste de caminar al paradero óptimo, pero esto es más complicado, consume más memoria, y además era necesario de ya entrenar para avanzar con la memoria. Esta idea se descartó.

Se decide con aplicar una penalización de 5 minutos al tiempo inicial para evitar valores nulos. Esta es una heurística razonable.

4.2.3 Entrenamiento diario

Primero, para agilizar el entrenamiento, se analizó día por día. El día miércoles no estaba disponible en los datos de RED. Entrenar semanalmente permite identificar cambios en los parámetros dependiendo del día. Ver la tabla 4.3 a modo resumen de los parámetros obtenidos.

En esta sección no se mostrarán las métricas diarias, pues consumirían mucho espacio y no son atingentes, pero en la sección siguiente, se mostrarán las métricas de desempeño semanal. La precisión promedio trivial (es decir, contando a los sets de alternativas con cardinalidad uno) de todos los días fue del 91% tanto en el split de validación como en el de entrenamiento. La precisión no trivial fue del 89%.

Día	intercept	wait_time	viajar_cost	cost_to_go	first_walk_min	is_initial_transfer	zero_onboard	ASC_metro
lunes	5.81×10^{-13}	-0.75	0.06	-5.41	-0.10	-0.19	2.02	1.70×10^{-13}
martes	5.79×10^{-13}	-0.99	0.03	-5.63	-0.07	-0.15	2.13	1.75×10^{-13}
jueves	4.95×10^{-13}	-0.83	-0.01	-5.59	-0.07	-0.15	2.07	1.50×10^{-13}
viernes	5.94×10^{-13}	-0.96	-0.04	-5.64	-0.10	-0.20	2.13	1.79×10^{-13}
sábado	7.64×10^{-13}	-1.09	-0.16	-5.36	-0.09	-0.19	2.28	2.38×10^{-13}
domingo	7.59×10^{-13}	-0.50	-0.23	-5.11	-0.08	-0.15	2.28	2.16×10^{-13}
Promedio	6.29×10^{-13}	-0.85	-0.06	-5.46	-0.09	-0.17	2.15	1.88×10^{-13}

Tabla 4.3: Coeficientes del modelo MNL entrenado para distintos días de la semana. La última fila muestra el promedio de cada columna. Este entrenamiento fue por cada día.

4.2.4 Entrenamiento semanal completo

Para tener una vista general, se opta por hacer un entrenamiento general de la semana completa, tomando un split del 20% de los datos totales. La tabla 4.4 muestra la cantidad de decisiones. Con ello, se obtienen las siguientes métricas (4.3) y parámetros (4.5).

Split de Datos	Cantidad
Datos cargados (tuplas)	(5 609 970, 25)
Después de limpieza (tuplas)	(4 620 455, 25)
Train (filas)	3 696 362
Val (filas)	924 093
Decisiones (train)	814 793

Tabla 4.4: Resumen de datos y particionado

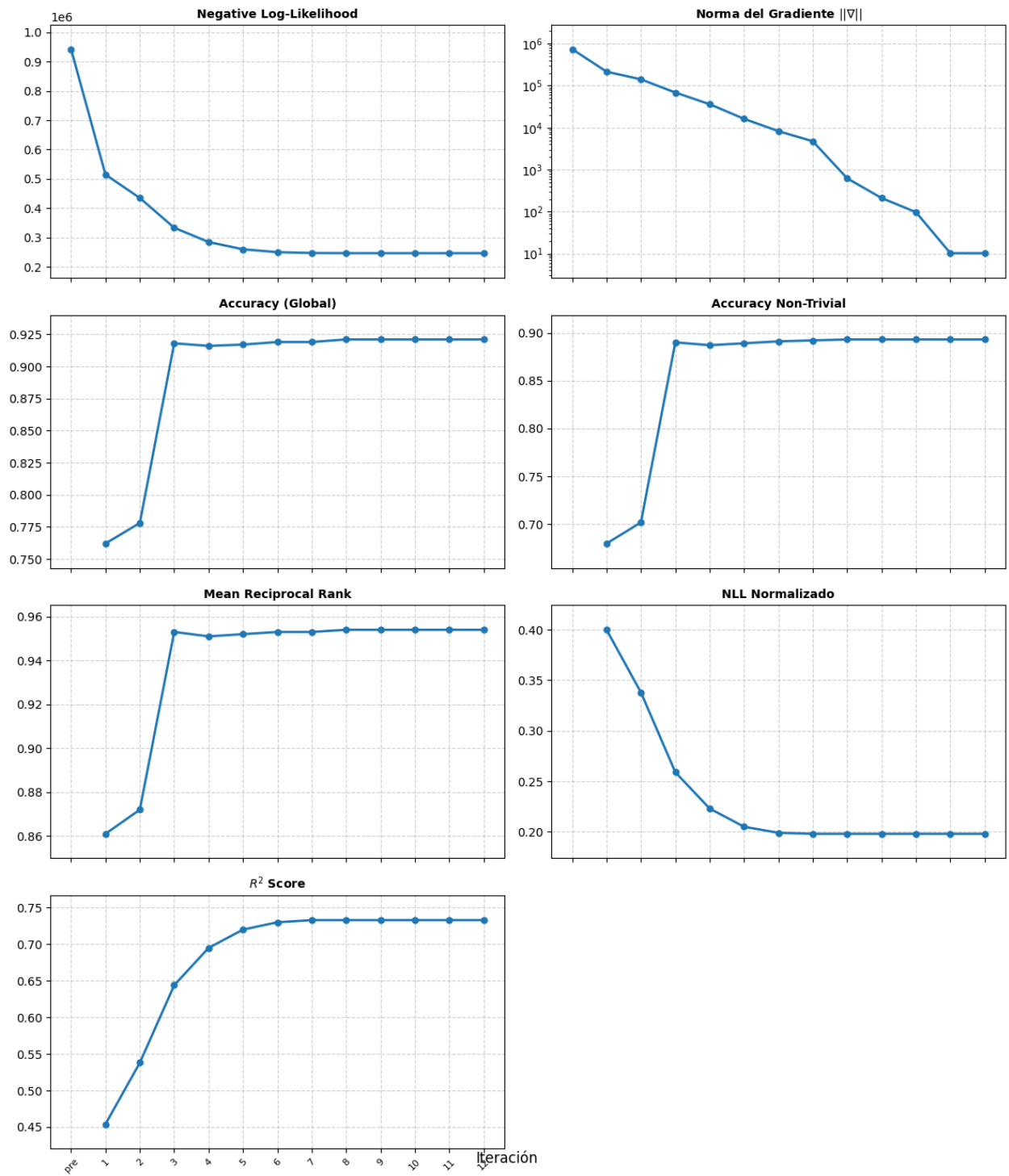


Figura 4.3: Resultados del modelo MNL. En todos los gráficos, el eje x es la época y el eje y la métrica correspondiente.

Parámetro	Valor
intercept	6.10933635e-13
wait_time	-9.01719464e-01
viajar_cost	8.10947587e-02
cost_to_go	-5.61161490e+00
first_walk_min	-9.16079751e-02
is_initial_transfer	-1.83215950e-01
zero_onboard	2.22291522e+00
ASC_metro	1.85152930e-13

Tabla 4.5: Coeficientes del modelo MNL

Métrica	Train	Val
NLL	0.3029	0.3078
Acc	0.922	0.921
AccNT	0.895	0.893
MRR	0.954	0.954
NLLn	0.196	0.198
R^2	0.738	0.733

Tabla 4.6: Métricas de la mejor época de la MNL Semanal

Se observa que el modelo fue entrenado exitosamente, con una precisión del 92% en ambas particiones. La precisión no trivial es del 89%, evidenciando que el modelo supera significativamente la predicción aleatoria.

Notar como afecta más a la utilidad tener un *cost_to_go* alto que un tiempo de viaje alto. Con esto se puede concluir que los usuarios prefieren alternativas que le acerquen lo más que puedan al destino, inclusive pagando más coste de viaje que otro servicio alimentador.

Se obtuvieron constantes positivas en el coste de viajar. Una colinealidad entre el coste restante (*cost to go*) y el tiempo de viajar puede ser una señal de esto. Si se mira desde un punto de vista de comodidad, un coste restante menor indica que el viaje tiene menos transbordos probablemente y es más directo. Entonces, un coste restante menor es más atractivo. Para tener un costo restante menor, es necesario viajar más tiempo en el primer servicio.

No se logran observar diferencias sustanciales entre los días de semana y fines de semana. Un análisis usando más semanas debe de ser imperativo para extraer conclusiones en cuanto a este tema.

