

Departamento de Urbanismo



**VALORACIÓN DE INMUEBLES
URBANOS:
COMPARATIVA MODELO
REGRESIÓN MULTIVARIABLE
VERSUS REDES NEURONALES
ARTIFICIALES PARA LA CIUDAD DE
MORELIA, MICHOACÁN, MÉXICO.**

TESIS DOCTORAL

Presentado por:

D. José Carlos Preciado Carrillo

Dirigido por:

Director Dr. D. Francisco Guijarro Martínez

Valencia, Octubre de 2015

Agradecimientos

En primer lugar quiero expresar mi gratitud a mi director Francisco Guijarro Martínez, así como a Jerónimo Aznar Bellver, por el gran apoyo que con generosidad me brindaron, aunado al interés y asesoría, para la realización de esta tesis, además de sus atinados comentarios y la particular atención que en todo momento tuve de ellos para llevar a buen puerto el presente texto.

Doy gracias a Dios, ya que ha sido el medio que ha propiciado mi formación como profesional.

Asimismo a mis padres, que pesar de haberse adelantado en el camino, sin ellos yo no hubiera existido físicamente, pues gracias a que en aquel tiempo, en todo momento, desde la infancia hasta la madurez, me dieron la base de mi formación, brindándome lo mejor de ellos.

A Gaby mi esposa, con amor y cariño, por sus palabras de aliento en los momentos más difíciles, por su apoyo y comprensión para seguir adelante, pero además por los dos hijos maravillosos.

A mis hijos Carlos y David, que son el estímulo fundamental en mi esfuerzo para salir adelante y así formar para ellos un futuro mejor.

A mis hermanos, Micaela, Antonio, Gloria, Victoria, María, Luis y Gilberto por esos grandes momentos de la infancia y la adolescencia que juntos pasamos.

A mis compañeros, por su amistad y los momentos de convivencia, también por su confianza en la realización de una de mis metas: la conclusión de mi trabajo de tesis.

A todos mis maestros, por su esmerada atención y los conocimientos que me han compartido, por su esfuerzo y preocupación para hacer de los sueños de sus alumnos realidades.

RESUMEN

Obtener el valor de mercado para inmuebles urbanos es necesario para la toma de decisiones en diferentes escenarios: tratándose de instituciones financieras, su interés radica en conocer este valor para el financiamiento de préstamo con garantía hipotecaria, pues los diferentes órganos de gobierno para cobrar los impuestos catastrales se sirven de este valor como base para la realización de diferentes trámites ante entidades públicas o particulares. Para una buena estimación del precio que explique las condiciones del mercado, es necesario construir modelos capaces de estimar dicho precio. Actualmente los métodos sintéticos que usan información escasa son los más utilizados por los tasadores, así como los modelos econométricos que usan un gran número de observaciones. Los primeros, bajo la denominación de modelos sintéticos, son un caso particular de los modelos econométricos.

Por otro lado, las redes neuronales artificiales, y en concreto la red multicapa, pueden ser una alternativa para la estimación del precio. En la presente tesis doctoral se

realiza una comparativa entre el modelo econométrico y la red neuronal, demostrando que las redes obtienen mejores resultados en la estimación de precio. El estudio se ha realizado en la ciudad de Morelia, Michoacán, México, para varios segmentos de vivienda.

RESUM

Obtindre el valor de mercat per a immobles urbans és necessari per a la presa de decisions en diferents escenaris: tractant-se d'institucions financeres, el seu interès està en conèixer aquest valor per al finançament de préstec amb garantia hipotecària, als diferents ordres de govern per cobrar els impostos cadastrals. Així aquest valor servirà de base per a la realització de diferents tràmits davant entitats públiques o particulars. Per a una bona estimació del preu o valor que expliqui les condicions del mercat, cal construir models capaços d'estimar aquest valor. Actualment els mètodes sintètics que fan servir informació escassa són els més utilitzats pels valuadors, així com els models econòmics -també denominats de regressió lineal múltiple, que fan servir moltes observacions. En realitat els models econòmics són una generalització dels mètodes sintètics.

D'altra banda, les xarxes neuronals artificials, i en concret la xarxa multicapa, poden ser una alternativa per a l'estimació del preu. En el present treball es realitza una comparativa entre el model econòmic seleccionat i la

xarxa neuronal, demostrant que les xarxes obtenen millors resultats en l'estimació de preu. L'estudi s'ha realitzat a la ciutat de Morelia, Michoacán, Mèxic, per a diversos segments d'habitatge.

ABSTRAC

To obtain the market value for urban properties, the following is necessary for the decision making in different scenarios: in case of dealing with financial institutions, the focus will be in acquiring knowledge of the value of the property to finance a mortgage. One of the different government orders are to collect property taxes, so this profit will serve as the foundation for proceedings before public and private organizations. For an accurate estimate of the price or value that will describe the market conditions, it is necessary to build models that will be capable of calculating such value. The synthetic methods that currently use a minor amount of information are the ones that are mostly used by appraisers, as well as the econometric models. Also, they are considered as the concept of multiple linear regression that utilize abundant observations. In reality the econometric models are a generalization of synthetic methods.

On the other hand, artificial neural networks and in particular the multi-layer network, can be an alternative for a price estimate. Currently a comparison is being

performed between the selected econometric model and the neural network, revealing that the networks give a better outcome for price estimates. This study has been achieved in Morelia, Michoacán, Mexico for several housing segments.

Key words: Economic valuation, real estate assets, multiple linear regression, artificial neural networks.

Índice

| | |
|---|-----|
| I. Introducción y objetivos | 13 |
| 1. Introducción y objetivos | 14 |
| II. Antecedentes. Evolución histórica de la valoración | 23 |
| 2. Antecedentes. Evolución histórica de la valoración.. | 24 |
| III. Modelización de la tasación. Modelos econométricos | 40 |
| 3. Modelización de la tasación. Modelos econométricos | 41 |
| 3.1. Concepto de tasación..... | 41 |
| 3.2. Terminología relacionada con el valor..... | 41 |
| 3.3. Modelos de regresión simple | 44 |
| 3.4. Modelos multivariantes..... | 65 |
| 3.5. Modelo de precios hedónicos | 67 |
| IV. Las redes neuronales artificiales | 78 |
| 4. Las redes neuronales artificiales | 79 |
| 4.1. Introducción..... | 79 |
| 4.2. Concepto de una red neuronal artificial (RNA)... | 81 |
| 4.3. Antecedentes históricos | 85 |
| 4.4. Elementos de una RNA..... | 93 |
| 4.5. Principales funciones de activación | 96 |
| 4.6. Operación de la RNA..... | 104 |
| 4.7. El perceptrón simple | 111 |
| 4.7.1. Qué es un perceptrón | 111 |
| 4.8. El perceptrón multicapa | 112 |
| 4.9. Topología del perceptrón multicapa..... | 114 |
| 4.10. Parámetros que determinan el aprendizaje de la red | 115 |
| 4.11. Método de aprendizaje del perceptrón multicapa | 116 |
| 4.12. La red con conexiones hacia adelante (<i>back propagation</i>)..... | 123 |

| | |
|---|-----|
| 4.13. La regla delta generalizada | 125 |
| 4.14. Estructura y aprendizaje de la red <i>back propagation</i> | 131 |
| V. Estudio comparativo entre los modelos de regresión múltiple y las redes neuronales multicapa aplicadas a la valoración inmobiliaria | 136 |
| 5. Estudio comparativo entre los modelos de regresión múltiple y las redes neuronales multicapa aplicadas a la valoración inmobiliaria..... | 137 |
| 5.1. Introducción..... | 137 |
| 5.2. Normativa urbanística de la ciudad de Morelia | 138 |
| 5.3. Fuente de datos..... | 144 |
| 5.4. Modelización del precio a través de la regresión múltiple | 156 |
| 5.5. Modelización del precio mediante red neuronal artificial perceptrón multicapa | 179 |
| VI. Resultados de la aplicación de los modelos de regresión y redes neuronales artificiales | 317 |
| 6. Resultados de la aplicación de los modelos de regresión y redes neuronales artificiales..... | 318 |
| VII. Conclusiones finales y trabajos futuros | 325 |
| 7. Conclusiones finales y trabajos futuros | 326 |
| VIII. Anexos | 335 |
| 8. Anexos | 336 |
| IX. Bibliografía | 359 |
| 9. Bibliografía..... | 360 |

I. Introducción y objetivos

1. Introducción y objetivos

La generación económica que produce el ramo de la construcción versus la creación de riqueza de un país o región está fuera de toda duda. Si se revisa el producto interno bruto (PIB), en específico el de este ramo de cualquier país del mundo, se observará que la participación en la generación de riqueza es muy alto, y se pudiera decir que siempre va camino hacia arriba, a pesar de los tiempos de crisis.

La actividad de la construcción afecta a un gran número de sectores, como el inmobiliario, al generar empleos directos e indirectos, y otros como la industria del cemento o del acero, incluida la captación de impuestos para los gobiernos; que a su vez son ingresos que se desprenden de estas ramas de la economía.

El déficit de vivienda en el mundo difícilmente se podrá abatir, desde la vivienda subsidiada o protegida por los gobiernos de cada país, pasando por la vivienda de tipo medio, residencial o de lujo, siempre habrá interés por adquirir una vivienda ya sea nueva o usada; así pues, habrá necesidad de conocer los valores de mercado de las mismas.

La valoración en todo tipo de activos es imprescindible. En cualquier sociedad moderna es necesario conocer el precio de los activos implicados para un gran número de trámites o actos económicos. Empezando por expropiaciones del gobierno, pago de impuestos catastrales, compra-venta, seguros, crédito hipotecario, re-expresión de estados financieros, escisión, juicios y tomas de decisiones de empresas, entre otras.

Día a día la sociedad demanda gran cantidad de servicios de valoración de inmuebles de naturaleza urbana en todos los países, por tanto el mercado va exigiendo metodologías más consistentes que reflejen la dinámica de los precios de mercado.

El estudio acerca de la manera más objetiva de estimar los precios de mercado de los inmuebles ha suscitado un gran interés desde hace algunas décadas. Algunos autores como Caballer *et al.* (2007), afirman que desde los primeros indicios de la existencia de la actividad de la tasación o valoración vinculada a la asignación de un valor económico a un activo, bien o servicio -que se puede situar en el funcionamiento de los nilómetros del Antiguo Egipto, 3500 años a.C., hasta la aparición de los tratados específicos sobre la materia-, la formulación conceptual y

metodológica se reducía a unos elementos muy rudimentarios, susceptibles de ser reducidos al concepto de proporcionalidad o vinculados a la renta de la tierra, mediante ciertas normas empíricas, más o menos precisas, recogidas en tratados normativos de índole jurídica o religiosa. López (2013) en su estudio sobre “Sistemas de información geográfica y econometría espacial en la tasación de inmuebles urbanos”, comenta los cambios que en los últimos años ha habido en el sector inmobiliario, debido en primer lugar a los altos precios de los bienes inmuebles. Esto hace que al asignar un valor a una vivienda cada vez se exige mayor precisión ya que de ello pueden depender decisiones relevantes. Otro comentario adicional es que los profesionales en valoración no sólo deben conocer los procedimientos matemáticos, sino que también deben tener conocimientos sólidos en teorías económicas, financieras, derecho y estadística.

La valoración juega un papel importante, pues sin ella sería muy complejo realizar los diferentes trámites, por ejemplo el pago del impuesto catastral o el dejar garantía ante una institución financiera, por mencionar algunos de entre otros tantos usos que tiene el hecho de conocer el precio de un inmueble.

Aznar *et al.* (2008) menciona que cuanto mayor es el incremento económico, mayor es la necesidad de una mejor y más ajustada valoración de los activos de esa sociedad.

A través del tiempo se han utilizado en sus diferentes etapas multitud de métodos de valoración, entre los que se encuentran los denominados clásicos. Entre ellos se pueden mencionar los métodos sintéticos o factores de corrección directa, que consisten en comparar a través de diferentes variables explicativas una serie de muestras de mercado con el activo o inmueble a valorar. En buena medida los resultados a los que se llega no son justificables, pues se basan en un número de comparables muy limitado y en la subjetividad del valorador a la hora de ponderar las diferentes variables que influyen en el valor de mercado del activo. Existe la necesidad académica y profesional de investigar y proponer nuevas metodologías que ayuden a resolver estas deficiencias.

En la actualidad se aplican con carácter general tres métodos básicos de valoración para la obtención del precio de un inmueble. El primero de ellos es el denominado método de coste, consistente en obtener el valor del terreno, más el de su construcción con sus

respectivas depreciaciones por edad, uso, conservación, etc., en caso de no ser vivienda nueva, y sus instalaciones u obras complementarias (baldas perimetrales, jardines, etc.).

El segundo método es el de actualización de flujos de caja (actualización de rentas), en el que el valor se obtiene descontando los flujos de caja que resultan por el alquiler del inmueble. Está basado en el principio económico de actualización, y considera valores en relación con el valor presente de beneficios futuros derivado del ingreso por renta (alquiler) de la propiedad, utilizando una tasa de actualización, apoyada por algún instrumento financiero, menos gastos inherentes o de mantenimiento del activo.

Por último se destaca el método comparativo o de mercado, el cual consiste en estimar el precio aplicando factores por corrección de bienes iguales o similares al bien objeto en estudio que han sido vendidos o que se encuentran compitiendo en el mercado.

La mayor parte de las investigaciones empíricas desarrolladas sobre temas de valoración se basan en el método comparativo o de mercado, utilizando como

muestras principalmente precios de oferta, pues tratándose de un mercado que no es transparente y tan heterogéneo se recurre a este tipo de información.

Como parte fundamental de este análisis, se hará la modelización de la valoración de inmuebles residenciales.

El objetivo de la presente tesis doctoral es la comparación entre un modelo econométrico (regresión lineal múltiple) y una red neuronal del tipo multicapa *back propagation* para la estimación y predicción de precios de inmuebles sobre los diferentes segmentos de mercado inmobiliario. La hipótesis es que las redes neuronales artificiales tienen la ventaja de ser capaces de encontrar relaciones más complejas entre las variables y, por tanto, mayor capacidad explicativa de los precios de los inmuebles frente a los modelos clásicos econométricos.

Otro de los objetivos planteados está relacionado con la selección de un modelo econométrico para la modelización del precio haciendo una comparación entre el modelo de regresión lineal múltiple, el modelo de regresión no lineal y la red neuronal multicapa *back propagation* con salida lineal y no lineal; demostrando que

estos últimos presentan mejores resultados en las diferentes modelizaciones.

En el capítulo 2 se realiza una revisión histórica de la evolución de la valoración. Desde el Antiguo Egipto hasta la época del Imperio Romano, ha existido la necesidad de asignar un precio a los diferentes bienes, como base para el cobro de impuesto por el uso y goce de los mismos. Se comentan las aportaciones que han hecho diferentes autores para una mejora en el proceso de la estimación en los precios de los inmuebles. El objetivo es identificar si realmente la valoración por estas metodologías es más objetiva que los métodos tradicionales denominados sintéticos para los diferentes segmentos y calidades de inmuebles residenciales.

En el capítulo 3 se revisan los principales modelos de la familia econométrica (regresión lineal y no lineal). Se detallan modelos univariantes y multivariantes, aunque este último suele tener más aplicación por la cantidad de variables independientes que se incluyen en los modelos.

En el capítulo 4 se expone una visión general acerca de las redes neuronales: cómo se construyen y su proceso de aprendizaje. Posteriormente se profundiza en

las redes neuronales artificiales multicapa *back propagation*, indicando sus principales características así como su algoritmo de aprendizaje, la regla delta generalizada. Esta red será utilizada para el estudio comparativo, pues se trata de una las redes más utilizadas, no sólo en valoración, dada su capacidad de generalización.

En el capítulo 5, se estudia para varios segmentos o calidades de inmuebles residenciales el modelo econométrico (modelo de regresión multivariante), en sus diferentes modalidades lineal y no lineal, comparando entre ellos mismos los resultados. En una segunda parte se plantea una red neuronal multicapa *back propagation* para la explicación de precios, realizando una comparativa con el modelo econométrico. El capítulo concluye con los resultados obtenidos en función de los parámetros utilizados para la comparación de ambos modelos.

En el capítulo 6 se plantea un estudio para la comparativa entre un modelo econométrico y la red neuronal multicapa *back propagation* para determinar los resultados. Y se concluye con los datos obtenidos.

En el capítulo 7 se presentan las principales conclusiones que se han encontrado del modelo econométrico y la red neuronal, para la diferente calidad de los inmuebles residenciales. Los resultados obtenidos verifican que la red neuronal mejora los resultados de la explicación de precios. También se presentan líneas para investigación futura a partir de los resultados de este trabajo.

Finalmente se recoge la relación de referencias bibliográficas consultadas para la elaboración de la presente tesis doctoral, así como un anexo que resume algunos de los cálculos realizados en los programas estadísticos SPSS y Statgraphics.

II. Antecedentes. Evolución histórica de la valoración

2. Antecedentes. Evolución histórica de la valoración

La valoración dio inicio por primera vez en la agrimensura. Caballer (2008) afirma que la valoración es tan antigua, como la historia del hombre. El Antiguo Egipto, considerado como tal desde el cuarto milenio antes de Cristo hasta la aparición del Imperio Romano, es conocido como uno de los países más prósperos de su época, debido a la riqueza agrícola, proveniente de la influencia del desbordamiento del río Nilo sobre las tierras agrícolas del valle. Ello les permitía la producción con elevados rendimientos de cereales, como el trigo y la cebada, así como hortalizas (cebollas, pepinos, lechugas) y frutas (granadas, uvas, dátiles e higo).

De ahí surgió la necesidad de valorar las parcelas para asignarles un precio, pues las tierras más cercanas al río Nilo obtenían mejores cosechas y, por tanto, un mayor precio.

Aznar *et al.* (2008), en su libro *Valoración inmobiliaria, métodos y aplicaciones*, comentan los

evidentes desarrollos en países como Italia desde finales de la Edad Media con autores tan importantes como Elia del Re, Serperi y Medici. En España se tiene a autores como Ruiz Rochera en el siglo IX, o Torrejón y Boneta ya en el siglo XX. Estos autores afirman que la investigación sobre el tema de valoración comienza en el último tercio de siglo pasado, fundamentalmente por los profesores Ballesteros y Roca Cladera.

A lo largo del siglo XX la ciencia de la valoración ha adquirido un notable desarrollo en el seno de distintas escuelas. Destacan la escuela italiana, la española, la anglosajona y la hispanoamericana.

En México, citado por Aznar *et al.* (2008), se dice que existen indicios de valoración en la época prehispánica. Así se puede afirmar que después de la fundación de México-Tenochtitlán en 1325, los aztecas inician la conquista del territorio y la imposición de tributos, y la ciudad cuenta con un aparato administrativo fiscal integrado, entre otros, por el equivalente a un ministro de hacienda (*chihuacohualt*), y un tesorero general (*hueycalpixque*), este último entendido como un grupo numeroso de recaudadores.

En cuanto a los modelos econométricos, también denominados con el concepto de regresión multivariante, dentro de la valoración por regresión hedónica se puede citar a Hass (1922), pionero de un estudio hedónico que aplicó mediante el análisis de precios de explotaciones agrarias, con dos variables independientes: una de ellas fue la distancia entre el centro de explotación y el centro de la ciudad la ciudad, y la otra variable fue el tamaño de la ciudad. Hass realizó su estudio en el condado de Blue Earth, Minnesota, en los Estados Unidos de América. Por otra lado Wallace (1926), citado por García (2006), en su estudio denominado “Una aproximación a la aplicación de la metodología hedónica espacial” referencia al caso del mercado de la vivienda, siendo el segundo investigador en utilizar esta metodología. Por otro lado Waught (1928, 1929), analiza los precios de cierta producción agrícola, concretamente de los espárragos ofertados en el mercado de Boston durante los meses de mayo y julio del año 1927, intentando explicar los determinantes de las diferencias en el precio de venta de espárragos. Para ello Waught (1928, 1929) desarrollo una estimación de parámetros de una regresión lineal de la forma:

$$\text{Precio} = f(X_1, X_2, X_3, X_n)$$

donde el precio es función de las diferentes variables que explican el precio de un bien.

Caridad *et al.* (1997), en su trabajo denominado “Análisis intraurbano del precio de las características de la vivienda en Córdoba: los barrios menos favorecidos”, afirman la especial importancia que en los últimos años han ido cobrando los modelos derivados del de Tiebout, especialmente por las mejoras metodológicas introducidas desde la valoración por regresión hedónica (Griliches, 1961; Lancaster, 1966; Muth, 1969; y otros), y en particular el trabajo de Rosen (1974). Estos autores, al aplicar un modelo de regresión hedónica se apoyaron en un análisis de componentes principales para la reducción de factores o variables, encuentran los siguientes resultados: las viviendas más nuevas tienen mayor demanda, las personas de renta media-baja demandan viviendas grandes. Otra de sus conclusiones es que las viviendas de grandes dimensiones son más valoradas por los ciudadanos que otros atributos.

Olmo *et al.* (2007)” encuentran que el precio de la vivienda es una variable muy correlacionada espacialmente. Su estructura de variabilidad muestra dos escalas de variación que representan las variaciones

locales y regionales, respectivamente. Dichas estructuras tienen alcances de 300 a 400 metros para las variaciones locales y de 2000 metros para las regionales. El grado del polinomio lo determinaron tomando en cuenta la distribución espacial de los precios en la ciudad con resultados de un polinomio de forma acampanada, con valores altos en el centro de la ciudad, que van disminuyendo hacia la periferia. Las variables independientes explican el 70,5% de las variaciones de los precios de las viviendas.

Por otro lado Coremberg (2000), mediante un modelo de análisis econométrico realizado en Argentina, afirma que existen evidencias empíricas para el principal centro urbano de aquel país que confirman la hipótesis de que el precio real de la vivienda se encuentra determinado inversamente por el costo real de uso y el índice de valor de la bolsa a precios constantes, en tanto que se relaciona directamente con el costo de la construcción a largo plazo.

Fernández *et al.* (2008), mediante un modelo de retardo espacial o modelo de regresión espacial, han estimado la estructura de dependencia espacial de los inmuebles. De esta forma, la información aportada por

dicha estructura de correlación espacial mejorará las estimaciones de precios realizadas.

Figueroa y Lever (1992) utilizan un modelo de regresión hedónica lineal logarítmica, y encuentran una disminución de valor a mayor superficie del lote valorado. Afirman que mediante el estudio de éstas diferencias, se puede subir el precio de un metro cuadrado hasta siete veces.

García (2007) encuentra evidencia sobre la importancia de incluir dentro de los valores hedónicos externalidades (ruido, microlocalización, macrolocalización, etc.), e internalidades (superficie construida, edad, número de baños, número de recámaras, etc.).

García (2008) manifiesta que las distintas características de una vivienda se agrupan en tres categorías: las estructurales propias de la vivienda, las de localización y las características del vecindario y la calidad del entorno físico donde se ubica la vivienda. Para su trabajo ha utilizado la regresión hedónica semilogarítmica y ha obtenido resultados donde afirma que las características que más influyen en el precio de la vivienda

en la ciudad de Málaga son las de tipo estructural (superficie construida, número de baños, etc.).

Otra aportación es el caso de Komarova (2009). Mediante el uso del modelo de regresión hedónica, con 20.000 observaciones del precio de pisos donde incluye variables estructurales (internas) y variables espaciales (externas), con las que analiza la importancia de la variable distancia al centro de la ciudad, obtiene los siguientes resultados: por cada kilómetro al centro de la ciudad, el precio de las viviendas disminuye en 5.400 dólares; después de la variable superficie cubierta, la segunda más importante es la variable contaminación ambiental.

Un caso similar al anterior lo describe Kryvobokov *et al.* (2007), quienes mediante regresión hedónica, usaron variables estructurales de los inmuebles y además las variables espaciales (distancia al centro de la ciudad Donetsk –Ucrania-, distancia a la parada del autobús, parques, áreas verdes, etc.), con un universo de 325 transacciones de inmuebles, de tal manera que las variables espaciales resultaron ser más importantes que las variables estructurales del inmueble.

Nikolaos *et al.* (2011) confirman la influencia importante de las variables espaciales, en específico la variable espacios abiertos y proximidad a parques.

Por otro lado Hansen *et al.* (2013), mediante un modelo similar al anterior, encuentran una alta influencia de la vista hacia el lago o el océano. Otros trabajos como los de Milton (1984), alta influencia en el precio de parcelas cerca de la una bahía; Smith (1995), Lansford *et al.* (1995) y Doss *et al.* (1996), concluyen alta plusvalía a vista de un inmueble a un lago.

Lisi (2012) afirma que debido a la heterogeneidad en los precios de las viviendas y las características, la función de valoración nunca se comporta linealmente, sino que casi siempre tiende a una función logarítmica.

Kroll *et al.* (2010) emplean un modelo de regresión hedónico con la introducción de variables como espacios abiertos, plantación de nuevos árboles, uso de techos de color claro; obtienen resultados de la reducción de gases de efecto invernadero por debajo de 80% y el incremento del oxígeno de 1990 al año 2020 en el estado de California.

Payton *et al.* (2008), mediante la regresión hedónica, muestran que el efecto de la vegetación más verde que rodea a una propiedad es un aumento positivo en el precio.

Otra de las investigaciones sobre el ruido provocado por los aeropuertos que están cerca de los desarrollos residenciales ha demostrado que los precios se ven muy disminuidos por este hecho (Dubin 1988, 1992; Can 1990). Bowen *et al.* (2011) encontraron evidencia de cómo disminuyen los precios de las viviendas por el ruido y la contaminación debido a la cercanía a los aeropuertos.

Kim *et al.* (2003) han demostrado que con una mejora de la calidad del aire se obtiene un ahorro económico en los hogares de 2.333 dólares anuales que representa un 1,43% del valor medio de los hogares en Seúl, Corea, y afirman que el valor de la calidad del aire capitalizado en precios de la casa es el valor actual obtenido mediante este concepto.

Dentro del análisis de regresión se puede citar los trabajos de Aznar y Guijarro (2005), Aznar y López (2008), en los cuales han aplicado la regresión multivariada a valoraciones industriales e inmobiliarias.

Existen multitud de investigaciones donde se aplica la regresión hedónica a inmuebles urbanos, utilizando dos tipos de variables independientes, como son las estructurales (internas de las viviendas) y las espaciales (externas), macro localización y micro localización. Dentro de estas variables se encuentran: nivel socioeconómico, distancia a centro de ocio, distancia a centro de la ciudad, contaminación del aire, etc. Se pueden citar trabajos como Gao *et al.* (2002), “Evaluación empírica del modelo de regresión hedónica”, por otro lado los modelos econométricos se usan en diferentes áreas de la investigación, en la valoración inmobiliaria existen infinidad de aplicaciones, Páez *et al.* (s.f.), con su trabajo “Una comparación empírica de técnicas de modelación”, Chaudhry *et al.* (2013), sobre “Valoración de servicios ambientales para la ciudad Chadigarh, India”, Anseli (1999), “Econométrica espacial”, Segura *et al.* (1998), “Modelos econométricos de valoración aplicados a la valoración fiscal”, Bengochea (2002), “Una valoración urbana de las áreas verdes urbanas”. Los beneficios futuros con base en variables sociales y económicas, es actualmente otra de las aplicaciones, véase Blight *et al.* (2002).

Otros autores que coinciden con los anteriores en la aplicación del modelo de regresión hedónica son Beamonte (2008), Mendieta y Perdomo (2007), Segura *et al.* (1997) y Quispe (2012). En sus diferentes trabajos de investigación han utilizado la regresión hedónica semilogarítmica y logarítmica, para ello han utilizado variables estructurales y variables geoespaciales obteniendo resultados satisfactorios.

Tratándose de redes neuronales artificiales con aplicaciones a la valoración, en el año 1991 Borst realizó trabajos en este campo, en la ciudad de Nueva Inglaterra, Estados Unidos de América.

Otra investigación sobre redes neuronales artificiales y regresión hedónica es la de Tabales *et al.* (2009), sobre el tema denominado “Propuesta metodológica para valoración de inmuebles urbanos para la ciudad de Córdoba en España”, donde hace el contraste de la regresión hedónica y redes neuronales artificiales por modelo multicapa aplicado a valoración masiva, utilizando 1.189 muestras, obteniendo mejores resultados con el modelo de redes, que con la regresión hedónica.

Otra de las aplicaciones es la de Pitarque *et al.* de la Universidad de Valencia (España) acerca de la valoración inmobiliaria mediante redes neuronales artificiales.

Fernández *et al.* (2012) utiliza un modelo con tres capas: una de entrada, otra intermedia y una capa de salida con red de retropropagación, con 1.442 muestras entre viviendas nuevas y de segunda mano, clasificando tres categorías -alta, media y baja-.

Por otro lado Gallegos (2004) emplean un modelo de redes multicapa con salida hacia adelante con tres capas, una de entrada, intermedia y una capa de salida, con 12 variables que explican el precio de venta de las viviendas. En una primera prueba utilizan 100 muestras, de las cuales 85 son para el entrenamiento y 15 para la validación. También comparó este modelo con un modelo de regresión múltiple con las mismas 85 muestras. El resultado por regresión múltiple resultó de 15%, con una diferencia de 7%, pues de acuerdo con estos estudios encuentra evidencias donde las redes neuronales artificiales obtienen buenos modelos de aproximación en la valoración inmobiliaria.

Pi-ying (2011) desarrolla un modelo de propagación hacia atrás de la red neuronal artificial. Contrasta el modelo RNA contra un modelo de regresión hedónica, y en sus conclusiones afirma que el modelo de redes neuronales artificiales presenta errores menores con respecto de la regresión hedónica entre 10 y 15 por ciento.

Otra aplicación es el caso de Amabile *et al.* (1998). En su investigación con uso de redes neuronales en el análisis espacial de valores de propiedad, utiliza dos modelos de redes denominados perceptrón multicapa (MLP) y de función de base radial (RBF), con una muestra de 100 transacciones hechas en la ciudad de Treviso (Italia), tomando para ello el 80% de la muestra para el aprendizaje y el 20% para la validación. Obtuvo un resultado de 2,3% en valores por debajo del umbral de evaluación de 10%, con una media cuadrática de menos de 5% para la red de base radial con una alta capacidad explicativa del valor en ambos modelos.

Por otra parte Limsombunchai (2004) utiliza 200 observaciones, aplicando ambos modelos, y encontrando evidencias con resultados más favorables a las redes neuronales.

Amri *et al.* (2003) emplean un modelo de red neuronal con múltiples capas y otro modelo neuroborroso, y con base en la lógica borrosa con variables estructurales y geoespaciales; los resultados obtenidos en ambos modelos son muy parecidos, con conclusiones afines y positivas.

Rossini (1998), con el apoyo de estudiantes de la Universidad del Sur de Australia, ha estimado modelos de red neuronal artificial usando un método de propagación hacia atrás y una función sigmoide, con dos estructuras en cada caso. La primera estructura con tres capas: capa de entrada, capa oculta y la capa de salida. La capa de entrada tiene una neurona para cada variable independiente. La capa de salida sólo tiene una neurona de salida y es el modelo de estimaciones. Otro modelo utilizado es el análisis de regresión múltiple, donde encuentra evidencias de que a mayor cantidad de observaciones utilizadas en el modelo de regresión se presentan mejores resultados; en cambio al utilizar menos cantidad de observaciones las redes neuronales muestran mejores resultados, para lo que han utilizado 1.940 muestras en la ciudad de Adelaida (Australia).

Zurada *et al.* (2006) utilizan una red neuronal perceptrón completamente conectada, con una capa oculta, la capa de salida conteniendo un solo nodo, lo que representa el precio estimado. La capa de entrada tiene 13 nodos que representan las 13 entradas de valores. Probaron varias redes neuronales con diferentes números de neuronas en la capa oculta y una neurona en la capa de salida que representa el precio. Resultó que la red con tres neuronas en la capa oculta produce el menor error en las unidades de prueba. La función de activación de tangente hiperbólica se utilizó en las neuronas. La normalización de la desviación estándar se utilizó para las variables. Esta normalización resta la media y la divide por la desviación estándar, de manera que los valores resultantes tienen una media de cero y una desviación estándar de uno. Otro modelo utilizado es el de regresión múltiple, para la reducción de variables o factores, a través del uso del análisis de componentes principales. De tener 13 variables se redujeron a tres. Para llevar a cabo esta investigación se usaron observaciones de la ciudad de Louisville, Kentucky, Estado Unidos de América. El conjunto de datos fue de 360 casas unifamiliares de los años 1982 hasta 1992.

Otra aplicación del modelo de red neuronal artificial es en la simulación de inundaciones usando GIS; a las inundaciones provocadas por el río Johor, Malasia, en el área de la cuenca, Bakhtyari Kia *et al.* (2001), mediante el modelo de red neuronal perceptrón multicapa *back propagation*, utilizaron una capa de entrada que contiene siete neuronas (una por elevación, pendiente topográfica, la acumulación de flujo, datos de geología, uso de la tierra, el suelo y las precipitaciones), cada una representando un factor causal que contribuye a la aparición de la inundación en la cuenca de captación. La capa de salida contiene una sola neurona que representa el curso del agua. Las capas ocultas de neuronas se utilizan para definir la compleja relación entre las variables de entrada y de salida. Una vez definidas las variables de entrada y de salida con salidas no lineales como la ecuación sigmoide, encontraron coeficientes de determinación que van desde 0,93 hasta 0,99.

III. Modelización de la tasación. Modelos econométricos

3. Modelización de la tasación. Modelos econométricos

3.1. Concepto de tasación

Se puede definir la ciencia de la tasación o valoración como aquella parte de la economía cuyo objeto es la estimación de un determinado valor o valores, con arreglo a determinadas hipótesis, con fines determinados y a través de procesos de cálculo basados en informaciones de carácter técnico (Caballer, 1998).

3.2. Terminología relacionada con el valor

Antes de profundizar más en el tema de la valoración, es necesario definir conceptos como valor de mercado, precio, costo y valor de un activo, que son fundamentales para cualquier proceso de valoración.

El valor de mercado de un bien es una representación de la utilidad reconocida por el mercado más que su condición física. La utilidad de los activos para

una determinada organización o individuo, puede diferir de aquella que un mercado concreto le otorgue.

El precio es la cuantía que se solicita, y se paga por un bien o servicio. El precio de venta es un hecho histórico, ya sea de conocimiento público o confidencial. Debido a las capacidades financieras, a las motivaciones o intereses especiales de un determinado comprador o vendedor, es el precio que se paga por los bienes, y puede o no tener relación con el valor que otros individuos asignen a esos bienes o servicios. El precio es en general un indicador de un valor relativo que el comprador o el vendedor concreto le asignan a los bienes o servicios en determinadas circunstancias.

El coste es la cantidad requerida para producir el bien o servicio. Una vez elaborado dicho bien o servicio, el coste se convierte en un hecho histórico. El precio pagado por un bien o servicio se convierte en un coste para el comprador.

El valor es un concepto económico que se refiere al precio al que con mayor probabilidad realizarán las transacciones los compradores y vendedores de un bien o servicio disponible para su adquisición. El valor no es un

hecho, sino una estimación del precio probable que se pagará por los bienes o servicios en un momento dado, de acuerdo con una definición concreta de valor. El concepto económico de valor refleja la opinión del mercado sobre los beneficios que obtendrá quien posea el bien o reciba el servicio a la fecha efectiva de la valuación.

En el caso que se revisa, la valoración de inmuebles de uso habitacional será el objetivo del trabajo, pues existen otros rubros como son los inmuebles industriales, comerciales, etc.; estos últimos no se abordarán.

Bien inmueble se define como el terreno y cualquier otro elemento que sea parte del mismo, por ejemplo árboles o minerales y cualquier otra cosa que le haya sido añadido por el ser humano. También se consideran bienes inmuebles todos los anexos que sean instalaciones permanentes como cañerías, cableado eléctrico, refrigeración, calefacción, ascensores bajo rasante, etc. (IVSC, 2005).

3.3. Modelos de regresión simple

3.3.1. Introducción

El modelo de regresión simple fue ideado por el investigador sir Francis Galton, quien publicó su trabajo *Natural Inheritance* en el año 1889. Como su nombre indica, relaciona una variable independiente (exógena) o explicativa, y una variable dependiente denominada endógena. Se puede decir que Y está en función de X; se trata de una relación lineal de una variable Y con una variable X.

$$Y = f(x) = b_0 + b_1X \quad [3.1]$$

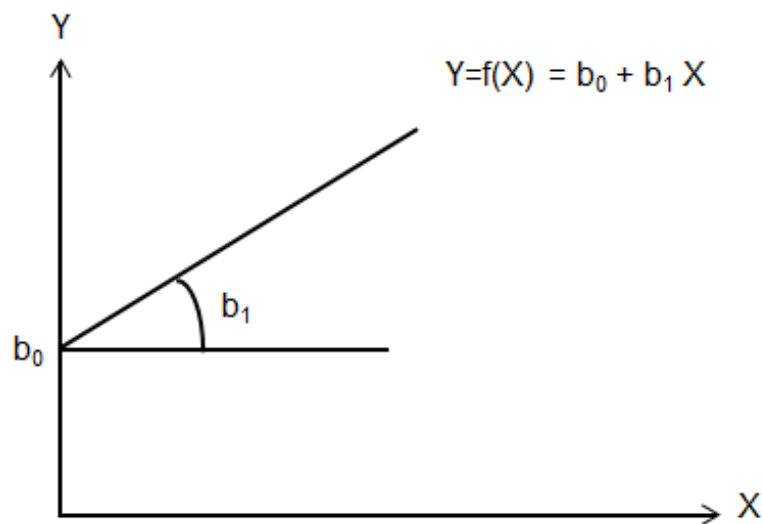


Figura 3.1. Fuente: Elaboración propia, readaptada de Guijarro (2013).

Donde b_0 es una constante llamada constante o intercepto, y b_1 es una constante llamada pendiente de la línea de regresión que denota el cambio que tendría Y por cada variación unitaria de X.

3.3.2. Estimación del modelo de regresión simple

En el modelo de regresión simple, los coeficientes del intercepto y la pendiente se calculan a partir de los valores muestrales de X e Y. Para clarificar la forma de obtener los coeficientes se utilizará el siguiente ejemplo en el cual figuran el precio y el área de 16 viviendas u observaciones que se distribuyen como se aprecia en la siguiente figura.

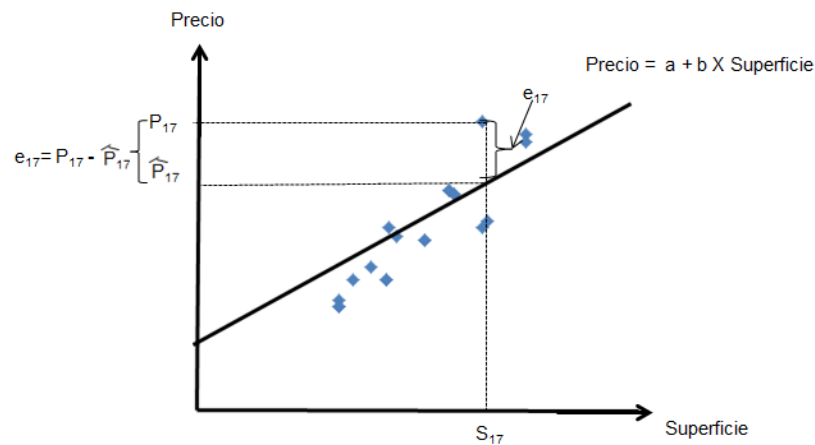


Figura 3.2. Fuente: Elaboración propia, readaptada de Guijarro (2013).

El modelo de regresión que se obtuvo, toma como variable dependiente el precio y como variable independiente la superficie, servirá para realizar predicciones o estimaciones de precios de viviendas según su superficie. Significa que, cuando se pretenda valorar una vivienda con superficie S_{17} , su valor estimado de \hat{P}_{17} será:

$$\hat{P}_{17} = a + bS_{17}$$

Teniendo en cuenta que se expresó la constante como a y la pendiente como b . Si se observa en la figura 3.2, se podría obtener el precio estimado de las 16 observaciones a partir de la recta de regresión. Supóngase por ejemplo el caso de la vivienda núm. 17, con precio P_{17} y superficie S_{17} . Al aplicar la expresión de la recta, su precio estimado sería:

$$\hat{P}_{17} = a + bS_{17}$$

No tiene por qué coincidir el precio real u observado P_{17} , con el precio estimado por el modelo de regresión. Se observa que el precio estimado de la recta de regresión es inferior al precio observado. La diferencia entre ambos precios recibe el nombre de residuo u error.

$$e_{17} = P_{17} - \hat{P}_{17} \quad [3.2]$$

Tal y como señala Guijarro (2013), la distancia al cuadrado entre el precio estimado y su proyección vertical sobre la recta es la medida que se quiere minimizar. De ahí la denominación de regresión por mínimos cuadrados o mínimo cuadrático, como en ocasiones se refieren al análisis de regresión. Los parámetros a y b son aquellos que hacen que la suma de los residuos al cuadrado sea lo más pequeña posible. El modelo de regresión se puede representar de la siguiente manera, optimizando la función objetivo y tantas restricciones de igualdad como observaciones se tengan en la muestra.

$$\begin{aligned} \text{Min } e_i^2 \quad & \text{for } i=1 \text{ to } n \\ e_i = P_i - \hat{P}_i = P_i - a + bS_i \end{aligned} \quad [3.3]$$

Volviendo al modelo más general, se intenta explicar una variable Y a partir de una variable X:

$$\text{Min } e_i^2 \quad i=1 \quad n \quad [3.4]$$

$$e_i = Y_i - a + bX_i$$

En la ecuación anterior la solución se obtiene de las siguientes expresiones de b y a:

$$b = \frac{\sigma_{xy}}{\sigma_x^2} \quad [3.5]$$

$$e_i = \bar{Y}_i - a + b\bar{X}_i \quad [3.6]$$

Donde σ_{xy} representa la covarianza entre las variables X e Y. σ_x^2 es la varianza de X. \bar{X} barra e \bar{Y} barra son los valores medios de X e Y, respectivamente.

Calculando el coeficiente de correlación lineal de Pearson, se mide en qué grado o qué tan relacionada está la variable dependiente y la independiente.

Mediante los siguientes factores se miden el grado de asociación o relación de la variable dependiente y la

variable independiente. En la siguiente tabla se observan las escalas de correlación lineal.

Tabla 3.1. Interpretación del coeficiente de correlación

| Valor absoluto mínimo | Valor absoluto máximo | Correlación |
|-----------------------|-----------------------|-------------|
| 0 | | Nula |
| 0.001 | 0.499 | Mala |
| 0.5 | 0.699 | Regular |
| 0.7 | 0.799 | Buena |
| 0.8 | 0.899 | Muy buena |
| 0.90 | 0.999 | Excelente |
| 1.00 | | Perfecta |

Fuente: Elaboración propia, readaptada de Ortiz (2003)

Mediante la siguiente ecuación se obtiene dicho coeficiente de correlación lineal.

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]} \quad [3.7]$$

Donde:

n es el número de pares de observaciones.

$\sum X$ es la suma de las variables X .

$\sum Y$ es la suma de las variables Y .

$(\sum X^2)$ es la suma de los cuadrados de la variable X .

$(\sum X)^2$ es la suma de las variables X , elevada al cuadrado.

$(\sum Y^2)$ es la suma de los cuadrados de la variable Y .

$(\sum Y)^2$ es la suma de las variables Y , elevada al cuadrado.

$\sum XY$ es la suma de los productos de X e Y .

Otra de las pruebas que se debe realizar es el cálculo del coeficiente de determinación, con el cual se trata de medir en qué proporción las variables o la variable independiente X , explica el valor de la variable dependiente Y .

Tabla 3.2. Ayuda para interpretar el coeficiente de determinación

| Valor absoluto mínimo[r ²] | Valor absoluto máximo[r ²] | Correlación |
|--|--|-------------|
| 0 | | Nula |
| 0.001 | 0.499 | Mala |
| 0.5 | 0.699 | Regular |
| 0.7 | 0.799 | Buena |
| 0.8 | 0.899 | Muy buena |
| 0.90 | 0.999 | Excelente |
| 1.00 | | Perfecta |

Fuente: Elaboración propia, readaptada de Ortiz (2003)

Mediante la siguiente ecuación se obtendrá el coeficiente de determinación que mide la proporción de la variación de Y, variable explicada y X variable explicativa independiente.

$$r^2 = \frac{SCR}{SCT} \sum_{i=1}^n (\hat{Y}_i - \bar{Y}_i)^2 / \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 \quad [3.8]$$

SCR: Suma cuadrática de los residuos

SCT: Suma cuadrática total

\hat{Y}_i : Precio estimado de las observaciones

\bar{Y}_i : Precio medio de observados

Y_i : Precios observados

r^2 : Coeficiente de determinación

r^2 ajustada.

Si se quieren comparar diferentes modelos, para verificar cual explica mejor el precio no se debe realizar a través del estadístico r^2 , sino a través de una variante del mismo; el estadístico r^2 corregido o ajustado, que toma en cuenta tanto el número de observaciones como el número de variables explicativas del modelo. Pues un modelo será mejor que otro si el r^2 ajustado es mayor, con independencia del r^2 . Por lo tanto a partir de ahora se debe observar el r^2 ajustado, que a diferencia del r^2 sí tiene en cuenta el diferente número de variables y observaciones entre dos modelos de regresión.

Su expresión matemática es la siguiente.

$$r^2_{ajustada} = 1 - \frac{\text{Suma de cuadrados del error (SCE)}}{\text{Suma de cuadrados residuos del modelo (SCR)}} \quad [3.9]$$

Donde:

r^2 = r cuadrada ajustada.

SCE = Suma de los cuadrados del error = $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$,
diferencia entre valor observado y valor estimado.

Y = Valor observado.

\hat{Y} = Valor estimado por el modelo.

SCR = Suma de los cuadrados del modelo = $\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$,
suma de los cuadrados de las diferencias entre cada valor
estimado y el valor promedio de las observaciones.

Otra forma de medir la bondad entre dos o varios
modelos, es determinar cual explica mejor el precio de los
inmuebles, lo cual es posible saber mediante el error
absoluto medio entre el precio observado y el precio
explicado. Su ecuación es la siguiente.

$$EAM = \sum_{i=1}^n \frac{|\hat{Y}_i - Y_i|}{n} \quad [3.9A]$$

Donde:

EAM = Error absoluto medio

$\sum_{i=1}^n$ = Sumatoria desde i hasta n .

\hat{Y}_i = Precio estimado.

Y_i = Precio observado.

n = Número de observaciones.

Otra de las bondades que se medirán entre el precio observado y el precio estimado, para comparar entre un modelo y otro, es la raíz del error cuadrático medio, utilizando la siguiente ecuación.

$$RECM = \sqrt{\frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{n}} \quad [3.10]$$

Donde:

RECM= Raíz del error cuadrático medio

$\sum_{i=1}^n$ = Sumatoria desde i hasta n.

\hat{Y}_i = Precio estimado.

Y_i = Precio observado.

$\sum_{i=1}^n (\hat{Y}_i - Y_i)^2$ = Sumatoria de la diferencia entre el precio estimado y el precio observado elevado al cuadrado.

n= Número de observaciones.

A continuación se definen algunos términos estadísticos con sus respectivas ecuaciones.

Los estadísticos de dispersión informan sobre la variabilidad o heterogeneidad en la distribución de datos. En cualquier tipo de investigación o toma de decisión, una vez que se tengan resultados se deben analizar con la intención de verificar qué tan homogéneos son esos resultados, para ello se deben usar las siguientes medidas.

Varianza. Es el promedio de la suma de todos los cuadrados de las desviaciones de una de la muestra o población.

La varianza se define a través de la siguiente expresión.

$$\text{Varianza}(X) = \sigma^2_x = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \quad [3.11]$$

En donde:

σ^2 = varianza

X_i = una observación individual en el conjunto

\bar{X} = media de la muestra

n = número de observaciones en la muestra

Tabla 3.3. Ejemplo de las viviendas con los siguientes valores correspondientes a un cierto centro urbano de una ciudad

| Número de vivienda | Precio (\$) X_i | $X_i - \bar{X}$ | $(X_i - \bar{X})^2$ |
|--------------------|-------------------|-----------------|---------------------|
| 1 | 2040000 | -503333 | 253344444444 |
| 2 | 3400000 | 856667 | 733877777778 |
| 3 | 3500000 | 956667 | 915211111111 |
| 4 | 1900000 | -643333 | 413877777778 |
| 5 | 2420000 | -123333 | 152111111111 |
| 6 | 2000000 | -543333 | 295211111111 |

En la tabla 3.3, se observan 6 muestras que se han recogido con precios que aparecen en la segunda columna. La media es el precio que toma el siguiente valor:

$$\bar{X} = \frac{2040000 + 3400000 + \dots + 2000000}{6} = 2,443,333.00 (\$)$$

La tercera columna refleja el precio de cada inmueble, respecto a la media ($X_i - \bar{X}$), en la cuarta y última columna, se elevan al cuadrado los valores de la columna anterior $(X_i - \bar{X})^2$.

Uno de los problemas que se le atribuye a la varianza es que su unidad de medida no es la misma que la unidad de medida de la variable sobre la que se ha

hecho su cálculo. Si el precio de un producto se mide en euros, la varianza del precio no se mide en euros. La unidad de medida que resulta al aplicar la varianza para el caso sería euros al cuadrado, ya que la varianza se ha construido como un promedio de las diferencias cuadráticas entre precios. Es precisamente lo que hace evidente que los valores obtenidos puedan parecer exagerados.

Por esta razón para medir de mejor manera la dispersión de una variable es más habitual emplear la desviación típica.

Desviación típica. La desviación típica, también denominada desviación estándar, se define como la raíz cuadrada de la varianza.

$$\text{Desviación típica } (X) = \sigma_x = \sqrt{\text{Varianza } (x)} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} \quad [3.12]$$

La principal virtud es que mantiene la unidad de medida de la variable a la que se aplica su cálculo. Si al ejemplo anterior de la tabla 3.3 se aplica la desviación típica.

$$\text{Desviación típica } (X) = \sigma_x = \sqrt{2,543,333} = 1,595$$

Como se aprecia su unidad de medida se expresa para este caso en euros. La desviación típica es la función más ampliamente utilizada en inferencia estadística.

Inferencia estadística. Al identificar a la población, se define una unidad observable, también denominada muestra, se recoge información de la misma para interpretar o deducir resultados de alguna o algunas características particulares sobre el caso estudiado.

Suponga que los precios de vivienda en un sector de alguna ciudad siguen una distribución normal y que una muestra obtenida es representativa del comportamiento de los precios en el resto del sector de la ciudad. En este caso, se puede inferir que la media del precio de la vivienda en toda la ciudad coincide con la media de la muestra tomada (Guijarro, 2013)

$$\text{Media (ciudad)} = \text{Media (muestra)} = 2,543,333 (\$)$$

Pero además se puede acotar entre qué rangos de valores se mueve una gran cantidad de viviendas (Guijarro, 2013). Se puede afirmar, también, que el precio de estas viviendas estará dentro de los siguientes rangos para los niveles de confianza de 90%, 95% y 99%:

$$\begin{array}{l} \text{Para un nivel de confianza del 90\%:} \\ \bar{X} \pm 1.645\sigma_x = [\bar{X} - 1.645\sigma_x, \bar{X} + 1.645\sigma_x] \end{array} \quad [3.13]$$

$$\begin{array}{l} \text{Para un nivel de confianza del 95\%:} \\ \bar{X} \pm 1.96\sigma_x = [\bar{X} - 1.96\sigma_x, \bar{X} + 1.96\sigma_x] \end{array} \quad [3.14]$$

$$\begin{array}{l} \text{Para un nivel de confianza del 99\%:} \\ \bar{X} \pm 2.57\sigma_x = [\bar{X} - 2.57\sigma_x, \bar{X} + 2.57\sigma_x] \end{array} \quad [3.15]$$

Reiterando, los cálculos anteriores serán válidos si 1) los precios siguen una distribución normal y 2) la muestra es una representación del conjunto de viviendas de toda la ciudad, se podrá afirmar que la media de los precios se mueve en los siguientes intervalos (Guijarro, 2013).

Nivel de confianza al 90%

$$\bar{X} \pm 1.645\sigma_x =$$

$$[2,543,333.33 - 1.645 \times 724,808.00 \quad 2,543,333.33 + 1.645 \times 724,808.00]$$

$$[1,351,024 \quad 3,735,643]$$

Nivel de confianza al 95%

$$\bar{X} \pm 1.96\sigma_x =$$

$$[2,543,333.33 - 1.96 \times 724,808.00 \quad 2,543,333.33 + 1.96 \times 724,808.00]$$

$$[1,122,710 \quad 3,963,957]$$

Nivel de confianza al 99%

$$\bar{X} \pm 2.57\sigma_x =$$

$$[2,543,333.33 - 2.57 \times 724,808.00 \quad 2,543,333.33 + 2.57 \times 724,808.00]$$

$$[680,577 \quad 4,406,090]$$

Si se extrae otra muestra de la misma ciudad, el precio medio de las viviendas estará dentro del intervalo [1'122,710 3'963,957] con una probabilidad o nivel de confianza de 95%. No todas las muestras tendrán el precio de sus viviendas dentro de este rango, pero se prevé que en 95% de los casos sí ocurrirá (Guijarro, 2013).

