

Predicción del Estado del Tráfico Usando Redes Neuronales y Datos FCD del Vehículo Conectado

Daniel Sempere-García, Sergio Montero, Miguel Sepulcre, Javier Gozalvez

daniel.sempere02@goumh.umh.es, smontero@umh.es, msepulcre@umh.es, j.gozalvez@umh.es

UWICORE lab. Universidad Miguel Hernández de Elche (UMH). Avda. Universidad s/n, 03202, Elche.

Abstract- Road traffic management is a key area for any transportation authority for road traffic planning and control. One of its challenges is the accurate prediction of the traffic status to anticipate and prevent traffic congestion situations through traffic re-routing or using dynamic message signs. To date, the use of cameras and induction loops has played a major role in the data acquisition process for the prediction of the traffic status. However, this type of sensors has high installation and maintenance costs, which limits their deployment. In this context, this paper proposes the use of connected vehicle data and artificial neural networks to improve the accuracy and ubiquity of traffic status predictions. Connected vehicles will become mobile sensors that provide information about their position and speed by means of Floating Car Data (FCD). Using real FCD, the results obtained show the potential of connected vehicles to overcome some of the limitations of static sensors.

Palabras Clave- Vehículo Conectado, Floating Car Data (FCD), Inteligencia Artificial, Redes Neuronales Artificiales, Predicción de Tráfico.

I. INTRODUCCIÓN

El incremento del número de vehículos en circulación de las últimas décadas ha dado lugar a la aparición de problemas en la circulación vial. Uno de estos problemas son las situaciones de congestión del tráfico, que llevan asociados un importante incremento en los niveles de emisión de CO₂, así como una considerable disminución del nivel de bienestar de los usuarios. Con el objetivo de minimizar las situaciones de congestión en carretera, es crucial realizar una gestión del tráfico vial dinámica y eficiente. Uno de los principales elementos empleados para ello es la estimación del estado del mismo en un futuro cercano, ya que una predicción fiable de posibles situaciones de congestión permite al gestor del tráfico anticiparse y tomar las medidas pertinentes (p. ej., anuncio de cambios de rutas mediante paneles de mensajería) para reducir o evitar los efectos negativos de las mismas.

Durante las últimas décadas se han desarrollado métodos y técnicas que proporcionan estimaciones precisas y fiables del estado del tráfico. Debido a la naturaleza no lineal de los datos y de las características del tráfico, las Redes Neuronales Artificiales o RNA son una de las técnicas de inteligencia artificial más empleadas. En el ámbito de la estimación del estado del tráfico, las RNA se han empleado principalmente como un método de análisis de datos debido a su capacidad de trabajar con cantidades masivas de datos, su flexibilidad de modelado, su capacidad de aprendizaje y generalización de problemas, su adaptabilidad y su buena capacidad de predicción [1]. Además, las RNA no requieren un conocimiento previo del comportamiento del tráfico para

su uso, ya que son capaces de aprender mediante los datos de tráfico proporcionados.

Para poder utilizar las RNA para la predicción del estado del tráfico se requiere un proceso de aprendizaje utilizando un conjunto de datos de entrenamiento. Las fuentes de datos de tráfico utilizadas hasta la fecha son fundamentalmente sensores fijos ubicados en carretera (espiras inductivas, cámaras de vídeo, radares, etc.) [2]. Estos sensores proporcionan información sobre la velocidad de los vehículos y la intensidad del tráfico en aquellos puntos donde se instalan, entre otras métricas, y presentan ciertas limitaciones, como la falta de ubicuidad de los datos, así como los elevados costes de despliegue y mantenimiento [3].

Como alternativa a los sensores fijos se plantea la utilización de datos en tiempo real procedentes de vehículo conectado, también conocidos como datos FCD (*Floating Car Data*). Los datos FCD son datos de posicionamiento geográfico y velocidad generados periódicamente por los vehículos en circulación, y enviados a un sistema de datos central para su procesado y almacenamiento. Mediante los datos FCD, el vehículo conectado proporciona información muy útil sobre el estado del tráfico de forma ubicua y sin necesidad de desplegar sensores fijos en carretera. Aunque su principal limitación a día de hoy es la baja penetración del vehículo conectado, se prevé que los datos FCD sean un complemento o alternativa a los datos proporcionados por sensores fijos para la estimación del tráfico vial [3].

