

Metodología de precios hedónicos vs. Redes Neuronales Artificiales como alternativas a la valoración de inmuebles. Un caso real

José M^a Caridad y Ocerín

Catedrático de la Universidad de Córdoba

Julia M. Núñez Tabales

Profesora de la Universidad de Córdoba

Nuria Ceular Villamandos

Profesora de la Universidad de Córdoba

En este trabajo se examinan dos métodos de valoración aplicables al mercado inmobiliario: la Metodología de Precios Hedónicos (MPH) –que analiza el precio del bien vivienda en función de sus principales características– frente a las Redes Neuronales Artificiales (RNA) –que tratan de superar la inflexibilidad y linealidad de los modelos hedónicos tradicionales–.

En el apartado empírico se trata de determinar el precio de la vivienda en Córdoba comparando ambas metodologías, concluyendo que la red presenta un mayor poder de predicción que las estimaciones obtenidas a través del modelo hedónico.

El presente trabajo aborda el tema de la determinación de los precios dentro del mercado inmobiliario.

Los precios de la vivienda en propiedad en España han experimentado desde mediados de los años ochenta y, especialmente, en la última década un crecimiento espectacular. Esto ha provocado que los consumidores 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 viviendas frente al alquiler. Por consiguiente, un mecanismo

objetivo de valoración 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.

A nivel internacional podrían situarse los orígenes del estudio del mercado inmobiliario 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. Este tipo de análisis ha sido relativamente frecuente fuera de nuestro país, 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 la metodología hedónica y las redes para determinar empíricamente los precios de los inmuebles en la ciudad de Córdoba y efectuar una comparación entre el poder de predicción de ambas técnicas para terminar sugiriendo cuál es más apropiada para la valoración de una vivienda.

Metodología

Metodología de Precios Hedónicos (MPH)

La idea subyacente a la técnica de los precios hedónicos es que, si un bien está en realidad constituido por un conjunto de atributos, entonces su precio de mercado deberá ser un agregado de los precios individuales

de todos ellos. Mediante técnicas estadísticas de regresión se puede llegar a una cuantificación monetaria de aquellos aspectos que conforman el activo analizado y estimar su contribución al valor global de mercado.

Con el objeto de explicar la heterogeneidad inherente al bien vivienda –diferente tamaño, calidad, características ambientales, localización...– numerosos análisis referentes a este mercado han considerado a la vivienda en términos hedónicos. De tal manera, que la unidad de vivienda es conceptualizada no como bien homogéneo e indivisible, sino como una cesta de atributos individuales cada uno de los cuales contribuye a la provisión de uno o más servicios de vivienda.

La mayoría de los autores sitúan el origen de la metodología de precios hedónicos (MPH) en los trabajos realizados por Court (1939) para la determinación de precios en el mercado automovilístico (1). No obstante, otros autores –como Colwell y Dillmore– señalan que el verdadero origen de los modelos hedónicos es preciso situarlo diecisiete años antes, en 1922, cuando Haas aplica esta metodología al cálculo de precios de la tierra de cultivo. Wallace (1926) continúa esta misma línea de investigación en Iowa. También encontramos en 1929 una aplicación de la MPH en el estudio de la calidad de las legumbres realizada por Waught.

Los estudios posteriores sobre MPH son atribuibles a Lancaster, que a mediados de los sesenta desarrolla la denominada Nueva Teoría del Consumidor, según la cual la utilidad se deriva de las características de los bienes y no de los bienes en sí mismos.

La primera aplicación de esta metodología al mercado de la vivienda la hallamos en los trabajos de Ridker y Henning (1967) que aportaron evidencia empírica de que la polución afectaba al precio de la vivienda (2). Por tanto, las características de una

(1) SIRMANS (2005)

(2) AGUILÓ SEGURA (2002)

vivienda pueden agruparse al menos en dos categorías bien diferenciadas: por un lado, las características estructurales de la vivienda y, por otro, las características relacionadas con la localización y el entorno físico.

En 1971 destaca la aportación a esta teoría realizada por Griliches. Pero es Shervin Rosen en 1974 el primero en proporcionar un tratamiento unificado del modelo teórico de los mercados implícitos subyacentes en la MPH. A partir de este momento, el modelo desarrollado por Rosen ha llegado a ser generalmente acep-

tado como el paradigma del enfoque hedónico.

