

Análisis de la calidad del servicio en el transporte público mediante redes neuronales artificiales

M^a Concepción Garrido Rodríguez

TRYSE Research Group. ETSICCP Granada, Universidad de Granada, España

Juan de Oña López

TRYSE Research Group. ETSICCP Granada, Universidad de Granada, España

RESUMEN

El éxito de un servicio de transporte público reside en su capacidad para captar nuevos pasajeros y de fidelizar el uso de los actuales. De ahí que las administraciones y los gestores de transporte se preocupen cada vez más por conocer cuál es la calidad percibida por los usuarios. Aunque existen numerosos métodos matemáticos que ya han sido utilizados para analizar la calidad del servicio en el transporte público, es necesario seguir avanzando en el estudio de nuevas técnicas que sean válidas para su estudio.

En esta ponencia se aplica la técnica de las redes neuronales artificiales al estudio de la calidad del servicio en el transporte público. La metodología novedosa desarrollada al aplicar las redes neuronales artificiales mitiga considerablemente las principales limitaciones que éstas presentan, por lo que además de aportar una nueva herramienta para análisis de la calidad del servicio en el transporte público, pone a disposición de los especialistas en esta técnica un nuevo procedimiento de aplicación de las redes neuronales artificiales aplicable a cualquier campo de estudio en el que esta técnica sea aplicada.

1. INTRODUCCIÓN

La calidad del servicio en el transporte público es un concepto complejo, difuso y abstracto, que en los últimos años ha pasado a ocupar una posición primordial en el estudio de cualquier sistema de transporte público, debido a la relevancia de la información que aporta a la hora de comprender la percepción y/o expectativas que los usuarios actuales y/o potenciales tienen de él.

La forma más frecuente de recoger información sobre la calidad del servicio es por medio de encuestas, a través de las cuales puede obtenerse una valoración global de ésta o una identificación de los diferentes atributos que la caracterizan. Precisamente esta información es la base a partir de la cual se pueden desarrollar estrategias tendentes a conseguir una mejora de la calidad percibida.

Hasta el momento se han venido aplicando diversas técnicas matemáticas para analizar la calidad del servicio en el transporte público, como son: el análisis factorial (Kim y Lee,

2011), los modelos de regresión (Kim y Lee, 2011; Aksoy et al., 2003; Tyrinopoulos y Aifadopoulo, 2008; Tyrinopoulos y Antoniou, 2008; Dell’Olio et al., 2010, Huse y Evangelho, 2007; Henser et al., 2010; Castillo y Benítez, 2012), los modelos de elección discreta y las ecuaciones estructurales (Andreasen, 1995; Eboli y Mazzulla, 2007; Friman y Gärling, 2001; Friman et al., 2001; Joewono y Kubota, 2007; Lai y Chen, 2011; Minser y Webb, 2010). Todos ellos presentan limitaciones relacionadas con la dependencia existente entre las variables, la necesidad de partir de un modelo inicial predefinido o la normalidad de los datos. Aunque se ha conseguido salvar muchas de estas limitaciones, es necesario seguir trabajando en la búsqueda de nuevas técnicas. Así, se han aplicado recientemente y de forma exitosa los árboles de decisión (de Oña et al., 2012), una técnica no paramétrica de minería de datos que salva algunas de las limitaciones presentes en los métodos anteriormente señalados.

Las redes neuronales artificiales (ANN) son una técnica matemática que simula el comportamiento biológico del cerebro humano (Martín del Brío y Sanz, 2006). Su principal fortaleza reside en su gran capacidad de predicción (aunque exista ruido o falta de información en los datos), y de tolerancia a fallos, así como en la posibilidad de captar las relaciones complejas y altamente no lineales existentes entre las variables predictoras consideradas (Abdelwahab y Abdel-Aty, 2001). Además, no es necesario definir a priori ningún modelo matemático (Mohammadipour y Alavi, 2009). Esto ha llevado a usar las ANN en numerosos campos de estudio, obteniéndose unos resultados de exactitud en la predicción muy elevados. Sin embargo, presenta dos limitaciones importantes: la dificultad para conocer explícitamente las relaciones existentes entre las variables predictoras y las variables dependientes y su dificultad para encontrar la ANN que presenta un comportamiento óptimo (Cao y Qiao, 2008). Debido a la primera de las limitaciones mencionada, a las ANN se las ha denominado “cajas negras”, y aunque se han desarrollado numerosos métodos para determinar la importancia relativa de las variables, no existe consenso sobre cuál de ellos presenta un mejor comportamiento y además siguen presentando problemas de inestabilidad (Cao y Qiao, 2008).

Recientemente, se ha aplicado de forma novedosa la técnica de las ANN para analizar la calidad del servicio en el transporte público (De Oña y Garrido, 2014; Garrido et al., 2014), y además se ha realizado una comparación entre las técnicas de ANN y árboles de decisión aplicadas a la calidad del servicio en el transporte público (De Oña et al., 2015; 2016); hasta entonces, las ANN habían sido utilizadas con éxito en otros campos de la ingeniería del transporte. La aplicación de ANN al transporte proliferó a partir de los años 90 (Dougherty, 1995), y desde entonces han surgido numerosos estudios en los que se ha aplicado esta técnica a las distintas ramas del mismo. En esta década se elaboraron estudios sobre el comportamiento del conductor (Lyons y Hunt, 1993; Hunt y Lyons, 1993; Pant y Balakrishman, 1994) en los que se analizaban aspectos como la toma de decisiones ante los condicionantes que presentaban las distintas alternativas de viaje, el cambio de conducta cuando se proporcionaba a los conductores determinada información o comportamientos

ante las condiciones de conducción programadas en simuladores. Así, Lee et al. (2010) aplicaron ANN y árboles de decisión para analizar los factores que afectan a las decisiones de los conductores al seleccionar la ruta de viaje entre varias alternativas, y Xie et al. (2003) modelaron diferentes modos de viajes al trabajo usando tres técnicas diferentes: ANN, árboles de decisión y modelos logísticos multinomiales.

Numerosos han sido también los estudios que han utilizado esta técnica para analizar los flujos de datos de tráfico, a través de matrices origen-destino, e incluso combinándola con otras técnicas como los algoritmos genéticos (Xiong y Schneider, 1992), por medio de los cuales se ha modelado el estado del tráfico (Bullock et al., 1993).

Costa y Marquellos (1997) analizaron la eficiencia en términos de productividad del metro de Londres entre los años 1970 y 1994, y los estudios llevados a cabo por Morato et al. (2001) utilizaron ANN para predecir el comportamiento del conductor ante las variaciones del tráfico. Del Hoyo et al. (2009) usaron las tipologías de ANN perceptrón multicapa, redes difusas y redes neuronales auto-organizadas para medir la calidad del servicio ofrecida por diversas aplicaciones de video y voz que pueden descargarse a través de internet y que indican en tiempo real el estado del tráfico de diferentes rutas alternativas de viaje. Las ANN también han sido usadas con éxito para estudiar la calidad del servicio en campos no relacionados con la ingeniería, obteniéndose unos ajustes y exactitud de los modelos entrenados muy elevados. Principalmente se han utilizado en el sector terciario para medir la calidad percibida por los usuarios por el servicio percibido, debido fundamentalmente a la gran importancia que tiene para la viabilidad financiera y la competitividad en el mercado de la compañías del sector, el ofrecer un valor añadido a sus productos y conseguir que sus clientes queden satisfechos (Armstrong y Kotler, 2000; Weitz y Jap, 1995). En la literatura aparecen estudios en los que se analizan los factores más influyentes en la calidad del servicio de comunicación plantilla – cliente para empresas del sector servicios (Lin, 2007; Bejou et al., 1996), el grado de satisfacción de los clientes de hoteles en función del trato recibido por el personal (Deng et al., 2008; Deng y Pei, 2009), de los estudiantes usuarios de servicios de concesiones de restaurantes en colegios, institutos y universidades (Larasati et al., 2012). Moutinho et al. (1996) evaluaron el grado de satisfacción de los compradores de una marca de vehículos y su fidelidad hacia la misma; Azadeh et al. (2011) aplicaron ANN para medir el grado de satisfacción de los trabajadores de diversas compañías de la industria de la refinación de gas. Sin embargo, y a pesar de la elevada exactitud de los modelos entrenados, todos estos estudios presentan las limitaciones de inestabilidad en el cálculo de importancia relativa de las variables y de dificultad para encontrar la red de comportamiento óptimo, apuntadas anteriormente.