4.3 GNN

4.3.1 Resultados

A continuación se presentan los resultados de la GNN en sus dos modos, el primero para cuando se agrega a los *embeddings* las características del MNL, y posteriormente cuando se omiten estas características. Recordar que este entrenamiento se hizo de la misma manera que el MNL semanal, es decir, el split y el sampleo de datos (0.1 de los reales) con la misma semilla.

GNN con características del MNL

Elemento	Valor
Embedding destinos	10 687 (dim = 128)
Scorer dinámico	$393 \rightarrow 128 \rightarrow 1$
Parámetros entrenables	$\approx 1\,717\,765$

Tabla 4.7: Resumen del modelo

Historial de Entrenamiento: Train vs Validation

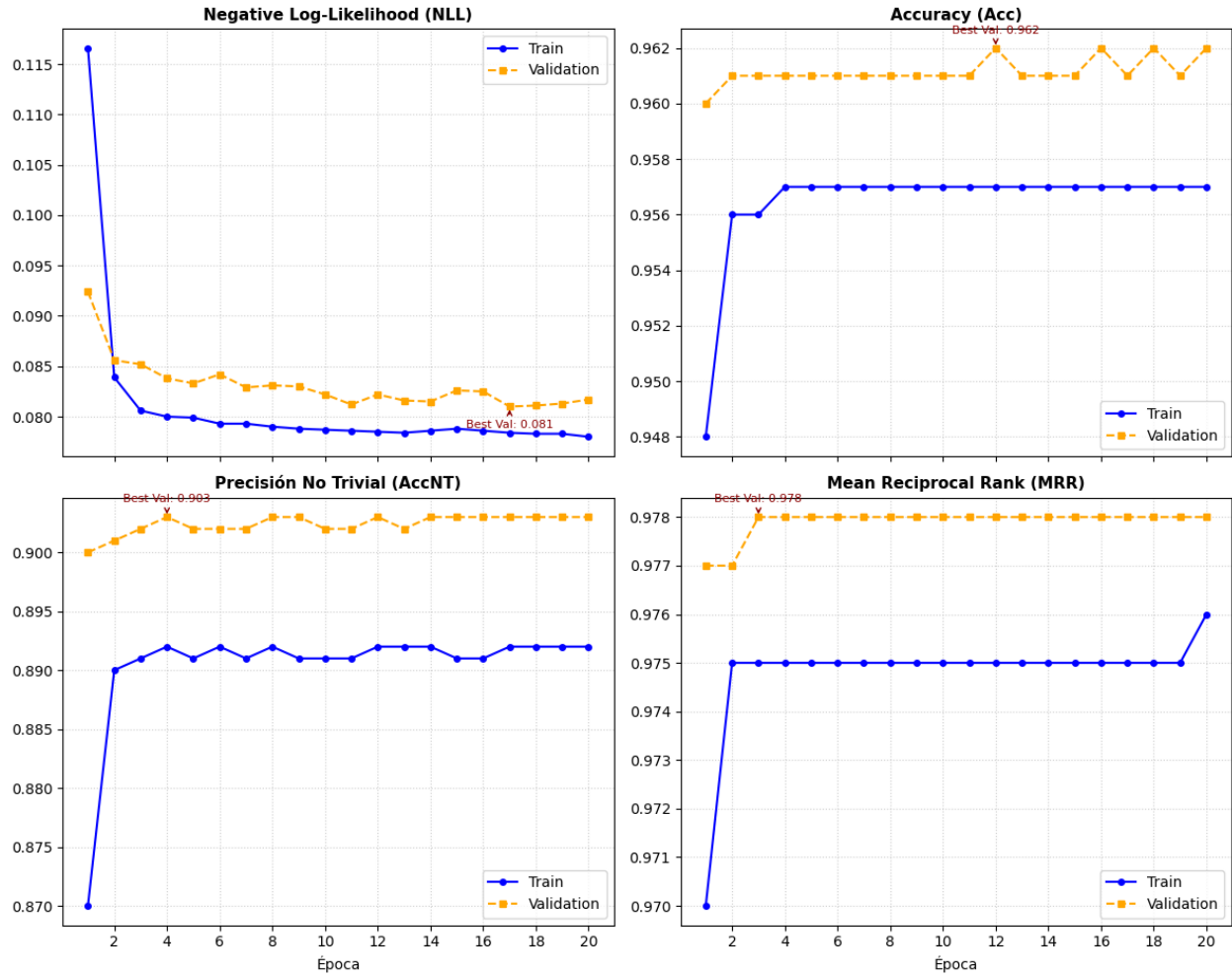


Figura 4.4: Resultados de entrenamiento de la GNN con características del MNL activadas

Con ello se obtiene un modelo en que en su mejor época tiene unas métricas de Precisión del 96.2% y Precisión No Trivial del 90.3%.

GNN sin características del MNL

Elemento	Valor
Embedding destinos	10687 (dim = 128)
Scorer dinámico	384 \rightarrow 128 \rightarrow 1
Parámetros entrenables	\approx 1,716,613

Tabla 4.8: Resumen del modelo

Historial de Entrenamiento: Train vs Validation

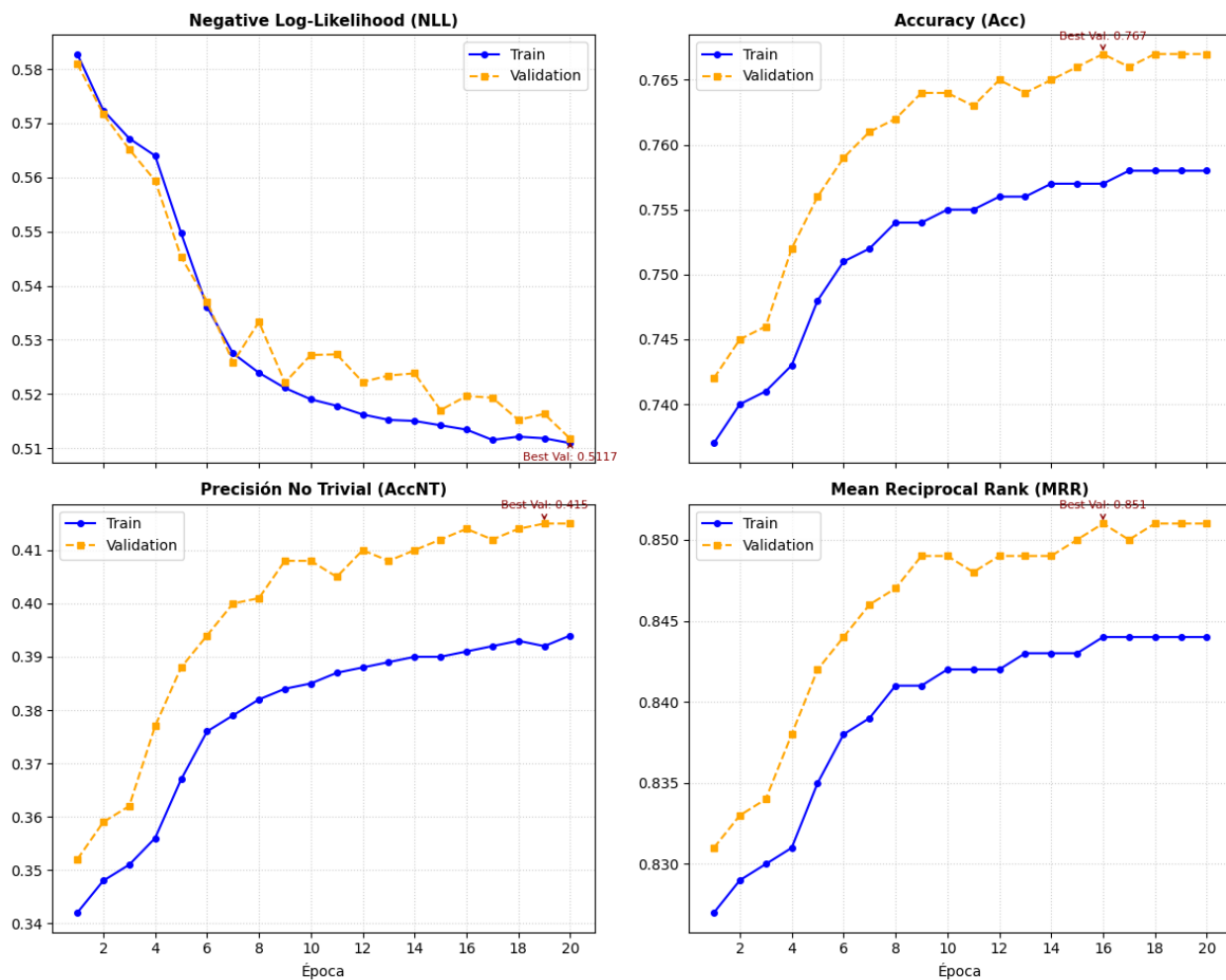


Figura 4.5: Resultados de entrenamiento de la GNN sin características del MNL

Mejor época (validación)	20 (Val NLL: 0.5117)
Test NLL	0.5169
Test Acc	0.766
Test AccNT	0.415
Test MRR	0.850

Tabla 4.9: Mejor modelo y evaluación en test

Discusión de los resultados

- Claramente se observa una diferencia entre colocar una capa de MNL versus no colocarla. Sin los datos del motor de costes del MNL, el modelo no consigue buenos resultados. Una precisión no trivial del 41.5% es mejor que el azar, pero peor que el MNL y que el GNN con las características del MNL.
- Comparado con el MNL, la GNN con las características del MNL tuvo una precisión no trivial del 90% (+1% en comparación con el MNL) y una precisión del 96% (+3% en comparación con el MNL), hay una leve mejora aportada por los *embeddings*. Se concluye que agregar los *embeddings* y los parametros tau ayudaron, pero estos por si solos no son suficientes para conseguir buenas métricas con el set de datos que se tiene.

