

Neural Network with Long Short Term Memory for Predicting Traffic Flow on 30 intersections

Paul Corella, *Universidad San Francisco de Quito*,

Resumen—This work presents a forecasting pipeline for urban traffic that predicts future vehicular flow at multiple intersections from multivariate time series. The approach targets 30 locations along California’s Interstate 5 (I-5) using historical occupancy, flow, and speed sampled at 5-minute intervals. We employ a many-to-many Long Short-Term Memory (LSTM) network to capture temporal dynamics behaviors. Inputs are framed with a four-step temporal window, and the model outputs concurrent flow forecasts for all intersections. Using an 80/20 train–test split, the proposed method achieves accuracy comparable to traditional traffic-forecasting baselines while reproducing recurrent patterns. We discuss scalability as the primary limitation: extending to larger road networks substantially increases computational demands and may require distributed orchestration. The implementation (PyTorch) is reproducible in Google Colab, and the repository includes processed data (30-node subset) and a slide deck summarizing methodology and results, facilitating reuse and extension in intelligent transportation applications.

Index Terms—Urban Traffic Management, Multivariate Time Series, Temporal Dependencies, Long Short-Term Memory (LSTM).

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[1]. 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 [2].

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 [3]. 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 [4]. 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 [5]. 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 [6]. El propósito de estos dispositivos es proveer datos históricos para determinar patrones más precisos. Los sistemas de transporte inteli-

Paul Corella and Karen Rosero was with Colegio de Ciencias e Ingenierías, Universidad San Francisco de Quito, Quito, e-mail: (pcorellam@estud.usfq.edu.ec and krosero@asig.com.ec

Manuscript received September 13, 2025; revised September 13, 2025. modified September 13, 2025.

gentes proporcionan un enfoque prometedor que permite lidiar con la problemática de las dependencias.



Figura 1. Sensores IoT para redes viales [7].

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.

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:

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 [8]. 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.

Modelos basados en edes neuronales convolucionales (CNNs)

Las redes neuronales convolucionales fueron diseñadas en principio para resolver problemas de visión computacional;

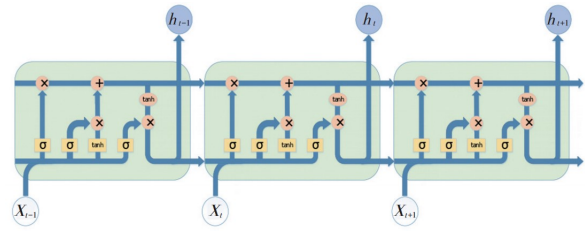


Figura 2. Diagrama estructural de red neuronal LSTM [8].

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 [9]. 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.

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 [10]. 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.

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 [11]. Con este estudio se demostró que el uso de GNNs es una técnica robusta para la predicción de configuraciones complejas de redes viales.

Modelos basados en series temporales multivariantes)

Los modelos autorregresivos como ARIMA han demostrado ser herramientas eficaces para modelar series temporales

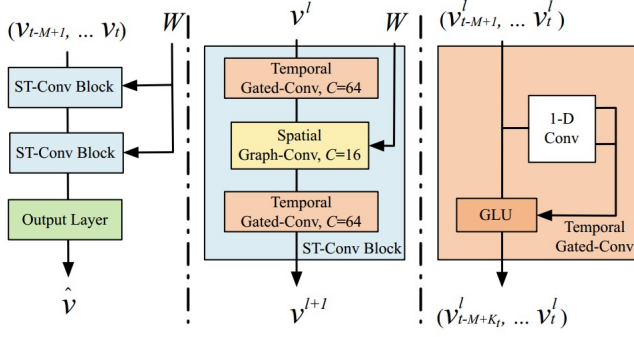


Figura 3. Diagrama estructural de red neuronal STGCN [11].

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 [12]. 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.

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 [13]. La arquitectura del modelo propuesto se muestra en la Figura 4.

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 propuesto es predecir el estado futuro de alguna característica del tráfico, en este caso el flujo 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

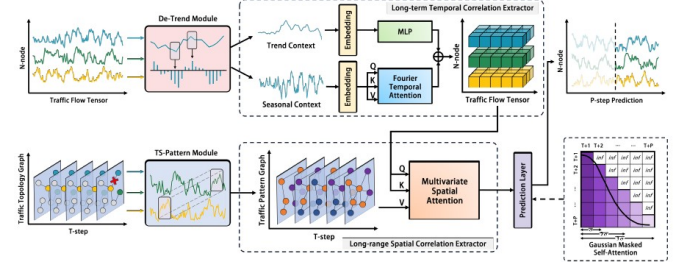


Figura 4. Diagrama estructural de red neuronal espacio temporal con el uso de attention transformers [13].

como su dependencia temporal. El enfoque de las Redes Neuronales Recurrentes Long Short Term Memory (LSTM) abordan este desafío puesto que se enfocan en identificar las dependencias temporales.

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.

METODOLOGÍA PROPUESTA

El presente estudio propone el uso de las Redes Neuronales Recurrentes Long Short Term Memory (LSTMs) de manera similar al estudio desarrollado por Ling Zhao el cual utiliza Gated Recurrent Unit (GRU) [14]. El propósito de esta arquitectura de modelos es diseñar una red neuronal que sea capaz de capturar las relaciones y temporales para proporcionar una predicción precisa y acorde al comportamiento de cada intersección de una sola característica partiendo de tres. Por ejemplo, la el flujo vehicular de una intersección. El uso de una LSTM many to many 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.

