

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Posgrados

**Neural Network with Graph Convolutional Network and Long Short Term
Memory for Predicting Average Urban Traffic Speed**

Maestría en Ciencia de Datos

Paul Santiago Corella Medina

Israel Pineda Ph.D.

Director de Trabajo de Titulación

Trabajo de titulación de posgrado presentado como requisito para la obtención del título de
Magíster en Ciencia de Datos

Quito, 01 de diciembre de 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ
COLEGIO DE POSGRADOS

HOJA DE APROBACIÓN DE TRABAJO DE TITULACIÓN

**Neural Network with Graph Convolutional Network and Long Short Term
Memory for Predicting Average Urban Traffic Speed**

Paul Santiago Corella Medina

Nombre del Director del Programa: Felipe Grijalva
Título académico: Ph.D. en Ingeniería Eléctrica
Director del programa de: Ciencia de Datos

Nombre del Decano del colegio Académico: Eduardo Alba
Título académico: Doctor en Ciencias Matemáticas
Decano del Colegio: Ciencias e Ingenierías

Nombre del Decano del Colegio de Posgrados: Dario Niebieskikwiat
Título académico: Doctor en Física

Quito, diciembre 2024

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombre del estudiante: Paul Santiago Corella Medina

Código de estudiante: 00339073

C.I.: 0401748603

Lugar y fecha: Quito, 1 de diciembre de 2024.

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETheses>.

UNPUBLISHED DOCUMENT

Note: The following graduation project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETheses>.

DEDICATORIA

A Roosevelt Arévalo, quien con su generosidad y confianza hizo posible que este sueño se materializara. Su apoyo fue fundamental para alcanzar esta meta, y por ello siempre tendrá mi más profundo agradecimiento. A mis profesores de la maestría, por compartir su conocimiento, por su guía y por motivarme a superar mis límites. Sus enseñanzas no solo marcaron mi desarrollo académico, sino también mi crecimiento personal. A mis amigos de la maestría, compañeros de viaje en esta enriquecedora experiencia, por su apoyo incondicional, su amistad y los momentos compartidos que hicieron de este camino algo inolvidable. Y, por supuesto, a mi familia, mi pilar y fuente de fortaleza inagotable. Gracias por creer en mí, por su paciencia y amor, y por ser la razón de mis mayores esfuerzos y logros. A todos ustedes, esta tesis está dedicada con gratitud y cariño eterno

AGRADECIMIENTOS

Expreso mi profundo agradecimiento a Roosevelt Arévalo, cuyo auspicio hizo posible mi formación académica en esta maestría. Su apoyo fue esencial para el desarrollo de este trabajo de titulación. A los profesores Israel y Felipe, por su guía académica, su experiencia compartida y el rigor intelectual que contribuyeron significativamente al análisis y desarrollo de esta investigación. A mis compañeros de la maestría, cuya colaboración y apoyo enriquecieron esta experiencia y el proceso de aprendizaje. Finalmente, agradezco a mi familia, cuyo respaldo constante y motivación fueron la base para superar cada desafío de esta etapa académica. Reconozco a todas las personas e instituciones que, de manera directa o indirecta, contribuyeron en la generación, recolección o análisis de los datos que sustentan este trabajo. Su aporte fue invaluable para la culminación de este proyecto.

RESUMEN

La gestión del tráfico urbano enfrenta desafíos crecientes debido al aumento poblacional, la saturación vial y las deficiencias en la infraestructura. Este proyecto presenta un modelo híbrido que combina Redes Convolucionales de Grafos (GCN) y Memoria a Largo y Corto Plazo (LSTM) para predecir la velocidad promedio del tráfico en intersecciones urbanas, aprovechando las dependencias espaciales y temporales. El componente espacial (GCN) modela las relaciones entre intersecciones como un grafo, mientras que el componente temporal (LSTM) captura patrones históricos del tráfico. El modelo se evaluó utilizando un conjunto de datos del Departamento de Transporte de California, que incluye 30 intersecciones de la interestatal 5 con una resolución temporal de cinco minutos durante una semana (diciembre de 2023). Los datos se dividieron en entrenamiento (80%) y prueba (20%). Los resultados clave demuestran que el modelo supera los métodos tradicionales de predicción de tráfico, con limitaciones en la escalabilidad para redes viales más grandes y un alto costo computacional. El estudio utilizó Google Colab como entorno de prueba, procesando datos originales en un conjunto de datos filtrado de 28 intersecciones en la semana de 3 de diciembre de 2023 hasta el 7 de diciembre de 2023. El modelo híbrido integra datos espaciales y temporales de manera efectiva, ofreciendo una herramienta prometedora para mejorar la predicción y gestión del tráfico urbano. El conjunto de datos y el código están disponibles en un repositorio público de GitHub para garantizar la reproducibilidad.

Palabras clave: Gestión del Tráfico Urbano, Dependencias Espaciales, Dependencias Temporales, Modelo Híbrido, Redes Convolucionales de Grafos (GCN) y Memoria a Largo y Corto Plazo (LSTM)

ABSTRACT

Urban traffic management faces growing challenges due to population growth, road saturation, and infrastructure deficiencies. This project presents a hybrid model that combines Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM) to predict the average traffic speed at urban intersections, leveraging spatial and temporal dependencies. The spatial component (GCN) models relationships between intersections as a graph, while the temporal component (LSTM) captures historical traffic patterns. The model was evaluated using a dataset from the California Department of Transportation, including 30 intersections along Interstate 5 with a temporal resolution of five minutes over a one-week period (December 2023). The data were split into training (80%) and testing (20%). Key results demonstrate that the model outperforms traditional traffic prediction methods, with limitations in scalability for larger road networks and high computational costs. The study used Google Colab as the testing environment, processing raw data into a filtered dataset of 28 intersections during the week of December 3, 2023, to December 7, 2023. The hybrid model effectively integrates spatial and temporal data, offering a promising tool to improve urban traffic prediction and management. The dataset and code are available in a public GitHub repository to ensure reproducibility.

Key words: Urban Traffic Management, Spatial Dependencies, Temporal Dependencies, Hybrid Model, Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM)

TABLA DE CONTENIDO

I	Introducción	12
II	Trabajos Previos	13
	II-A Modelos basados en redes neuronales recurrentes (RNN)	13
	II-B Modelos basados en edes neuronales convolucionales (CNNs)	13
	II-C Modelos combinados (Spatio – Temporal models)	13
	II-D Modelos basados en redes de grafos (Graph Neural Networks GNNs)	13
	II-E Modelos basados en series temporales multivariables)	14
	II-F Modelos de transformadores espacio – temporales	14
III	Descripción del problema	14
	III-A Dependencia espacial	14
	III-B Dependencia temporal	14
IV	Metodología Propuesta	14
	IV-A Descripción del componente espacial (GCN)	15
	IV-B Descripción del componente temporal (LSTM Units)	16
	IV-C Descripción del modelo combinado GCN+LSTM	16
	IV-D Descripción del método de evaluación del modelo	16
V	Experimentación	17
	V-A Descripción de los datos disponibles	17
	V-B Selección de los parámetros del modelo	17
	V-C Definición del set de entrenamiento y prueba	18
VI	Resultados	18
VII	Conclusiones	18
	References	18