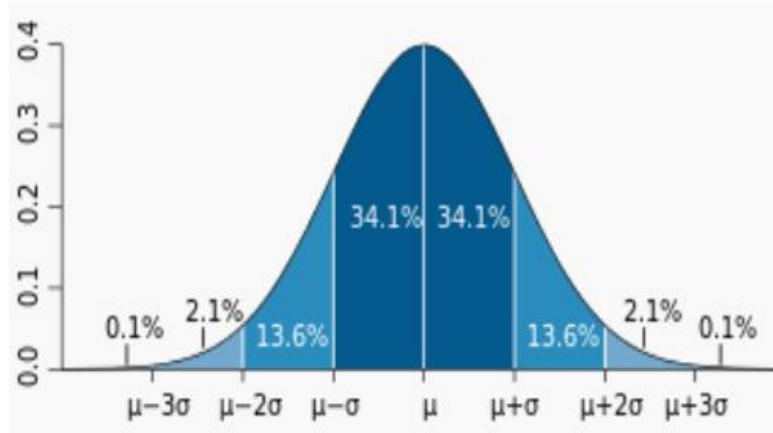
Dentro de los diferentes tipos de distribuciones que existen, como poisson, beta, gama, tao, normal, etc., la que más se usa es la distribución normal, pues es la que mejor representa los fenómenos físicos en la investigación científica; es por ello que se utiliza para la presente investigación.

Distribución normal estándar.

La regla empírica asume que:

1. Cerca de 68% del área bajo la curva se encuentra $\mu + (-)1$ desv. estándar
2. Cerca de 95% del área bajo la curva se encuentra $\mu + (-) 2$ desv. estándar
3. Cerca de 99.5% del área bajo la curva se encuentra $\mu + (-) 3$ desv. estándar

Figura 3.3. Representación de la distribución normal



Fuente: Imagen tomada a partir del buscador Google¹

La figura 3.3 representa la distribución normal, también denominada campana de Gauss. La desviación típica también se utiliza habitualmente para medir datos atípicos, también denominado *outliers*; y niveles de confianza de una población o muestra de observaciones, pues de esta manera se puede estimar si una muestra

¹https://www.google.com.mx/search?q=distribucion+normal&biw=1167&bih=546&tbm=isch&imgil=74YCzM7UBEE3M%253A%253BKQulledMTsEkEM%253Bhttps%25253A%25252F%25252Fes.wikipedia.org%25252Fwiki%25252FDistribuci%25252525C3%25252525B3n_normal&source=iu&pf=m&fir=74YCzM7UBEE3M%253A%252CKQulledMTsEkEM%252C_&usg=__XUB9f3T6Luv4qakAoGQUs0yJhDc%3D&ved=0CDsQyjdqFQoTCNTs25O58sYCFRcPkgod7f0BzQ&ei=Gn6xVdT3OpeeyATt-4foDA#imgsrc=74YCzM7UBEE3M%3A&usg=__XUB9f3T6Luv4qakAoGQUs0yJhDc%3D

representa características particulares de una población normal.

Datos atípicos o anómalos. Son aquellos que no representan, ni son representados por la tendencia central, pues se separan mucho del comportamiento medio de las observaciones (Guijarro, 2013).

La ecuación para identificarlos es la siguiente:

$$e_i^{estandar} = \frac{e_i - \bar{e}}{\sigma_e} \quad [3.16]$$

donde:

$e_i^{estandar}$ = Residuo estandarizado de la observación i – ésima

e_i = Residuo de la observación i – ésima

\bar{e} = desviación típica de los residuos

Estos residuos estandarizados permiten identificar los datos atípicos, mediante modelo regresión o modelo matemático, y tienen la característica de tener media cero y desviación estándar 1. La expresión anterior se simplifica a:

$$e_i^{estandar} = \frac{e_i}{\sigma_e} \quad [3.16A]$$

donde:

$e_i^{estandar}$ = Residuo estandarizado de la observación i – ésima

e_i = Residuo de la observación i – ésima

σ_e = desviación típica de los residuos

3.4. Modelos multivariantes

3.4.1. Introducción

El modelo de regresión múltiple es una extensión del modelo de regresión simple. En lugar de explicar la variable dependiente con una sola variable independiente, en este modelo se incluyen más de dos variables independientes. Al aumentar la cantidad de variables independientes, ayuda al modelo a tener una mejor capacidad explicativa, gracias a un mayor valor en el estadístico R^2 . Al adicionar más variables independientes suele aumentarse la cantidad de observaciones para una mejor explicación del mismo modelo.

3.4.2. Descripción de la ecuación

La forma descriptiva general de la ecuación de regresión lineal múltiple es la siguiente:

$$Y = f X = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \gamma \quad [3.17]$$

Donde Y es la variable dependiente, que en el caso que se estudia sería el precio, β_0 es la constante o

intercepto en el eje de la ordenada, X_i es la i -ésima variable explicativa y β_i es el coeficiente asociado con la i -ésima variable explicativa.

La interpretación de los coeficientes varía con respecto de las variables explicativas asociadas en el modelo.

El coeficiente β_0 es la intercepción o regresando de Y , cuando todas las variables explicativas toman valor cero. El β_i es la cantidad en que Y cambia cuando la variable X_i se incrementa en una unidad, manteniendo constante el resto de las variables, *ceteris paribus* γ representa el error cometido entre el precio observado y el precio estimado.

Al aplicar la regresión lineal múltiple, es necesario realizar las pruebas estadísticas que se explicaron en la regresión lineal simple, en este caso debe realizarse para más de una variable independiente.

Se ha explicado anteriormente la aplicación del análisis por regresión. Véanse los trabajos realizados por Aznar y Guijarro (2005), Aznar y López (2008), Guijarro (2013), Caballer y Guadalajara (2005), quienes al usar la regresión multivariable con diferentes variables cuantitativas obtuvieron buenos resultados.

3.5. Modelo de precios hedónicos

3.5.1. Introducción

La presencia de bienes con características propias con otros que podrían ser diferenciados, cercanos a ellos, es muy habitual en el mercado, véase García (2006). Las empresas para hacer más atractivos sus productos, van incorporando atributos que hacen estos bienes diferenciados, que a su vez los hace fácilmente identificables, con una infinidad de oferta los consumidores pueden adquirir estos productos para satisfacer sus necesidades, sus deseos o decisiones de compra. Estos efectos de análisis económicos han motivado la creación de nuevas metodologías para interpretar mediante la oferta y la demanda los precios marginales que están dispuestos a pagar los consumidores.

A mediados del siglo XX se empieza a utilizar el análisis económico para la investigación relacionada con la utilidad o satisfacción de algún producto de consumo con características heterogéneas o diferenciadas, conocidas con el nombre de teorías de precios hedónicos.

3.5.2. Aplicaciones recientes diferentes investigadores

La formación espacial de los precios de inmuebles ha sido una de las principales líneas de investigación de la economía urbana, Richardson (1971), citado por Humarán y Roca (2010). En general, los valores de los inmuebles se desagregan en dos componentes endógenos: el valor del suelo y el valor de las mejoras que se componen de las edificaciones e instalaciones y las obras complementarias.

Caridad *et al.* (1997), Olmo y Cuervos (2007), Fernández y Larráz (2008), Nikolaos *et al.* (2011), García (2004) García (2007), Komarova (2009), Kryvobokov *et al.* (2007), en sus diferentes trabajos acerca del modelo hedónico espacio-temporal y el análisis variográfico del precio de la vivienda, encuentran evidencias de que el precio de la vivienda es una variable muy correlacionada espacialmente. Su estructura de variabilidad muestra dos escalas de variación espacial fruto de los factores micro y macrolocalizativos.

Entendiéndose los factores microlocalizativos como todas aquellas variables externas que son propias del vecindario o barrio donde se ubica la vivienda, estatus

social, calidad del entorno físico, las distancias geoespaciales al centro del barrio que generen valor.

Los factores macrolocalizativos son variables externas propias de las distancias geoespaciales al centro de la ciudad o población que generan una diferencia en el precio, distancia a las escuelas, parada de autobuses, centro de ocio, hospitales, centros de trabajo, bancos, ruido, calidad de aire y otras tantas más variables que pueden influir en el precio de los inmuebles.

Estos dos factores que se han mencionado se denominan factores externos.

Variables internas o estructurales propias del inmueble o vivienda como son superficie construida, superficie del terreno, antigüedad, número de baños, número de recámaras, calidad de los acabados, espacio de garaje, y otras variables que se pudieran incluir en el modelo dependiendo de cada ciudad, que expliquen el precio integrador de la vivienda.

3.4.1. Fundamentos teóricos

En cuanto a los modelos de precios hedónicos, se puede citar a Hass (1922) quien es pionero de un estudio

hedónico que aplicó mediante un análisis de precios de explotaciones agrarias, analizando dos variables independientes, una de ellas la distancia entre el centro de explotación y de la ciudad. Hass realizó su estudio en el condado de Blue Earth, Minnesota, Estados Unidos de América, y la otra variable fue el tamaño de la ciudad.

Por otro lado Wallace (1926) hace referencia al caso del mercado de la vivienda, y fue el segundo investigador en utilizar esta metodología. Según Colwell y Dilmore (1999), el segundo investigador que utilizó esta metodología para analizar también el valor de las explotaciones agrarias. Este estudio fue realizado en el estado de Iowa, Estados Unidos de América.

Waught (1928, 1929) realiza un análisis de precios de cierta producción agrícola, concretamente de los espárragos ofertados en el mercado de Boston durante los meses de mayo a julio de 1927, intentando explicar los determinantes de las diferencias, los mismos precios medios de una unidad de venta en espárragos. Por ello Waught implementó una estimación de parámetros de una regresión lineal de la forma.

$$\text{Precio} = f(\text{Atributos o características del bien})$$

$$\text{Precio} = f(V_i + V_{ema} + V_{emi})$$

Donde V_i son las variables internas o estructurales, V_{ema} las variables externas macrolocalizativas y V_{emi} las variables externas microlocalizativas.

Los primeros intentos por construir una formulación teórica de este enfoque fueron concretados por Houthakker (1952) y por Tinbergen (1956), citados por García (2007), si bien no es hasta mediados de los años sesenta cuando se realiza, por parte de Lancaster (1966, 1971 y 1979), un desarrollo de la teoría del comportamiento del consumidor orientado hacia la demanda de bienes heterogéneos con características identificables y valorables objetivamente. De esta forma, y como señala Hulten (2003), citado por García (2007), se le puede considerar a Lancaster (1966), el primero en aportar la literatura básica de la economía elaborada para la comprensión de los precios de los atributos de un bien. Él afirma que un bien por sí mismo no genera utilidad para los consumidores, sino que ésta proviene de las múltiples características o cualidades intrínsecas que el bien posee.

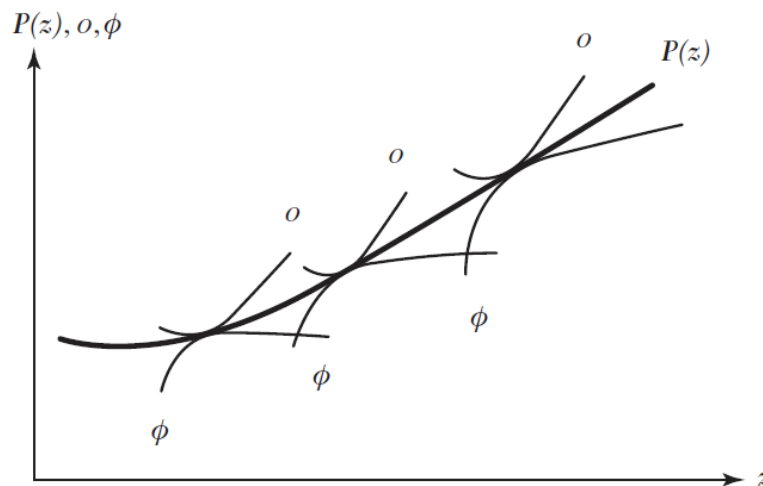
En 1974, con el artículo de Sherwin Rosen (1974), se establecen por primera vez las bases teóricas

consistente y metodológicamente suficiente para justificar la utilización de la técnica. Rosen analiza varios tipos de bienes diferenciados que pueden ser descritos por un vector de características: $Z = (z_1, z_2, \dots, z_n)$, donde z_i mide la cantidad de la característica i . De la observación del precio de cada bien y la cantidad de características asociadas a él, resulta un conjunto particular de precios hedónicos o implícitos. Oferentes y demandantes distinguen perfectamente entre cantidades de las características, eso implica que toda una variedad de alternativas (diferentes tipos de combinaciones de cada z_i) esté disponible. Así los mercados de estos productos implícitamente revelan una función $P(Z) = P(Z_1, \dots, Z_n)$, que relaciona precios y características o atributos, por tanto el consumidor y el productor toman sus decisiones de consumo y de producción en ese precio.

Rosen definió los precios hedónicos como aquellos precios implícitos con base en las características o atributos reflejados por los agentes económicos a partir de precios observados en el mercado de productos diferenciados y de las cantidades específicas con sus respectivas características.

Por otra parte es entendible que estos precios hedónicos no son fáciles de estimar, debido a las observaciones $P(Z)$, que representan un conjunto de funciones de valoración y otras de funciones de oferta lo que lleva a un problema de identificación, véase figura (3.4).

Función de precio hedónico



Fuente: Quispe (2012).

En resumen una función de precios hedónicos (FPH), es la función de equilibrio de mercado provocada por la interacción de las funciones de precios de los compradores y las funciones de precios de los vendedores. En el caso que aquí se revisa, si la vivienda es un bien heterogéneo diferenciado por un conjunto de atributos $H = (h_1, h_2, \dots, h_k)$, la función de precios hedónicos establece la

relación funcional entre los gastos observados de los hogares en vivienda $P(H)$ y el nivel de características incluidas en el vector H . Esta igualdad se puede representar como sigue:

$$P(H) = f(h_1, h_2, \dots, h_k)$$

Por tanto, el precio de cualquier atributo o característica contenido en H , $P_k (= \chi P(H) / \chi h_k)$ puede ser considerado como el precio implícito de equilibrio de ese atributo. Con la debida especificación funcional de la FPH, los coeficientes estimados proporcionarán los precios marginales estimados de las características. Por tanto una manera de interpretar la función de precios hedónicos es asumiendo que si el precio de una vivienda es 2000 unidades monetarias mayor que otra con las mismas características, excepto el garaje, entonces esta diferencia sería el precio marginal o implícito que estarían dispuestos a pagar los demandantes por tal diferencia.

De este modo a partir del desarrollo empírico se pretende estimar la relación existente entre los diferentes precios de las distintas viviendas de un mismo bien diferenciado o heterogéneo $P(Z) = P(Z_1, \dots, Z_n)$ y las distintas características o atributos que componen cada

una de esas viviendas $P(H) = f(h_1, h_2, \dots, h_k)$, utilizando para ello la técnica de regresión. Se intentará estimar la relación.

$$P(H) = f(h_{1i}, h_{2i}, h_{3i}, \dots, h_{ni}, \gamma_i) \quad [3.18]$$

A partir de la ecuación anterior de esta relación estimada entre el precio y las características (relación hedónica), se pueden obtener las valoraciones marginales implícitas (precios hedónicos) de cada uno de los atributos, aplicando la derivadas parciales de cada característica: $\chi P(H)/\chi h_k$.

O expresado en forma de regresión:

$$Y = fP = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \gamma \quad [3.19]$$

Donde P es la variable dependiente, que en este caso sería el precio, β_0 es la constante o intercepto en el eje de la ordenada, X_i es la i-ésima variable explicativa o atributo y β_i es el coeficiente asociado a la i-ésima variable explicativa o pesos marginales de los atributos o características.

La interpretación de los coeficientes varía con respecto de las variables explicativas asociadas con el modelo.

El coeficiente β_0 es la intercepción o regresando de Y, cuando todas las variables explicativas toman valor cero. El β_i es la cantidad en qué Y cambia cuando la variable X_i se incrementa en una unidad, y se mantiene constante el resto de las variables, *ceteris paribus*, γ representa el error cometido entre el precio observado y el precio estimado.

Tratándose de bienes heterogéneos, como es el caso de las viviendas, la mayoría de las investigaciones aplican el modelo de forma semilogarítmica, tal como se muestra a continuación:

$$\ln(P) = \beta X + \gamma \quad [3.20]$$

Donde P representa el vector de precio de venta de las viviendas observadas, β es la matriz de coeficientes, es la cantidad en qué P cambia cuando la variable X_i se incrementa en una unidad, manteniendo constante el resto de las variables, *ceteris paribus*, γ representa el error cometido entre el precio observado y el precio estimado.

La ecuación [3.20] al trabajar con más de dos variables independientes se transforma en la siguiente ecuación:

$$\ln(P) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \gamma \quad [3.21]$$

IV. Las redes neuronales artificiales

4. Las redes neuronales artificiales

4.1. Introducción

Redes neuronales biológicas. El sistema neuronal del ser humano y el de los animales está compuesto por el sistema nervioso y el hormonal. El diseño general del sistema nervioso está conformado por la neurona, unidad funcional básica del sistema nervioso central, el cual se estima que contiene 100,000 millones de neuronas (Gayton-Hall 2009).

Estructura de una neurona biológica típica.

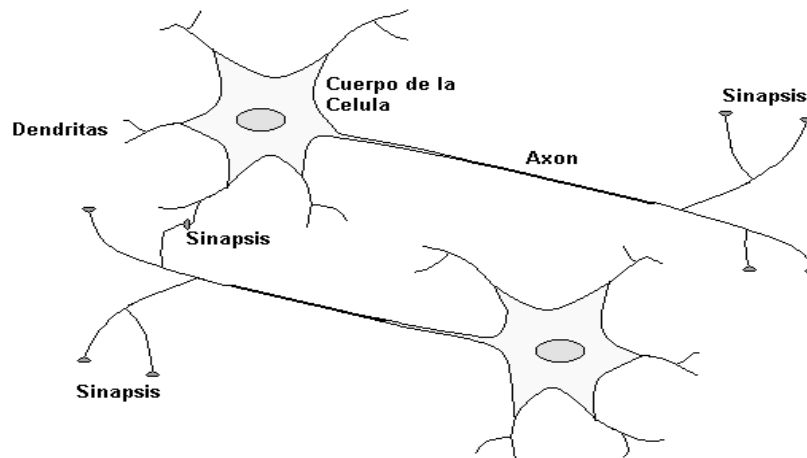


Figura 4.1. Fuente: Tomada de la tesis de Aguilar (2008).

En todas las neuronas se identifican tres partes principales:

- Cuerpo celular o soma.
- Dendritas.
- Axón.

La palabra sinapsis significa la unión de dos neuronas.²

Dendritas son las ramas fibrosas que emanan del cuerpo celular, Pérez (2003).

Axón son las fibras principales que emanan el cuerpo celular. Es el canal transmisor de los impulsos generados por las células. Se ramifica en su extremo final para conectar con otras neuronas, a través de las dendritas de éstas, que actúan como canales receptores de información, Pérez (2003).

Uno de los grandes misterios que han preocupado al hombre desde los tiempos ancestrales, es el de su propia naturaleza.

² La palabra sinapsis viene del griego *synapto* que significa unión o conexión estrecha.

Este misterio está asociado con el de la inteligencia artificial; dentro de ella está el ser inteligente (Isasi-Galván 2008, p. 1).

En la medida en que la ciencia y la tecnología avanzan en este mundo globalizado, uno de los retos más importantes a los que el ser humano en la actualidad se enfrenta es de la construcción de sistemas inteligentes; estos dispositivos físicos y lógicos que día con día se abren más campos o áreas donde se aplica la inteligencia artificial, tanto así que lleva un camino de desarrollo de forma exponencial, pues la sociedad demanda mejores resultados en todas las áreas.

4.2. Concepto de una red neuronal artificial (RNA)

Se han recogido varias definiciones sobre la noción de inteligencia artificial (IA).

1. La IA son algoritmos implementados en forma de programas informáticos o modelos electrónicos, basados en el funcionamiento del cerebro humano. Corchado *et al.* (2000, p. 3).
2. La IA son modelos computacionales que surgieron como un intento de conseguir formalizaciones

matemáticas acerca de la estructura del cerebro. Las RNA imitan la estructura hardware del sistema nervioso, centrándose en el funcionamiento del cerebro humano, y basándose en el aprendizaje a través de la experiencia. Flores y Fernández (2008, p. 16).

3. La IA es un sistema de computación compuesto de un gran número de elementos simples, elementos de proceso muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas. Hecht-Niese (1988), citado por Hilera y Martínez (1995, p. 9).
4. La IA es la ciencia e ingeniería de hacer máquinas inteligentes, especialmente programas de cómputo que realizan funciones muy diversas, imitando el comportamiento de un humano. John McCarthy (1956).

Dentro de la inteligencia artificial se pueden distinguir tres grandes áreas.

- Lógica difusa
- Redes neuronales artificiales
- Algoritmos genéticos

La lógica borrosa o difusa, se entiende como una de las ramas de la inteligencia artificial que le permite a una computadora analizar información del mundo real en una escala entre lo falso o lo verdadero. Los sistemas borrosos son una alternativa a las nociones de pertenencia y lógica que se iniciaron en la Grecia antigua. Ponce (2010, p. 2).

Para esta rama de la inteligencia artificial su mayor aplicabilidad es la robótica, donde se apoya del álgebra de Boole (booleana), haciendo uso de un sistema binario; un ejemplo claro sería la forma de operar precisamente un televisor.

Está encendido, verdadero (1).

Está apagado, falso (0).

Las redes neuronales artificiales (RNA) se pueden definir como un dispositivo diseñado a imitación de los sistemas nerviosos de los animales, consistentes en una interconexión de unidades, denominadas neuronas artificiales o elementos de proceso, cuyo funcionamiento se inspira en el de las neuronas biológicas. Pérez (2003, p. 13).

Un algoritmo genético (AG), es una búsqueda iterativa, inspirada en los principios de selección natural.

Los algoritmos genéticos no buscan modelar la evolución biológica, sino derivar estrategias de optimización. El concepto se basa en la generación de poblaciones de individuos mediante la reproducción de los padres. Ponce (2010, p. 12).

La inteligencia artificial (IA) va adquiriendo mayor importancia conforme los científicos en todo el mundo publican resultados de pruebas sobre nuevas aplicaciones; esto por la necesidad o demanda de hacer mejor las cosas en el menor tiempo posible y con mayor precisión.

Se trata de un campo multidisciplinario en donde han tenido que ver investigadores de diferente formación profesional; hasta el día de hoy al pinchar redes neuronales artificiales en una página web como www.google.com.mx, se despliegan 640,000 aplicaciones. Dentro de éstas, están integradas las aplicaciones a la valoración inmobiliaria.

4.3. Antecedentes históricos

Se puede considerar que los primeros pasos a la inteligencia artificial se dieron desde hace mucho tiempo por Aristóteles (384-322 a.C.) y Platón (427-347 a.C.), y se continuaron en la edad media por Descartes (1596-1650) y por algunos otros filósofos empiristas del siglo XVII.

La mayoría de los estudios sobre RNA citan los trabajos de Warren McCulloch y Walter Pitts como los iniciadores en el año de 1943, cuando realizaron su primer modelo matemático sobre redes neuronales artificiales.

Tal modelo se basa en la idea de que las neuronas operen mediante impulsos binarios. Este modelo introduce la idea de una función de paso por umbral, utilizada posteriormente por muchos modelos como las redes neuronales artificiales discretas. Isasi y Galván (2008, p. 18, citan a Kosko 1988).

Donald Hebb desarrolló posteriormente un procedimiento matemático de aprendizaje. Los estudios de Hebb sobre las neuronas y las condiciones clásicas de aprendizaje en su libro *Organization of Behavior*. Isasi y Galván (2008, p. 18, citan a Hebb 1949).

Los lógicos del siglo XX, incluidos Kurt Gödel, Stephen Kleene, Emil Post, Alonzo Church y Alan Turing formalizaron y clasificaron lo que puede (y lo que no puede) ser hecho por medio de los sistemas lógicos y computacionales. Posteriormente computólogos como Stephen Cook y Richard Karp identificaron las clases de cómputos, que siendo realizables mediante computadoras, son prohibitivos en términos de espacio o tiempo. Entre ellos, fue Alan Turing quien escribió el primer artículo moderno sobre la posibilidad de crear sistemas inteligentes artificiales (en el año de 1950). Aragón (2002, p. 19).

Hasta el año 1955, la mayoría de los investigadores consideran la IA como la primera generación, del año 1956 se ha considerado como la evolución de la segunda generación y ésta comienza con una conferencia pronunciada en Dartmouth, que organizó John McCarthy y Marvin Minsky en 1956, donde se fundaron las bases de la IA moderna.

En el año de 1957 Frank Rosenblatt comienza su desarrollo con el perceptrón y sus aplicaciones las enfocó al reconocedor de patrones.

En 1959, Bernard Widrow y Marcial Hoft de la Universidad de Stanford, desarrollaron el modelo Adaline (*Adaptative Linear Element*). Uno de los mayores investigadores de las redes neuronales desde los años sesenta hasta nuestros días es Stephen Grossberg (Universidad de Boston). A partir de su extenso conocimiento fisiológico, ha escrito numerosos libros y desarrollado modelos de redes neuronales. Estudió los mecanismos de la percepción y la memoria. Grossber realizó en 1967 una red, Avalancha, que consistía en elementos discretos con actividad que varía con el tiempo, que satisface ecuaciones diferenciales continuas para resolver actividades tales como reconocimiento continuo del habla y aprendizaje del movimiento de los brazos en un robot. Hiler y Martínez (1995, p. 4, citan a Grossberg 1982).

Numerosas críticas que surgieron entre los años 1969 y 1982, frenaron las investigaciones acerca de las redes neuronales artificiales. Marvin Minsky y Seymour Papert, del Instituto Tecnológico de Massachusetts (MIT), publicaron un libro denominado *El perceptrón* (el perceptrón multinivel poseía sólo una capa).

En 1982 coincidieron numerosos eventos que hicieron resurgir el interés por las redes neuronales. John Hopfield presentó su trabajo (Hopfield 1982). En el trabajo, describe con claridad y rigor matemático una red a la que ha dado su nombre, que es una variación del asociador lineal, pero además mostró que tales redes pueden trabajar y lo que pueden hacer. También en 1982 se celebró la *U.S.-Japan Joint Conference on Cooperative /Competitive Neuronal Networks* y Fujitsu comenzó el desarrollo en *computadores pensantes para la aplicación en robótica*. Hilera y Martínez (1995, p. 5).

Cada día aparecen nuevas aplicaciones y aportaciones científicas de las redes neuronales artificiales en todas las ramas del conocimiento.

De acuerdo con Núñez *et al.* (2009), en el campo de la valoración inmobiliaria entre los más destacados se puede citar a Bort en 1991, con un estudio realizado en Nueva Inglaterra, Estados Unidos, Tay y Ho en 1992 en Singapur, Do y Grudnitski en 1992 en California, Estados Unidos, Collins y Evans durante 1994 en el Reino Unido, Worzala, Lenk y Silva en 1995 en Colorado (Estados Unidos), Rossini en 1997 al sur de Australia, Haynes y Tan en el año 1998 en Gold Coast, Australia, Cechin en 2000,

Porto Alegre (Brasil), Karakozova en 2000, Helsinki (Finlandia) y Nguyen y Cripps en el año 2001, Tennessee (Estados Unidos).

Muchos de los estudios antes mencionados establecen la comparación entre los sistemas de IA y los métodos tradicionales de valoración, especialmente con la regresión múltiple, manifiestan ventajas de los modelos de IA contra los modelos de regresión en rangos medios de 5 a 10%.

Los modelos de IA ofrecen errores entre 5 y 10%, mientras los modelos de regresión múltiple entre 10 y 15%.

Otro caso es el de Zurada *et al.* (2006), en su investigación denominada “Enfoques no convencionales”. Utilizando una red neuronal perceptrón completamente conectado, con una capa oculta, la capa de salida contiene un solo nodo, lo que representa el precio estimado. La capa de entrada tiene 13 nodos que representan la entrada de 13 valores, probaron varias redes neuronales con diferentes números de neuronas en la capa oculta y una neurona en la capa de salida que representa el precio estimado. Resultó que la red con 3 neuronas en la capa oculta produce el menor error en las

unidades de prueba. La función de activación de tangente hiperbólica se utilizó para neuronas. La normalización de la desviación estándar se utilizó para las variables. Esta normalización resta la media y la divide por la desviación estándar, de manera que los valores que resultan tienen una media de cero y una desviación estándar de uno. Otro modelo que se usó fue el de regresión múltiple, para la reducción de variables o factores, mediante el uso del análisis de componentes principales de 13 variables, redujeron a tres variables y luego a seis variables, para estimar los precios, los resultados obtenidos en ambos modelos son muy parecidos; para llevar a cabo esta investigación se utilizaron observaciones hechas en la ciudad de Louisville, Kentucky, Estado Unidos, el conjunto de datos fue de 360 casas unifamiliares en los años de 1982 a 1992.

Otras aplicaciones, mediante un modelo de red neuronal artificial para la simulación de inundaciones usando GIS; a las inundaciones que suceden en la cuenca del río Johor, Malasia, Bakhtyari Kia *et al.* (2001), mediante el modelo de red neuronal perceptrón multicapa back propagación utilizaron una capa de entrada que contenía siete neuronas (una por la elevación, la pendiente topográfica, la acumulación de flujo, los datos

de geología, el uso de la tierra, el suelo y las precipitaciones), cada uno representando un factor causal que contribuye a la aparición de la inundación en la cuenca de captación. La capa de salida contiene una sola neurona y representa el curso del agua. Las capas ocultas de neuronas se utilizan para definir la compleja relación entre las variables de entrada y de salida. Una vez definidas las variables de entrada y de salida con salidas no lineales, como la ecuación sigmoide, los autores han encontrado coeficientes de determinación que van desde 0.93 hasta 0.99 para cada una de las variables aplicadas en el estudio.

Razi *et al.* (2005), “Índice de predicción de precios para inmuebles, al usar la red neuronal artificial versus análisis de regresión múltiple (MRA)”, en su investigación predicen el índice de precios con el uso de la red neuronal artificial; mediante el análisis de regresión múltiple versus la red neuronal artificial con múltiples análisis de regresión (MRA).

La variable dependiente fue el índice de precios de casa, como variables independientes utilizó la tasa de interés, el ingreso familiar, la tasa de empleo y el tamaño de la población. Los datos para la variable dependiente

cubrieron el período desde el año 2000 hasta 2011 y se obtuvieron de diversas fuentes. Los datos del índice de precios de la vivienda de Malasia fueron recogidos del Centro de Información de Bienes Nacionales (NAPIC), en la siguiente tabla se observan los resultados de los dos modelos que se utilizaron.

| Training Data (%) | Testing Data (%) | R ² | r ² | Mean squared error | Mean absolute error | Correlation coefficient r |
|-------------------|------------------|----------------|----------------|--------------------|---------------------|---------------------------|
| 90 | 10 | 0.9667 | 0.9747 | 3.986 | 1.413 | 0.9873 |
| 80 | 20 | 0.9932 | 0.9932 | 0.819 | 0.718 | 0.9966 |
| 70 | 30 | 0.9923 | 0.9924 | 0.922 | 0.766 | 0.9962 |
| 60 | 40 | 0.9923 | 0.9924 | 0.921 | 0.774 | 0.9962 |

Los resultados muestran la combinación óptima de datos compuesta por 80% de datos de entrenamiento y 20% de los datos de prueba con la más alta $r = 0,9966$ y la media más baja de error al cuadrado de 0,819. El R^2 también muestra que los datos de entrenamiento y de prueba seleccionados fueron capaces de reflejar 99,32% del valor real, se asume una vez más en este estudio la mejora de los resultados de las redes neuronales artificiales con respecto del análisis de regresión múltiple.

4.4. Elementos de una RNA

Un conjunto de conexiones o sinapsis caracterizadas cada una de ellas por un peso sináptico, de forma que la señal de entrada X_j , presente en la neurona k , se verá multiplicada por pesos sinápticos W_{jk} . Este peso puede variar en un rango que incluye tanto valores negativos (entrada inhibidora) como valores positivos (entrada excitadora). Corchado *et al.* (2000, p. 6).

La regla de propagación que determina la entrada efectiva o el nivel de excitación de la neurona k , denotada por s_k , a partir de todas las entradas individuales a la misma, donde se considera como entrada efectiva a la suma todas las señales de entrada X_j a la neurona k , ponderadas por su respectivos pesos sinápticos w_{jk} . Corchado *et al.* (2000, p. 6).

“ F_k , función de activación que determina la salida y_k de la neurona a partir de su nivel de excitación” Corchado *et al.* (2000, p. 6).

El modelo incluye la entrada externa adicional, denominada polarización o "bías" y denotada por b_k , su

función es aumentar o disminuir el poder de excitación de la neurona dependiente, si es un valor positivo o negativo. Véase la figura 4.2.

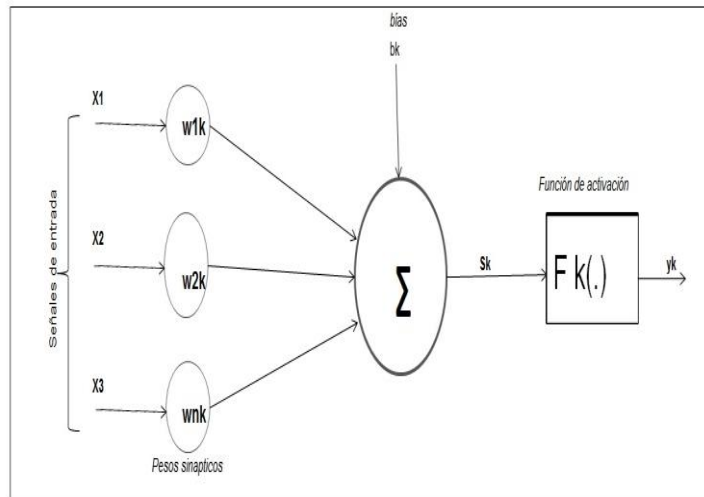


Figura 4.2 Fuente: Tomada de Corchado *et al.* (2000, p. 6), y adaptación propia

El modelo matemático se describe de la siguiente manera.

$$S_k = \sum_{i=1}^n (W_{jk})(X_j) + b_k \quad [4.1]$$

Donde \$W\$, representa el peso sináptico de la neurona, \$X\$ las señales de entrada, \$b\$ la polarización o bía: disminuir o aumentar el umbral de excitación de la neurona.

Si se considera la polarización b_k como peso sináptico de una entrada adicional en el cual su valor siempre será 1, se reescribe la ecuación a una forma homogénea.

Todo modelo neuronal debe disponer de una regla de propagación que combine la salida de cada neurona con su correspondiente ponderación por el patrón de conexión.

$$S_k = \sum_{i=k}^n (W_{jk})(X_j) \quad [4.2]$$

Siendo en este caso $W_k = [b_k, w_{1k}, \dots, w_{jk}, \dots, w_{nk}]^T$ el vector columna de pesos y $x = [1, x_1, \dots, x_N]^T$ el vector columna de entrada, esto a su vez se puede escribir de forma vectorial tal que: $y_k = F_k[w_k^T, x]$. Corchado *et al.* (2000, p. 7).

El algoritmo anterior puede también representarse como una notación matricial, en este caso se puede considerar una W de dimensiones $N \times N$ que representará todos los pesos de la red. Véase la figura 4.3.

$$W = \begin{pmatrix} W_{11} & W_{12} & W_{13} & \dots & W_{1n} \\ W_{21} & W_{22} & W_{23} & \dots & W_{2n} \\ W_{31} & W_{32} & W_{33} & \dots & W_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ W_{m1} & W_{m2} & W_{m3} & \dots & W_{mn} \end{pmatrix}$$

Figura 4.3. Fuente propia. Representa una matriz con todos los pesos de los patrones de entrada a la red.

Dependiendo del tipo de salida, las neuronas pueden tomar valores de 0 o 1 si son discretas. Otro rango de valores que va desde [0, +1] o [-1, +1], si son continuas.

4.5. Principales funciones de activación

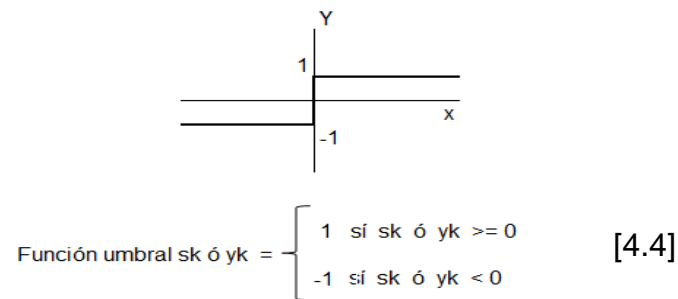
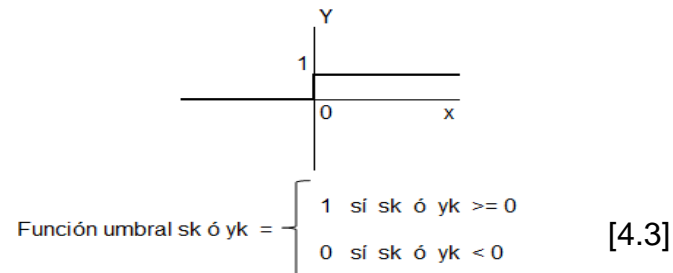
Existen cuatro funciones principales de transferencia típicas que determinan diferentes tipos de neuronas: La función escalón, la función lineal y mixta, la sigmoideal y la función gaussiana.

4.5.1. Función de activación escalón

Función escalón o umbral se utiliza cuando la salida de red es de tipo binario, dos posibles valores (0, 1).

La neurona se activará cuando el valor neto de entrada sea igual o mayor al umbral y no se activará en caso contrario. Véanse las ecuaciones 4.3 y 4.4, así como la figura 4.4.

Figura 4.4. Representación de la función escalón



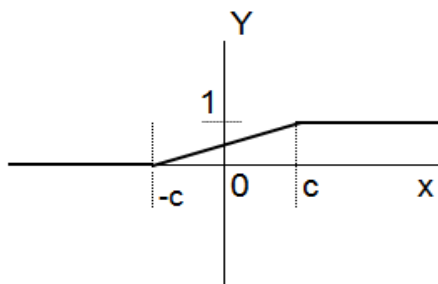
Fuente: Propia, readaptada de Hilera y Martínez (1995).

La figura 4.4 representa dos funciones escalón en ambos casos se toma en cuenta el umbral cero, en caso de que no lo fuera, el escalón quedaría desplazado.

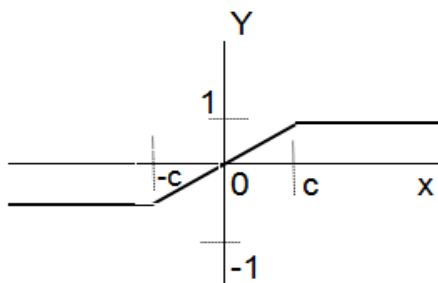
4.5.2. Función de activación lineal y mixta

Esta expresión responde a la expresión sk ó $yk(x) = x$, las neuronas con función mixta, si la suma de señales de entrada es menor que un límite inferior, la activación se define como 0 (0-1). Si dicha suma es mayor o igual que el límite superior, entonces la activación es 1. Si la suma de entrada está comprendida entre ambos límites, superior e inferior, entonces la activación se define como una función lineal de la suma de las señales de entrada. Véanse las ecuaciones 4.5 y 4.6, así como la figura 4.5.

Figura 4.5. Representación de la función de activación lineal y mixta.



$$sk \text{ ó } yk(x) = \begin{cases} 0 & \text{sí } x < -c \\ 1 & \text{sí } x > c \\ x/(2c) + 1/2 & \text{en otro caso} \end{cases} \quad [4.5]$$



$$sk \text{ ó } yk(x) = \begin{cases} -1 & \text{sí } x < -c \\ 1 & \text{sí } x > c \\ a.x & \text{en otro caso} \end{cases} \quad [4.6]$$

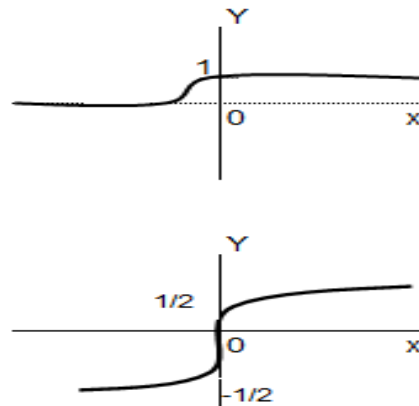
Fuente: Propia, readaptada de Hilera y Martínez (1995).

La figura 4.5 representa dos funciones: una lineal y otra mixta, en ambos casos se toma en cuenta el umbral de acuerdo con las ecuaciones [4.5 y 4.6].

4.5.3. Función de activación continúa sigmoideal

Es cualquier función definida en el intervalo de posibles valores de entrada con un incremento monotónico y que tenga ambos límites superiores e inferiores (sigmoideal o arco tangente). El valor obtenido de salida debe estar comprendido en la zona alta o baja del sigmoide, su derivada es siempre positiva y cercana a cero para valores grandes positivos o negativos. Véanse la ecuación 4.7 y figura 4.6.

Figura 4.6. Representación de la función de activación sigmoideal



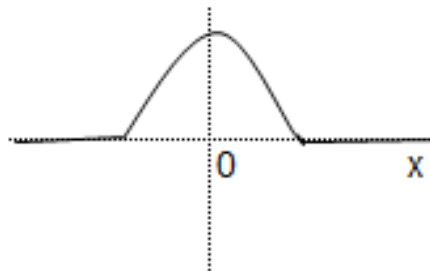
$$sk \text{ ó } yk (x) = 1 / (1 + e^{-x}) \quad [4.7]$$

Fuente: Propia, readaptada de Hilera y Martínez (1995).

4.5.4. Función de activación gaussiana

Los centros y anchura de estas funciones pueden ser adaptados. Mapeos que suelen requerir dos niveles ocultos (neuronas en la red que se encuentran entre las de entrada y las de salida), utilizando neuronas con dos funciones de transferencia sigmoidales; algunas veces se pueden realizar con un solo nivel en redes con neuronas de función gaussiana. Hilera y Martínez (1995). Véanse la ecuación 4.8 y la figura 4.7.

Figura 4.7. Representación de la función de activación gaussiana



$$s_k \text{ ó } y_k(x) = A \cdot e^{-Bx^2} \quad [4.8]$$

Fuente: Propia, readaptada de Hilera y Martínez (1995).

Normalmente estas funciones, antes detalladas, suelen considerarse determinísticas y son monótonas crecientes y continuas como se ha observado en las neuronas biológicas. Existe otro tipo de funciones, pero las más usadas son las que se han descrito anteriormente.

Una RNA consta de varios elementos de proceso conectados de alguna forma, generalmente organizados en grupos denominados **capas**. Existen dos capas típicas en toda red, que contienen conexiones con el exterior:

- La **capa de entrada o patrón**, que son los datos que se presentan a la red.
- La **capa de salida**, que muestra la respuesta de la red a una entrada.

En la figura 4.8 se muestra un ejemplo.

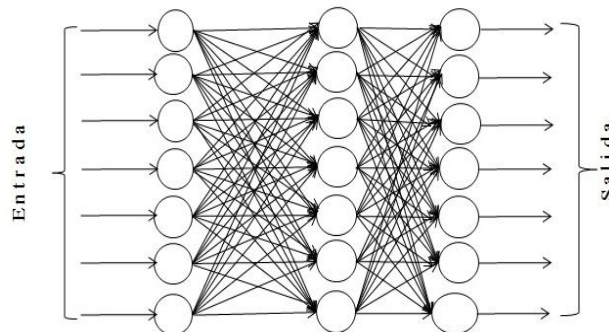


Figura 4.8. Fuente propia. Muestra la entrada, la red y las salidas de patrones o resultados.

Como se observa en la figura 4.8, cada una de las entradas sirve para distribuir las entradas de la red, dentro de la misma figura existe una capa intermedia denominada también capa oculta, ésta no tiene conexión directa con el entorno.

Pueden existir varios tipos de conexiones entre las neuronas:

- **Conexiones intracapa o laterales:** se trata de neuronas que producen conexiones entre las neuronas de una misma capa.
- **Conexiones intercapa:** se producen entre neuronas de diferentes capas.
- **Conexiones realimentadas:** tienen sentido contrario al de entrada-salida.
- **Conexiones autorrecurrentes:** se realimentan de una neurona consigo misma.

Las conexiones entre las neuronas pueden ser excitatorias (en el caso de pesos positivos) o inhibitorias (en el caso de pesos negativos). En general no se define una conexión explícitamente como de un tipo u otro, sino que por medio del aprendizaje se obtiene un valor para el peso que incluye signo y magnitud. Pérez (2003, p. 24).

4.6. Operación de la RNA

Se clasifican dos formas de operación en un sistema neuronal: uno es el modo del recuerdo o ejecución, y otro el modo de aprendizaje o entrenamiento.

4.6.1. Fase de aprendizaje

Es el proceso por el cual se produce un ajuste en los parámetros libres de la red con la finalidad de estimular y producir una salida de la red.

Si la salida no es la deseada, habrá que realizar otras iteraciones; estas iteraciones normalmente se efectúan cambiando las ponderaciones o lo que es lo mismo al ir variando los pesos sinápticos de los patrones de entrada. Por ejemplo si los pesos de esta ecuación $W_{jk} = [b_k, w_{1k}, \dots, w_{jk}, \dots, w_{nk}]^T$.

Fueran $W_{jk} = [1, .80, \dots, .80, \dots, .60, \dots, w_{nk}]^T$, y mediante el aprendizaje o iteraciones los resultados no son los deseados, habría que hacer cambios de los mismos hasta que arrojen los valores deseados. Dentro del tipo de aprendizaje existen dos básicamente: el tipo supervisado y el no supervisado.

Aprendizaje supervisado.

Existe un instructor o profesor que dispone de un conjunto de elementos, patrones o valores de entrenamiento, formados por la entrada y la salida a la red. El profesor realizará iteraciones con las diferentes entradas y se generará una salida que por lo general en la primera iteración no coincidirá con la salida esperada. Se calcula el error de la salida y en función de este error, se van actualizando los pesos, hasta que el error sea menor y obtener el valor esperado.

Aprendizaje no supervisado.

Como su nombre lo indica, se trata de aprendizaje no supervisado en donde no se cuenta con un instructor, se dispone de un conjunto de elementos, donde no existe una salida deseada de la red. Se representará las entradas que harán las iteraciones con las diferentes entradas y se generará una salida. Estas salidas se agruparán por patrones para verificar su similitud (*clustering*), de esta manera se obtendrán los resultados.

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 global de la red, lo que permitirá actualizar los pesos. Como en este caso no se puede determinar la fracción de error correspondiente a cada salida, es necesario el uso de algoritmos de mayor complejidad, además de un conjunto de entrenamiento de mayor tamaño.

En este medio camino entre el aprendizaje supervisado y el no supervisado, se utilizará una medida del error cometido, que en este caso es un valor global que indica el error cometido por la red, no por cada elemento individual. Al igual que en el aprendizaje no supervisado, no se indica explícitamente la salida deseada.

Aprendizaje híbrido.

En una misma red se utilizan el aprendizaje supervisado y el no supervisado, normalmente en distintas capas de la red.

Aprendizaje hebbiano.

Este tipo de aprendizaje fue propuesto por Donal O. Hebb en 1949, según Hebb dice que *“Cuando un axón de una celda A esta lo suficientemente cerca como para conseguir excitar una celda B y repetida o persistentemente toma parte en su activación. Algún proceso de crecimiento o cambio metabólico tiene lugar en una o ambas celdas, de tal manera que la eficiencia de A. Cuando la celda a activar es B, aumenta. Por celda”*. Hebb entiende al conjunto de neuronas fuertemente interconectadas a través de una estructura compleja. La eficiencia podría identificarse con la intensidad y magnitud de las conexiones; es decir con el peso.

El aprendizaje hebbiano consiste en el ajuste de los pesos de las conexiones de acuerdo con la correlación de activación existente entre los valores (salidas) de las neuronas conectadas.

$$\Delta W_{ij} = s_k_i s_k_j \quad [4.9]$$

Donde ΔW_{ij} es la variación en los pesos en las conexiones de las neuronas i y j ; s_k_i es la salida de la neurona i ; s_k_j es la salida de la neurona j .

Esto indica que si las dos neuronas están activas positivamente, se produce un reforzamiento de la conexión. Por el contrario cuando una es activa y la otra es pasiva (negativa), se produce un debilitamiento de la conexión. Se trata de una regla de activación no supervisada. Pues los pesos se modifican en función de los estados (salidas) de las neuronas obtenidos después de la presentación de cierto estímulo, es decir información de entrada que afecta a la red sin considerar si se deseaba obtener o no esos estados de activación.

Este tipo de aprendizaje lo utilizó Hopfield en la red conocida como Red Hopfield, introducida en el año 1982.

Los algoritmos de aprendizaje generalmente se basan en métodos numéricos iterativos que tratan de minimizar una función de coste, lo que en ocasiones puede dar problemas en la convergencia del algoritmo. Estos aspectos no pueden tratarse de un modo general, sino que se deben estudiar para cada algoritmo concreto. La convergencia es una manera de comprobar si una determinada arquitectura, junto a su regla de aprendizaje, es capaz de resolver un problema, pues el grado de error que se mide durante el proceso de aprendizaje describe la precisión del ajuste de la red.

“La red tardará un tiempo variable en aprender, pudiendo en ocasiones no conseguirlo. Esto puede deberse a varias causas” (Pérez 2003, p. 26).

- El modelo que se ha seleccionado no es el adecuado.
- Las características del modelo no son las adecuadas (número de capas, número de neuronas y tipo de función de activación).
- Que los patrones o entradas utilizadas para entrenar la red no expliquen bien el problema de interés, lo anterior puede suceder porque no son los patrones más importantes del problema a resolver.

4.6.2. Evaluación del aprendizaje de la red

Uno de los aspectos importantes en la construcción de las RNA, es probar los resultados que se obtienen de las salidas, que arroje valores fiables; esto se obtiene al ir probando los errores de las diferentes iteraciones.

La forma de probar estos resultados es a través de la aplicación de las técnicas estadísticas como son el coeficiente de correlación lineal, si es el caso, en ocasiones puede que se esté aplicando una salida no lineal, entonces se está en un escenario en donde se tiene que aplicar la técnica de una correlación no lineal a la salida de la red.

Al aplicar estas técnicas lo que se obtiene en realidad es la relación que existe entre las variables independientes o patrones de entrada y el resultado que es la variable dependiente, lo cual significa qué grado de asociación existe entre ellas.

4.6.3. Según el tipo de respuesta

4.6.3.1.- Las redes autoasociativas se entrenan para que asocien un patrón consigo mismas, de forma tal que, al presentar dicho patrón con un ruido superpuesto o de forma parcial, sean capaces de identificar el patrón.

4.6.3.2.- Las redes heteroasociativas se entrenan para que ante la presencia de un patrón A respondan con otro diferente B.

“La auto-asociación implica aprendizaje no supervisado, mientras que la heteroasociación supone aprendizaje supervisado” (Pérez 2003, p. 26).

4.7. El perceptrón simple

4.7.1. Qué es un perceptrón

Una forma de comenzar a entender qué es un perceptrón, es asociarlo con un sensor, bien sea de temperatura, humedad, nivel de líquido, grado de acidez, coloración, densidad, etc. Es en esencia, un dispositivo que, ante la presencia de uno o varios fenómenos de entrada, permite representarlo mediante una señal de salida fácilmente reconocible. Ahora bien, si a este simple dispositivo se le asignan varios canales de entrada (dos o más), se le habrá agregado una notable mejoría ya que podrá discriminar o diferenciar entre distintos fenómenos de entrada y emitir una salida que representará el criterio diferenciador o resultado de la interacción entre las entradas.

Este modelo neuronal fue introducido por Rosenblatt a finales de los años 1950. El perceptrón

simple es un modelo unidireccional, compuesto por dos capas de neuronas, una sensorial o de entrada y otra de salida. La operación de una red de este tipo, se diseñó para trabajar con patrones de entrada y salida de tipo binario ([0,1] o [-1,1]) (Martín del Brío y Sanz 2007, p. 47).

Las neuronas de entrada no realizan ningún cómputo, únicamente envían la información a la siguiente capa. La neurona de salida tiene una función de activación de tipo escalón e incluye un umbral o entrada externa adicional, denominada polarización o "bias" y está denotada por b_k , su función es aumentar o disminuir el poder de excitación de la neurona independiente según si es un valor positivo o negativo, de tal manera que calcula su entrada neta como una suma ponderada de las entradas por sus pesos, a la que se le resta o se le suma el umbral.

Véase la figura 4.2. Se muestra la red de un perceptrón simple unidireccional.

4.8. El perceptrón multicapa

En 1969 Minsky y Papert publicaron su libro *Perceptrón. Una introducción a la geometría computacional*, se trata de una generalización del

perceptrón simple, comentado líneas atrás, surgió como consecuencia de las limitaciones de su arquitectura en lo referente del problema de separabilidad no lineal.

El perceptrón multicapa, inicialmente desarrollado por P. Werbos (1974), permite resolver este problema. Posee una estructura con al menos una capa oculta y su algoritmo de entrenamiento es del tipo corrección de error. Se basa en el cálculo del gradiente distribuido en los diferentes componentes de la red. Sánchez y Alanís (2006, p. 63) citan a Werbos (1974).

El perceptrón multicapa crea modelos a partir de multiplicadores, sumadores y funciones. El entrenamiento consiste en la presentación de pares de vectores en las capas de entrada y salida (vectores de entrada y salida deseada). La red crea un modelo al ajustar sus pesos en función de los vectores de entrenamiento, de forma que a medida que pasan estos patrones hacia cada vector de entrada, la red producirá un valor de salida más similar al vector de salida esperado. Estas redes también se llaman de retropropagación (*back propagation*), este nombre se les ha asignado por el tipo de aprendizaje que utilizan. Corchado *et al.* (2000, p. 87 y 88).

4.9. Topología del perceptrón multicapa

Las redes de este tipo tienen una capa de entrada, una capa intermedia y otra de salida. La figura 4.9 muestra el diagrama y su topología³. La información se propaga de capa en capa (de derecha a izquierda), por medio de las neuronas de cada capa.