Los objetivos de este estudio son, por una parte, demostrar que las ANN son una técnica válida para analizar la calidad del servicio en el transporte público, y por otra, desarrollar una nueva metodología que permita mitigar las principales limitaciones que actualmente presenta esta técnica, y que sea de aplicación a cualquier estudio en el que se apliquen las

ANN.

2. CASO DE ESTUDIO

2.1 Descripción del sistema de transporte

Granada es una ciudad cuyo crecimiento está constreñido por una parte, por la abrupta orografía de su sierra y por otra, por las protecciones medioambientales de su vega. Paradójicamente, los pueblos de la vega que rodean a la ciudad no tienen impuesta ningún tipo de limitación a su crecimiento, por lo que en las dos últimas décadas se ha producido un aumento considerable en el tamaño y en la población de los municipios del área metropolitana, mientras que la ciudad central ha mantenido prácticamente tanto su tamaño como su población. Actualmente, la población de Granada capital (según datos INE 2015) es de 235.800 habitantes, mientras que la de los núcleos de su área metropolitana es de 252.595 habitantes. Las fuertes relaciones socioeconómicas que han surgido entre la zona metropolitana y la ciudad central han generado unas demandas de movilidad muy importantes, que han provocado situaciones de congestiones y atascos debidas al uso excesivo del vehículo privado. En el año 2003 se creó el Consorcio de Transporte Metropolitano del Área de Granada (CTMAG) para potenciar un sistema de transporte público por autobús eficaz que fuese una alternativa real al vehículo privado. El servicio de transporte público por autobús existente hasta el momento mejoró considerablemente mediante la puesta en funcionamiento de nuevas líneas y el refuerzo de expediciones de las existentes, de modo que en 2007 un total de 10.800.000 viajeros utilizaron este modo de transporte público, contando éste con 56 líneas, 13 operadores, una flota de 96 unidades y un sistema tarifario unificado. Tras la creación del CTMAG, el porcentaje de crecimiento del número de viajeros interanual pasó de ser del 1,17% al 5,01%. (Fig. 1).

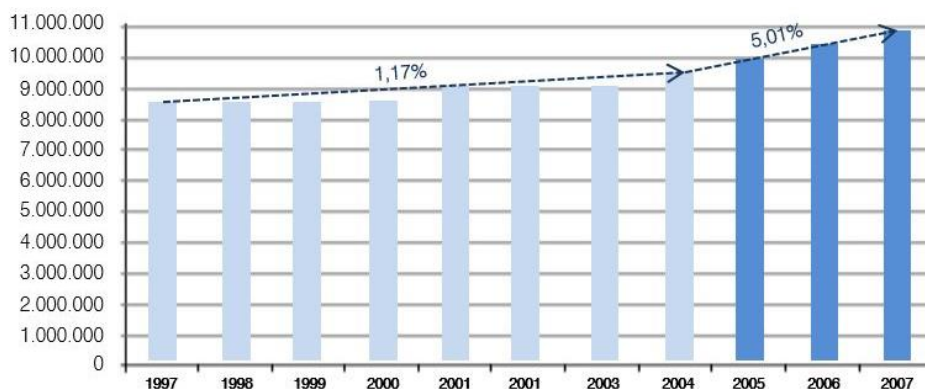


Fig. 1 – Evolución de viajeros en las líneas del área metropolitana de Granada

2.2 Datos

La encuesta de la que provienen los datos utilizados en esta investigación fue realizada a usuarios del servicio de transporte público en autobús que opera en el área metropolitana de Granada (España) en marzo de 2007. El CTMAG llevó a cabo entrevistas cara a cara durante los cinco días de la semana, en las principales paradas de autobús, de modo que finalmente

se recopilaron 858 encuestas.

El cuestionario se estructuró en dos secciones. El objetivo de la primera fue recopilar información relacionada con: información general del viaje (p.ej.e., parada de autobús, línea, operador, origen, destino), características socioeconómicas de los pasajeros (p. eje. sexo, edad, disponibilidad de vehículo privado) y hábitos de viaje (motivo del viaje, frecuencia, tipo de billete, modos complementarios usados por el usuario desde el origen hasta la parada de autobús y desde la parada de autobús hasta el lugar de destino).

La mayoría de los entrevistados fueron mujeres (Tabla 1). Más de la mitad tenían edades comprendidas entre los 18 y los 30 años, y sólo un 9,5% eran mayores de 60 años. La mayoría de los encuestados (61,1%) disponía de vehículo privado para hacer el viaje. Un 29,4% viajaba por motivos de trabajo y otro porcentaje similar por estudios, mientras que el resto de los encuestados lo hacía por otras razones, como visita al médico, compras, vacaciones, etc. La mayoría de ellos usaba el autobús con una frecuencia casi diaria, o con bastante frecuencia, mientras que sólo un 10% de la muestra encuestada lo usaba de forma esporádica. El modo de transporte complementario al autobús y que completaba sus trayectos era a pie. También indicaron otros modos que completaran el trayecto como el coche, el autobús urbano, la motocicleta, etc., pero su porcentaje de uso era mucho menor. En cuanto a la tipología de billete usado, el abono de transporte y el billete simple eran los más usados, representando entre ambos un porcentaje superior al 80%.

Variable	Distribución de respuestas
Sexo	Hombre (33,0%), Mujer (67,0%)
Edad	18-30 (56,5%), 31-60 (34,1%); >60 años (9,5%)
Disponibilidad de vehículo privado	Sí (38,9%), No (61,1%)
Motivo del viaje	Trabajo (29,4%), Estudios (22,9%), Médico (14,2%), Compras (4,4%), Actividades personales (18,7%), Vacaciones (0,2%), Ocio (8,6%), Otros (1,5%)
Frecuencia	Casi diaria (67,9%), Frecuentemente (20,6%), Ocasionalmente (9,0%), Esporádicamente (2,5%)
Modos de transporte complementarios desde el origen hasta la parada de autobús	A pie (77,6%), Coche (1,9%), Autobús urbano (16,9%), Motocicleta (0,5%), Otros (3,1%)
Modos de transporte complementarios desde la parada de autobús hasta el destino	A pie (94,5%), Coche (2,1%), Autobús urbano (2,3%), Motocicleta (0,2%), Otros (0,9%)
Tipo de billete	Abono de transporte (49,6%), Billete sencillo (41,2%), tarjeta de la tercera edad (4,8%), Otros (4,4%)

Tabla 1 – Características de la muestra encuestada

La segunda sección de la encuesta estaba más orientada a captar las opiniones de los usuarios sobre el servicio prestado, es decir, a obtener información sobre la percepción de calidad del

servicio prestado. Para ello, entre otras variables, se definieron 12 atributos de calidad del servicio que lo caracterizaban, y una variable de valoración global del servicio. Los datos de esta sección de encuesta que se utilizaron para realizar el estudio son los correspondientes a la valoración que los usuarios hicieron de cada uno de los atributos que definían la calidad del servicio en fase post-evaluación; cada atributo podía ser valorado por el usuario en una escala de valores del 0 al 10 (Tabla 2). Los atributos utilizados para caracterizar la calidad del servicio fueron: información, puntualidad, seguridad a bordo, amabilidad del conductor, limpieza del interior del autobús, espacio en el autobús, temperatura en el autobús, accesibilidad a/desde el autobús, tarifa, velocidad, frecuencia del servicio y proximidad a/desde el origen/destino.

Atributo	Ratio de percepción media
(INF) Información	6,86
(PUN) Puntualidad	7,41
(SEG) Seguridad a bordo	7,73
(TRA) Amabilidad del conductor	7,96
(LIM) Limpieza del interior del autobús	7,46
(ESP) Espacio en el autobús	7,21
(TEM) Temperatura en el autobús	7,43
(ACC) Accesibilidad a/desde el autobús	6,90
(PRE) Tarifa	6,44
(VEL) Velocidad	7,30
(FRE) Frecuencia	6,99
(PRO) Proximidad desde/a origen/destino	7,43
(SQ) Calidad global del servicio	7,10

Tabla 2 – Ratios de percepción media de los atributos de calidad del servicio

3. MÉTODOS

3.1 El Perceptrón Multicapa (MLP)

Una ANN está compuesta por unidades elementales de procesamiento de la información denominadas neuronas, que están organizadas en capas y que interactúan entre sí a través de los pesos sinápticos, en los que reside el conocimiento de la red (Palmer y Montaña, 2002). Las funciones de activación delimitan la amplitud de los posibles valores de las salidas de cada neurona (Haykin, 1999). Existen muchas tipologías de ANN, de las cuales la utilizada en este estudio es la denominada perceptrón multicapa (MLP) con algoritmo de aprendizaje de descenso por gradiente (BP), porque Funahashi (1989) y Hornik et al. (1989) demostraron, a través de grupos de investigación totalmente independientes, que una red MLP con una única capa oculta adecuadamente entrenada es un aproximador universal de funciones. Además, esta tipología es muy popular (Moghaddam et al., 2010) y es la usada en un mayor número de casos (un 70%) en los que se aplica la técnica de las ANN (Martín del Brío y Sanz, 2006).