El presente trabajo estudia el uso de RNA y analiza su precisión para la predicción del estado del tráfico utilizando datos reales FCD del vehículo conectado. Evolucionando trabajos anteriores como [3], en este trabajo se proponen diferentes métodos para pre-procesar los datos, y se estudia la dependencia temporal y espacial de los mismos. Los resultados obtenidos con FCD se han comparado con los que podrían obtenerse con sensores fijos, mostrando el potencial del vehículo conectado como sensor móvil para proporcionar información útil en la estimación del estado del tráfico.

II. REDES NEURONALES ARTIFICIALES

Existen multitud de tipos de RNA, de entre los cuales se ha seleccionado el Perceptron Multicapa por ser considerado como un aproximador universal capaz de aproximar e interpolar relaciones no lineales entre la entrada y salida del sistema. Es una de las arquitecturas más empleadas en la resolución de problemas en áreas como el reconocimiento de voz, la conducción de vehículos, la predicción de series temporales, etc. [4], y se ha demostrado su capacidad para

realizar estimaciones no lineales del estado del tráfico con sensores fijos [2], aunque son más limitados los estudios que las usan para la predicción del tráfico usando datos FCD [3].

Un Perceptron Multicapa consta de varias capas de neuronas interconectadas, tal y como muestra la Fig. 1: una capa de entrada, varias capas ocultas y una capa de salida. Las conexiones inter-neuronales del Perceptron Multicapa se establecen siempre en dirección hacia adelante, y cada conexión presenta un peso w_{ij}^c . La relación entre las entradas y salidas se obtiene propagando hacia adelante las variables de entrada, de forma que cada neurona procesa los datos recibidos en sus entradas para producir una respuesta o activación que se propaga a la siguiente capa. [4].

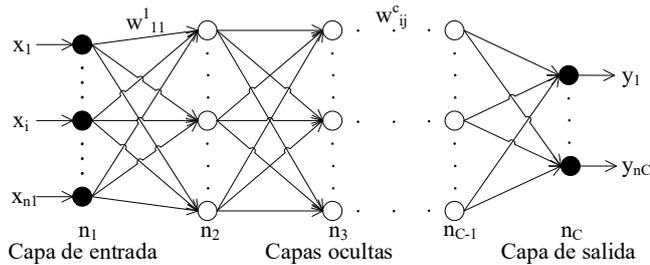


Fig. 1. Arquitectura general del Perceptron Multicapa.

El número de entradas, el número de neuronas de la capa oculta y el número de capas ocultas óptimos dependen en general del problema a resolver y de la cantidad de datos que se tengan para entrenar la red. Se han realizado pruebas usando 1, 2 y 3 capas ocultas (no incluido por falta de espacio), y los resultados muestran la reducida diferencia entre las distintas configuraciones. Por ello, el Perceptron Multicapa utilizado en este trabajo utiliza tres capas: la de entrada, una capa oculta, y la de salida. Pese a ser la arquitectura más sencilla, se ha demostrado su capacidad para realizar estimaciones no lineales a partir de datos de tráfico con una complejidad computacional menor [3][5].

Al trabajar con datos de tráfico es deseable tener en cuenta las posibles relaciones temporales y espaciales existentes. Para ello, en este trabajo se introducen como valores de entrada a la RNA los datos actuales y del pasado reciente en cada momento, considerando además los correspondientes al segmento en cuestión y sus colindantes en el caso de FCD. En función de la configuración temporal y espacial utilizada para el pre-procesado de los datos varía la cantidad de entradas de la red neuronal, así como el número de neuronas de la capa oculta. La capa de salida en este trabajo presenta una única neurona, correspondiente con el valor que se desea estimar. Los valores de entrada y salida dependen del tipo de datos de tráfico usados, por lo que son distintos a la hora de entrenar la RNA propuesta con datos FCD (velocidad) o de sensores fijos (velocidad e intensidad).

Para la utilización de la RNA en la predicción del tráfico, se requiere realizar un proceso de entrenamiento previo. Para ello, se usa un conjunto de datos de entrenamiento y otro de test. El conjunto de entrenamiento se emplea en el proceso de aprendizaje para obtener los pesos w_{ij}^c , que son validados posteriormente con el conjunto de datos de test, que debe ser distinto al de entrenamiento para asegurar la capacidad de generalización de la red entrenada.