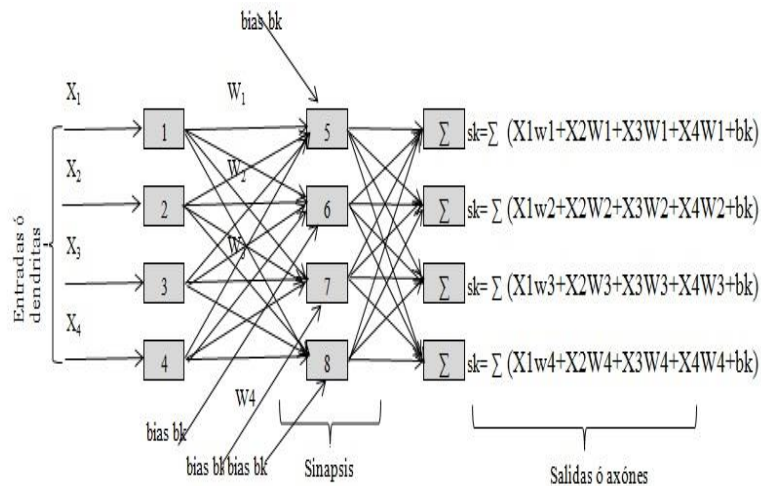


Figura 4.9. Fuente: Propia, muestra la topología típica.

En la figura 4.9 se muestra una red de un perceptrón

³ Topología: Propiedades o características que posee un cuerpo geométrico.

multicapa, se trata de un modelo unidireccional compuesto por tres capas, en la primer capa recibe cuatro entradas, también denominadas biológicamente como dendritas, aquí solamente se encarga de recibir información no realiza ningún otro proceso, en la segunda capa oculta denominada también sinapsis se realiza la ponderación de pesos y la tercer capa, conocida también como axones, es la que conduce la salida de la multiplicación de la ponderación de los pesos por las entradas, incluidas las bias, la polarización o el umbral para cada entrada, dando un resultado denominado sk. Su ecuación matemática se ha denotado anteriormente como [4.1].

4.10. Parámetros que determinan el aprendizaje de la red

Cuando se trabaja con redes neuronales artificiales siempre es necesario tomar decisiones, por ejemplo.

- Cómo representar la información en las entradas de entrenamiento.
- Cuántas capas debe tener la red neuronal artificial y cuántas neuronas cada capa.
- De qué manera se obtendrán los pesos de los patrones de entrada.
- Qué valores deben tomar las bias o la polarización.
- Cómo debe realizarse el entrenamiento.

4.11. Método de aprendizaje del perceptrón multicapa

Para explicar cuál es la manera en que un perceptrón multicapa o red multicapa aprende, se utiliza una red con tres capas (véase figura 4.10): una de entrada, una intermedia y una de salida.

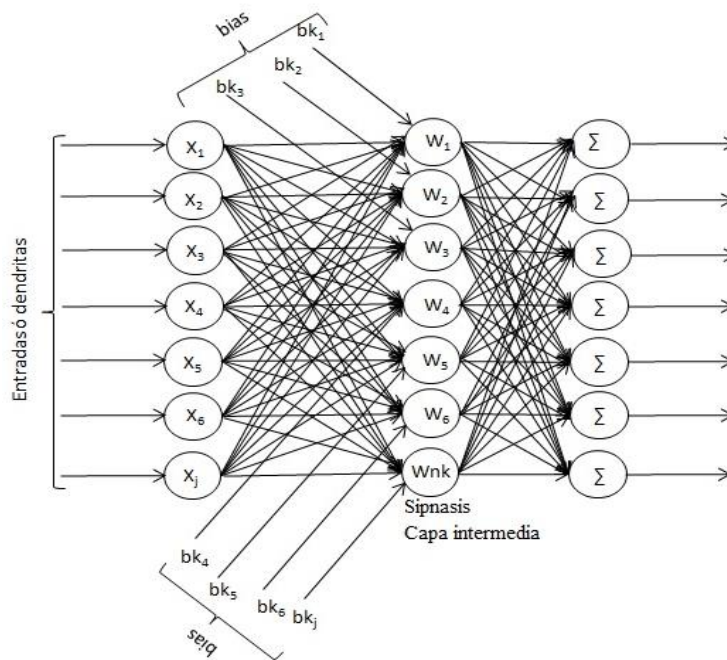


Figura 4.10. Fuente: Propia, muestra la topología típica.

En la figura 4.10 se muestra una red de un perceptrón multicapa, se trata de un modelo unidireccional compuesto por tres capas, en el ejemplo anterior en la

primer capa recibe siete entradas, pero pueden ser más, o también denominadas biológicamente dendritas, aquí solamente se encarga de recibir no realiza ningún proceso, en la segunda capa oculta denominada también sinapsis se realiza la ponderación de pesos y en la tercer capa denominada también axones, es por donde se conduce la salida de la multiplicación de la ponderación de los pesos por las entradas, incluyendo las bias, polarización o umbral para cada entrada, dando un resultado denominado s_k .

$$S_k = \sum_{i=k}^n (W_{jk})(X_j) + bk \quad [4.10]$$

Donde W representa el peso sináptico de la neurona, X , las señales de entrada y b la polarización o la bia contribuyendo a disminuir o aumentar el umbral de excitación de la neurona.

Si se considera la polarización b_k como peso sináptico de una entrada adicional en el cual su valor siempre será 1, se reescribe la ecuación a una forma homogénea.

Todo modelo neuronal debe disponer de una regla de propagación que combine la salida de cada neurona con su correspondiente ponderación por el patrón de conexión.

$$S_k = \sum_{i=k}^n (W_{jk})(X_j) \quad [4.11]$$

Siendo en este caso $W_k = [b_k, w_{1k}, \dots, w_{jk}, \dots, w_{nk}]^T$ el vector columna de pesos y $x = [1, x_1, \dots, x_N]^T$ el vector columna de entrada, esto a su vez se puede escribir de forma vectorial tal que: $y_k = F_k[w_k^T \cdot x]$. Corchado *et al.* (2000, p. 7).

El algoritmo anterior puede también representarse como una notación matricial, en este caso se puede considerar una W de dimensiones $N \times N$ que representará todos los pesos de la red.

Algunos autores no utilizan las bias o la polarización, de tal manera que la red queda representada de la siguiente manera (véase figura 4.10).

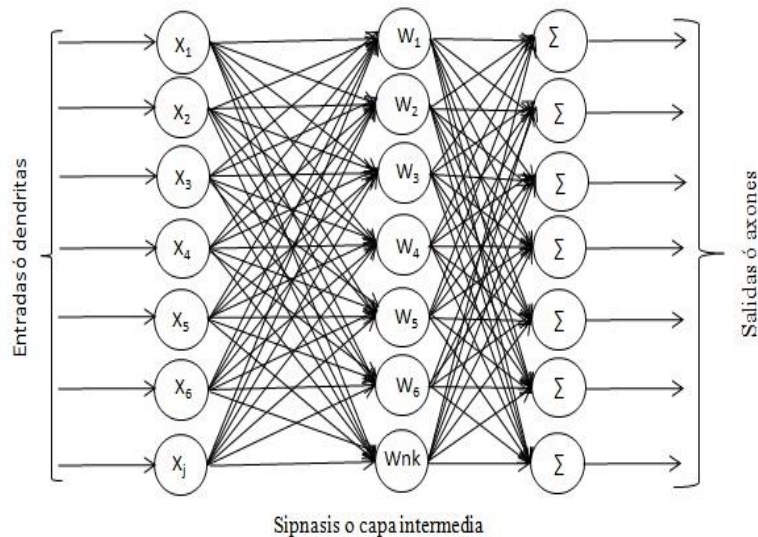


Figura 4.11. Fuente: Propia, muestra la topología típica.

Como se observa en la figura 4.11, se muestra una red de un perceptrón multicapa o red multicapa; se trata de un modelo unidireccional compuesto por tres capas, en la primer capa recibe siete entradas, aquí solamente se encarga de recibir no realiza ningún proceso, en la segunda capa oculta denominada también sinapsis se realiza la ponderación de pesos y en la tercer capa la salida de la multiplicación de la ponderación de los pesos por las entradas, aquí ya no se incluyen las bias, la polarización o el umbral para cada entrada, dando un resultado denominado s_k , de acuerdo con la ecuación [4.11].

En este caso el vector de pesos queda representado de la siguiente manera, $W_k = [w_{1k}, \dots, w_{jk}, \dots, w_{nk}]^T$ el vector columna de pesos y $x = [1, x_1, \dots, x_N]^T$ el vector columna de entrada, al escribir de forma vectorial que: $y_k = F_k[w_k^T x]$.

La notación anterior puede también representarse como una notación matricial, en este caso se puede considerar una W de dimensiones $n \times n$ que representará todos los pesos de la red.

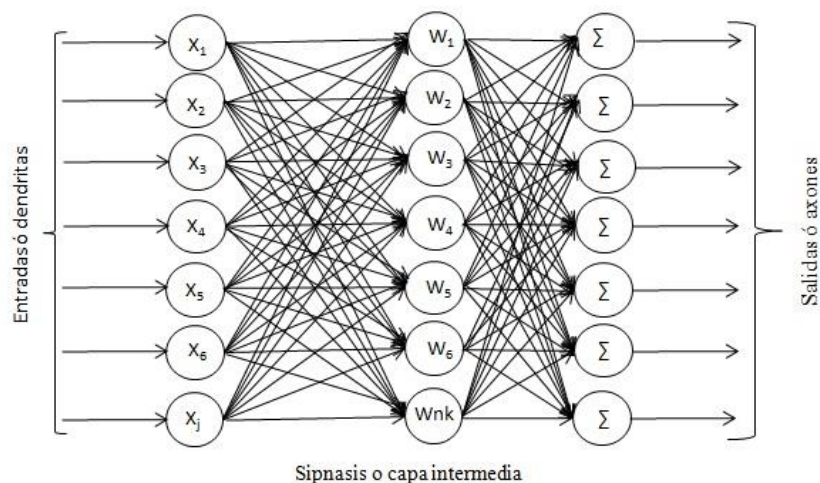
La manera general de representar una matriz se ilustra en la figura 4.12.

$$A_{ij} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

Figura 4.12. Elaboración propia.

La figura anterior representa una matriz como un arreglo rectangular de números donde cada elemento de la matriz está identificado por la posición de su renglón y su columna tal que la notación a_{mn} representa el valor que se encuentra en el renglón m y la columna n de la matriz.

Figura 4.13. Estructura de la red multicapa *back propagation*



Fuente: Elaboración propia.

La figura 4.13 representa la estructura de una red multicapa *back propagation*, con tres capas: la primera de entrada de neuronas a la red, la segunda capa oculta donde se procesa la activación de las neuronas o proceso matemático y la tercer capa de salida donde se muestran los resultados.

4.12. La red con conexiones hacia adelante (*back propagation*)

La red neuronal con conexiones hacia adelante (*back propagation*), propuesta en el año de 1986 por Rumelhart, Hinton y Williams (Rumelhart, 1996), basándose en trabajos de Werbos (1974) y Parker (1982), formalizaron un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada a la red y las clases correspondientes, utilizaron más niveles de neuronas de los que usó Rosenblatt para desarrollar el perceptrón. Este método, conocido en general como *back propagation* (propagación del error hacia atrás), se basa en la generalización de la regla delta y, a pesar de sus propias limitaciones, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales.

Este algoritmo de propagación hacia atrás, o retropropagación, es un regla de aprendizaje que se puede aplicar en modelos de redes con más de dos capas de neuronas. Una característica importante de este algoritmo es la representación interna del conocimiento

que es capaz de organizar la capa intermedia de las células para conseguir cualquier correspondencia entre la entrada y la salida de la red.

De forma simplificada, el funcionamiento de una red *back propagation* consiste en el aprendizaje de un conjunto entradas-salidas, al emplear un ciclo de propagación-adaptación de dos fases. En una primera fase se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, que a su vez se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado obtenido en las neuronas de salida con la salida que se desea obtener y se calcula un valor del error para cada neurona de salida. Después, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Con base en el valor del error recibido, se reajustan los pesos de conexión de cada neurona de manera que la siguiente ocasión que se presente el mismo patrón, la salida estará

más cercana a la deseada; es decir, que el error disminuye.

La importancia de la red *back propagation* consiste en su capacidad de auto-adaptar los pesos de las neuronas de las capas intermedias que incrementan el potencial de aprendizaje, acerca de la relación que existe entre un conjunto de patrones de entradas y sus salidas. Esta capacidad generalizada se exige a los sistemas de aprendizaje, para dar salidas satisfactorias a entradas en el sistema de entrenamiento.

4.13. La regla delta generalizada

La regla delta fue propuesta por Widrow en 1960, y su uso se ha extendido a redes con capas intermedias (regla delta generalizada) con conexiones hacia adelante (*feedforward*) y cuyas neuronas tienen funciones de activación continuas (lineales o sigmoidales), lo que da origen al algoritmo de retropropagación (*back propagation*). Utiliza una función o superficie de error asociada con la red, busca el momento estable de mínima energía o de mínimo error a través del camino descendente de la superficie del error, luego realimenta el error del sistema para realizar la modificación de los pesos con un valor

proporcional al gradiente decreciente de dicha función de error. Hilera y Martínez (1995).

El proceso de aprendizaje para ajustar los pesos es exactamente el mismo que el de la regla delta utilizado en el perceptrón y Adaline, es decir al actualizar los pesos en forma proporcional a la delta, o diferencia entre la salida deseada y la obtenida realmente, mediante la siguiente ecuación:

$$\Delta w_{jk}(t + 1) = \alpha \delta_{pj} s_{k_{pj}} \quad [4.12]$$

Donde:

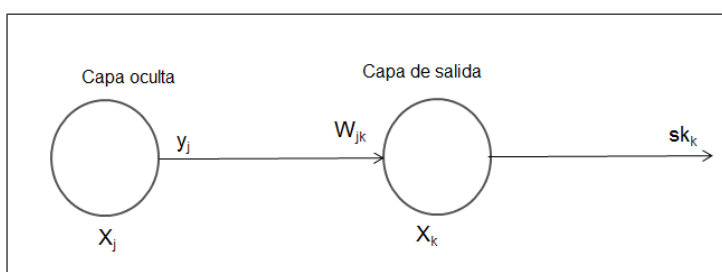
$\Delta W_{jk}(t+1)$: es la variación entre la neurona j y la k, una vez que reciben impulso que las active.

α : factor de aprendizaje que tendrá un valor en el intervalo entre 0 y 1, determina la velocidad de aprendizaje de la neurona (valor constante).

$S_{k_{pj}}$: valor de la salida de la neurona k, bajo el aprendizaje j.

δ_{pj} : valor delta o diferencia entre la salida deseada y la obtenida realmente (valor concreto).

Figura 4.14. Conexión entre una neurona de una capa oculta y una neurona de salida



Fuente: Elaboración propia.

La figura 4.14 representa lo siguiente:

Dada una neurona X_j y la salida que produce, el cambio de peso de la conexión que une la salida de dicha neurona con la unidad X_k (W_{jk}), para un patrón de aprendizaje j determinado, es el resultado de la ecuación [4.12].

La diferencia entre la regla delta y delta generalizada estriba en el valor concreto de δ_{ik} que se obtiene. La regla delta se define como la diferencia cuadrática del error entre la salida de la red con los pesos actuales y las salidas deseadas.

En cambio la regla delta generalizada actualiza los pesos de forma proporcional, la diferencia entre la salida deseada de la neurona k y el valor realmente obtenido por la neurona k ($\delta_k - s_{kj}$). Sin embargo para la regla delta generalizada, la delta se define como:

$$\delta_{pj} = (d_{pj} - sk_{pj})X_j \quad [4.13]$$

Donde:

p: representa un patrón de entrenamiento.

d_{pj} : es el valor de salida deseado de la neurona k para el patrón j.

sk_{pj} : es el valor de salida de la neurona k para el patrón j.

X_j : función de entrada neta de información que recibe la neurona j.

En la regla delta, cuándo una neurona no sea de salida de la red, es decir, sea una capa intermedia, el error que se genere será función del error cometido por las neuronas que reciben como entrada la salida de dicha neurona. Esto es lo que se denomina anteriormente la propagación del error hacia atrás. En este caso la delta generalizada se expresará como sigue:

$$\delta_{pj} = \left(\sum_k \delta_{pk} - W_{jk} \right) X_j \quad [4.14]$$

Donde:

k: cubre todas aquellas neuronas a las que está conectada la salida.

W_{jk} : representa todos los pesos de conexión de todas las neuronas

X_j : representa las neuronas de entrada a la red o patrones de entrada.

δ_{pj} : valor concreto, diferencia entre valor de salida deseada y la obtenida realmente.

Se ha explicado en la ecuación [4.12], el factor α , que es el parametro de aprendizaje, determina la velocidad de aprendizaje de la neurona. A mayor tasa de aprendizaje mayores son los cambios de los pesos a cada iteración y por tanto más rápido será el aprendizaje. Sin embargo, para cambios grandes los resultados podrian empeorar. Rumelhart *et al.* (1986), citado por Hilera y Martínez (1995), proponen se agregue en el incremento de los pesos un término (momento), β quedando como sigue:

$$W_{pj}(t+1) = W_{kj}(t) + \alpha \delta_{pj} s k_{pj} + \beta (W_{jk}(t) - W_{jk}(t+1)) =$$

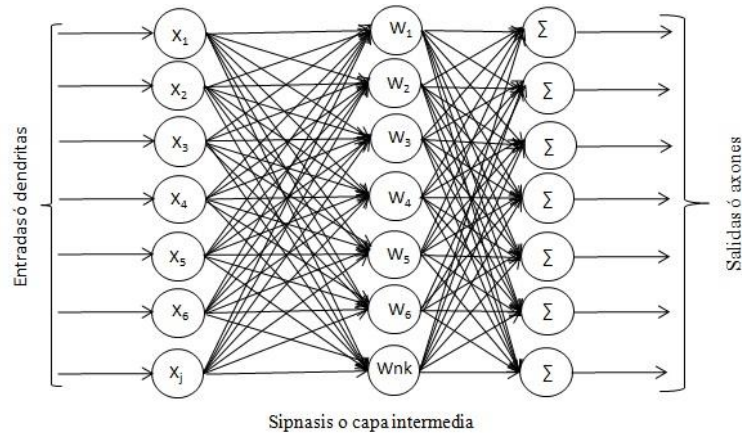
$$\Delta W_{jk}(t+1) = \alpha \delta_{pj} s k_{pj} + \Delta W_{jk}(t) \quad [4.15]$$

Donde β es una constante (momento), que determina el efecto en (t+1) del cambio de los pesos en el instante t.

Dicho momento lo que consigue es una convergencia de la red en un menor número de iteraciones, si en t el incremento de pesos es positivo y en t+1, también lo es, entonces el descenso por la superficie de error en t+1 es mayor, si en t el incremento es positivo y en t+1 es negativo, el paso que se da en t+1 es más pequeño, lo cual es adecuado, y significa que se ha pasado por un mínimo y los pasos deben ser menores para poder alcanzar el aprendizaje.

4.14. Estructura y aprendizaje de la red *back propagation*

Figura 4.15. Estructura de la red *back propagation*



Fuente: Elaboración propia.

En la figura 4.15 se aprecia la estructura de una red *back propagation*. En este caso presenta una capa de entrada, una capa oculta y una capa de salida, la cual cumple con el hecho de que sólo se requiere como mínimo una capa oculta.

La aplicación del algoritmo *back propagation* o regla delta en el caso del perceptrón requiere que las neuronas tengan una activación continua y diferenciable, por lo que generalmente se utiliza una función de tipo sigmoidea, aunque también puede ser una de tipo lineal.

Enseguida se detallan en forma resumida los pasos y las fórmulas para aplicar el algoritmo de entrenamiento.

Paso 1

Iniciar los pesos en la red con valores pequeños aleatorios.

Paso 2

Presentar un patrón de entrada $X: X_1, X_2, \dots, X_j$ y especificar la salida deseada que debe generar la red d_1, d_2, \dots, d_j (si la red se utiliza como un clasificador, todas las salidas deseadas tendrán el valor de cero, salvo una que será de la clase a la que pertenece el patrón de entrada).

Paso 3

Calcular la salida actual de la red, primero se presentan las entradas a la red y luego se calculan las salidas que presenta cada capa, hasta llegar a la capa de salida, ésta será la salida de la red sk_1, sk_2, \dots, sk_j . Las etapas son las siguientes.

Se calculan las entradas netas para las neuronas ocultas que provienen de las neuronas de entrada.

Para una neurona j oculta:

$$sk_{pj}^h = \sum_{j=1}^n W_{jk} X_j + bk^h \quad [4.16]$$

En donde el índice h se refiere a las magnitudes de la capa oculta, el subíndice j , al j -ésimo vector de entrenamiento, k , a la k -ésima neurona oculta. El término bk puede ser opcional, actúa como una entrada más en la red.

Se calculan las salidas de las neuronas ocultas:

$$sk_{pj} = sk_{pj}^h \quad [4.17]$$

Se realizan los mismos cálculos para obtener las neuronas de salida de la capa oculta.

$$sk_{jk}^0 = \sum_{j=1}^n W_{jk} X_j^0 + bk^0 \quad [4.18]$$

$$sk_{jk} = sk_{jk}^0 \quad [4.19]$$

Paso 4

Calcular los términos de error de todas las neuronas.

Si la neurona k es una neurona de la capa de salida, el valor de la delta es:

$$\delta_{jk} = (d_{jk} - s_{k_{jk}})X_{jk} \quad [4.20]$$

Paso 5

Actualización de los pesos.

Para ello se utiliza el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente.

$$\begin{aligned} W_{jk}(t+1) &= W_{jk}(t) + \Delta W_{jk}(t+1) \\ \Delta W_{jk}(t+1) &= \alpha \delta_{kj} s_{k_{kj}} \end{aligned} \quad [4.21]$$

Y para los pesos de las neuronas de la capa oculta:

$$\begin{aligned} W_{jp}(t+1) &= W_{kj}(t) + \Delta W_{kj}(t+1) \\ \Delta W_{kj}(t+1) &= \alpha \delta_{jp} s_{k_{jp}} \end{aligned} \quad [4.22]$$

En ambos casos, para acelerar el proceso de aprendizaje, se puede adicionar un término de momento de valor:

$$\beta_{kj}^0(t) - W_{kj}^0(t-1) \quad [4.23]$$

Para una neurona de salida y de valor:

$$\beta_{kj}^h(t) - W_{kj}^h(t-1) \quad [4.24]$$

Para una neurona oculta.

Paso 6

El proceso se repite hasta que el término de error.

$$Ep = \frac{1}{2} \sum_{j=1}^n \delta_{jk}^2 \quad [4.25]$$

Resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

V. Estudio comparativo entre los modelos de regresión múltiple y las redes neuronales multicapa aplicadas a la valoración inmobiliaria

5. Estudio comparativo entre los modelos de regresión múltiple y las redes neuronales multicapa aplicadas a la valoración inmobiliaria

5.1. Introducción

El primer paso consiste en obtener toda la información necesaria para plantear el problema. El objetivo que se persigue en este paso es adquirir un conocimiento técnico profundo de las observaciones para validar los modelos, la procedencia de las observaciones (ciudad) donde se han obtenido, con sus diferentes características o atributos que se tomarán en cuenta.

Las observaciones se obtuvieron en la ciudad de Morelia, municipio del mismo nombre, perteneciente al estado de Michoacán o provincia de Michoacán, en México; el organismo de gobierno que regula los asuntos relacionados con los predios urbanos es la Secretaría de Desarrollo Urbano y Medio Ambiente del municipio de Morelia.

5.2. Normativa urbanística de la ciudad de Morelia

En el año 2008 se promulgó el Código de Desarrollo Urbano del estado de Michoacán, con el objeto de regular, administrar y controlar todos los predios urbanos, adelante se detallan los artículos más importantes acerca de tal normativa urbana.

Características de diseño urbano e infraestructura de los desarrollos.

Artículo 313.

Los Fraccionamientos Habitacionales Urbanos, son aquellos que el Ayuntamiento podrá autorizar, ubicados dentro de los límites de un centro de población y sus lotes se aprovechen predominantemente para vivienda (Michoacán, 2008, p. 91).⁴

Artículo 314.

Los Fraccionamientos Habitacionales Urbanos tipo residencial, son aquéllos que se ubican en áreas cuya densidad de población puede ser mayor de 51 habitantes

⁴ Cuadernos Michoacanos de Derecho. Código de Desarrollo Urbano, 7ª Edición, 2008, Michoacán, México.

por hectárea pero no mayor de 150 habitantes por hectárea, deberán tener como mínimo, las características siguientes:

- I. Lotificación. Sus lotes no podrán tener una superficie menor de 300 metros cuadrados, sus frentes serán de 15.00 metros cuando se ubiquen sobre vialidades colectoras y principales y de 12.00 metros cuando tengan frente a vialidades secundarias;
- II. Usos del suelo. El aprovechamiento predominante será de vivienda unifamiliar y se permitirá solamente el diez por ciento de la superficie vendible para áreas comerciales o de servicios en las zonas autorizadas;

En este tipo de Desarrollos se permitirá la construcción de viviendas multifamiliares en un máximo del diez por ciento de la superficie vendible, en las zonas autorizadas, sin sobrepasar la densidad máxima permitida (Michoacán, 2008, p. 92).⁵

⁵ Ídem, p. 91.

Artículo 315.

Los Fraccionamientos Habitacionales Urbanos tipo medio, son aquellos ubicados en áreas cuya densidad población puede ser mayor de 151 habitantes por hectárea pero no mayor de 300 habitantes por hectárea y deberá tener como mínimo las características siguientes:

- I. Lotificación. Sus lotes no podrán tener una superficie menor de 200 metros cuadrados, sus frentes serán de 10.00 metros cuando se ubiquen sobre vialidades colectoras y principales y de 8.00 metros cuando tengan frente a vialidades secundarias;
- II. Uso del suelo. El aprovechamiento predominante será de vivienda unifamiliar y se permitirá solamente hasta el diez por ciento de la superficie vendible para áreas comerciales o de servicios en las zonas autorizadas (Michoacán, 2008, p. 93).⁶

Artículo 316.

Los Fraccionamientos Habitacionales Urbanos tipo interés social, son aquellos que se ubican en áreas cuya

⁶ Ídem. p. 91.

densidad de población puede ser mayor de 301 habitantes por hectárea, pero no mayor de 500 habitantes por hectárea y deberán tener como mínimo las características siguientes:

- I. Lotificación: Sus lotes no podrán tener una superficie menor de 96 metros cuadrados; sus frentes serán de 7.00 metros cuando se ubique sobre vialidades colectoras y principales y de 6.00 metros cuando tengan frente a vialidades secundarias y su fondo mínimo será de 16 metros;
- II. (*sic*)
- III. Usos del suelo: El aprovechamiento predominante será de vivienda y se permitirá solamente el quince por ciento de la superficie vendible para áreas comerciales o de servicios, en las zonas autorizadas;

En este tipo de Desarrollos Habitacionales Urbanos se permitirá la construcción de viviendas multifamiliares en un máximo del cincuenta por ciento de la superficie vendible en las zonas autorizadas, sin sobrepasar la densidad máxima permitida (Michoacán, 2008, p. 93).⁷

⁷ Ídem Pág.91

Artículo 317.

Los Fraccionamientos Habitacionales Urbanos tipo popular, son aquellos que por condiciones especiales de la zona en que se ubicarán, por la limitada capacidad económica de quienes vayan a habitarlos y por la urgencia inmediata de resolver problemas de vivienda, puede ser autorizados por el Ayuntamiento, con los requisitos mínimos de urbanización que se establecen en la fracción IV, incisos a, b, d, e, i y j del artículo anterior.

Artículo 321.

Las características mínimas que deberán cumplir los Fraccionamientos Habitacionales Suburbanos tipo campestre, son las siguientes;

- I. Lotificación. Sus lotes no podrán tener un frente menor de 30 metros, ni una superficie menor a 1,200 metros cuadrados.

Las construcciones deberán remeterse 5 metros a partir del parámetro; superficie que se dejará como área libre. Se permitirá la construcción como máximo en el

cuarenta y cinco por ciento de la superficie del lote y el resto se provechará en áreas verdes y espacios abiertos;

- II. Usos del suelo: El aprovechamiento predominante será de vivienda unifamiliar o para recreación y huertos familiares; cuando menos se deberá destinar un diez por ciento y no más del quine por ciento del área vendible para usos comerciales y de servicios.

Artículo 322.

Los fraccionamientos Habitacionales Suburbanos rústicos tipo granja deberán tener, como mínimo, las características siguientes:

- I. Lotificación. Sus lotes no podrán tener un frente menor de 35 metros, ni una superficie menor a 3,000 metros cuadrados; Se permitirá la construcción para habitación como máximo, en el quinde por ciento de la superficie del lote y el resto se aprovechara en actividades agropecuarias.
- II. Usos del suelo: el aprovechamiento predominante será para actividades agropecuarias. Cuando menos se deberá destinar un dos por ciento y no más del cinco del área vendible para usos comerciales o de servicios;

- III. Vialidad: las vialidades principales deberán tener como mínimo una anchura de 16 metros, medida de paramento a paramento y las vialidades secundarias una anchura de 12 metros, medida de paramento a paramento; las banquetas serán de 2 metros de ancho (Michoacán, 2008, p. 94-96).

5.3. Fuente de datos

En el subtítulo 5.1 se ha manifestado cómo se han obtenido las observaciones para validar los modelos, así pues falta definir la selección de los *inputs* o variables de entrada, donde se ha observado la influencia de diversos factores en la explicación del precio.

En la valoración inmobiliaria existen variables que influyen en el valor directamente y otras que lo hacen de forma inversa; lo que significa, por ejemplo: A mayor antigüedad menor valor, a menor antigüedad mayor valor.

Si se presenta este caso en alguna variable se debe cuidar desde principio al formular el modelo la transformación de esta variable, esto se corrige dividiendo $1/x$. La recomendación es usar variables con influencia directa.

Por su tipo, las variables se clasifican en:

- Nominales o categóricas.
- Ordinales.
- Numéricas.

VARIABLES NOMINALES O CATEGÓRICAS.- “Se definen a partir de diferentes categorías que no guardan relación (orden) entre sí. Ejemplo: Código postal, Orientación de la vivienda” (Guijarro 2013, p. 5).⁸

| Vivienda | Orientación |
|----------|-------------|
| 1 | Este |
| 2 | Oeste |
| 3 | Este |
| 4 | Norte |
| 5 | Norte |
| 6 | Este |
| 7 | Oeste |
| 8 | Sur |

VARIABLES ORDINALES.- “Vienen definidas por

⁸ Guijarro Martínez Francisco, *Estadística aplicada a la valoración. Modelos multivariantes de valoración*. Universidad Politécnica de Valencia, 2013.

diferentes niveles o categorías que guardan un orden” (Guijarro, 2013, p. 6).⁹

Por ejemplo: “Entorno comercial” con 3 diferentes niveles:

- Muy bueno
- Bueno
- Deficiente

| Vivienda | Entorno comercial |
|----------|-------------------|
| 1 | Muy bueno |
| 2 | Bueno |
| 3 | Deficiente |

Variables numéricas.- Se definen como variables cuantitativas, frente a la variables cualitativas que se han mencionado con anterioridad. Por ejemplo: Número de dormitorios, superficie de terreno, distancia a hospitales, etcétera.

En valoración inmobiliaria el precio se explica por medio de variables numéricas denominadas también cuantitativas y cualitativas.

⁹ Ídem, p. 5.

Adelante se detalla en que consiste cada una de las variables independientes.

Longitud. Ángulo formado entre el meridiano que pasa por el punto considerado y el meridiano fundamental de Greenwich elegido, por convención, como origen de esta coordenada geográfica. Lo que da como resultado la distancia en el eje “Y”.

Ejemplo: Columna F en el archivo en Excel.
101.08293 o 101° 08'2.93”.

Latitud. Columna G en el archivo en Excel.
19.41123 o 19° 41'12.3”.

El centro de la ciudad de Morelia, Michoacán, México (enseguida se muestran dos imágenes obtenidas de Google Earth), tiene las siguientes coordenadas: 19°42'08.82"- 101°11'35.07" y 1,923 metros sobre el nivel del mar (msnm). Véanse las figuras 5.1 y 5.2.

Figura 5.1. Centro de la ciudad

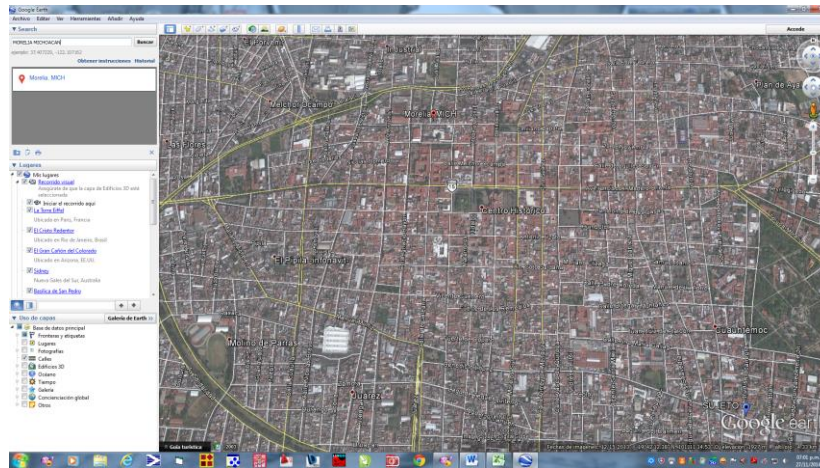


Figura 5.2. Centro de la ciudad



Así la latitud se define como la distancia que existe entre un punto cualquiera y el Ecuador, medida sobre el meridiano. Lo que da como resultado la distancia en el eje “X”.

Estas dos medidas de geoposicionamiento espacial dan como resultado la siguiente variable.

DCMacro.

Se trata de la variable a la que se ha denominado: macrolocalización, la distancia euclídea o distancia entre dos puntos, resultado de la distancia elevada al cuadrado de X e Y. Se ha observado en la investigación de los precios de los inmuebles que a medida que el inmueble está más alejado del centro de la ciudad, disminuye su precio y a medida que está más cerca, aumenta el precio.

Columnas J y K calculadas de dos formas: una la distancia euclídea y la otra como la suma de las distancias Y e X.

DCMicro.

Se trata de la variable que se ha denominado de microlocalización, la distancia euclídea, resultado de la

distancia elevada al cuadrado de X e Y, obtener raíz cuadrada al centro del barrio, la distancia medida en metros lineales al centro del barrio donde se ubica el inmueble que se observó. Se sabe gracias a la investigación de los precios de los inmuebles que al centro del barrio el valor de éstos es más alto, según estén más cerca del mismo.

Columnas N y O. De igual forma que la anterior la distancia euclídea y la suma de las distancias X e Y.

El centro histórico de la ciudad de Morelia está dividido en cuatro sectores.

Los sectores están divididos por la avenida Morelos norte y la avenida Francisco I. Madero oriente. Véase Figura 5.3.

Figura. 5.3. Sectores de la ciudad



La distancia MCMacro es cada una de las mediciones de GPS de cada observación o muestra al centro de la ciudad, debido a que el centro de la ciudad genera un mayor valor a los inmuebles. MCMacro significa macrolocalización.

Ejemplo: Macrolocalización de un inmueble al centro de la ciudad.

Superficie de terreno.

Se trata de la variable independiente, misma que se cuantifica en metros cuadrados.

Superficie construida o vendible.

Variable independiente que se cuantifica en metros cuadrados, se trata de la distribución de todos los espacios de la vivienda.

Número de baños.

Variable independiente en la que se cuantifica el número de baños completos y medios baños.

Número de dormitorios.

Variable independiente que cuantifica el número de dormitorios del inmueble.

Antigüedad.

Variable independiente que cuantifica la edad en años que tiene el inmueble.

Número de niveles.

Variable independiente que cuantifica el número de niveles con el que fue construido el inmueble (1, 2, 3, etcétera).

Infraestructura.

Variable independiente cualitativa, porque incluye los servicios como agua potable, drenaje, alcantarillado, energía eléctrica, banquetas, guarniciones y tipo de pavimentos en las calles; algunos corresponden a variables numéricas y otras a variables ordinales.

Para las variables nominales se usará la siguiente clasificación.

- 3 Muy bueno
- 2 Bueno
- 1 Suficiente
- 0 Deficiente

En la cual están consideradas las siguientes variables.

- Abastecimiento de agua.
- Drenaje sanitario.
- Electrificación.
- Alumbrado público.
- Transporte público.
- Recolección de basura.
- Vigilancia.
- Otros servicios.
- Guarniciones.
- Banquetas.

Equipamiento urbano o entorno urbano

En algunos países, principalmente los que están en vías de desarrollo -también conocidos como emergentes-, a esta variable se le suele denominar atributos de sustentabilidad de la vivienda; en los países europeos se le denomina entorno urbano. Por la preocupación global que existe del cambio climático derivado de la contaminación de todo tipo, esta variable en realidad es una metavariante que está compuesta por muchos atributos, algunos de ellos cuantitativos y otros cualitativos, que influyen directamente y otros que lo hacen inversamente proporcional al precio.

Los atributos que integran esta variable son los siguientes.

- Escuelas.- Primarias, secundarias, preparatorias, universidades etc., su medición será cuantificable en metros lineales.
- Consultorios o centros médicos.- Distancia en metros lineales hasta la vivienda a valorar.
- Servicio de transporte urbano.- Distancia a la parada (estación del transporte colectivo urbano) y tiempo en espera del mismo.

- Mercado o supermercados.- Medible en metros lineales.
- Plaza o jardines.- Distancia en metros lineales.
- Centros deportivos.- Distancia en metros lineales.
- Centros comunitarios.- Distancia en metros lineales.
- Contaminación general.
- Servicios (hospitales, restaurantes, bancos, escuelas, etcétera).
- Calidad de los comercios.
- Ancho de calle.
- Zonas verdes.
- Coeficiente de optimización del suelo (C.O.S.).
- Coeficiente de uso del suelo (C.U.S.).
- Densidad habitacional en la zona medida en la cantidad de habitantes por hectárea.
- Distancia al centro de la ciudad.

Las variables antes mencionadas son en principio las que originalmente se investigaron en campo. Con posterioridad se realizará un análisis estadístico, con el fin de verificar en cuáles de ellas habría que hacer una transformación, en particular las que expliquen con mayor certeza el precio de los inmuebles.

5.4. Modelización del precio a través de la regresión múltiple

En los apartados anteriores se ha realizado una revisión de la familia de los modelos de regresión múltiple, lineales y no lineales. Suele ser conveniente su uso para estimar el precio de los inmuebles, a partir de una muestra obtenida del mercado, conociendo bien los precios de transacción o bien los precios de oferta. Para ello se parte de una muestra compuesta por 75% de la venta cerrada o consumada y el resto se refiere a precios de oferta con factor de negociación de 5%, lo anterior con el fin de transformarlas todas a ventas consumadas. Respecto del factor de negociación se ha consultado a varios corredores inmobiliarios, quienes han proporcionado dicho porcentaje.

La muestra se compuso de observaciones tomadas desde el mes de octubre del año 2013 hasta noviembre de 2014.

En primer lugar se aplicará el modelo de regresión múltiple. Con una muestra de 961 inmuebles en sus diferentes segmentos, tomados de la ciudad de Morelia, Michoacán, México.

Para poder realizar la comparación entre el modelo econométrico y las redes neuronales artificiales perceptrón multicapa, se analizarán los coeficientes de correlación, coeficientes de determinación, el r^2 ajustado, el error estándar de estimación y el error cuadrático medio.

Mediante SPSS, se obtuvo la información que a continuación se detalla.

Tabla 5.1. Estadísticos descriptivos de la construcción

| | N | Mínimo | Máximo | Media | Desviación estándar |
|----------------------|-----|--------|---------|----------|---------------------|
| construcción | 985 | 34.00 | 1600.00 | 137.3924 | 112.98573 |
| N válido (por lista) | 985 | | | | |

Fuente: Elaboración propia

En las transacciones de inmuebles habitacionales, la variable que más consideran los demandantes, es la superficie construida o cubierta privativa. Depurando observaciones atípicas, se cuenta con 961 observaciones de inmuebles a utilizar para el estudio, con superficie construida mínima de 34 metros cuadrados y superficie máxima de 1.600 metros cuadrados, con una media de 137,39 metros cuadrados y una desviación estándar de 112,98 metros cuadrados.

Tabla 5.2. Estadísticos descriptivos del nivel de renta

| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 1.00 | 420 | 43.7 | 43.7 | 43.7 |
| | 2.00 | 389 | 40.5 | 40.5 | 84.2 |
| | 3.00 | 152 | 15.8 | 15.8 | 100.0 |
| | Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

En la tabla 5.2 se cuenta con 961 observaciones de inmuebles, donde se detalla la variable nominal denominada nivel de renta, especificada en tres categorías: nivel deficiente, bueno y muy bueno; en la tabla aparecen con su nivel de frecuencia y su porcentaje de participación del total de las observaciones.

Tabla 5.3. Estadísticos descriptivos de la infraestructura

| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 1.00 | 80 | 8.3 | 8.3 | 8.3 |
| | 2.00 | 608 | 63.3 | 63.3 | 71.6 |
| | 3.00 | 273 | 28.4 | 28.4 | 100.0 |
| | Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

De igual forma que el caso anterior, en la tabla 5.3, se cuenta con 961 observaciones de inmuebles, donde se detalla la variable nominal denominada infraestructura, especificada en tres categorías: nivel deficiente, bueno y muy bueno con su nivel de frecuencia y su porcentaje de

participación del total de las observaciones.

Tabla 5.4. Estadísticos descriptivos del número de garajes

| | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|------------|------------|-------------------|----------------------|
| Válido | | | | |
| 1.00 | 27 | 2.8 | 2.8 | 2.8 |
| 2.00 | 588 | 61.2 | 61.2 | 64.0 |
| 3.00 | 289 | 30.1 | 30.1 | 94.1 |
| 4.00 | 50 | 5.2 | 5.2 | 99.3 |
| Total | 7 | .7 | .7 | 100.0 |
| Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

La tabla 5.4 considera 961 observaciones de inmuebles, donde se detalla la variable nominal denominada número de garajes por vivienda, desde uno hasta cuatro con sus respectivas frecuencias por número de garajes, detallando el porcentaje del total de las observaciones.

Tabla 5.5. Estadísticos descriptivos del número de niveles

| | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|------------|------------|-------------------|----------------------|
| Válido | | | | |
| 1.00 | 34 | 3.5 | 3.5 | 3.5 |
| 2.00 | 283 | 29.4 | 29.4 | 33.0 |
| 3.00 | 609 | 63.4 | 63.4 | 96.4 |
| 4.00 | 17 | 1.8 | 1.8 | 98.1 |
| 5.00 | 8 | .8 | .8 | 99.0 |
| Total | 10 | 1.0 | 1.0 | 100.0 |
| Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

La tabla 5.5 muestra en cuantos niveles se encuentra construida cada vivienda, con un intervalo desde 1 y hasta 5 niveles, con su respectiva frecuencia por número de niveles, detallando el porcentaje del total de las observaciones.

Tabla 5.6. Estadísticos descriptivos del número de baños

| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 1 | 349 | 36.3 | 36.3 | 36.3 |
| | 1.5 | 192 | 20.0 | 20.0 | 56.3 |
| | 2 | 75 | 7.8 | 7.8 | 64.1 |
| | 2.5 | 203 | 21.1 | 21.1 | 85.2 |
| | 3 | 58 | 6.0 | 6.0 | 91.3 |
| | 3.5 | 53 | 5.5 | 5.5 | 96.8 |
| | 4 | 20 | 2.1 | 2.1 | 98.9 |
| | 4.5 | 8 | .8 | .8 | 99.7 |
| | 5 | 3 | .3 | .3 | 100.0 |
| | Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

La tabla 5.6 muestra el rango del número de baños por vivienda desde 1 hasta 5, y en las demás columnas se pueden observar la frecuencia y el porcentaje de participación del total de viviendas observadas.

Tabla 57. Estadísticos descriptivos del número de dormitorios

| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 1 | 39 | 4.1 | 4.1 | 4.1 |
| | 2 | 347 | 36.1 | 36.1 | 40.2 |
| | 3 | 471 | 49.0 | 49.0 | 89.2 |
| | 4 | 84 | 8.7 | 8.7 | 97.9 |
| | 5 | 13 | 1.4 | 1.4 | 99.3 |
| | 6 | 7 | .7 | .7 | 100.0 |
| | Total | 961 | 100.0 | 100.0 | |

Fuente: Elaboración propia

La tabla 5.7 muestra el rango del número de dormitorios por vivienda desde 1 hasta 6. En las demás columnas se observan la frecuencia y el porcentaje de participación del total de viviendas observadas.

Tabla 5.8. Estadísticos descriptivos del equipamiento urbano

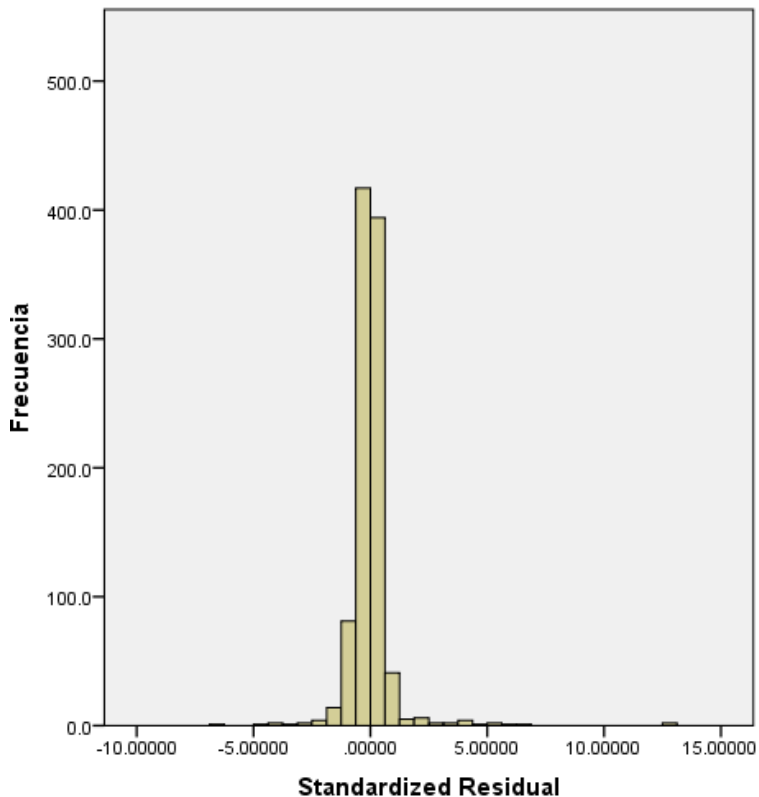
| | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
|--------|-------|------------|------------|-------------------|----------------------|
| Válido | 1 | 116 | 12.1 | 12.1 | 12.1 |
| | 2 | 521 | 54.2 | 54.2 | 66.3 |
| | 3 | 324 | 33.7 | 33.7 | 100.0 |
| | Total | 961 | 100.0 | 100.0 | |

De igual forma que el caso de la tabla 5.3, en la tabla 5.8 se detalla la variable nominal denominada equipamiento urbano, que contiene tres categorías: nivel deficiente, bueno y muy bueno, con su nivel de frecuencia y su porcentaje de participación del total de las observaciones.

Originalmente se contaba sólo con 985 observaciones y se han excluido del modelo 24, por ser atípicas en algunas de sus variables analizadas.

De acuerdo con las ecuaciones [3.15] y [3.16], anteriormente explicadas, y con base en la curva de distribución normal, se tomó la decisión de eliminar aquellas observaciones que se encuentren fuera del intervalo $[\mu+3\sigma; \mu-3\sigma]$, analizando la figura siguiente:

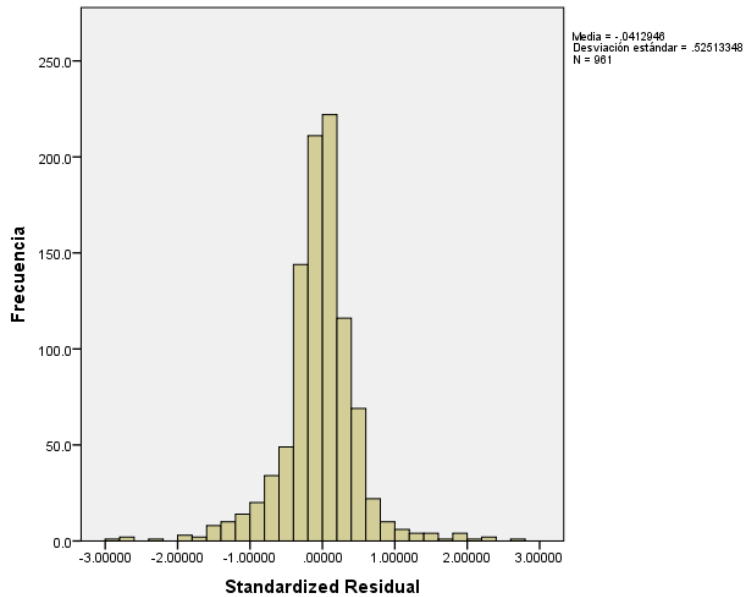
Figura 5.4. Grafica de residuos de muestras atípicas



Fuente: Elaboración propia

En la Figura 5.4 se observan datos atípicos de muestras que están fuera de los rangos establecidos para este trabajo de investigación, pues el intervalo elegido es de $[\mu+3\sigma; \mu-3\sigma]$. Una vez que se excluyeron las observaciones atípicas, se realizó la gráfica que se muestra en la figura siguiente.

Figura [5.5]. Gráfica de residuos estandarizados



Fuente: Elaboración propia

La figura 5.5 contiene resultados de los que se eliminaron 24 observaciones de la muestra original; una vez descartados los datos anómalos, los residuos estandarizados quedaron compuestos por 961 observaciones que cumplen con la distribución normal en el rango $[\mu+3\sigma; \mu-3\sigma]$, que es la muestra que se utilizará para comparar y validar los modelos antes explicados.

Para poder comparar entre el modelo econométrico

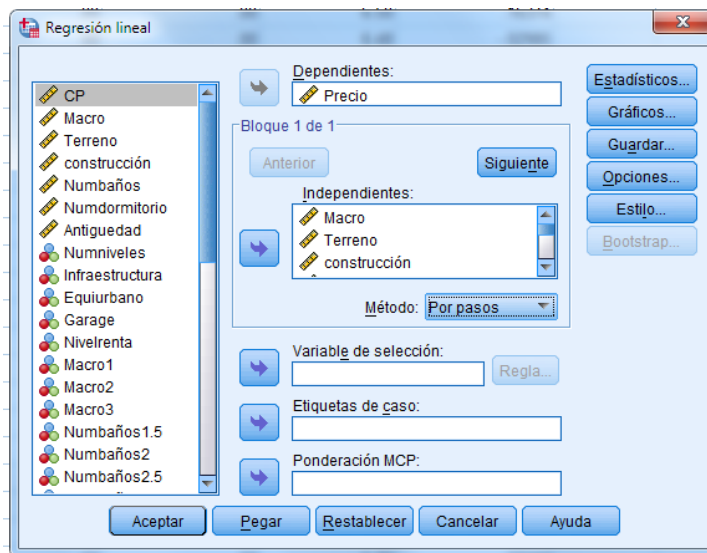
y las redes neuronales artificiales, se analizarán los errores de un modelo y otro.

Primero, se aplicará la regresión múltiple utilizando las variables anteriormente explicadas.

Para la aplicación de las ecuaciones [3.7], [3.8], [3.9] y [3.19] se utilizó el programa SPSS versión 22, de la empresa IBM.

Se han introducido todas las variables, como se observa en la figura siguiente.

Figura 5.6. Contenido de variable programa SPSS



Fuente: Elaboración propia

La figura 5.60 contiene información de las 961 observaciones utilizadas para la implementación del modelo de regresión múltiple. Aplicando el método de pasos sucesivos, se han guardado los residuos estandarizados, para graficarlos y verificar datos anómalos o atípicos y así no tomarlos en cuenta.

Al aplicar el modelo de regresión múltiple se obtuvieron los siguientes resultados:

Tabla 5.9. Resumen del modelo

| Modelo | R | R cuadrado | R cuadrado ajustado | Error estándar de la estimación |
|--------|------|------------|---------------------|---------------------------------|
| 1 | 0.95 | .902 | .900 | 314609.4772 |

Fuente: Elaboración propia

La tabla 5.9 contiene los resultados del modelo de regresión, en la columna 2 se calculó el coeficiente de correlación lineal de Pearson de acuerdo con la ecuación [3.7].

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}$$

Este coeficiente mide el grado de relación entre la variable dependiente con respecto de las variables independientes, en este caso es del [R], de 95%, que se considera como un coeficiente excelente. En la columna 3 se observa el resultado del coeficiente de determinación, obtenido como el resultado del cuadrado del coeficiente de correlación, de acuerdo con la ecuación [3.8], $[R_{\text{cuadrado}}]$, de 90.20%. Se puede considerar un coeficiente excelente, ya que trata de medir en qué proporción las variables o la variable independiente X explica la variabilidad de la variable dependiente Y. En la columna 4 se observa el $[R_{\text{cuadrado ajustado}}]$, obtenido con base en la ecuación [3.9].

$$r^2_{ajustada} = 1 - \frac{\text{Suma de cuadrados del error (SCE)}}{\text{Suma de cuadrados residuos del modelo (SCR)}}$$

Tabla 5.10. Tabla Anova, análisis de la varianza

| Modelo | | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
|--------|-----------|-------------------|-----|------------------|---------|-------------------|
| 1 | Regresión | 8.6.E+14 | 16 | 5.38E+13 | 543.184 | .000 ^a |
| | Residuo | 9.3.E+13 | 944 | 9.90E+10 | | |
| | Total | 9.5.E+14 | 960 | | | |

Fuente: Elaboración propia.

La tabla [5.10] muestra los resultados del análisis de la varianza, en la segunda columna se observan [8.6E+14], la suma de los cuadrados de la regresión,

[9.3E+14] como suma de los residuos o error de los cuadrados, y [9.5E+14] suma total de los cuadrados; en la columna 3, números de variables, los grados de libertad y el total de las observaciones; en la columna 4 la media cuadrática; en la columna 6 informa del grado de significación estadística, que en el caso de estudio está por debajo de 5% del modelo en su conjunto; dicho de otra manera el modelo arroja un nivel de confianza estadística del al menos 95%.