Asimismo, Freeman (1979) facilitó la primera justificación teórica para la aplicación de esta metodología al mercado de la vivienda.

A partir de Rosen las aplicaciones de la MPH se han sucedido, especialmente en los países anglosajones. Siguiendo un orden cronológico en el cuadro 1 se relacionan los autores más destacables que han aplicado esta metodología, ya sea con el objetivo de obtener el precio de la vivienda o con el de obtener índices de precios.

Cuadro 1
Utilización de la MPH en valoración inmobiliaria

A. Objetivo: Obtener el precio de la vivienda
Ridker y Henning (1967); Wilkinson (1973); Kain y Quiley (1975, 1979); Tang (1975); Freeman (1979); Wheaton (1979); Witte y otros (1979); Li y Brown (1980, 1985); Palmquist (1992); Caridad y Brañas (1996); So y otros (1996); Clapp y Giacotto (1998); Chesire y Sheppard (1998); Adair y otros (1996,2000); Chattopadhyay (1999); Bilbao Terol (2000); Caridad y Ceular (2000); Bateman y otros (2001); Boyle y Kiel (2001); Hidano (2002); Aguiló (2002); Bengochea (2003); Malpezzi (2003); Goodman y Theriault (2003); Thibodeau (2003); Harding y otros (2003); Cervero y Duncan (2004); Fletcher y otros (2004); McMillen (2004); Bond y otros (2005).
B. Objetivo: Obtener índices de precios
Straszheim (1975); Goodman (1978); Palmquist (1980); Butler (1982); Case y Shiller (1987); Englund y otros (1988); Case y otros (1991); Haurin y otros (1991); Mills y Simenauer (1996); Wallace (1996); Meese y Wallace (1997,2003); Wolverton y Senteza (2000); Bover y Velilla (2001)

Fuente: elaboración propia.

En concreto, para el mercado inmobiliario español destacan las aportaciones de: Caridad y Brañas (1996), Caridad y Ceular (2000) –ambas para el mercado inmobiliario cordobés–, Bilbao Terol (2000) en Asturias, Bover y Velilla (2001) en Madrid, Aguiló Segura (2002) en las Islas Baleares y Bengochea Morancho (2003) en Castellón.

Estas aportaciones ponen en evidencia, a pesar de que el modelo no está exento de

limitaciones (3), la utilidad de la metodología hedónica para identificar los factores

(3) Entre los problemas que puede presentar un modelo hedónico pueden destacarse la multicolinealidad entre las variables predeterminadas del modelo –debido a las características intrínsecas de dichas variables–. Por otro lado, en lo que al análisis de las perturbaciones aleatorias se refiere, el estimar con datos de corte transversal, conlleva la posible presencia de heterocedasticidad.

determinantes del precio de un inmueble y para la cuantificación de los mismos.

Redes neuronales artificiales (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 procesamiento 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.

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.

La Inteligencia Artificial comienza a aplicarse en la valoración de inmuebles hace aproximadamente quince años. 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 fuera de nuestras fronteras cabe citar los que se recogen en el cuadro 2.

Cuadro 2
Utilización de sistemas de Inteligencia Artificial (IA) en valoración inmobiliaria

Autor	Año	Área de Estudio
Borst	1991	Nueva Inglaterra
Tay y Ho	1992	Singapur
Quang Do y Grudnitski	1992	California (EE.UU.)
Collins y Evans	1994	Reino Unido
Worzala, Lenk y Silva	1995	Colorado (EE.UU.)
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 (EE.UU.)
Kauko	2003	
Limsombunchai y otros	2004	Nueva Zelanda

Fuente: Gallego (2004) y elaboración propia.

A nivel nacional, dentro del ámbito de valoración de inmuebles destacan las aportaciones de Ceular y Caridad (2000), Mohamed (2002), Fuentes Jiménez (2003), García Rubio (2004), Gallego (2004) y Lara (2005), aplicaciones relativas al mercado inmobiliario de la ciudad de Córdoba, Cádiz, Melilla, Albacete, Madrid y Jaén, respectivamente.

McClelland y Rumelhart (1985) 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 toleran-

cia a fallos. Por otra parte, el número de neuronas ocultas interviene en la eficacia del aprendizaje y generalización de la red (Hilera, 1995). 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 entrenamiento, 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, el algoritmo de entrenamiento o regla de aprendizaje (BP estándar y sus variantes como término de momento, gradientes conjugados, Quickprop, regla Delta-var-Delta...) (4) 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, y el segundo se utilizará para comprobar la eficiencia de la misma.

