

ESTIMACIÓN DEL PRECIO DE LA VIVIENDA MEDIANTE REDES NEURONALES ARTIFICIALES (RNA) EN DIFERENTES MARCOS TEMPORALES

Julia M. Núñez Tabales⁽¹⁾; Jose M^a Caridad y Ocerin⁽²⁾; Nuria Ceular Villamandos⁽³⁾ Antonio Miguel Fuentes Jiménez⁽⁴⁾

- (1) Dpto. Estadística, Econometría, I.O. y Org. Empresas. Universidad de Córdoba. jnt@uco.es
- (2) Dpto. Estadística, Econometría, I.O. y Org. Empresas. Universidad de Córdoba. ccjm@uco.es
- (3) Dpto. Estadística, Econometría, I.O. y Org. Empresas. Universidad de Córdoba. nuria.ceular@uco.es
- (4) Dpto. Estadística e Investigación Operativa. Universidad de Granada. fuentesj@ugr.es

RESUMEN: La obtención de un mecanismo objetivo de valoración inmobiliaria resulta de interés para numerosos colectivos. El propósito último de este trabajo es investigar la posibilidad de aplicar la metodología de Redes Neuronales Artificiales (RNA) a la valoración de inmuebles, para ello en el apartado empírico se estima el precio de la vivienda en la ciudad de Córdoba. Dicha metodología emerge como alternativa a los tradicionales modelos de regresión múltiple, tratando de superar la inflexibilidad característica de estos últimos.

Para la obtención de los datos de trabajo se recurre a Agencias de la Propiedad Inmobiliaria (APIs) situadas en la capital cordobesa. La información suministrada refleja transacciones reales, por tanto el precio suministrado recoge el verdadero precio de mercado. La base de datos contiene un total de 10.124 registros correspondientes a pisos de venta libre vendidos desde el año 2002 hasta el primer semestre de 2006 y en la misma se recogen, además del precio de venta, una serie de características -tanto *internas* del inmueble, como *externas* del edificio donde éste se ubica-. Se obtiene en un primer momento un modelo RNA utilizando los registros del primer semestre de 2006 y posteriormente se proponen los modelos correspondientes a los ejercicios anteriores (2002 a 2005) y se efectúan comparaciones entre los mismos.

Palabras Clave: vivienda, redes neuronales artificiales (RNA), valoración urbana.

ABSTRACT. Several studies oriented to the estimation of prices in real estate markets have been published in the last decade. There are several potential users in need of an objective tool to estimate the price of dwellings: owners of houses and flats, builders, property agents, investment societies, insurance companies, or the public sector department as the fiscal or municipal authorities. In paragraph empirical it comes to determining the price of housing in Cordoba using Artificial Neural Networks (ANN). Neural networks models provide a more flexible tool that can adapt better than hedonic models to nonlinearities of the real estate markets, and with larger forecasting power.

The modelling of property value has been developed with a sample of sales in the city of Cordoba, that is, in an urban zone. From a series of attributes relating to the housing it will try to provide of an objective form which is the price of market. Some exogenous variables are considered, related to the property, the building and its environment. The market price has been recorded for 10124 sales from 2002 until the first semester of 2006; these have been provided by the Real Estate Agencies involved in the study. Categorical data have been summarized through some ad hoc functions that estimate the importance of several related characteristics of each unit.

KEY WORDS: Housing, Artificial Neural Networks (ANN), Real estate appraisal.

Recibido: 25 de noviembre 2008

Revisado: 8 de enero 2009

Aceptado: 17 de febrero 2009

1. Introducción

Desde mediados de los años ochenta y, especialmente, en la última década los precios de la vivienda en propiedad en España han experimentado un crecimiento espectacular. Esto ha originado que las familias españolas destinen una parte muy significativa del presupuesto familiar a la adquisición del bien vivienda. También es preciso tener presente que existe una evidente preferencia sociológica de los españoles hacia la compra de una vivienda frente al alquiler, tendencia alimentada en gran medida por la legislación existente al respecto. Por consiguiente, un mecanismo objetivo de valoración, a partir de los atributos del inmueble, resulta de interés para numerosos colectivos, entre los que cabe citar los propios propietarios, constructores, agentes de la propiedad inmobiliaria, inversores, entidades tasadoras, financieras, aseguradoras o la propia Administración.

Los orígenes del estudio del mercado inmobiliario podrían situarse a nivel internacional en la década de los cincuenta, siendo a partir de los estudios de Rosen, a mediados de los setenta, cuando comienza a analizarse el precio del inmueble en los núcleos urbanos aportando un marco metodológico completo utilizando modelos hedónicos. Fuera de nuestro país ha sido relativamente frecuente encontrar este tipo de análisis, pero comparativamente existen muy pocos estudios al respecto en España, donde en cualquier caso éstos se realizan desde el punto de vista macroeconómico, explicando eminentemente el comportamiento del agregado correspondiente a la inversión en construcción residencial o a las inversiones en vivienda.

Por otra parte, existe un limitado número de trabajos en este ámbito que aplican la técnica de Redes Neuronales Artificiales (RNA). Este trabajo utiliza dicha metodología para determinar empíricamente los precios de los inmuebles en la ciudad de Córdoba en diferentes marcos temporales. En primer lugar, se obtiene un modelo de RNA utilizando los registros del primer semestre de 2006 y posteriormente se propondrán los modelos correspondientes a los ejercicios anteriores (2002 a 2005) y se efectuarán comparaciones.

2. Origen y funcionamiento de una Red Neuronal Artificial (RNA)

La Inteligencia Artificial (IA) se ha desarrollado paralelamente al progreso en el estudio del funcionamiento del cerebro humano, de manera que conocidos los principios básicos de funcionamiento del cerebro se intentaron construir sistemas artificiales que funcionaran de forma análoga y en pequeña escala.

La IA es un campo multidisciplinar del conocimiento, ya que ha sido aplicada a la medicina (para el diagnóstico, detección y predicción de enfermedades), en el procesado de señales, en la economía (para la detección del fraude en tarjetas de crédito, tendencias a corto y medio plazo en bolsa de valores, concesión de créditos, etc.) y el medio ambiente (predicción de irradiación solar o niveles tóxicos de ozono), entre otros.

La Inteligencia Artificial comienza a aplicarse en la valoración de inmuebles a principios de la década de los noventa. Desde entonces hasta la actualidad han surgido numerosas experiencias y la creación de nuevos modelos va en aumento.

Entre los autores que han desarrollado los sistemas más destacables cabe citar los que se recogen en la tabla 1.

En España, dentro del ámbito de valoración de inmuebles destacan las aportaciones de Caridad y Ceular (2001), García Rubio (2004), Gallego (2004) y Lara (2005), aplicaciones relativas al mercado inmobiliario de la ciudad de Córdoba, Albacete, Madrid y Jaén, respectivamente.

Los orígenes de los sistemas neuronales artificiales se remontan a los trabajos de McCulloch y Pitts (1943) con el fin de resolver problemas de carácter cognitivo que no son fáciles de programar de modo algorítmico. Sin embargo, existen trabajos anteriores que abrieron el camino a estos investigadores, entre los que podemos destacar el realizado por Kart Lashley en los años 20.