En la tabla siguiente se detallan las variables explicativas de los precios de los inmuebles de acuerdo con la ecuación [3.19], expresada de la siguiente manera.

$$Y = f P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \gamma$$

Con la cual se obtienen los resultados que se detallan en la tabla siguiente.

Tabla 5.11. Resumen de los coeficientes de las variables explicativas del precios de los inmuebles

| Modelo | | Coeficientes no estandarizados | | Coeficientes estandarizados | t | Sig. |
|--------|-----------------|--------------------------------|----------------|-----------------------------|-------|------|
| | | B | Error estándar | Beta | | |
| 1 | (Constante) | -518793.21 | 70260.95 | | -7.38 | .000 |
| | construcción | 7317.29 | 235.75 | .615 | 31.04 | .000 |
| | Nivelrentaalta | 352399.00 | 38533.82 | .129 | 9.15 | .000 |
| | Infraestructura | 120419.43 | 25680.44 | .069 | 4.69 | .000 |
| | Garage | 280536.89 | 37852.79 | .188 | 7.41 | .000 |
| | Garage2 | -278499.25 | 43322.96 | -.128 | -6.43 | .000 |
| | Terreno | 980.59 | 158.71 | .092 | 6.18 | .000 |
| | Numniveles3 | -423183.09 | 79627.03 | -.056 | -5.31 | .000 |
| | Numbañõs5 | 1011584.58 | 190220.02 | .057 | 5.32 | .000 |
| | Antigüedad | -7384.90 | 1416.62 | -.060 | -5.21 | .000 |
| | Numbañõs4 | 385868.09 | 78655.03 | .055 | 4.91 | .000 |
| | Nivelrentabaja | -85291.50 | 27284.30 | -.042 | -3.13 | .002 |
| | Numdormitorio6 | -356332.40 | 122263.06 | -.030 | -2.91 | .004 |
| | Numbañõs2 | -104478.40 | 38415.92 | -.028 | -2.72 | .007 |
| | Garage3 | -237462.32 | 86847.20 | -.053 | -2.73 | .006 |
| | Numbañõs4.5 | 296949.62 | 119440.68 | .027 | 2.49 | .013 |
| | Equiurbano | 49521.80 | 23183.10 | .032 | 2.14 | .033 |

a. Variable dependiente: Precio

Fuente: Elaboración propia

En la columna 1 se encuentra el nombre de cada una de las variables explicativas; en la columna 2 están los coeficientes de cada una de las variables que explican el precio; en la columna 3 muestra el error estándar de cada una de las variables, incluida la constante; en la columna 4 se aprecian los coeficientes estandarizados de cada variable; en la columna 5 se refiere al valor t y en la columna 6 está la significación de cada una de las variables, todas por debajo de 5%, con un nivel de

confianza del 95%.

La ecuación resultante sería la siguiente:

$$Y = -518793.21 + 7317.29 * \text{Superficie construida} + 352399 * \text{Nivelrentaalta} + 120419.43 * \text{infraestructura} + 280536.89 * \text{Garage} - 79499 * \text{Garage2} + 980.59 * \text{Terreno} - 423183 * \text{Numniveles3} + 1011586 * \text{Numbaños5} - 7385 * \text{Antigüedad} + 385868 * \text{Numbaños4} - 85292 * \text{Nivelrentabaja} - 356332 * \text{Numdormitorio6} - 104478 * \text{Numbaños2} - 237462 * \text{Garage3} + 296950 * \text{Numbaños4.5} + 49522 * \text{Equiurbano}$$

Mediante la ecuación anterior se estimará el precio de un inmueble de la muestra con las siguientes características:

Cuenta con 240 metros cuadrados de terreno, 250 metros cuadrados de construcción, 3.5 baños, 3 dormitorios, 0 años de antigüedad (nueva), 2 niveles, nivel 3 de infraestructura, nivel 3 de equipamiento, 3 garage, 0 baños, 0 garage2, 1 garage3, 0 nivel de renta baja, 1 nivel de renta alta, 0 baños4, 0 baño 4.5, 0 baños5, 0 dormitorio6 y 0 numniveles3.

A continuación se expone a modo de ejemplo el precio estimado para una vivienda de determinadas

características, a partir de la ecuación anteriormente obtenida.

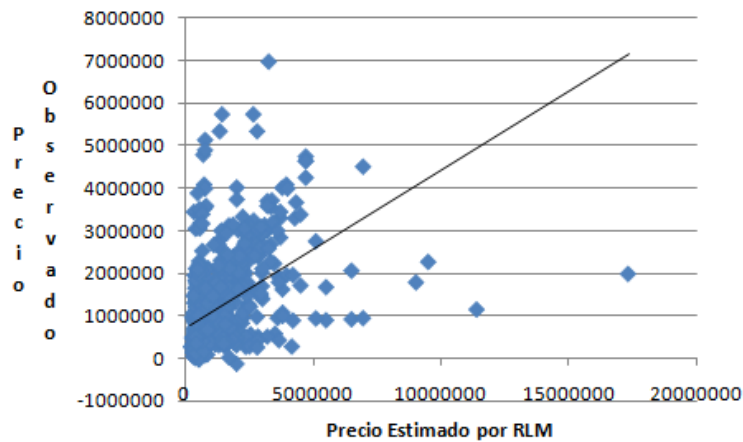
Tabla. 5.12. Representa la estimación de valor de un inmueble

| Nombre de la variable | Coficiente | Cantida de la variable | Valor parcial |
|-----------------------------|------------|------------------------|----------------|
| (Constante) | -518793 | 1 | -518793 |
| construcción | 7317 | 250 | 1829322 |
| Nivelrentaalta | 352399 | 1 | 352399 |
| Infraestructura | 120419 | 3 | 361258 |
| Garage | 280537 | 3 | 841611 |
| Garage2 | -278499 | 0 | 0 |
| Terreno | 981 | 240 | 235342 |
| Numniveles3 | -423183 | 0 | 0 |
| Numbaños5 | 1011585 | 0 | 0 |
| Antigüedad | -7385 | 0 | 0 |
| Numbaños4 | 385868 | 0 | 0 |
| Nivelrentabaja | -85291 | 0 | 0 |
| Numdormitorio6 | -356332 | 0 | 0 |
| Numbaños2 | -104478 | 0 | 0 |
| Garage3 | -237462 | 1 | -237462 |
| Numbaños4.5 | 296950 | 0 | 0 |
| Equiurbano | 3 | 23183 | 69549 |
| Valor estimado del inmueble | | | 2933226 |

Fuente. Elaboración propia

Acerca del comportamiento de los precios observados y estimados véase la siguiente figura:

Figura 5.7. Gráfica de los precios observados y estimados mediante regresión lineal múltiple



Fuente. Elaboración propia.

En la figura [5.7.] se ilustra el comportamiento de los precios observados y los precios estimados mediante regresión lineal múltiple.

De la relación entre estas dos variables se deduce un problema de heterocedasticidad. Para mitigarlos, se hará uso de la variante logarítmica para la definición de algunas variables.

De nuevo se aplicará una regresión múltiple, en su variante semilogarítmica, empleando las mismas variables

independientes, pero con la diferencia de que como variable dependiente se utiliza el logaritmo del precio, de acuerdo con la ecuación [3.20].

$$\ln(P) = \beta X + \gamma$$

Reacomodando las variables independientes, queda de la siguiente manera.

$$\ln(P) = \beta_0 + \beta_1 * \text{Superficie Construida} + \beta_2 * \text{Nivelrentaalta} + \dots + \beta_n * \text{Equiurbano} + \gamma.$$

Tabla 5.13. Resumen del modelo.

| Modelo | R | R cuadrado | R cuadrado ajustado | Error estándar de la estimación |
|--------|------|------------|---------------------|---------------------------------|
| 1 | .944 | .891 | .890 | .10092 |

Fuente: Elaboración propia.

La tabla [5.13] contiene los resultados del modelo de regresión.

El r^2 ajustado en la columna 4 es de 89%, también muy bueno.

Tabla 5.14. Resultado ANOVA. Análisis de la varianza

| Modelo | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
|-----------|-------------------|-----|------------------|---------|-------------------|
| 1 | | | | | |
| Regresión | 77.000 | 10 | 7.700 | 755.989 | .000 ^a |
| Residuo | 9.432 | 926 | .010 | | |
| Total | 86.432 | 936 | | | |

Fuente: Elaboración propia

La tabla [5.14] muestra los resultados del análisis de la varianza, en la segunda columna se observan [77.00], la suma de los cuadrados de la regresión, [9.432] como suma de los residuos o error de los cuadrados, y [86.432] suma total de los cuadrados; en la columna 3, números de variables, los grados de libertad y el total de las observaciones; en la columna 4 la media cuadrática; en la columna 6 informa del grado de significación estadística, que en el caso de estudio está por debajo de 5% del modelo en su conjunto; dicho de otra manera el modelo arroja un nivel de confianza estadística del al menos 95%.

Tabla 5.15. Resumen de los coeficientes de las variables explicativas del precio de los inmuebles

| Modelo | | Coeficientes no estandarizados | | Coeficientes estandarizados | t | Sig. |
|--------|-----------------|--------------------------------|----------------|-----------------------------|---------|-------|
| | | B | Error estándar | Beta | | |
| 1 | (Constante) | 5.259 | .017 | | 301.748 | 0.000 |
| | construcción | .002 | .000 | .577 | 29.743 | .000 |
| | Nivelrentaalta | -.120 | .015 | -.138 | -8.061 | .000 |
| | Infraestructura | .041 | .008 | .076 | 5.343 | .000 |
| | Terreno | .000 | .000 | .083 | 5.068 | .000 |
| | Antigüedad | -.002 | .000 | -.061 | -5.223 | .000 |
| | Numniveles3 | -.071 | .025 | -.031 | -2.835 | .005 |
| | Nivelrenta | .171 | .009 | .396 | 19.777 | .000 |
| | Numbaños4.5 | -.095 | .038 | -.029 | -2.527 | .012 |
| | Numdormitorio | -.145 | .039 | -.041 | -3.693 | .000 |
| | Equiurbano3 | .048 | .009 | .073 | 5.156 | .000 |

Fuente: Elaboración propia.

En la columna 1 se encuentra el nombre de cada una de las variables explicativas; en la columna 2 están los coeficientes de cada una de las variables que explican el precio; en la columna 3 muestra el error estándar de cada una de las variables, incluida la constante; en la columna 4 se aprecian los coeficientes estandarizados de cada variable; en la columna 5 se refiere al valor t y en la columna 6 está la significación de cada una de las variables, todas por debajo de 5%, con un nivel de confianza del 95%.

La ecuación quedaría de la siguiente manera:

$$Y = 5.259 + .002 * \text{Superficie} - .120 * \text{Nivelrentaalta} + .04 * \text{Infraestructura} + .000 * \text{Terreno} -$$

.002*Antigüedad-.071*Numniveles3+ .171* Nivel renta-
 .095*Numbaños4.5- .145*Numdormitorio+.048*
 Equiurbano3.

Los resultados se detallan en la tabla siguiente.

Tabla 5.16. Representa la estimación del precio de un inmueble

| Nombre de la variable | Coefficiente | Cantidad de la variable | Valor parcial |
|------------------------------|--------------|-------------------------|---------------|
| (Constante) | 5.259 | 1 | 5.259 |
| Construcción | 0.002 | 250 | 0.5 |
| Nivelrentaalta | -0.12 | 1 | -0.12 |
| Infraestructura | 0.041 | 3 | 0.123 |
| Terreno | 0 | 240 | 0 |
| Antigüedad | -0.002 | 0 | 0 |
| Numniveles3 | -0.071 | 0 | 0 |
| Nivelrenta. | 0.171 | 3 | 0.513 |
| Numbaños4.5 | -0.095 | 0 | 0 |
| Numdormitorio6 | -0.145 | 0 | 0 |
| Equiurbano3 | 0.048 | 1 | 0.048 |
| Suma de logaritmo | | | 6.323 |
| Precio estimado del inmueble | | | 2103778 |

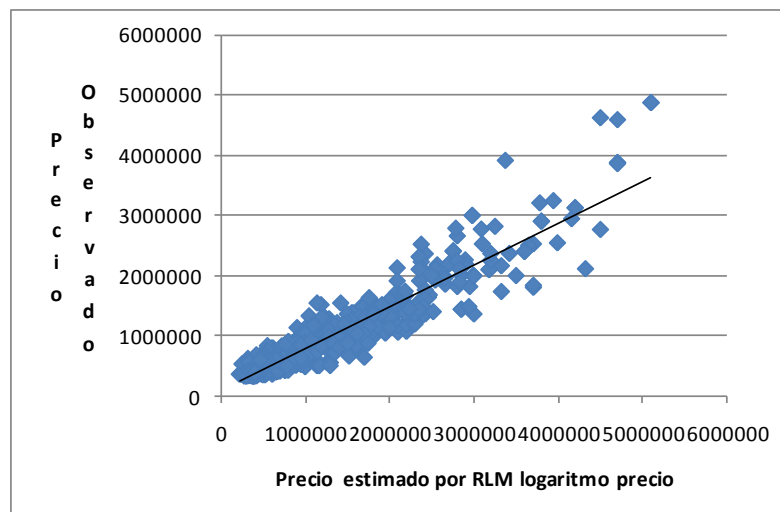
Fuente: Elaboración propia.

La tabla [5.16] en la primer columna contiene las variables del modelo, en la segunda los coeficientes de logaritmo de cada variable, en la tercer columna la

cantidad de la variable y en la cuarta columna el valor parcial del logaritmo de cada variable, aplicando antilogaritmo en la suma de la cuarta columna de la siguiente manera $y = 10^{6.323}$, da como resultado la cantidad de 2,103,778, que corresponde al valor estimado del inmueble.

La siguiente figura es un gráfico que refleja el comportamiento del precio.

Figura 5.8. Gráfica de los precios observados y estimados mediante el logaritmo del precio



Fuente: Elaboración propia.

La figura [5.8.] representa el gráfico de los precios observados y los precios estimados mediante el logaritmo

del precio, se observa que ya no presenta problemas de homocedasticidad.

Puede concluirse, por tanto, que el modelo de regresión multivariable es significativo en general y en cada una de sus variables.

En el modelo de regresión logarítmica es significativo, en general y en cada una de sus variables, solo que se han utilizado 936 y 10 variables explicativas, 24 observaciones menos y 6 variables menos que en el modelo de regresión lineal, para que pudiera resultar significativo con un nivel de confianza del 95%.

5.5. Modelización del precio mediante red neuronal artificial perceptrón multicapa

En este apartado se explicará el precio a través del modelo perceptrón multicapa (red multicapa), mediante una red neuronal multicapa *back propagation*, ya descrita anteriormente en sus diferentes funciones de activación y las diferentes funciones en la capa de salida.

La corrección de los errores se basa en un aprendizaje de tipo supervisado, por lo que para el entrenamiento se han indicado las entradas o *inputs* que se tienen en la red y las salidas que se desean obtener para dichas entradas.

Para la simulación y entrenamiento de la red se ha utilizado el programa SPSS versión 22, de la empresa IBM.

En cuanto a capas ocultas (intermedias), este programa permite la existencia de varias capas ocultas (para la existencia de la red *back propagation*, es necesario como mínimo una capa oculta).

En cuanto al ratio de aprendizaje, para este caso se utilizará una capa oculta, con 16 variables en la capa

oculta y 961 neuronas (observaciones), con una tasa de aprendizaje inicial muy pequeña, hasta alcanzar una tasa del 0.40 y diferentes impulsos. Adicionalmente se iniciará con intervalos de tiempo desde 10 minutos hasta 100 minutos y se irá observando el mejoramiento o empeoramiento de los resultados de salida de la red.

La flexibilidad del programa permite la posibilidad de que las neuronas puedan utilizar funciones de transferencia de forma distinta. Así pues se utilizan funciones lineales, tangente hiperbólica y sigmoideas.

A continuación se estimarán los precios de inmuebles utilizando los modelos de redes neuronales artificiales, en específico el perceptrón multicapa en sus diferentes maneras de activación a la entrada de patrones e igualmente en sus salidas a la red. Utilizando las ecuaciones [4.2], [4.10], [4.12] y [4.15].

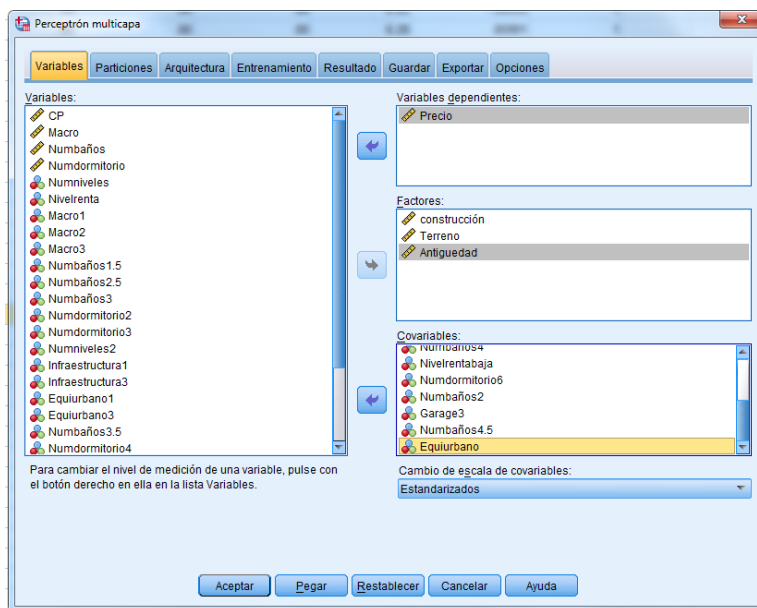
Con las mismas observaciones y variables utilizadas en los modelos de regresión, para no dar ventaja a ninguno, se hará la estimación de los precios de los inmuebles.

Se iniciará con el entrenamiento de una red con

entradas de neuronas en la capa oculta, utilizando una función tangente hiperbólica y una salida identidad (lineal), con diferentes intervalos de tiempo y se evaluarán los resultados para a su vez verificar la mejoría de la misma.

La siguiente figura contiene información de las variables a utilizar en el modelo, mediante el programa SPSS.

Figura 5.9. Variables a utilizar en el modelo RNA

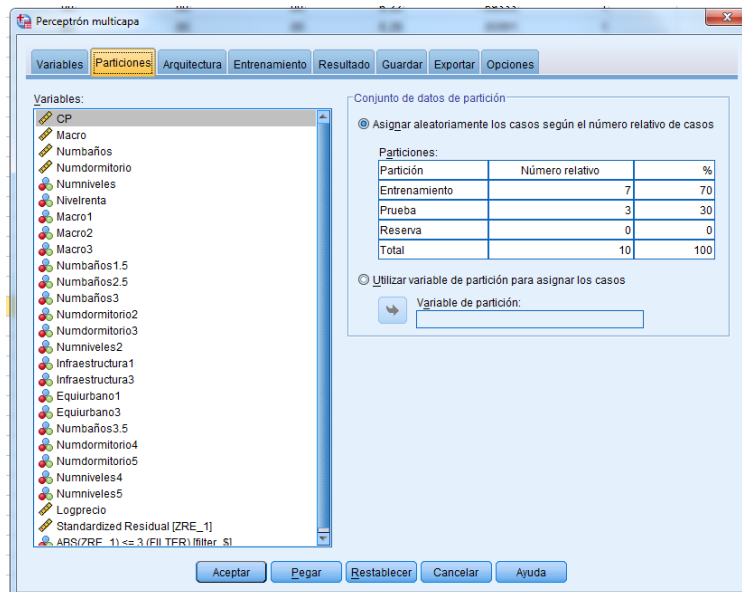


Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red: 75% de las observaciones se

utilizarán para entrenamiento del modelo y 25% para validar que tan bien explica el precio el modelo. Se trata de las 16 variables independientes y una variable dependiente, las mismas variables utilizadas en los modelos de regresión, para no dar ventaja a ningún modelo en la obtención de resultados.

Figura. 5.10. Patrones de entrada a la red, entrenamiento y prueba del mismo

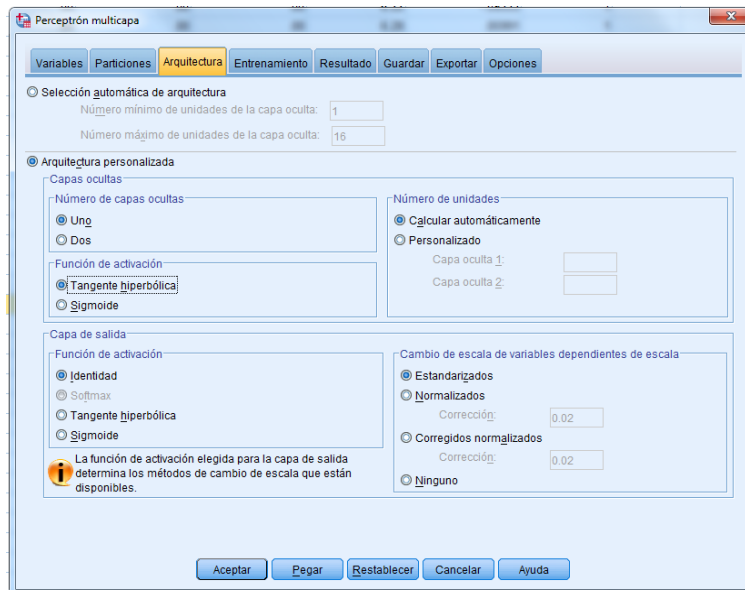


Fuente: Elaboración propia

La figura [5.10] contiene información sobre la arquitectura de la red, se ha seleccionado una capa oculta con un máximo de 16 unidades o patrones en la misma, con una función de activación a la entrada tangente

hiperbólica y una función de salida de la red denominada identidad (lineal).

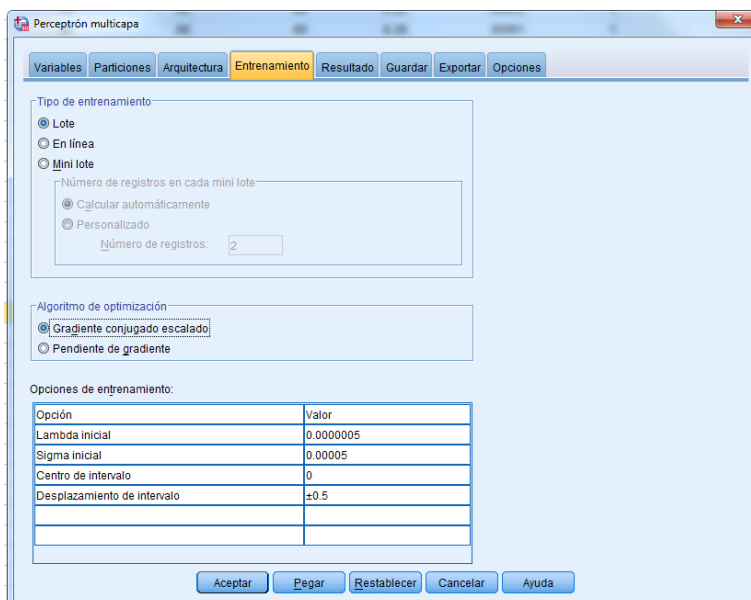
Figura 5.11. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura [5.11] se observa el tipo de entrenamiento aquí denominado en lote, con sus diferentes opciones de entrenamiento.

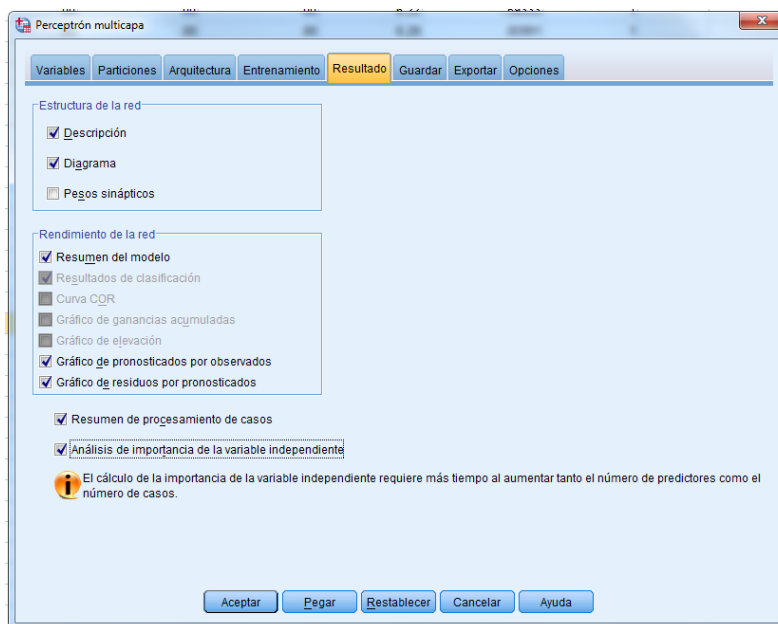
Figura 5.12. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura [5.12], los resultados de la red, en este caso descripción, diagramas, resumen del modelo, gráficos de pronósticos observados, gráficos de residuos por pronósticos, resumen de procesamientos de casos y análisis de importancia de la variable independiente.

Figura 5.13. Resultado de la RNA



Fuente: Elaboración propia

Una vez ejecutado el entrenamiento de la red como se explicó, se han obtenido los resultados que se detallan en las siguientes tablas.

Tabla 5.17. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 667 | 75.6% |
| | Pruebas | 215 | 24.4% |
| Válido | | 882 | 100.0% |
| Excluido | | 79 | |
| Total | | 961 | |

Fuente: Elaboración propia

Se puede observar en la tabla [5.14], que de las 961 observaciones, 75.60% (667) se han utilizado para el entrenamiento del modelo, en 24.40% (215) se estimaron los precios para validar el modelo y se han excluido 79, por ser atípicas.

Tabla 5.18. Errores de predicción de la RNA

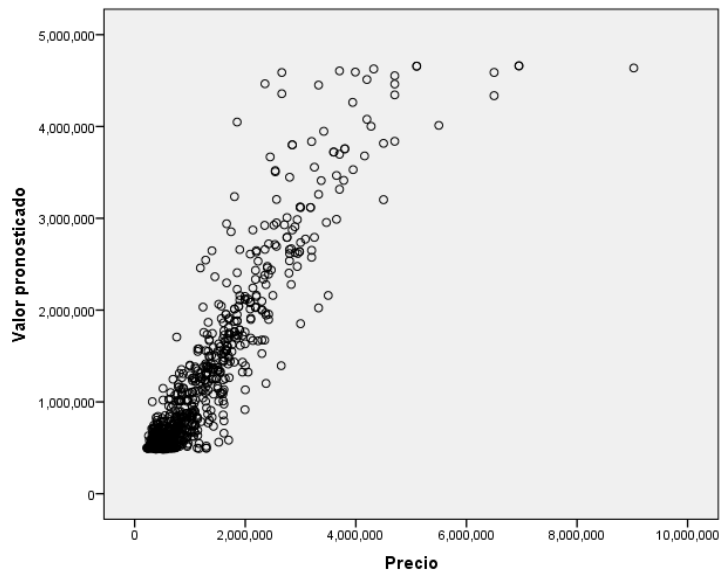
| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 46.846 |
| | Error relativo | .141 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:00.53 |
| Pruebas | Error de suma de cuadrados | 24.179 |
| | Error relativo | .195 |

Fuente: Elaboración propia

En la tabla [5.18] se muestra la suma de errores de 46.84%, un error relativo de 0.141%. Estos errores están contruidos con las muestras de entrenamiento de la red, los errores mediante las pruebas, refiriéndose a las muestras en donde se han calculado los precios de los inmuebles. Además está la suma de errores de cuadrados

de 24.179% y un error relativo de 0.195%. A continuación se muestra una gráfica de precio observado y precio pronosticado.

Figura 5.14. Gráfica de precio observado y precio pronosticado

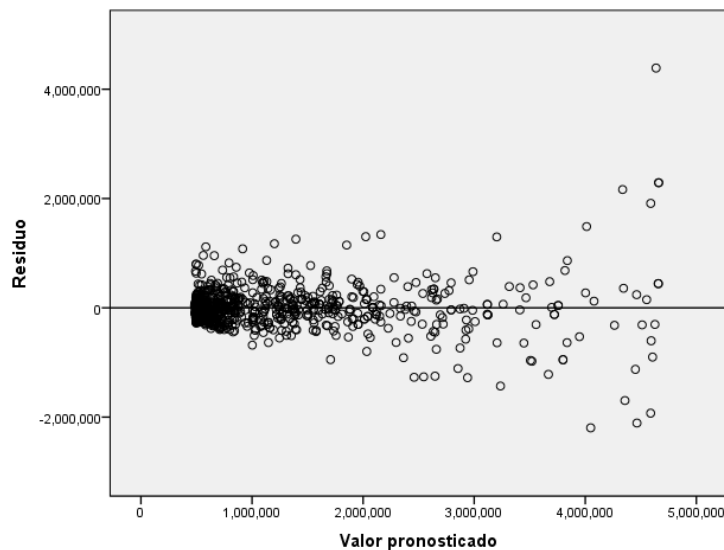


Fuente: Elaboración propia

En la figura [5.14] se muestra la gráfica que representa en el eje x, el precio de los inmuebles observados y en el eje de la y, los precios explicados o pronosticados por el modelo, además se observa la relación directamente proporcional entre ellos.

Continuando en la figura [5.15], se destaca lo siguiente.

Figura 5.15. Gráfica de residuos y valor pronosticado



Fuente: Elaboración propia

La figura [5.15] representa un gráfico de los residuos contra los valores pronosticados; existe un comportamiento normal a excepción de algunas muestras alejadas de la nube de puntos, pero se reitera que ya fue realizada la prueba de normalidad y los datos cumplen con la misma en el rango $[\mu+3\sigma; \mu-3\sigma]$. La explicación estaría nuevamente en un comportamiento heterocedástico de la muestra.

En la tabla siguiente se puede notar la importancia de las variables independientes.

Tabla 5.19. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| construcción | .109 | 53.0% |
| Terreno | .143 | 69.6% |
| Antigüedad | .058 | 28.2% |
| Nivelrentaalta | .021 | 10.1% |
| Infraestructura | .001 | .5% |
| Garage | .205 | 100.0% |
| Garage2 | .026 | 12.7% |
| Numniveles3 | .034 | 16.5% |
| Numbaños5 | .062 | 30.3% |
| Numbaños4 | .062 | 30.2% |
| Nivelrentabaja | .056 | 27.4% |
| Numdormitorio6 | .008 | 4.1% |
| Numbaños2 | .007 | 3.5% |
| Garage3 | .035 | 17.2% |
| Numbaños4.5 | .147 | 71.8% |
| Equiurbano | .026 | 12.5% |

Fuente: Elaboración propia

La tabla [5.19] indica las variables independientes en la columna 1, en la columna 2 la importancia de cada la variable y en la columna 3 la importancia normalizada de las variables independientes.

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenido los siguientes resultados.

Tabla 5.20. *Test* de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Traning(entrenamiento) | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|------------------------|---------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Identidad | 667 | 215 | 79 | 10 minutos | 89.07% | 79.33% | 78.95% | 448511 | 448002 |
| Tangente hiperbólica | Identidad | 667 | 205 | 89 | 20 minutos | 93.37% | 87.17% | 86.93% | 340906 | 340515 |
| Tangente hiperbólica | Identidad | 651 | 229 | 81 | 30 minutos | 88.36% | 78.07% | 77.66% | 466684 | 466154 |
| Tangente hiperbólica | Identidad | 669 | 227 | 65 | 40 minutos | 75.64% | 57.22% | 56.44% | 630402 | 629698 |
| Tangente hiperbólica | Identidad | 660 | 219 | 82 | 50 minutos | 91.56% | 83.83% | 83.53% | 408509 | 408044 |
| Tangente hiperbólica | Identidad | 677 | 201 | 83 | 60 minutos | 90.50% | 81.90% | 81.56% | 388264 | 387822 |
| Tangente hiperbólica | Identidad | 671 | 223 | 67 | 80 minutos | 90.58% | 82.05% | 81.73% | 428209 | 427730 |
| Tangente hiperbólica | Identidad | 676 | 215 | 70 | 100 minutos | 87.13% | 75.92% | 75.48% | 481093 | 480553 |

Fuente: Elaboración propia

La tabla 5.20 contiene los resultados del entrenamiento de la red con un factor de activación en la capa oculta con una función de entrada tangente hiperbólica y un factor de activación en la capa de salida con una función identidad. En la columna tres se observa el número de observaciones utilizadas para el entrenamiento (*training*), en la columna cuatro el número de observaciones para realizar pruebas (*test*). En la columna cinco se observan las observaciones que han

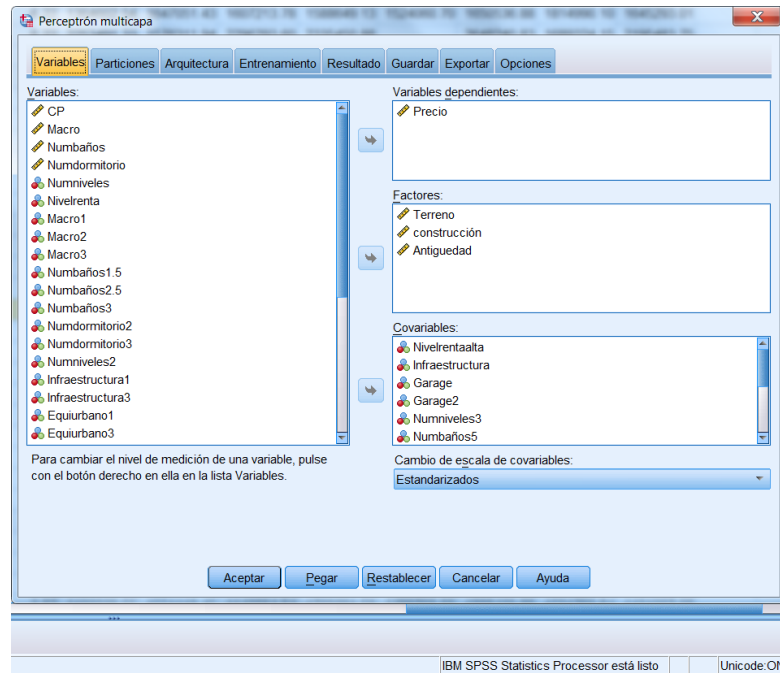
sido excluidas por el modelo por ser atípicas, en la columna seis se describe el tiempo de entrenamiento de la red, en la columna siete el coeficiente de correlación, en la columna ocho se observa el resultado del coeficiente de determinación, en la columna nueve el r^2 cuadrado ajustado, y en la columna diez el error cuadrático medio.

Se puede concluir en este modelo que de los ocho resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el *test* de entrenamiento con tiempo de 20 minutos es en el que mejor se estimó el precio; para ello en el modelo se utilizaron 667 observaciones de entrenamiento, 205 para prueba y se eliminaron 89 muestras, lo cual arroja un coeficiente de correlación de 93%, y un coeficiente de determinación de 86.93%.

Utilizando los mismos parámetros del modelo anterior, ahora con una red de activación en las entradas ocultas utilizando una función tangente hiperbólica y una función de activación con salida tangente hiperbólica.

La siguiente figura contiene la información de las variables usadas en el modelo, mediante el programa SPSS.

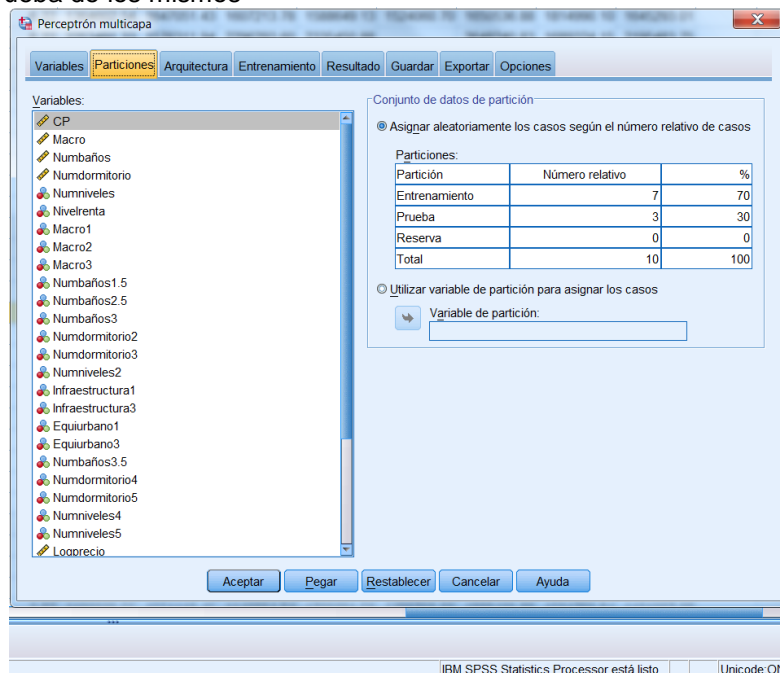
Figura 5.14. Variables a utilizar en el modelo RNA



Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red, 70% de las observaciones se utilizarán para entrenamiento del modelo y 30% para validar que tan bien explica el precio el modelo. Se trata de 16 variables independientes y una variable dependiente.

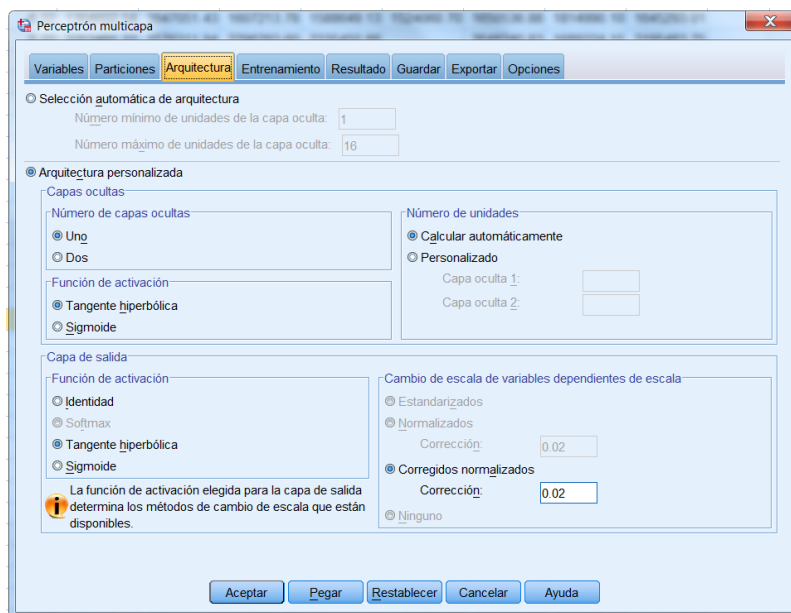
Figura. 5.15. Patrones de entrada a la red, entrenamiento y prueba de los mismos



Fuente: Elaboración propia

La figura 5.16 contiene información sobre la arquitectura de la red, en este caso se ha seleccionado una capa oculta de un máximo de 16 unidades o patrones en la misma, con una función de activación a la entrada tangente hiperbólica y una función de salida de la red denominada tangente hiperbólica.

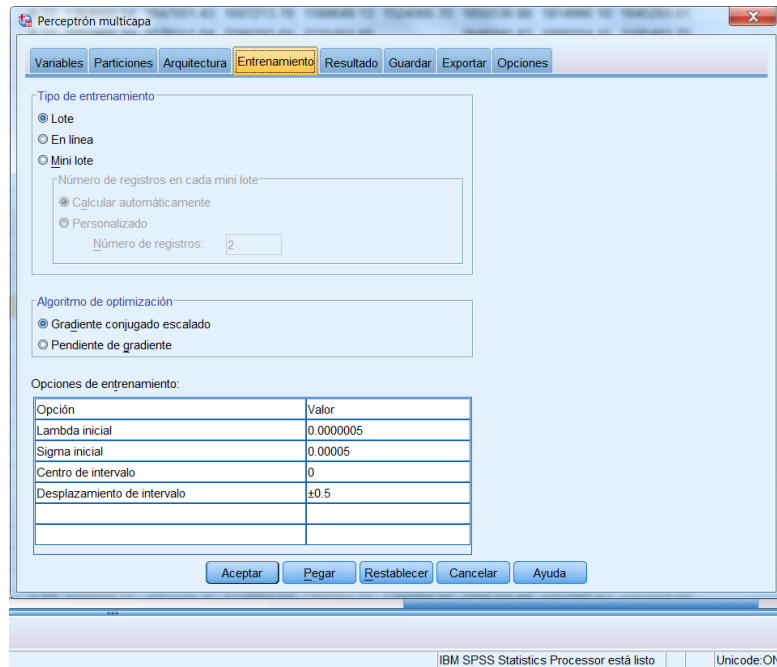
Figura 5.16. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.17 se observa el tipo de entrenamiento denominado en lote, con sus diferentes opciones de entrenamiento.

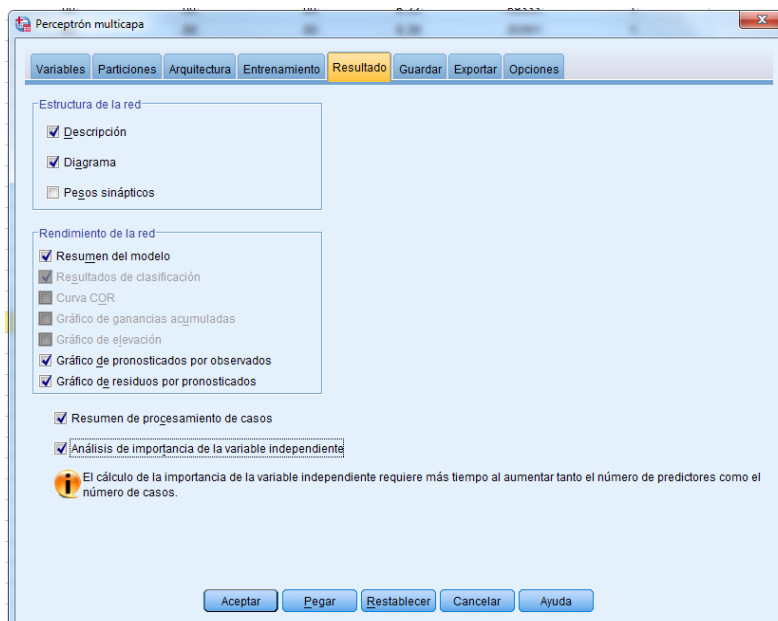
Figura 5.17. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.18 están los resultados de la red: descripción, diagramas, resumen del modelo, gráficos de pronósticos por observados, gráficos de residuos por pronósticos, resumen de procesamientos de casos y análisis de importancia de la variable independiente.

Figura 5.18. Resultado de la RNA



Fuente: Elaboración propia

Una vez ejecutado el entrenamiento de la red, se han obtenido los resultados que se detallan en las siguientes tablas:

Tabla 5.21. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 671 | 75.1% |
| | Pruebas | 223 | 24.9% |
| Válido | | 894 | 100.0% |
| Excluido | | 67 | |
| Total | | 961 | |

Fuente: Elaboración propia

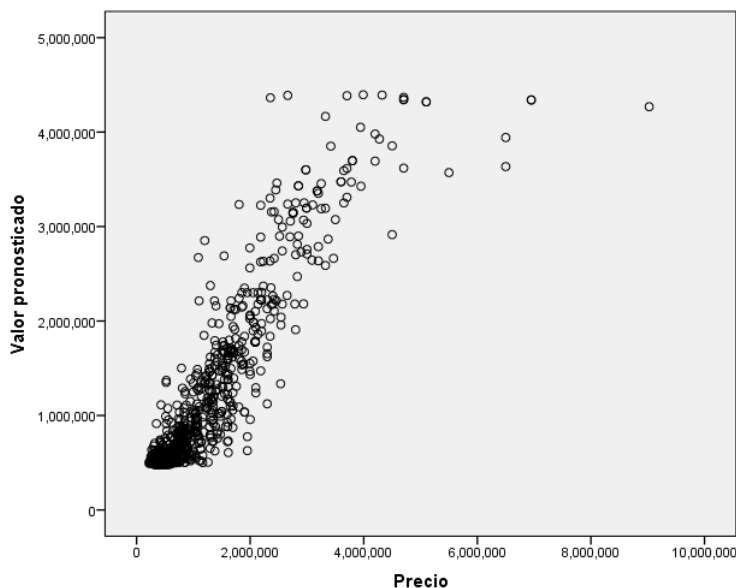
Tabla 5.22. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 2.347 |
| | Error relativo | .136 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:01.44 |
| Pruebas | Error de suma de cuadrados | 1.658 |
| | Error relativo | .264 |

Fuente: Elaboración propia

En la tabla 5.22 se muestran la suma de errores de 2.347, un error relativo del 0.136. A continuación se muestra una gráfica de precio observado y precio pronosticado.

Figura 5.19. Gráfica de precios pronosticados y precios observados

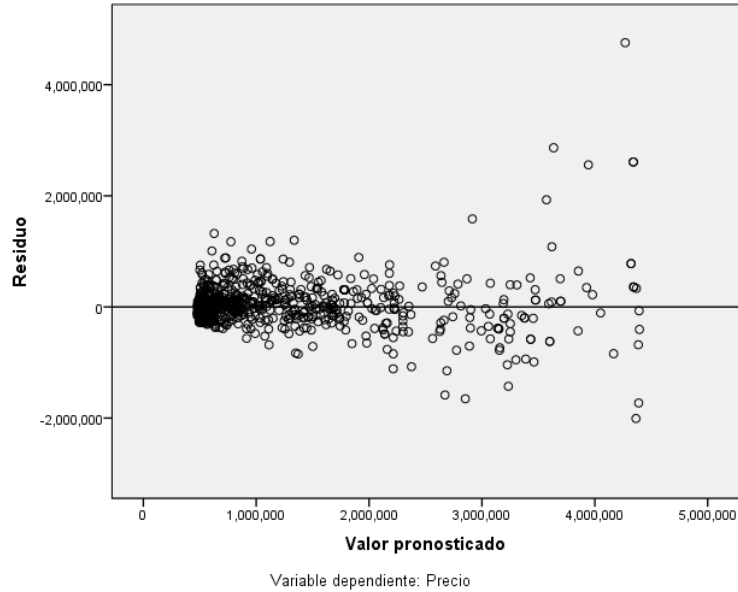


Fuente: Elaboración propia

En la figura 5.19 está la gráfica que representa en el eje X, el precio de los inmuebles observados, y en el eje de la Y los precios pronosticados por el modelo. Además se distingue una relación directamente proporcional entre ellos, pero de nuevo con algún problema de heterocedasticidad.

Continuando en la figura 5.20 se observa lo siguiente.

Figura 5.20. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura 5.20 representa un gráfico del residuo contra el valor pronosticado; con un comportamiento normal a excepción de algunas muestras alejadas de la nube de puntos.

En la tabla siguiente se describe la importancia de las variables independientes.

Tabla 5.21. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|----------------|-------------|-------------------------|
| Terreno | .123 | 49.9% |
| construcción | .121 | 49.2% |
| Antigüedad | .065 | 26.6% |
| Infraestructur | .022 | 8.8% |
| Equiurbano | .021 | 8.5% |
| Garage | .192 | 78.3% |
| Numbaños2 | .000 | .2% |
| Garage2 | .015 | 6.3% |
| Garage3 | .023 | 9.3% |
| Nivelrentabaja | .049 | 19.9% |
| Nivelrentaalta | .018 | 7.1% |
| Numbaños4 | .023 | 9.3% |
| Numbaños4.5 | .246 | 100.0% |
| Numbaños5 | .039 | 15.7% |
| Numdormitorio | .013 | 5.4% |
| Numniveles3 | .031 | 12.5% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenido los siguientes resultados.

Tabla 5.22. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Training(entre namiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef.de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|--------------------------|--------|-----------|---------------------|----------------------|-----------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Tangente hiperbólica | 667 | 215 | 79 | 10 minutos | 86% | 75% | 74% | 515606 | 515021 |
| Tangente hiperbólica | Tangente hiperbólica | 667 | 205 | 89 | 20 minutos | 87% | 76% | 76% | 454621 | 454099 |
| Tangente hiperbólica | Tangente hiperbólica | 651 | 229 | 81 | 30 minutos | 87% | 76% | 75% | 478003 | 477459 |
| Tangente hiperbólica | Tangente hiperbólica | 669 | 227 | 65 | 40 minutos | 83% | 68% | 68% | 574366 | 573724 |
| Tangente hiperbólica | Tangente hiperbólica | 660 | 219 | 82 | 50 minutos | 89% | 79% | 78% | 462812 | 462286 |
| Tangente hiperbólica | Tangente hiperbólica | 677 | 201 | 83 | 60 minutos | 85% | 72% | 71% | 501615 | 501043 |
| Tangente hiperbólica | Tangente hiperbólica | 671 | 223 | 67 | 80 minutos | 91% | 83% | 82% | 417265 | 416798 |
| Tangente hiperbólica | Tangente hiperbólica | 676 | 215 | 70 | 100 minutos | 86% | 74% | 74% | 481281 | 480740 |

Fuente: Elaboración propia

La tabla 5.22 contiene los resultados del entrenamiento de la red con un factor de activación en la capa oculta, una función de entrada tangente hiperbólica y un factor de activación en la capa de salida con una función tangente hiperbólica.

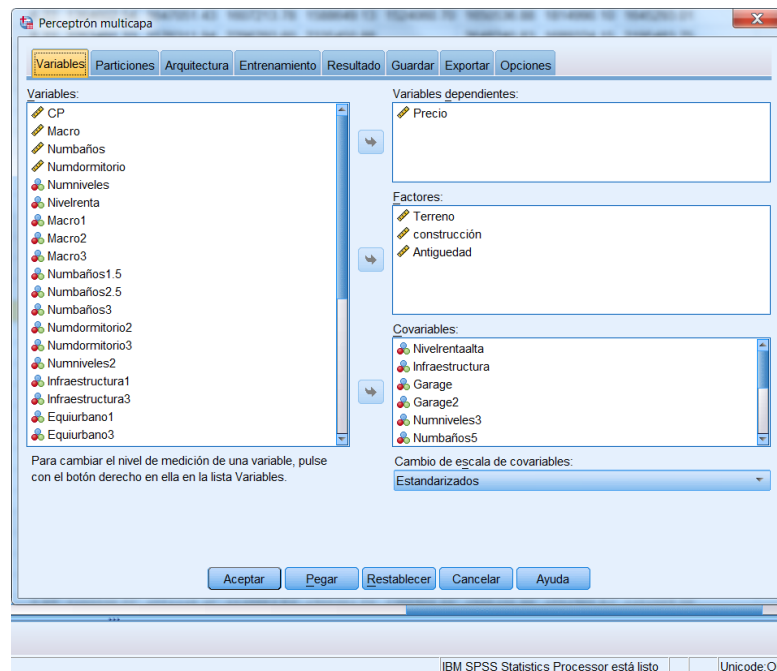
Se puede concluir en este modelo que de los ocho resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el *test* de entrenamiento con tiempo de 80 minutos es el que mejor ha estimado el precio: arroja un coeficiente de correlación de 91%, un coeficiente de determinación de 83%, y un r^2 ajustado de 82%.

Nuevamente se utilizan los mismos parámetros del

modelo anterior, ahora una red de activación en sus entradas oculta con una función tangente hiperbólica y una función de activación con salida sigmoïdal.

En la siguiente figura se muestra la informaci3n de las variables a utilizar en el modelo, mediante el programa SPSS.

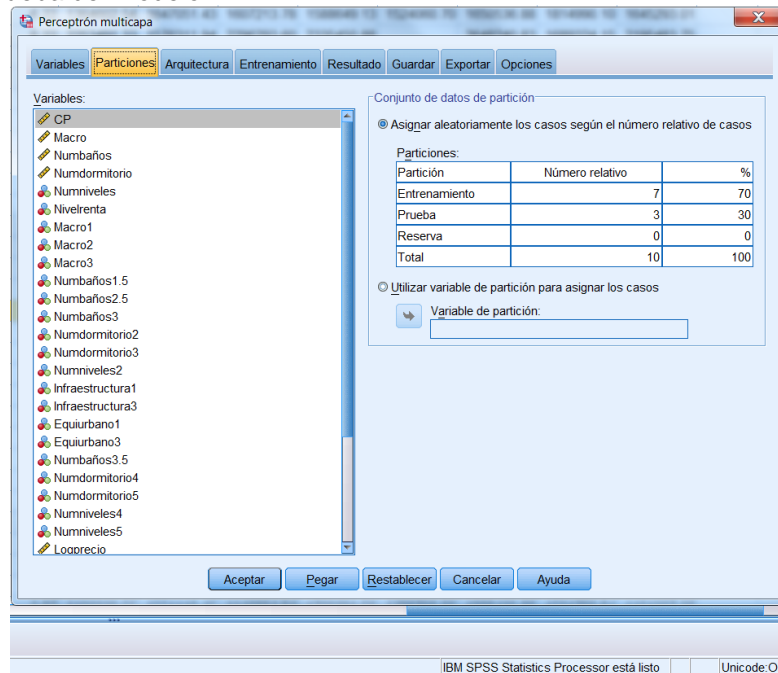
Figura 5.21. Variables a utilizar en el modelo RNA



Fuente: Elaboraci3n propia

En la siguiente figura se observan el tipo de patrones de entrada a la red:

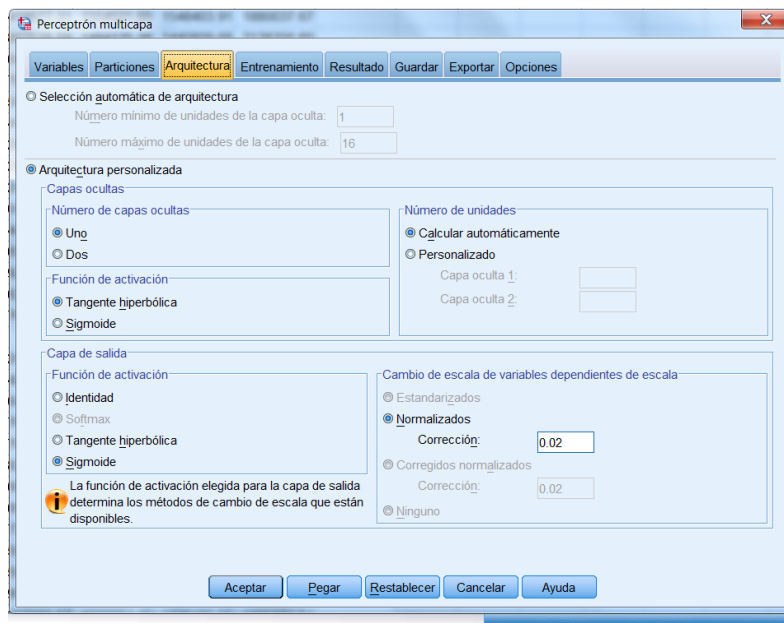
Figura 5.22. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura 5.23 contiene información acerca de la arquitectura de la red.