La base de datos es dividida aleatoriamente en tres conjuntos: entrenamiento, validación y

test. El conjunto de datos de entrenamiento permite a la red aprender y capturar las relaciones inherentes en las variables de estudio, es decir, se parte de una red vacía de conocimiento con valores iniciales aleatorios de todos los pesos sinápticos, y durante la fase de entrenamiento, la ANN debe modificar iterativamente los valores de sus pesos sinápticos para aprender de la información que se le proporciona por medio de una regla de aprendizaje, hasta que la salida obtenida tiende a ser la deseada. Durante esta fase de entrenamiento, la diferencia entre los valores obtenidos tras presentarle a la red un nuevo caso y los esperados se cuantifica por medio de una función de error, de modo que el entrenamiento avanza conforme se desciende a lo largo de la superficie de error hasta que se alcanza un mínimo que puede ser local o global. Precisamente esa imposibilidad de saber si ese mínimo es local o global imposibilita saber si la ANN tipo MLP entrenada es la que presenta el mejor comportamiento. Además, en el algoritmo se incluyen dos parámetros: el momentum, que es un factor constante que multiplica los valores de los pesos sinápticos durante el entrenamiento, y que tiene la función de mitigar las oscilaciones seguidas por la trayectoria de descenso a lo largo de la superficie de error, mediante la modificación del ritmo de aprendizaje fijado en función del signo de las actualizaciones de los pesos, y el parámetro ratio de aprendizaje, que modifica el ratio de cambio de los pesos sinápticos en cada iteración. Ambos parámetros aceleran la convergencia del entrenamiento (Hagan et al., 1996).

Se aplica una técnica de validación cruzada para evitar que el MLP memorice la información proporcionada por el conjunto de datos de entrenamiento e incurra en sobreaprendizaje. Así, el conjunto de validación evita que se produzca sobreajuste del modelo. En una última fase, el conjunto test mide objetivamente la exactitud del modelo de ANN entrenado o error de generalización (Bishop, 1995; Haykin, 1999).

3.2 Métodos de cálculo de importancia relativa de las variables

Los métodos seleccionados en este estudio para determinar la importancia relativa de las variables en un modelo de ANN, han sido propuestos y aplicados por numerosos autores en diversos campos de investigación (Olden y Jackson, 2002; Gevrey et al., 2003). Éstos han sido: Perturb, Profile, Connection Weights (CW) y Derivadas Parciales (PaD).

3.2.1 Método Perturb

Este método se basa en el principio de perturbar o introducir ruido a una de las entradas mientras que las restantes variables mantienen sus valores originales. Posteriormente, se compara el error cuadrático medio (MSE) entre las salidas obtenidas antes y después de la perturbación (Yao et al., 1998; Scardy y Harding, 1999).

Un ruido δ es progresivamente aplicado a cada variable en cinco pasos: 10%, 20%, 30%, 40% y 50% de su valor original. Así, la variable x_i cambia sus valores a $x_i = x_i + \delta$. Cuanto mayor sea el error medido al introducir las perturbaciones, mayor será la importancia relativa de la variable en ese modelo de red.

3.2.2 Método Profile

Este método analiza la evolución de cada variable de entrada a lo largo de un rango de valores, mientras las variables restantes se quedan bloqueadas en unos valores fijados (Lek et al., 1995; 1996a; 1996b).

Cada variable predictora x_i adopta 11 valores diferentes que son resultado de la división del rango comprendido entre sus valores máximo y mínimo en diez intervalos iguales. Seguidamente, todas las variables excepto una fijan inicialmente sus valores en su mínimo, y sucesivamente en su primer cuartil, mediana, tercer cuartil y valor máximo. En consecuencia, se obtienen 5 valores de la variable respuesta para cada uno de los 11 valores adoptados por x_i , y se calcula la mediana de esos 5 valores. Finalmente, para cada variable se representa una curva con el perfil de variación. Cuanto mayor sea la variación relativa de los valores representados en el perfil para cada variable, mayor será la importancia relativa de esa variable en el modelo estudiado.

3.2.3 Método Connection Weights (CW)

Este método determina la importancia relativa de las variables predictoras del modelo en función de los pesos sinápticos de la ANN, según la expresión matemática (Olden y Jackson, 2002):

$$R_{ik} = \sum_{j=1}^L w_{ij} \cdot w_{jk} \quad (1)$$

Siendo R_{ik} el valor de la importancia relativa de la variable predictora x_i respecto a la variable de la capa de salida k , w_{ij} el peso sináptico que conecta la neurona i de la capa de entrada con la neurona j de la capa oculta, y w_{jk} el peso sináptico que conecta la neurona j anterior con la neurona k de la capa de salida.

3.2.4 Método de Derivadas Parciales (PaD)

Este método analiza los efectos de primer orden de las variables predictoras del modelo respecto a la variable salida, utilizando para ello el conjunto de datos de entrenamiento (Dimopoulos et al., 1995). La salida proporcionada por una neurona de la capa oculta de una ANN tipo MLP con funciones de activación tipo sigmoideo sigue la ecuación:

$$h_j = \frac{1}{(1+e^{-net_j})} \quad (2)$$

$$net_j = \sum_i w_{ij} \cdot x_i - \theta_j \quad (3)$$

donde h_j es la salida de la neurona j de la capa oculta, x_i es el valor de la variable predictora considerada en la capa de entrada, w_{ij} es el peso sináptico que conecta la variables predictora x_i y la neurona j de la capa oculta, y θ_j es el bias de la neurona j de la capa oculta.

La salida de una neurona k de la capa de salida vendrá dada por las siguientes expresiones

$$y_k = \frac{1}{(1+e^{-net_k})} \quad (4)$$

$$net_k = \sum_k w_{jk} \cdot h_j - \theta_k \quad (5)$$

donde w_{jk} es el valor del peso sináptico entre la neurona j de la capa oculta y la neurona k de la capa de salida, θ_k es el bias de la neurona k de la capa de salida e y_k es la salida que arroja la neurona k de la capa de salida.

La expresión que relaciona la variación de los valores de salida y_k respecto a la variación de la variable predictora x_i se obtiene con la aplicación de la regla de la cadena:

$$\frac{dy_k}{dx_i} = \sum_j \frac{dy_k}{dnet_k} \cdot \frac{dnet_k}{dh_j} \cdot \frac{dh_j}{dnet_j} \cdot \frac{dnet_j}{dx_i} = \sum_j f'_k \cdot w_{jk} \cdot f'_j \cdot w_{ij} \quad (6)$$

$$f'_k = y_k \cdot (1 - y_k) \quad (7)$$

$$f'_j = h_j \cdot (1 - h_j) \quad (8)$$

El valor de la sensibilidad de cada variable x_i viene dado por la expresión:

$$L_i = \frac{\sum_P \frac{dy_k}{dx_i}}{P} \quad (9)$$

siendo P el número total de casos de entrenamiento.

Cuanto mayor sea el valor de L_i , mayor será la importancia relativa de la variable x_i en el modelo de ANN entrenado.

3.2 Metodología

En este estudio se ha trabajado con conjuntos de ANN en lugar de con una única. Cada conjunto de ANN está compuesto por una serie de ANN con la misma arquitectura, que son entrenadas utilizando idénticos algoritmos de aprendizaje, con una partición aleatoria de los datos para cada entrenamiento en conjuntos de entrenamiento, validación y test en una proporción 70/15/15, funciones de activación (tipo logaritmo sigmoideo), valores de momentum de 0,9 y ratio de aprendizaje de 0,1 (Hagan et al., 1996) y 20.000 iteraciones. La única diferencia entre cada una de las ANN incluidas en el mismo conjunto reside en los valores iniciales de los pesos con los que comienza su entrenamiento, ya que en cada una de ellas estos valores se obtienen de forma aleatoria. Una vez que cada una de las ANN ha sido entrenada, se aplican los cuatro métodos (CW, Perturb, Profile y PaD) descritos en el apartado 3.1, obteniéndose un ranking de importancia relativa de las variables consideradas

por el modelo para cada uno de los métodos. Debido a la inestabilidad de los resultados de importancia relativa obtenidos cuando se aplica un único método a una única red (Cao y Qiao, 2008), este estudio ha desarrollado un procedimiento basado en el cálculo del ranking de importancia relativa para cada método en función de los valores de importancia media obtenidos a partir de cada conjunto de ANN. El software utilizado ha sido MATLAB (Beale et al., 2007).