Tabla 1.

Utilización de sistemas de Inteligencia Artificial (IA) en valoración inmobiliaria

AUTOR	AÑO	ÁREA DE ESTUDIO
Borst	1991	Nueva Inglaterra
Tay y Ho	1992	Singapur
Do y Grudnitski	1992	California (U.S.A.)
Collins y Evans	1994	Reino Unido
Worzala, Lenk y Silva	1995	Colorado (U.S.A.)
Mc Cluskey	1996	
Rossini	1997	Sur de Australia
Haynes y Tan	1998	Gold Coast en Australia
Bonissone	1998	
Cechin	2000	Porto Alegre (Brasil)
Karakozova	2000	Helsinki (Finlandia)
Nguyen y Cripps	2001	Tennessee (U.S.A.)
Kauko	2003	
Limsombunchai y otros	2004	Nueva Zelanda
Liu, Zhang and Wu	2006	
Selim	2009	Turquía

Fuente: Gallego (2004) y elaboración propia

Rumelhart y McClelland (1986) definen la Red Neuronal Artificial (RNA) como una red compuesta de varios operadores simples -también denominados *elementos de proceso, nodos o neuronas*-, dotados de pequeña cantidad de memoria. Dichos nodos están conectados mediante canales de comunicación unidireccionales -*axones*-, que transportan datos numéricos -no simbólicos-. Los nodos únicamente operan sobre sus datos locales y sobre las entradas que reciben a través de los axones. Entre los nodos existe un patrón de conexión representado por una *matriz de pesos o ponderaciones sinápticas (W)*, en la que w_{ij} representa la intensidad y el sentido con el que la salida de la neurona *i-ésima* afecta a la activación de la neurona *j-ésima*. En todo modelo neuronal es preciso disponer de una regla de propagación que combine las salidas de cada neurona con las correspondientes ponderaciones

establecidas por el patrón de conexión para especificar como se valoran las entradas que reciba de cada neurona.

El modelo básico de red está estructurado jerárquicamente en tres partes: la capa de entrada *-input-*, la capa o capas ocultas -comúnmente referenciadas como “caja negra”- y la capa de salida *-output-*. Una cuestión importante es la selección del número de capas ocultas de la red y el número de neuronas por capa. A este respecto no se establecen reglas concretas que permitan la determinación del número de neuronas o el número de capas ocultas de una red para resolver problemas concretos. Hay que tener en cuenta que un número excesivo de capas puede generar ruido, pero se puede conseguir una mejor tolerancia a fallos (Bonilla, 1997). Por otra parte, el número de neuronas ocultas interviene en la eficacia del aprendizaje y generalización de la red¹. Así, en general, pocas neuronas ocultas provocan que la red no distinga bien las características del problema provocando que en la fase de entrenamiento no se alcancen los límites especificados, mientras que la definición de demasiadas neuronas ocultas, sin embargo, aunque no lleve a resultados erróneos, genera procesos tediosos de entrenamiento² además de producir la pérdida de capacidad de generalización de la red. En definitiva, determinar el tamaño de las capas no es una tarea sencilla y suele estar basado en la experiencia, se efectuará en función del problema en estudio y de la precisión esperada de la red.

Otras consideraciones a tener en cuenta serán la determinación de la *regla o función de activación* y del *algoritmo o regla de aprendizaje*. La *función de activación* de las neuronas de las capas ocultas es una función (lineal, sigmoide o logística, hiperbólica...) que, a partir del estado actual de activación de una neurona y de su entrada neta (determinada por el patrón de conexión y la regla de propagación), evalúa el estado de activación siguiente. Por su parte, los *algoritmos de entrenamiento o reglas de aprendizaje* (BP estándar y sus variantes como término de momento, gradientes conjugados, Quickprop, regla Delta-var-Delta...) ³ es el procedimiento por el cual se procede a la modificación del patrón de conexión.

Otro aspecto es la división del conjunto de datos muestrales disponibles en dos subconjuntos: el primero de ellos se tomará como patrón ejemplo para entrenar la red y ajustar los pesos (*fase de aprendizaje*⁴), y el segundo se utilizará para comprobar la eficiencia de la misma (*fase de ejecución*).

¹ Hilera (1995).

² Una capa oculta con muchas neuronas hace que la red tarde menos iteraciones en aprender, pero cada iteración llevará más tiempo, puesto que habrá que calcular más pesos (Pérez y Martín, 2003).

³ Sobre los algoritmos de entrenamiento hay que tener presente que constantemente aparecen nuevas propuestas en la literatura y que, a menudo, un método ideal para acelerar un entrenamiento puede producir un pobre rendimiento en una aplicación diferente (Freeman y Skapura, 1993).

⁴ El tipo de aprendizaje de una red viene determinado por la forma en que se adaptan los pesos o parámetros, distinguiéndose los siguientes tipos de aprendizaje:

Considerando todas estas cuestiones previas se diseña la red comenzando por la elección de un vector de observaciones de las variables exógenas *-inputs* o variables de entrada-, con sus correspondientes observaciones de la variable endógena *-output* o salida deseada- e introducción de dicho vector en la red. Dicha información se propagará hasta la capa superior obteniendo el vector salida. Este proceso se efectúa mediante la función de activación que produce la transformación de las entradas netas de cada neurona de una misma capa en salidas de neuronas de dicha capa a la siguiente, constituyendo así las entradas de la capa superior. Una vez completado este proceso con los n vectores de entrada se procede a calcular la salida global de la red y, por diferencia con la salida deseada, se obtiene el error global por unidad de salida. A continuación, se procede a determinar la contribución relativa de los nodos a dicho error y mediante un algoritmo de entrenamiento los errores serán modificados de forma repetitiva para cada vector del conjunto de patrones de entrenamiento hasta que el error global obtenido se minimice.

Mediante el desarrollo de este comportamiento se consigue la autoadaptación de la red, permitiendo que cuando se le presente una nueva entrada pueda proporcionar una salida adecuada -es lo que se denomina *capacidad de generalización*-.