4.4 Experimentos

Se realizaron experimentos para mostrar redistribución de demanda. Para ello, se usaron los coeficientes de la tabla 4.10 obtenidos de un día viernes. No se usó la GNN debido a que tomaba más tiempo y que los resultados obtenidos eran muy parecidos a los de la MNL cuando tenía la capa de las características del MNL activada. Como ambos tienen resultados parecidos, la MNL es más interpretable y además tarda menos tiempo en generar resultados, se decide por usar la MNL como motor de utilidad.

Coeficiente	Valor
intercept	5.94×10^{-13}
wait_time	-0.96
viajar_cost	-0.04
cost_to_go	-5.64
first_walk_min	-0.10
is_initial_transfer	-0.20
zero_onboard	2.13
ASC_metro	1.79×10^{-13}

Tabla 4.10: Coeficientes del modelo MNL usado para los experimentos

Estos valores reflejan la importancia relativa de cada atributo en la elección de alternativas de viaje según el modelo MNL entrenado. Para el predictor solo se usarán los coeficientes `wait_time`, `viajar_cost`, `cost_to_go` y `first_walk_min`, debido a que los otros valores son constantes, muy cercanos a cero, o derivados de los coeficientes principales.

4.4.1 Disminución de oferta de un servicio

Se tiene un paradero P y Q conectados por un set de servicios S para un bin b . En el *baseline* (la oferta real) se obtiene una distribución de probabilidad dada. Si se modifica la oferta de uno de los servicios, por ejemplo, aumentando el doble el tiempo de espera (disminuyendo la cantidad de buses que operan el servicio) se obtiene una comparación entre las distribuciones de probabilidades para antes y después del cambio de oferta. Se muestran dos ejemplos ilustrativos.

Ejemplo 1: Ir desde PJ394 a PA300

Ambos paraderos tienen de servicios disponibles que dejan directo en el destino, el 503 y el 517. Entonces, el costo restante o `cost_to_go` es 0, ya que dejan directamente en el destino del usuario. Ver Figura 4.6 que ilustra los tiempos de cada servicio del paradero. El experimento consiste en aumentar al doble el tiempo de espera del 517.

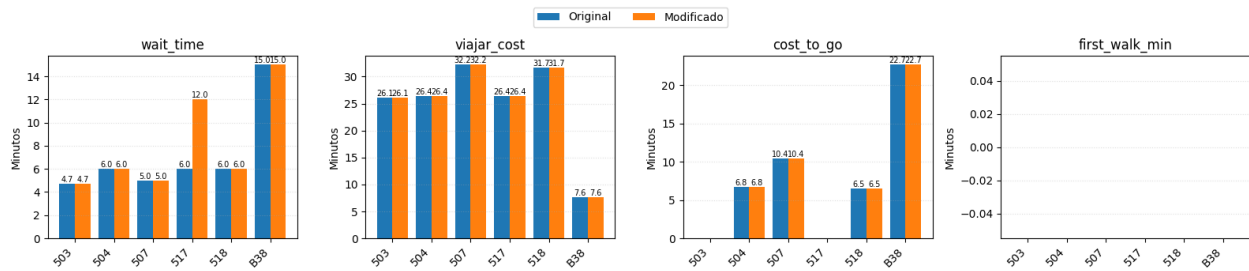


Figura 4.6: Costes para ir desde PJ394 a PA300. En azul se muestran los costes reales y en naranja los cambios de oferta. Notar como el tiempo de espera del servicio 517 aumenta (la barra naranja es mas alta que la azul).

Si se ejecuta el predictor, se obtiene una redistribución de probabilidades como la mostrada en la Figura 4.7.

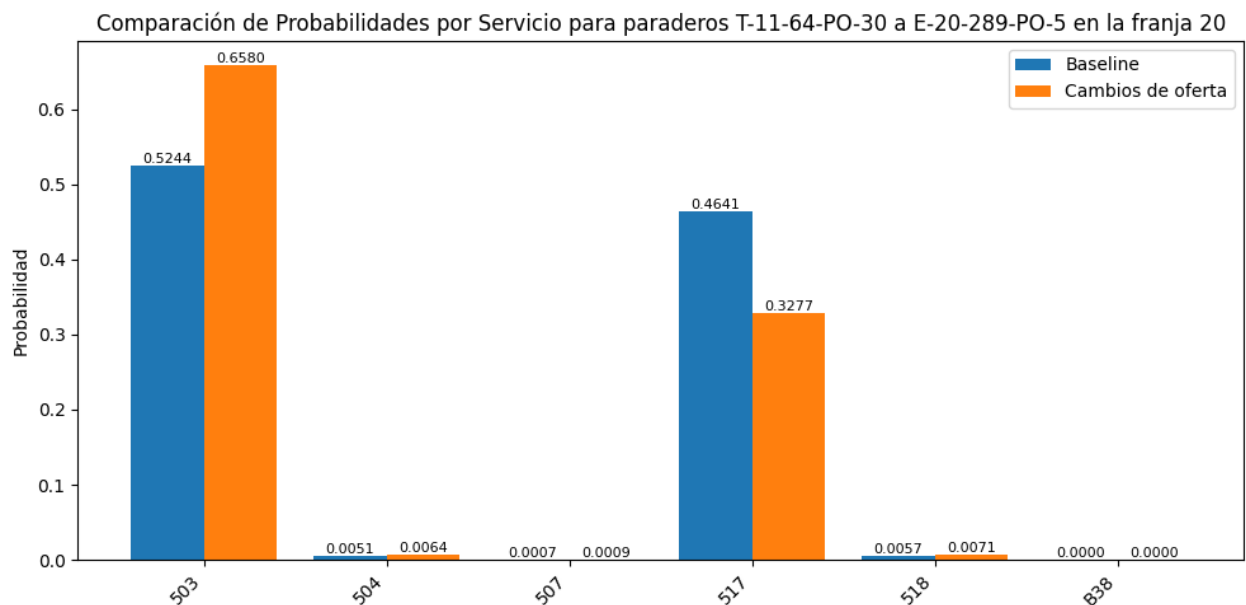


Figura 4.7: Distribución de probabilidades para alternativas de viaje antes y después del cambio de oferta

Notar como el servicio 503 pierde probabilidad y el 517 la gana. Pero no es una transferencia directa. Los otros servicios igual ganan un poco de atractivo al perderlo el 503.

Ejemplo 2: Ir desde PJ394 a PA433

Este ejemplo es distinto. A diferencia del anterior, efectivamente solo un servicio llega directamente al destino, el 507. El resto entonces, tiene un costo restante mayor que cero. Ver Figura 4.8 que ilustra los tiempos de cada servicio del paradero.

Al ejecutar el MNL, se obtuvo una redistribución de probabilidades como la mostrada en la Figura 4.9.

Notar que no cambia mucho la probabilidad del servicio 507. A pesar de que su tiempo de espera se duplica, sigue siendo la mejor alternativa.