ÍNDICE DE TABLAS

I	Información de las carreteras analizadas.	18
---	---	----

ÍNDICE DE FIGURAS

1	Sensores IoT para redes viales [1].	13
2	Diagrama estructural de red neuronal LSTM [2].	13
3	Diagrama estructural de red neuronal STGCN [3].	14
4	Diagrama estructural de red neuronal espacio temporal con el uso de attention transformers [4].	14
5	Componente GCN en aplicado a tres nodos	15
6	Componente LSTM en aplicado a una ventana de cuatro unidades de tiempo	16
7	Arquitectura GCN + LTSM del modelo propuesto para tres nodos con una ventana de tiempo de cuatro unidades	17
8	Comparativa entre valores reales y valores predichos para la intersección “715947” (S OF 710) en el set de entrenamiento	19
9	Comparativa entre valores reales y valores predichos para la intersección “715947” (S OF 710) en el set de prueba	19
10	Comparativa entre valores reales y valores predichos para la intersección “716930” (FERRIS) en el set de entrenamiento	19
11	Comparativa entre valores reales y valores predichos para la intersección “716930” (FERRIS) en el set de prueba	20
12	Comparativa entre valores reales y valores predichos para la intersección “718085” (TRIGGS) en el set de entrenamiento	20
13	Comparativa entre valores reales y valores predichos para la intersección “718085” (TRIGGS) en el set de prueba	20

Neural Network with Graph Convolutional Network and Long Short Term Memory for Predicting Average Urban Traffic Speed

Paul Corella, *Universidad San Francisco de Quito*, Israel Pineda, *Universidad San Francisco de Quito*,

Abstract—Urban traffic management faces significant challenges due to population growth, road saturation, and infrastructure deficiencies. This study presents a hybrid model combining Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM) to predict average traffic speed at urban intersections by leveraging spatial and temporal dependencies. Evaluated on a dataset from the California Department of Transportation, including 30 intersections along Interstate 5 with five-minute temporal resolution over one week in December 2023, the model demonstrated superior performance compared to traditional methods, despite scalability and computational cost limitations. Conducted in Google Colab, the study processed raw data into a filtered dataset of 28 intersections, effectively integrating spatial and temporal data for improved traffic prediction. The dataset and code are publicly available to ensure reproducibility.

Index Terms—Urban Traffic Management, Spatial Dependencies, Temporal Dependencies, Hybrid Model, Graph Convolutional Networks (GCN) and Long Short-Term Memory (LSTM).

I. INTRODUCCIÓN

EL tráfico vehicular representa un desafío significativo para las redes viales urbanas. Principalmente, la problemática se evidencia en velocidades reducidas, largos tiempos de espera de los usuarios en los atascos, etc. Este fenómeno es producto de una saturación vial y sobre todo de una mala gestión de los embotellamientos.

Dentro de las principales causas que intensifican la congestión vehicular se puede destacar el crecimiento demográfico conforme lo que se indica en estudio realizado por Ian Thomson y Alberto Bull para la CEPAL[5]. Inherentemente un aumento de la población implica un crecimiento en el parque automotor que sobrecarga las redes viales. Por otro lado, también se suma al problema una infraestructura vial pobre con redes viales ineficientes que no permiten la gestión adecuada de un gran volumen de vehículos [6].

Por otro lado, existen grandes metrópolis que se ven afectadas por las características y limitaciones geográficas. Actualmente, se cuenta con pocos estudios que se enfoquen en la simulación del comportamiento del tráfico [7]. Hoy en día, la gestión y planificación del tráfico tiene que

hacer frente a desafíos como la dependencia espacial y la dependencia temporal del flujo de autos. Por esta razón, todos los modelos que se diseñen para brindar una solución a este problema deben tomar en cuenta esta relación espacial temporal.

Para aquello relacionado a la relación espacial, se indica que esta se refiere a la influencia entre las intersecciones centrales con sus similares adyacentes. Principalmente, se enfoca en describir la forma en la cual una acción suscitada en un cruce vial puede afectar en cadena a los demás. Por ejemplo, algunos estudios detallan la relación dinámica entre intersecciones donde los semáforos buscan optimizar en conjunto la eficiencia de la red vial [8]. En estos estudios se indica que una apropiada toma de decisiones permite optimizar el flujo entre intersecciones. Con esto en mente, es crucial que el modelo que se proponga como solución a la gestión del tráfico capture adecuadamente el dinamismo y la dependencia entre intersecciones para de ese modo establecer un patrón que describa correctamente el comportamiento del sistema.

Por otro lado, la dependencia temporal, según se menciona en trabajo desarrollado por Octavio Salcedo, César Augusto Hernández y Andrés Escobar, corresponde a las condiciones actuales del tráfico vehicular que se ven influenciadas por los estados previos [9]. La evolución temporal en las comúnmente conocidas horas pico, mismas que corresponden a franjas horarias que se caracterizan por la acumulación de vehículos, son el producto de la acumulación de autos en un cruce vial con el paso del tiempo. Entonces, los retrasos por la congestión vehicular son consecuencia de interacciones dinámicas de estados previos del sistema. Un claro ejemplo de esta dependencia es una intersección que a medida que pasa el tiempo va cargándose de más autos. Por lo tanto, los modelos predictivos que se propongan deben ser capaces de identificar los patrones a lo largo del tiempo utilizando datos históricos.

Para abordar una solución apropiada de la gestión del tráfico, es fundamental que el modelo predictivo se diseñe con base en el aprendizaje de las interacciones profundas espacio-temporales del sistema. El modelo debe ser capaz de representar el patrón de comportamiento y los factores que influyen en este. Estudios como el desarrollado por Nautiyal y Bathla sugieren que se integren sensores IoT como los que se muestran en la Figura 1 [10]. El propósito de estos

Paul Corella and Israel Pineda was with Colegio de Ciencias e Ingenierías, Universidad San Francisco de Quito, Quito, e-mail: (pcorella@estud.usfq.edu.ec and ipineda@usfq.edu.ec

dispositivos es proveer datos históricos para determinar patrones más precisos. Los sistemas de transporte inteligentes proporcionan un enfoque prometedor que permite lidiar con la problemática de las dependencias.



Figure 1. Sensores IoT para redes viales [1].

En resumen, la gestión del tráfico requiere de modelos, tecnologías y datos que en conjunto permiten tener una comprensión profunda de los patrones tomando en cuenta la relación entre los cruces viales y el comportamiento a lo largo del tiempo.