III. DATOS FCD DE VEHÍCULO CONECTADO

En este trabajo se han utilizado datos FCD reales de vehículo conectado proporcionados por la empresa INRIX.

INRIX es un proveedor de datos FCD con una red de más de 250 millones de vehículos conectados. INRIX recolecta los datos sobre la velocidad y posición geográfica de forma individual para cada uno de sus vehículos, y los proporciona a sus clientes de forma agregada temporalmente en intervalos de 1 minuto y espacialmente organizando las carreteras en segmentos. Los segmentos propuestos por INRIX son los segmentos XD, cuya longitud aproximada es de 1 milla. De este modo, cada minuto proporciona para cada segmento XD la velocidad media, la velocidad de referencia del segmento (velocidad de flujo libre), y el tiempo de recorrido.

Los datos FCD utilizados para la realización del presente trabajo corresponden a 9 carreteras ubicadas en el País Vasco, sumando un total de 505 segmentos XD. El conjunto temporal de datos empleado es de 5 semanas, con un total de aproximadamente 50000 datos por cada segmento. El 80% de estos datos se ha empleado para el entrenamiento de la RNA propuesta y el 20% restante para el correspondiente test.

La calidad y características de los datos pueden influir significativamente en la precisión de la predicción con RNA. Las 9 carreteras seleccionadas presentan diferentes características del comportamiento del tráfico y distintos índices de penetración de la flota de vehículos conectados, influyendo así en la cantidad de situaciones de congestión sucedidas o detectadas, respectivamente. Para que la RNA sea capaz de estimar y detectar las situaciones de congestión, esta debe entrenarse con un conjunto de datos que presente el mayor número de ejemplos posible de este tipo de situaciones. El análisis de los datos de cada carretera ha permitido clasificarlas en dos amplios grupos: carreteras con una cantidad aceptable de situaciones de congestión, y carreteras con datos con pocas situaciones de congestión (carreteras tipo 1 y tipo 2, respectivamente).

Las características de los datos FCD son diferentes a las de los sensores fijos. En primer lugar, los sensores fijos proporcionan medidas bastante precisas de velocidad y conteo de todos los vehículos que pasan por su ubicación, mientras que FCD solo proporciona el valor de la velocidad de los vehículos conectados. La velocidad se calcula de un modo diferente para FCD (valor medio espacial) y para los sensores fijos (valor medio temporal). Dada la baja penetración a día de hoy del vehículo conectado, se puede afirmar que los datos de sensores fijos son más precisos que los de FCD, aunque la principal ventaja que proporcionan estos últimos es la ubicuidad con respecto a los primeros.

IV. MÉTRICAS DE RENDIMIENTO

El rendimiento de una RNA en la estimación de un parámetro se mide habitualmente utilizando un conjunto de datos de test del que se conoce los valores reales a predecir y se comparan con las estimaciones obtenidas. Con datos FCD el parámetro a estimar es la velocidad de los vehículos. La precisión con la que se estima dicha velocidad se suele evaluar calculando distintas métricas del error absoluto o relativo, tales como la media o la mediana. Al tratarse de una cantidad muy elevada de valores existe la posibilidad de que alguno de ellos aisladamente proporcione errores disparados que afecten al valor de la media, por lo que en este trabajo se utiliza la mediana, asegurando que el 50% de los valores son como máximo el valor de error indicado.

En este trabajo se pretende ir un paso más allá y evaluar el rendimiento de la RNA propuesta para predecir el estado del tráfico haciendo uso de datos FCD. Para ello, se sigue el

criterio que utiliza INRIX para clasificar el estado del tráfico en un segmento XD en 4 niveles, en base al ratio entre la velocidad media del segmento y su velocidad de referencia. Cuanto mayor es la velocidad media frente a la velocidad de referencia, menor nivel de congestión se observa. En la Tabla I se muestran los 4 niveles de congestión que define INRIX, donde el nivel D corresponde al nivel de mayor congestión. Para estimar el estado del tráfico en un instante futuro, la RNA propuesta estima la velocidad media en dicho instante, y con ella y la Tabla 1 se predice el nivel de congestión correspondiente. Dado que cada segmento posee una velocidad de referencia, los valores de velocidad en las entradas y la salida de la RNA son valores relativos respecto a la velocidad de referencia del segmento correspondiente.