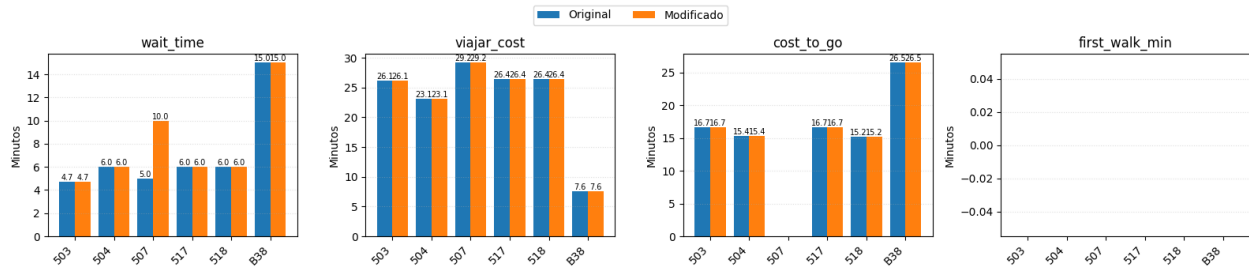


Figura 4.8: Costes para ir desde PJ394 a PA433

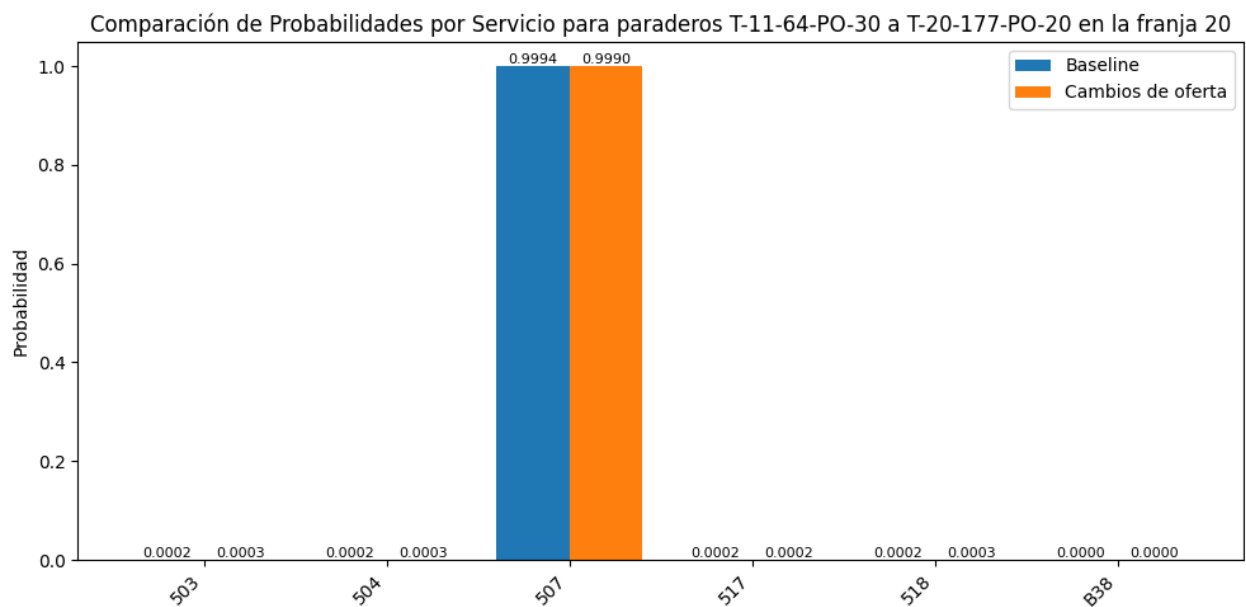


Figura 4.9: Distribución de probabilidades para alternativas de viaje antes y después del cambio de oferta (Experimento 2)

Se observa que:

- Dos servicios *compiten* cuando tienen costes de viaje y restantes parecidos. Es decir, dos servicios que llevan directamente al destino o tienen trayectos parecidos. Un ejemplo de esto es el 503 y el 517. El coeficiente del costo restante evidencia este comportamiento. Viajes sin transbordos son altamente atractivos para los usuarios.
- Cuando hay servicios que compiten, se obtiene una redistribución de la demanda más notoria como fue el caso del 503 vs el 517. En el caso del 507, al no tener ningún servicio que compita directamente, un mayor tiempo de espera no influye en lo atractivo del servicio.

Se sigue aumentando el tiempo de espera, hasta un 1500% más grande que el original (esto hace que el tiempo de espera pase a 100 minutos). Se obtiene una redistribución de probabilidades como la que sigue en la Figura 4.10.

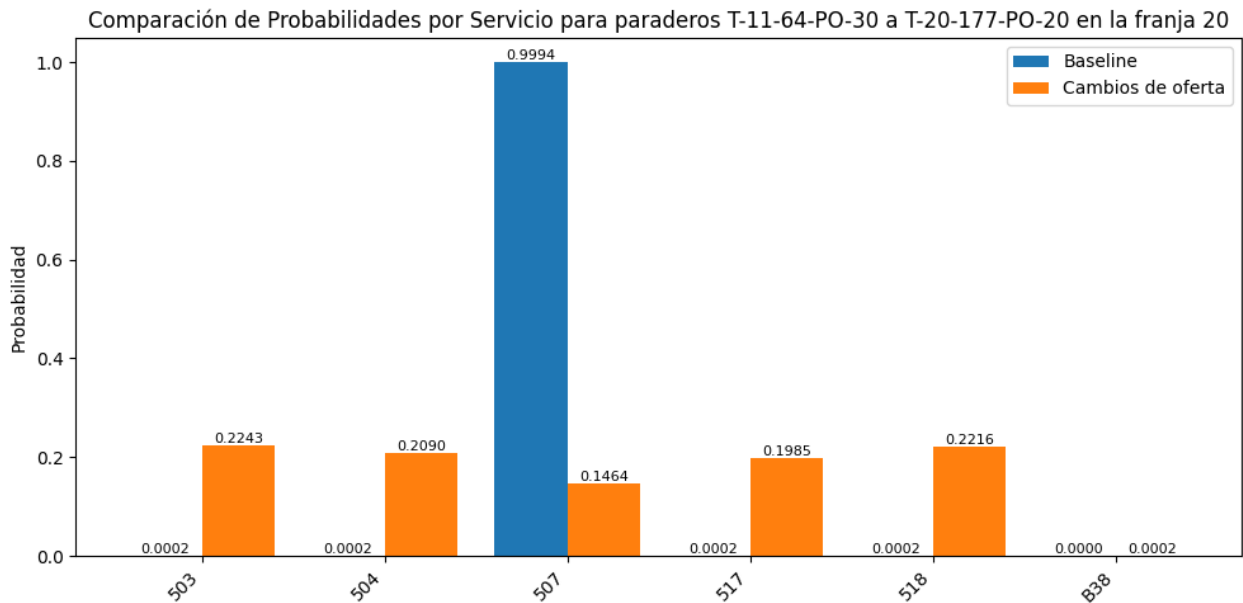


Figura 4.10: Distribución de probabilidades para alternativas de viaje con un aumento del 1500% en el tiempo de espera del servicio 507

Esta redistribución de demanda causará un efecto dominó que cambiará los transbordos siguientes. Por un efecto de simplicidad, el siguiente paso de decisión será determinístico tal como se ha asumido durante este trabajo. Cuando un usuario se baje en un paradero dado para hacer transbordo, se tomará el siguiente servicio de manera segura, y no con probabilidades. (Si no, sería una cadena de probabilidades condicionales que complicaría mucho el problema).

Para cada alternativa, no solo aumentará la demanda del servicio dado, sino que su transbordo aumentará también de demanda. En el caso de ir de PJ394 a PA433, los servicios que aumentaron su demanda alimentarán a los siguientes servicios en su transbordo. Para ello, se verán los caminos de cada servicio obtenidos por Dijkstra. La tabla 4.11 muestra los caminos que toma cada alternativa junto con la diferencia de probabilidad entre el *baseline* y el cambio de oferta. Un análisis indica que los servicios que aumentaron su demanda propagaran este aumento de demanda a los transbordos, en este caso, fijarse en 503, 504, 517 y 518. Estos recorridos dejan a usuarios en L2 en Santa Ana, por lo que es sensato concluir que un aumento de tiempo de espera en 507 provoca un aumento de demanda de L2 sujeto a que las personas se suban a PJ394.

Una vez se bajen en Parque Ohiggins, un efecto importante ocurre en el paradero aledaño a la estación de Metro. El AD predice que se tomará el 506, pero realmente es el servicio más probable a ser tomado, no necesariamente todos lo tomarán. Esta nueva demanda redistribuida se repartirá en los servicios que pasan por este paradero y que llevan a Beauchef, en este caso, el 506, 506v, 506e y el 507. Ver la tabla 4.11 para visualizar los trayectos y su cambio en su probabilidad.