II. TRABAJOS PREVIOS

Los modelos de predicción de tráfico que actualmente existen permiten identificar los complejos patrones de flujo vehicular en las redes viales urbanas. El principal objetivo de estos modelos es capturar el comportamiento dinámico de la infraestructura vial incluyendo las variaciones temporales y las interacciones espaciales. A continuación, se muestran los modelos existentes:

A. Modelos basados en redes neuronales recurrentes (RNN)

Este tipo de redes neuronales, con su variante Long-Short Term Memory (LSTM), son excepcionales y ampliamente usadas para modelar los fenómenos temporales y sus dependencias. Este tipo de redes neuronales son capaces de procesar secuencias de datos históricos que permiten predecir un patrón de comportamiento partiendo de la periodicidad de los datos. Particularmente, el uso de LSTMs, como se muestra en la Figura 2, ha resultado ser eficiente al manejar datos históricos del tráfico que son altamente variables y no dinámicos. En el estudio desarrollado para ZTE Communications se emplea un modelo desarrollado a partir de esta arquitectura de redes recurrentes del tipo LSTM. Este modelo permitió que pueda capturar las relaciones de los datos históricos y luego proporcionar un escenario en el futuro [2]. Se destaca que este modelo propuesto es particularmente efectivo ya que permite identificar los patrones temporales a corto y largo plazo. La conclusión del estudio enfatiza en resaltar que la dependencia temporal es crucial para modelar sistemas dinámicos temporales como el tráfico vehicular.

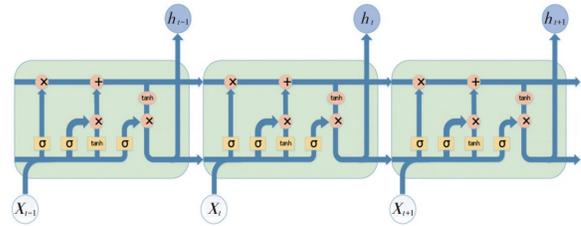


Figure 2. Diagrama estructural de red neuronal LSTM [2].

B. Modelos basados en redes neuronales convolucionales (CNNs)

Las redes neuronales convolucionales fueron diseñadas en principio para resolver problemas de visión computacional; sin embargo, estas han sido adaptadas para capturar un patrón que represente el comportamiento del tráfico en la red vial. Su principio de funcionamiento aplicado para el propósito de estudio del tráfico vehicular pretende tratar a la red como un gráfico, esto a su vez permite modelar las condiciones de ubicación de un vehículo y trazar su influencia en las zonas adyacentes. Un estudio desarrollado por Gonzalvez en el 2017 utiliza redes CNN que permiten analizar datos espaciales del tráfico evidenciando la interacción entre los segmentos adyacentes a través de relaciones convolucionales [11]. En este trabajo se procesa una “Imagen del tráfico” para detectar un patrón que describe el comportamiento del sistema. La red propuesta mostró tener una alta efectividad para predecir condiciones críticas de tráfico.

C. Modelos combinados (Spatio – Temporal models)

Estos modelos son el resultado de combinar los dos modelos descritos anteriormente. Su principio de funcionamiento se basa en combinar las CNNs para capturar las dependencias espaciales y las RNNs para capturar la parte temporal. Estas arquitecturas combinadas se caracterizan por integrar tanto las variaciones de espacio y tiempo. En el 2017 un estudio desarrollado propuso un modelo de red híbrido que combina CNNs y LSTMs [12]. Este enfoque resultó, según sus experimentos, ser sumamente efectivo para la predicción del tráfico de una red vial grande y a su vez se destaca que ya toma en cuenta los dos factores de dependencias el espacio y el tiempo.

D. Modelos basados en redes de grafos (Graph Neural Networks GNNs)

Estas arquitecturas de redes neuronales, las mismas que se presentan en la Figura 3, son especialmente óptimas para modelar la infraestructura vial puesto que son capaces de representar de forma apropiada la red vial a través de intersecciones (nodos) y de carreteras (aristas). La ventaja del uso de estas redes es que incorporan información estática e información dinámica. El modelo desarrollado por Yu y Yin utiliza una red de grafos con una capa recurrente [3]. Con este estudio se demostró que el uso de GNNs es una técnica robusta para la predicción de configuraciones complejas de redes viales.

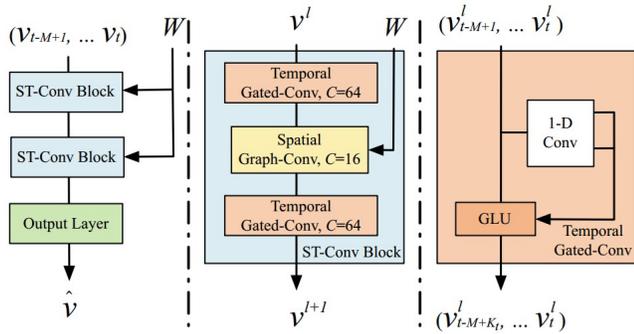


Figure 3. Diagrama estructural de red neuronal STGCN [3].

E. Modelos basados en series temporales multivariantes)

Los modelos autorregresivos como ARIMA han demostrado ser herramientas eficaces para modelar series temporales debido a su capacidad para capturar patrones históricos y tendencias. Sin embargo, en el contexto del tráfico vehicular, donde las interacciones entre segmentos viales son fundamentales, los modelos tradicionales presentan limitaciones al enfocarse en datos univariados. Para abordar esta complejidad, se han desarrollado variantes híbridas que combinan ARIMA con información espacial, permitiendo considerar las interdependencias entre diferentes tramos de una red vial. El enfoque híbrido propuesto por Vanajakshi y Kumar integra datos espaciales, como densidad vehicular o flujo entre segmentos conectados, dentro de la estructura autorregresiva de ARIMA [13]. Este modelo se basa en una serie de pasos: preprocesamiento de datos recolectados de sensores, ajuste de modelos ARIMA para capturar patrones temporales en segmentos específicos y la incorporación de matrices espaciales que reflejan la influencia entre carreteras adyacentes. Esto mejora significativamente la precisión en predicciones, facilitando una mejor gestión del tráfico y planificación de rutas.

F. Modelos de transformadores espacio – temporales

Este tipo de modelos se inspira en el uso de arquitecturas conocidas como Transformers comúnmente utilizadas en el procesamiento de lenguaje natural. Su principal característica en este campo es que permiten incorporar y manejar de manera adecuada la dependencia a lo largo del tiempo. En 2024, se propuso un modelo basado en Transformers Espacio-Temporales para la predicción del tráfico vehicular. La ventaja de este estudio radica en que su enfoque permite capturar la relación espacial y temporal con una alta resolución y brindando predicciones más precisas [4]. La arquitectura del modelo propuesto se muestra en la Figura 4.

III. DESCRIPCIÓN DEL PROBLEMA

Predecir el comportamiento del tráfico vehicular a través de una modelo de red neuronal es una solución adecuada al crítico problema de la gestión del transporte inteligente. Principalmente el objetivo del modelo híbrido propuesto es

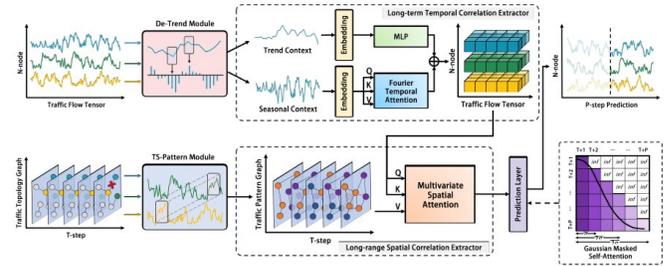


Figure 4. Diagrama estructural de red neuronal espacio temporal con el uso de attention transformers [4].

predecir el estado futuro de alguna característica del tráfico, en este caso la velocidad promedio en cada intersección a través del uso de datos históricos. El mayor reto al que se enfrenta este modelo es la naturaleza no lineal de los datos, así como su dependencia temporal y espacial. El enfoque combinado de las Redes Neuronales Graph Convolutional Network (GCN) y las Redes Neuronales Recurrentes Long Short Term Memory (LSTM) abordan este desafío ya que cada componente se encarga de integrar las dependencias espaciales y temporales respectivamente.