(4) 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).

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 (Multi Layer Perceptron o MLP), diseñado para la aplicación empírica de este estudio. Es una 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–.

En la primera fase se obtienen las salidas de las unidades de la capa oculta como respuesta a un vector de entradas. La fórmula que calcula la entrada neta que recibe el elemento j de la capa oculta es (5):

$$a_j = \sum_{i=0}^n \omega_{ij} x_i$$

donde se ha considerado un peso adicional (ω_{0j}) que actúa como umbral y pondera a una entrada fija, x_0 , de valor unitario.

La respuesta de la unidad se obtiene introduciendo la entrada neta calculada dentro de la función de activación:

$$z_j = f(a_j)$$

En una segunda fase se obtiene la salida global de la red. Por ejemplo, el elemento k de la capa de salida, que recibe como entradas las salidas proporcionadas por m nodos ocultos tendría una entrada neta equivalente a:

$$a'_k = \sum_{j=0}^m \omega'_{jk} z_j; z_0 = 0$$

Por tanto, la salida global de la red sería:

$$y_k = g(a'_k) = g\left(\sum_{j=0}^m \omega'_{jk} z_j\right) = g\left[\sum_{j=0}^m \omega'_{jk} f\left(\sum_{i=0}^n \omega_{ij} x_i\right)\right]$$

Objeto del estudio y materiales utilizados

La finalidad del presente estudio es la de modelizar el valor de transacción de la vivienda libre y multifamiliar (tipo piso) situada en la ciudad de Córdoba en zonas

(5) X_i , $i=0, \dots, n$; z_j , $j=0, \dots, m$ e y_k , $k=1, \dots, l$, constituyen los nodos de la capa de entrada, oculta y de salida, respectivamente.

ω_{ij} representa la ponderación del nodo i de entrada al nodo j de la capa oculta y ω'_{jk} la ponderación del elemento oculto j en el nodo de salida k.

urbanas no periféricas (6), 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. A partir de los datos sobre construcción de nuevas viviendas publicados por el Ministerio de Vivienda se llegó 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 y las principales sociedades de tasación. No obstante, ambos manejan valores de tasación inmobiliaria y no exactamente precios de mercado (7). Asimismo, el Ministerio de Vivienda suministra información

(6) 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.

(7) 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.

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 significativos como la zona concreta en la que se ubica la vivienda (8), 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 –18 puntos de venta en total–.

La información suministrada por las APIS refleja transacciones reales correspondientes al primer semestre de 2006. Por haber tenido lugar en ese período la transacción real del inmueble, 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.

La base de datos contiene un total de 2.888 registros correspondientes a pisos de venta libre y en la misma se recogen una serie de características, tanto internas del inmueble, como externas del edificio donde se ubica, para las viviendas objeto de compra-venta en el período descrito. Los

(8) 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.

Cuadro 3. Clasificación de variables

Internas de la Vivienda			Externas del Edificio	
Básicas	Superficie const. Dormitorios Baños Aseos Terraza (*) Teléfono (*) Armarios empotrados(*) Garaje(*) Trastero(*) Climatización		Generales	Año edificación Ascensor(*) Tendedero(*)
Generales	Calidad	Solados Carpintería ext. Carpintería int. Amueblado(*) Cocina amueblada.(*)	Extras	Piscina(*) Tenis(*) Jardines(*)
	Reforma	Reformado(*)		
Orientación	Exterior(*)			
Económicas	Gastos de comunidad Precio de mercado		Localización	Zona ubicación

campos suman en total 26 y se clasifican en el Cuadro 3.

Todas las variables marcadas en el Cuadro 3 con asterisco son dicotómicas (presencia/ausencia de una determinada variable). 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 solearía 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 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 índices que recojan a la vez varias caracte-

rísticas con el objetivo de poder atribuirles una interpretación cuantitativa (Saura, 1995; Jaén y Molina, 1995).

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 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

Cuadro 4. 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.

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 el Cuadro 4 los índices construidos.

Resultados

Metodología de precios hedónicos (MPH)

La variable dependiente del modelo que se pretende construir es el precio final (11) estimado de transacción del inmueble corrobó expresado en euros.