Dentro de las modalidades de red destaca el Perceptrón Multicapa (*Multilayer Perceptrón* o *MLP*), diseñado para el apartado empírico de este trabajo. Es una

-
- Aprendizaje *Supervisado*: Existe un profesor o supervisor que dispone un conjunto de patrones de entrenamiento, formados por una entrada a la red y la salida esperada para dicha entrada. El profesor presentará a la red de forma iterativa las diferentes entradas. Para cada entrada se producirá una salida que, generalmente, no coincidirá con la salida esperada, calculándose en ese caso el error de salida. En función de este error, se actualizarán los pesos sinápticos, de manera que al volverse a presentar la misma entrada, la respuesta de la red presente un menor error.
 - Aprendizaje *No Supervisado*: En este caso partimos de un conjunto de ejemplos para los que no existen una respuesta deseada de la red. Se determinarán las entradas de forma iterativa a fin de que la red, mediante su regla de aprendizaje, pueda descubrir las regularidades subyacentes en esos datos de entrada, organizándolos en clases no determinadas a priori. Consiste en la estimación de la función densidad de probabilidades, que describe la distribución de patrones pertenecientes al espacio de entrada, por lo que se pueden reconocer regularidades en el conjunto de entradas, extraer rasgos o agrupar patrones según su similitud (clustering).
 - Aprendizaje *Reforzado*: Se dispone de un conjunto de entradas, para cada una de las cuales se obtiene una salida de la red. Se calcula una medida del éxito o fracaso de la red a nivel global, que permitirá determinar los pesos. Como en este caso no se puede determinar la fracción de error correspondiente a cada salida, son necesarios algoritmos de mayor complejidad, así como con conjunto de entrenamiento de un mayor tamaño. Se trata de un mecanismo intermedio entre el supervisado y el no supervisado, ya que al igual que el primero utiliza una medida del error cometido, aunque en este caso sea un valor global indicativo del error de la red. Ahora bien tiene en común con el aprendizaje no supervisado, el no indicar explícitamente la salida deseada.
 - Aprendizaje *Híbrido*: Se trata de utilizar en una misma red el aprendizaje supervisado y el no supervisado, normalmente en distintas capas de la misma.

red alimentada hacia delante -tipo *feedforward*-, que en el supuesto más simple posee una única capa oculta y su funcionamiento se puede dividir en dos fases: propagación de la información desde la capa de entrada a la capa oculta -con el cálculo de los nodos o neuronas ocultas- y propagación de la información desde la capa oculta a la salida -con el cálculo de la salida global de la red-.

3. Finalidad del estudio y material utilizado

La finalidad del presente estudio es la de modelizar el valor de transacción de la vivienda libre y multifamiliar (piso) situada en la ciudad de Córdoba en zonas urbanas no periféricas⁵, es decir, a partir de una serie de atributos referentes al inmueble se tratará de proporcionar de una forma objetiva cuál es su precio de mercado.

Con respecto a la población de vivienda en el municipio cordobés, según datos del último Censo de Población y Vivienda, Córdoba contaba en 2001 con un total de 130.563 viviendas, de las cuales 100.778 eran viviendas familiares principales -viviendas utilizadas durante todo el año o la mayor parte de él como residencia habitual o permanente-, 11.434 viviendas secundarias, 16.429 viviendas vacías y 1.922 de otro tipo. A partir de los datos sobre construcción de nuevas viviendas publicados por el Ministerio de Vivienda se ha llegado a la actualización de la población a 31 de diciembre de 2005 obteniéndose una población total de 135.920 viviendas. No obstante, la población objeto de estudio es algo más reducida que dicha cifra, ya que se limita a vivienda libre multifamiliar (pisos).

Si se desea obtener información relativa a precios de la vivienda se dispone tanto de fuentes de información primaria como secundaria.

Entre las fuentes de información secundaria cabe destacar el Ministerio de Vivienda (antes Ministerio de Fomento) y las principales sociedades de tasación. No obstante, ambos manejan valores de tasación inmobiliaria y no exactamente precios de mercado⁶. Asimismo, el Ministerio de Vivienda suministra información atendiendo únicamente a variables tales como la superficie de la vivienda o la antigüedad de la misma, pero sin tener en cuenta otro tipo de atributos que en principio podrían ser

⁵ *La razón por la que no se incluyen viviendas rurales (o incluso urbanas periféricas) ni unifamiliares es debida a que las características de este tipo de viviendas son muy distintas en cuanto a superficies, localización y entorno frente a las viviendas urbanas tipo pisos, de forma que la relación entre el precio y las características es muy distinta.*

⁶ *Se entiende por “precio de mercado”, la cuantía económica vinculada a una transacción, medida en términos monetarios. Por el contrario, el “valor de mercado” –también conocido como valor venal- se define, en el campo de la valoración y tasación como el estimador del precio de mercado. Se trata de una cifra o un intervalo obtenido a partir del análisis estadístico y económico, totalmente teórico, aunque basado en bases de datos sobre precios y otras variables y que puede aproximarse o diferir, más o menos, del precio real que ha ocurrido o del precio futuro que va a ocurrir si la transacción tiene lugar.*

significativos como la zona concreta en la que se ubica la vivienda⁷, la posesión o no de garaje, trastero, etc.

Es por ello por lo que se recurre en este estudio a fuentes de información primaria. La información a procesar procede de Agencias de la Propiedad Inmobiliaria (APIs) situadas en la capital cordobesa, en concreto destacar por su grado de penetración en el mercado el Grupo Inmobiliario Barin, que cuenta con la red de agencias inmobiliarias más amplia de Córdoba (un total de 18 puntos de venta).

La base de datos contiene un total de 10.124 registros correspondientes a pisos de venta libre vendidos desde el año 2002 hasta el primer semestre de 2006 (año 2002: 772 registros, año 2003: 1685 registros, año 2004: 1399 registros, año 2005: 3380 registros y primer semestre de 2006: 2888 registros) y en la misma se recogen, además del precio de venta, una serie de características -tanto internas del inmueble, como externas del edificio donde éste se ubica-. Los campos, en total 26, se clasifican en la Tabla 2.

La información suministrada por las APIs refleja transacciones reales -dado que ha tenido lugar la sustracción efectiva del inmueble-, por lo tanto el precio suministrado recoge el verdadero precio de mercado. En definitiva, no se están tratando transacciones potenciales, con lo cual evitamos contemplar precios de oferta que quizás en algún momento se convertirían en precios reales de mercado si realmente tuviera lugar la transacción efectiva.

Por otra parte, el hecho de seleccionar viviendas efectivamente vendidas, con precios reales, evita la duplicidad que podría tener lugar si se trabajara con precios de oferta ante la posibilidad de que una vivienda fuera ofertada por varias agencias al mismo tiempo.

En el siguiente epígrafe de este trabajo se obtendrán en un primer momento un modelo de valoración mediante RNA para el primer semestre de 2006 y posteriormente se propondrán los modelos correspondientes a los ejercicios anteriores (2002 a 2005).

Todas las variables marcadas en la tabla 2 con asterisco son dicotómicas (sí/no). Son variables cuantitativas la superficie construida, el número de dormitorios, el número de baños, el número de aseos, los gastos de comunidad y el precio de mercado. Han sido categorizadas para poder ser tratadas estadísticamente las calidades de la solería y la carpintería exterior e interior del inmueble, así como el nivel de climatización existente en el mismo. La zona de ubicación corresponde al barrio/s en el que la vivienda se ubica, de modo que en total la ciudad de Córdoba ha sido dividida en 33 zonas. Como se observa en la tabla 2 dentro de las variables que caracterizan a una vivienda se distinguen dos tipos: las numéricas o cuantitativas y las no numéricas o cualitativas. Estas últimas deben ser tratadas por medio de índi-

⁷ No bastaría con la consideración de las zonas por códigos postales que tiene en cuenta el Ministerio de Vivienda, pues en la mayor parte de los casos esto da lugar a áreas muy heterogéneas.

ces que recojan a la vez varias características con el objetivo de poder atribuirles una interpretación cuantitativa⁸.