A. Dependencia espacial

El tráfico en una intersección está siempre bajo la influencia de lo que sucede a su alrededor en los otros nodos. Por ejemplo, si un alto flujo vehicular se presenta en una intersección, las rutas alternativas aledañas van a ser directamente afectadas. Para representar este fenómeno, se utiliza el componente GCN, el cual permite modelar esta relación entre intersecciones de la red vial a través de un grafo donde cada nodo representa una intersección y cada arista es la carretera que une dos o más nodos.

B. Dependencia temporal

El patrón del tráfico varía a lo largo del tiempo con fluctuaciones en el corto plazo como son patrones horarios hasta fluctuaciones a largo plazo como las tendencias diarias o semanales. Las redes propuestas LSTM son eficaces para capturar las series de tiempo con dependencias temporales de larga duración.

IV. METODOLOGÍA PROPUESTA

El presente estudio propone la combinación de dos modelos, las Redes Convolucionales de Grafos (GCNs) y las Redes Neuronales Recurrentes Long Short Term Memory (LSTMs) de manera similar al estudio desarrollado por Ling Zhao el cual combina el componente GCN y Gated Recurrent Unit (GRU) [14]. El propósito de esta combinación de modelos para formar un único modelo híbrido es diseñar una arquitectura de red neuronal que sea capaz de capturar las relaciones espaciales y temporales para proporcionar una predicción precisa y acorde al comportamiento de cada intersección de una sola característica. Por ejemplo, la velocidad promedio de una intersección. El uso combinado de GCN+LSTM es en principio una solución prometedora

para brindar un soporte técnico en la predicción del tráfico vehicular. Sin embargo, es crucial que se tome en cuenta que el modelo debe afrontar retos importantes como la normalización de los datos, la limitación de las unidades de procesamiento, y la complejidad del modelo híbrido.

A. Descripción del componente espacial (GCN)

En este estudio se utilizan las Redes Neuronales Convolucionales (GCN) como un mecanismo que permite compartir la información entre los nodos. Su principio de funcionamiento se basa en extender la idea de la convolución al dominio de los grafos. Si bien la convolución es más conocida en el análisis de imágenes, también se aplica para el procesamiento de datos estructurados como grafos mediante la interacción entre sus nodos y sus aristas. El objetivo de una estructura GCN es generar representaciones embebidas de los nodos. Esto a su vez permite capturar sus características individuales y también las características de la estructura del grafo. El concepto de convolución se refiere a el proceso que agrega información de los nodos vecinos para actualizar la información del nodo central; de este modo se comparte la información a lo largo de toda la estructura. Los módulos clave de una GCN corresponden a:

Matriz de características Se representa con X donde N es el número de nodos y F es la cantidad de características por nodo. Matemáticamente se expresa como:

Una matriz de características $X \in \mathbb{R}^{N \times F}$:

- N : Número de nodos.
- F : Número de características por nodo.

Matriz de adyacencia Se representa con A donde N es el número de nodos. Esta matriz facilita representar la conexión entre nodos. Principalmente, se compone valores de 1 para indicar que efectivamente existe una conexión entre dos intersecciones y de 0 cuando no hay una carretera que una las intersecciones en cuestión. Matemáticamente se expresa como:

$A[i, j] = 1$ si hay una arista entre el nodo i y el nodo j , y 0 en caso contrario.

Una vez definida la matriz de características y la matriz de adyacencia como entradas del modelo, el esquema de funcionamiento de la GCN es el siguiente:

El primer paso corresponde a la normalización de la matriz de adyacencia. Para ello se hace uso de la matriz laplaciana. Matemáticamente se expresa de la siguiente manera: La matriz \tilde{A} se define como:

$$\tilde{A} = D^{-1/2}(A + I)D^{-1/2}$$

Donde:

- A es la matriz de adyacencia.
- I es la matriz identidad que permite incluir los self-loops en cada nodo y, de esa manera, preservar la información del nodo central.

- D es la matriz diagonal de grados, de dimensión $N \times N$, que permite normalizar la matriz de adyacencia. Matemáticamente se representa:

$$D[i, i] = \sum_j A[i, j].$$

Para la propagación de características, la información de los nodos se actualiza mediante la siguiente expresión matemática:

$$H^{(l+1)} = \sigma(\tilde{A}H^{(l)}W^{(l)})$$

Donde:

- $H^{(l)}$ corresponde a la matriz de características en la capa l . Esta matriz tiene la misma dimensión que la matriz de características X , que contiene los datos de entrada para el modelo.
- $W^{(l)}$ corresponde a la matriz de pesos con dimensión $N \times N$.
- σ corresponde a la función de activación del modelo. Esta puede ser una función sigmoide, una función ReLU o cualquier otra función de activación.
- \tilde{A} corresponde a la matriz de adyacencia normalizada, de dimensiones $N \times N$.

Cabe destacar que se pueden agregar tantas capas convolucionales como se considere apropiado. Cada capa puede tener su propia función de activación. Para el presente estudio se utilizan dos capas GCN que se representan a través de la siguiente ecuación:

$$f(X, A) = \sigma(\hat{A} \text{ReLU}(\hat{A}XW_0)W_1)$$

En resumen, este componente del modelo es el encargado de aprender las dependencias espaciales entre los nodos. Su principal función es combinar las características individuales de cada nodo para aprender su representación en toda la estructura. En la Figura 5 se muestra una representación de este componente aplicado para tres nodos. En esencia lo que se busca con esto es que cada nodo comparta su información con sus nodos adyacentes y viceversa. Este componente se aplica y se repite para cada uno de los vectores en cada tiempo t .

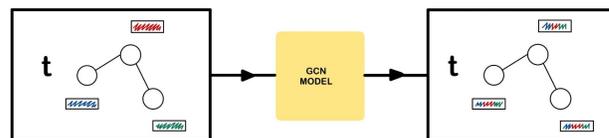


Figure 5. Componente GCN en aplicado a tres nodos

B. Descripción del componente temporal (LSTM Units)

Para este componente se utiliza el modelo de Red Neuronal Recurrente LSTM. La unidad LSTM es un tipo de célula diseñada para identificar patrones a lo largo del tiempo que contengan secuencias de datos, evitando problemas comunes de las Redes Neuronales Recurrentes (RNNs) como el desvanecimiento del gradiente o la explosión del gradiente.