La secuencia de pasos seguidos para desarrollar esta metodología ha sido:

1. Entrenamiento de cada una de las H arquitecturas de ANN consideradas, con 12 neuronas en la capa de entrada (una por cada atributo de calidad del servicio considerado), $H \in [1, N]$ neuronas en la capa oculta, y 1 neurona en la capa de salida (la calidad global del servicio) (Fig. 2) un número M de veces, y usando diferentes valores aleatorios para los pesos iniciales. Así, se obtienen NxM ANN entrenadas. En este estudio se consideraron unos valores $N=30$ y $M=50$, resultando un total de 1.500 ANN entrenadas.

2. Cálculo de la exactitud o capacidad de generalización de las NxM ANN entrenadas, a través de la fórmula del MAPE (Delen et al., 2006):

$$MAPE = \frac{1}{T} \cdot \sum_{i=1}^T \text{abs}\left(\frac{\text{Valor actual}_i - \text{Valor objetivo}_i}{\text{Valor objetivo}_i}\right) \quad (10)$$

Siendo T el número de casos incluidos en el grupo test.

3. Cálculo del valor mínimo del MAPE para cada una de las arquitecturas (o conjuntos) de ANN, y selección del conjunto que obtiene el mínimo valor global de MAPE.

4. Aplicación de los cuatro métodos de cálculo de importancia relativa a cada una de las M ANN de la arquitectura (conjunto) seleccionado en el paso anterior. (paso 3).

5. Cálculo del valor medio de importancia relativa para cada variable de cada una de las M ANN pertenecientes a la arquitectura considerada.

6. Determinación del ranking de importancia de las 12 variables consideradas en el estudio, en función de los valores obtenidos en el paso anterior para cada uno de los métodos aplicados.

7. Comparación de los resultados obtenidos según cada método.

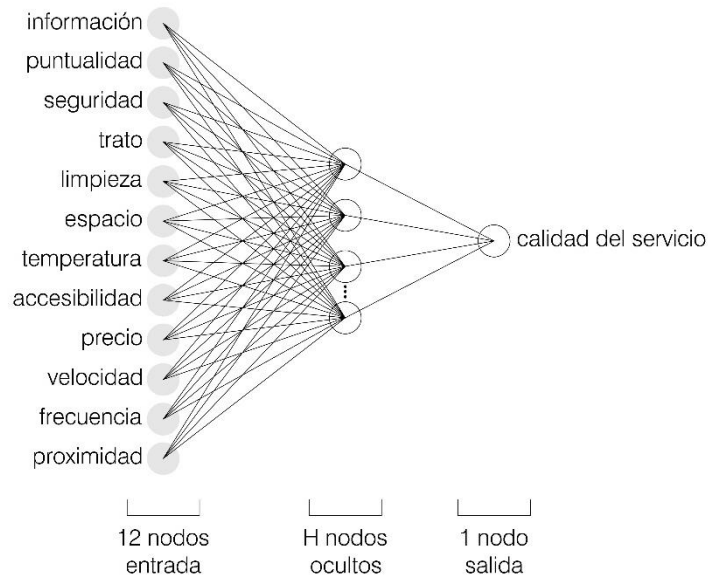


Fig. 2 – Arquitectura de los conjuntos de ANN tipo MLP desarrollados

4. RESULTADOS

4.1 Exactitud de los modelos de ANN entrenados

Para medir la exactitud y capacidad de generalización de los modelos de ANN calibrados en cada conjunto H, se calculó el valor del MAPE para cada una de ellas, obteniéndose el rango de valores para cada uno de los conjuntos que se recogen en la Tabla 3.

Todas las ANN entrenadas alcanzaron una exactitud muy elevada, superior al 90% en todos los casos. Estos resultados concuerdan con los obtenidos por otros autores que han aplicado la técnica de las ANN para evaluar la contribución relativa de las variables que definen la calidad del servicio en otros campos no relacionados con el transporte, como el sector de la educación (Mahapatra y Khan, 2006), el sector servicios (Lin, 2007; Deng et al., 2008; Deng y Pei, 2009; Larasati et al., 2012), o el grado de satisfacción en el trabajo percibido por los empleados de varias compañías de refinería de gas (Azadeh et al., 2011). En todos estos estudios los modelos de ANN alcanzaron una exactitud muy elevada. Estos resultados, junto con el obtenido en esta investigación, refuerzan la idea defendida por Garver (2002) de que las ANN son una técnica adecuada para evaluar la importancia relativa de las variables que definen la satisfacción de los clientes en cualquier sector, incluido por tanto el del transporte. Además, la exactitud alcanzada por todas las ANN entrenadas en este estudio es significativamente superior a la obtenida del estudio elaborado por de Oña et al. (2012), en el que se utilizaron árboles de decisión para analizar la contribución relativa de los atributos de calidad del servicio en el ámbito del transporte público, que estuvo comprendida entre el 59,72% y el 62,16%.

Aunque no existen grandes diferencias de exactitud en las arquitecturas entrenadas, aquella con H=6 neuronas en la capa oculta presenta menores valores mínimo y medio de MAPE (Tabla 3), así que fue la seleccionada para desarrollar la siguiente fase de este estudio.

H	Mínimo	Media	Desviación estándar	Máximo
1	0,0400	0,0531	0,0082	0,0874
2	0,0401	0,0532	0,0090	0,0943
3	0,0438	0,0525	0,0045	0,0679
4	0,0372	0,0519	0,0067	0,0668
5	0,0424	0,0528	0,0062	0,0765
6	0,0325	0,0475	0,0064	0,0580
7	0,0417	0,0514	0,0052	0,0643
8	0,0437	0,0520	0,0044	0,0858
9	0,0436	0,0513	0,0036	0,0620
10	0,0368	0,0530	0,0074	0,0744
11	0,0399	0,0505	0,0051	0,0674
12	0,0384	0,0520	0,0086	0,0938
13	0,0396	0,0526	0,0106	0,0997
14	0,0427	0,0517	0,0043	0,0616
15	0,0354	0,0526	0,0065	0,0706
16	0,0365	0,0514	0,0087	0,0853
17	0,0379	0,0513	0,0056	0,0610
18	0,0416	0,0513	0,0070	0,0761
19	0,0406	0,0513	0,0046	0,0600
20	0,0419	0,0516	0,0049	0,0635
21	0,0395	0,0508	0,0056	0,0611
22	0,0346	0,0524	0,0078	0,0863
23	0,0424	0,0518	0,0700	0,0742
24	0,0393	0,0497	0,0047	0,0591
25	0,0385	0,0508	0,0067	0,0715
26	0,0405	0,0524	0,0063	0,0708
27	0,0417	0,0520	0,0078	0,0923
28	0,0375	0,0507	0,0056	0,0633
29	0,0427	0,0534	0,0075	0,0789
30	0,0413	0,0517	0,0091	0,0890

Tabla 3 – Rango de valores de MAPE para los conjuntos de ANN entrenados

4.2 Importancia relativa de las variables

Los valores de importancia relativa obtenidos por cada uno de los métodos fueron escalados en el rango [0,100], es decir, tras conocer el rango de valores de importancia relativa obtenido para todas las variables de un método, se le asignó un valor de 100 al mayor de ellos, mientras que la importancia relativa de las restantes variables fueron escaladas en función del rango de valores de dicho método. Esta transformación es necesaria para comparar y analizar adecuadamente los resultados entre métodos, ya que el rango de valores de importancia relativa obtenido por cada método difiere de los demás incluso en varios órdenes de magnitud.

Los diagramas de cajas representados en las Fig. 3 a 6 muestran las importancias relativas

de las 12 variables independientes obtenidas tras aplicar cada uno de los cuatro métodos seleccionados para este estudio (Perturb, Profile, CW y PaD) a cada una de las 50 ANN pertenecientes al conjunto con H=6 neuronas en su capa oculta. Las figuras muestran que el rango de valores de importancia relativa de cada variable predictora es muy amplio, no sólo tras aplicar diferentes métodos, sino al aplicar el mismo método a dos ANN distintas entrenadas partiendo de pesos iniciales aleatorios diferentes. Algunos autores (Zhou et al., 2002; Cao and Qiao, 2008) ya pusieron de manifiesto este problema, y sugirieron trabajar con conjuntos de ANN para mitigar los problemas de estabilidad derivados de los análisis de sensibilidad. Estas propuestas han reforzado la idea desarrollada en esta investigación de trabajar, no sólo con una única ANN, sino con un conjunto de ellas.