Tabla 2.
Clasificación de atributos y variables

INTERNAS DE LA VIVIENDA		EXTERNAS DEL EDIFICIO	
BÁSICAS	Superficie const.	GENERALES	Año edificación Ascensor(*) Tendedero(*)
	Dormitorios Baños Aseos Terraza (*) Teléfono (*) Armarios empotrados(*) Garaje(*) Trastero(*) Climatización		
GENERALES	CALIDAD	EXTRAS	Piscina(*) Tenis(*) Jardines(*)
	Solados Carpintería ext. Carpintería int. Amueblado(*) Cocina amue.(*)		
	REFORMA	Reformado(*)	
ORIENTACIÓN	Exterior(*)		
ECONÓMICAS	Gastos de comunidad Precio de mercado	LOCALIZACIÓN	Zona ubicación

Para la elaboración de dichos índices se han tenido en cuenta dos aspectos: Por un lado, que la interpretación de los mismos sea suficientemente clara, evitando en la medida de lo posible el uso de ponderaciones complicadas que dieran lugar a una interpretación excesivamente subjetiva de los resultados y, por otro lado, cuál es la situación del mercado y las percepciones de los consumidores. La adaptación a la realidad de estos índices se ha contrastado pidiendo la opinión a expertos, tales como los propios agentes de la propiedad inmobiliaria.

El valor que podrán tomar los índices que van a tratarse seguidamente oscilará entre 0 y 1 con la finalidad de homogeneizarlos y de que todos tengan a priori la misma importancia relativa. De tal manera que si el valor de un determinado índice se sitúa próximo a la unidad tendría lugar una situación óptima de las variables que componen dicho índice, por el contrario un valor cercano a cero pondría de manifiesto una situación desfavorable entre las variables que lo componen.

Teniendo presente lo anterior, se recogen en la tabla 3 los índices construidos.

⁸ Richardson (1973); Saura (1995); Jaén y Molina (1995)

Tabla 3.

Indicadores Sintéticos Aritméticos

DENOMINACIÓN ÍNDICE	VARIABLES UTILIZADAS PARA SU CONSTRUCCIÓN
ÍNDICE DE CALIDADES	Solería, carpintería exterior, cocina amueblada y reformas realizadas
ÍNDICE ARQUITECTÓNICO EXTERNO	Disponibilidad de piscina, pista de tenis y jardines
ÍNDICE ARQUITECTÓNICO INTERNO	Antigüedad del edificio, disponibilidad de ascensor y de terraza en la azotea del mismo
ÍNDICE DE APERTURA	Orientación de la vivienda y existencia de terraza en al misma
ÍNDICE DE ANEJOS	Disponibilidad de garaje y trastero
ÍNDICE DE UBICACIÓN	Barrio donde se ubica el edificio ⁹ y nivel de renta ¹⁰ de la zona

Fuente: Elaboración Propia

4. Resultados

4.1. Análisis de Datos

En el proceso de elaboración de una red con óptima capacidad de predicción del precio total del inmueble cordobés se selecciona una red tipo Perceptrón Multicapa -MLP, *Multi Layer Perceptron*-, pues siguiendo investigaciones previas en este campo es la que mejores resultados proporciona con diferencia en este tipo de análisis¹¹.

En lo que respecta a las variables de entrada *-inputs-* a considerar, a priori es deseable incluir un número no demasiado elevado de este tipo de variables fundamentalmente porque la inclusión de un número elevado de variables no origina importantes mejoras en el poder explicativo de la red, de manera que si no se incluyen aquellas variables que poco puedan aportar esto tiene un reducido efecto en la significación global del modelo. De modo que ante varias redes con un poder explicativo similar se elegirá la más simple. Finalmente los *inputs* seleccionados fueron los siguientes:

⁹ Córdoba no es una ciudad monocéntrica, puesto que en ella aparecen zonas distantes cuyas viviendas poseen altos precios, por ello puede dividirse a la ciudad en diferentes centros que podrían ser apreciados por el comprador potencial de este tipo de bien.

¹⁰ Fue posible ampliar la información referente al barrio donde la vivienda está ubicada incorporando datos de los estudios realizados por el Ayuntamiento de la ciudad y la Gerencia de Urbanismo referentes a la renta media de los vecinos de cada uno de los barrios que componen la ciudad. Finalmente se obtuvieron cinco distritos claramente diferenciados que se consideran con niveles de renta y riqueza análogos.

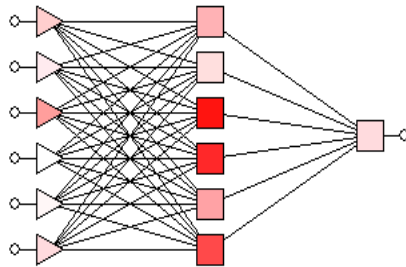
¹¹ Freeman y Skapura (1993), Haykin (1999), García Rubio (2004).

- *Superficie*, que mide las dimensiones de la vivienda y está expresada en metros cuadrados construidos.
- *Antigüedad*, que recoge el número de años que tiene el edificio en el que se ubica la vivienda.
- *IUbicación*, índice de ubicación del inmueble, que pondera la situación geográfica del inmueble junto con el nivel de renta de la zona.
- *IAnejos*, índice de anejos, que recoge la existencia de garaje y trastero.
- *Comun*, gastos de comunidad mensuales expresados en euros.
- *Sol^Car*, que es la interacción entre la calidad de la solería y la calidad de la carpintería exterior de la vivienda.

La salida *-output-* de la red que se pretende construir es el precio final¹² estimado de transacción del inmueble cordobés expresado en euros.

En lo que se refiere a la arquitectura, se diseñan numerosas redes combinando diferente número de capas ocultas, distinto número de nodos en éstas, funciones de activación y algoritmos de aprendizaje variados. Finalmente, se selecciona un Perceptrón Multicapa (MLP) con una capa oculta que presenta la estructura 6:6-6-1:1 (véase figura 1). Esto significa que el número de inputs es de seis¹³, que existen seis neuronas o nodos en la capa oculta y que el número de unidades en la capa de salida se reduce a una, puesto que se tiene una única variable dependiente -el precio total de la vivienda-.

Figura 1.
Estructura de la red perceptrón multicapa 6:6-6-1:1



Para la estimación de la red neuronal se ha utilizado el programa Trajan Neural Networks.

¹² Este precio incluye las comisiones y gastos de gestión de la empresa inmobiliaria que actúa de intermediaria en la transacción. No obstante, no están incluidos impuestos tales como el IVA de los gastos de gestión o el Impuesto sobre Transmisiones Patrimoniales.

¹³ 6:6 significa que el número de neuronas de entrada es de 6 y que tras el preprocesamiento de los datos esta cifra se mantiene ya que no se han utilizado codificaciones especiales –como el esquema 1 de N- vinculadas a las variables cualitativas ante la ausencia de éstas.

Se ha seleccionado una función de activación lineal para la capa de entrada, sin embargo a la capa oculta y la capa de salida se les asigna una función logística o sigmoidea¹⁴. Como función de error se eligió la Suma de Cuadrados de los Errores.