TABLA I
NIVELES DE CONGESTIÓN UTILIZADOS

Niveles	D	C	B	A
Vel. media / Vel. referencia	0-31%	32-62%	63-92%	>93%

La comparativa con sensores fijos en cuanto a la mediana del error relativo es análoga a la de FCD. Sin embargo, cabe destacar que para los sensores fijos la clasificación del estado del tráfico en distintos niveles se efectúa en base al HCM (*Highway Capacity Manual*), obteniendo un total de 6 niveles (A, B, C, D, E, F). Por tanto, el análisis y los valores proporcionados se corresponden con el porcentaje de acierto en la estimación del máximo nivel de congestión, que es el nivel D para FCD y el nivel F para sensores fijos.

V. EVALUACIÓN

A. Configuración y pre-procesado

El rendimiento de la RNA propuesta con datos FCD se ha evaluado empleando diversos métodos de pre-procesado de los datos. Un primer método consiste en entrenar una RNA por cada segmento XD de cada carretera (método Seg), siendo este el que se utiliza habitualmente en la literatura. Sin embargo, este método puede tener problemas de escalabilidad en redes de carreteras de gran tamaño y con un gran número de segmentos. Para tratar de reducir este problema se propone el método Car, que consiste en el entrenamiento de un modelo de red por cada carretera, agrupando los datos de todos los segmentos que forman la misma. Esta nueva metodología implica una mayor sencillez en la gestión de los modelos cuando la red de carreteras es muy extensa e incluye una gran cantidad de segmentos, por lo que se plantea como posible solución para dichos casos.

Estos dos métodos de pre-procesado se han evaluado además aplicando técnicas de balanceo de datos. Este tipo de técnicas son útiles para lograr que dentro del conjunto de datos empleado exista la misma cantidad de datos asociados a cada uno de los 4 niveles de congestión definidos. Para aplicar balanceo a un conjunto de datos se ha estudiado en cada caso la cantidad de instancias de entrenamiento asociadas a cada nivel, se ha seleccionado aquel nivel que menor cantidad de instancias presenta, y se han reducido las de los niveles restantes eliminando aleatoriamente las sobrantes para igualar la cantidad de instancias de todos ellos. Dado que la eliminación de instancias se realiza aleatoriamente, este proceso se ha repetido numerosas veces promediado los resultados obtenidos. A los métodos Seg y Car se les denomina SegB y CarB al aplicar balanceo.

La RNA propuesta para trabajar con datos FCD se ha evaluado utilizando múltiples configuraciones para generar

los conjuntos de entrenamiento y de test, con el objetivo de analizar la dependencia espacial y temporal de los datos con un barrido más amplio y completo que en [3]. El barrido efectuado para conservar la relación temporal consiste en la toma de datos de los 30, 60 o 90 minutos previos al instante actual como entradas de la RNA, mientras que la relación espacial se ha mantenido usando las configuraciones siguientes: datos de cada segmento por separado, datos de cada segmento, el previo y el posterior, o datos de cada segmento, los dos previos y los dos posteriores. Por tanto, se analizan 9 configuraciones, identificadas con el formato “(n° de seg. previos) - (n° de seg. posteriores) - (tiempo pasado en minutos)” (p. ej., 1-1-60). La predicción obtenida consiste en estimar la velocidad en un futuro próximo de 15 minutos.

El rendimiento de la RNA propuesta para datos FCD se ha comparado con el que se obtiene con datos de sensores fijos. Estos sensores fijos proporcionan datos de velocidad e intensidad del tráfico (vehículos/hora) agregados cada 5 minutos sobre el punto concreto donde estén desplegados. En particular, se han utilizado los datos disponibles de 2 sensores fijos recogidos durante 2 semanas. Se ha realizado un barrido temporal considerando distintas cantidades de datos previos al actual en cada caso, y a partir del análisis realizado (no incluido por falta de espacio) se ha seleccionado la configuración con mejores resultados. Cabe destacar que los valores incluidos se usan como referencia para comparar el comportamiento al trabajar con FCD, y que han sido calculados con los datos disponibles en este caso, por lo que pueden diferir en cuanto a calidad con respecto a otras investigaciones realizadas en este campo.