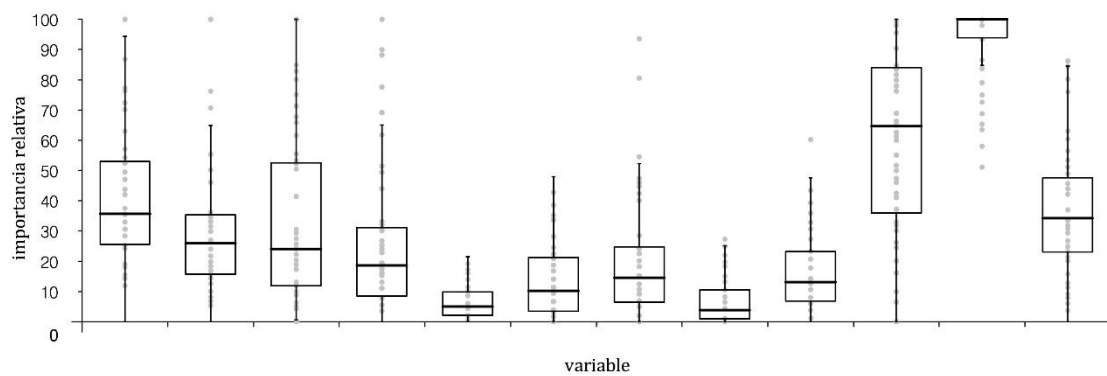


Fig. 3 – Diagrama de cajas de importancia relativa de las variables predictoras según el método PERTURB

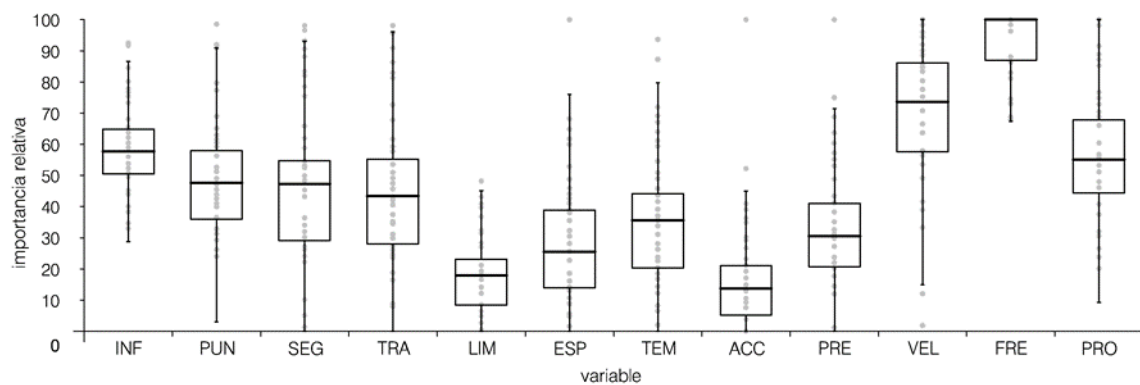


Fig. 4 – Diagrama de cajas de importancia relativa de las variables predictoras según el método PERTURB

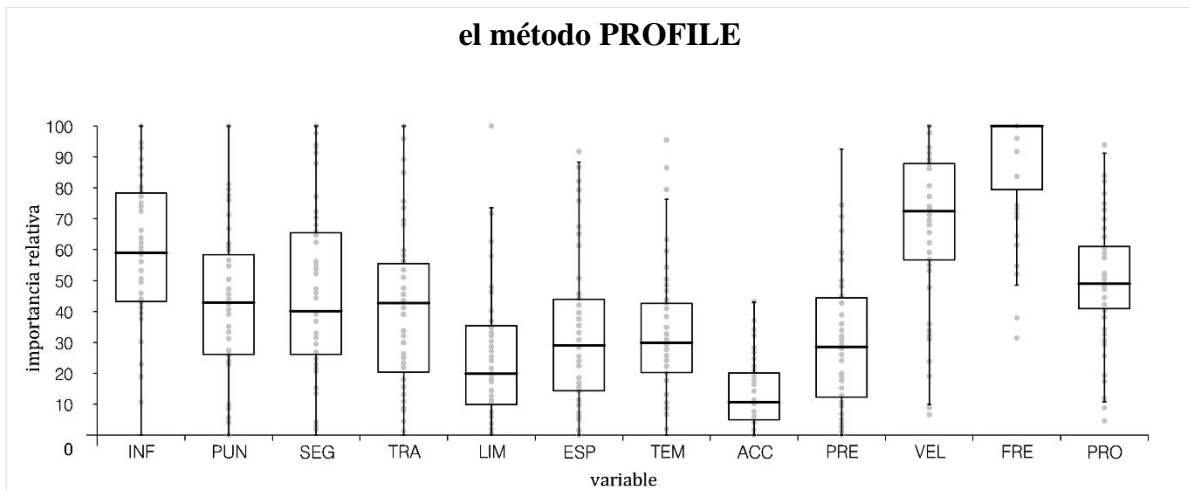


Fig. 5 – Diagrama de cajas de importancia relativa de las variables predictoras según el método CW

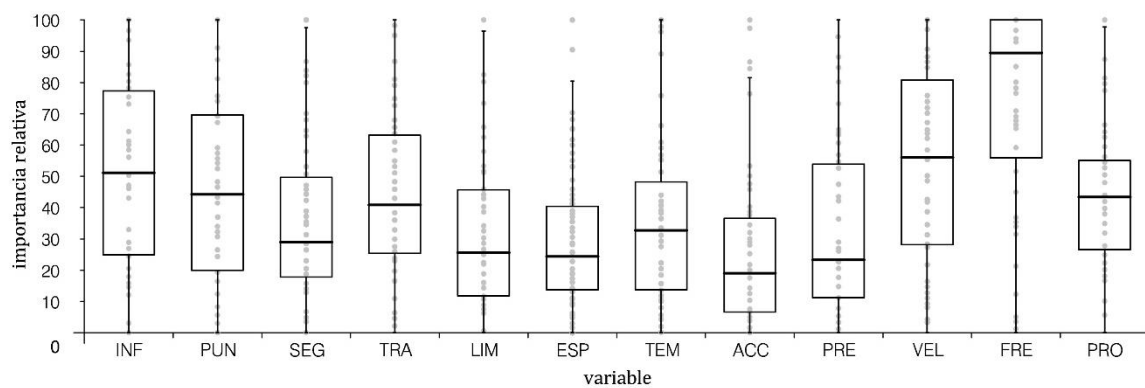


Fig. 6 – Diagrama de cajas de importancia relativa de las variables predictoras según el método PaD

Los métodos Profile y Perturb proporcionan rangos de valores de importancia relativa más restringidos para ciertas variables, como la FRECUENCIA, que siempre tiene una importancia relativa superior al 70% según el método Profile, y superior al 55% según el método Perturb; o la variable LIMPIEZA, cuya importancia relativa siempre está por debajo del 50% según el método Profile, y por debajo del 20% según el método Perturb.

Los resultados de los otros dos métodos, especialmente del PaD, muestran que cualquier variable puede alcanzar cualquier valor de importancia relativa comprendido entre el 0% y el 100%. En cualquier caso, el ranking de importancia obtenido tras aplicar los cuatro métodos para todas las variables es demasiado impredecible, lo que supone una limitación muy importante para determinar de forma adecuada un valor de importancia relativa y un orden de magnitud del mismo. Sin embargo, si se calculan los valores medios de importancia relativa obtenidos para cada variable por cada uno de los métodos, los resultados obtenidos son mucho más homogéneos.