La estructura de una unidad LSTM incluye dos estados y tres compuertas que en conjunto permiten controlar el apropiado flujo de información. A continuación, se describe cada uno de ellos:

- **Estado de celda c_t :** Corresponde a la memoria a largo plazo que transporta información a través de cada iteración. Este valor se actualiza con cada paso de tiempo a través de las compuertas.
- **Estado oculto h_t :** Corresponde a la salida de la unidad LSTM. El valor de este estado contiene la información inmediata y relevante del paso en el tiempo t .
- **Compuerta de entrada:** Esta compuerta se encarga de determinar qué porcentaje de nueva información se almacenará en el estado de celda. Su comportamiento se describe mediante la ecuación:

$$i_t = \sigma(W_i \cdot \mathbf{x}_t + U_i \cdot \mathbf{h}_{t-1} + b_i)$$

- **Compuerta de olvido:** Esta compuerta determina qué porcentaje de información debe ser descartada. Su principio de funcionamiento se describe como:

$$f_t = \sigma(W_f \cdot \mathbf{x}_t + U_f \cdot \mathbf{h}_{t-1} + b_f)$$

- **Compuerta de salida:** Esta compuerta establece cuánta información del estado de celda c_t influye en el estado oculto h_t . Su relación matemática es:

$$o_t = \sigma(W_o \cdot \mathbf{x}_t + U_o \cdot \mathbf{h}_{t-1} + b_o)$$

El estado de celda c_t y el estado oculto h_t se actualizan en cada paso de tiempo según las siguientes ecuaciones:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c \cdot \mathbf{x}_t + U_c \cdot \mathbf{h}_{t-1} + b_c)$$

$$h_t = o_t \odot \tanh(c_t)$$

Donde:

- \mathbf{x}_t : Entrada en el paso de tiempo t .
- W, U, b : Son matrices de pesos y sesgos correspondientes a las compuertas.
- σ : Función sigmoide.
- \odot : Producto elemento a elemento.
- \tanh : Función tangente hiperbólica.

En resumen, la principal característica de la unidad LSTM son la persistencia de la memoria, es decir que el uso de las compuertas permite determinar la información que se preserva y la información que se descarta. Partiendo desde lo más básico una LSTM es una Red Neuronal con

compuertas que permite memorizar información a largo plazo. Para conformar una red LSTM se requiere combinar en serie cuantas unidades LSTM se desee analizar. En este estudio en específico se trabaja con una ventana de cuatro unidades hacia el pasado que permiten predecir el valor presente. En la Figura 6 se muestra como la unión secuencial de cuatro unidades LSTM permiten predecir el quinto valor.

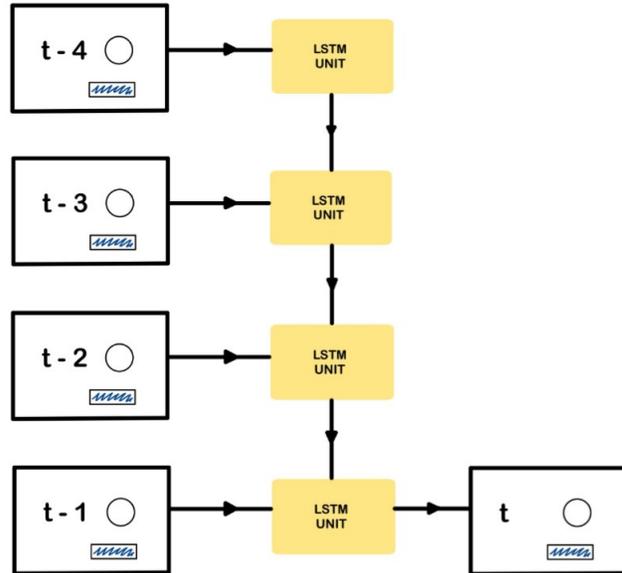


Figure 6. Componente LSTM en aplicado a una ventana de cuatro unidades de tiempo

C. Descripción del modelo combinado GCN+LSTM

En esencia el modelo combinado utiliza el componente GCN para compartir la información entre nodos en cada uno de los pasos de tiempo. Esto garantiza que exista la dependencia espacial entre los nodos. Como salida del componente GCN se tiene vectores de características con información compartida entre toda la infraestructura. Luego se encadenan tantas unidades LSTM como se deseen. Para este caso de estudio en específico se utilizan cuatro unidades LSTM que representan una ventana de cuatro unidades de tiempo (t-4, t-3, t-2, t-1). El estado oculto de la cuarta unidad LSTM es la salida de la cadena y corresponde al valor predicho (t). Este proceso se replica en cada uno de los nodos.

En la Figura 7 se representa la arquitectura completa del modelo híbrido.

D. Descripción del método de evaluación del modelo

El modelo híbrido GCN+LSTM se evalúa a través de la función de pérdida (Loss). Esta propuesta consiste en un modelo de aprendizaje supervisado donde las etiquetas son las mediciones en el tiempo t . Se utiliza el método

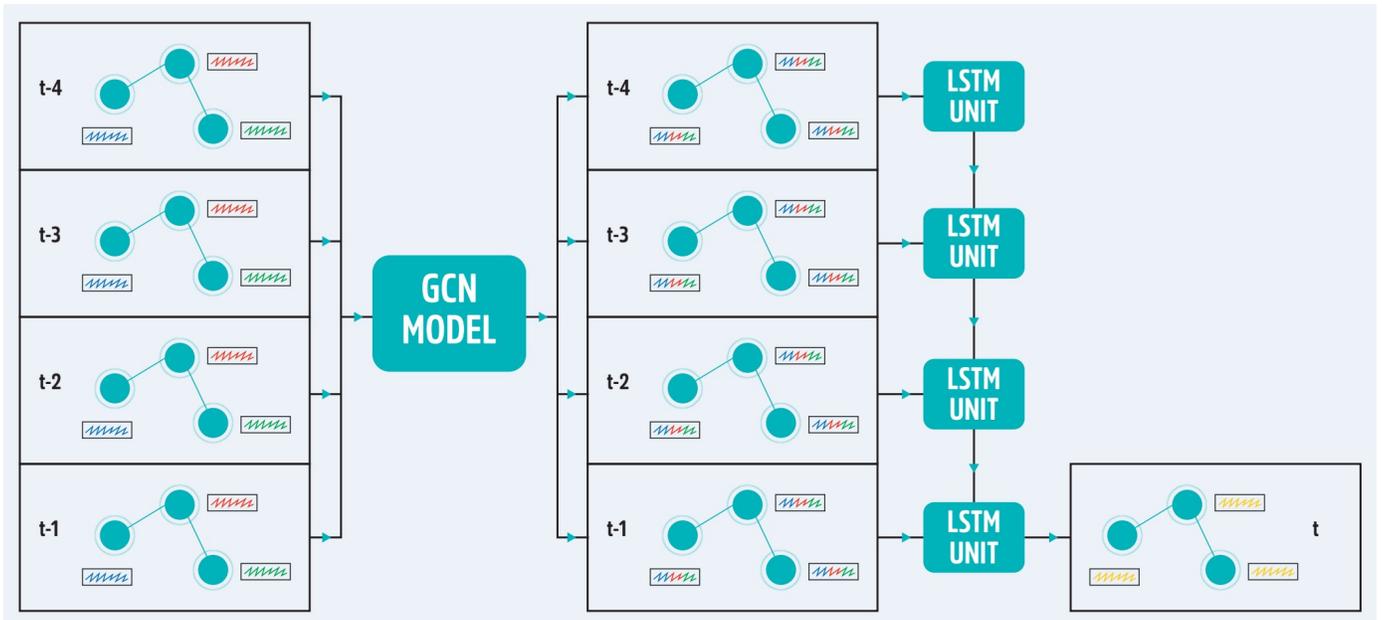


Figure 7. Arquitectura GCN + LSTM del modelo propuesto para tres nodos con una ventana de tiempo de cuatro unidades