Alternativa (Porcentaje en comparación al *baseline*)	Itinerario
B38 (+ 0% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio B38 (Ida) en T-11-64-PO-30</p> <p>Bajar en T-8-64-PO-30</p> <p>Subir al servicio 507 (Ida) en T-8-64-PO-30</p> <p>Bajar en T-20-177-PO-20</p>
503 (+ 22% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio 503 (Ida) en T-11-64-PO-30</p> <p>Bajar en E-20-289-PO-5</p> <p>Caminar de E-20-289-PO-5 a METRO CAL Y CANTO</p> <p>Subir al servicio L2 (Metro) en METRO CAL Y CANTO</p> <p>Bajar en METRO PARQUE OHIGGINS</p> <p>Caminar de METRO PARQUE OHIGGINS a E-20-189-OP-40</p> <p>Subir al servicio 506 (Ret) en E-20-189-OP-40</p> <p>Bajar en T-20-177-OP-8</p> <p>Caminar de T-20-177-OP-8 a T-20-177-PO-20</p>
504 (+ 20% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio 504 (Ida) en T-11-64-PO-30</p> <p>Bajar en T-20-188-NS-10</p> <p>Caminar de T-20-188-NS-10 a METRO SANTA ANA</p> <p>Subir al servicio L2 (Metro) en METRO SANTA ANA</p> <p>Bajar en METRO PARQUE OHIGGINS</p> <p>Caminar de METRO PARQUE OHIGGINS a E-20-189-OP-40</p> <p>Subir al servicio 506 (Ret) en E-20-189-OP-40</p> <p>Bajar en T-20-177-OP-8</p> <p>Caminar de T-20-177-OP-8 a T-20-177-PO-20</p>
517 (+ 19% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio 517 (Ida) en T-11-64-PO-30</p> <p>Bajar en E-20-289-PO-5</p> <p>Caminar de E-20-289-PO-5 a METRO CAL Y CANTO</p> <p>Subir al servicio L2 (Metro) en METRO CAL Y CANTO</p> <p>Bajar en METRO PARQUE OHIGGINS</p> <p>Caminar de METRO PARQUE OHIGGINS a E-20-189-OP-40</p> <p>Subir al servicio 506 (Ret) en E-20-189-OP-40</p> <p>Bajar en T-20-177-OP-8</p> <p>Caminar de T-20-177-OP-8 a T-20-177-PO-20</p>
518 (+ 22% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio 518 (Ida) en T-11-64-PO-30</p> <p>Bajar en T-20-203-NS-20</p> <p>Caminar de T-20-203-NS-20 a METRO SANTA ANA</p> <p>Subir al servicio L2 (Metro) en METRO SANTA ANA</p> <p>Bajar en METRO PARQUE OHIGGINS</p> <p>Caminar de METRO PARQUE OHIGGINS a E-20-189-OP-40</p> <p>Subir al servicio 506 (Ret) en E-20-189-OP-40</p> <p>Bajar en T-20-177-OP-8</p> <p>Caminar de T-20-177-OP-8 a T-20-177-PO-20</p>
507 (- 86% de probabilidades)	<p>Inicia en paradero T-11-64-PO-30</p> <p>Subir al servicio 507 (Ida) en T-11-64-PO-30</p> <p>Bajar en T-20-177-PO-20</p>

Tabla 4.11: Itinerarios de las alternativas desde T-11-64-PO-30 hasta T-20-177-PO-20

Notar que si hay 100 personas que quieren ir a Beauchef en un día, las 100 tomarían el 507 en el caso base. En el caso modificado, aumentaría la demanda del día en 80 para L2 y para 506. Esto es el efecto dominó del que se comentó al comienzo del informe que se debería de analizar.

Cuantificar los cambios de demanda en cuando hay cambios de oferta se vuelven interesantes cuando se prueban situaciones más realistas. Se puede por ejemplo, cortar la línea 1. Esto es lo que se hizo en el siguiente ejemplo.

4.4.2 Suspensión de un servicio

Para el siguiente experimento, se colocó un indicador booleano que desactivó todas las aristas SUBIR, BAJAR Y VIAJAR de la L1. Haciendo esto, tomar el caso de alguien que quiera ir de San Pablo a Baquedano.

Las figuras 4.11 y 4.12 muestran las probabilidades y costes para cada alternativa. Notar como las probabilidades de usar la L1 bajan y las de la L5 suben. Nótese que estas probabilidades están condicionadas a que el usuario haya decidido ir a Baquedano.

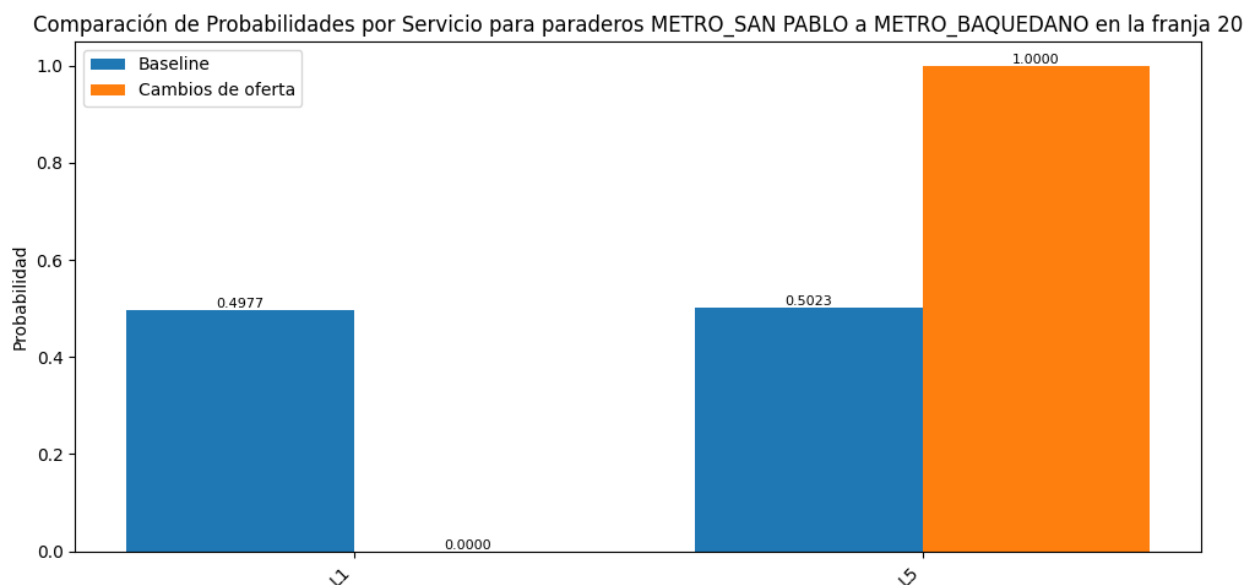


Figura 4.11: Distribución de probabilidades para alternativas de viaje con suspensión de la Línea 1

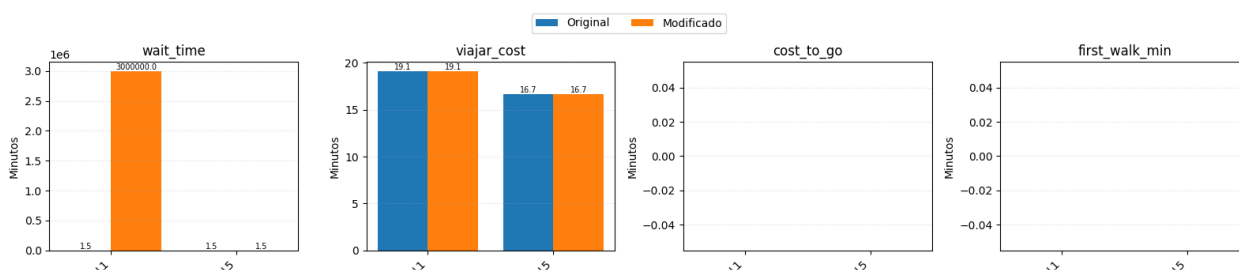


Figura 4.12: Costes para ir de San Pablo a Baquedano con suspensión de la Línea 1, antes y después de colocar un tiempo de espera en L1 muy grande.

Algo más interesante pasa cuando se quiere ir a una estación de L1 que no combine con L5, es decir, ambos servicios no compiten. Por ejemplo, ir de San Pablo a Tobalaba.