En el método Profile, con los resultados de importancia relativa obtenidos en cada una de las N=50 ANN pertenecientes al conjunto con H=6 neuronas en la capa oculta, para cada una de las variables predictoras de la calidad del servicio, se generó un rango de 50 perfiles de variación en el intervalo [0,1]. Así, la importancia relativa de cada una de las variables viene determinada por la diferencia entre los valores máximo y mínimo (diferencia en el eje de ordenadas) de la línea que representa la media de valores del perfil de variación. Los resultados obtenidos indican que la variable FRECUENCIA alcanza la importancia relativa más elevada (100,00%), seguida de la VELOCIDAD (77,72%), la INFORMACIÓN (64,15%) y la PROXIMIDAD (60,24%), por lo que estas variables tienen una importancia global en el modelo elevada. Un segundo nivel de importancia, considerado como medio-alto, puede ser asignado a las variables PUNTUALIDAD (54,45%), SEGURIDAD (53,28%) y TRATO (48,59%), seguido por un nivel de importancia media que contiene a las variables TEMPERATURA (38,44%), TARIFA (36,40%) y ESPACIO (27,22%). En el último escalón de importancia relativa quedarían las variables ACCESIBILIDAD (17,34%) y LIMPIEZA (3,36%).

Tras aplicar el método Perturb, también se obtienen 50 perfiles de variación para cada variable, que representan la variación del error MSE, en función de cómo afecta el porcentaje de ruido o perturbación introducida a la capacidad de respuesta de la ANN. De nuevo, la diferencia entre los valores máximo y mínimo de la línea que representa los valores medios de cada perfil, indica la importancia relativa de las variables, de modo que cuanto mayor sea esta diferencia, mayor será la importancia relativa de la variable. Así, la FRECUENCIA (100,00%) es la variable globalmente más importante, seguida de la VELOCIDAD (63,60%). Las variables INFORMACIÓN (42,88%), PUNTUALIDAD (32,92%), SEGURIDAD (32,92%), TRATO (30,33%), PROXIMIDAD (23,28%) y TEMPERATURA (22,23%) se sitúan en un segundo nivel de importancia media. En las últimas posiciones del ranking de importancia quedan las variables PRECIO (17,90%), ESPACIO (14,51%), ACCESIBILIDAD (7,58%) y LIMPIEZA (7,12%).

Según el método CW, la variable con la mayor importancia relativa es la FRECUENCIA (100,00%), seguida de la VELOCIDAD (75,98%) y de la INFORMACIÓN (66,68%). En un segundo nivel de importancia, quedarían incluidas las variables PROXIMIDAD (55,49%), SEGURIDAD (51,38%), PUNTUALIDAD (51,35%) y TRATO (47,81%), seguido por el grupo de importancia media que contiene a las variables TEMPERATURA (36,65%), ESPACIO (36,45%), PRECIO (31,98%) y LIMPIEZA (27,39%). La variable con menor importancia relativa es la ACCESIBILIDAD (14,56%).

El método PaD considera como variables más importantes la FRECUENCIA (100,00%), la VELOCIDAD (72,90%), la INFORMACIÓN (71,01%), la PUNTUALIDAD (64,13%), el TRATO (60,67%) y la PROXIMIDAD (58,30%). La importancia relativa del resto de variables puede considerarse como media: TEMPERATURA (49,05%), SEGURIDAD (48,08%), PRECIO (45,43%), LIMPIEZA (44,01%), ESPACIO (43,35%) y

ACCESIBILIDAD (37,41%).

La Fig. 7 muestra una comparación del ranking de importancia relativa de las 12 variables consideradas, que ha sido determinada por cada método. Esta figura refleja que las posiciones de importancia relativa son muy similares, y que por tanto aplicando la metodología propuesta en esta investigación se mitiga considerablemente la elevada inestabilidad inherente en estos métodos cuando se aplican a una única ANN. Los cuatro métodos coinciden en que las variables más influyentes son la FRECUENCIA, VELOCIDAD e INFORMACIÓN, y en que las variables menos influyentes son la LIMPIEZA, ACCESIBILIDAD y ESPACIO, con la única excepción del método CW, que considera a las variables ACCESIBILIDAD, LIMPIEZA y PRECIO como las menos importantes, seguidas de la variable ESPACIO. Las seis variables restantes ocupan posiciones intermedias en el ranking. Los métodos Profile y Perturb proporcionan una jerarquía de importancia similar, mientras que el método PaD es el que presenta mayores discrepancias.

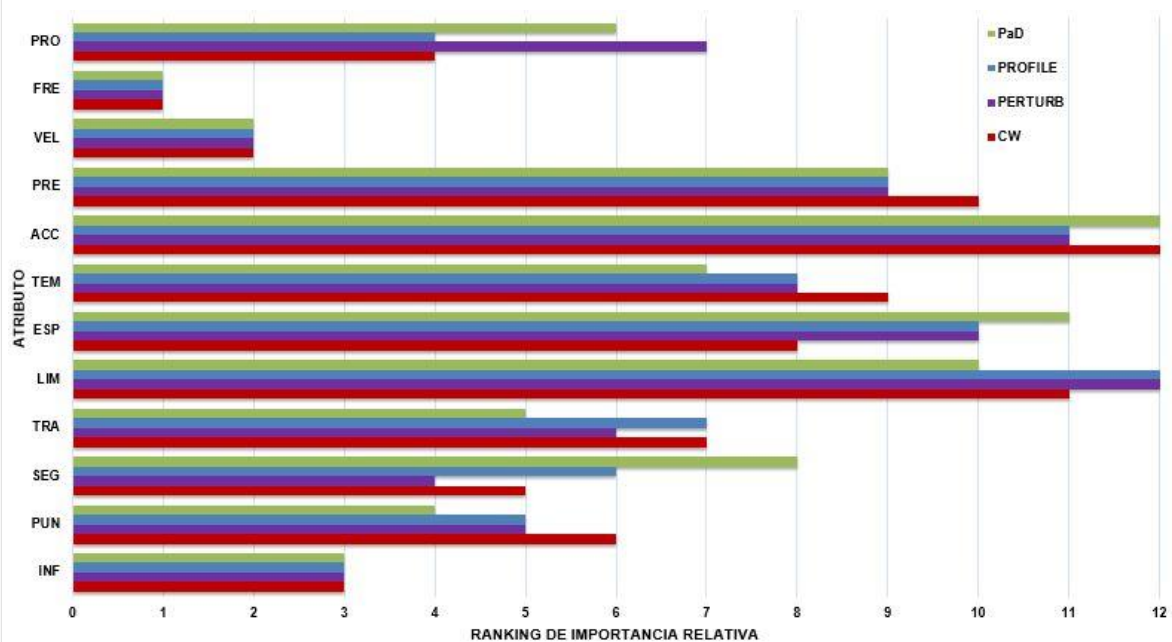


Fig. 7 – Ranking de importancia relativa de los atributos de calidad del servicio según los métodos PROFILE, PERTURB, CW y PaD

Respecto al grado de importancia relativa asignado a cada variable y expresado en tanto por ciento, los métodos Profile y Perturb permiten establecer cuatro escalones de importancia: muy alta, alta, media y baja. El método CW establece tres niveles claros de importancia: alta, media y baja, y el método PaD establece unos porcentajes de influencia que pueden diferenciarse en dos niveles de importancia: alto y medio, ya que el menor valor de importancia relativa queda por encima del 37%.

5. CONCLUSIONES

En este estudio se ha aplicado la técnica de ANN para analizar la calidad del servicio de transporte público por autobús. Esta metodología, aplicada por primera vez a este campo concreto del transporte, permite predecir la calidad global del servicio, así como determinar la influencia de los atributos o variables características que se han utilizado para su definición, lo que proporciona a los operadores de transporte público una nueva herramienta para analizar las fortalezas y debilidades que ofrecen al usuario a través del servicio prestado, lo que a su vez permite plantear mejoras concretas en aquellos aspectos que son detectados como mejorables o mal valorados. Con ello, podrá ofrecerse un mejor servicio a los actuales pasajeros y mejorar la competitividad de este modo de transporte frente a otros.

Esta investigación, al igual que otros estudios previos (Mussone et al., 1999; Delen et al., 2006; Mohammadipour y Alavi, 2009; Moghaddam et al., 2010; Akin y Akbaç, 2010), confirma la elevada capacidad de predicción y generalización de la técnica de las ANN, puesto que se obtiene un error de generalización, medido a través del MAPE, del orden de 0,05, lo que indica que los modelos de ANN entrenados son capaces de proporcionar al menos un 95% de salidas correctas.