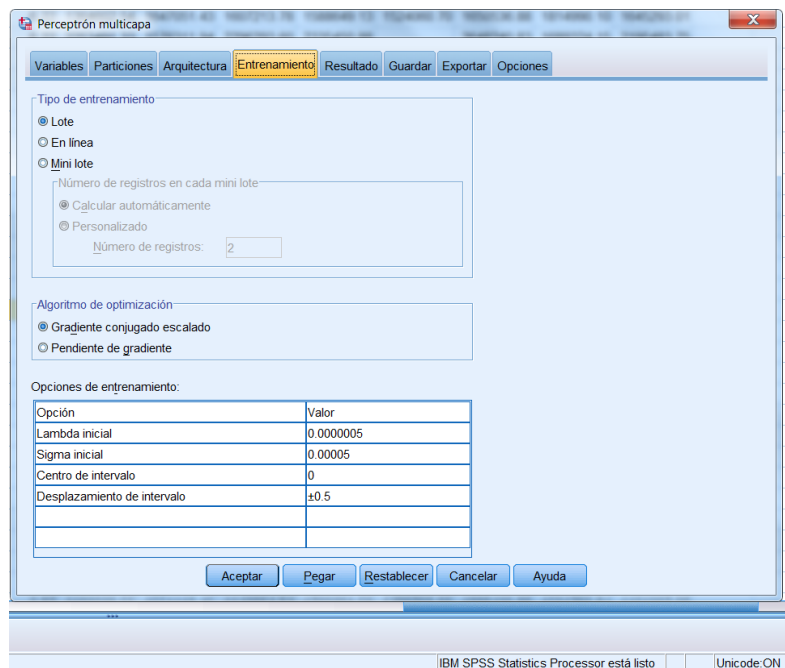
Figura 5.23. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.24 se observa el tipo de entrenamiento denominado en lote, con sus diferentes opciones de entrenamiento

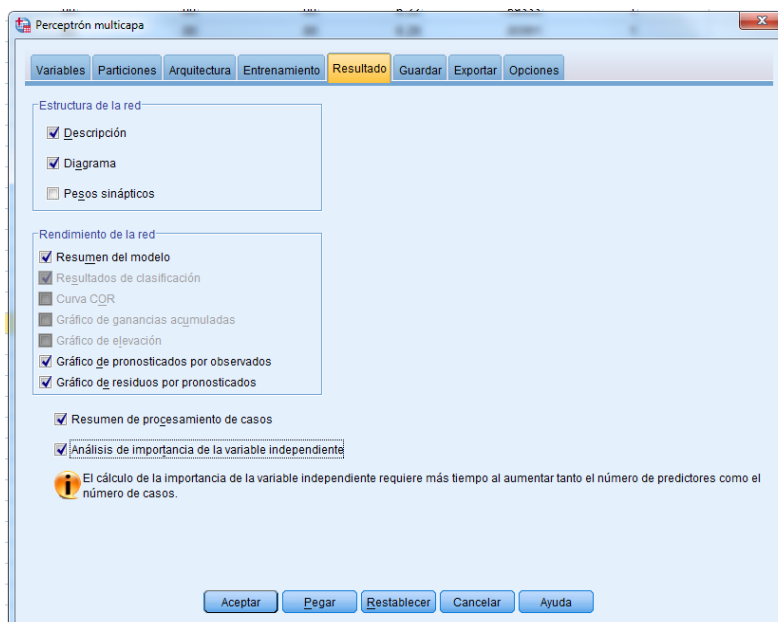
Figura 5.24. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la imagen de la figura 5.25 se muestran los resultados de la red: descripción, diagramas, resumen del modelo, gráficos de pronósticos por observados, gráficos de residuos por pronósticos, resumen de procesamientos de casos y análisis de importancia de la variable independiente.

Figura 5.25. Resultado de la RNA



Fuente: Elaboración propia

Una vez ejecutado el entrenamiento de la red, se han obtenido los resultados que en las tablas siguientes se detallan.

Tabla 5.23. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 654 | 74.9% |
| | Pruebas | 219 | 25.1% |
| Válido | | 873 | 100.0% |
| Excluido | | 88 | |
| Total | | 961 | |

Fuente: Elaboración propia

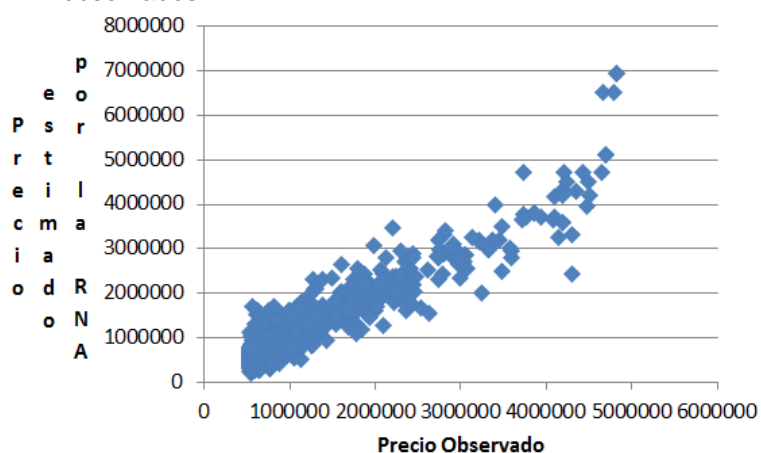
Tabla 5.24. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | .579 |
| | Error relativo | .089 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:00.43 |
| Pruebas | Error de suma de cuadrados | .423 |
| | Error relativo | .172 |

Fuente: Elaboración propia.

En la tabla 5.24 se muestran la suma de errores de 0.579, un error relativo de 0.089.

Figura 5.26. Gráfica de precios pronosticados y precios observados

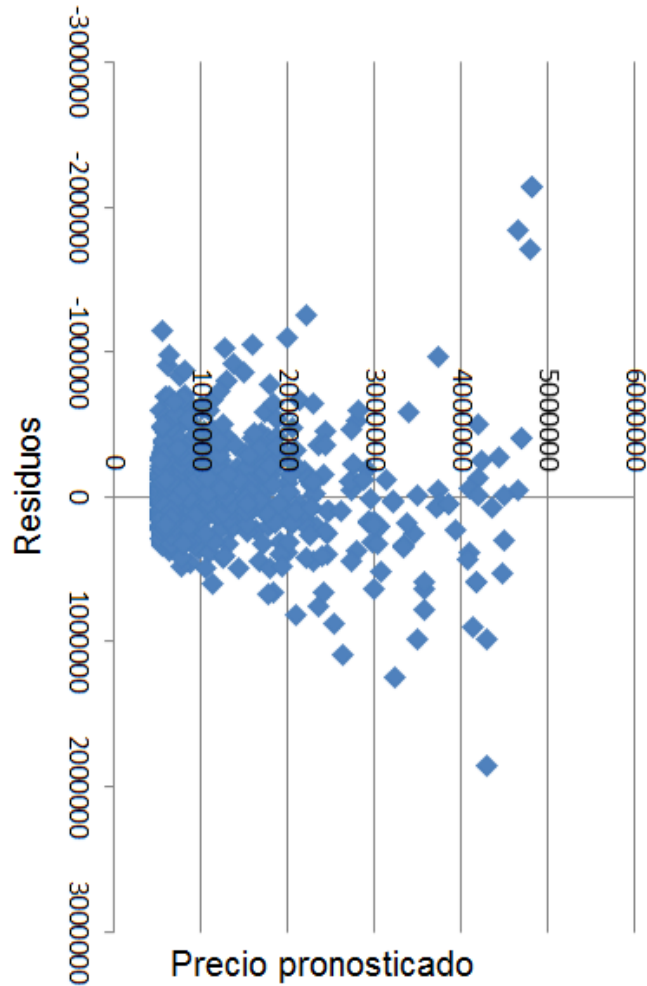


Fuente: Elaboración propia.

La gráfica de la figura 5.26 representa en el eje X el precio de los inmuebles observados, y en el eje de la Y los precios pronosticados por el modelo.

Continuando en la figura 5.27, se observa lo siguiente.

Figura 5.27. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura 5.27 representa un gráfico de los residuos contra el valor pronosticado.

En la tabla siguiente se observa la importancia de

las variables independientes.

Tabla 5.28. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .080 | 30.6% |
| construcción | .092 | 34.8% |
| Antigüedad | .056 | 21.4% |
| Nivelrentaalta | .007 | 2.5% |
| Infraestructura | .015 | 5.8% |
| Garage | .106 | 40.4% |
| Garage2 | .004 | 1.4% |
| Numniveles3 | .043 | 16.5% |
| Numbaños5 | .257 | 97.6% |
| Numbaños4 | .012 | 4.7% |
| Nivelrentabaja | .038 | 14.5% |
| Numdormitorio | .000 | .1% |
| Numbaños2 | .005 | 2.0% |
| Garage3 | .002 | .6% |
| Numbaños4.5 | .263 | 100.0% |
| Equiurbano | .019 | 7.1% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se llegó a los siguientes resultados.

Tabla 5.29. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Traning(entrenamiento) | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|------------------------|---------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Sigmoidea | 665 | 208 | 88 | 10 minutos | 90.79% | 82.43% | 82.10% | 395307 | 394853 |
| Tangente hiperbólica | Sigmoidea | 688 | 197 | 76 | 20 minutos | 84.07% | 70.67% | 70.13% | 550456 | 549834 |
| Tangente hiperbólica | Sigmoidea | 666 | 220 | 75 | 30 minutos | 75.65% | 57.23% | 56.44% | 594329 | 593658 |
| Tangente hiperbólica | Sigmoidea | 692 | 212 | 57 | 40 minutos | 83.53% | 69.78% | 69.23% | 536594 | 536000 |
| Tangente hiperbólica | Sigmoidea | 654 | 219 | 88 | 50 minutos | 94.19% | 88.71% | 88.50% | 322748 | 322378 |
| Tangente hiperbólica | Sigmoidea | 668 | 211 | 82 | 60 minutos | 90.97% | 82.76% | 82.44% | 426072 | 425587 |
| Tangente hiperbólica | Sigmoidea | 669 | 222 | 70 | 80 minutos | 86.60% | 75.00% | 74.54% | 490805 | 490254 |

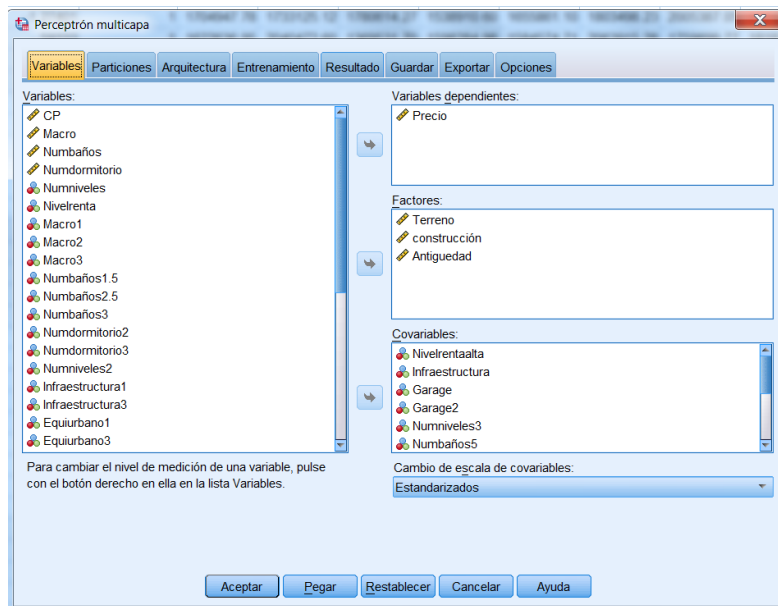
Fuente: Elaboración propia.

La tabla 5.29 contiene los resultados del entrenamiento de la red con un factor de activación en la capa oculta, una función de entrada tangente hiperbólica y un factor de activación en la capa de salida con una función sigmoidea (sigmoidea).

Se puede concluir en este modelo que de los ocho resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el test de entrenamiento con tiempo de 50 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 94%, coeficiente de determinación de 89%, y r^2 ajustado de 89%.

En la siguiente figura se detalla la información de las variables a utilizar en el modelo, mediante el programa SPSS.

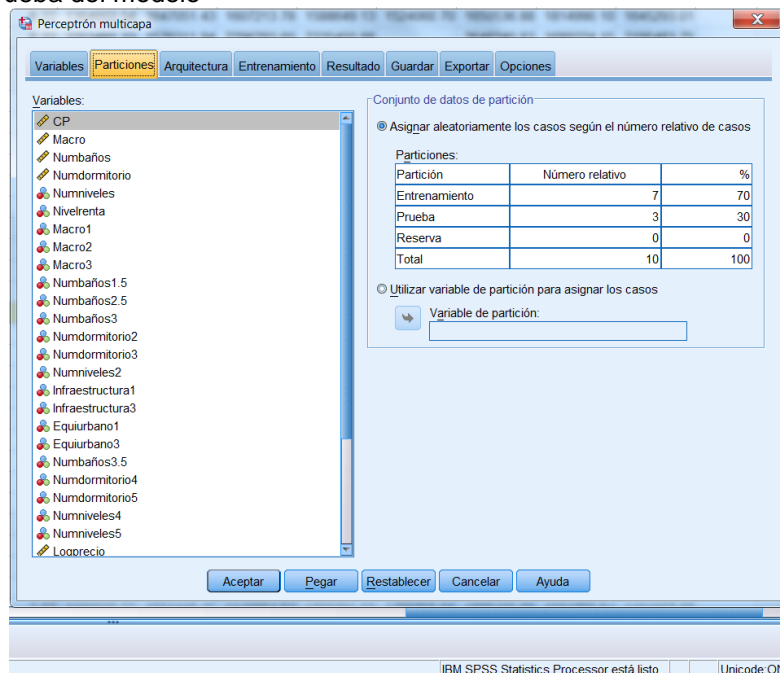
Figura 5.28. Variables a utilizar en el modelo RNA



Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red.

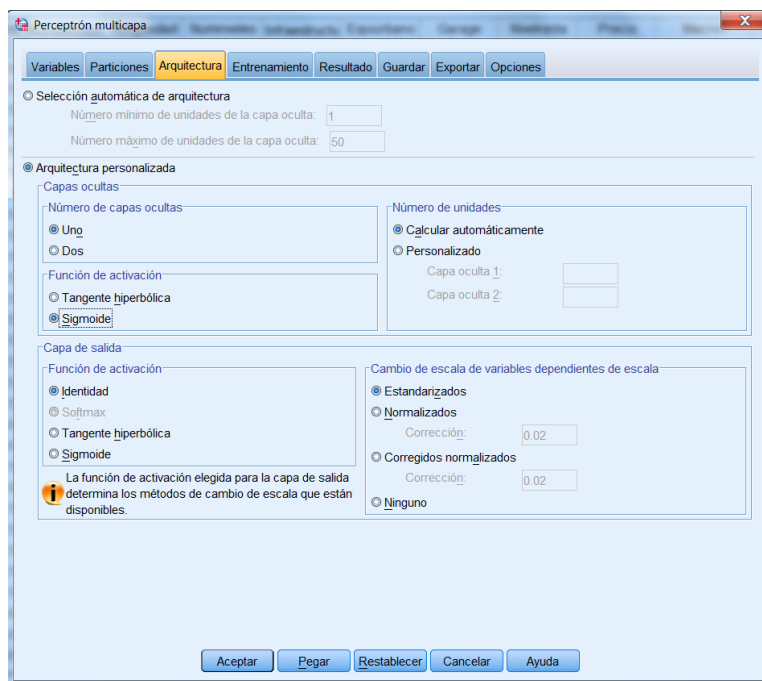
Figura. 5.29. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura [5.30] contiene información sobre la arquitectura de la red: una capa oculta con un máximo de 16 unidades o patrones en la misma, con una función de activación sigmoideal a la entrada y una función de salida de la red denominada identidad (lineal).

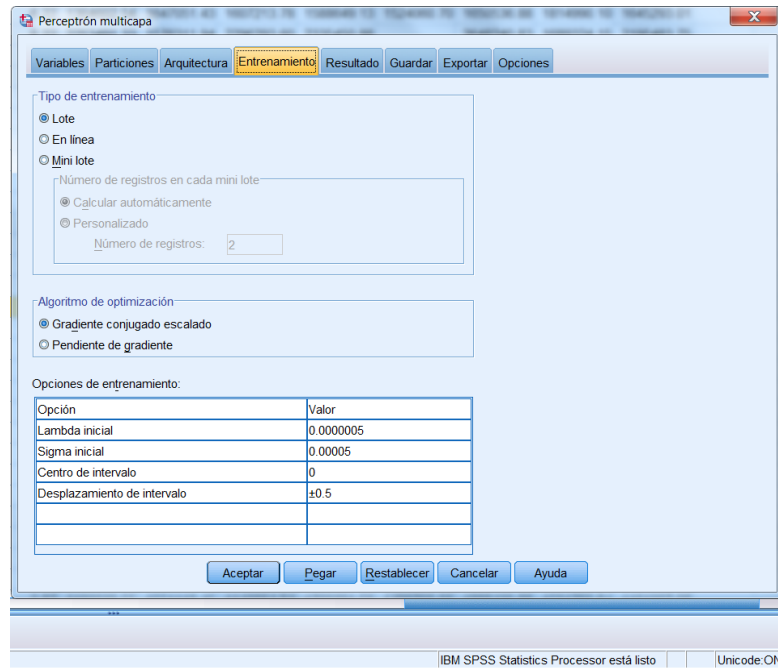
Figura 5.30. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.31 se observa el tipo de entrenamiento denominado en lote con sus diferentes opciones de entrenamiento.

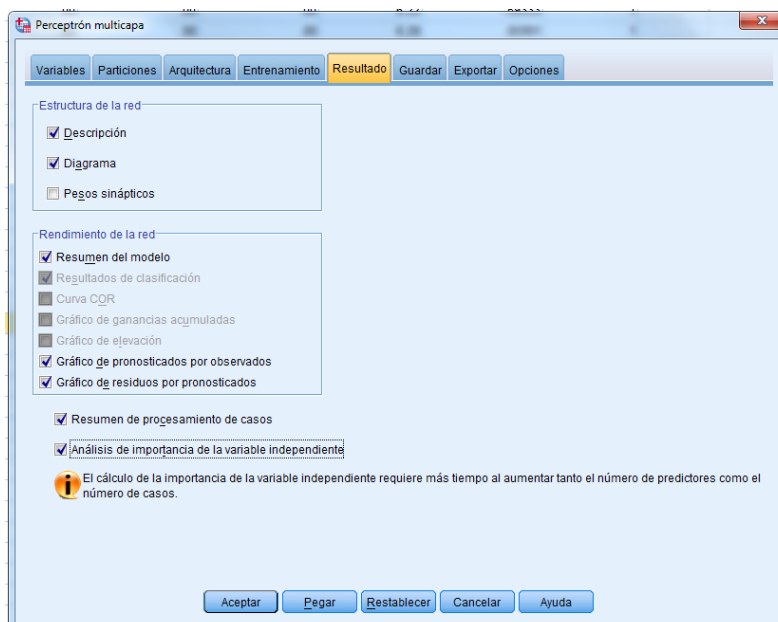
Figura 5.31. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.32 se aprecian los resultados de la red.

Figura 5.32. Resultado de la RNA



Fuente: Elaboración propia

Una vez ejecutado el entrenamiento de la red, se han obtenido los siguientes resultados que se detallan en las tablas.

Tabla 5.30. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 677 | 76.2% |
| | Pruebas | 211 | 23.8% |
| Válido | | 888 | 100.0% |
| Excluido | | 73 | |
| Total | | 961 | |

Fuente: Elaboración propia

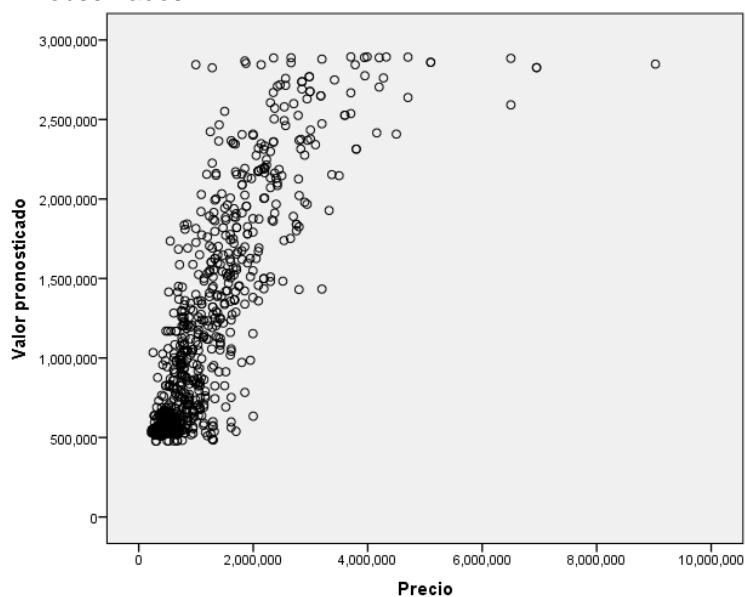
Tabla 5.31. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 1.416 |
| | Error relativo | .329 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:00.78 |
| Pruebas | Error de suma de cuadrados | .329 |
| | Error relativo | .320 |

Fuente: Elaboración propia

En la tabla 5.31 se muestran la suma de errores del 1.416, y un error relativo de 0.329.

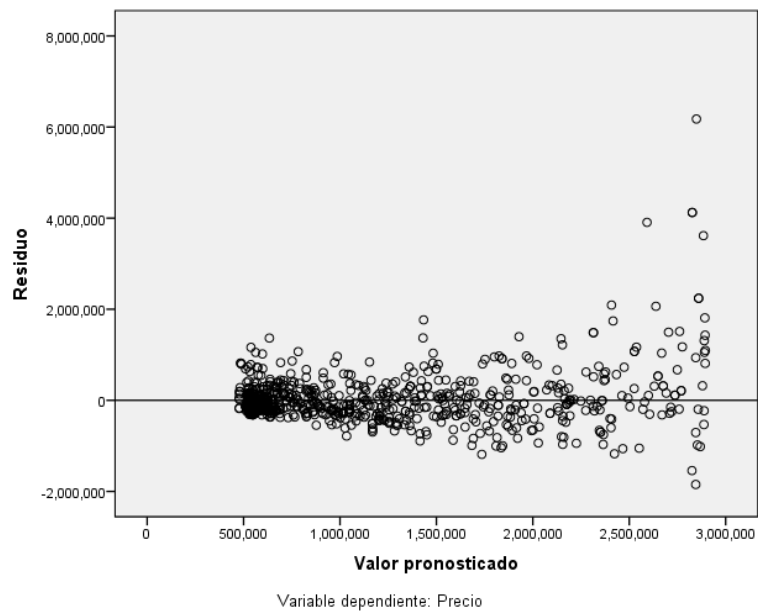
Figura 5.33. Gráfica de precios pronosticados y precios observados



Fuente: Elaboración propia

Continuando en la figura [5.34], se observa lo siguiente:

Figura 5.34. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura [5.34] representa un gráfico de los residuos contra los valores pronosticados.

En la tabla siguiente se observa la importancia de las variables independientes.

Tabla 5.32. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .061 | 31.5% |
| construcción | .045 | 23.0% |
| Antigüedad | .027 | 14.1% |
| Nivelrentaalta | .038 | 19.3% |
| Infraestructura | .021 | 11.0% |
| Garage | .034 | 17.6% |
| Garage2 | .030 | 15.3% |
| Numniveles3 | .077 | 39.4% |
| Numbaños5 | .121 | 62.1% |
| Numbaños4 | .163 | 83.4% |
| Nivelrentabaja | .072 | 37.0% |
| Numdormitorio | .020 | 10.3% |
| Numbaños2 | .002 | 1.2% |
| Garage3 | .082 | 42.1% |
| Numbaños4.5 | .195 | 100.0% |
| Equiurbano | .011 | 5.4% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenidos los siguientes resultados.

Tabla 5.33. *Test* de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Traning(entr enamiento) | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|-------------------------|------------|-----------|---------------------|----------------------|-----------------------|-------------------------|------------------------------|------------------------|
| Sigmoidea | Identidad | 680 | 210 | 71 | 10 minutos | 96% | 92% | 92% | 269756 | 269447 |
| Sigmoidea | Identidad | 677 | 211 | 73 | 20 minutos | 91% | 83% | 83% | 442795 | 442294 |
| Sigmoidea | Identidad | 671 | 202 | 88 | 30 minutos | 80% | 63% | 63% | 545783 | 545167 |
| Sigmoidea | Identidad | 693 | 201 | 67 | 40 minutos | 88% | 77% | 77% | 481722 | 481189 |
| Sigmoidea | Identidad | 677 | 211 | 73 | 50 minutos | 97% | 94% | 94% | 234693 | 234424 |
| Sigmoidea | Identidad | 662 | 221 | 78 | 60 minutos | 83% | 68% | 68% | 564525 | 563883 |
| Sigmoidea | Identidad | 698 | 189 | 74 | 80 minutos | 90% | 81% | 80% | 437670 | 437179 |
| Sigmoidea | Identidad | 691 | 203 | 67 | 100 minutos | 90% | 80% | 80% | 432360 | 431864 |

Fuente: Elaboración propia

La tabla 5.33 contiene los resultados del entrenamiento de la red con función de activación en la capa oculta denominada sigmoidea y una función de activación en la capa de salida denominada identidad (lineal).

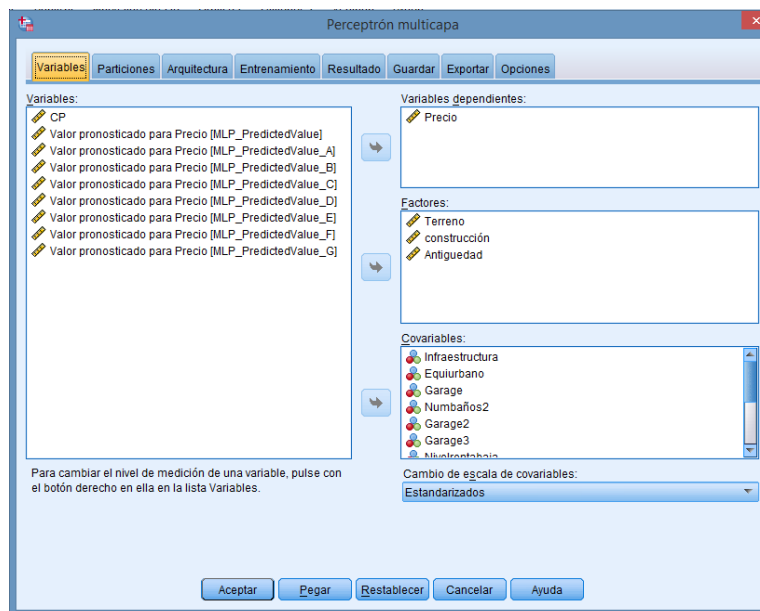
Se puede concluir en este modelo que de los 10 resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el *test* de entrenamiento con tiempo de 50 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 97%, coeficiente de determinación de 94%, un r^2 ajustado de 94%.

Continuando con el modelo de RNA, utilizando los mismos parámetros del modelo anterior, ahora con una red de activación en sus entradas, una capa oculta que utiliza una función sigmoidea (sigmoidea) y una función de

activación con salida tangente hiperbólica.

En la siguiente figura está la información de las variables a utilizar en el modelo, mediante el programa SPSS.

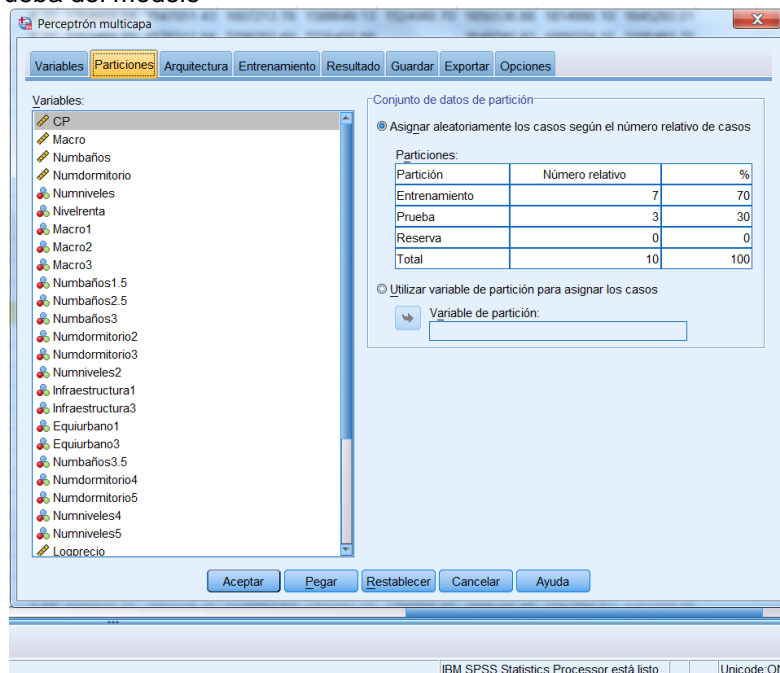
Figura 5.35. Variables a utilizar en el modelo de la RNA.



Fuente: Elaboración propia.

En la siguiente figura se observan el tipo de patrones de entrada a la red.

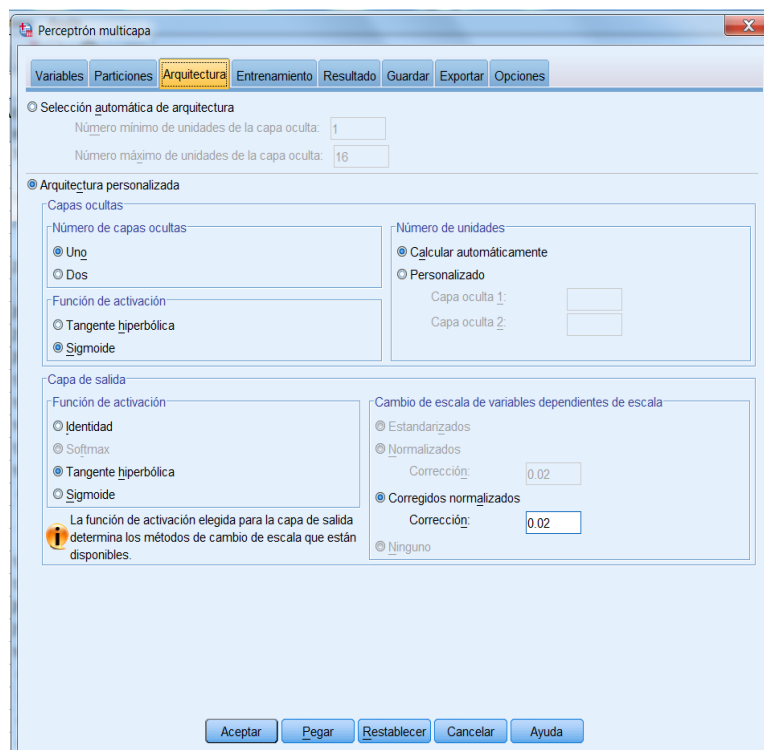
Figura. 5.36. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura 5.36 contiene información acerca de la arquitectura de la red. Se seleccionó una capa oculta con un máximo de 16 unidades o patrones en la misma, con una función de activación sigmoïdal en la entrada y una función de salida de la red denominada tangente hiperbólica.

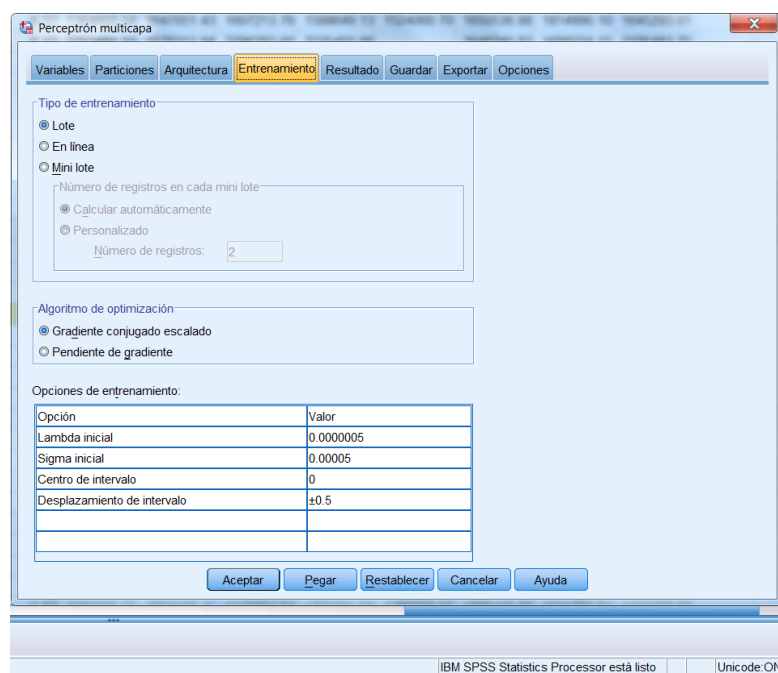
Figura 5.37. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.38 se observa el tipo de entrenamiento denominado en lote, con sus diferentes opciones de entrenamiento.

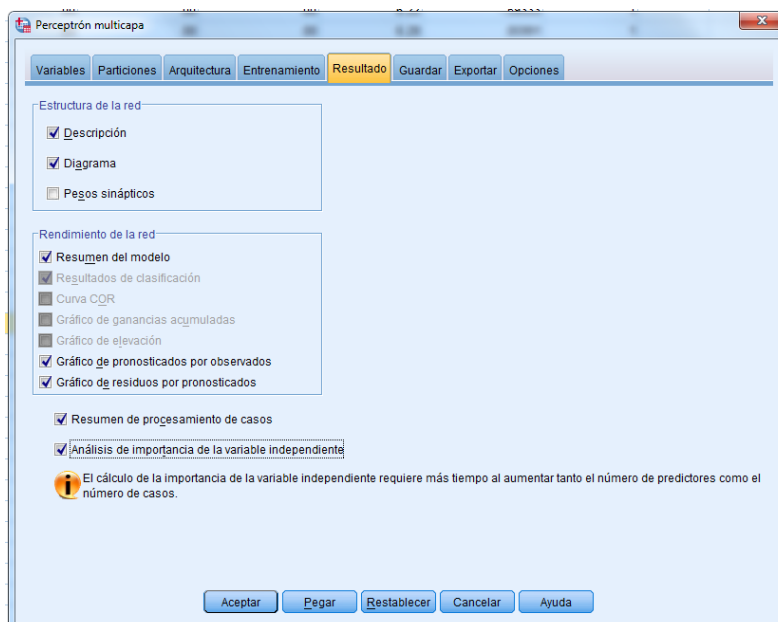
Figura 5.38. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.39 se presentan los resultados de la red.

Figura 5.39. Resultado de la RNA



Fuente: Elaboración propia

De nuevo, una vez que se ejecutó el entrenamiento de la red, se han obtenido los siguientes resultados.

La tabla 5.34. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 660 | 75.1% |
| | Pruebas | 219 | 24.9% |
| Válido | | 879 | 100.0% |
| Excluido | | 82 | |
| Total | | 961 | |

Fuente: Elaboración propia

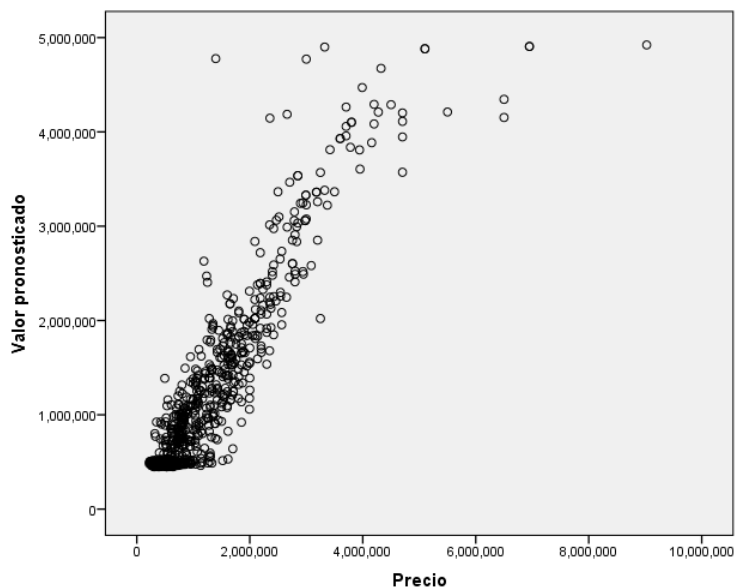
Tabla 5.31. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 1.416 |
| | Error relativo | .329 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:00.78 |
| Pruebas | Error de suma de cuadrados | .329 |
| | Error relativo | .320 |

Fuente: Elaboración propia

En la tabla 5.34 se muestra la suma de errores de 1.416, y un error relativo de 0.329.

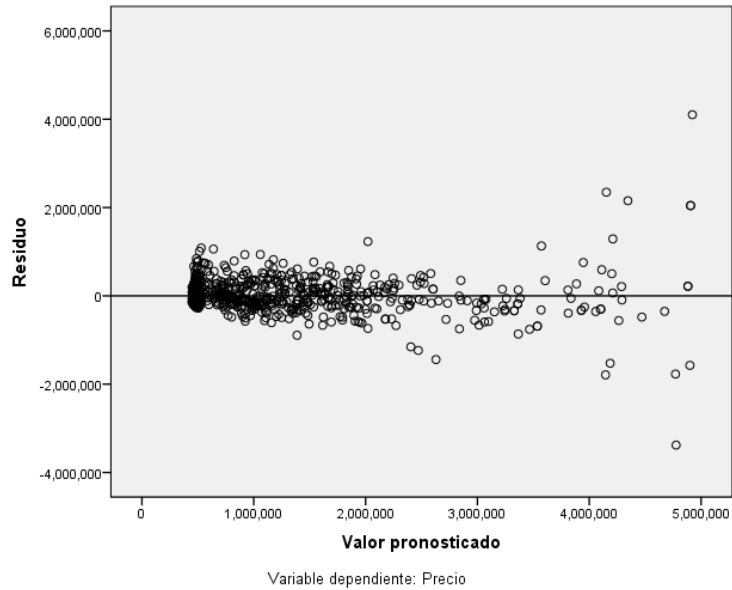
Figura 5.40. Gráfica de precios pronosticados y precios observados



Fuente: Elaboración propia

Continuando con la figura 5.34, se observa lo siguiente:

Figura 5.41. Gráfica de residuos y valor pronosticado



Fuente: Elaboración propia

La figura [5.41] representa un gráfico de los residuos contra el valor pronosticado.

En la tabla siguiente, se observa la importancia de las variables independientes.

Tabla 5.42. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .050 | 28.9% |
| construcción | .054 | 31.2% |
| Antigüedad | .052 | 29.7% |
| Infraestructura | .035 | 19.9% |
| Equiurbano | .005 | 3.1% |
| Garage | .174 | 100.0% |
| Numbaños2 | .010 | 5.9% |
| Garage2 | .003 | 1.9% |
| Garage3 | .039 | 22.5% |
| Nivelrentabaja | .073 | 41.6% |
| Nivelrentaalta | .057 | 32.7% |
| Numbaños4 | .061 | 34.8% |
| Numbaños4.5 | .174 | 99.6% |
| Numbaños5 | .121 | 69.5% |
| Numdormitorio6 | .017 | 10.0% |
| Numniveles3 | .073 | 41.7% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenidos los siguientes resultados.

Tabla 5.43. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación de salida | Traning(entr enamiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|-----------------------------|-------------------------|------------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Sigmoidea | Tangente hiperbólica | 667 | 215 | 79 | 10 minutos | 88% | 77% | 76% | 530697 | 530095 |
| Sigmoidea | Tangente hiperbólica | 667 | 205 | 89 | 20 minutos | 86% | 74% | 73% | 487788 | 487228 |
| Sigmoidea | Tangente hiperbólica | 651 | 229 | 81 | 30 minutos | 84% | 70% | 70% | 520246 | 519654 |
| Sigmoidea | Tangente hiperbólica | 669 | 227 | 65 | 40 minutos | 86% | 75% | 74% | 505593 | 505028 |
| Sigmoidea | Tangente hiperbólica | 660 | 219 | 82 | 50 minutos | 92% | 86% | 85% | 390055 | 389611 |
| Sigmoidea | Tangente hiperbólica | 677 | 201 | 83 | 60 minutos | 87% | 76% | 75% | 455917 | 455397 |
| Sigmoidea | Tangente hiperbólica | 671 | 223 | 67 | 80 minutos | 91% | 83% | 83% | 415859 | 415394 |
| Sigmoidea | Tangente hiperbólica | 676 | 215 | 70 | 100 minutos | 85% | 73% | 72% | 501195 | 500632 |

Fuente: Elaboración propia

La tabla [5.43] contiene los resultados del entrenamiento de la red con la función de activación en la capa oculta denominada sigmoidea y una función de activación en la capa de salida denominada tangente hiperbólica.

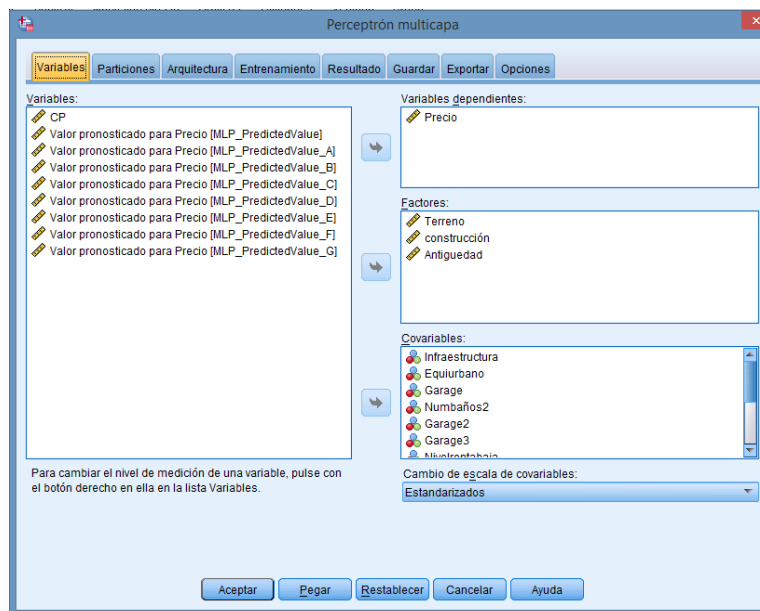
Se puede concluir en este modelo que de los 10 resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el test de entrenamiento con tiempo de 50 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 92%, coeficiente de determinación de 86%, y un r^2 ajustado de 85%.

En el siguiente modelo de RNA utiliza los mismos parámetros del modelo anterior, pero ahora con una red de activación en sus entradas, una capa oculta que usa

una función sigmoideal (sigmoidea) y una función de activación con salida sigmoideal (sigmoidea).

La siguiente figura contiene información de las variables a utilizar en el modelo, mediante el programa SPSS.

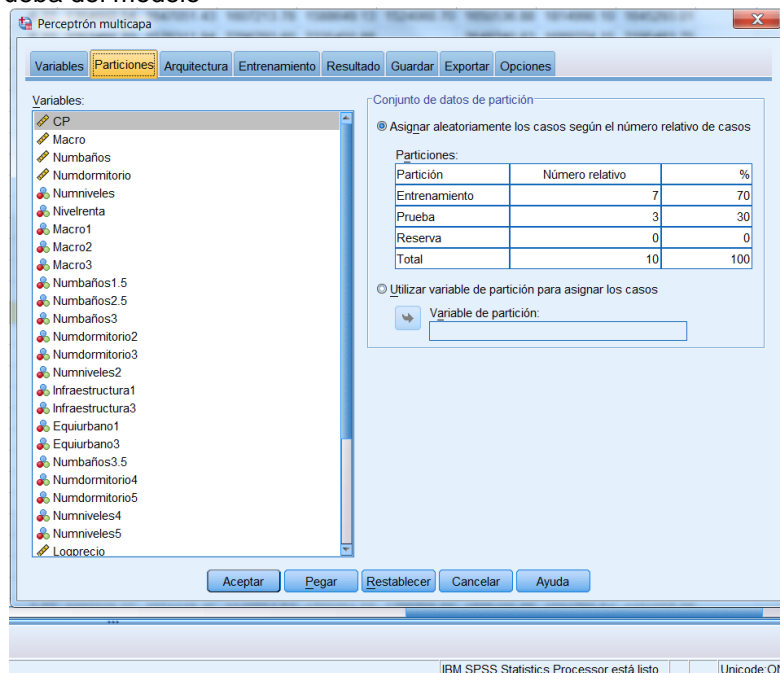
Figura 5.42. Variables a utilizar en el modelo de la RNA



Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red.

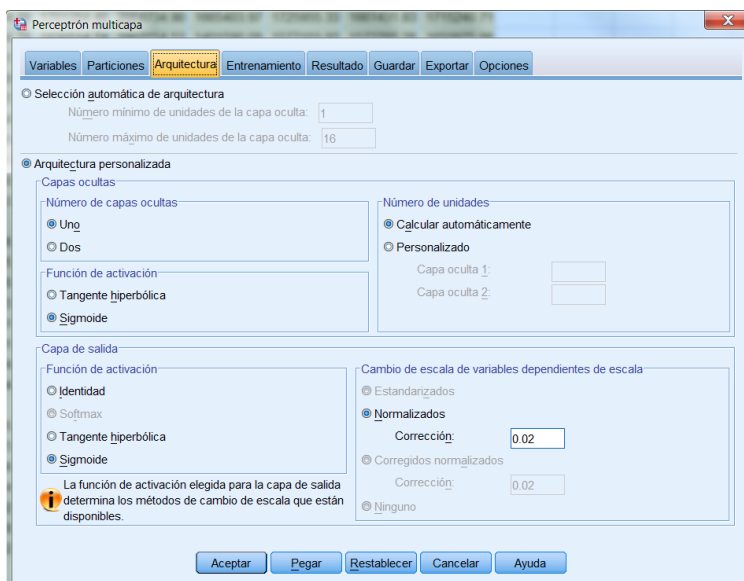
Figura. 5.43. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura 5.43 contiene información acerca de la arquitectura de la red. Se ha seleccionado una capa oculta con un máximo de 16 unidades o patrones en la misma, con una función de activación sigmoidea a la entrada y una función de salida de la red denominada sigmoidea (sigmoidea).

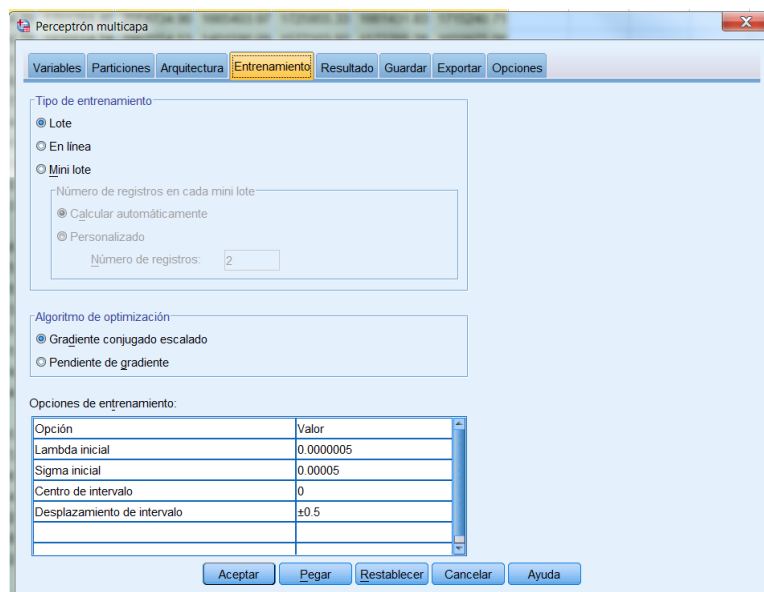
Figura 5.44. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura [5.45] se observa el tipo de entrenamiento denominado en lote, con sus diferentes opciones de entrenamiento.

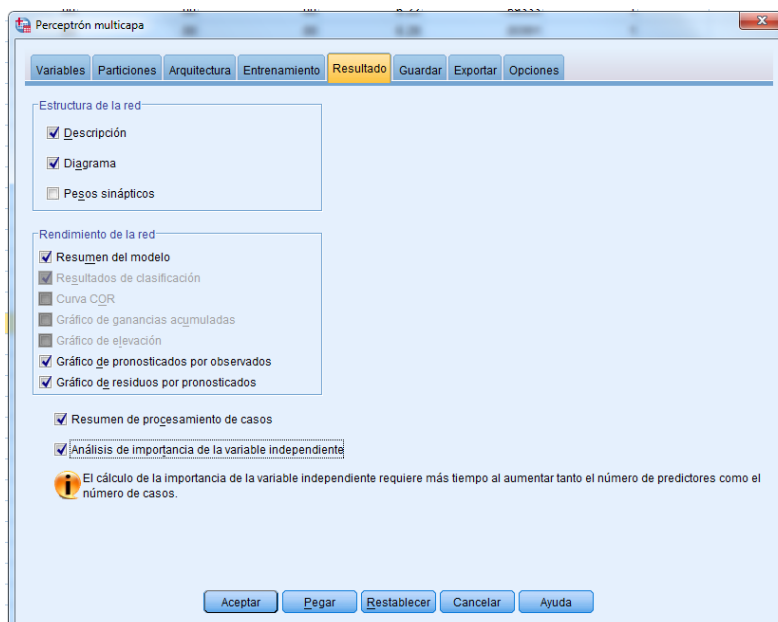
Figura 5.45. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.46 están los resultados de la red: diagramas.

Figura 5.46. Resultado de la RNA



Fuente: Elaboración propia

Nuevamente una vez que se ejecutó el entrenamiento de la red, se han obtenido los resultados, mismos que se detallan en la siguiente tabla.

Tabla 5.44. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 671 | 75.1% |
| | Pruebas | 223 | 24.9% |
| Válido | | 894 | 100.0% |
| Excluido | | 67 | |
| Total | | 961 | |

Fuente: Elaboración propia.

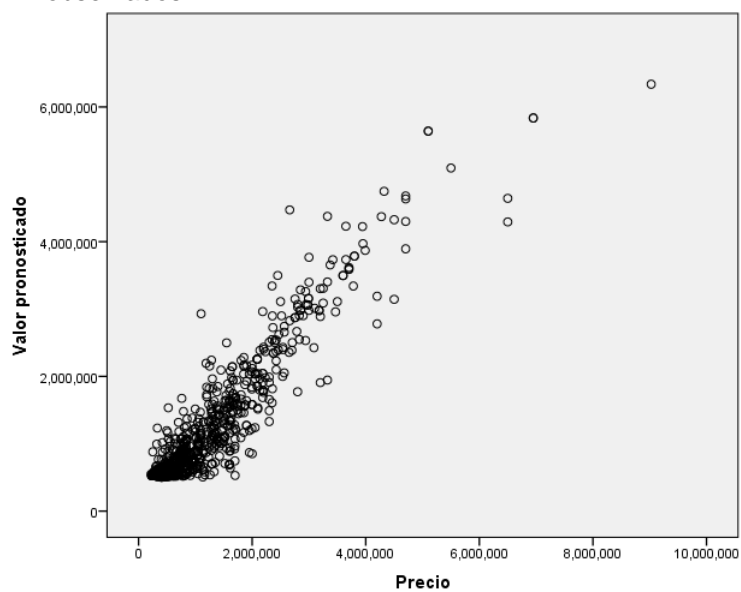
Tabla 5.45. Errores de predicción de la RNA

| | | |
|-------------------|---------------------------------|---|
| Entrenami ento | Error de | .392 |
| | Error | .091 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de | 0:00:00.49 |
| Pruebas | Error de | .328 |
| | Error | .209 |

Fuente: Elaboración propia.

En la tabla 5.45 se muestran la suma de errores de 0.392, y un error relativo de .09.

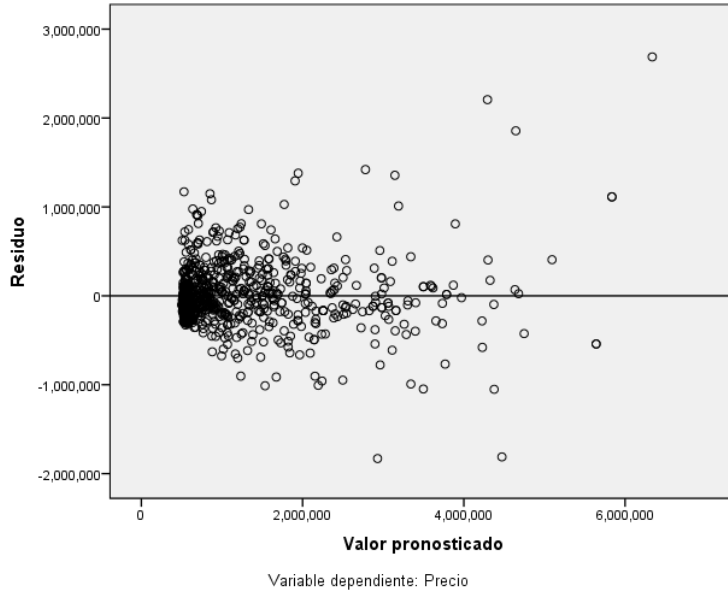
Figura 5.46. Gráfica de precios pronosticados y precios observados



Fuente: Elaboración propia

Continuando en la figura 5.47, se distingue lo siguiente.