Viaje en alternativa: L5

- **Inicio:** Paradero METRO_SAN PABLO

- Subir al servicio **L5** (sentido Metro) en METRO_SAN PABLO
- Bajar en METRO_BAQUEDANO
- Caminar desde METRO_BAQUEDANO hasta E-20-53-PO-115
- Subir al servicio **503** (sentido Ida) en E-20-53-PO-115
- Bajar en E-14-170-NS-5
- Caminar desde E-14-170-NS-5 hasta METRO_TOBALABA

• Fin del camino

El algoritmo que se tiene, por construcción tomará el camino con el coste más bajo para ir desde Baquedano hacia Tobalaba si es que el Metro está desactivado. Se puede hacer este análisis corriendo el algoritmo del MNL desde el paradero E-20-53-PO-115 (el más cercano a Baquedano que tiene servicios que dejan cerca, según el enrutador). Las figuras 4.13 y 4.14 muestran los costes y probabilidades para cada alternativa. Notar como el servicio 503 es el más atractivo, ya que es el que tiene el menor coste total. Una suspensión de la L1 entre Baquedano y Tobalaba sugiere que todo el volumen de pasajeros que usualmente toma este tramo se redistribuirá en los servicios en superficie con las probabilidades de más abajo.

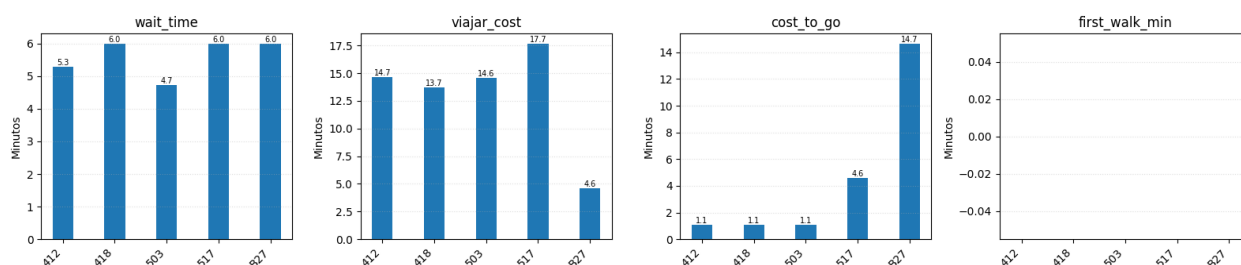


Figura 4.13: Costos del paradero E-20-53-PO-115 a Tobalaba con suspensión de la Línea 1

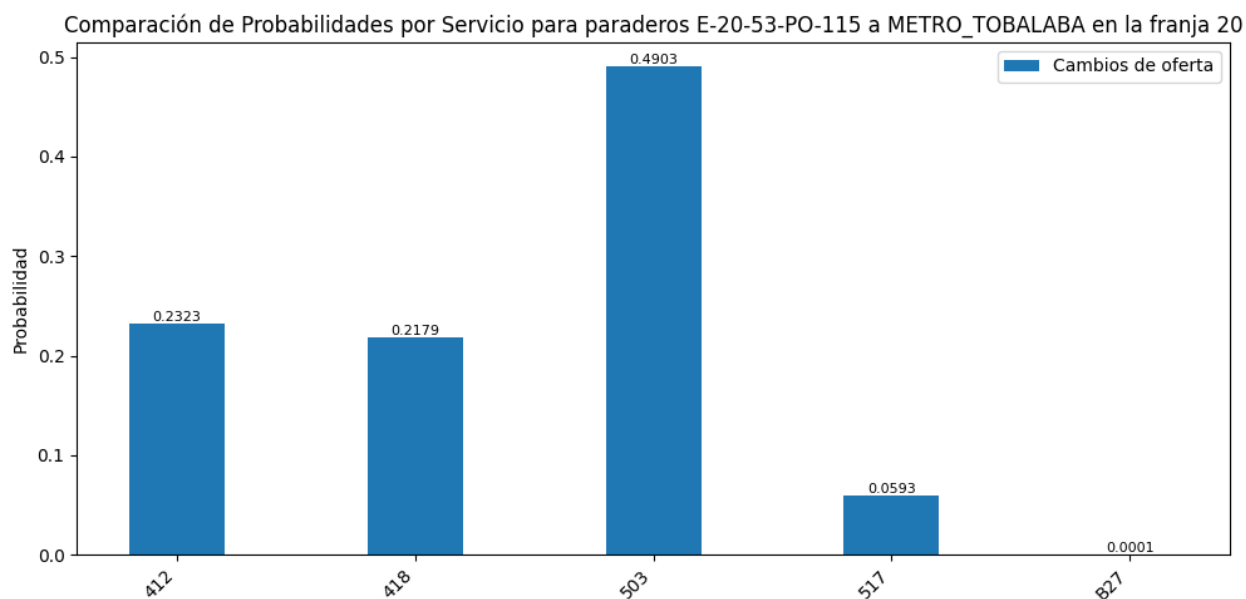


Figura 4.14: Distribución de probabilidades para alternativas con suspensión de la Línea 1

Al ver esta redistribución, la primera medida a tomar sería reforzar con recorridos como el 412, 418 y 503 pues tienen un `cost_to_go` bajo (es decir, acercan más a Tobalaba), en cambio, no reforzar servicios con recorridos similares al 517 y B27, pues por probabilidad, no deberían de ser elegidos para llegar a Tobalaba.

Este análisis solo funciona para las personas que quieran tomar el metro en Baquedano. Un análisis más profundo debería de considerar a todas las personas que quieran tomar algún tramo de la Línea 1 y se encuentren con un corte de servicio. Por ejemplo, una distribución de usuarios desde Baquedano a Tobalaba por cada estación intermedia tomaría distintos servicios que le dejen en Tobalaba dependiendo del origen y sus estrategias de elección.

4.4.3 Agregar Línea 7

Para el siguiente experimento, se agregó la L7 y se obtuvieron las siguientes métricas:

Número de etapas promedio

Nº Etapas Promedio (Real): 1.73 Nº Etapas Promedio (Sintético): 2.15

Hay un aumento de etapas. Esto debido a la inclusión de un servicio nuevo. La L7 actúa como un servicio intermedio interesante para los usuarios.

Uso de servicios (Líneas de Metro)

Se procede a comparar el uso antes y después de la L7. Sintético se refiere a este último.

Servicio	Etapas (Real)
L1	880 685
L5	624 255
L2	435 869
L4	371 041
L3	365 217
L6	173 456
L4A	71 110
T506 00I	10 168
T506 00R	9859
T556 00R	9715

Tabla 4.12: Uso de Servicios (Etapas Reales)

Servicio	Etapas (Sintético)
L1	1 173 755
L5	791 909
L2	557 354
L3	492 883
L4	480 813
L6	193 936
L4A	90 910
506	56 032
406	48 794
L7	40 516

Tabla 4.13: Uso de Servicios (Etapas Sintéticas)

Se observa un aumento sustancial en el uso del Metro en general. Esto puede ser causa de los siguientes puntos:

- L7 disminuye el uso del servicio de buses. Estos usuarios antes llegaban directamente a su destino usando una micro. Ahora, usando la L7, inyectan demanda directamente al servicio subterráneo, haciendo que la demanda nueva ganada por L7 se distribuya por toda la red.
- Las líneas de metro comienzan a alimentar la L7, ya que esta está bien conectada con el resto de la red, haciendo más apetecibles los viajes en metro en general.

Uso de la L7 como primera etapa

No se registraron viajes con L7 en primera etapa, esto por la restricción del paradero inicial y el coste de transbordo inicial. Se ahondará más en esta decisión en la discusión final.

Servicios alimentadores de L7

Los servicios alimentadores se definen como los que son transbordo directo a L7.

Notar como servicios que comparten recorrido con la L7 (por ejemplo, el 508 que recorre junto a la L7 desde Mapocho con Huelén hasta Salvador) sirven como alimentadores. Esto debido a que la L7 tiene estaciones más espaciadas que la L1, haciendo que es recomendable tomar un recorrido que acerque a caminar.

Servicios alimentadores como el J01 son del tipo que recogen a usuarios los cuales no viven cerca del trayecto de la L1. Por ejemplo, J01 recorre toda la avenida Neptuno desde Carrascal hasta General Bonilla. La mayoría de los usuarios tomaba J01 para acercarse a Metro San Pablo. Ahora, la usarán para acercarse a Metro Cerro Navia. Es esperable notar un aumento de demanda en J01.

Servicio Alimentador (Etapa k-1)	N° de Viajes a L7 (Etapa k)
L1	5203
L5	4053
406	3311
L2	2533
426	1814
L3	1656
405	1617
508	1465
C01	1186
J01	1142

Tabla 4.14: Servicios Alimentadores Principales de L7 (Transbordos)

Comparación de carga total (número de aristas viajar recorridas por servicio)

Línea	Total Estaciones (Real)	Total Estaciones (Sintético)	Diferencia
L6	948 250	984 713	36 463
L4A	371 047	407 921	36 874
L4	3 585 554	3 633 894	48 340
L2	3 928 697	4 020 191	91 494
L5	6 095 884	6 228 347	132 463
L1	7 546 426	7 725 747	179 321
L3	2 464 663	2 734 873	270 210
L7	–	383 674	383 674

Tabla 4.15: Comparación de Carga de Metro (Total Estaciones-Pasajero Recorridas)

Hubo un aumento de carga en toda la red por lo anteriormente dicho. La L7 ganó viajes, pero no tanto como se esperaba.