Por otra parte, se aborda el problema de la inestabilidad que presentan los métodos de cálculo de la importancia relativa de las variables predictoras consideradas en los modelos ANN, ya expuesto previamente por varios autores (Gedeon, 1997; Sung 1998; Palmer y Montaña, 2002; Olden y Jackson, 2002; Gevrey et al., 2003; Paliwal y Kumar, 2011). Hasta el momento no ha sido posible establecer un consenso sobre qué método es más estable, exacto o robusto, ni cuál proporciona resultados más ajustados a la realidad, lo que ha sembrado la duda de si estos métodos, que determinan la relación causa-efecto entre cada variable predictora – variable dependiente, son realmente válidos para clarificar la función que cada variable representa en un modelo de ANN (también denominado modelo de caja negra), y aún más importante, si es realmente útil aplicar esta técnica en aquellos campos de investigación en los que los resultados obtenidos pueden tener consecuencias serias y comprometedoras. Los valores de importancia relativa obtenidos tras aplicar los métodos Perturb, Profile, CW y PaD a una única ANN, muestran una variabilidad muy elevada, no sólo cuando se aplican diferentes métodos, sino cuando se aplica el mismo método a una misma arquitectura de ANN entrenada partiendo de diferentes valores de pesos iniciales aleatorios. Por tanto, trabajando de esta forma no existe garantía de que los resultados de importancia relativa obtenidos sean válidos.

Sin embargo, en este estudio se ha trabajado con conjuntos de ANN con la misma arquitectura sobre los que se han aplicado diferentes métodos de cálculo de la importancia relativa, extrayéndose valores medios. Este procedimiento o “modus operandi” consigue obtener resultados de importancia relativa similares, independientemente del método

utilizado, especialmente similares para los métodos Perturb y Profile.

Siguiendo esta metodología, cualquiera de los métodos de cálculo de la importancia relativa es válido, ya que todos ellos coinciden en identificar las variables más importantes y las menos importantes, aunque el método PaD es el menos recomendado de los cuatro analizados, puesto que muestra una mayor variabilidad en los valores de importancia relativa extraídos. Además, otro argumento que refuerza la robustez de esta nueva metodología es el hecho de que las variables FRECUENCIA, VELOCIDAD, PUNTUALIDAD y PROXIMIDAD son clasificadas como las más importantes tras aplicar los cuatro métodos, lo que concuerda con los resultados obtenidos por otros autores (Eboli y Mazulla, 2008; Eboli y Mazulla 2010; Dell’Olio et al.; 2010; Dell’Olio et al.; 2011; De Oña et al., 2012) que han aplicado otras técnicas, como modelos de regresión multinomial logística, modelos de elección discreta, modelos probit, árboles de decisión o ecuaciones estructurales.

Otro escollo que salva esta la metodología de cálculo de importancia relativa de las variables es la dificultad de seleccionar una ANN o conjunto de ANN óptimo o sub- óptimo, ya que, como es bien conocido en la literatura científica a través de los trabajos desarrollados por diversos autores (Cao y Qiao, 2008; Paliwal y Kumar, 2011), existen numerosas posibilidades de entrenamiento y de obtención de resultados diferentes de los modelos de ANN entrenados, sea cual sea su tipología, con lo que es muy complicado afirmar que en una determinada investigación se ha encontrado y entrenado el mejor modelo. Con este nuevo procedimiento, se ha demostrado que al trabajar con conjuntos de ANN, las oscilaciones en cuanto al porcentaje de importancia relativa de las variables predictoras del modelo se mitiga notablemente, y se pueden conseguir resultados parecidos. La principal debilidad de esta metodología es que, aunque los cuatro métodos proporcionan un ranking de importancia similar, los porcentajes de importancia relativa difieren fundamentalmente en las variables de importancia media, siendo necesaria por tanto la opinión de un experto para decidir qué método ofrece unos resultados más similares a los esperados. En general, el método Perturb asigna una importancia relativa más baja a las variables, mientras que los tres métodos restantes presentan valores más elevados y similares entre sí.

AGRADECIMIENTOS

Al Departamento de Ingeniería Civil de la ETSICCP de Granada y al Consorcio de Transportes del Área Metropolitana de Granada por proporcionar los datos necesarios para la elaboración de esta investigación.

REFERENCIAS

- ABDELWAHAB, H.T. y ABDEL-ATY, M.A. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record* 1746, pp. 6–13.
- AKIN, D. y AKBAÇ, B. (2010). A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics. *Scientific Research*

and Essays 5(19), pp. 2837-2847.

AKSOY, S., ATILGAN, E. y AKINCI, S. (2003). Airline services marketing by domestic and foreign firms: differences from the customers' viewpoint. *Journal of Air Transport Management* 9, pp. 343-351.

ANDREASSEN, T.W. (1995). Satisfaction with public services: The case of public transportation. *J. Services Marketing* 9(5), pp. 30-41.

ARMSTRONG, G., y KOTLER, P. (2000). *Marketing: An introduction*. New York: Prentice-Hall.

AZADEH, A., ROUZBAHMAN, M., SABERI, M. y FAM, I. M. (2011). An adaptative neural network algorithm for assessment and improvement of job satisfaction with respect to HSE and ergonomics program: the case of a gas refinery. *Journal of Loss Prevention in the Process Industries* 24, pp. 361-370.

BEALE, M.H., HAGAN, M.T. y DEMUTH, H.B. (2007). *Neural Network Toolbox 7. User's Guide*. MathWorks, Inc. 3 Apple Hill Drive Natic, MA 01760-2098.

BEJOU, D., WRAY, B. y INGRAM, T.N. (1996). Determinants of relationships quality: an artificial neural network analysis. *Journal of Business Research* 36, pp. 137-143.

BISHOP, C.M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

BULLOCK, D., GARRETT, J. y HENDRICKSON, C. (1993) A neural network of image-based vehicle detection. *Transportation Research Part C* 1, pp. 235-247.

CAO, M. y QIAO, P. (2008). Neural network committee-based sensitivity analysis strategy for geotechnical engineering problems. *Neural Computing and Applications* 17, pp. 509-519.

CASTILLO, J.M. y BENÍTEZ, F.G. (2012). Determining a public transport satisfaction index from user surveys. *Transportmetrica A* 9(8), pp.713-741.

COSTA, A. y MARKELLOS, R.N. (1997). Evaluating public transport efficiency with neural network models. *Transportation Research, part C: Emerging Technologies*, pp. 301-312.

DELEN, D., SHARDA, R. y BESSONOV, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis and Prevention* 38, pp. 434-444.

DENG, W-J., CHEN, W-C. y PEI, W. (2008). Backpropagation neural network based importance-performance analysis for determining critical service attributes. *Expert Systems with Applications* 34, pp. 1115-1125.

DENG, W-J. y PEI, W. (2009). Fuzzy neural based importance-performance analysis for determining critical services attributes. *Experts Systems with Applications* 36, pp. 3774-3784.

DE OÑA, J., DE OÑA, R. y CALVO, F. J. (2012). A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications* 39, pp. 11164-11171.

DE OÑA, J., DE OÑA, R. y GARRIDO, C. (2016). Extraction of attribute importance from satisfaction surveys with data mining techniques: a comparison between neural networks and decision trees. *Transportation Letters*. DOI: 10.1080/19427867.2015.1136917.

DE OÑA, J. y DE OÑA, R. (2015). Quality of service in public transport based on customer

- satisfaction surveys: a review and assessment of methodological approaches. *Transportation Science* 49(3), pp. 605-622.
- DE OÑA, J. y GARRIDO, C. (2014). Extracting the contribution of independent variables in neural network models: a new approach to handle instability. *Neural Computing and Applications* 25 (3-4), pp. 859-869.
- DEL HOYO, R., MARTÍN DEL BRÍO, B., MEDRANO, N. y FERNÁNDEZ-NAVAJAS, J. (2009). Computational intelligence tools for next generation quality of service management. *Neurocomputing* 72, pp. 3631-3639.
- DELL'OLIO, L., IBEAS, A. y CECÍN, P. (2010). Modelling user perception of bus transit quality. *Transport Policy* 17(6), pp. 388-397.
- DELL'OLIO, L., IBEAS, A. y CECÍN, P. (2011). The quality of service desired by public transport users. *Transport Policy* 18(1), pp. 217-227.
- DENG, W-J., CHEN, W-C. y PEI, W. (2008). Backpropagation neural network based importance-performance analysis for determining critical service attributes. *Expert Systems with Applications* 34, pp. 1115-1125.
- DENG, W-J. y PEI, W. (2009). Fuzzy neural based importance-performance analysis for determining critical services attributes. *Experts Systems with Applications* 36, pp. 3774-3784.
- DIMOPOULOS, Y., BOURRET, P. y LEK, S. (1995). Use of some sensitivity criteria for choosing networks with good generalization ability. *Neural Processing Letters* 2, pp. 1-4.
- DOUGHERTY, M. (1995). A review of neural networks applied to transport. *Transportation Research, Part C* 3(4), pp. 247-260.
- EBOLI, L. y MAZZULLA, G. (2007). Service quality attributes affecting customer satisfaction for bus transit. *Journal of Public Transportation* 10(3), pp. 21-34.
- EBOLI, L. y MAZZULLA, G. (2008). A Stated Preference Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology* 31(5), pp. 509-523.
- EBOLI, L. y MAZZULLA, G. (2010). How to capture the passengers' point of view on a transit service through rating and choice opinions. *Transport Review* 30, pp. 435-450.
- FRIMAN, M., EDVARDSSON, B. y GÄRLING, T. (2001). Frequency of negative critical incidents and satisfaction with public transport services, I. *Journal of Retailing and Consumer Service* 8(2), pp. 95-104.
- FRIMAN, M. y GÄRLING, T. (2001). Frequency of negative critical incidents and satisfaction with public transport services, II. *Journal of Retailing and Consumer Services* 8(2), pp. 105-114.
- FUNAHASHI, K.I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks* 2, pp. 183-192.
- GARRIDO, C., DE OÑA, R. y DE OÑA, J. (2014). Neural networks for analyzing service quality in public transportation. *Expert Systems with Applications* 41(15), pp. 6830-6838.
- GARVER, M.S. (2002). Using data mining for customer satisfaction research. *Marketing research* 14(1), pp. 8-12.
- GEDEON, T.D. (1997). Data mining of inputs: analyzing magnitude of functional measures.