Mean Squared Error (MSE) puesto que, al tratarse de una comparación entre dos puntos, el valor real y el valor predicho, resulta mejor obtener la diferencia entre ellos y minimizarla al máximo. Entonces, matemáticamente la función de pérdida se expresa a través de la siguiente ecuación:

$$MSE = \frac{1}{MN} \sum_{j=1}^M \sum_{i=1}^N (y_i^j - \hat{y}_i^j)^2$$

Donde:

- y_i^j representa la medición real de la característica del tráfico en el tiempo t .
- \hat{y}_i^j representa la predicción de la característica del tráfico en el tiempo t .
- M es el número de series de tiempo analizadas por cada lote, para este caso $M = 16$.
- N es el número de nodos en cada lote, para este caso $N = 28$.

V. EXPERIMENTACIÓN

Los datos y resultados de los experimentos realizados en el presente trabajo pueden ser encontrados en el repositorio de GitHub con el siguiente enlace https://github.com/PauIsantiagooc/Modelo_GCNLSTM

A. Descripción de los datos disponibles

Para este estudio en específico se cuenta con datos provistos por el Departamento de Transporte de California. Los datos corresponden a 30 intersecciones de la interestatal 5 de California (01A: Los Angeles I-5 – Primary Direction) y tienen información provista por sensores con una resolución

de cinco minutos entre mediciones sobre características como el Flujo Vehicular (Flow), el nivel de ocupación vehicular (Occupancy) y la velocidad promedio vehicular (Speed). El set de datos con el cual se trabaja contiene la información con las características descritas anteriormente de todo el mes de diciembre del 2023. No obstante, para el caso de estudio del presente documento se utiliza la primera semana con días ordinarios de todo el conjunto de datos, es decir, los datos que corresponden del 3 de diciembre de 2023 hasta el 7 de diciembre de 2023. Dado que el modelo propuesto se enfoca en capturar los patrones de la dinámica vehicular de una sola característica como lo es la velocidad promedio en cada intersección, el set de datos se enfoca en estos datos. Para el rango de tiempo mencionado, se tiene que los datos de la intersección “716896” (PHOEBE) y la intersección “762347” (ARTESIA) no cuentan con datos por lo cual estas son excluidas del análisis. El resto de los datos no presenta ninguna inconsistencia. Finalmente cabe destacar que las intersecciones se encuentran unidas de manera consecutiva y cada una cuenta con una medición desde las 0h00 del 3 de diciembre de 2023 hasta las 23h55 del 7 de diciembre de 2023 con una resolución de cinco minutos entre medición. En la Tabla 1 se muestran las intersecciones con las cuales en el presente estudio se realiza la investigación.

B. Selección de los parámetros del modelo

En esta sección se describen los parámetros relevantes del modelo híbrido GCN+LSTM. Para este caso de estudio se utiliza una tasa de aprendizaje (learning rate) de 0.1, un tamaño del lote (batch size) de 16 unidades y una cantidad de épocas (epoch) de 500. Dada la naturaleza del modelo, se tiene que este utiliza dos capas GCN y cuatro unidades LSTM como se indica en el apartado de la metodología propuesta.

Table I
INFORMACIÓN DE LAS CARRETERAS ANALIZADAS.

ID	Name	Lat	Long
715947	S OF 710	34.015325	-118.17127
716930	FERRIS	34.013529	-118.166192
718085	TRIGGS	34.011229	-118.161247
716928	ATLANTIC	34.007592	-118.157266
718364	GASPAR	34.002455	-118.151164
716924	WASHINGTON 1	33.994935	-118.14469
763980	MALT	33.992211	-118.142079
716922	GARFIELD	33.986224	-118.136014
768523	GREENWOOD	33.981723	-118.130845
716920	SLAUSON	33.97646	-118.125953
715929	GUATEMALA	33.971707	-118.123095
716895	PARAMOUNT	33.963867	-118.11987
716918	LAKEWOOD 2	33.958225	-118.11239
716916	LAKEWOOD 1	33.956897	-118.110532
763990	GARNISH	33.952949	-118.105163
715920	S OF 605	33.938544	-118.094941
715916	TINA	33.924281	-118.084952
716912	PIONEER	33.920763	-118.082501
716911	IMPERIAL	33.916643	-118.079557
716908	SAN ANTONIO/NORWALK	33.911074	-118.071686
759610	SILVER BOW	33.907584	-118.067796
716906	FIRESTONE	33.903961	-118.064046
716907	ROSECRANS	33.900669	-118.059422
763706	FIDEL	33.896666	-118.052294
716902	CARMENITA	33.892489	-118.044573
763748	SPRING	33.89061	-118.041088
716898	VALLEY VIEW	33.882892	-118.026822
716896	PHOEBE	33.879952	-118.021355
762398	OSMOND	33.876316	-118.014605
762347	ARTESIA	33.875077	-118.012367

C. Definición del set de entrenamiento y prueba

Con base en la descripción de los datos, se separa el set de datos de la siguiente forma, 80% de los datos corresponde al set de entrenamiento (Training set) y el 20% restante corresponde al set de prueba (Testing Set). Esto resulta que los datos que abarcan desde las 0h00 del 3 de diciembre de 2023 hasta las 23h55 del 6 de diciembre del 2023 corresponden al set de entrenamiento (Training Set) mientras que todas la mediciones desde las 0h00 hasta las 23h55 del 7 de diciembre del 2023 corresponden al set de prueba (Testing set)

VI. RESULTADOS

A continuación, se presentan los resultados de tres intersecciones. La Figura 8 corresponde a la comparación de los valores reales con los valores predichos en el set de entrenamiento y la Figura 9 muestra la comparación de los valores reales con los valores reales y predichos de la intersección “715947” (S OF 710). La misma representación se muestra en la Figura 10 y la Figura 11 para la intersección “716930” (FERRIS). Finalmente, la Figura 12 y la Figura 13 contienen la comparación de la intersección “718085” (TRIGGS). El resto de las figuras con esta comparativa se muestran en el repositorio de GitHub que forma parte integral de este estudio. El enlace del repositorio es el siguiente https://github.com/Paulsantiagooc/Modelo_GCNLSTM.

VII. CONCLUSIONES

El modelo GCN+LSTM propuesto resulta ser una herramienta útil para capturar la dependencia espacial y temporal de una red vial. El componente GCN efectivamente captura la conectividad entre las intersecciones, mientras que el componente LSTM captura la dinámica temporal que permite identificar los patrones de comportamiento en cada una de las intersecciones con el paso del tiempo.