Se observan ciertas desviaciones a lo que se espera de esta línea y su desempeño. Esto era, una baja en la demanda de L1. Razones para esto tienen que ver con la solución propuesta y como introduce artefactos. El coste del transbordo inicial y la restricción de los puntos extremos fijos (el origen y el destino) que hacían como anclas del camino final indujeron desviaciones. En una situación real, una persona podría cambiar su paradero de viaje inicial para llegar a L7 y tomarla como etapa 1. Este comportamiento es difícil de replicar, pues no se sabe donde vive cada persona.

De todas maneras, un efecto interesante descubierto, y no predicho en las hipótesis, es como la L7 le quita demanda a servicios de buses. Esta redistribución inyecta más usuarios a la red de Metro.

También se nota un aumento de usos de servicios alimentadores, tal como se mencionó en la introducción y en la literatura.

4.5 Discusión final de la solución

A continuación se discuten aciertos y limitaciones de los aspectos claves de la solución entregada.

Transbordos determinísticos

Los transbordos o viajes con más de una etapa fueron tratados de manera determinista en sus etapas posteriores a la inicial, esto quiere decir que después de bajarse, el costo restante es definido de manera estricta. Una solución interesante puede ser concatenar varios MNL para cada etapa, pero esto complica mucho el problema. Notar que este enfoque habría hecho el valor *cost_to_go* no determinado, si no que una distribución o valor esperado. En esta memoria el *cost_to_go* es el mínimo dado que el usuario se baja en el paradero óptimo y elija el servicio óptimo. Es una simplificación fuerte, pero que funciona en gran parte de las decisiones, ya que el entrenamiento incluyó a etapas intermedias, es decir, que se obtuvo un 89% de precisión no trivial en todas las etapas, por lo que es seguro decir que el modelo predice correctamente cualquier etapa del viaje. Además recordar que gran parte de los viajes son de una etapa (el promedio de 1.75 etapas en los datos históricos lo demuestran), por lo que los transbordos son menos comunes de lo que se podría esperar.

Distribución probabilística

Un acierto claro fue la inclusión de probabilidades en la elección de alternativas. Esto permite modelar la incertidumbre y variabilidad de los usuarios al elegir alternativas que tengan utilidades parecidas. Ejemplos de ello se observan en los experimentos, donde servicios que compiten directamente, como lo son la L1 y L5 en San Pablo a Baquedano, o el 503 y 517 en PJ394 a PA300, muestran redistribuciones de demanda más equiprobables.

Algunos usuarios prefieren esperar mas para llegar directamente, otros prefieren tomar el primer bus que llegue para esperar menos. Todas estas decisiones modifican el parámetro beta que acompaña a estas características, haciendo que el MNL incluya muchas formas de decisión de los usuarios.

Grafo Bipartito

El grafo bipartito resultó ser conveniente para modelar la red de transporte. La separación entre paraderos y servicios permitió modelar de mejor manera los tiempos de espera y los costos de viaje. Además, la inclusión de aristas de caminar. Fue sencillo obtener caminos con los algoritmos de ruteo, ya que cada arista denotaba una acción clara y definida, además de poder separar los transbordos, evento que en un grafo común sería imposible de modelar.

Correlación espacial (Para el MNL)

Tal como se mencionó al inicio de la memoria, un usuario puede decidir en base a *clusters* o paraderos cercanos que le proveen una mejor oferta. El caso de preferir paraderos con más servicios competitivos que otros puede ser un factor importante, tanto, que el usuario puede preferir a caminar más para tener un tiempo de espera más corto o confiable. Efectos como los de preferir el Metro por sobre otros modos de viaje debido a su fiabilidad y su alta frecuencia no son modelados. En un principio la GNN buscaba capturar estos efectos, y la mejora en las métricas lo evidencia, pero no logró una mejora sustancial como para que valiera la pena su complejidad y

uso en los experimentos.

Elección de paradero inicial y final

Esta limitación viene más por el lado de los datos de ADATRAP. Lógicamente no se sabe donde vive la gente, solo su paradero de inicio del viaje y el del final de éste. Por ello, el viaje ya viene condicionado a que se eligió un paradero determinado desde el comienzo. Esto causa que nuevos servicios agregados nunca tengan demanda en la etapa 1, artefacto que causó una notable diferencia de uso de la línea 7 con lo que se esperaba, pues ningún viaje comienza en paraderos que recorren la línea, a no ser que sean buses, caso que no se exploró en esta memoria. También el paradero final actúa como una ancla, no tan fuerte como el paradero inicial, pues un usuario puede caminar hasta el paradero final para terminar el viaje, pero eso agrega un costo que algunas veces no vale la pena para tomar ese camino. El hecho de que existan estas anclas, hace que sea complicado quitarle demanda a los servicios iniciales y finales, pues en el caso del metro, son el único servicio en el paradero la mayor parte de las veces.

Un ejemplo analizado fue ir desde Neptuno con Mapocho hasta Pedro de Valdivia (L7). En este caso, para el estado actual de la red, un usuario probablemente tome un bus, sea el J01 o el J08 para llegar a San Pablo y tomar el metro que estime conveniente. Cuando se agrega la L7, teniendo en cuenta que en Neptuno con Mapocho se instalará la estación Cerro Navia (L7), el usuario podría preferir tomar la L7 desde Cerro Navia. Pero esto no se modela, pues el paradero inicial es fijo por los datos de ADATRAP. Entonces, el usuario no puede elegir tomar la L7 en etapa 1, pues su paradero inicial no está en Cerro Navia (L7), entonces el modelo compara entre caminar (cinco minutos de penalización) o ir a San Pablo en J01 o J08. Una forma de solucionar esta penalización sería desanclar al usuario de su paradero inicial y dejarlo elegir un paradero inicial en un radio alrededor del paradero inicial real. Esto permitiría elegir un paradero no existente en los datos reales y comenzar con viajes en etapa 1 en L7, pero podría producir artefactos o desviaciones de la realidad, ya que la distancia euclidiana recta es distinta a la real caminable. Un paradero a trescientos metros caminando puede estar realmente mucho mas lejos, ya que se debe cruzar una autopista, un barrio entero o una avenida sin cruces peatonales.

Coste de la primera transferencia

El modelo penaliza de la misma manera caminar diez minutos reales para elegir el paradero al inicio del viaje, que caminar un minuto. Esto causa que la elección de la Línea 7 sea más acotada, pues los usuarios comienzan su viaje en el paradero real que eligieron, y luego cambian a la Línea 7, aunque quede a un par de metros, sea más costoso que tomar un servicio y combinar más adelante. Esto contamina la elección real más probable.

Comodidad

Un factor importante no modelado. La comodidad puede verse afectada dinámicamente según la hora y el momento del recorrido. Un servicio que va lleno es menos cómodo que uno que va vacío. Modelar el llenado de los buses requeriría saber exactamente el tamaño de los buses, el tipo de flota que tiene cada servicio y que tipo de buses saca cada servicio dependiendo del bin, datos que no se encontraban en el programa de operaciones, salvo un contador general de capacidad de todos los buses, pero no individual a cada vehículo que circula.

Velocidad promedio v/s velocidad por arista

Los datos entregados por red muestran datos de velocidad promedio en todo el recorrido. Claramente hay trazos del recorrido más lentos que otros, probablemente los más lentos son en áreas céntricas mientras que los rápidos son en áreas suburbanas. Esto puede afectar localmente en viajes cortos, tomando costos de viaje más pequeños que los reales. Este efecto se puede amortiguar en viajes cortos cuando los servicios que compiten en las alternativas comparten el eje de circulación, pero por ejemplo un servicio que no usa avenidas puede verse perjudicado ante uno que alcanza velocidades mayores.

Tarifas

No se tuvo en cuenta el coste del viaje. Esto afecta claramente el uso del servicio. Personas pueden preferir evitar el metro pues es más costoso.

ADATRAP

Al depender de las predicciones de ADATRAP, se heredan sus errores y sus casos de uso límites. Por ejemplo, aproximadamente el 60% de los viajes tenían el punto de bajada guardado, lo que ya limita los datos de entrenamiento.