Para la evaluación de la RNA propuesta se ha empleado el software de código abierto Weka (*Waikato Environment for Knowledge Analysis*), que incluye una colección de algoritmos de aprendizaje automático y herramientas para el procesamiento de datos, entre otros, y Matlab para pre- y post-procesar los datos generados por Weka.

B. Predicción de la velocidad

Con el objetivo de mostrar de un modo visual los resultados de la estimación que se pretende efectuar con los datos FCD y técnicas planteadas, en la Fig. 2 se muestra un ejemplo que compara los valores reales y predichos de velocidad que se han obtenido para un segmento concreto. Como se puede observar, la estimación se ajusta con bastante precisión al valor real. La mediana del error relativo obtenido para este segmento concreto es de 2.23%.

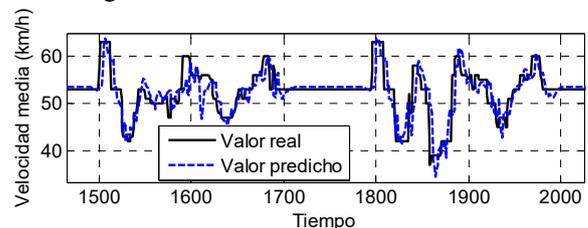


Fig. 2. Ejemplo de estimación del valor de la velocidad.

El análisis efectuado respecto a los valores de la mediana del error relativo se recoge en la Fig. 3 (a) (carreteras tipo 1) y (b) (carreteras tipo 2) para el barrido de configuraciones que analiza la dependencia espacial y temporal de los datos. Para la obtención de cada punto en todas las gráficas siguientes se calcula la media de los valores obtenidos para todas las carreteras de cada tipo, habiendo calculado también la media para todos los segmentos de cada carretera en los

métodos Seg y SegB. De modo general, para ambos tipos de carreteras se observa que para los métodos Seg y Car los valores de la mediana del error relativo obtenidos se encuentran en un rango similar a los obtenidos para los sensores fijos, es decir, siempre por debajo del 10%, aunque el método Seg produce resultados ligeramente mejores a Car. Los resultados obtenidos no permiten a priori obtener una clara indicación sobre la configuración óptima respecto a la dependencia temporal y espacial de los datos.

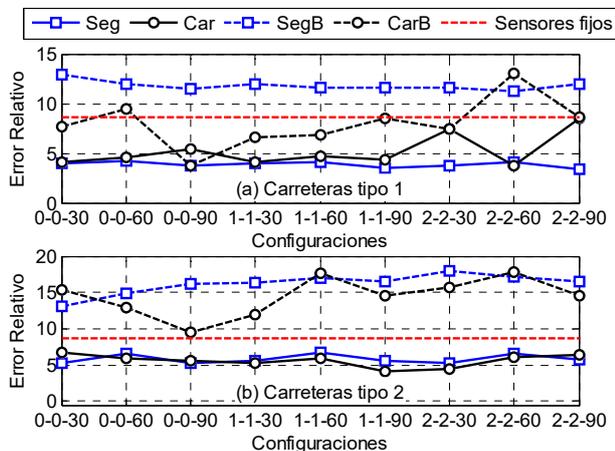


Fig. 3. Mediana del error relativo de la velocidad.

Cuando aplicamos técnicas de balanceo de datos (métodos SegB y CarB), se observa que los valores de error relativo se incrementan. Esto es debido principalmente a la disminución del tamaño de los conjuntos de entrenamiento. Aunque el balanceo de datos pueda parecer contraproducente para la predicción de la velocidad, se verán más adelante sus beneficios para la detección de situaciones de congestión.

C. Predicción del estado del tráfico

El estado del tráfico se ha estimado a partir de la predicción de la velocidad realizada con la RNA propuesta. Dado que un operador de tráfico está fundamentalmente interesado en anticiparse a posibles situaciones de congestión de tráfico, a continuación se presentan los resultados para analizar la precisión con la que la RNA propuesta es capaz de predecir correctamente una situación de máxima congestión. Estos resultados se muestran en Fig. 4 (a) (carreteras tipo 1) y (b) (carreteras tipo 2).