El modelo propuesto tiene un mejor rendimiento que los modelos tradicionales puesto que permite integrar tanto dimensiones temporales como espaciales. Otros modelos abordan este tema exclusivamente desde una de las dos perspectivas.

Este modelo presenta en esencia un alto grado de complejidad. Si bien el modelo es escalable puesto que puede adaptarse fácilmente a las distintas arquitecturas de la red vial, su complejidad aumenta significativamente a medida que se aumentan las intersecciones. Esto deriva en un gran consumo de recursos de cómputo.

La calidad de las predicciones depende de los datos con los que se cuenta. Este modelo, por ejemplo, solo fue entrenado con cuatro días y 28 intersecciones. Si bien el modelo es capaz de adaptarse adecuadamente a datos de las mismas intersecciones, es robusto únicamente ante cambios pequeños de los patrones. En caso que los cambios sean grandes el modelo puede requerir reentrenamiento.

REFERENCES

- [1] ITS Answers, “Adaptive signal system,” 2024, accessed: 2024-11-25. [Online]. Available: <https://www.itsanswers.com/project/adaptive-signal-system/>
- [2] W. Shihao, Z. Qinzheng, Y. Han, L. Qianmu, and Q. Yong, “A network traffic prediction method based on lstm,” *ZTE Communications*, vol. 17, no. 2, pp. 19–29, 2019, published online June 19, 2019. [Online]. Available: <http://kns.cnki.net/kcms/detail/34.1294.TN.20190619.0902.002.html>
- [3] B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*. Stockholm, Sweden: International Joint Conferences on Artificial Intelligence, 2018, pp. 3634–3640. [Online]. Available: <https://www.ijcai.org/proceedings/2018/0505.pdf>
- [4] Q. Yu, W. Ding, H. Zhang, Y. Yang, and T. Zhang, “Rethinking attention mechanism for spatio-temporal modeling: A decoupling perspective in traffic flow prediction,” in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, ser. CIKM ’24. New York, NY, USA: Association for Computing Machinery, 2024, p. 3032–3041. [Online]. Available: <https://doi.org/10.1145/3627673.3679571>
- [5] I. Thomson and A. Bull, “La congestión del tránsito urbano: causas y consecuencias económicas y sociales,” CEPAL - División de Recursos Naturales e Infraestructura, Santiago de Chile, Tech. Rep. 25, 2001, n° de venta: S.01.II.G.105.
- [6] F. Bedoya-Maya, A. Calatayud, and V. G. Mejia, “Estimating the effect of urban road congestion on air quality in latin america,” Inter-American Development Bank, Tech. Rep., 2022.
- [7] J. P. Orellana, I. Pineda, and R. Armas, “Distributed architecture for large scale simulation to estimate co vehicles emissions,” in *2022 IEEE Sixth Ecuador Technical Chapters Meeting (ETCM)*, 2022, pp. 1–6.

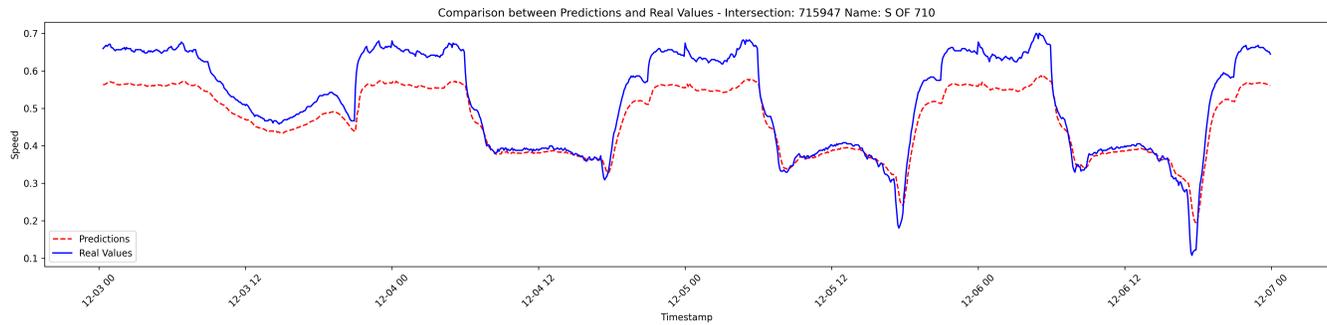


Figure 8. Comparativa entre valores reales y valores predichos para la intersección “715947” (S OF 710) en el set de entrenamiento

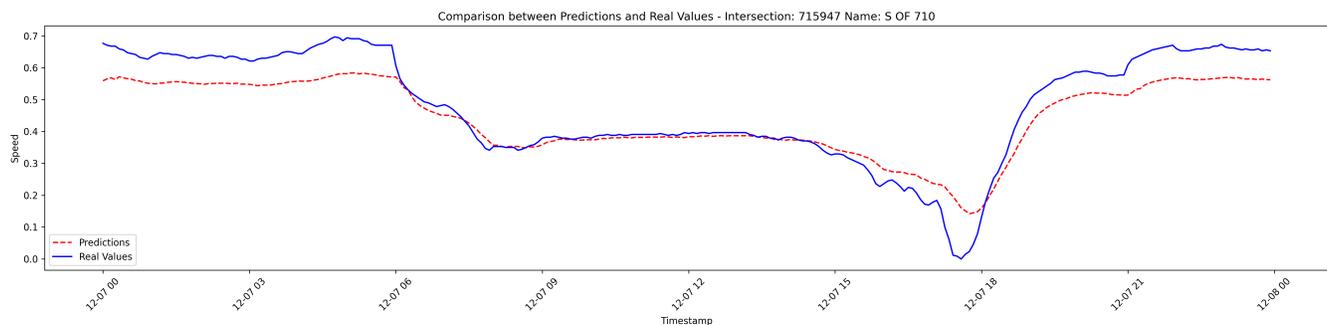


Figure 9. Comparativa entre valores reales y valores predichos para la intersección “715947” (S OF 710) en el set de prueba

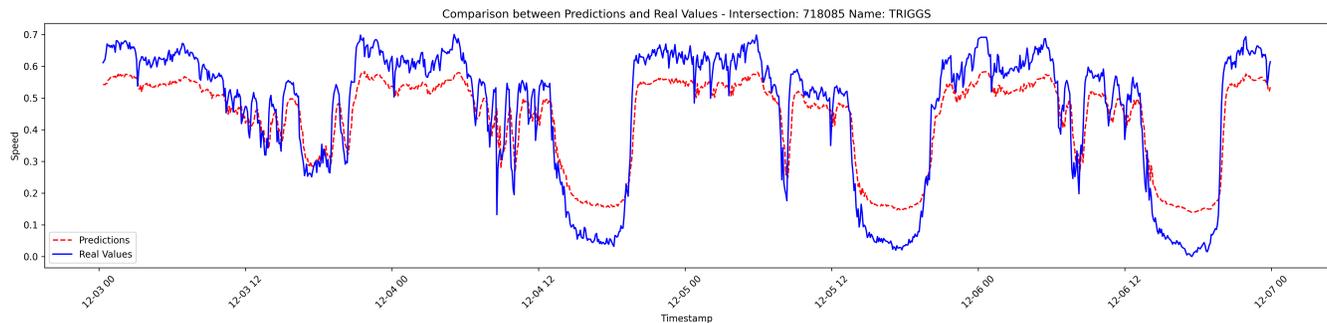


Figure 10. Comparativa entre valores reales y valores predichos para la intersección “716930” (FERRIS) en el set de entrenamiento