(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) 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.

Una primera decisión a tomar es la elección de la forma funcional adecuada para estimar la ecuación hedónica, que puede ser lineal o no. Dicha elección se reduce a una cuestión empírica, ya que las distintas aportaciones existentes sobre este tema no han establecido un criterio para seleccionar aquella forma funcional que ofrezca mejores resultados. Por ello se ensaya con diferentes formas funcionales y se elige aquella que proporciona un mejor ajuste. Pueden destacarse entre las formas funcionales más utilizadas tradicionalmente la forma lineal, la semilogarítmica y la doblemente logarítmica. En este caso se ha seleccionado la forma lineal.

En lo que respecta a las variables explicativas a incluir en la ecuación hedónica, a priori es deseable incluir un número no demasiado elevado de este tipo de variables fundamentalmente por dos razones. En primer lugar, porque las variables independientes suelen estar relacionadas entre sí apareciendo en ese caso problemas de multicolinealidad. En segundo lugar, porque la inclusión de un número elevado de variables no origina importantes mejoras en el poder explicativo del modelo, de manera que si no se incluyen aquellas variables que poco puedan aportar esto tiene un reducido efecto en los coeficientes de las variables clave y en la significación global del modelo. De modo que ante varios modelos con un poder explicativo similar se elegirá el más simple en virtud del principio de parsimonia.

Dado que la oferta no es aditiva, es posible añadir en el modelo términos que recogen interacciones entre las variables, así como formas cuadráticas.

Así pues, una especificación que incluiría todas las posibles interacciones es:

$$Pr\acute{e}cio_i = b_0 + \sum_{j=1}^k b_j x_j + \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k b_{ij} x_i x_j$$

$$Pr\acute{e}cio = \beta_0 + \beta_1 Superficie + \beta_2 Antigüedad + \beta_3 Ubicación + \beta_4 IAnejos + \beta_5 Comun + \beta_6 Sol^{\wedge}Car$$

Las variables explicativas (12) de la ecuación hedónica son las siguientes:

- *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.

aunque evidentemente los modelos resultantes gozarán de mayor simplicidad.

Para la estimación del modelo hedónico se han utilizado los paquetes econométricos EViews 4.0 y SPSS 12.0.

A continuación se procede a la estimación de varios modelos hedónicos alternativos por el método de Mínimos Cuadrados Ordinarios, de los cuales se selecciona tras el proceso de validación el siguiente:

- *Sol^{\wedge}Car*, que es la interacción o producto entre el nivel de calidad de la solería y el nivel de calidad de la carpintería exterior de la vivienda.

Inicialmente (Cuadro 5) se detectó la presencia de heterocedasticidad tras aplicar el test de White. Esta característica fue corregida mediante el método propuesto por ese mismo autor, de modo que la ecuación hedónica se concreta finalmente como sigue:

$$Pr\acute{e}cio = 193.679 + 1109.951Superficie - 1067.449Antigüedad + 64297.29IUbicación + 18458.66IAnejos + 1296.708Comun + 5117.270Sol^{\wedge}Car$$