Para garantizar la capacidad generalizadora de la red el conjunto de observaciones de la muestra¹⁵ ha sido dividido de forma aleatoria en dos subconjuntos

- El de *entrenamiento*, con un total de 952 registros (que suponen el 80% de la muestra).
- El de *test*, que contiene 237 registros.

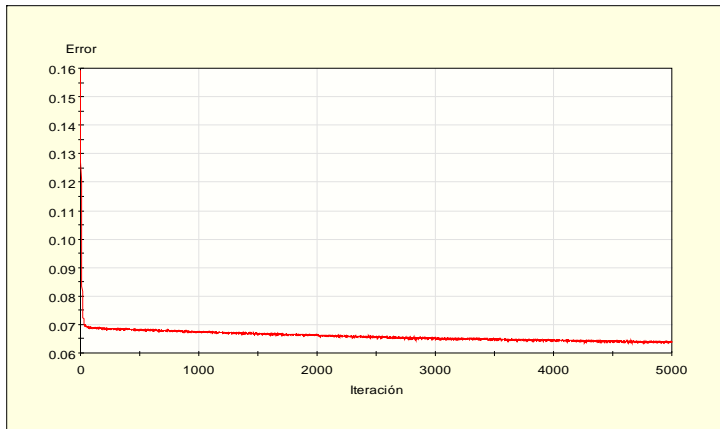
En lo que a la determinación del algoritmo de entrenamiento se refiere se seleccionó el algoritmo de retropropagación de errores -BP o Backpropagation- con las siguientes puntualizaciones:

- Inicialización aleatoria de los pesos y umbrales.
- Máximo número de iteraciones: 5000.
- Ratio de aprendizaje: 0,1.
- Término de momento: 0,1.

El gráfico 1 muestra la evolución del error a lo largo del proceso de entrenamiento. Es importante tener en cuenta al respecto que un período excesivo de entrenamiento puede dar lugar al sobreajuste o sobreaprendizaje de la red que impediría la capacidad de generalización de la misma. En el caso que nos ocupa el mínimo error se alcanza en la iteración 4624, por tanto el proceso de entrenamiento finaliza en ese punto.

Gráfico 1.

Evolución del error durante el proceso de entrenamiento



¹⁴ La función logística o sigmoidea presenta un rango de salida en el intervalo $[0,1]$ y su expresión es: $f(x) = 1/(1 + e^{-x})$, donde x es la entrada neta de cada neurona.

¹⁵ Se utilizan sólo las observaciones completas contando, por tanto, con un total de 1189 registros.

4.2. Modelos Resultantes

Las tablas 4 y 5 muestran las matrices de pesos y umbrales que conectan las neuronas de la capa de entrada con la oculta y las neuronas de ésta con la capa de salida, respectivamente. Las seis neuronas correspondientes a la capa oculta aparecen denotadas con un 2 seguido del número de neurona correspondiente. La neurona de la capa de salida queda reflejada en 3.1.

Tabla 4.

Matriz de pesos de neuronas entre capa de entrada y capa oculta

	2.1	2.2	2.3	2.4	2.5	2.6
Umbral	-0,8637	-5,3193	0,8587	1,3877	-0,8773	-2,9494
SUPERFICIE	1,4581	-2,4698	-3,1902	0,8425	-0,3820	-0,6004
COMUN	3,4714	-1,6389	1,7720	-0,5584	0,6896	-1,0893
ANTIGÜEDAD	-0,0090	-0,1852	0,8528	0,7126	-1,2623	2,4097
IANEJOS	2,1949	-1,4684	-0,0846	-0,2235	-2,8127	1,4203
IUBICACIÓN	-1,1403	-0,7805	-5,0060	-1,5345	0,3433	-2,8736
SOL^CAR	0,3414	-0,7721	0,2906	1,4205	-1,0161	-0,0384

Tabla 5.

Matriz de pesos de neuronas entre capa oculta y capa de salida

	3.1
Umbral	-0,9370
2.1	3,9678
2.2	-3,9225
2.3	-3,7264
2.4	1,8743
2.5	0,9303
2.6	-2,5079

A continuación se procede a observar la influencia de cada uno de los *inputs* sobre el precio total de mercado del inmueble. Esto puede apreciarse a través

del *análisis de sensibilidad*, que nos permite realizar una ordenación de las variables por su poder explicativo tal y como queda reflejada en la tabla 6. El ratio de error de cada *input* representa el cociente entre el error del modelo sin incluir la variable y el error incluyéndola. A la vista de los resultados puede afirmarse que la variable que más contribuye a explicar el precio de venta total de la vivienda es la superficie¹⁶ (1,2953), seguida del índice de ubicación de la misma (1,2009) y de los gastos de comunidad mensuales expresados en euros (1,1577). Para las tres variables explicativas restantes -antigüedad, índice de anejos e interacción entre la solería y la carpintería- se observan ratios muy inferiores a los de las tres variables principales y, además, presentan ratios muy parecidos que oscilan aproximadamente entre el 1,04 y el 1,06.

Tabla 6.
Análisis de sensibilidad de los inputs

INPUT	RATIO	ORDEN
SUPERFICIE	1.2953	1
COMUN	1.1577	3
ANTIGÜEDAD	1.0395	6
IANEJOS	1.0414	5
IUBICACIÓN	1.2009	2
SOL^CAR	1.0644	4

El hecho de obtener ratios tan reducidos para tres de los *inputs* puede hacer pensar en la posibilidad de podar la red para prescindir de la información proporcionada por los mismos. No obstante, se decide mantenerlos básicamente, porque se probaron algunos modelos en los que se prescindía de estas variables y los resultados empeoraron con respecto a los obtenidos para el modelo completo.

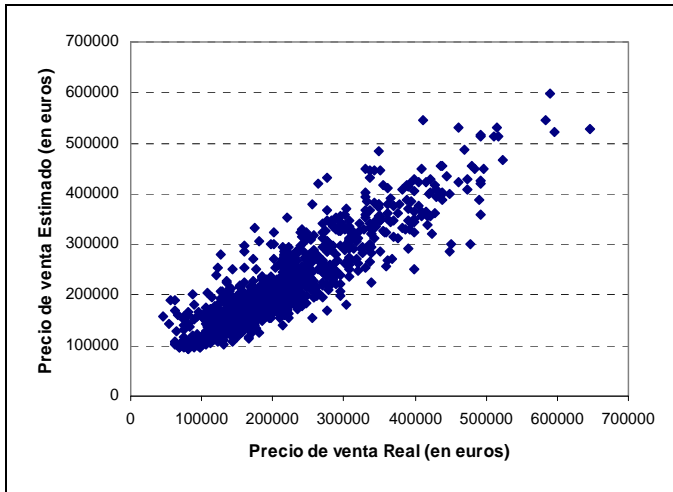
Los resultados obtenidos en el proceso de validación de la red neuronal muestran un alto grado de ajuste en las estimaciones del precio de la vivienda utilizando la red propuesta, puesto que se ha alcanzado un *coeficiente de determinación* (R^2) de 86,05%. La *Raíz del Error Cuadrático Medio* -RECM- (expresado en €) toma un valor de 39.540,36. Por otra parte, la *media del error absoluto* -sin compensaciones entre sobrevaloraciones e infravaloraciones de viviendas- toma un valor de 28.551,34 € y el *error medio relativo* -cociente expresado en porcentaje que recoge el error medio absoluto entre el precio medio real de mercado- asciende a 13.69%.