Figura 5.47. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura 5.47 representa un gráfico de los residuos contra los valores pronosticado.

En la tabla siguiente se aprecia la importancia de las variables independientes.

Tabla 5.48. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .110 | 49.7% |
| construcción | .144 | 65.0% |
| Antigüedad | .078 | 35.2% |
| Infraestructura | .003 | 1.2% |
| Equiurbano | .027 | 12.1% |
| Garage | .222 | 100.0% |
| Numbaños2 | .010 | 4.6% |
| Garage2 | .005 | 2.4% |
| Garage3 | .019 | 8.6% |
| Nivelrentabaja | .078 | 35.2% |
| Nivelrentaalta | .029 | 13.1% |
| Numbaños4 | .017 | 7.6% |
| Numbaños4.5 | .055 | 24.7% |
| Numbaños5 | .157 | 70.7% |
| Numdormitorio | .012 | 5.5% |
| Numniveles3 | .033 | 15.0% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenidos los siguientes resultados.

Tabla 5.49. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Traning(entr enamiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Erro estándar de estimación | Error cuadrático medio. |
|------------------------------|---------------------------------|-------------------------|--------|-----------|---------------------|----------------------|------------------------|-------------------------|-----------------------------|-------------------------|
| Sigmoidea | Sigmoidea | 667 | 215 | 79 | 10 minutos | 93% | 87% | 86% | 374815 | 374391 |
| Sigmoidea | Sigmoidea | 667 | 205 | 89 | 20 minutos | 93% | 86% | 85% | 351311 | 350909 |
| Sigmoidea | Sigmoidea | 651 | 229 | 81 | 30 minutos | 88% | 77% | 77% | 470785 | 470250 |
| Sigmoidea | Sigmoidea | 669 | 227 | 65 | 40 minutos | 77% | 59% | 59% | 579126 | 578480 |
| Sigmoidea | Sigmoidea | 660 | 219 | 82 | 50 minutos | 92% | 85% | 85% | 382097 | 381663 |
| Sigmoidea | Sigmoidea | 677 | 201 | 83 | 60 minutos | 89% | 79% | 79% | 418334 | 417858 |
| Sigmoidea | Sigmoidea | 671 | 223 | 67 | 80 minutos | 93% | 87% | 87% | 353758 | 353362 |
| Sigmoidea | Sigmoidea | 676 | 215 | 70 | 100 minutos | 85% | 73% | 73% | 494085 | 493531 |

Fuente: Elaboración propia

La tabla [5.49] contiene los resultados del entrenamiento de la red con la función de activación en la capa oculta denominada sigmoidea y una función de activación en la capa de salida denominada sigmoidea.

Se puede concluir en este modelo que de los 10 resultados de entrenamiento con incremento de tiempo en intervalos de 10 minutos, el *test* de entrenamiento con tiempo de 80 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 93%, un coeficiente de determinación de 87%, y un r^2 ajustado de 87%.

Se han obtenido resultados de seis modelos de la RNA con las siguientes formas de activaciones.

Función de activación tangente hiperbólica en la capa oculta con una función de activación en la capa de salida identidad (lineal); en un segundo modelo con

función de activación en la capa oculta tangente hiperbólica, función de activación en la capa de salida tangente hiperbólica; en un tercer modelo se ha utilizado una función sigmoidea en la capa oculta con una función identidad en la capa de salida; en un cuarto modelo se ha utilizado una función sigmoidea en la capa oculta y una función tangente hiperbólica en la capa de salida, y por último la misma función sigmoidea en la capa oculta con una función en la capa de salida sigmoidea. En todos los casos se utilizaron 16 variables independientes y una variable dependiente, en su arquitectura se seleccionaron intervalos que desde 1 hasta 50 unidades en su capa oculta, en todos los casos se utilizó un tipo de entrenamiento en lote, con un tipo de gradiente escalado conjugado, obteniéndose los siguientes resultados que son los mejores en cada una de sus formas de activación, véase tabla 5.50.

Tabla 5.50. Test de resultados de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Training/entrenamiento | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R2ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|------------------------|---------|-----------|---------------------|----------------------|------------------------|------------|------------------------------|------------------------|
| Tangente hiperbólica | Identidad | 667 | 205 | 89 | 20 minutos | 93% | 87% | 87% | 340906 | 340515 |
| Tangente hiperbólica | Tangente hiperbólica | 671 | 223 | 67 | 80 minutos | 91% | 83% | 82% | 417265 | 416798 |
| Tangente hiperbólica | Sigmoidea | 654 | 219 | 86 | 50 minutos | 94% | 89% | 88% | 322748 | 322378 |
| Sigmoidea | Identidad | 677 | 211 | 73 | 50 minutos | 97% | 94% | 94% | 234693 | 234424 |
| Sigmoidea | Tangente hiperbólica | 660 | 219 | 82 | 50 minutos | 92% | 86% | 85% | 390055 | 389611 |
| Sigmoidea | Sigmoidea | 671 | 223 | 67 | 80 minutos | 93% | 87% | 87% | 353758 | 353362 |

Fuente: Elaboración propia

En la tabla [5.50] se observan los resultados de las diferentes formas de activación de las redes neuronales artificiales. Los seis modelos presentan buenos resultados, sin embargo el mejor de ellos es la red con función de activación en la capa oculta sigmoidea y una función de activación en la capa de salida denominada identidad (lineal), con 677 muestras de entrenamiento, 211 para pruebas o validación y 73 excluidas por ser atípicas, tiempo de entrenamiento 50 minutos, coeficiente de correlación entre variables de 97%, coeficiente de determinación de 94%, r^2 ajustado de 94%.

Una vez validados los modelos y analizado su comportamiento en términos generales, se realiza la aplicación para estimar del precio de un inmueble en un caso concreto, es decir mediante la ecuación implícita que se ha desarrollado en la red estimada.

En todos los casos se trata de un modelo denominado perceptrón multicapa (MLP), red alimentada hacia adelante con única capa oculta, con 16 variables independientes y una dependiente. A continuación se detalla la ecuación, explicada anteriormente [4.12].¹⁰

¹⁰ Siendo en este caso $W_k = [b_k, w_{1k}, \dots, w_{jk}, \dots, w_{nk}]^T$ el vector columna de pesos y $x = [1, x_1, \dots, x_{nj}]^T$ el vector columna de entrada, esto a su vez se puede escribir de forma vectorial tal que: $y_k = F_k[w_k^T x]$.

$$S_k = \sum_{j=1}^n (W_{jk}) (X_j)$$

Para este caso concreto las funciones de activación utilizadas en la capa oculta han sido la sigmoidea y salida identidad (lineal).

En este caso se cuenta con la siguiente información de un inmueble fuera de la muestra con las siguientes características: Se trata de un inmueble que cuenta con 240 metros cuadrados de terreno, 250 metros cuadrados de construcción, 3.5 baños, 3 dormitorios, 0 años de antigüedad (nueva), 2 niveles, nivel 3 de infraestructura, nivel 3 de equipamiento, 3 garaje, 0 baños, 0 garaje2, 1 garaje3, 0 nivel de renta baja, 1 nivel de renta alta, 0 baños4, 0 baño 4.5, 0 baños5, 0 dormitorio6 y 0 numniveles3.

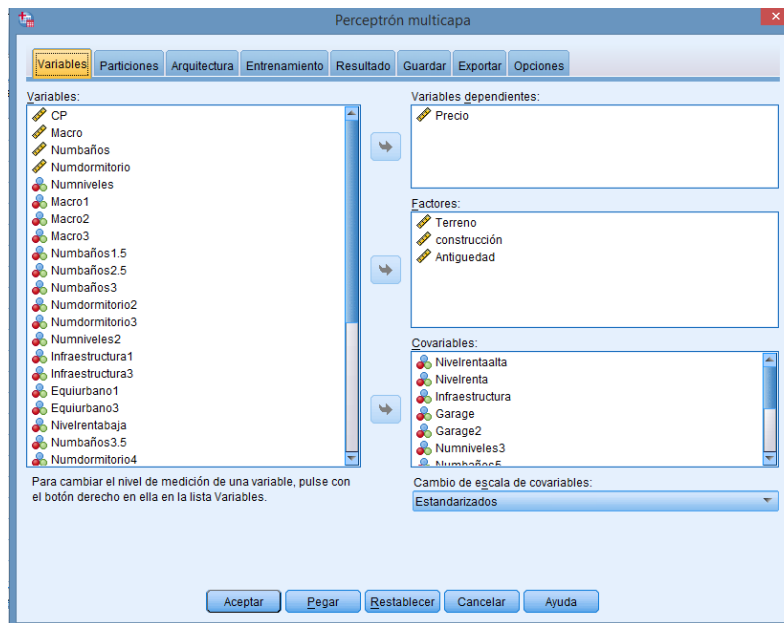
En el programa se incluirá como una muestra más, para que se estime el precio.

Reiterando que se utilizará para la prueba un modelo con una red de activación en sus entradas oculta

utilizando una función sigmoïdal (sigmoïdea) y una función de activación con salida identidad.

En la siguiente figura, se muestra la información de las variables a utilizar en el modelo, mediante el programa SPSS.

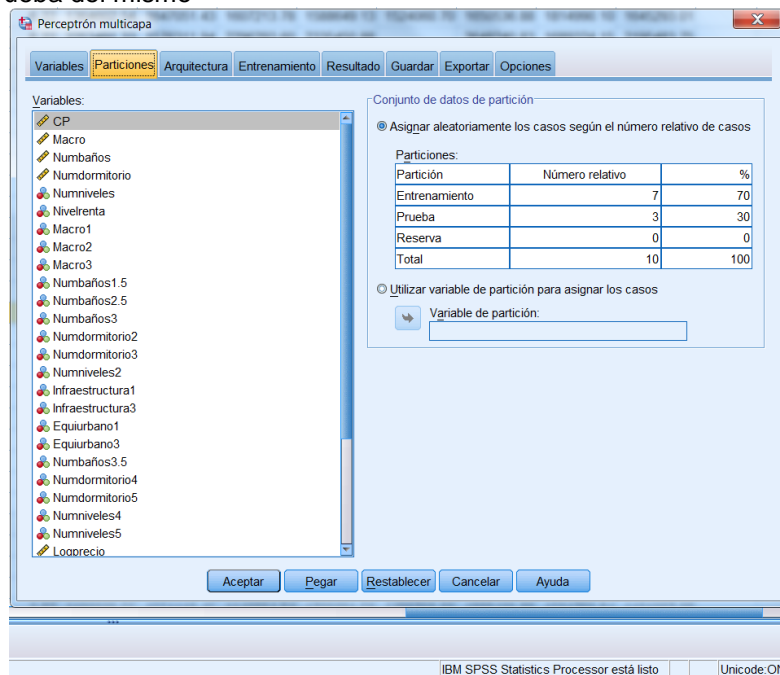
Figura 5.47 Variables a utilizar en el modelo RNA



Fuente: Elaboración propia

En la siguiente figura, se observan el tipo de patrones de entrada a la red.

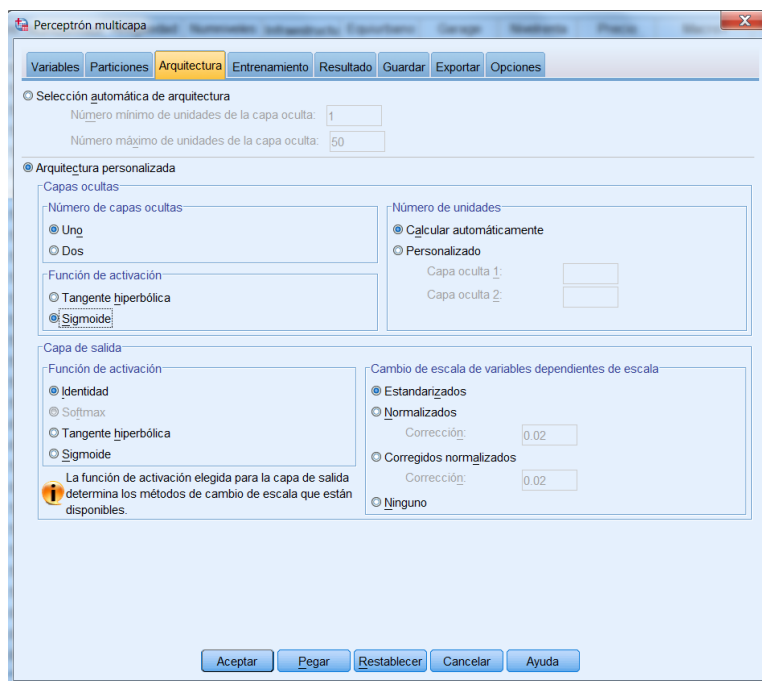
Figura. 5.48. Patrones de entrada a la red, entrenamiento y prueba del mismo



Fuente: Elaboración propia

La figura 5.49, contiene información sobre la arquitectura de la red, en este caso se ha seleccionado una capa oculta un máximo de 16 unidades o patrones en la misma, con una función de activación a la entrada sigmoideal y una función de salida de la red denominada identidad (lineal).

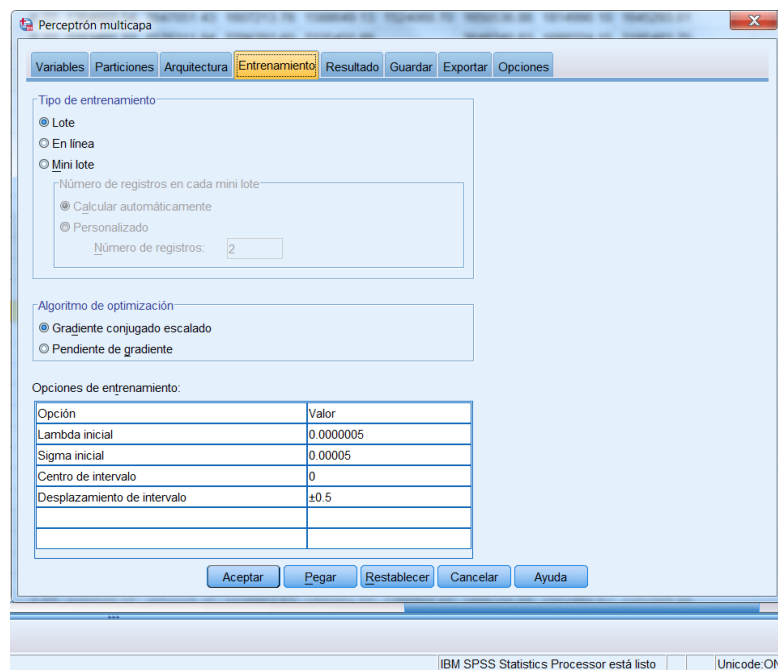
Figura 5.49. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.50 se observa el tipo de entrenamiento denominado en lote, con sus diferentes opciones de entrenamiento.

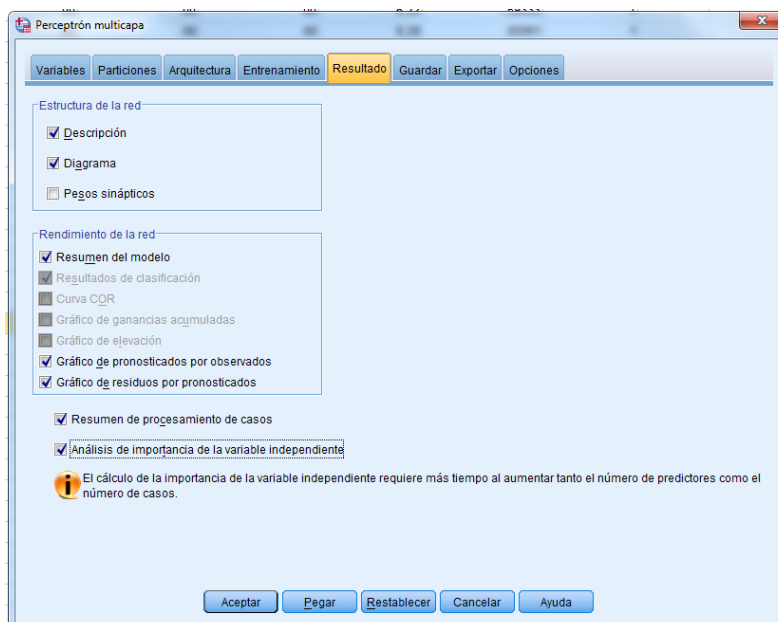
Figura 5.50. Contiene tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.51, los resultados de la red.

Figura 5.51. Resultado de la RNA



Fuente: Elaboración propia

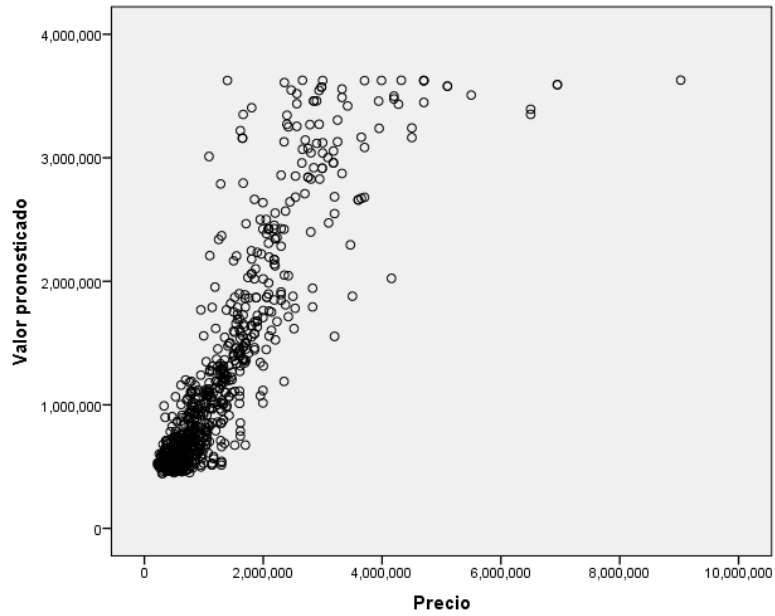
Una vez ejecutado el entrenamiento de la red como ya se ha explicado, se obtuvieron los siguientes resultados que se detallan adelante.

Tabla 5.51. Resumen de procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 663 | 75.0% |
| | Pruebas | 221 | 25.0% |
| Válido | | 884 | 100.0% |
| Excluido | | 78 | |
| Total | | 962 | |

Fuente: Elaboración propia

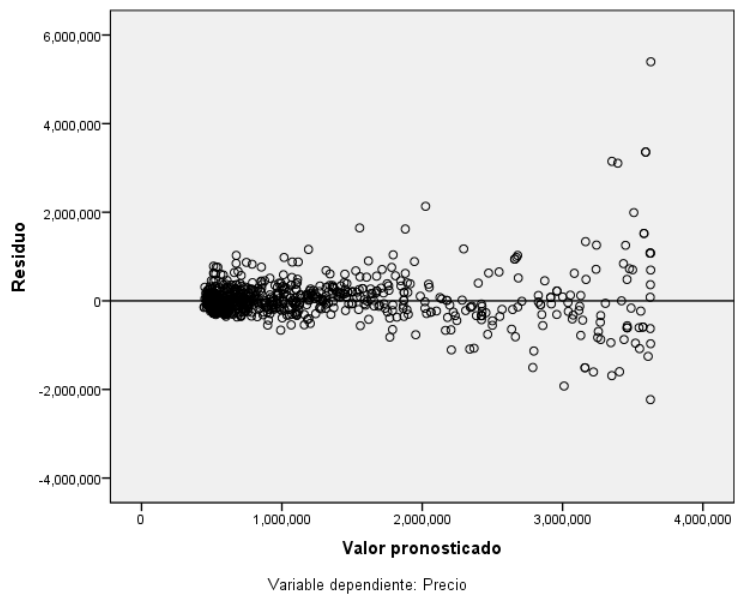
Figura 5.52 Gráfica de precios pronosticados y precios observados



Fuente: Elaboración propia

Continuando en la figura [5.53], se observa lo siguiente.

Figura 5.53. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura 5.53 representa un gráfico de los residuos contra el valor pronosticado.

En la tabla siguiente, se muestra la importancia de las variables independientes.

Tabla 5.52. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .087 | 43.2% |
| construcción | .055 | 27.3% |
| Antigüedad | .048 | 24.0% |
| Nivelrentaalta | .019 | 9.4% |
| Nivelrenta | .101 | 50.1% |
| Infraestructura | .013 | 6.4% |
| Garage | .201 | 100.0% |
| Garage2 | .026 | 12.9% |
| Numniveles3 | .038 | 19.1% |
| Numbaños5 | .192 | 95.7% |
| Numbaños4 | .005 | 2.4% |
| Numdormitorio6 | .006 | 3.1% |
| Numbaños2 | .010 | 5.0% |
| Numbaños4.5 | .166 | 82.6% |
| Garage3 | .025 | 12.4% |
| Equiurbano | .009 | 4.4% |

Fuente: Elaboración propia

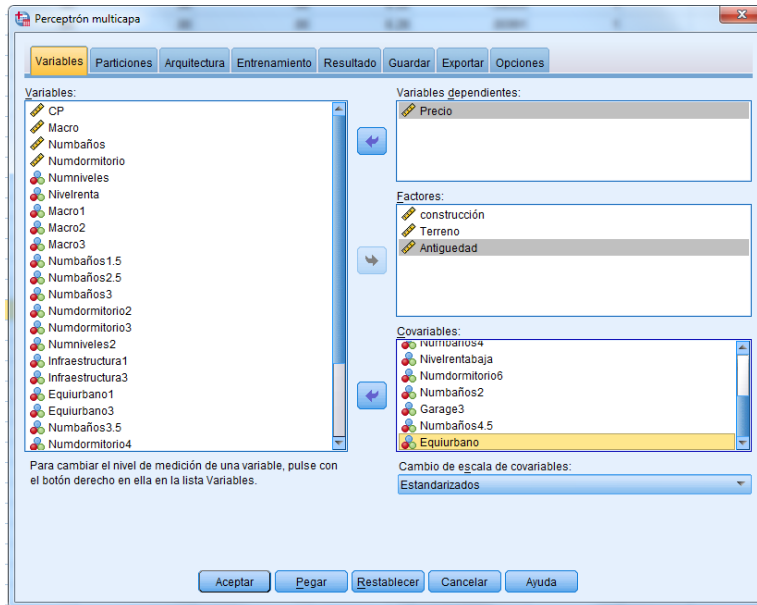
Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado: el precio observado es de \$3,182,500.00 y el estimado por la red de \$3,055,124, con un error de \$127,376 entre el precio observado y el precio estimado, resultando 4% por debajo del precio observado.

De igual forma que los seis modelos anteriores, tomando como base las ecuaciones [4.2], [4.10], [4.12] y [4.15], mediante redes neuronales artificiales, en específico el perceptrón multicapa en sus diferentes maneras de activación a la entrada de patrones e igualmente en sus salidas a la red, con las mismas observaciones y variables utilizadas en los modelos de regresión, para no dar ventaja a ninguno, se ha realizado la estimación de los precios de los inmuebles, solo que ahora se modificara la arquitectura en los seis modelos anteriores. El cálculo de las unidades en la capa oculta el programa SPSS lo realizaba automáticamente. Ahora, para los siguientes modelos, se personalizará el número de unidades en la capa oculta a 16 variables, manteniendo igual el resto de la información.

Se iniciará con el primer modelo de entrenamiento con una red con entradas de neuronas en la capa oculta, utilizando una función tangente hiperbólica y una salida identidad (lineal), con diferentes intervalos de tiempo.

La siguiente figura contiene información de las variables a utilizar en el modelo, mediante el programa SPSS.

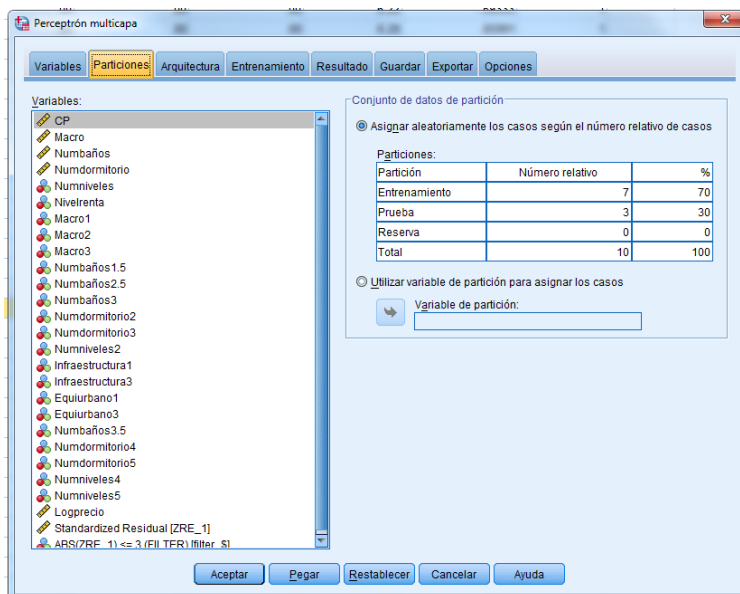
Figura 5.54. Variables a utilizar en el modelo RNA



Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red.

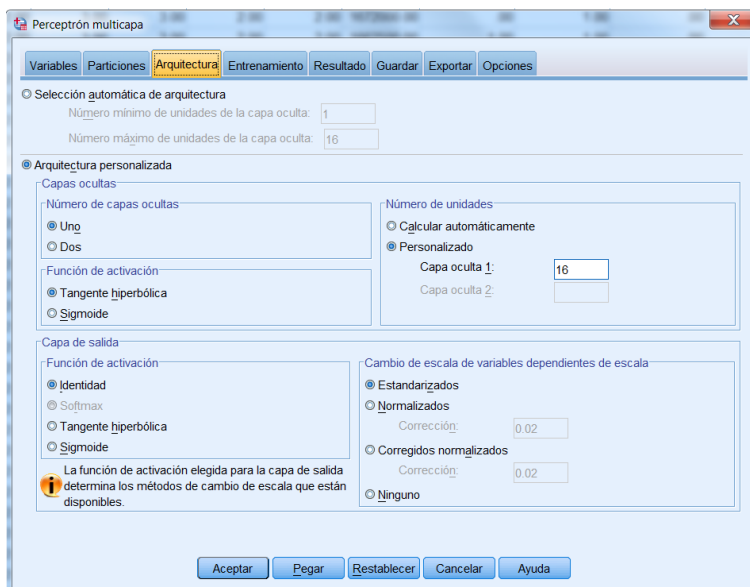
Figura 5.55. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura 5.55 contiene información sobre la arquitectura de la red: se ha seleccionado una capa oculta un máximo de 16 unidades o patrones en la misma, con una función de activación a la entrada tangente hiperbólica y una función de salida de la red denominada identidad (lineal).

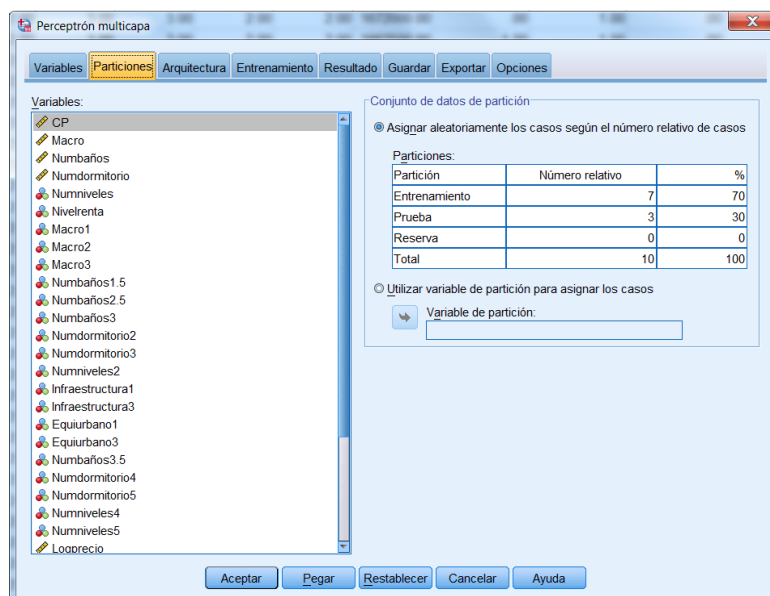
Figura 5.56. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.56 se observa el tipo de entrenamiento denominado en lote con sus diferentes opciones de entrenamiento.

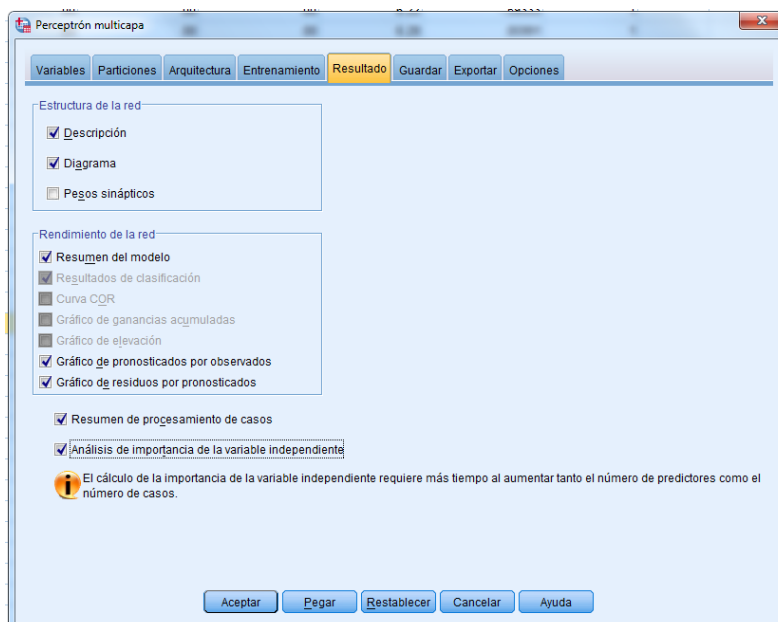
Figura 5.57. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia.

En la figura 5.58 los resultados de la red.

Figura 5.58. Resultado de la RNA



Fuente: Elaboración propia

Una vez ejecutado el entrenamiento de la red, se han obtenido los siguientes resultados que se detallan en las tablas.

Tabla 5.53. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 667 | 76.5% |
| | Pruebas | 205 | 23.5% |
| Válido | | 872 | 100.0% |
| Excluido | | 89 | |
| Total | | 961 | |

Fuente: Elaboración propia

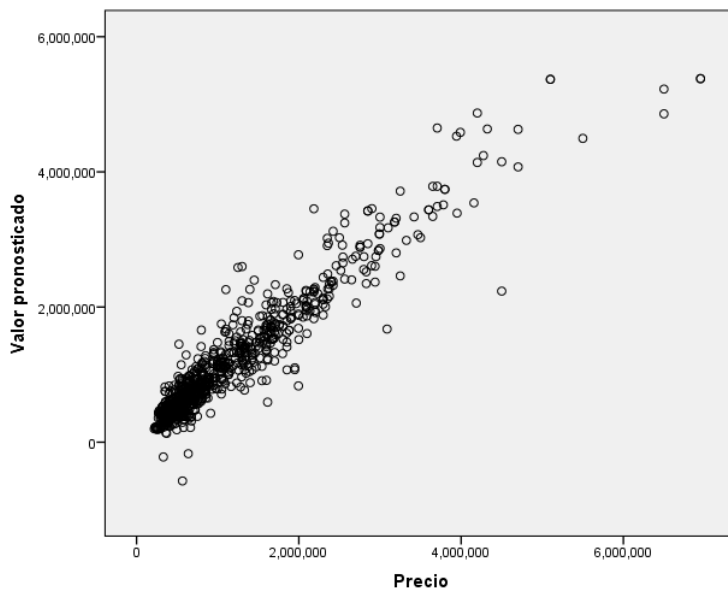
Tabla 5.54. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 19.072 |
| | Error relativo | .057 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:02.13 |
| Pruebas | Error de suma de cuadrados | 21.763 |
| | Error relativo | .232 |

Fuente: Elaboración propia

En la tabla 5.52 se muestran la suma de errores de 19.072%, y un error relativo de 0.057%.

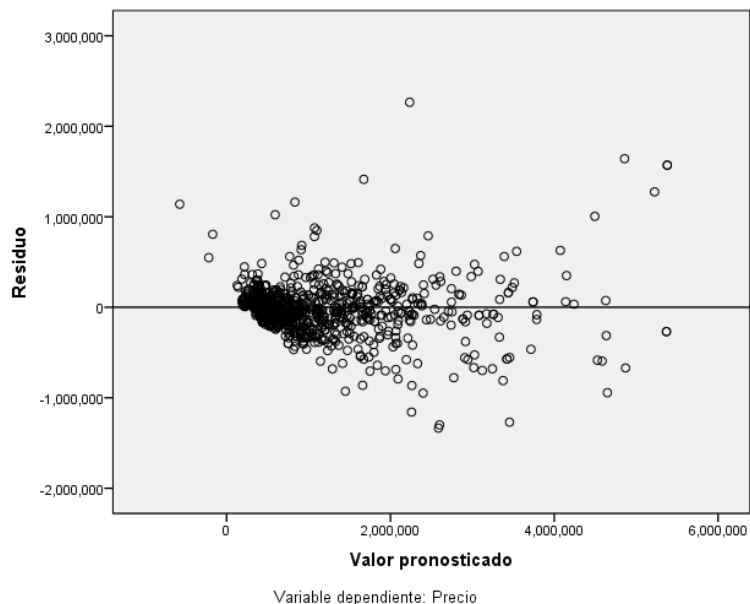
Figura 5.59. Gráfica de precio observado y precio pronosticado



Fuente: Elaboración propia

Continuando en la figura 5.60 se aprecia lo siguiente:

Figura 5.60. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

La figura 5.60 representa un gráfico de los residuos contra los valores pronosticados.

En la tabla siguiente se ilustran los valores de la importancia de las variables independientes.

Tabla 5.55. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .095 | 75.8% |
| construcción | .080 | 64.3% |
| Antigüedad | .055 | 43.7% |
| Nivelrentaalta | .047 | 37.6% |
| Infraestructura | .024 | 19.1% |
| Garage | .117 | 94.1% |
| Garage2 | .023 | 18.8% |
| Numniveles3 | .047 | 37.8% |
| Numbaños5 | .098 | 78.7% |
| Numbaños4 | .076 | 60.5% |
| Nivelrentabaja | .035 | 28.2% |
| Numdormitorio6 | .049 | 38.9% |
| Numbaños2 | .050 | 39.8% |
| Garage3 | .058 | 46.5% |
| Numbaños4.5 | .125 | 100.0% |
| Equiurbano | .021 | 16.9% |

Fuente: Elaboración propia

Una vez que la red se ha entrenado en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, se han obtenido los siguientes resultados.

Tabla 5.56. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa salida | Número de unidades en la capa oculta | Traning(entre namiento) | Pruebas | Excluidos | T. de entrenam ento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|------------------------------|--------------------------------------|-------------------------|---------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Identidad | 16 | 667 | 215 | 79 | 10 minutos | 89% | 79% | 79% | 469711 | 469178 |
| Tangente hiperbólica | Identidad | 16 | 667 | 205 | 89 | 20 minutos | 95% | 91% | 91% | 296471 | 296131 |
| Tangente hiperbólica | Identidad | 16 | 651 | 229 | 81 | 30 minutos | 90% | 81% | 80% | 424288 | 423806 |
| Tangente hiperbólica | Identidad | 16 | 669 | 227 | 65 | 40 minutos | 87% | 76% | 76% | 486067 | 485525 |
| Tangente hiperbólica | Identidad | 16 | 660 | 219 | 82 | 50 minutos | 94% | 89% | 89% | 313486 | 313129 |
| Tangente hiperbólica | Identidad | 16 | 677 | 201 | 83 | 60 minutos | 94% | 89% | 89% | 313486 | 313129 |
| Tangente hiperbólica | Identidad | 16 | 671 | 223 | 67 | 80 minutos | 94% | 89% | 88% | 337817 | 337438 |
| Tangente hiperbólica | Identidad | 16 | 676 | 215 | 70 | 100 minutos | 90% | 80% | 80% | 431355 | 430871 |

Fuente: Elaboración propia

En tabla 5.56 se han obtenido resultados del entrenamiento de la red.

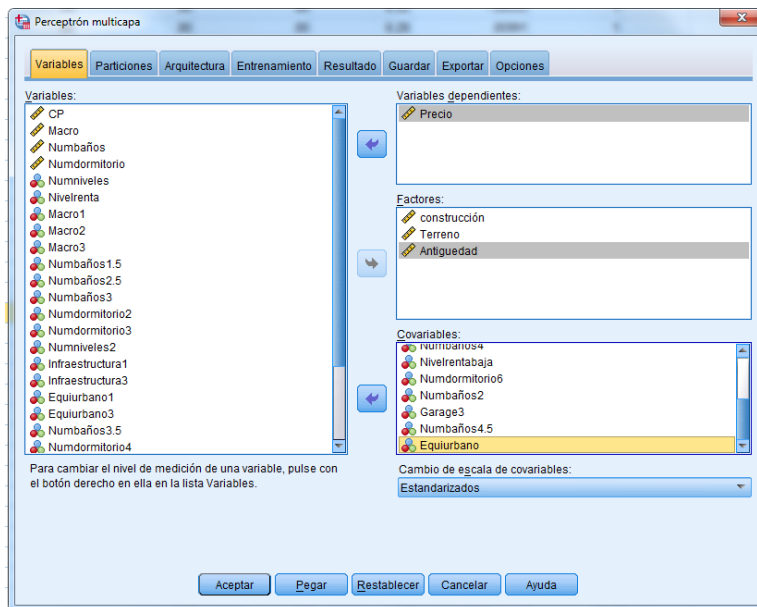
Las conclusiones de este modelo de acuerdo con los resultados obtenidos, una vez que se entrenó a la red con intervalos de tiempo de 10 minutos hasta llegar a 100 minutos; utilizando 16 variables independientes en la capa oculta, el *test* de entrenamiento con tiempo de 20 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 95%, coeficiente de determinación de 91%, y un r^2 ajustado de 91%.

El modelo de entrenamiento se repite con una red con entradas de neuronas en la capa oculta, utilizando una función tangente hiperbólica y una salida tangente hiperbólica, con diferentes intervalos de tiempo y se evaluarán los resultados para su vez verificar la mejoría de

la misma.

La siguiente figura contiene la información de las variables a utilizar en el modelo, mediante el programa SPSS.

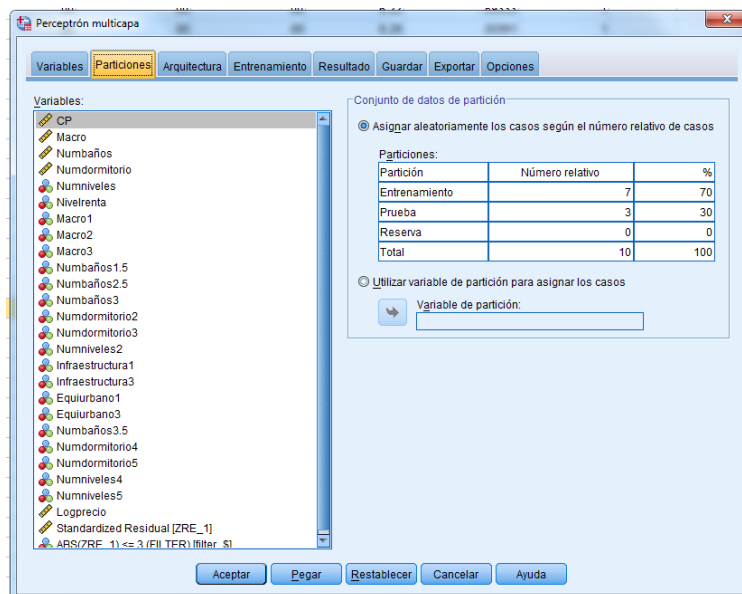
Figura 5.55. Variables a utilizar en el modelo de la RNA



Fuente: Elaboración propia

En la siguiente figura se observan el tipo de patrones de entrada a la red.

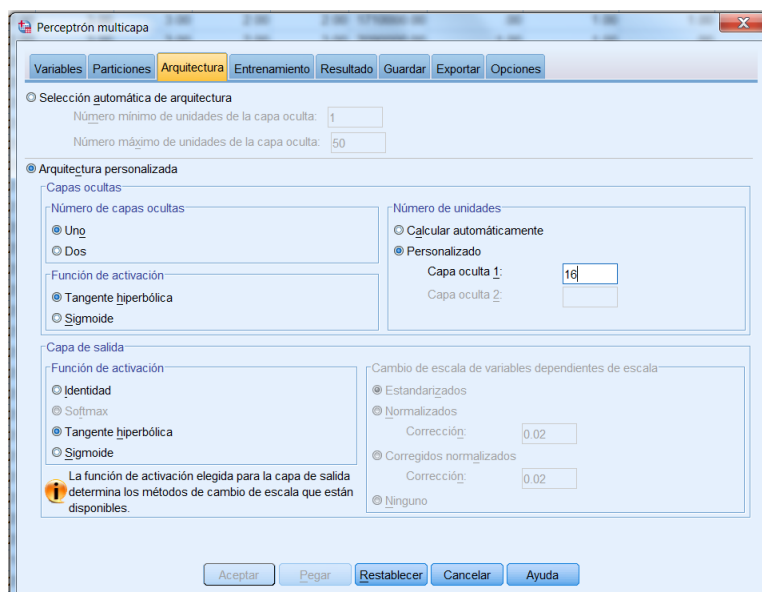
Figura 5.56. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

La figura 5.56 contiene información acerca de la arquitectura de la red.

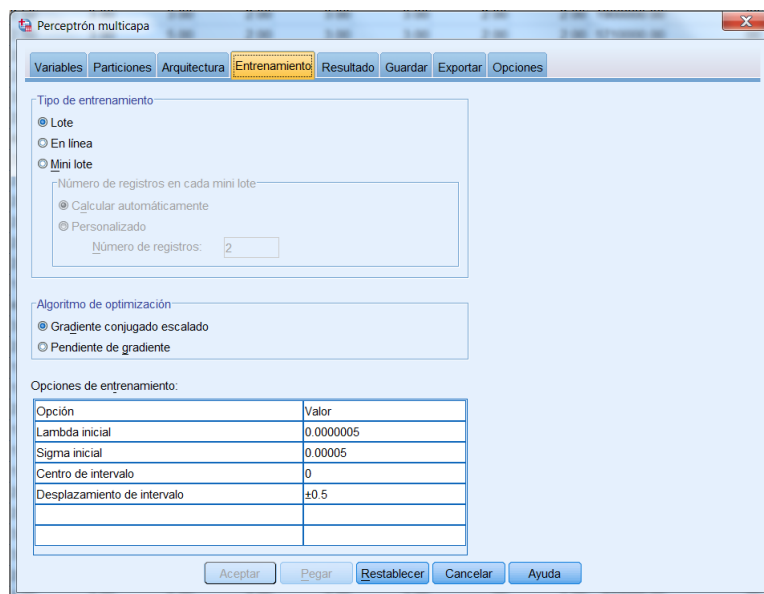
Figura 5.57. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



Fuente: Elaboración propia

En la figura 5.58 se observa el tipo de entrenamiento denominado en lote con sus diferentes opciones de entrenamiento.

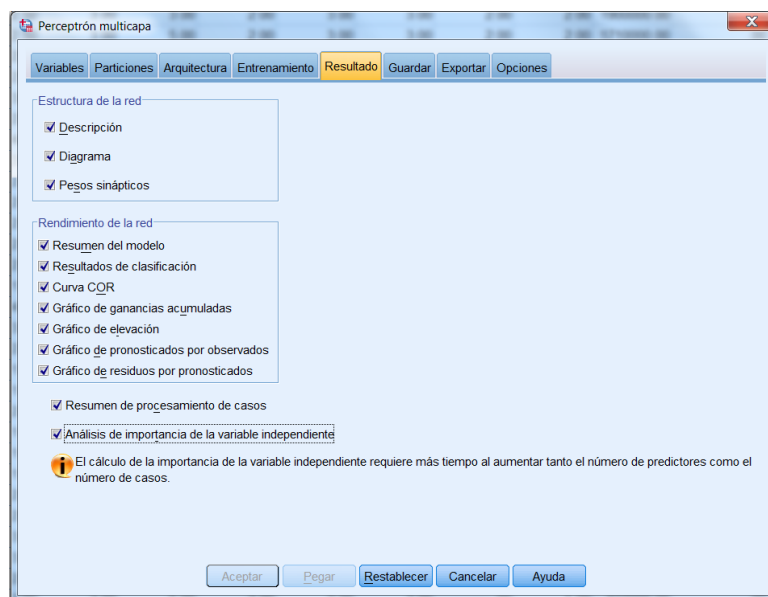
Figura 5.58. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

En la figura 5.59 se muestran los resultados de la red.

Figura 5.59. Resultado de la RNA



Fuente: Elaboración propia

Al ejecutar el entrenamiento de la red, se obtuvieron los resultados que se detallan en las siguientes tablas siguientes.

Tabla 5.57. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 698 | 78.7% |
| | Pruebas | 189 | 21.3% |
| Válido | | 887 | 100.0% |
| Excluido | | 74 | |
| Total | | 961 | |

Fuente: Elaboración propia

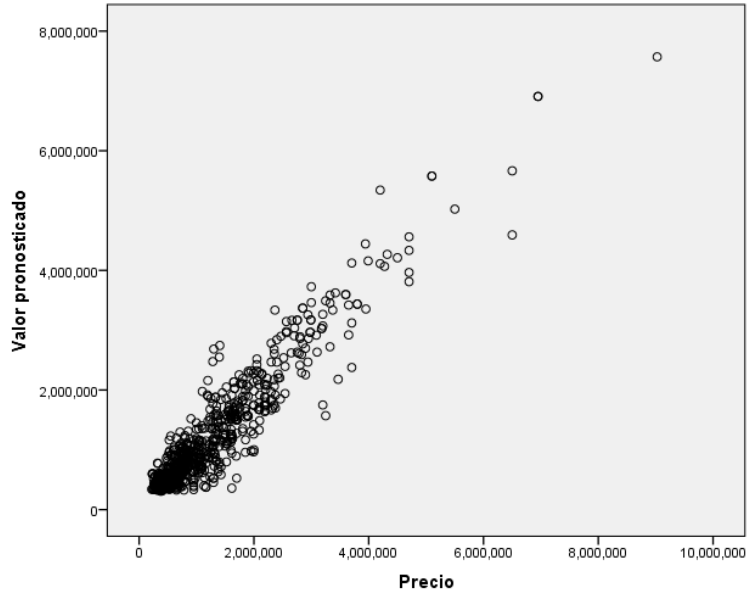
Tabla 5.58. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 1.617 |
| | Error relativo | .091 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:04.22 |
| Pruebas | Error de suma de cuadrados | .801 |
| | Error relativo | .159 |

Fuente: Elaboración propia

La tabla 5.58 se muestra la suma de errores de 1.617%, y un error relativo de 0.091%.

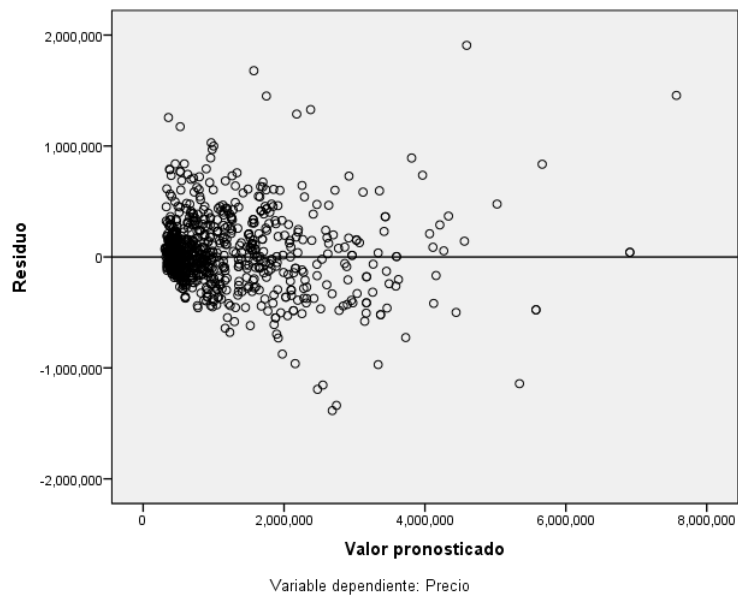
Figura 5.60. Gráfica de precio observado y valor pronosticado



Fuente: Elaboración propia

Continuando en la figura [5.61], se observa lo siguiente.

Figura 5.61. Gráfica de residuos y valor pronosticado



Fuente: Elaboración propia

En la tabla siguiente se observa la importancia de las variables independientes.

Tabla 5.59. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .065 | 37.7% |
| construcción | .070 | 40.1% |
| Antigüedad | .052 | 30.0% |
| Nivelrentaalta | .055 | 31.6% |
| Infraestructura | .028 | 16.2% |
| Garage | .157 | 90.5% |
| Garage2 | .021 | 12.0% |
| Numniveles3 | .044 | 25.3% |
| Numbaños5 | .174 | 100.0% |
| Numbaños4 | .042 | 24.1% |
| Nivelrentabaja | .037 | 21.3% |
| Numdormitorio6 | .045 | 25.9% |
| Numbaños2 | .021 | 11.9% |
| Garage3 | .042 | 24.4% |
| Numbaños4.5 | .119 | 68.7% |
| Equiurbano | .028 | 15.9% |

Fuente: Elaboración propia

Se reiteran aquí los comentarios hechos en los modelos anteriores una vez entrenada la red en intervalos de tiempo de 10 minutos con incrementos de 10 minutos hasta obtener el mejor resultado, sin sobreentrenar a la red, y se obtuvieron los siguientes resultados. Véase la tabla siguiente.

Tabla 5.60. *Test* de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación de capa salida | Número de unidades en la capa oculta | Training (entrenamiento) | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|--------------------------------------|--------------------------|---------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Tangente hiperbólica | 16 | 680 | 210 | 71 | 10 minutos | 90% | 82% | 81% | 411020 | 410557 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 677 | 211 | 73 | 20 minutos | 93% | 87% | 87% | 355893 | 355492 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 671 | 202 | 88 | 30 minutos | 89% | 79% | 79% | 423863 | 423377 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 693 | 201 | 67 | 40 minutos | 87% | 77% | 76% | 475813 | 475080 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 677 | 211 | 73 | 50 minutos | 92% | 84% | 84% | 389433 | 388994 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 662 | 221 | 78 | 60 minutos | 90% | 81% | 81% | 436655 | 436160 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 698 | 189 | 74 | 80 minutos | 95% | 90% | 90% | 325477 | 325110 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 691 | 203 | 67 | 100 minutos | 88% | 78% | 77% | 466332 | 465810 |

Fuente: Elaboración propia

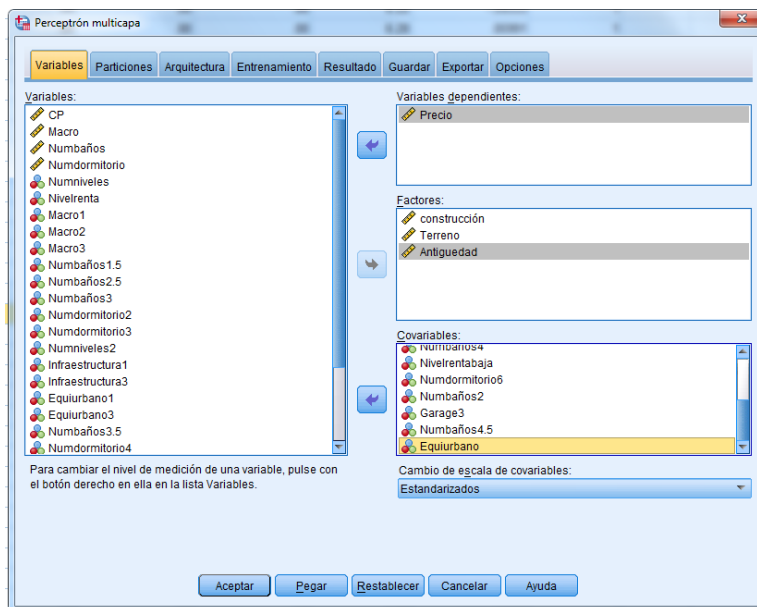
Las conclusiones acerca de este modelo de acuerdo con los resultados obtenidos, una vez que se entrenó la red con intervalos de tiempo de 10 minutos hasta llegar a 100 minutos; utilizando 16 variables independientes en la capa oculta, el *test* de entrenamiento con tiempo de 80 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 95%, un coeficiente de determinación de 90%, y un r^2 ajustado de 90%.

Ahora se presentan los resultados utilizando el modelo con una función de activación en la capa oculta de entrada de sus patrones tangente hiperbólica y en sus salidas de la red, con las mismas observaciones, la función denominada sigmoidea; y se ha personalizado el número de unidades en la capa oculta a 16 variables, toda

la demás información será igual.