Evasión

El porcentaje de evasión del transporte público en Santiago es alto, aproximadamente del 30% en las últimas estimaciones. No se tienen distribuciones probables de evasión por paradero, lo que puede afectar la demanda real en ciertos servicios.

Capítulo 5

Conclusión

Para concluir, es importante recordar el Problema y el Objetivo general de esta memoria:

Partiendo desde el problema definido como *Generar datos sintéticos de demanda de transporte público*.

Se tiene como objetivo *Diseñar e implementar un modelo que prediga demanda de transporte dado un escenario (definido como una configuración de red y su respectiva infraestructura urbana); y usar este modelo para predecir demanda en distintos escenarios para medir el impacto de intervenciones en el escenario actual*.

Desde ese contexto y objetivo, se abordan dos métodos. Uno del MNL y otro el MNL/GNN. Para ambos enfoques, la solución se enfocó en predecir una alternativa a usar. Esto, en base a su origen, destino (inamovibles) y el día. Para ello fue necesaria una representación de los datos cómoda y versátil. El grafo bipartito permitió modelar de buena manera la red de transporte, separando paraderos y servicios, y permitiendo modelar tiempos de espera y costos de viaje de manera clara y transparente al enrutador, no teniendo que agregar costes de transbordo de manera artificial.

El MNL permitió un análisis cuantitativo interpretable acerca de las variables que el autor de la memoria consideró importantes. Estos son, el tiempo de viaje, el coste restante de viaje al transbordar y el tiempo de espera. En este ámbito, se observó algo interesante. Las personas prefieren viajar más tiempo si eso significa minimizar el coste restante, la variable que más pesaba al seleccionar una alternativa. En otras palabras, las personas evitan hacer transbordos.

El GNN presenta mejores resultados para predecir, en parte gracias a sus *embeddings* y su correlación espacial, aunque no se usó en los experimentos, debido a su lentitud. De todas maneras, la precisión del MNL es suficientemente parecida (uno por ciento de diferencia) como para poder generalizar sin preocupaciones.

Los experimentos muestran tanto las fortalezas y las debilidades de la solución expuesta. Primero, no se observó una clara diferencia entre los coeficientes para distintos días. Un análisis de más semanas es necesario para ver una tendencia. Por otro lado, dos servicios compiten (tienen probabilidades parecidas) en general cuando tienen costes restantes parecidos. Servicios no compiten cuando uno tiene coste restante muy bajo y otro muy alto. Para que estos servicios

se vuelvan competidores entre si, el coste de espera del dominante debe de aumentar bastante para que valga la pena hacer el transbordo del otro servicio. Esto resume la preferencia para minimizar el coste restante y su dominancia sobre los atributos necesarios para una decisión.

La redistribución de demanda es difícil de analizar de manera local. Esto debido a la determinancia de las decisiones después de seleccionar el primer transbordo. Un análisis local sobre la redistribución revela como hay un aumento de demanda en servicios que combinan con los servicios elegidos.

De manera global, la redistribución de demanda en el caso de la Línea 7 mostró resultados no esperados. Por un lado, el aumento del uso del metro en general, un hallazgo interesante. Esto gracias a la captación de demanda de buses que hacían el recorrido largo. Esto se puede ver reflejado en el aumento de etapas en promedio. Además, se lograron observar los alimentadores y como se relacionan con la L7. Estos eran los buses que se esperaba que aumentaran su demanda al llevar a personas hacia la L7. El aumento de carga en la L1 es un hallazgo que probablemente se deba a artefactos en la solución. Se puede explicar por las restricciones del coste inicial de transferencia y como también las líneas del metro ahora son más llamativas para el usuario.

Las limitaciones de este trabajo pasan por las condiciones de borde de los viajes (origen y destino fijos), el coste de transbordo inicial y por las predicciones de bajada de ADATRAP. Estas decisiones, consecuencia de las condiciones de los datos y su extensibilidad, deben de ser tomadas en cuenta para interpretar los resultados. Es interesante como decisiones de diseño de la solución pueden condicionar de tal manera los resultados obtenidos y su alcance analítico.

Para finalizar, se concluye que se logró modelar la decisión de las personas, y estudiar la redistribución de la demanda causada por cambios en la oferta, sujeta a decisiones en la solución que pueden condicionar el análisis posible sobre estos resultados.

Para trabajo futuro, sería interesante desacoplar a los usuarios de los paraderos iniciales y finales, para establecer zonas o cuadrantes en donde tienen la libertad de elegir el paradero inicial al que ir. Esto solucionaría el problema encontrado en las etapas 1 de la L7 (que eran inexistentes) y se podría eliminar la simplificación del coste inicial.

Bibliografía

- [1] Cayul, L.H.C. 2017. *Desarrollo y aplicación de modelo de simulación basada en agentes a gran escala para la ciudad de santiago*. Universidad de Chile.
- [2] Dirección de Transporte Público Metropolitano 2024. Modelos de demanda. <https://dtpm.cl/index.php/documentos/modelos-de-demanda>.
- [3] Hamilton, W.L., Ying, R. and Leskovec, J. 2018. Inductive representation learning on large graphs.
- [4] Jiang, Q. 2022. GMM clustering based on WOA optimization and space-time coupled urban rail traffic flow prediction by CEEMD-SE-BiGRU-AM. *Mobile Information Systems*. 2022, 1 (2022), 7846630.
- [5] Li, W., Zhou, M., Dong, H., Wu, X. and Zhang, Q. 2021. Forecast of passenger flow of urban rail transit based on the DNNC model. *2021 33rd chinese control and decision conference (CCDC)* (2021), 4615–4620.
- [6] Li, Y., Zhang, J., Wang, J. and Wang, Y. 2022. Deep learning-based short-term traffic flow prediction considering spatial-temporal correlation. (2022). DOI:<https://doi.org/https://doi.org/10.1049/itr2.12018>.
- [7] Liu, L., Chen, J., Wu, H., Zhen, J., Li, G. and Lin, L. 2020. Physical-virtual collaboration modeling for intra-and inter-station metro ridership prediction. *IEEE Transactions on Intelligent Transportation Systems*. 23, 4 (2020), 3377–3391.
- [8] Massobrio, R. and Nesmachnow, S. 2020. Urban mobility data analysis for public transportation systems: A case study in montevideo, uruguay. *Applied Sciences*. 10, 16 (2020), 5400.
- [9] Morris, C., Ritzert, M., Fey, M., Hamilton, W.L., Lenssen, J.E., Rattan, G. and Grohe, M. 2021. Weisfeiler and leman go neural: Higher-order graph neural networks.
- [10] Ortúzar, J. de D. and Willumsen, L.G. 2011. *Modelling transport*. John Wiley & Sons.
- [11] Ramírez, Á.E.T. 2020. *Análisis espacial de los impactos en la demanda de transporte público producto de una nueva línea de metro utilizando datos masivos*. Universidad de Concepción.
- [12] SmartcitySantiagoChile 2025. ADATRAP: Herramienta para análisis de datos masivos de transporte público.
- [13] Soto, F.J.M. 2023. *Estimación y análisis de modelos de demanda agregada para el transporte público en santiago de chile*. Universidad de Chile.
- [14] Torrepadula Franca, R. di et al. 2024. Machine learning for public transportation demand prediction: A systematic literature review. *Engineering Applications of Artificial Intelligence*. (2024). DOI:<https://doi.org/https://doi.org/10.1016/j.engappai.2024.109166>.
- [15] Wang, Y., Yin, H., Chen, T., Liu, C., Wang, B., Wo, T. and Xu, J. 2021. Passenger mobility prediction via representation learning for dynamic directed and weighted graphs. *ACM Transactions on Intelligent Systems and Technology (TIST)*. 13, 1 (2021), 1–25.
- [16] Wei, J., Cheng, Y., Chen, K., Wang, M., Ma, C. and Hu, X. 2022. Nonlinear model-based subway station-level peak-hour ridership estimation approach in the context of peak deviation. (2022). DOI:<https://doi.org/https://doi.org/10.1177/03611981221075624>.
- [17] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Yu, P.S. 2021. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*. 32, 1 (2021), 4–24. DOI:<https://doi.org/10.1109/TNNLS.2020.2978386>.

- [18] Ye, J., Xu, Z. and Gou, X. 2022. An adaptive grey-markov model based on parameters self-optimization with application to passenger flow volume prediction. *Expert Systems with Applications*. 202, (2022), 117302.
- [19] Zhao, X., Guan, H., Sun, H. and Lu, J. 2022. A prophet-based passenger flow prediction model on IC card data. *2021 6th international conference on intelligent transportation engineering (ICITE 2021)* (Singapore, 2022), 1082–1092.