International Journal of Neural Systems 8(2), pp. 209-218.

GEVREY, M., DIMOPOULOS, I. y LEK, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling* 160, pp. 249-264.

HAGAN, M.T., DEMUTH, H.B. y BEALE, M.H. (1996). *Neural network design*. Campus Publishing Service. Colorado University Bookstore. ISBN 0-9717321-0-8.

HAYKIN, S. (1999). *Neural networks. A comprehensive Foundation*. 2nd Edition. Pentence-Hall.

HENSER, D.A., MULLEY, C. y YAHYA, N. (2010). Passenger experience with quality-enhanced bus service: the Tyne and Wear “superoute” services. *Transportation* 37(2), pp. 239-256.

HORNIK, K., STICHCORBE, M. y WHITE, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks* 2, pp. 359-366.

HUNT, J. G. y LYONS, G. D. (1993). *Modelling dual carriageway lane changing using neural networks*. Informing Technologies for Construction Civil Engineering and Transport, Brunei.

HUSE, C. y EVANGELHO, F. (2007). Investigating business traveler heterogeneity: low-cost vs full-service airline users?. *Transportation Research Part E* 43, pp. 259-268.

JOEWONO, T.B. y KUBOTA, H. (2007). Exploring negative experiences and user loyalty in paratransit. *Transportation Research Record* 2034, pp. 134-142.

KIM, Y.K. y LEE, H.R. (2011). Customer satisfaction using low cost carriers. *Tourism Management* 32(2), pp. 235-243.

LAI, W. y CHEN, C. (2011). Behavioral intentions of public transit passengers, the roles of service quality, perceived value, satisfaction and involvement. *Transport Policy* 18(2), pp. 318-325.

LARASATI, A., DE YONG, C. y SLEVITCH, L. (2012). The application of neural network and logistics regression models on predicting customer satisfaction in a student-operated restaurant. *Procedia Social and Behavioral Sciences* 65, pp. 94-99.

LEE, CH., RAN, B., YANG, F. y LOH, W.Y. (2010). A hybrid tree approach to modeling alternate route choice behavior with online information. *Journal of Intelligent Transportation Systems* 14(4), pp. 209–219.

LEK, S., BELAND, A., DIMOPOULOS, I., LAUGA, J. y MOREAU, J. (1995). Improved estimation, using neural networks, of the food consumption of fish populations. *Marine and Freshwater Research* 46(8), pp. 1229-1236.

LEK, S., DELACOSTE, M., BARAN, P., DIMOPOULOS, I., LAUGA, J. y AULAGNIER, S. (1996a). Application of neural networks to modeling nonlinear relationships in ecology. *Ecological Modelling* 90, pp. 39-52.

LEK, S., BELAND, A., BARAN, P., DIMOPOULOS, I. y DELACOSTE, M. (1996b). Role of some environmental variables in trout abundance models using neural networks. *Aquatic Living Resources* 9, pp. 23-29.

LIN, W-B. (2007). An empirical of service quality model from the viewpoint of management. *Expert Systems with Applications* 32, pp. 364-375.

- LYONS, G. y HUNT, J. (1993). *Traffic modelling-a role for neural networks?*. Proceedings Third International Conference on the Application of Artificial Intelligence to Civil and Structural Engineering, Edinburgh, U.K.
- MAHAPATRA, S.S. y KHAN, M.S. (2006). *A methodology for evaluation of service quality using neural networks*. Presented at the International Conference on Global Manufacturing and Innovation, Rourkela.
- MARTÍN DEL BRÍO, B. y SANZ, A. (2006). *Redes neuronales y sistemas borrosos*. Editorial Ra-Ma.
- MINSER, J. y WEBB, V. (2010). Quantifying the benefits: application of customer loyalty modeling in public transportation context. *Transportation Research Record* 2144, pp. 111-120.
- MOHAMMADIPOUR, A.H. y ALAVI, S.H. (2009). The optimization of the geometric cross-section dimensions of raised pedestrian crosswalks: a case of study in Qazvin. *Accident Analysis and Prevention* 41, pp. 314-326.
- MOGHADDAM, F.R., AFANDIZADEH, S. y ZIYADI, M. (2010). Prediction of accident severity using artificial neural networks. *International Journal of Civil Ingeniering* 9(1), pp. 41-49.
- MORATO, D., ARACIL, J. y DÍEZ, L.A. (2001). On linear prediction of Internet traffic for packet and burst switching network. *Proceedings of 10th International Conference on Computer Communications Networks*, pp. 138-143.
- MOUTNHO, L., DAVIES, F. y CURRY, B. (1996). The impact of gender on car buyer satisfaction and loyalty. *Journal of Retailing and Customer Services* 3(3), pp. 135-144.
- MUSSONE, L., FERRARI, A. y ONETA, M. (1999). An analysis of urban collisions using an artificial intelligence model. *Accident Analysis and Prevention* 31, pp. 705-718.
- OLDEN, J.D. y JACKSON, D.A. (2002). Illuminating the “black-box”: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling* 154, pp. 135-150.
- PALIWAL, M. y KUMAR, U.A. (2011). Assessing the contribution of variables in feed forward neural network. *Applied Software Computing* 11, pp. 3690-3696.
- PALMER, A. y MONTAÑO, J.J., (2002). Redes neuronales artificiales: abriendo la caja negra. *Metodología de las ciencias del comportamiento* 4(1), pp. 77-93.
- PANT, P. D. y BALAKRISHMAN, P. (1994). Neural network for gap acceptance at stop-controlled intersections. *Journal of Transportation Engineering* 120(3), pp. 432-446.
- SCARDI, M. y HARDING, L.W. (1999). Developing an empirical model of phytoplankton primary production: a neural networks case study. *Ecological Modelling* 120(2-3), pp. 213-223.
- SUNG, A.H. (1998). Ranking importance of input parameters of neural networks. *Expert Systems with Applications* 15, pp. 405-411.
- TYRINOPOULOS, Y. y ANTONIOU, C. (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy* 15(4), pp. 260–272.
- TYRINOPOULOS, Y. y AIFADOPOULOU, G. (2008). A complete methodology for the quality control of passenger services in the public transport business. *European Transport*

38, pp. 1-16.

WEITZ, B. A. y JAP, S.D. (1995). Relationship marketing and distribution channels. *Journal of the Academy of Marketing Science* 23(4), pp. 305–320.

XIE, C., LU, J. y PARKANY, E. (2003). Work travel mode choice modelling with data mining. *Transportation Research Record: Journal of the Transportation Research Board* 1854, pp. 50-61.

XIONG, Y. y SCHNEIDER, J. B. (1992). Transportation network design using a cumulative genetic algorithm and neural network. *Transportation Research Record* 1364, pp. 31-44.

YAO, J., TENG, N., POH, H.L. y TAN, C.L., (1998). Forecasting and analysis of marketing data using neural networks. *Journal of Information Science and Engineering* 14, pp. 843-862.

ZHOU, Z.H., WU, J. y TANG, W. (2002). Ensembling neural networks: many could be better than all. *Artificial Intelligence* 137(1-2), pp. 239-263.