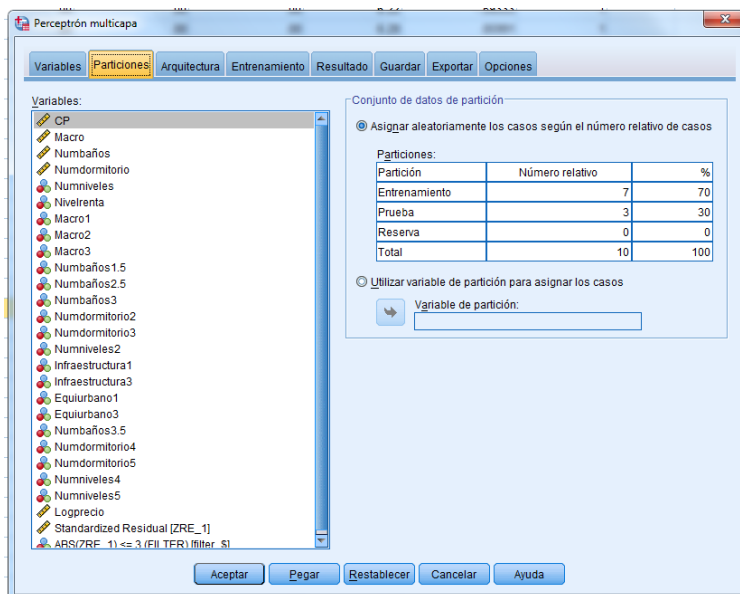
La siguiente figura contiene información de las variables a utilizar en el modelo, mediante el programa SPSS.

Figura 5.62. Variables a utilizar en el modelo de la RNA



Fuente: Elaboración propia

Figura. 5.63. Patrones de entrada a la red, entrenamiento y prueba del modelo



Fuente: Elaboración propia

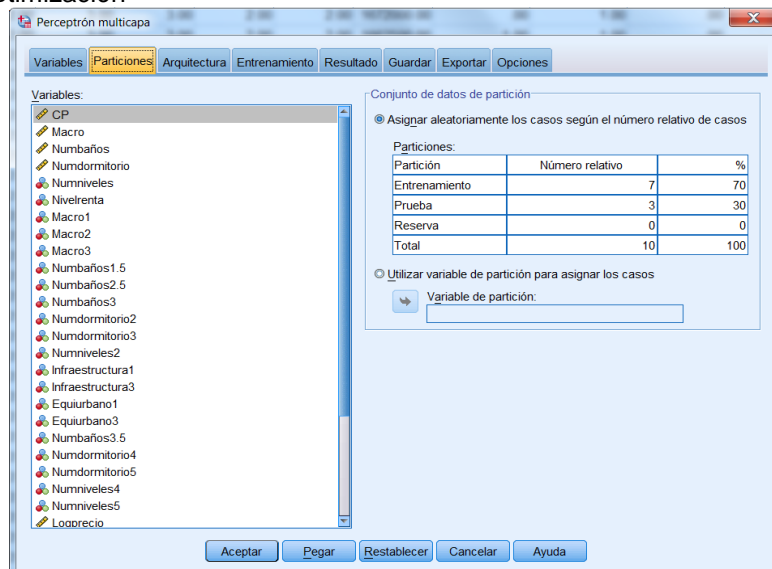
Figura 5.64. Arquitectura, número de capas ocultas, función de entrada y de salida de la red

The screenshot shows the 'Perceptrón multicapa' window with the 'Arquitectura' tab selected. The interface includes the following elements:

- Selección automática de arquitectura:** Radio button selected. Input fields for 'Número mínimo de unidades de la capa oculta' (1) and 'Número máximo de unidades de la capa oculta' (16).
- Arquitectura personalizada:** Radio button selected.
 - Capas ocultas:** 'Número de capas ocultas' set to 'Uno' (1). 'Número de unidades' set to 'Personalizado' with 'Capa oculta 1' at 16 and 'Capa oculta 2' empty.
 - Función de activación:** 'Tangente hiperbólica' selected.
 - Capa de salida:** 'Sigmoide' selected.
 - Cambio de escala de variables dependientes de escala:** 'Normalizados' selected with 'Corrección' set to 0.02.
- Función de activación de la capa de salida:** 'Sigmoide' selected. A warning icon and text state: 'La función de activación elegida para la capa de salida determina los métodos de cambio de escala que están disponibles.'
- Buttons:** 'Aceptar', 'Pegar', 'Restablecer', 'Cancelar', and 'Ayuda' at the bottom.

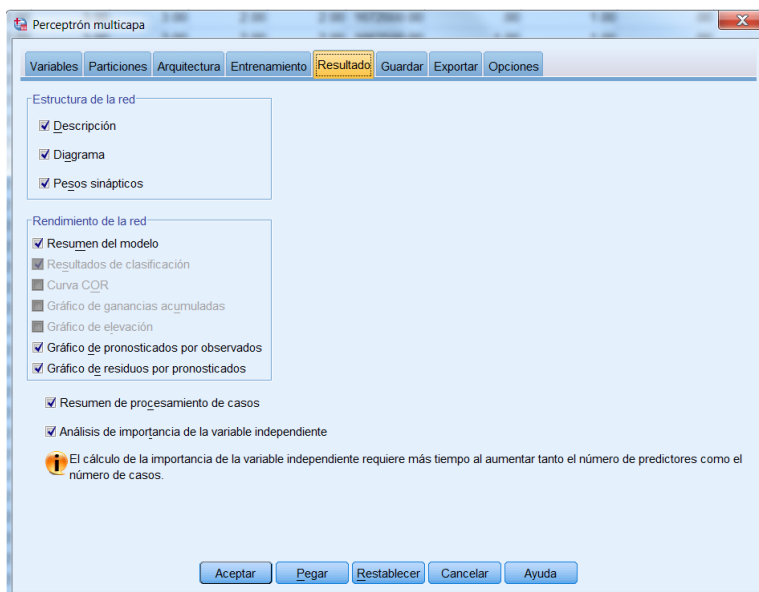
Fuente: Elaboración propia

Figura 5.65. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

Figura 5.67. Resultado de la RNA



Fuente: Elaboración propia

Tabla 5.61. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 663 | 75.9% |
| | Pruebas | 210 | 24.1% |
| Válido | | 873 | 100.0% |
| Excluido | | 88 | |
| Total | | 961 | |

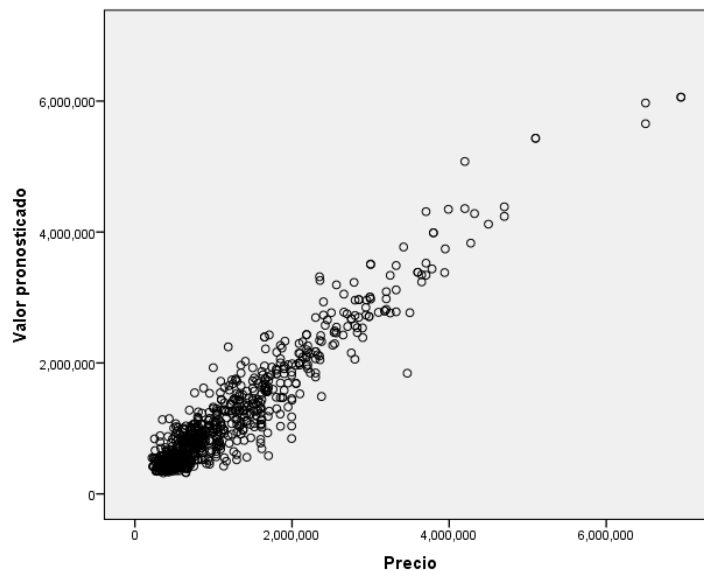
Fuente: Elaboración propia

Tabla 5.62. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | .437 |
| | Error relativo | .066 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:03.64 |
| Pruebas | Error de suma de cuadrados | .317 |
| | Error relativo | .157 |

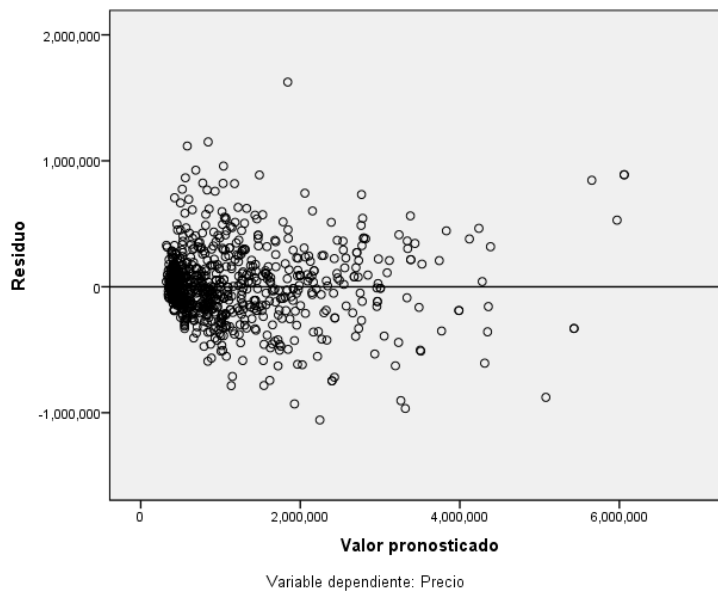
Fuente: Elaboración propia

Figura 5.68. Gráfica de precio observado y precio pronosticado



Fuente: Elaboración propia

Figura 5.69. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

Tabla 5.63. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .077 | 41.4% |
| construcción | .071 | 38.3% |
| Antigüedad | .047 | 25.3% |
| Nivelrentaalta | .052 | 28.0% |
| Infraestructura | .021 | 11.5% |
| Garage | .157 | 85.1% |
| Garage2 | .012 | 6.7% |
| Numniveles3 | .042 | 22.8% |
| Numbaños5 | .074 | 40.0% |
| Numbaños4 | .102 | 54.9% |
| Nivelrentabaja | .038 | 20.3% |
| Numdormitorio | .055 | 29.7% |
| Numbaños2 | .020 | 10.6% |
| Garage3 | .033 | 17.7% |
| Numbaños4.5 | .185 | 100.0% |
| Equiurbano | .015 | 8.3% |

Fuente: Elaboración propia

Tabla 5.64. *Test* de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa salida | Número de unidades en la capa oculta | Traning(entrenamiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|------------------------------|--------------------------------------|------------------------|--------|-----------|---------------------|----------------------|-----------------------|-------------------------|------------------------------|------------------------|
| Tangente hiperbólica | Sigmoidea | 16 | 665 | 208 | 88 | 10 minutos | 96% | 91% | 91% | 280030 | 279709 |
| Tangente hiperbólica | Sigmoidea | 16 | 688 | 197 | 76 | 20 minutos | 93% | 86% | 85% | 386385 | 385948 |
| Tangente hiperbólica | Sigmoidea | 16 | 666 | 220 | 75 | 30 minutos | 93% | 87% | 87% | 352728 | 352330 |
| Tangente hiperbólica | Sigmoidea | 16 | 692 | 212 | 57 | 40 minutos | 91% | 83% | 83% | 408733 | 408280 |
| Tangente hiperbólica | Sigmoidea | 16 | 654 | 219 | 88 | 50 minutos | 95% | 90% | 90% | 310852 | 310496 |
| Tangente hiperbólica | Sigmoidea | 16 | 668 | 211 | 82 | 60 minutos | 93% | 86% | 86% | 373685 | 373240 |
| Tangente hiperbólica | Sigmoidea | 16 | 669 | 222 | 70 | 80 minutos | 95% | 90% | 90% | 323430 | 323067 |
| Tangente hiperbólica | Sigmoidea | 16 | 663 | 210 | 88 | 100 minutos | 94% | 89% | 89% | 320197 | 319830 |

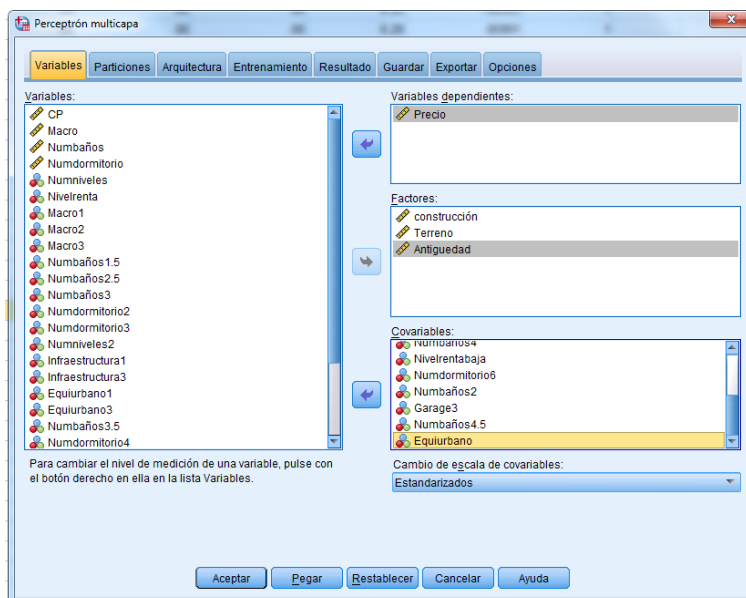
Fuente: Elaboración propia

Se puede concluir que de los ocho intervalos de tiempo, una vez que se entrenó a la red con intervalos de tiempo de 10 minutos hasta llegar a 100 minutos, utilizando 16 variables independientes en la capa oculta, el *test* de entrenamiento con tiempo de 10 minutos es el que mejor ha estimado el precio: coeficiente de correlación de 96%, un coeficiente de determinación de 91%, y un r^2 ajustado de 91%.

Continuando ahora con otro modelo, se usará una función de activación en la capa oculta de entrada de sus patrones sigmoidea (sigmoidea) y en sus salidas de la red, con las mismas observaciones, función denominada identidad (lineal); se personalizó el número de unidades en la capa oculta a 16 variables, toda la demás información será igual.

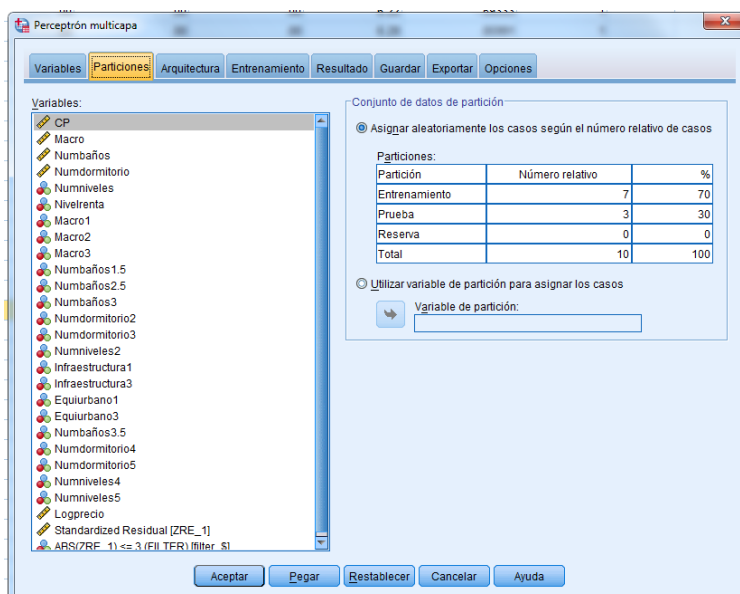
Se iniciará el entrenamiento de la red con entradas de neuronas en la capa oculta, mediante una función sigmoidea y una salida identidad, con diferentes intervalos de tiempo y se irán evaluando los resultados para a su vez verificar la mejoría de la misma.

Figura 5.70. Variables a utilizar en el modelo de la RNA



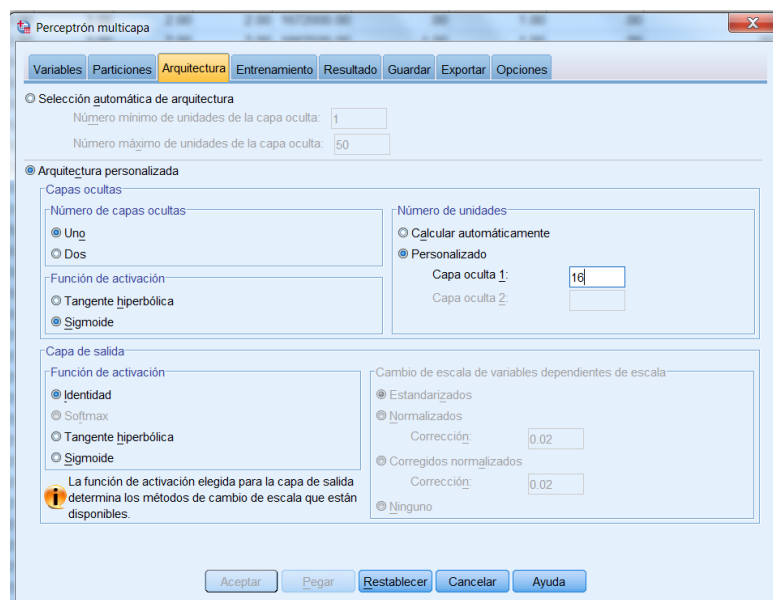
Fuente: Elaboración propia

Figura 5.71. Patrones de entrada a la red, entrenamiento y prueba del mismo



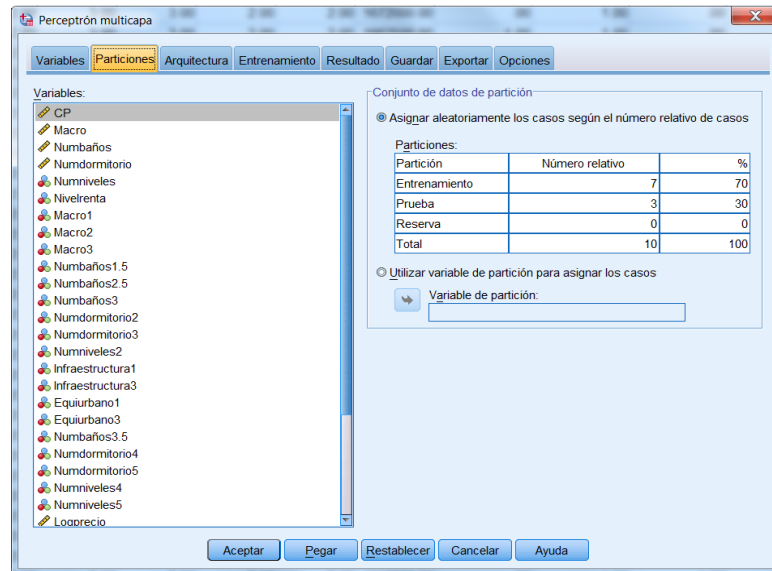
Fuente: Elaboración propia

Figura 5.72. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



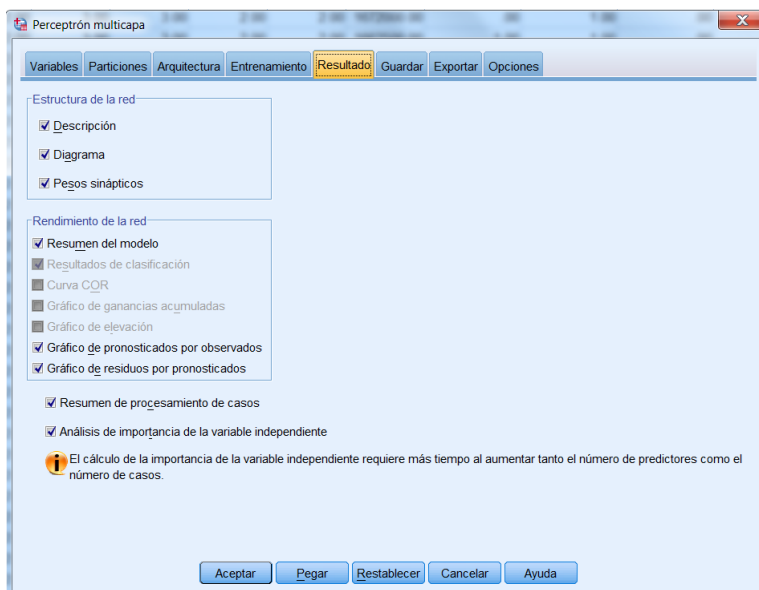
Fuente: Elaboración propia

Figura 5.73. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

Figura 5.73. Resultado de la RNA



Fuente: Elaboración propia

Tabla 5.65. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 677 | 76.2% |
| | Pruebas | 211 | 23.8% |
| Válido | | 888 | 100.0% |
| Excluido | | 73 | |
| Total | | 961 | |

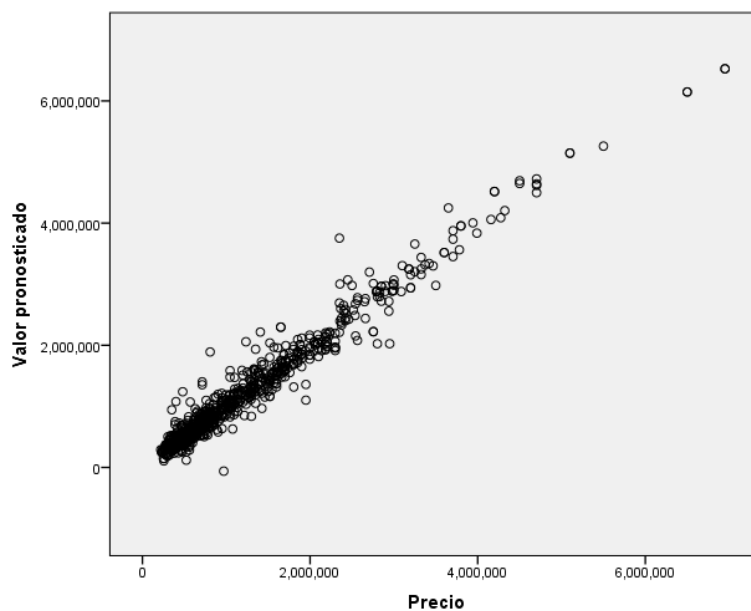
Fuente: Elaboración propia

Tabla 5.66. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 5.050 |
| | Error relativo | .015 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:04.10 |
| Pruebas | Error de suma de cuadrados | 10.421 |
| | Error relativo | .104 |

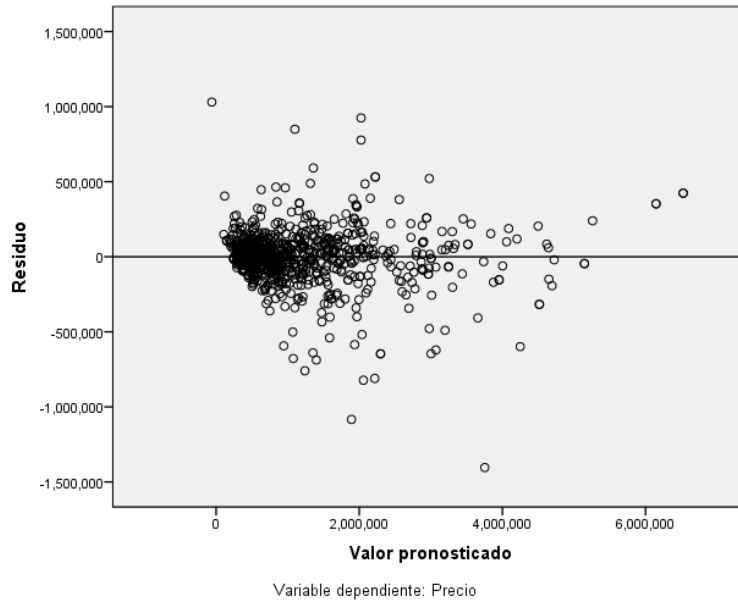
Fuente: Elaboración propia

Figura 5.74. Gráfica de precio observado y precio pronosticado



Fuente: Elaboración propia

Figura 5.75. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

Tabla 5.67. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|----------------|-------------|-------------------------|
| Terreno | .164 | 100.0% |
| construcción | .159 | 96.9% |
| Antigüedad | .050 | 30.8% |
| Nivelrentaalta | .044 | 26.8% |
| Infraestructur | .023 | 13.7% |
| Garage | .103 | 63.1% |
| Garage2 | .025 | 15.3% |
| Numniveles3 | .056 | 34.2% |
| Numbaños5 | .062 | 38.1% |
| Nivelrentabaja | .026 | 16.0% |
| Numdormitorio | .087 | 52.9% |
| Numbaños2 | .036 | 22.2% |
| Garage3 | .038 | 23.2% |
| Numbaños4.5 | .107 | 65.6% |
| Equiurbano | .018 | 11.3% |

Fuente: Elaboración propia

Tabla 5.68. Test de entrenamiento de la RNA

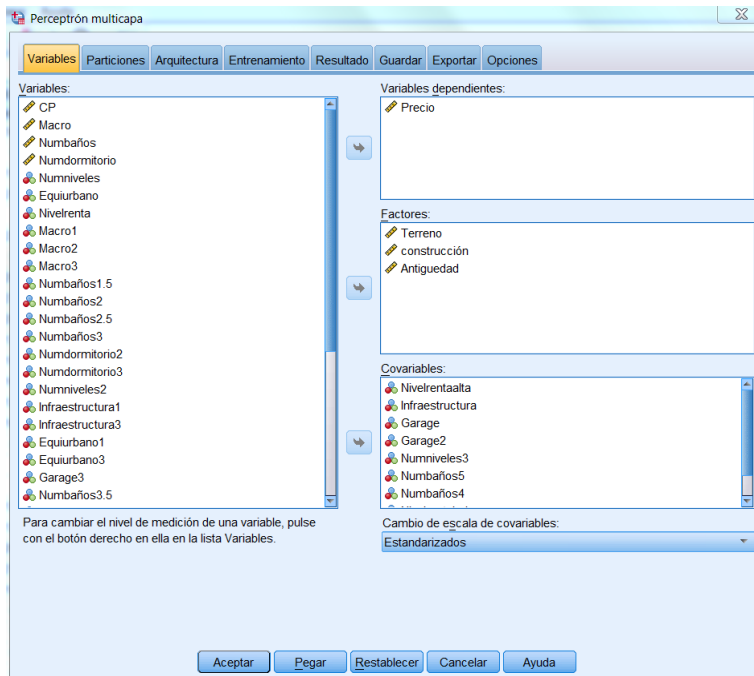
| F. de activación capa oculta | F. de activación capa salida | Número de unidades en la capa oculta | Traning(entre namiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|------------------------------|--------------------------------------|-------------------------|--------|-----------|---------------------|----------------------|-----------------------|-------------------------|------------------------------|------------------------|
| Sigmoidea | Identidad | 16 | 677 | 211 | 73 | 10 minutos | 98% | 97% | 96% | 183206 | 183000 |
| Sigmoidea | Identidad | 16 | 671 | 202 | 88 | 20 minutos | 88% | 78% | 77% | 442307 | 441800 |
| Sigmoidea | Identidad | 16 | 693 | 201 | 67 | 30 minutos | 84% | 70% | 69% | 561851 | 561222 |
| Sigmoidea | Identidad | 16 | 677 | 211 | 73 | 40 minutos | 94% | 88% | 87% | 346382 | 345992 |
| Sigmoidea | Identidad | 16 | 662 | 221 | 78 | 50 minutos | 92% | 85% | 85% | 410185 | 409720 |
| Sigmoidea | Identidad | 16 | 698 | 189 | 74 | 60 minutos | 98% | 96% | 96% | 190132 | 189917 |
| Sigmoidea | Identidad | 16 | 691 | 203 | 67 | 80 minutos | 85% | 72% | 71% | 563051 | 562421 |
| Sigmoidea | Identidad | 16 | 665 | 208 | 88 | 100 minutos | 98% | 96% | 96% | 183597 | 183386 |

Fuente: Elaboración propia

Una vez analizados los resultados de los ocho intervalos de tiempo de entrenamiento de la red, el *test* de entrenamiento de 10 minutos es el que mejores datos estadísticos refleja, reiterando un tiempo de 10 minutos, utilizando 16 variables independientes en la capa oculta: coeficiente de correlación entre variables de 98%, un coeficiente de determinación de 97%, y un r^2 ajustado de 96%.

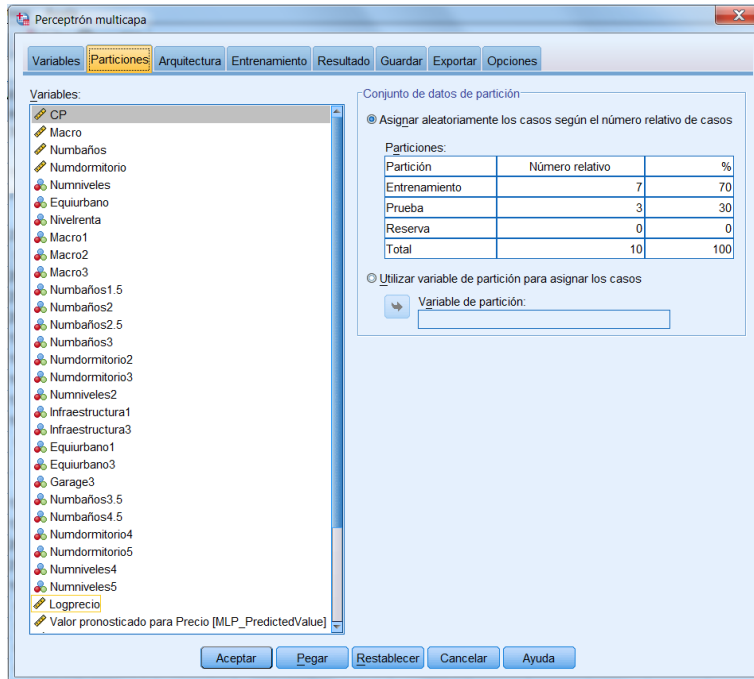
Ahora se utilizará una función de activación en la capa oculta de entrada de los patrones sigmoidea (sigmoidea) y en las salidas de la red, con las mismas observaciones, una función denominada tangente hiperbólica; personalizando el número de unidades en la capa oculta a 16 variables, toda la demás información será igual.

Figura 5.76. Variables a utilizar en el modelo RNA



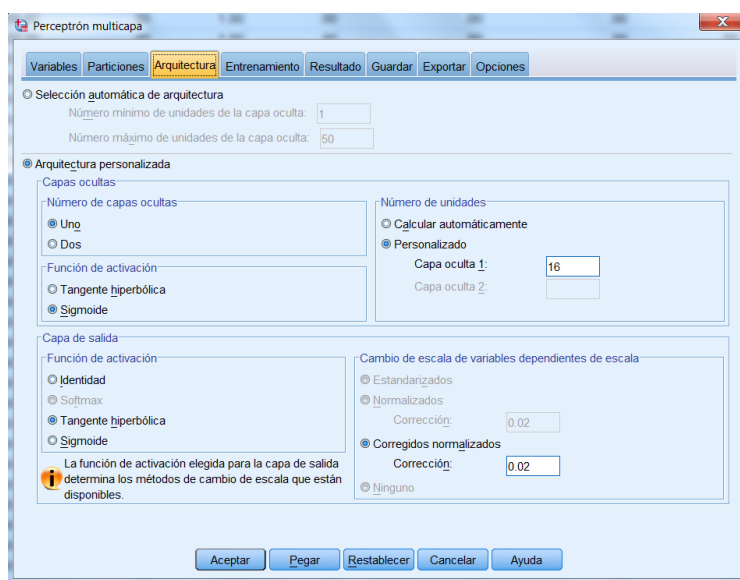
Fuente: Elaboración propia

Figura 5.77. Patrones de entrada a la red, entrenamiento y prueba del modelo



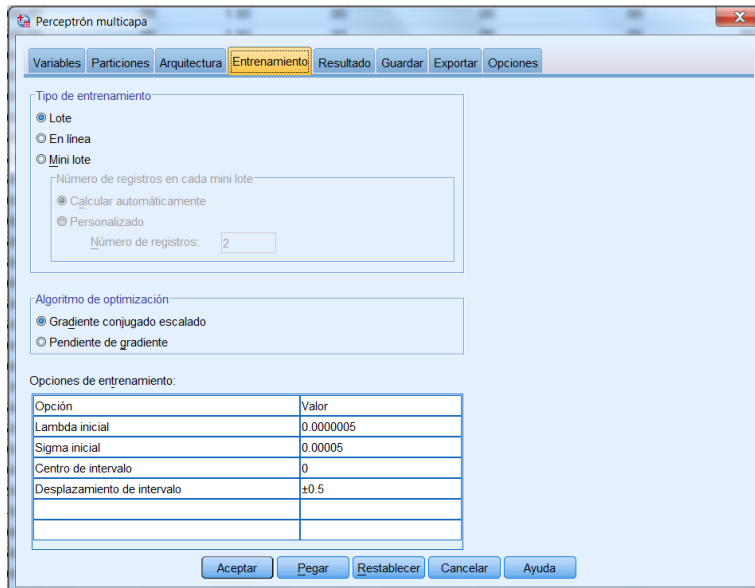
Fuente: Elaboración propia

Figura 5.78. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



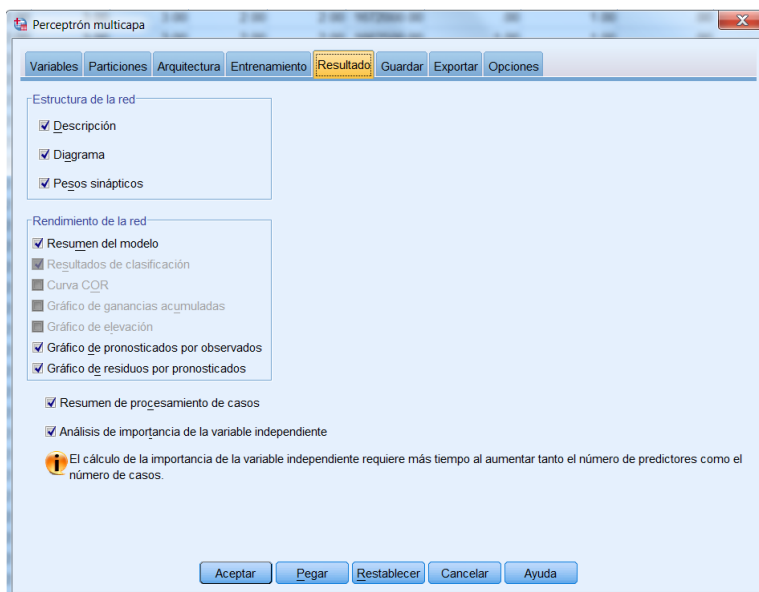
Fuente: Elaboración propia

Figura 5.79. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

Figura 5.80. Resultado de la RNA



Fuente: Elaboración propia

Tabla 5.69. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 677 | 77.1% |
| | Pruebas | 201 | 22.9% |
| Válido | | 878 | 100.0% |
| Excluido | | 83 | |
| Total | | 961 | |

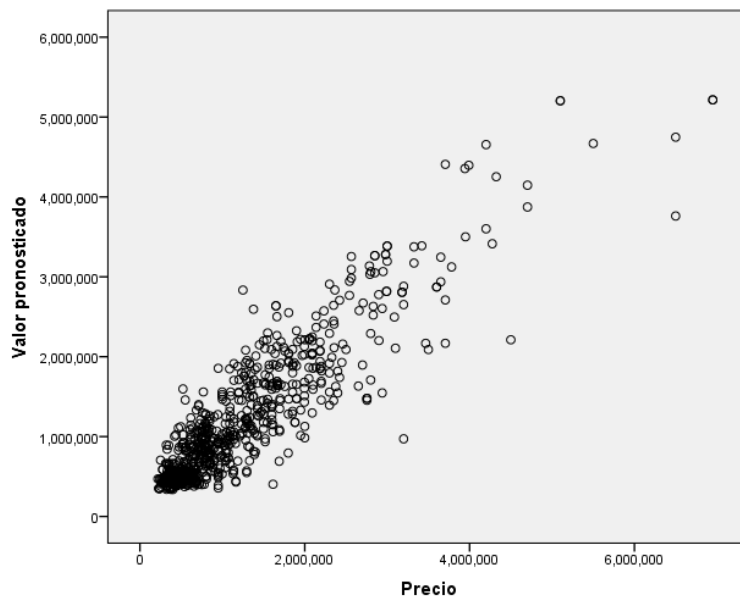
Fuente: Elaboración propia

Tabla 5.70. Errores de predicción de la RNA

| | | |
|----------------------|-----------------------------------|---|
| Entrenamiento | Error de suma de cuadrados | 4.174 |
| | Error relativo | .158 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:05.80 |
| Pruebas | Error de suma de cuadrados | 1.662 |
| | Error relativo | .211 |

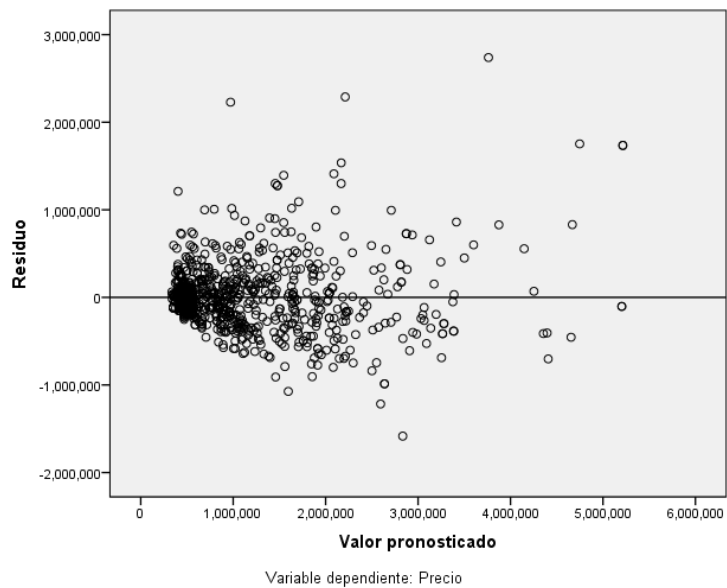
Fuente: Elaboración propia

Figura 5.81. Gráfica de precio observado y precio pronosticado



Fuente: Elaboración propia

Figura 5.82. Gráfica de residuos y valor pronosticado



Fuente. Elaboración propia

Tabla 5.71. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .064 | 39.7% |
| construcción | .066 | 40.8% |
| Antigüedad | .056 | 34.9% |
| Infraestructura | .033 | 20.5% |
| Equiurbano | .025 | 15.6% |
| Garage | .161 | 100.0% |
| Numbaños2 | .010 | 6.1% |
| Garage2 | .021 | 12.9% |
| Garage3 | .047 | 29.0% |
| Nivelrentabaja | .068 | 42.3% |
| Nivelrentaalta | .043 | 26.5% |
| Numbaños4 | .070 | 43.5% |
| Numbaños4.5 | .100 | 62.0% |
| Numbaños5 | .152 | 94.6% |
| Numdormitorio | .048 | 29.7% |
| Numniveles3 | .037 | 23.3% |

Fuente: Elaboración propia

Tabla 5.72. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación de capa salida | Traning(entre namiento) | Prueba | Excluidos | T. de entrenaminto | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|-------------------------|--------|-----------|--------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Sigmoidea | Tangente Hiperbólica | 667 | 215 | 79 | 10 minutos | 85% | 73% | 72% | 539795 | 539183 |
| Sigmoidea | Tangente Hiperbólica | 667 | 205 | 89 | 20 minutos | 89% | 80% | 79% | 449896 | 449379 |
| Sigmoidea | Tangente Hiperbólica | 651 | 229 | 81 | 30 minutos | 88% | 78% | 77% | 470898 | 470363 |
| Sigmoidea | Tangente Hiperbólica | 669 | 227 | 65 | 40 minutos | 85% | 72% | 72% | 513542 | 512968 |
| Sigmoidea | Tangente Hiperbólica | 660 | 219 | 82 | 50 minutos | 91% | 83% | 82% | 431269 | 430778 |
| Sigmoidea | Tangente Hiperbólica | 677 | 201 | 83 | 60 minutos | 91% | 83% | 82% | 387704 | 387262 |
| Sigmoidea | Tangente Hiperbólica | 671 | 223 | 67 | 80 minutos | 89% | 79% | 79% | 454622 | 454113 |
| Sigmoidea | Tangente Hiperbólica | 676 | 215 | 70 | 100 minutos | 86% | 74% | 74% | 475098 | 474565 |

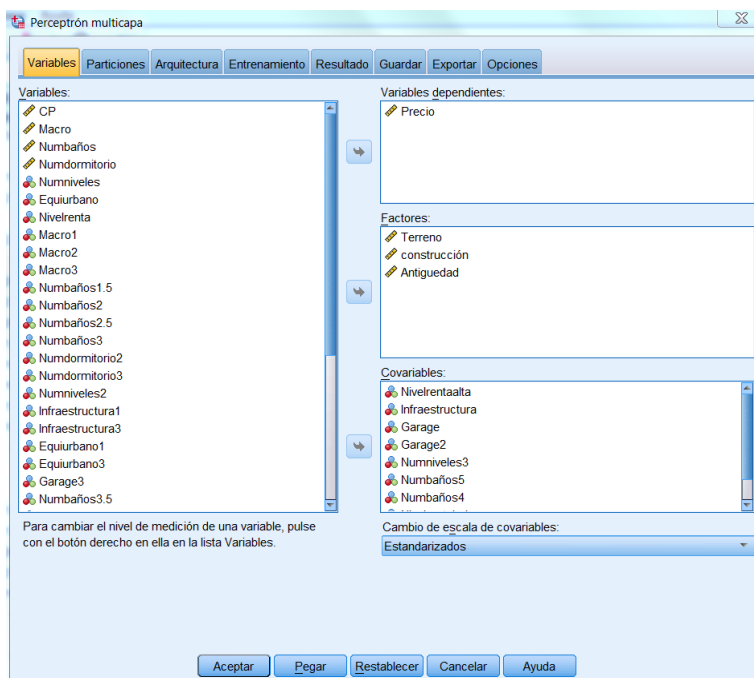
Fuente: Elaboración propia

Una vez analizados los resultados de los ocho intervalos de tiempo de entrenamiento de la red, el *test* de entrenamiento de 60 minutos es el que mejores datos estadísticos refleja, reiterando un tiempo de 60 minutos, utilizando 16 variables independientes en la capa oculta: coeficiente de correlación entre variables de 91%, un coeficiente de determinación de 83%, y un r^2 ajustado de 82%.

Continuando con el modelo, pero ahora utilizando una función de activación en la capa oculta de entrada de los patrones tipo sigmoidea (sigmoidea) y en sus salidas de la red, con las mismas observaciones, una función también sigmoidea; personalizando el número de unidades en la capa oculta a 16 variables, en tanto la demás

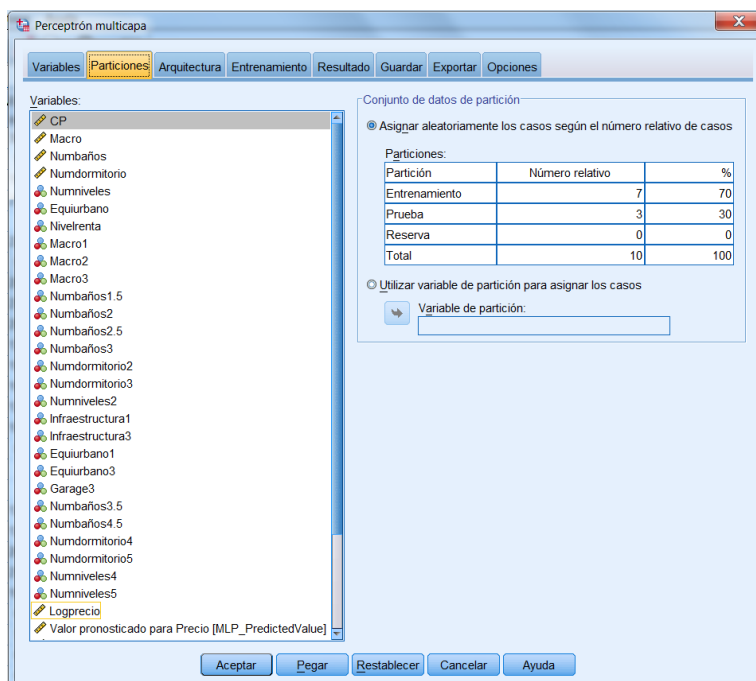
información será igual.

Figura 5.83. Variables a utilizar en el modelo RNA



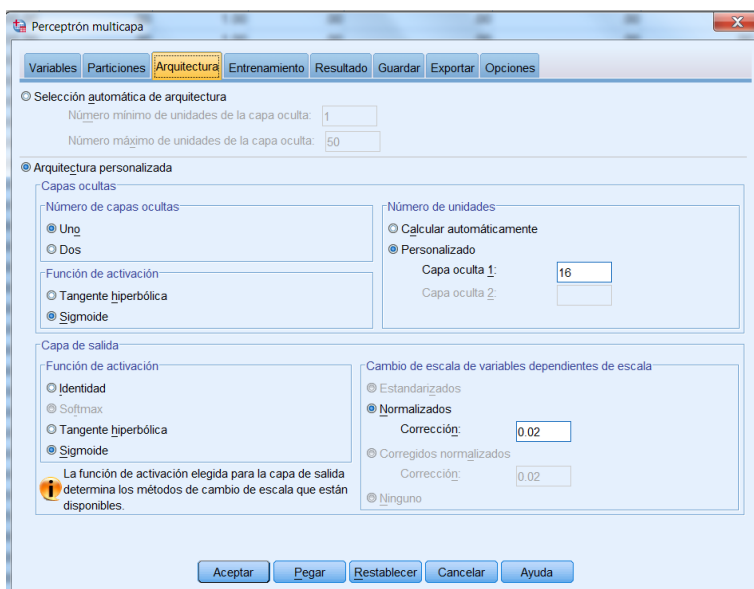
Fuente: Elaboración propia

Figura 5.84. Patrones de entrada a la red, entrenamiento y prueba del modelo



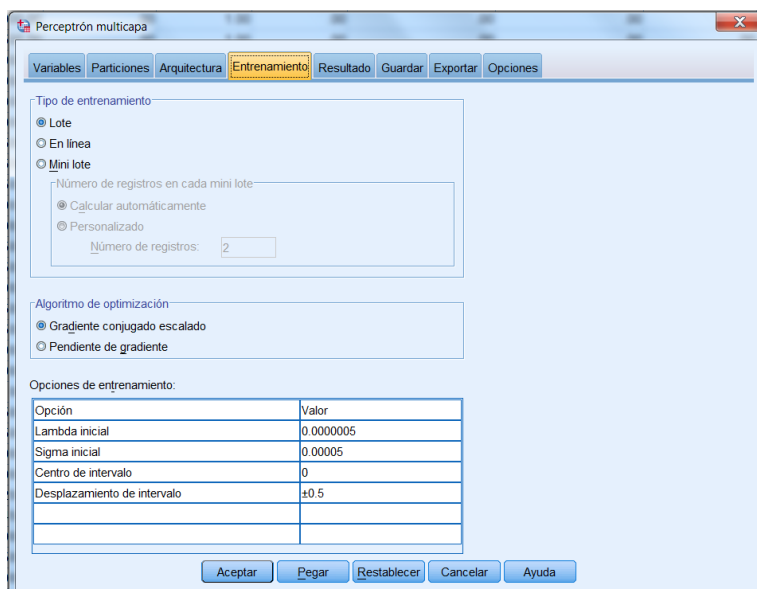
Fuente: Elaboración propia

Figura 5.85. Arquitectura, número de capas ocultas, función de entrada y de salida de la red



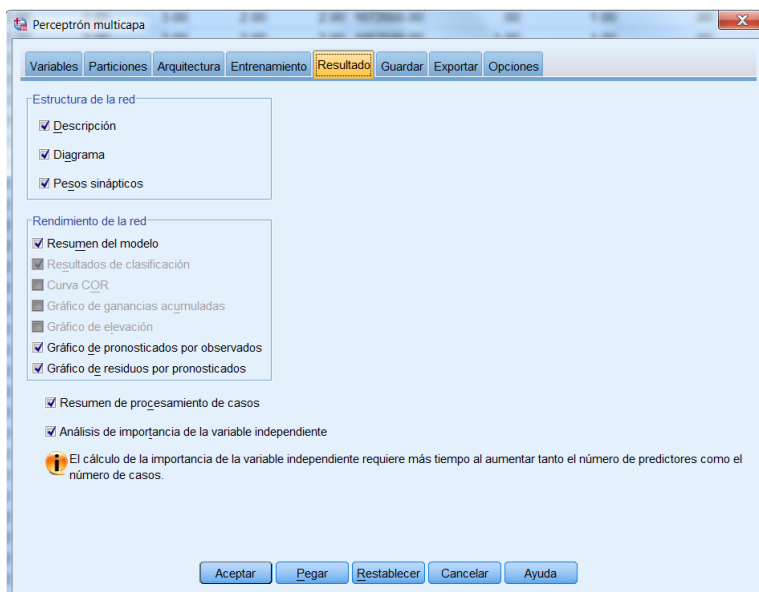
Fuente: Elaboración propia

Figura 5.86. Tipo de entrenamiento y algoritmo de optimización



Fuente: Elaboración propia

Figura 5.87. Resultado de la RNA



Fuente: Elaboración propia

La tabla 5.73. Resumen del procesamiento de casos

| | | N | Porcentaje |
|----------|---------------|-----|------------|
| Ejemplo | Entrenamiento | 660 | 75.1% |
| | Pruebas | 219 | 24.9% |
| Válido | | 879 | 100.0% |
| Excluido | | 82 | |
| Total | | 961 | |

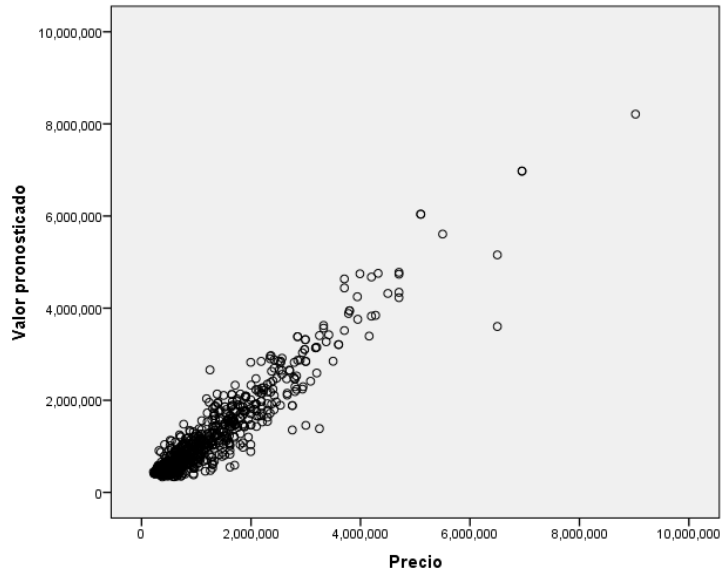
Fuente: Elaboración propia

Tabla 5.74. Errores de predicción de la RNA

| | | |
|---------------|----------------------------|---|
| Entrenamiento | Error de suma de cuadrados | .325 |
| | Error relativo | .077 |
| | Regla de parada utilizada | 1 pasos consecutivos sin disminución del error ^a |
| | Tiempo de preparación | 0:00:02.22 |
| Pruebas | Error de suma de cuadrados | .251 |
| | Error relativo | .173 |

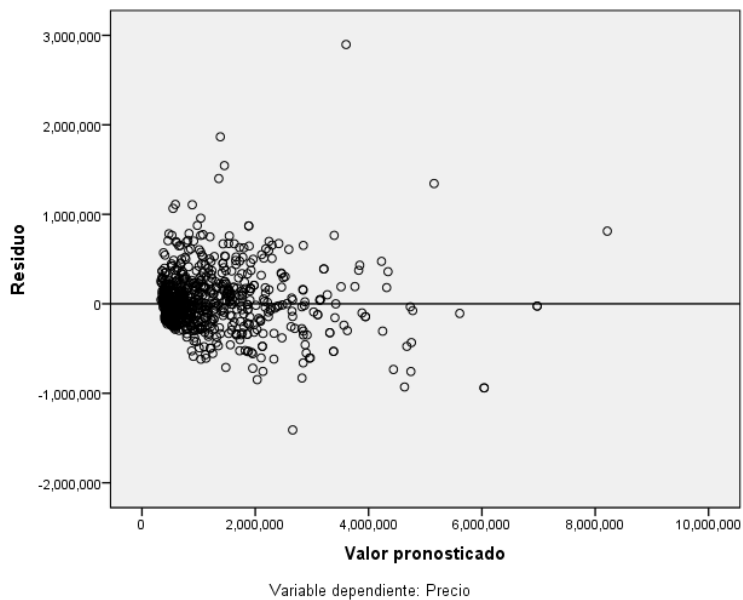
Fuente: Elaboración propia

Figura 5.88. Gráfica de precio observado y precio pronosticado



Fuente: Elaboración propia

Figura 5.89. Gráfica de residuo y valor pronosticado



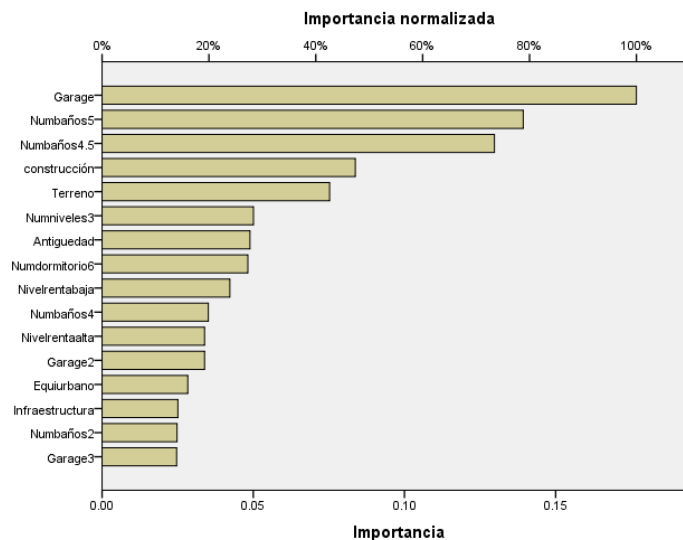
Fuente. Elaboración propia

Tabla 5.90. Importancia de las variables independientes

| | Importancia | Importancia normalizada |
|-----------------|-------------|-------------------------|
| Terreno | .075 | 42.6% |
| construcción | .084 | 47.4% |
| Antigüedad | .049 | 27.7% |
| Infraestructura | .025 | 14.2% |
| Equiurbano | .028 | 16.0% |
| Garage | .177 | 100.0% |
| Numbaños2 | .025 | 14.0% |
| Garage2 | .034 | 19.2% |
| Garage3 | .025 | 13.9% |
| Nivelrentabaja | .042 | 23.9% |
| Nivelrentaalta | .034 | 19.2% |
| Numbaños4 | .035 | 19.9% |
| Numbaños4.5 | .130 | 73.4% |
| Numbaños5 | .139 | 78.8% |
| Numdormitorio | .048 | 27.3% |
| Numniveles3 | .050 | 28.3% |

Fuente: Elaboración propia

Figura 5.90. Gráfica de importancia normalizada



Fuente: Elaboración propia

Tabla 5.92. Test de entrenamiento de la RNA

| F. de activación capa oculta | F. de activación capa de salida | Número de unidades en la capa oculta | Traning(entrenamiento) | Prueba | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|--------------------------------------|------------------------|------------|-----------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Sigmoidea | Sigmoidea | 16 | 667 | 215 | 79 | 10 minutos | 86% | 74% | 73% | 514595 | 514012 |
| Sigmoidea | Sigmoidea | 16 | 667 | 205 | 89 | 20 minutos | 93% | 86% | 86% | 356039 | 355631 |
| Sigmoidea | Sigmoidea | 16 | 651 | 229 | 81 | 30 minutos | 93% | 86% | 86% | 369815 | 369395 |
| Sigmoidea | Sigmoidea | 16 | 669 | 227 | 65 | 40 minutos | 87% | 76% | 75% | 482130 | 481592 |
| Sigmoidea | Sigmoidea | 16 | 660 | 219 | 82 | 50 minutos | 95% | 90% | 90% | 319403 | 319040 |
| Sigmoidea | Sigmoidea | 16 | 677 | 201 | 83 | 60 minutos | 89% | 79% | 79% | 415590 | 415117 |
| Sigmoidea | Sigmoidea | 16 | 671 | 223 | 67 | 80 minutos | 92% | 85% | 85% | 402596 | 402145 |
| Sigmoidea | Sigmoidea | 16 | 676 | 215 | 70 | 100 minutos | 86% | 74% | 73% | 481837 | 481296 |

Fuente: Elaboración propia

Una vez analizados los resultados de los ocho intervalos de tiempo de entrenamiento de la red, el *test* de

entrenamiento de 50 minutos fue el que mejores datos estadísticos reflejó, reiterando un tiempo de 50 minutos, utilizando 16 variables independientes en la capa oculta: coeficiente de correlación entre variables de 95%, un coeficiente de determinación de 90%, y un r^2 ajustado de 90%.