¹⁶ Con respecto al poder de cada uno de los *inputs* o factores señalar que se probó con otras combinaciones de *inputs* y en todos los casos resultaba la superficie con el ratio más elevado, seguida por el índice de ubicación en la siguiente posición.

Se procede a continuación a la representación gráfica del precio de mercado real y el precio de mercado estimado por la red (véase gráfico 2). Puede observarse que la nube de puntos presenta un ajuste a la bisectriz del primer cuadrante muy elevado, incluso en inmuebles con precios altos, lo que indica que la capacidad predictiva de la red para las observaciones de la muestra de viviendas de Córdoba es muy elevada.

Gráfico 2.

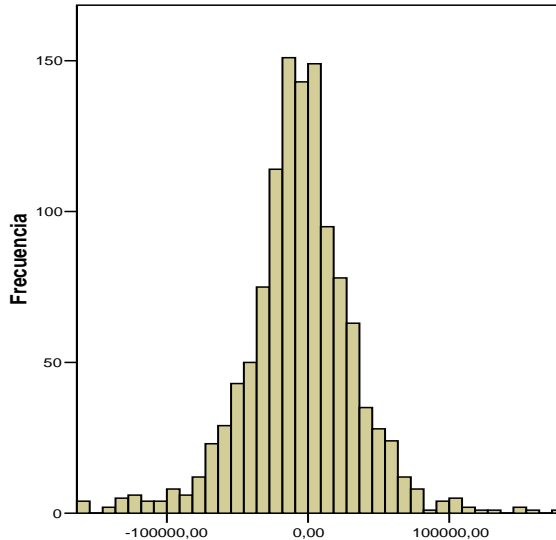
Precio de venta real vs. Precio de venta estimado por la red (en euros)



En el gráfico 3 se muestra un histograma de los errores cometidos por la red neuronal. Se aplica el test de Jarque-Bera para verificar que el término error sigue una distribución normal, obteniéndose un p-valor de 0,57 que permite aceptar la hipótesis nula de normalidad de los residuos. El coeficiente de asimetría arroja un valor de -0,124, por tanto su proximidad a 0 indica que la distribución es prácticamente simétrica. Por otra parte, el coeficiente de curtosis toma un valor de 3,39 lo que permite afirmar que la distribución es leptocúrtica o con alto grado de apuntamiento en el centro.

Con el fin de analizar la capacidad de generalización de la red multicapa obtenida a partir de patrones procedentes de la muestra correspondiente a las transacciones de vivienda efectuadas en la ciudad de Córdoba durante el año 2006, se procede a estimar, manteniendo la misma arquitectura de red, el precio de los inmuebles recogidos en las muestras de las transacciones efectuadas en los años 2002, 2003, 2004 y 2005.

Gráfico 3.

Histograma de errores cometidos por la red

De nuevo la variable dependiente *-output-* es el precio final de transacción del inmueble y como variables independientes *-inputs-* se mantienen las mismas que ya se utilizaron en la red propuesta para 2006 con el fin de poder establecer comparaciones.

Si lo que se pretende es observar el funcionamiento de la red diseñada para otros períodos se mantendrán por completo todas las características descritas para dicha red. Recordemos que estamos ante una red Perceptrón Multicapa -MLP-, con una arquitectura 6:6-6-1:1, con funciones de activación: lineal-logística-logística, algoritmo de entrenamiento BP y que se mantienen los parámetros fijados con anterioridad en cuanto a ratio de aprendizaje, término de momento, iteraciones y demás.

Al igual que se efectuó en el año 2006, el conjunto de observaciones de la muestra correspondiente a cada año ha sido dividido de forma aleatoria en dos subconjuntos: el de *entrenamiento* -que contiene el 80% de la muestra- y el de *test* -que contiene el resto-.

En la siguiente tabla 7 se recogen el número de registros completos -sin campos vacíos para los *inputs* seleccionados-, así como el número de registros que han sido utilizados para la fase de entrenamiento (80%) y la fase de test (20%) en cada uno de los años en que se ha aplicado la red.

Tabla 7.

Número de registros completos, en fase de entrenamiento y test

	REGISTROS COMPLETOS	ENTRENAMIENTO	TEST
2002	470	376	94
2003	914	731	183
2004	791	633	158
2005	1686	1349	337

Seguidamente se realiza el análisis de sensibilidad, es decir, se procede a observar la influencia de cada uno de los *inputs* sobre el precio total de mercado del inmueble. El gráfico 4 refleja que la ordenación de las variables según su poder explicativo no se mantiene constante en los distintos períodos estudiados, aunque sí conviene destacar determinados aspectos:

- La variable *superficie* se consolida como la de mayor influencia sobre el precio final estimado, ocupando en todos los años la primera posición.
- La segunda variable con mayor poder explicativo es siempre el *índice de ubicación*.
- El tercer puesto es ocupado por la *antigüedad* del inmueble en los años 2002 y 2003. Mientras que en los restantes años la variable que ocupa la tercera posición son los gastos de comunidad.
- El cuarto puesto es ostentado por variables con poder explicativo reducido - ratio inferior al 1,10-. En los años 2002 y 2003 se encuentra en este puesto el *índice de anejos*, en el año 2004 la *antigüedad* y en los restantes la *interacción entre la solería y la carpintería*.
- En quinto lugar se observa durante los dos primeros años los *gastos de comunidad*, en 2004 la *interacción entre la solería y la carpintería* y el resto de años aparece en esta posición el *índice de anejos*.
- Por último, la variable con menor poder explicativo es la *interacción de la solería y la carpintería* en 2002 y 2003, el *índice de anejos* en 2004¹⁷ y la *antigüedad* en los dos últimos ejercicios.

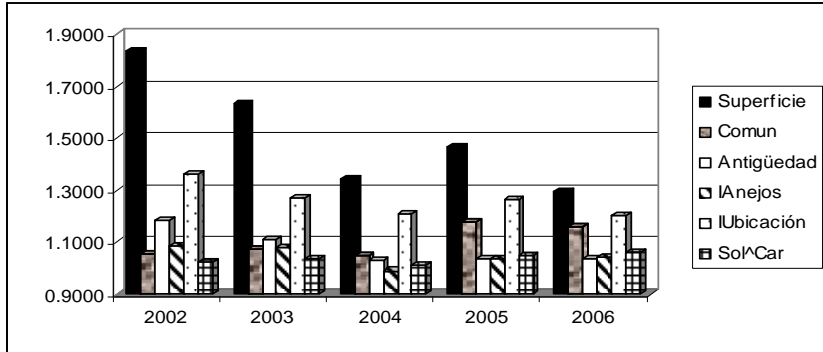
En resumen, la ordenación de las variables se mantiene en los años 2002 y 2003 y, por otro lado, en los años 2005 y 2006. Las variables con mayor poder explicativo son siempre la *superficie* seguida del *índice de ubicación*. Se observa además que las varia-

¹⁷ El *índice de anejos* en el año 2004 posee un ratio de 0,9917 lo que significa que el error de la red prescindiendo de esta variable es menor que si la incluimos. A pesar de ello se ha mantenido dado que su valor está muy próximo a la unidad y, además, porque se pretende efectuar una comparación con el modelo hedónico estimado anteriormente que sí incluía esta variable.

bles *antigüedad* e *índice de anejos* van perdiendo poder a favor de la variable *gastos de comunidad* y, en menor medida, de la *interacción entre la solería y la carpintería*.