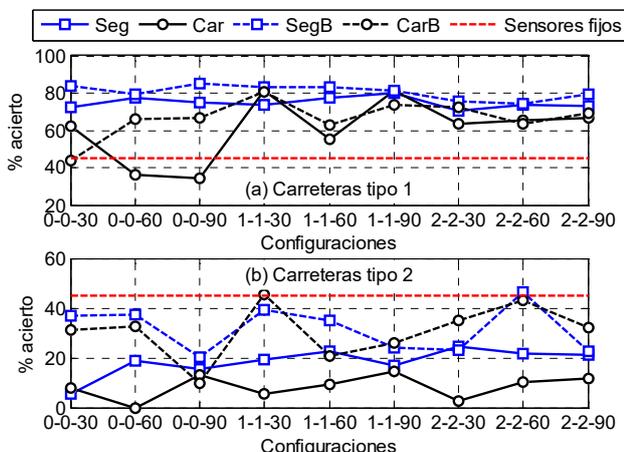


Fig. 4. Porcentaje de acierto al estimar el nivel de máxima congestión.

Para las carreteras tipo 1 (Fig. 4 (a)), se observa que el porcentaje de acierto al estimar el nivel de máxima congestión con datos FCD (independientemente del método

usado) supera el logrado para la prueba realizada con los datos de sensores fijos disponibles (45%). Asimismo, los resultados obtenidos muestran que el método Seg presenta mayores porcentajes de acierto para todas las configuraciones (entre el 70% y el 80%) que el método Car (propuesto en este trabajo). Sin embargo, existen ciertas configuraciones mediante las cuales se logran alcanzar estos mismos porcentajes de acierto con el método Car, demostrando su potencial para la predicción del tráfico. Los resultados obtenidos también demuestran que si se aplican técnicas de balanceo al conjunto de datos de entrenamiento (métodos SegB y CarB) se logra una mejora general de los porcentajes de acierto con respecto a los de los métodos Seg y Car. La mejora se produce especialmente para algunas configuraciones y se obtienen valores de porcentaje de acierto de hasta un 85% (método SegB).

Para las carreteras tipo 2 (Fig. 4 (b)), se comprueba que los porcentajes de acierto del nivel de máxima congestión obtenidos son menores que para las carreteras tipo 1, debido a la reducida cantidad de situaciones de congestión contenidas en los conjuntos de datos utilizados. De hecho, los resultados obtenidos se encuentran por debajo del valor de referencia de los sensores fijos (45%). Además, se observa que en este caso los porcentajes asociados al método Seg son de nuevo superiores a los del método Car. Sin embargo, estos reducidos valores consiguen elevarse si se aplica el balanceo del conjunto de datos de entrenamiento a estos métodos (métodos SegB y CarB), pasando a tomar valores que prácticamente duplican a los anteriores, y que se sitúan en la mayoría de casos entre el 25% y el 48%.

VI. CONCLUSIONES

En este trabajo se propone el uso de redes neuronales artificiales y analiza su precisión para la predicción del estado del tráfico usando datos reales FCD del vehículo conectado. El análisis realizado utilizando diversas formas de pre-procesado de datos muestra el potencial del vehículo conectado para la detección de situaciones de congestión vial frente al costoso despliegue de sensores fijos tradicionales. A medida que aumente la penetración del vehículo conectado en las carreteras, se espera incluso que la precisión de este tipo de sistemas aumente.

AGRADECIMIENTOS

Este trabajo ha sido realizado a través de una sub-contratación de la empresa IKUSI en el marco del proyecto CIVIC, financiado por el Ministerio de Industria, Energía y Turismo (TSI-100105-2014-201).

REFERENCIAS

- [1] M. Karlaftis y E. Vlahogianni, "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights," in *Transportation Research Part C: Emerging Technologies*, vol. 19, Issue 3, pp. 387-399, Junio 2011.
- [2] E. I. Vlahogianni et al., "Short-term traffic forecasting: Where we are and where we're going," in *Transportation Research Part C: Emerging Technologies*, vol. 43, Part 1, pp. 3-19, Junio 2014.
- [3] C. de Fabritiis et al., "Traffic Estimation and Prediction Based on Real Time Floating Car Data," *Proc. 11th International IEEE Conference on Intelligent Transportation Systems*, Octubre 2008, Beijing (China).
- [4] P. Isasi Viñuela e I. M. Galván León, *Redes de Neuronas Artificiales. Un Enfoque Práctico*, Prentice Hall, 2004.
- [5] K. Y. Chan et al., "Prediction of Short-Term Traffic Variables Using Intelligent Swarm-Based Neural Networks," in *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, Enero 2013.