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.

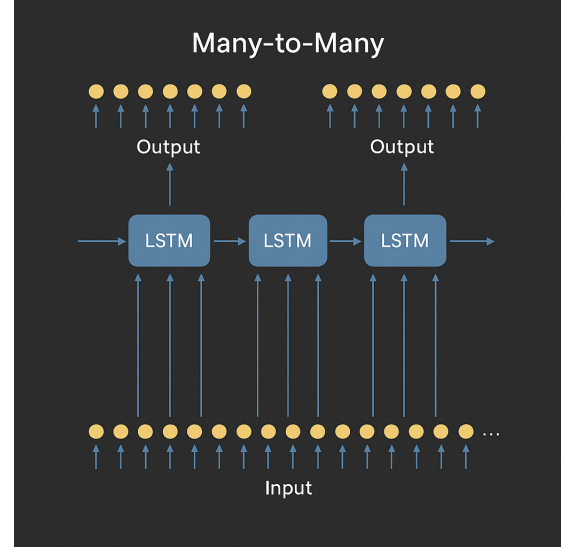


Figura 5. Componente LSTM en aplicado a una ventana de cuatro unidades de tiempo

Descripción del método de evaluación del modelo

El modelo 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 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:

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

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 $N = 28$.

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/Paulsantiagoc/Modelo_LSTM_many2many

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 registradas en el 2016. Dado que el modelo propuesto se enfoca en capturar los patrones de la dinámica vehicular de una sola característica como lo es el flujo vehicular en cada intersección, el set de datos se enfoca en esa variable. Los datos no presentan ninguna inconsistencia. Finalmente cabe destacar que las intersecciones se encuentran unidas de manera consecutiva y cada una cuenta con una medición desde las 0h00 hasta las 23h55 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.

Cuadro 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

Selección de los parámetros del modelo

En esta sección se describen los parámetros relevantes del modelo LSTM. Para este caso de estudio se utiliza una tasa de aprendizaje (learning rate) de 0.001, un tamaño del lote (batch size) de 256 unidades y una cantidad de épocas (epoch) de 50. Dada la naturaleza del modelo, se tiene que este utiliza cuatro unidades LSTM como se indica en el apartado de la metodología propuesta.

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).

RESULTADOS

A continuación, se presentan los resultados de 1 intersección. La Figura 8 corresponde a la comparación de los valores reales con los valores predichos. El resto de las figuras con esta comparativa se puede ser codificadas en el repositorio de GitHub que forma parte integral de este estudio. El enlace del repositorio es el siguiente https://github.com/Paulsantiagoc/Modelo_GCNLSTM.

CONCLUSIONES

El modelo LSTM propuesto resulta ser una herramienta útil para capturar la dependencia temporal de una red vial. 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 rendimiento equiparable al que los modelos tradicionales ofrecen (modelo estadísticos) puesto que permite integrar tanto dimensiones temporales no lineales.

Este modelo presenta en esencia un alto grado de complejidad. Si bien el modelo es escalable puesto que puede adaptarse fácilmente a una mayor cantidad de datos 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 a penas meses y solo 30 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.

REFERENCIAS

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.

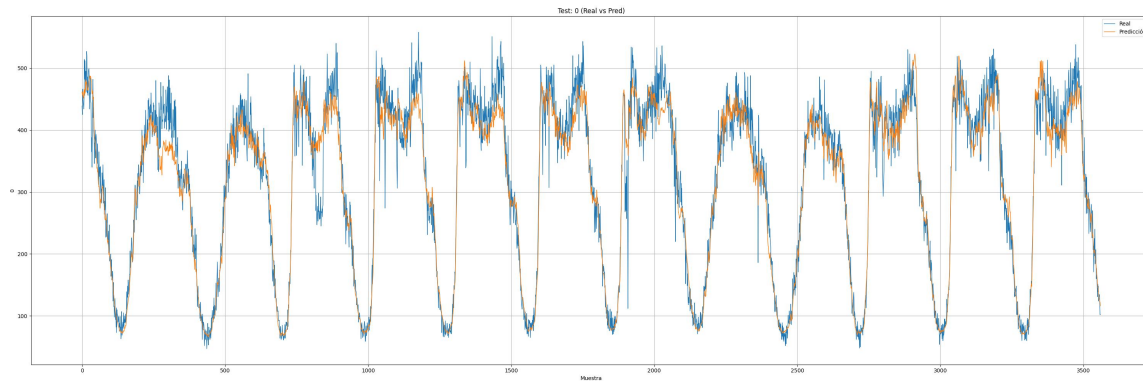


Figura 6. Comparativa entre valores reales y valores predichos para la intersección “715947” (S OF 710)

- [5] 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>
- [6] 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>
- [7] ITS Answers, “Adaptive signal system,” 2024, accessed: 2024-11-25. [Online]. Available: <https://www.itsanswers.com/project/adaptive-signal-system/>
- [8] W. Shihao, Z. Qinzhen, 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>
- [9] 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/>
- [10] 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.
- [11] 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>
- [12] 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>
- [13] 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>
- [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.