Cuadro 5. Estimación de la ecuación hedónica

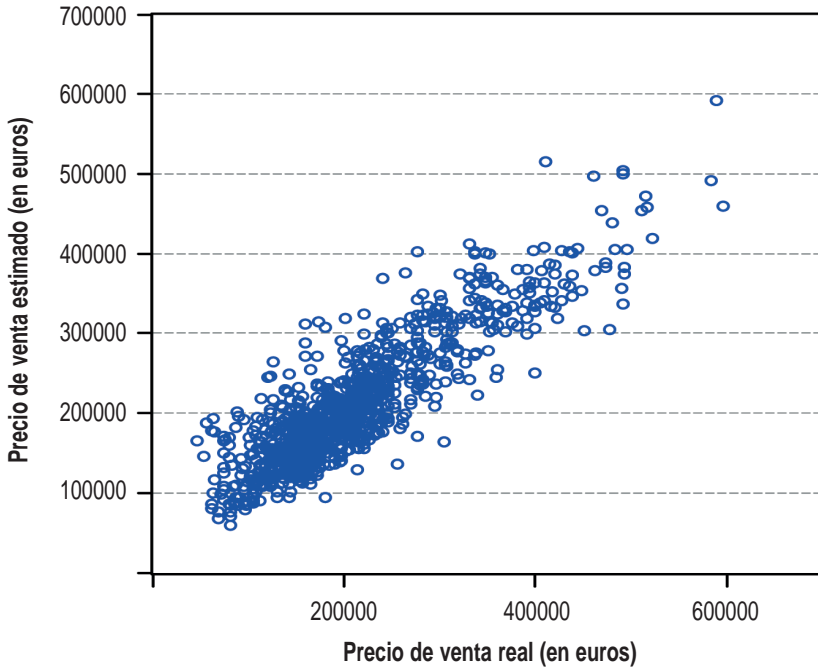
Variable Dependiente: Precio de venta				
Variable	Coficiente	Error estándar	Estadístico t-Student	Prob.
C	193.6790	10237.53	0.018919	0.9849
SUPERFICIE	1109.951	83.62050	13.27368	0.0000
ANTIG	-1067.449	173.5725	-6.149872	0.0000
IUBICACION	64297.29	5494.403	11.70232	0.0000
IANEJOS	18458.66	4572.246	4.037111	0.0001
COMUN	1296.708	105.9213	12.24218	0.0000
SOL^{\wedge}CAR	5117.270	504.0529	10.15225	0.0000

Fuente: Elaboración propia.

(12) En relación a las variables independientes o explicativas indicar que se estimaron otros modelos alternativos que incluían otras variables que en principio podrían resultar relevantes a la hora determinar el

precio de la vivienda –presencia de terraza en la vivienda, climatización, armarios empotrados, azotea transitable, ascensor...–, pero finalmente fueron rechazadas debido a que se conseguía un menor grado de ajuste.

Gráfico 1
Precio de venta real (en euros) vs. Precio de venta estimado por MPH



Fuente: Elaboración propia.

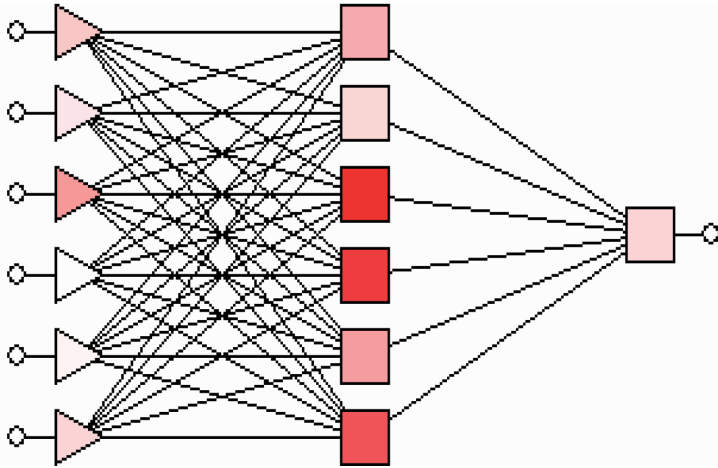
El contraste de los test usuales, incluyendo el test de estabilidad de Chow para los coeficientes es correcta. Con respecto a la multicolinealidad se calcula el valor del índice de condición k , que toma un valor de 16,6, lo cual denota la inexistencia de multicolinealidad si tenemos en cuenta que ésta se presenta para un valor de k superior a 20 (13). El coeficiente de determinación R^2 toma un valor de 77,38%. La raíz del error cuadrático medio (RECM) es

(13) Para la detección de multicolinealidad también se estimaron por Mínimos Cuadrados Ordinarios las regresiones de cada una de las variables explicativas con las restantes obteniéndose un R^2 máximo del 40% lo que corrobora la ausencia de multicolinealidad en el modelo planteado.

igual a 41.645,43 y el error medio relativo es 14,45%. Con objeto de determinar la capacidad predictiva del modelo se calcula el índice de desigualdad de Theil que toma un valor de 0,09. Su proximidad a cero muestra una elevada capacidad predictiva.

En el gráfico 1 se comparan los precios estimados a través del modelo hedónico seleccionado con los precios reales obtenidos en la muestra. La mayoría de los puntos del gráfico se posicionan muy próximos a la bisectriz del primer cuadrante. No obstante, puede apreciarse la existencia de un grupo de viviendas con precios altos –superiores a 400.000 euros– que quedan infravaloradas por el modelo, es decir, el modelo proporciona valores estimados inferiores al valor de mercado.