Gráfico 4.

Comparativa de análisis de sensibilidad de los inputs



En lo referente a la validación, los estadísticos obtenidos para la red descrita aplicada a los años 2002 a 2005 son bastante satisfactorios y quedan reflejados en la tabla 8. Para el *coeficiente de determinación* (R^2) se obtienen los mayores valores en los años 2002 y 2004 -ambos superan un grado de ajuste del 90%-, se observa una disminución de la *raíz del error cuadrático medio* (RECM) con respecto a los resultados obtenidos para 2006, así como de la *desviación típica residual* y de los *errores medios absolutos*. No obstante, el *error medio relativo* más bajo ha sido el obtenido para el primer semestre de 2006.

Tabla 8.

Comparación de los resultados obtenidos 2002-2006

	R^2	RECM	DESVIACIÓN TÍPICA RESI- DUAL	ERROR MEDIO ABSOLUTO	ERROR MEDIO RELATIVO
2002	90,21%	21.652,25	21.634,04	14.900,48	13,78%
2003	84,75%	25.884,77	25.738,99	18.999,93	14,75%
2004	90,23%	30.983,78	28.908,32	24.157,36	16,26%
2005	81,12%	32.825,44	32.755,04	23.562,59	13,82%
2006	86,05%	39.540,36	39.102,13	28.551,34	13,69%

5. Conclusiones

El desarrollo de la Inteligencia Artificial permite la utilización de sistemas de Redes Neuronales Artificiales (RNA) como alternativa a los métodos económicos de modelización tradicional.

La utilización de redes neuronales se muestra apropiada para evitar la rigidez propia de los planteamientos de los modelos tradicionales de regresión, en los que, entre otras limitaciones, destaca el elevado grado de multicolinealidad que suele estar presente entre las variables explicativas del precio del bien. Otra ventaja de una RNA es su capacidad para estimar el valor de las propiedades que presentan características significativamente diferentes de las de los bienes de su entorno (*outliers*), dado que este tipo de sistemas somete a las muestras a procesos matemáticos mucho más complejos que los tradicionales modelos. No obstante, aunque la mayoría de los trabajos señalan la superioridad de las redes es preciso señalar que algunos estudios llegaron a conclusiones bien distintas rechazando la red por no superar a los modelos tradicionales de regresión y por su naturaleza de “caja negra”¹⁸.

Formulada una red MLP (Perceptrón Multicapa) para tratar de determinar el precio de venta total de un inmueble, teniendo en cuenta las transacciones efectuadas en Córdoba durante el primer semestre de 2006, se aprecia que los resultados obtenidos en el proceso de validación de la red neuronal muestran un alto grado de ajuste en las estimaciones del precio de la vivienda.

En cuanto a los *inputs* utilizados indicar que la superficie, la ubicación, la antigüedad y la presencia de garaje suelen estar presentes en la mayor parte de los trabajos que han aplicado redes neuronales a la valoración de bienes inmuebles.

La *superficie* -ya sea útil o construida- suele ocupar la primera posición -o segunda en algunos casos¹⁹- al efectuar el análisis de sensibilidad, seguida por la *ubicación del inmueble*. Esta última variable viene recogida por algunos autores como “distancia al centro de la ciudad” en el caso de ciudades monocéntricas. También hubiera sido interesante precisar variaciones en precio ante variaciones concretas en la ubicación -por números en una calle o coordenadas geográficas-, pero ha de tenerse presente que los resultados obtenidos han sido tomados en función de la información disponible y las APIs por motivos de confidencialidad se negaron rotundamente a proporcionar la localización exacta de los inmuebles.

Desde un primer momento también se consideró relevante la presencia en el modelo de la variable *antigüedad* del inmueble. En la red propuesta dicho atributo ha sido expresado en número de años, no obstante en otros estudios la variable antigüedad ha sido sustituida por la variable “estado de conservación del inmueble” considerada más representativa. De hecho esta última puede estar relacionada con el input incluido en la red propuesta que considera la interacción entre la calidad de la

¹⁸Entre los trabajos que han puesto de manifiesto que las redes neuronales no son necesariamente superiores a los modelos de regresión se encuentran los de Allen and Zumwalt (1994) o Worzala, Lenk y Silva (1995).

¹⁹ García Rubio (2004)

solería y la calidad de la carpintería exterior -que ocupa un cuarto lugar en el análisis de sensibilidad-.

También la variable *garaje* suele ser un atributo común en este tipo de modelos²⁰, pues a priori el precio del inmueble será tanto mayor si lleva el garaje incluido, así como si se incluyen otros elementos tales como el típico cuarto trastero.

Destacar, por último, la inclusión en la red propuesta de los *gastos de comunidad*. Desde el principio la presencia de los gastos de comunidad está totalmente justificada si tenemos en cuenta que existe una elevada correlación entre esta variables y otros elementos arquitectónicos externos (como piscina, pista de tenis o jardines) que consiguen elevar significativamente la cuantía de los gastos de comunidad. Por otro lado, es lógico que el comprador de un inmueble esté dispuesto a pagar más por el mismo si se da una mayor calidad de la solería y de la carpintería exterior.

Finalmente, señalar que las redes en general, y la obtenida en particular, presentan un importante inconveniente: es difícil validar los resultados en las capas intermedias, además se carece de modelo para poder interpretar económicamente los resultados de los coeficientes obtenidos. Sin embargo, se ha de indicar que si bien dichos coeficientes son claramente obtenidos en el modelo de regresión, tampoco su interpretación resulta del todo correcta, dado que los componentes de la vivienda no aparecen de forma individual sino formando parte de un todo. Por ello, no parece apropiado en este sentido hablar de precios implícitos.

Para comprobar la capacidad generalizadora de la red diseñada, se aplica la misma red a las transacciones efectuadas en Córdoba durante los años 2002 a 2005 corroborando el buen funcionamiento de la red. Esta afirmación se apoya en que los coeficientes de determinación (R^2) calculados para las distintas redes toman valores en un rango que oscila entre el 81 al 90%. En concreto, destaca la red aplicada a las transacciones efectuadas en el año 2002 como la que mejores resultados ha arrojado.

Entre los usos posibles de la metodología analizada para valoración inmobiliaria señalar los siguientes:

- Utilización por los propios Agentes de la Propiedad Inmobiliaria como apoyo para ajustar de una forma relativamente fácil, rápida y barata el precio de la vivienda que desea ser vendida por un cliente.
- Los resultados también podrían ser tenidos en cuenta por los constructores de vivienda del ámbito en que se realiza el estudio, que habrán de prestar una mayor atención a aquellos atributos de mayor relevancia para la determinación del precio del inmueble.
- Sería interesante efectuar comparaciones, tanto a nivel espacial como temporal, del precio de las transacciones inmobiliarias y de los principales determinantes del mismo si se propugnaran desde la Administración este tipo de investigaciones.