- [8] L. C. González, C. S. Lozano, and J. F. J. Gordillo, “Modelación de un sistema inteligente de tráfico vehicular por medio de una simulación basada en agentes,” Pontificia Universidad Javeriana, Bogotá, Colombia, Tech. Rep., 2020, trabajo de grado, Facultad de Ingeniería, Ingeniería Industrial.
- [9] O. Salcedo, C. A. H. Suárez, and A. E. Díaz, “Diseño de un modelo de tráfico a través de series de tiempo para pronosticar tráfico wimax,” *Ingeniería*, vol. 12, no. 1, pp. 4–13, 2007, accessed: 2024-11-26. [Online]. Available: <http://www.redalyc.org/articulo.oa?id=498850164003>
- [10] A. P. Nautiyal and N. Bathla, “Smart transportation systems: Iot-enabled traffic management & vehicle-to-infrastructure communication,” *International Journal for Multidisciplinary Research (IJFMR)*, vol. 6, no. 5, pp. 1–23, 2024. [Online]. Available: <https://www.ijfmr.com>
- [11] J. Mena-Oreja and J. Gozalvez, “Predicción de la velocidad del tráfico basada en redes neuronales convolucionales,” *Performance Measurement System (PeMS)*, 2017, este trabajo utiliza datos del Departamento de Transporte de California. [Online]. Available: <https://pems.dot.ca.gov/>
- [12] J. Zhang, Y. Zheng, and D. Qi, “Deep spatio-temporal residual networks for citywide crowd flows prediction,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 31, no. 1, 2017.
- [13] S. V. Kumar and L. Vanajakshi, “Short-term traffic flow prediction using seasonal arima model with limited input data,” *European Transport Research Review*, vol. 7, no. 21, pp. 1–9, 2015. [Online]. Available: <https://link.springer.com/article/10.1007/s12544-015-0170-8>
- [14] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, “T-gcn: A temporal graph convolutional network for traffic prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020.

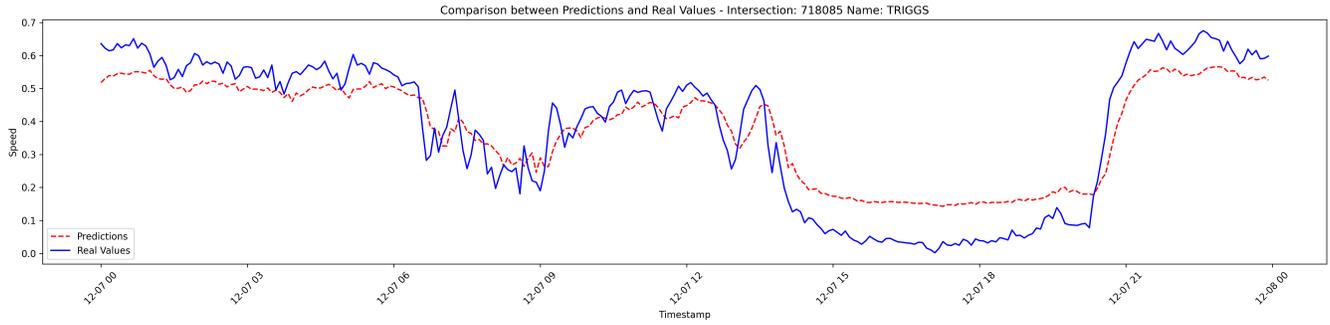


Figure 11. Comparativa entre valores reales y valores predichos para la intersección “716930” (FERRIS) en el set de prueba

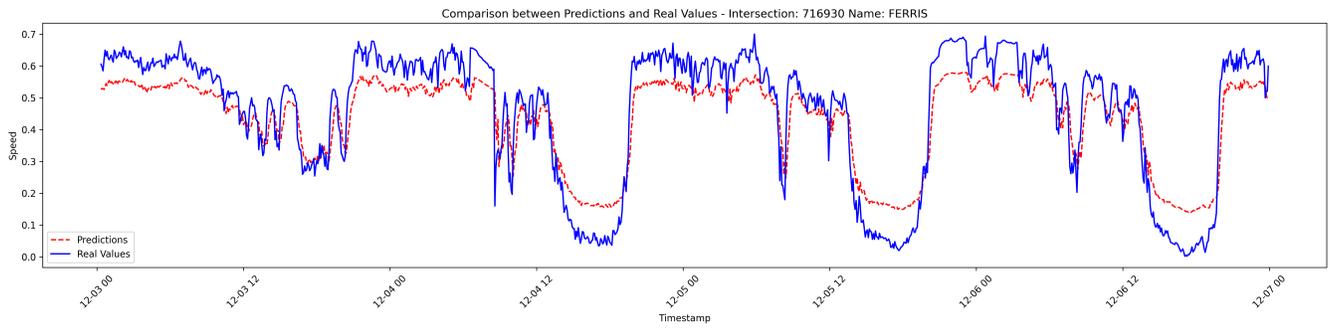


Figure 12. Comparativa entre valores reales y valores predichos para la intersección “718085” (TRIGGS) en el set de entrenamiento

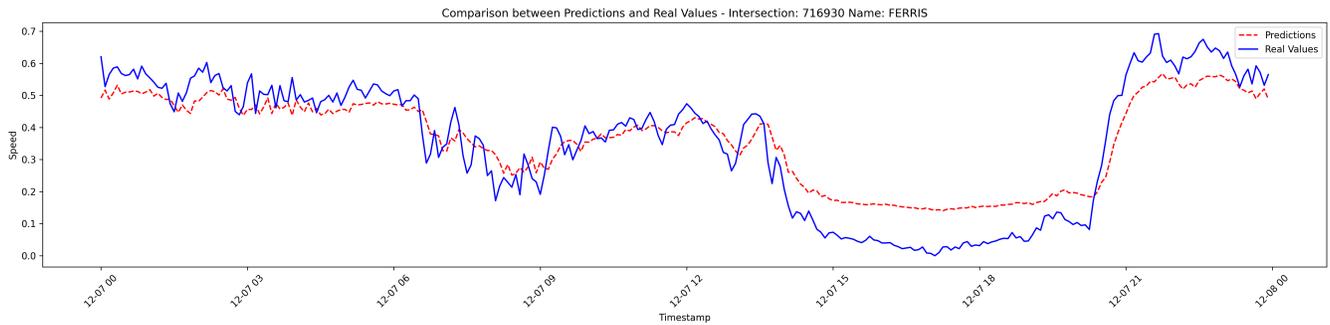


Figure 13. Comparativa entre valores reales y valores predichos para la intersección “718085” (TRIGGS) en el set de prueba