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



Redes Neuronales Artificiales (RNA)

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, pues siguiendo investigaciones previas en este campo es la que mejores resultados proporciona con diferencia en este tipo de análisis (véase p.ej. Haykin, 1999; Freeman y Skapura, 1993; García Rubio, 2004). La aproximación al precio de la vivienda a través de una red neuronal se realiza tomando como inputs las mismas variables internas y externas que fueron utilizadas en la estimación del modelo hedónico. Se pretende con ello efectuar comparaciones entre los resultados obtenidos con sendas estimaciones.

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 única capa oculta que presenta la estructura 6:6-6-1:1

(véase figura 1). Esto significa que el número de inputs es de seis (14), 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–.

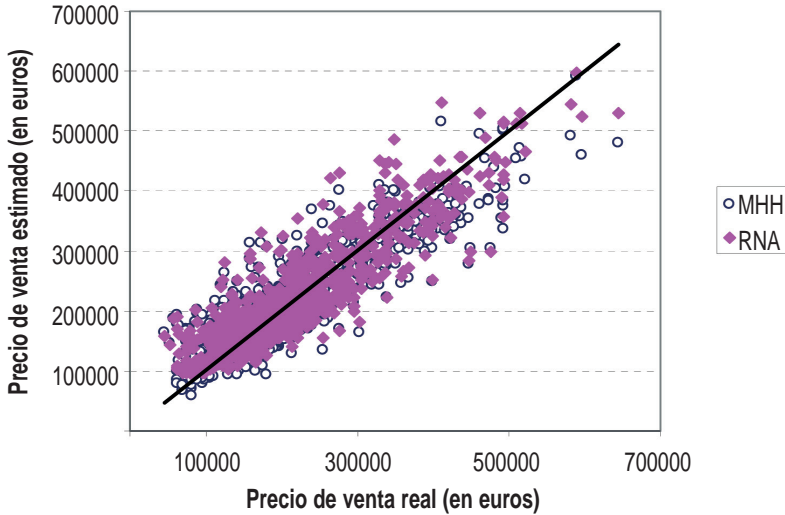
Para la estimación de la red neuronal se ha utilizado el programa Trajan Neural Networks.

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 sigmoidea (15). Como función de error se eligió la Suma de Cuadrados de los Errores.

(14) 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.

(15) 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.

Gráfico 2. Precio de venta real vs. Precio de venta estimado: modelo de precios hedónicos y Red Neuronal artificial



Para garantizar la capacidad generalizadora de la red el conjunto de observaciones de la muestra 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 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:

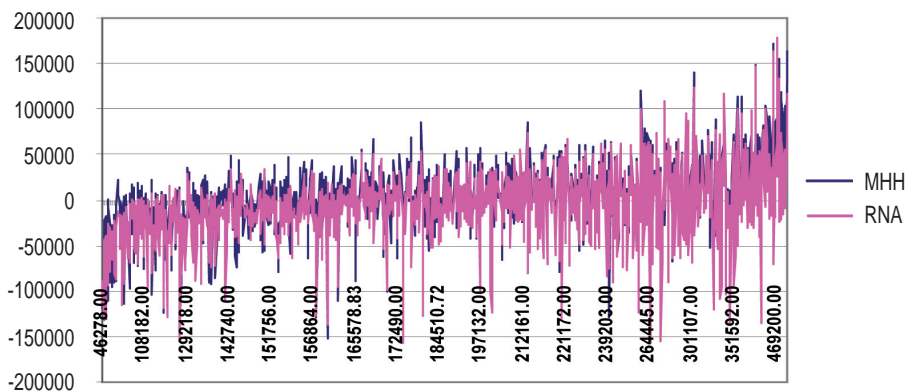
- Inicialización aleatoria de los pesos y umbrales.
- Máximo número de iteraciones: 5000.
- Ratio de aprendizaje: 0,1.
- Término de momento: 0,1.

En el gráfico 2 y en el cuadro 6 se presentan de forma comparativa los resultados obtenidos tras aplicar MPH y RNA.