²⁰ Véase Lara Cabeza (2005); Quang Do y Grudnitski (1992); García Rubio (2004); Lisombunchai (2004); Worzala et. al. (1995).

Referencias bibliográficas

- Allen, W.C. And Zumwalt, J.K. (1994). *Neural Networks: a word of caution. Working Paper. Colorado State University.*
- Bonilla, M^a y Puertas, R. (1997). Análisis de las redes neuronales: aplicación a problemas de predicción y clasificación financiera. *Quaderns de Treball 43. Dpto. de Economía Financiera y Matemática. Universidad de Valencia.*
- Bonissone et. Al. (1998). "Automated residential property valuation: An accurate and reliable approach based on soft computing", en R. Ribeiro; H. Zimmermann; R. Yager y J. Kacprzyk, J. (eds.). *Soft Computing in Financial Engineering. Heidelberg: Physica-Verlag.*
- Borst, R. (1991). "Artificial Neural Networks: The Next Modelling/ Calibration Technology for the Assessment Community?", *Property Tax Journal, IAAO, vol. 10 (1), pp. 69-94.*
- Caridad, J. M. y Ceular, N. (2001). "Un análisis del mercado de la vivienda a través de Sistemas de Redes Neuronales", *Revista de Estudios de Economía Aplicada, Vol. 18, pp. 67-81.*
- Cechin, A.L.; Souto, A. y González, M.A. (2000). Real estate value of Porto Alegre city usin artificial neural networks. *Paper presented at the Brazilian Symposium of Neural Networks (SBRN'2000). Rio de Janeiro, Brasil.*
- Collins, A. y Evans, A. (1994). "Neural Networks: an application to residential valuation in the U.K", *Journal of Property Valuation and Investment, vol. 11, n^o2, pp. 195-204.*
- Do, A. y Grudnitski, G. (1992). "A Neural Network Approach to Residential Property Appraisal", *The Real Estate Appraiser, Vol.58 (3), pp. 38-45.*
- Fan, G.; Ong, Z.S.E. and Koh, H.C. (2006). "Determinants of house price: A decision tree approach", *Urban Studies, Vol.43 (12), pp. 2301-2315.*
- Freeman, J. and Skapura, D.M. (1993). *Redes neuronales, algoritmos, aplicaciones y técnicas de programación. Wilmington Addison-Wesley Diaz de Santos, Barcelona, pp. 26-241.*
- Gallego Mora-Esperanza, J. (2004). "La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid", *Revista CT/Catastro, n^o 50, pp.51-67*
- García Rubio, N. (2004). Desarrollo y aplicación de redes neuronales artificiales al mercado inmobiliario: aplicación a la ciudad de Albacete. *Tesis Doctoral. Universidad de Castilla – La Mancha. España.*
- Haykin, S. (1999). *Neural networks: A comprehensive foundation. Prentice –Hall, Nueva York.*
- Haynes, J. y Tan, C.N. (1998). An artificial neural network real estate price simulator. *School of Information Tecnology. Bond University. Australia*
- Hilera, J. R. y Martínez, V. J. (1995). *Redes neuronales artificiales: Fundamentos, modelos y aplicaciones. Ra-Ma, Madrid.*
- Jaén, M. y Molina, A. (1995). Modelos econométricos de tenencia y demanda de vivienda. *Servicio de publicaciones de la Universidad de Almería.*

- Karakozova, O. (2000).* A comparison between neural network and multiple regression approaches: An application to residential valuation in Finland. *Tesis de Swedish School of Economics and Business Administration. Suecia.*
- Kauko, T. (2003).* "On current neural network applications involving spatial modelling of property prices", *Journal of Housing and the Built Environment, Vol.18 (2), pp.159-181.*
- Lara Cabeza, J. (2005).* "Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria", *Revista Mapping, n° 104, pp. 64-71.*
- Lashley, K.S. (1930).* "Basic Neural Mechanism in Behavior", *Psychological Review, n° 37, pp.1-24.*
- Limsombunchai, V., Gan, C. and Lee, M. (2004).* "House price prediction: Hedonic Price Model vs. Artificial Neural Network", *American Journal of Applied Sciences, vol. 1(3), pp. 193-201.*
- Liu, J.; Zhang, G.X. and Wu, W.P. (2006).* "Applications of fuzzy neural network for real estate prediction", *LNCS, n° 3973, pp. 1187-1191.*
- McCluskey, W. (1996).* "Predictive accuracy of machine learning models for mass appraisal of residential property", *Residential Valuer's Journal, July, pp. 41-47.*
- McCulloch, W. y Pitts, W (1943).* "A logical calculus of the idea immanent in nervous activity", *Bulletin of Mathematical Biophysics, n° 5, pp. 115-133.*
- Nguyen, N. y Cripps, A. (2001).* "Predicting housing value: a comparison of multiple regression analysis and artificial neural networks", *Journal of Real Estate Research, vol. 22 (3), pp. 314-336.*
- Peidró, J. M. (2004).* "La demanda que no cesa", *Boletín de Mercado 03, CB Richard Ellis, pp. 77-85.*
- Pérez Delgado, M. L. y Martín Martín, Q. (2003).* Aplicación de las redes neuronales artificiales a la estadística. *Ed. La Muralla. Madrid.*
- Quang Do, A. and Grudnitski, G. (1992).* "A neural network approach to residential property appraisal", *Real Estate Appraiser, Vol.58 (3), pp. 38-45.*
- Richardson, H. W. (1973).* Economía Regional. Teoría de la localización, estructuras urbanas y crecimiento regional. *Ed. Vicens Vives, Barcelona.*
- Rosen, S. (1974).* "Hedonic Prices and Implicit Markets: Product Differentiation in Pure competition", *Journal of Political Economy, n° 82, pp. 34-55.*
- Rossini, P. (1997).* Application of Artificial Neural Networks to the Valuation of Residential Property. *3rd Pacific Rim Real Estate Society Conference, New Zealand.*
- Rumelhart, D. and McClelland, J. (1986).* Parallel distributed processing: Explorations in the microstructure of cognition. *MIT, Cambridge.*
- Saura, P. (1995).* Demanda de características de la vivienda en Murcia. *Secretariado de Publicaciones de la Universidad de Murcia.*
- Selim, H. (2009).* "Determinants of house prices in Turkey: Hedonic regresión versus artificial neural network", *Expert Systems with Applications, n°36, pp.2843-2852*

Tay, D. and Ho, D. (1992). "Intelligence and the Mass Appraisal of Residential Apartments", Journal of Property Valuation and Investment, vol. 10, pp. 525-540.

Worzala, E.; Lenk, M. and Silva, A. (1995). "exploration of neural networks and its application to real estate valuation", Journal of Real Estate Research, vol. 10 (2), pp. 185-202.