Cuadro 6. Comparación de los resultados obtenidos con MPH y RNA

	Metodología de Precios Hedónicos	Red Neuronal Artificial
Coefficiente de Determinación (R ²)	77,38%	86,05%
Raíz del Error Cuadrático Medio (RECM)	41.645,43	39.540,36
Desviación típica residual	41.911,91	39.102,13
Error medio absoluto	30.579,18	28.551,34
Error medio relativo	14,45%	13,69%

Gráfico 3
 Errores cometidos por el Modelo de Precios Hedónico vs. Red Neuronal Artificial



El proceso de estimación mediante el uso de redes neuronales ofrece resultados más satisfactorios que la estimación mediante modelos hedónicos. Con la red se ha conseguido un grado de ajuste del 86% (R^2) frente al 77% alcanzado por el modelo hedónico de regresión. Además, al aplicar la red, la raíz del error cuadrático medio (RECM) disminuye –de 41.645,43 pasa a 39.540,36– y también se observa una clara disminución de la desviación típica residual, del error medio absoluto y del error medio relativo.

En el gráfico 2 se comparan los precios reales de venta de los inmuebles con los precios de venta estimados a partir de la red y del modelo hedónico seleccionado. En ambas metodologías la mayoría de los puntos se posicionan muy próximos a la bisectriz del primer cuadrante, no obstante pueden apreciarse algunas diferencias:

- El modelo hedónico presentaba un grupo de viviendas con precios superiores a los 400.000 euros que quedaban claramente infravaloradas y que con la aplicación de la red neuronal se han situado más próximas a la bisectriz del primer cuadrante.

- Existen determinados inmuebles con precios de venta reales inferiores a 180.000 euros que quedaron infravaloradas por el modelo de hedónico y para los que la red neuronal proporciona un precio estimado más próximo al real. No obstante, ésta por su parte sigue incurriendo, como el modelo hedónico, en cierta sobrevaloración para precios de venta reales inferiores a los 100.000 euros (véase también gráfico 3 comparativo de errores).
- En el gráfico 3 puede apreciarse que los errores positivos –por infravaloración– correspondientes al modelo hedónico suelen resaltar por encima de los errores que corresponden a la red, mientras que si nos fijamos en los errores negativos –por sobrevaloración– éstos son predominantemente atribuibles a la red y muy pocos de los picos negativos son relativos al modelo hedónico. En definitiva, la red corrige la infravaloración que realiza el modelo hedónico para determinados registros, pero por otra parte puntualmente excede las sobreestimaciones efectuadas por el hedónico.

Conclusiones

Teniendo presente los resultados obtenidos en este estudio para la determinación del precio total de la vivienda en Córdoba se propone la red neuronal estimada, puesto que posee una mayor capacidad para reconocer el proceso de formación de los precios.

No obstante, hay que hacer constar que la red neuronal propuesta ha sido obtenida tras un proceso de prueba y error. De no haber seguido esta táctica, los resultados podrían no haber indicado superioridad de la red. De hecho, algunos estudios previos llegaron a conclusiones bien distintas y rechazaron la red por su naturaleza de “caja negra”. Entre 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).

En cualquier caso, la utilización de redes neuronales se muestra apropiada para evitar la rigidez propia de los planteamientos de la óptica hedónica, en la que, entre otras limitaciones, destaca el elevado grado de multicolinealidad que suele estar presente entre las variables explicativas del precio del bien. Además, las redes ofrecen la posibilidad de introducir como inputs las variables originales —previa estandarización de las mismas—, mientras que el modelo hedónico necesita la generación de índices lo que implica pérdida de información y dificultades a la hora de interpretar los resultados. En este caso se han mantenido los índices como inputs de la red calculada no por necesidad, sino para poder efectuar comparaciones con el modelo hedónico. Otra ventaja de una RNA —que en este estudio se ha puesto de manifiesto— 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 el modelo hedónico.

Por último, 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 la ecuación hedónica, 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.

Bibliografía

- AGUILÓ SEGURA, P. M. (2002): El método de valoración de los precios hedónicos. Una aplicación al sector residencial de las Islas Baleares. Tesis Doctoral. Universidad de las Islas Baleares.
- BENGOCHEA MORANCHO, A. (2003): A hedonic valuation of urban green areas. *Landscape and Urban Planning*, 66: 35-41.
- BILBAO TEROL, C. (2000): Relación entre el precio de venta de una vivienda y sus características: un análisis empírico para Asturias. *Revista Asturiana de Economía*, 18:141-150.
- BOVER, O. Y VELILLA (2001): Precios hedónicos de la vivienda sin características: el caso de las promociones de viviendas nuevas. *Estudios Económicos*, nº 73. Banco de España. Madrid. 51 págs.
- CARIDAD, J. M. y BRAÑAS, P. (1996): Demanda de características de la vivienda en Córdoba: un modelo de precios hedónico. *Revista de Estudios Regionales*, 46: 139-153.
- 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*, 67-81.
- COLLINS, A. AND EVANS, A. (1994): Aircraft noise and residential property values. *Journal of Transport Economics and Policy*, Vol. 28(2): 175-197.
- COURT, A.T. (1939): Hedonic Price Indexes with Automotive Examples, *Dynamics of Automobile Demand*. General Motors Corporation, 99-117.

FREEMAN, J. and SKAPURA, D. M. (1993): Redes neuronales, algoritmos, aplicaciones y técnicas de programación. Ed. Wilmington Addison-Wesley Diaz de Santos, 26-241.

GALLEGO MORA-ESPERANZA, J. (2004): La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid. CT: Catastro nº 50, 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.

GRILICHES, Z. (1971): Introduction: Hedonic Price Indexes Revisited. In Price Indexes and Quality Changes: Studies in New Methods of Measurement. Griliches. Cambridge. Harvard University Press, 3-15.

HAYKIN, S. (1999): Neural networks: A comprehensive foundation. Ed. Prentice-Hall. 696 págs.

HILERA, J. R. y MARTÍNEZ, V. J. (1995): Redes neuronales artificiales: Fundamentos, modelos y aplicaciones. Ed. Ra-Ma. Madrid. 390 págs.

JAÉN, M. y MOLINA, A. (1995): Modelos econométricos de tenencia y demanda de vivienda. Servicio de publicaciones de la Universidad de Almería. 111 págs.

KAUKO, T. (2003): On current neural network applications involving spatial modelling of property prices. Journal of Housing and the Built Environment. Vol.18 (2): 159-181.

LANCASTER, K. J. (1966): A New Approach to Consumer Theory. Journal of Political Economy, 74: 132-157.

LARA CABEZA, J. (2005): Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria. Mapping, 104: 64-71.

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): 193-201.

Mc CLELLAND Y RUMELHART, D. (1985): Explorations in Parallel Distributed Processing, 1 y 2. MIT Press, Cambridge. 344 págs.

NÚÑEZ TABALES, J. et. al. (2007): Aproximación a la valoración inmobiliaria mediante la metodología de precios hedónicos (MPH). Actas de las XVII Jornadas Hispano-Lusas de Gestión científica celebradas en La Rioja.

NGUYEN, N. y CRIPPS, A. (2001): Predicting housing value: a comparison of multiple regression analysis and artificial neural networks. Journal of Real Estate Research, 22(3): 314-336.

PÉREZ DELGADO, M. L. y MARTÍN MARTÍN, Q. (2003): Aplicación de las redes neuronales artificiales a la estadística. Editorial La Muralla. Madrid. 140 págs.

QUANG DO, A. and GRUDNITSKI, G. (1992): A neural network approach to residential property appraisal. Real Estate Appraiser, Vol.58 (3): 38-45.

RIDKER, R. and HENNING, A. (1967): The determinants housing prices and the demand for clean air. Journal Environmental Economy Management, 5: 81-102.

ROSEN, S. (1974): Hedonic Prices and Implicit Markets: Product Differentiation in Pure competition. Journal of Political Economy, n. 82: 34-55.

SAURA, P. (1995): Demanda de características de la vivienda en Murcia. Secretariado de Publicaciones de la Universidad de Murcia. 157 págs.

SIRMANS, G.S. et. Al. (2005): The composition of Hedonic Pricing Models. Journal of Real Estate Literature. Cleveland. Vol. 13 (1): 3-43.

WALLACE, N. (1996): Hedonic –based price indexes for housing: Theory, estimation and index construction. Economic Review– Federal Reserve Bank of San Francisco, 34-48.

WAUGHT, F. V. (1929): Quality as a Determinant of Vegetable Prices. Columbia University Press. New York. 156 págs.

WORZALA, E.; LENK, M. and SILVA, A. (1995): An exploration of neural networks and its application to real estate valuation. Journal of REAL ESTATE RESEARCH, 10(2): 185-202. ■