De igual forma que en los primeros 6 modelos anteriores, se obtienen resultados de los segundos 6 modelos, sólo con la diferencia de que ahora se hará con 16 variables (unidades) en la capa oculta; el primero de ellos con una función de activación tangente hiperbólica en la capa oculta con una función de activación en la capa de salida identidad (lineal), en un segundo modelo con función de activación en la capa oculta tangente hiperbólica, función de activación en la capa de salida tangente hiperbólica, en un tercer modelo se ha utilizado una función sigmoidea en la capa oculta con una función identidad en la capa de salida, en un cuarto modelo se ha utilizado una función sigmoidea en la capa oculta y una función tangente hiperbólica en la capa de salida y por último la misma función sigmoidea en la capa oculta con una función en la capa de salida sigmoidea. En todos los casos se utilizaron 16 variables independientes y una variable dependiente, en su arquitectura se seleccionó

intervalo que va desde 1 a 50 unidades en su capa oculta, en todos los casos se utilizó un tipo de entrenamiento en lote, con un tipo de gradiente escalado conjugado, obteniéndose los siguientes resultados de cada una de sus formas de activación, véase tabla 5.93.

Tabla 5.93. Test de resultados de entrenamiento de la RNA, con 16 unidades (variables), capa oculta

| F. de activación capa oculta | F. de activación capa de salida | Número de unidades en la capa oculta | Traning(en capa o) | Pruebas | Excluidos | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R2ajustado | Error estándar de estimación | Error cuadrático medio |
|------------------------------|---------------------------------|--------------------------------------|--------------------|---------|-----------|---------------------|----------------------|------------------------|------------|------------------------------|------------------------|
| Tangente hiperbólica | Identidad | 16 | 667 | 205 | 89 | 20 minutos | 95% | 91% | 91% | 296471 | 296131 |
| Tangente hiperbólica | Tangente hiperbólica | 16 | 698 | 189 | 74 | 80 minutos | 95% | 90% | 90% | 325477 | 325110 |
| Tangente hiperbólica | Sigmoidea | 16 | 665 | 208 | 88 | 10 minutos | 96% | 91% | 91% | 280030 | 279709 |
| Sigmoidea | Identidad | 16 | 677 | 211 | 73 | 10 minutos | 98% | 97% | 96% | 183206 | 183000 |
| Sigmoidea | Tangente hiperbólica | 16 | 677 | 201 | 83 | 60 minutos | 91% | 83% | 82% | 387704 | 387262 |
| Sigmoidea | Sigmoidea | 16 | 660 | 219 | 82 | 50 minutos | 95% | 90% | 90% | 319403 | 319040 |

Fuente: Elaboración propia

La tabla [5.93], se observan los resultados de las diferentes formas de activación de las redes neuronal y artificial. Los 6 modelos presentan buenos resultados, siendo el mejor de ellos la red con función de activación en la capa oculta sigmoidea y una función de activación en la capa de salida denominada identidad (lineal), con 16 unidades en la capa oculta.

Una vez validados los 6 modelos del segundo grupo y ya que se analizado su comportamiento en términos generales, se procede a la aplicación para estimar el precio de un inmueble en un caso concreto, es decir

mediante la ecuación implícita que se ha desarrollado en la red estimada.

Se trata de un inmueble que cuenta con 240 metros cuadrados de terreno, 250 metros cuadrados de construcción, 3.5 baños, 3 dormitorios, 0 años de antigüedad (nueva), 2 niveles, nivel 3 de infraestructura, nivel 3 de equipamiento, 3 garaje, 0 baños, 0 garaje2, 1 garaje3, 0 nivel de renta baja, 1 nivel de renta alta, 0 baños4, 0 baño 4.5, 0 baños5, 0 dormitorio6 y 0 numniveles3.

El precio observado es de \$3,182,500.00 y el estimado por la red de \$3,228,373.00, con un error de \$45,873.00 entre el precio observado y el precio estimado, resultando 1.44% por arriba del precio observado, con un coeficiente de correlación de 97%, coeficiente de determinación de 95%, y un r^2 ajustado de 95%.

VI. Resultados de la aplicación de los modelos de regresión y redes neuronales artificiales

6. Resultados de la aplicación de los modelos de regresión y redes neuronales artificiales

A continuación se detallan resultados obtenidos de los diferentes modelos aplicados.

Tabla 6.1. Resultados de los modelos de regresión y redes neuronales artificiales

| Regresión Lineal Múltiple | | | | | | | | | | | | |
|---------------------------|---------------------------|--------------------------------------|--------------------------|--------|-----------|---------------------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| Regresión Lineal Múltiple | Regresión Lineal Múltiple | Número de unidades en la capa oculta | Training (entrenamiento) | Prueba | Excluidos | Total de la muestra | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
| RLM | RLM | 0 | 961 | 0 | 24 | 985 | 961 | 95% | 90% | 90% | 314610 | 311814 |
| RLM | Log Precio | 0 | 936 | 0 | 0 | 936 | 936 | 94.5% | 89% | 89% | 405060 | 404627 |

| Redes Neuronales Artificiales (RNA) | | | | | | | | | | | | |
|-------------------------------------|---------------------------------|--------------------------------------|--------------------------|--------|-----------|---------------------|---------------------|----------------------|------------------------|-------------------------|------------------------------|------------------------|
| F. de activación capa oculta | F. de activación capa de salida | Número de unidades en la capa oculta | Training (entrenamiento) | Prueba | Excluidos | Total de la muestra | T. de entrenamiento | Coef. de correlación | Coef. de determinación | R ² ajustado | Error estándar de estimación | Error cuadrático medio |
| Sigmoidea | Identidad | 16 | 677 | 221 | 73 | 961 | 50 minutos | 97% | 94% | 94% | 234693 | 234424 |
| Sigmoidea | Identidad | 16 | 677 | 211 | 73 | 961 | 10 minutos | 98% | 97% | 96% | 183206 | 234424 |

Fuente. Elaboración propia

En la tabla [6.1] se detalla en las columnas 1 y 2 el tipo de modelo para el caso de modelos de regresión y para el caso de las RNA la función de activación en la capa oculta y función de activación en la capa de salida, en la columna tres sólo en RNA el número de unidades en la capa oculta, columna cuatro número de muestra para el entrenamiento, columna cinco sólo para el caso de RNA el número de muestras para el entrenamiento de la red, columna seis se observan las muestras excluidas en todos los modelos por ser atípicas, columna siete el número de muestras total utilizadas en el estudio, columna ocho el

coeficiente de correlación de las variables, columna nueve coeficiente de determinación, columna diez r^2 ajustado, columna once el error estándar de estimación y columna doce el error cuadrático medio.

En los resultados de la tabla antes mencionada se observa que el modelo de red neuronal con 16 unidades en la capa oculta, supera en cada una de sus mediciones estadísticas a la regresión múltiple, logarítmica y al otro modelo de RNA propuesto.

En la siguiente tabla se muestran los resultados de la estimación del precio de un inmueble fuera de la muestra.

Para ello se han utilizado en el modelo de regresión lineal 985 observaciones, de las cuales se han eliminado 24, por resultar atípicas y 16 variables independientes, en el segundo modelo, solo que ahora tomando como el logaritmo del precio de la variable dependiente se han utilizado 936 observaciones y 10 variables independientes. En cambio en los modelos de redes neuronales se han utilizado las 961 observaciones, se utilizado 677 para el entrenamiento de la red, 112 para la validación o prueba de los modelos y se han excluido 73.

Tabla 6.2. Resultados de estimación del precio por modelos de regresión y de redes neuronales artificiales

| Regresión lineal múltiple | | | | | | | | |
|---------------------------|---------------------------|----------------------|------------------------|-------------------------|------------------|-------------------------------|----------------|------------------------|
| Regresión Lineal Múltiple | Regresión Lineal Múltiple | Coef. de correlación | Coef. de determinación | R ² ajustado | Precio observado | Precio estimado por el modelo | Error absoluto | Error cuadrático medio |
| RLM | RLM | 95% | 90% | 90% | 3,182,500 | 2,933,226 | 249274 | 499 |
| RLM | Log. Precio | 94.5% | 89% | 89% | 3,182,500 | 2,103,778 | 1078722 | 1039 |

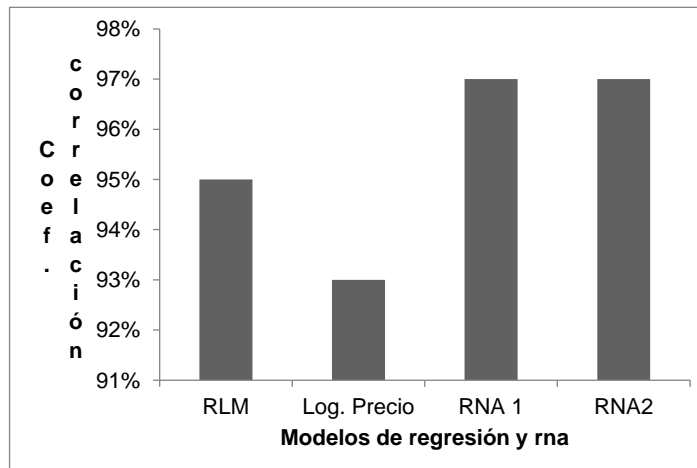
| Redes neuronales artificiales (RNA) | | | | | | | | |
|-------------------------------------|------------------------------------|----------------------|------------------------|-------------------------|------------------|-------------------------------|----------------|------------------|
| F. de activación de capa oculta | F. de activación de capa de salida | Coef. de correlación | Coef. de determinación | R ² ajustado | Precio observado | Precio estimado por el modelo | Error absoluto | Error cuadrático |
| Sigmoidea | Identidad | 97% | 94% | 94% | 3,182,500 | 3,055,124 | 127376 | 357 |
| Sigmoidea | Identidad | 97% | 95% | 95% | 3,182,500 | 3,228,373 | 45873 | 214 |

Fuente. Elaboración propia.

La tabla [6.2] muestra los resultados de la estimación del precio por los diferentes modelos, en la columna uno y dos los diferentes modelos y el tipo de función, en la columna tres se observan los resultados del coeficiente de correlación, en la columna cuatro el coeficiente de determinación, en la columna cinco el r^2 ajustado, columna seis el precio observado, columna siete el precio estimado por el modelo, columna ocho error absoluto y en la columna nueve el error cuadrático.

En las siguientes figuras se observan algunos gráficos que representan las bondades de cada uno de los modelos aplicados.

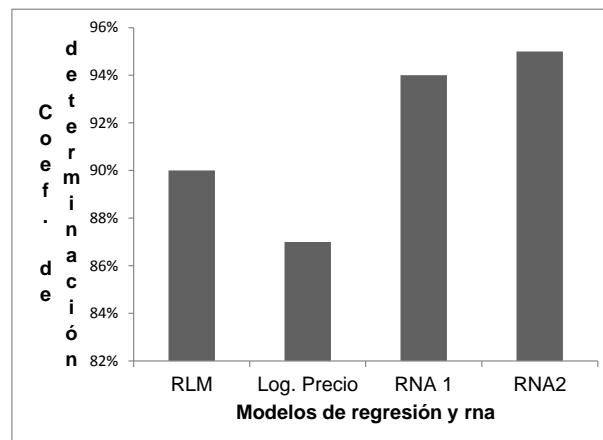
Figura 6.3. Coeficiente de correlación de los modelos utilizados



Fuente: Elaboración propia

La figura [6.3] representa los resultados del coeficiente de correlación entre el precio observado y el precio estimado por el modelo, explica que tanta relación existe entre las variables independientes y la variable dependiente.

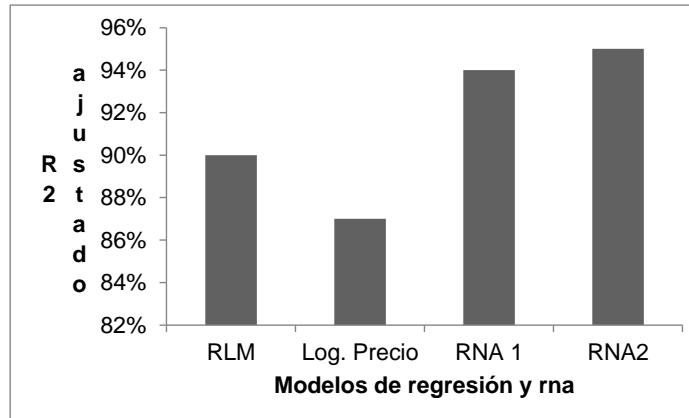
Figura 6.4. Coeficiente de determinación



Fuente: Elaboración propia

En la figura [6.4] se observan resultados del coeficiente de determinación en cada uno de los modelos, de regresión y redes.

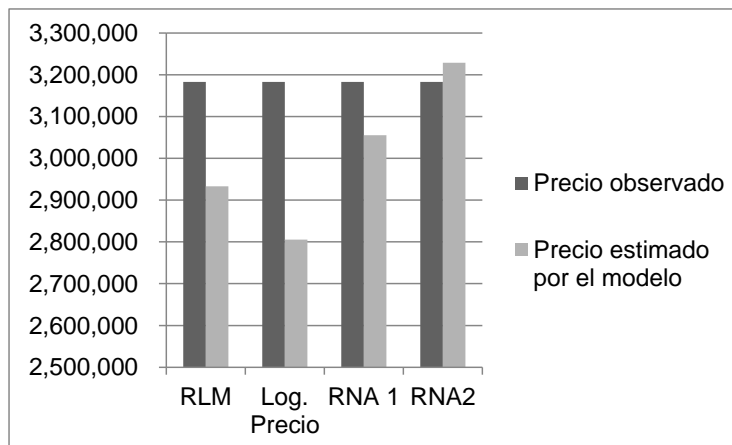
Figura 6.5. R² ajustado



Fuente: Elaboración propia

La figura [6.5] representa el r^2 ajustado de cada uno de los modelos utilizados.

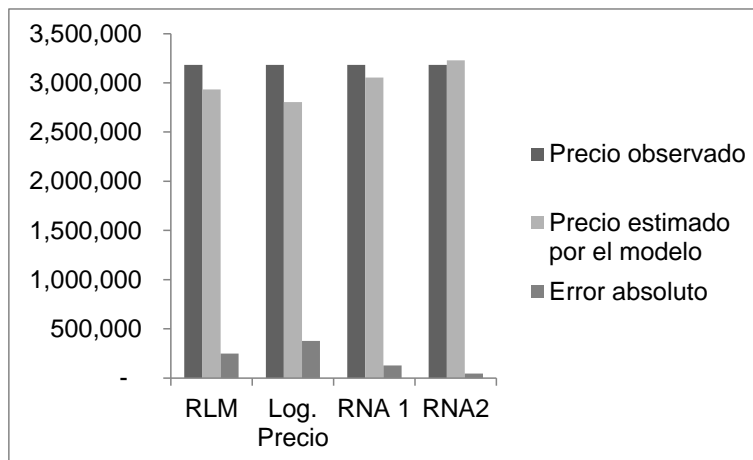
Figura 6.6. Precio observado y precio estimado por los modelos



Fuente: Elaboración propia

En la figura 6.6 se observan los resultados del precio observado y el precio estimado por cada uno de los modelos aplicados.

Figura 6.7. Precio observado, precio estimado y error absoluto



Fuente: Elaboración propia

En la figura [6.7] se representan los resultados entre el precio observado, precio estimado por cada modelo y el error absoluto entre ambos precios.

VII. Conclusiones finales y trabajos futuros

7. Conclusiones finales y trabajos futuros

La valoración de activos, y en este rubro se encuentran los inmuebles, ha sido de gran importancia desde tiempos remotos, sirve de base para la toma de decisiones en diferentes instrumentos financieros. La sociedad, gobiernos, instituciones financieras y empresas exigen cada día mejores estimaciones de precios. Los modelos de regresión en sus diferentes formas, y las redes neuronales artificiales como un activo metodológico más reciente, convergen en una nueva forma de modelizar y predecir el precio de los inmuebles con resultados muy prometedores.

Tras analizar el mayor y más relevante número de variables independientes, los profesionales del sector de la tasación persiguen obtener el precio de los inmuebles.

A modo de resumen, se expondrán las principales conclusiones de esta tesis, obtenidas a través de sus diferentes capítulos.

En el capítulo dos, se habla de las aportaciones que han hecho distintos y notables autores para una mejora de la estimación de precios de los inmuebles, mediante modelos econométricos y de redes neuronales artificiales.

En el tercer capítulo se han presentado los principales modelos de la familia econométrica (regresión lineal y no lineal). Se detallan modelos univariantes y multivariantes, siendo este último el de mayor aplicación por la cantidad de variables independientes que se incluyen en los modelos.

Para estimar el precio de los inmuebles se han utilizado el modelo de regresión lineal múltiple y el modelo de regresión logarítmica, obteniéndose el mejor resultado con el primero de estos dos modelos pese a los evidentes problemas de heterocedasticidad, comúnmente señalados por la mayoría de investigadores de este campo.

En el cuarto capítulo se han expuesto definiciones y funcionamiento de las redes neuronales artificiales en general y de la red neuronal *back propagation* en particular, pues es la que se usa en esta investigación. La selección de esta red se debe a su capacidad de autoadaptar los pesos durante el proceso de aprendizaje, así como su

generalización y flexibilidad, logrando soluciones favorables con muestras grandes o reducidas.

En esta investigación se ha utilizado el entrenamiento supervisado, con la ventaja de que se pueden verificar los resultados predictivos y además compararlos con los precios observados.

La regla delta generalizada es el algoritmo de aprendizaje de la red neuronal seleccionada, que utiliza el mecanismo para corregir el error e ir modificando los diferentes pesos de las diferentes neuronas, hasta obtener un resultado óptimo o deseado, pero además que cumpla con los parámetros estadísticos ideales.

En el capítulo cinco se hace una comparativa entre los modelos de regresión y las redes neuronales artificiales en específico el perceptrón multicapa.

Para llevar a cabo esta comparación, se ha obtenido una muestra de 985 inmuebles de una importante ciudad mexicana, compuesta por 688 inmuebles de venta consumada y 297 con precios de oferta con un factor de negociación de 5%, para de esta manera trabajar con inmuebles de venta consumada. Precisamente la

aplicación de estos modelos a un país como México es el principal aporte de esta tesis. El mercado inmobiliario de este país presenta una serie de particularidades, al igual que otros del espectro latinoamericano, que difieren de forma singular de los comúnmente observables en otras regiones del mundo. Así, la práctica totalidad de las investigaciones previas analizadas desarrollaban sus modelos sobre países europeos o en los Estados Unidos de América, donde la realidad inmobiliaria es diametralmente contrapuesta a la de México. De ahí el interés en conocer el comportamiento y aplicabilidad de los modelos econométricos y de red neuronal en un mercado inmobiliario como el mexicano.

Posteriormente se realizaron las pruebas estadísticas de normalidad para eliminar datos atípicos, cumpliendo con los rangos $[\mu+3\sigma; \mu-3\sigma]$. Al eliminar los datos atípicos y las variables irrelevantes, la muestra se reduce a 961 observaciones, tras lo que se procedió a la aplicación del primer modelo de regresión, hasta encontrar el mejor ajuste, obteniendo resultados de coeficiente de correlación de 95%, 90% de coeficiente de determinación, y 90% de r^2 ajustado.

Enseguida se aplicó un modelo logarítmico

exactamente igual que el anterior, sólo que en este caso se tomó como variable dependiente el logaritmo de los precios de la muestra. El motivo es que la relación entre el precio y la superficie suele presentar problemas de heterocedasticidad, cuestión ya reportada en investigaciones previas sobre países diferentes a México. Así, se obtuvieron los siguientes resultados: 93% de coeficiente de correlación, 87% de coeficiente de determinación, y 87% de r^2 ajustado.

Por otro lado las redes neuronales artificiales, en específico perceptrón multicapa, se aplicaron a dos grupos: en un primer grupo 6 modelos utilizando 16 variables independientes y 1 variable dependiente con las mismas 961 muestras que se utilizaron en los modelos de regresión para no dar ventaja a ningún modelo y así poder comparar los resultados. Se utilizaron una red de función de activación a la entrada función hiperbólica, con una función de activación en su salida lineal, hiperbólica y sigmoideal; posteriormente una función de activación en la entrada denominada sigmoideal y funciones de activación en su salida lineal, hiperbólica y sigmoideal. Del total de las observaciones, el 70% se tomaron para el entrenamiento (*training*) de la red y el 30% restante para la prueba (*test*).

En la arquitectura de la red se seleccionó la opción automática con un número máximo de unidades en la capa oculta de 16. Se optó por el entrenamiento en lote e intervalos de tiempo de entrenamiento de la red de 10 minutos, con incrementos de 10, hasta 100 minutos, siendo el mejor de estos primeros 6 la red con función de activación en la capa oculta sigmoidea y una función de activación en la capa de salida denominada identidad (lineal), tiempo de entrenamiento 50 minutos, coeficiente de correlación entre variables de 97%, coeficiente de determinación de 94%, y r^2 ajustado de 94%.

El otro grupo de 6 modelos de redes neuronales artificiales perceptrón multicapa, en estos casos se aplicaron con las siguientes condicionantes: utilizando 16 variables independientes y 1 variable dependiente con las mismas 961 observaciones que se utilizaron en los modelos de regresión para no dar ventaja a ningún modelo y así poder comparar los resultados. Se utilizaron una red de función de activación a la entrada función hiperbólica con una función de activación en su salida lineal, hiperbólica y sigmoidea, luego una función de activación en la entrada denominada sigmoidea y funciones de activación en su salida lineal, hiperbólica y sigmoidea. Del total de las observaciones, de nuevo el 70% se tomaron para el

entrenamiento de la red y el 30% restante para el *test*.

Para la arquitectura de la red se seleccionó la opción automática con un número máximo de unidades en la capa oculta de 50 y una arquitectura personalizada de 16 unidades en la capa oculta. Se optó por el entrenamiento en lote e intervalos de tiempo de entrenamiento de la red con tiempos de 10 minutos, con incrementos de 10, hasta 100 minutos, siendo el mejor de estos segundos 6 modelos la red con función de activación en la capa oculta sigmoidea y una función de activación en la capa de salida denominada identidad (lineal), con 16 unidades en la capa oculta de la arquitectura personalizada, tiempo de entrenamiento 10 minutos, coeficiente de correlación entre variables de 98%, coeficiente de determinación de 97%, y r^2 ajustado de 96%.

Una primera conclusión es que la red multicapa con una función de activación en la capa oculta sigmoidea y salida lineal con arquitectura personalizada de 16 unidades, y con tiempo de entrenamiento de 10 minutos, supera a los modelos de regresión lineal múltiple en 3% del coeficiente de correlación, 7% en el coeficiente de determinación, 6% en el r^2 ajustado. También supera al modelo de regresión logarítmica en 5% el coeficiente de

correlación, 10% en coeficiente de determinación, 9% en r^2 ajustado.

La diferencia entre los dos grupos de modelos de redes neuronal multicapa es que en los primeros 6 modelos se utilizó una arquitectura automática con 16 unidades máxima en la capa oculta, y en el segundo en la 50 unidades máxima en la capa oculta. Adicionalmente en esta segunda se empleó una arquitectura personalizada con 16 unidades en la capa oculta.

Con base en los resultados, en línea con los resultados de otros investigadores, se concluye que para el caso de México también las redes neuronales son capaces de explicar mejor la variabilidad del precio de los inmuebles que los modelos de regresión múltiple.

Una segunda conclusión al estimar el precio de un inmueble fuera de la muestra para validar la metodología, es que la red multicapa con una función de activación en la capa oculta sigmoidea y salida lineal con arquitectura personalizada de 16 unidades, supera a los modelos de regresión lineal múltiple en 2% del coeficiente de correlación, 5% en el coeficiente de determinación, 5% en el r^2 ajustado. Y también supera al modelo de regresión

logarítmica 3% en el coeficiente de correlación, 8% en el coeficiente de determinación, y 8% en r^2 ajustado.

Algunos investigadores mencionados al principio de esta tesis doctoral afirman que los modelos de redes neuronales artificiales ofrecen errores entre 5 % y 10%, mientras que los modelos de regresión entre 10% y 15%. Lo cual está en línea con los resultados obtenidos en esta tesis sobre un municipio mexicano, evidenciando que el comportamiento de los precios en este país puede ser modelizado de forma exitosa a través de las redes neuronales, no evidenciándose diferencias significativas con respecto de los resultados obtenidos en otras áreas geográficas.

Finalmente, se concluye la confirmación de la hipótesis de partida donde se afirma que las redes neuronales artificiales pueden superar a los modelos de regresión y explicar de forma más precisa los precios de los inmuebles en México.

VIII. Anexos

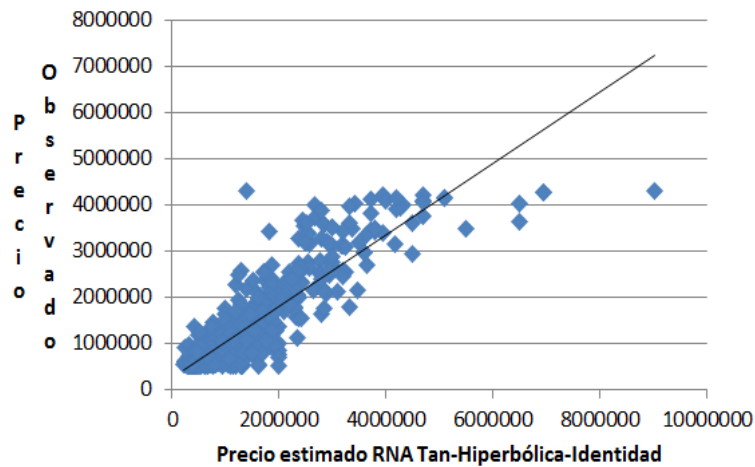
Valoración de inmuebles urbanos: Comparativa modelo regresión multivariable versus redes neuronales artificiales para la ciudad de Morelia, Michoacán, México.

D. José Carlos Preciado Carrillo
Dirigido por Dr. D. Francisco Guijarro Martínez

335

8. Anexos

Figura 8.1. Gráfica del precio observado y precio estimado por la RNA

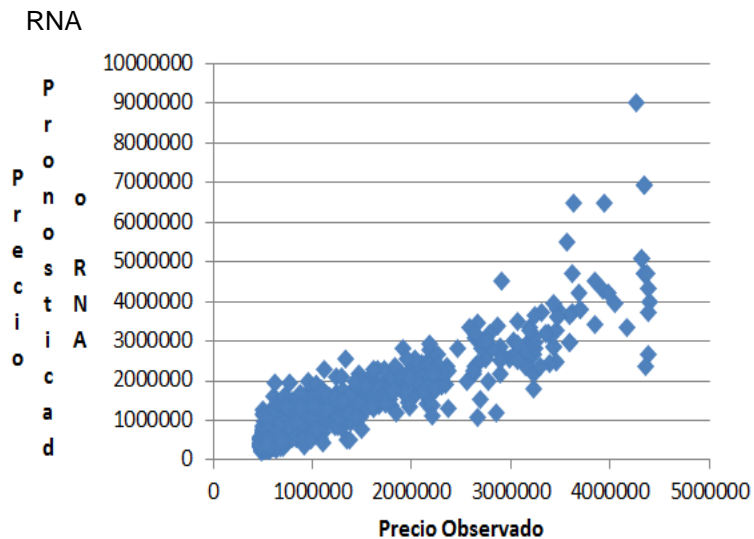


Fuente: Elaboración propia.

La figura [8.1] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con una entrada en la red oculta tangente hiperbólica-identidad, con tiempo de 20 minutos, utilizando las

muestras de entrenamiento de la red (*training*): 882 de un total de 961 observaciones.

Figura 8.2. Gráfica de precio observado y precio estimado por

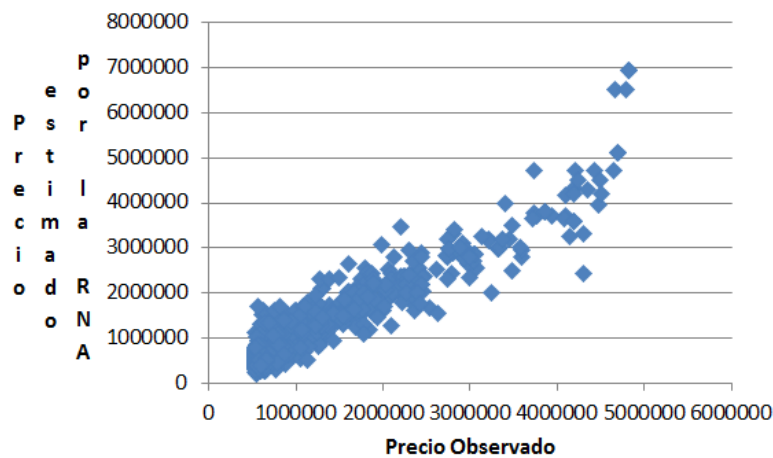


Fuente: Elaboración propia

La figura [8.2], contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA, con una función de entrada en la red oculta tangente hiperbólica y con salida tangente hiperbólica, con tiempo

de 80 minutos, se utilizaron 671 muestras de entrenamiento de la red (*training*), 223 para la prueba o validación, y se eliminaron del modelo 70 observaciones, por ser atípicas, de un total de 961.

Figura 8.3. Gráfica de precio observado y precio estimado por RNA

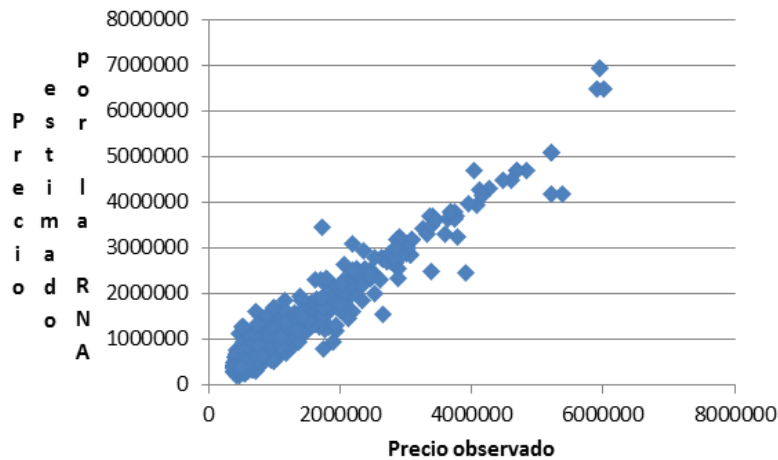


Fuente: Elaboración propia

La figura [8.3] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida sigmoïdal, con tiempo de 50 minutos, utilizando 654 muestras de entrenamiento para la red (*training*), 219 para

la prueba o validación, eliminando del modelo 88 observaciones por ser atípicas de un total de 961.

Figura 8.4. Gráfica de precio observado y precio estimado por RNA



Fuente: Elaboración propia

La figura [8.4] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA, con entrada en la red oculta sigmoidea-salida identidad, con tiempo de 50 minutos, utilizando 677 muestras de entrenamiento de la red (training), 211 para la prueba o validación, eliminando del modelo 73 observaciones por ser atípicas de un total de 961.

Figura 8.5. Gráfica de precio observado y precio estimado por RNA

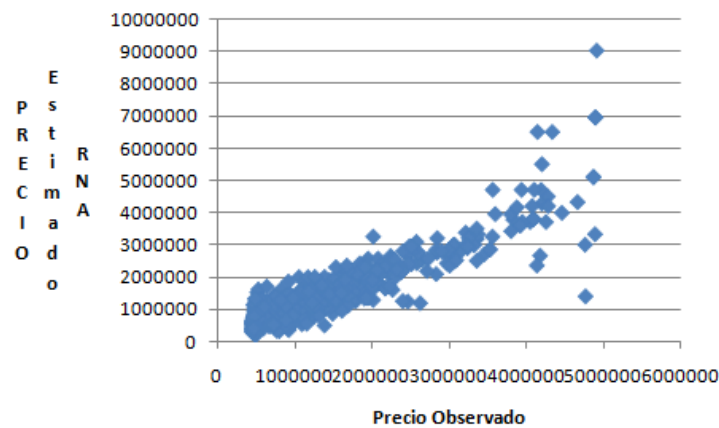
| | | | | |
|-----------------|---------------------------------|----|-------------------|-----|
| Capa de entrada | Factores | 1 | Terreno | |
| | | 2 | construcción | |
| | | 3 | Antigüedad | |
| | Covariables | 1 | Nivelrentaalta | |
| | | 2 | Infraestructura | |
| | | 3 | Garage | |
| | | 4 | Garage2 | |
| | | 5 | Numniveles3 | |
| | | 6 | Numbaños5 | |
| | | 7 | Numbaños4 | |
| | | 8 | Nivelrentabaja | |
| | | 9 | Numdormitorio6 | |
| | | 10 | Numbaños2 | |
| | | 11 | Garage3 | |
| | | 12 | Numbaños4.5 | |
| | | 13 | Equiurbano | |
| | Número de unidades ^a | | | 495 |
| | Método de cambio de escala | | Estandarizados | |
| Capas ocultas | Número de capas ocultas | | | 1 |
| | Número de unidades en la | | | 1 |
| | Función de activación | | Sigmoide | |
| Capa de salida | Variables dependientes | 1 | Precio | |
| | Número de unidades | | | 1 |
| | Método de cambio de escala | | Normalizada | |
| | Función de activación | | Sigmoide | |
| | Función de error | | Suma de cuadrados | |

Fuente: Elaboración propia

La figura [8.5] contiene los resultados de las variables o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida identidad, tiempo de entrenamiento de la red 50 minutos,

utilizando 677 observaciones de entrenamiento de la red (*training*), 211 para la prueba o validación, eliminando del modelo 73 observaciones, por ser atípicas, de un total de 961.

Figura 8.6. Gráfica de precio observado y precio estimado por RNA



Fuente: Elaboración propia

La figura [8.6] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida tangente hiperbólica, con tiempo de 50 minutos, utilizando 660 muestras de entrenamiento de la red (*training*), 219 para la prueba o validación, eliminando del modelo 82 observaciones, por ser atípicas, de un total de 961.

Figura 8.7. Gráfica de precio observado y precio estimado por RNA

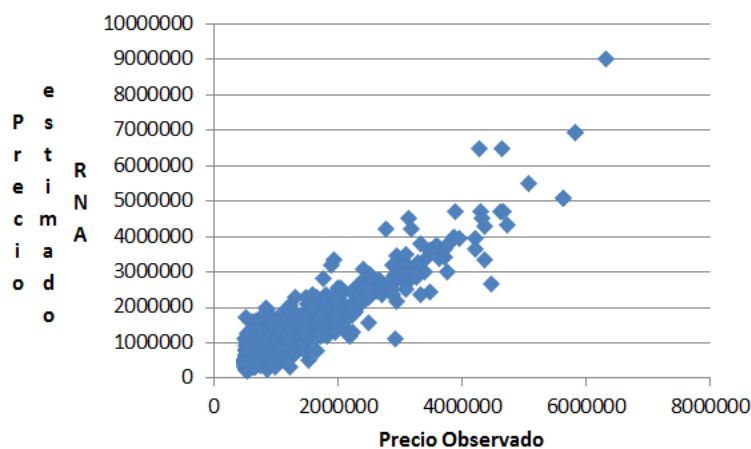
| | | | |
|-----------------|--------------------|----|----------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antiquedad |
| | Covariables | 1 | Infraestructur |
| | | 2 | Equiurbano |
| | | 3 | Garaqe |
| | | 4 | Numbaños2 |
| | | 5 | Garaqe2 |
| | | 6 | Garaqe3 |
| | | 7 | Nivelrentabaj |
| | | 8 | Nivelrentaalta |
| | | 9 | Numbaños4 |
| | | 10 | Numbaños4. |
| | | 11 | Numbaños5 |
| | | 12 | Numdormitori |
| | | | o6 |
| | | 13 | Numniveles3 |
| | Número de | | 497 |
| | Método de cambio | | Estandarizad |
| | de escala para las | | os |
| Capas ocultas | Número de capas | | 1 |
| | Número de | | 1 |
| | Función de | | Sigmoide |
| Capa de salida | Variables | 1 | Precio |
| | dependientes | | |
| | Número de | | 1 |
| | Método de cambio | | Corregidos |
| | de escala para las | | normalizados |
| | Función de | | Tangente |
| | activación | | hiperbólica |
| | Función de error | | Suma de |
| | | | cuadrados |

Fuente: Elaboración propia

La figura [8.7] contiene los resultados de las variables o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la

RNA con entrada en la red oculta sigmoidea-salida identidad, tiempo de entrenamiento de la red 50 minutos, utilizando 660 muestras de entrenamiento de la red (*training*), 219 para la prueba o validación, eliminando del modelo 82 observaciones, por ser atípicas, de un total de 961.

Figura 8.8. Gráfica de precio observado y precio estimado por RNA



Fuente: Elaboración propia

La figura [8.8] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida sigmoidea, con un tiempo de 80 minutos, utilizando 671 muestras de entrenamiento de la red (*training*), 223 para la prueba o validación, eliminando del modelo 67 observaciones, por ser atípicas, de un total de 961.

Figura 8.9. Gráfica de precio observado y precio estimado por la RNA

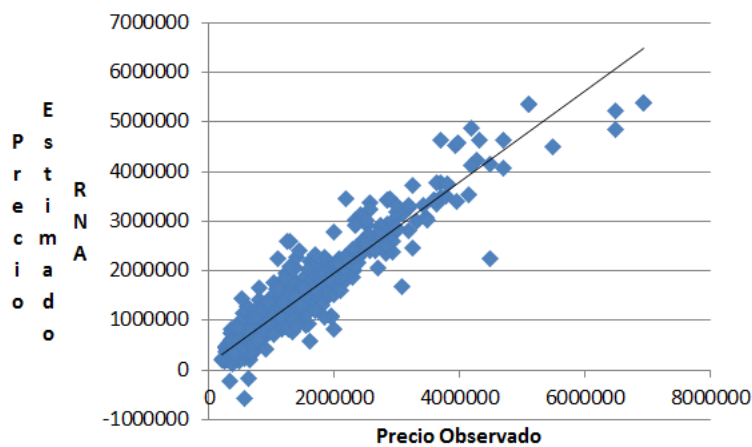
| | | | |
|-----------------|---------------------------------|----------------|-----------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Infraestructura |
| | | 2 | Equiurbano |
| | | 3 | Garage |
| | | 4 | Numbaños2 |
| | | 5 | Garage2 |
| | | 6 | Garage3 |
| | | 7 | Nivelrentabaja |
| | | 8 | Nivelrentaalta |
| | | 9 | Numbaños4 |
| | | 10 | Numbaños4.5 |
| | 11 | Numbaños5 | |
| | 12 | Numdormitorio | |
| | 13 | Numniveles3 | |
| | Número de unidades ^a | 497 | |
| | Método de cambio de | Estandarizados | |
| Capas ocultas | Número de capas | 1 | |
| | Número de unidades | 1 | |
| | Función de activación | Sigmoide | |
| Capa de salida | Variables | 1 | Precio |
| | Número de unidades | 1 | |
| | Método de cambio de | Normalizada | |
| | Función de activación | Sigmoide | |
| | Función de error | Suma de | |

Fuente: Elaboración propia

La figura [8.9] contiene los resultados de las variables o factores y covariables en la capa oculta de la

red, los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida sigmoidea, tiempo de entrenamiento de la red 80 minutos, utilizando 671 muestras de entrenamiento de la red (*training*), 223 para la prueba o validación, eliminando el modelo 67 observaciones por ser atípicas, de un total de 961.

Figura 8.10. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.10], contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida identidad (lineal), con tiempo de 20 minutos, utilizando 667 muestras de entrenamiento de la red (*training*), 205 para la prueba o validación,

eliminando el modelo 89 observaciones por ser atípicas, de un total de 961.

La figura [8.11] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida identidad (lineal), con tiempo de 20 minutos, con 16 variables independientes en la capa oculta, 667 muestras de entrenamiento de la red (*training*), 205 para la prueba o validación, eliminando el modelo 89 observaciones por ser atípicas, de un total de 961.

Figura 8.11. Gráfica de precio observado y precio estimado por la RNA

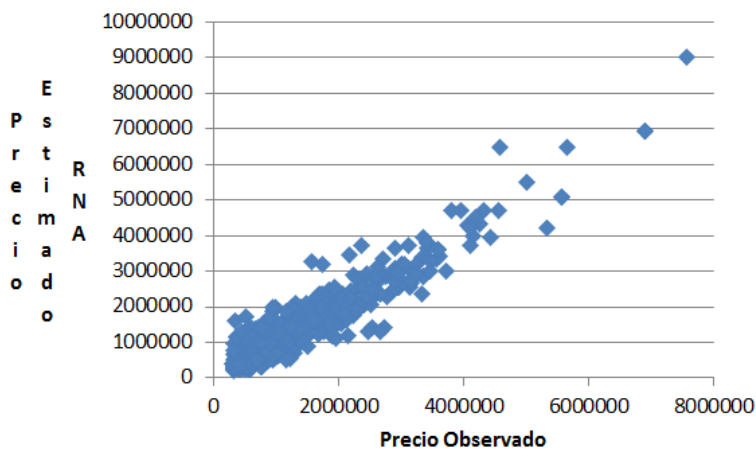
| | | | |
|-----------------|---|-------------------|----------------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Nivelrentaalta |
| | | 2 | Infraestructura |
| | | 3 | Garage |
| | | 4 | Garage2 |
| | | 5 | Numniveles3 |
| | | 6 | Numbaños5 |
| | | 7 | Numbaños4 |
| | | 8 | Nivelrentabaja |
| | | 9 | Numdormitorio6 |
| | | 10 | Numbaños2 |
| 11 | Garage3 | | |
| 12 | Numbaños4.5 | | |
| | 13 | Equiurbano | |
| | Número de unidades ^a | 474 | |
| | Método de cambio de escala para las covariables | Estandarizados | |
| Capas ocultas | Número de capas ocultas | 1 | |
| | Número de unidades en la Función de activación | 16 | |
| Capa de salida | Variables dependiente | 1 | Tangente hiperbólica |
| | | | Precio |
| | Número de unidades | 1 | |
| | Método de cambio de escala para las dependientes de | Estandarizados | |
| | Función de activación | Identidad | |
| | Función de error | Suma de cuadrados | |

Fuente: Elaboración propia

La figura [8.11] contiene los resultados de las variables o factores y covariables en la capa oculta de la

red, los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida sigmoidea, tiempo de entrenamiento de la red 20 minutos, utilizando 667 muestras de entrenamiento de la red (*training*), 205 para la prueba o validación, eliminando el modelo 89 observaciones por ser atípicas, de un total de 961.

Figura 8.12. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.12] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida tangente hiperbólica, con tiempo de 80 minutos, utilizando 698 muestras de entrenamiento de la red (*training*), 189

para la prueba o validación, eliminando el modelo 74 observaciones por ser atípicas, de un total de 961.

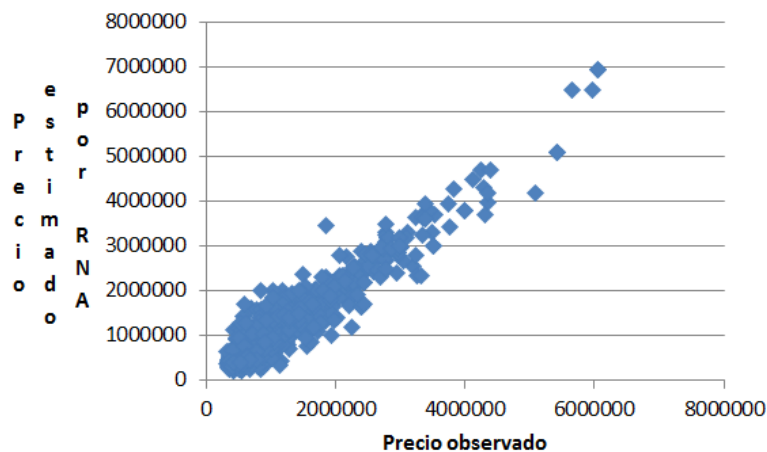
Figura 8.13. Tabla de precio observado y precio estimado por la RNA.

| | | | |
|-----------------|-------------------------------------|-------------------------|-----------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Nivelrentaalta |
| | | 2 | Infraestructura |
| | | 3 | Garage |
| | | 4 | Garage2 |
| | | 5 | Numniveles3 |
| | | 6 | Numbaños5 |
| | | 7 | Numbaños4 |
| | | 8 | Nivelrentabaja |
| | | 9 | Numdormitorio6 |
| | | 10 | Numbaños2 |
| 11 | Garage3 | | |
| 12 | Numbaños4.5 | | |
| 13 | Equiurbano | | |
| | Número de unidades ^a | 483 | |
| | Método de cambio de | Estandarizados | |
| Capas ocultas | Número de capas ocultas | 1 | |
| | Número de unidades en la | 16 | |
| | Función de activación | Tangente hiperbólica | |
| Capa de salida | Variables dependientes | 1 | Precio |
| | Número de unidades | 1 | |
| | Método de cambio de escala para las | Corregidos normalizados | |
| | Función de activación | Tangente hiperbólica | |
| | Función de error | Suma de cuadrados | |

Fuente: Elaboración propia.

La figura [8.13] contiene los resultados de las variable o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida tangente hiperbólica, tiempo de entrenamiento de la red 80 minutos, utilizando 698 muestras de entrenamiento de la red (*training*), 189 para la prueba o validación, eliminando el modelo 74 observaciones por ser atípicas, de un total de 961.

Figura 8.14. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.14] representa la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta

tangente hiperbólica-salida sigmoidea, con tiempo de 10 minutos, utilizando 665 muestras de entrenamiento de la red (*training*), 208 para la prueba o validación, eliminando el modelo 88 observaciones por ser atípicas, de un total de 961.

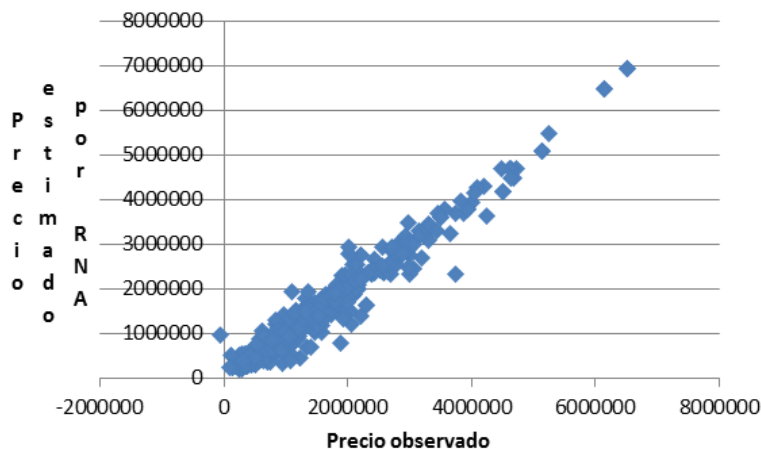
Figura 8.15. Gráfica de precio observado y precio estimado por la RNA

| | | | |
|-----------------|---------------------------------|----------------------|-----------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Nivelrentaalta |
| | | 2 | Infraestructura |
| | | 3 | Garage |
| | | 4 | Garage2 |
| | | 5 | Numniveles3 |
| | | 6 | Numbaños5 |
| | | 7 | Numbaños4 |
| | | 8 | Nivelrentabaja |
| | | 9 | Numdormitorio6 |
| | | 10 | Numbaños2 |
| 11 | Garage3 | | |
| 12 | Numbaños4.5 | | |
| 13 | Equiurbano | | |
| | Número de unidades ^a | 466 | |
| | Método de cambio de | Estandarizados | |
| Capas ocultas | Número de capas ocultas | 1 | |
| | Número de unidades en la | 16 | |
| | Función de activación | Tangente hiperbólica | |
| Capa de salida | Variables | 1 | Precio |
| | | Número de unidades | 1 |
| | Método de cambio de | Normalizada | |
| | Función de activación | Sigmoide | |
| | Función de error | Suma de cuadrados | |

Fuente: Elaboración propia

La figura [8.15] contiene los resultados de las variables o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida sigmoidea, tiempo de entrenamiento de la red 10 minutos, utilizando 665 muestras de entrenamiento de la red (*training*), 208 para la prueba o validación, eliminando el modelo 88 observaciones por ser atípicas, de un total de 961.

Figura 8.16. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.16] representa la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta

denominada sigmoidea-salida identidad (lineal), con tiempo de 10 minutos, utilizando 677 muestras de entrenamiento de la red (*training*), 211 para la prueba o validación, eliminando el modelo 73 observaciones por ser atípicas, de un total de 961.

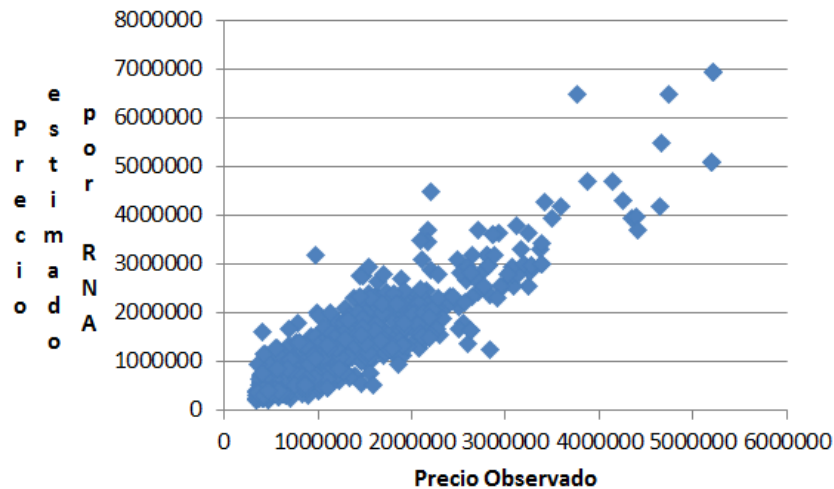
Figura 8.17. Gráfica de precio observado y precio estimado por la RNA

| | | | | |
|---------------------------------|-------------------------------|-------------------|---------|-----------------|
| Capa de entrada | Factores | 1 | Terreno | |
| | | 2 | | |
| | | 3 | | |
| | Covariables | 1 | | Antigüedad |
| | | 2 | | Nivelrentaalta |
| | | 3 | | Infraestructura |
| | | 4 | | Garage |
| | | 5 | | Garage2 |
| | | 6 | | Numniveles3 |
| | | 7 | | Numbaños5 |
| | | 8 | | Nivelrentabaja |
| | | 9 | | Numdormitorio6 |
| 10 | Numbaños2 | | | |
| 11 | Garage3 | | | |
| 12 | Numbaños4.5 | | | |
| Número de unidades ^a | | 465 | | |
| | Método de cambio de escala | Estandarizados | | |
| | | | | |
| Capas ocultas | Número de capas ocultas | 1 | | |
| | Número de unidades en la capa | 16 | | |
| | Función de activación | Sigmoide | | |
| Capa de salida | Variables | 1 | Precio | |
| | Número de unidades | 1 | | |
| | Método de cambio de escala | Estandarizados | | |
| | Función de activación | Identidad | | |
| | Función de error | Suma de cuadrados | | |

Fuente: Elaboración propia

La figura [8.17] contiene los resultados de las variable o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta tangente hiperbólica-salida sigmoidea, tiempo de entrenamiento de la red 10 minutos, utilizando 677 muestras de entrenamiento de la red (*training*), 211 para la prueba o validación, eliminando el modelo 73 observaciones por ser atípicas, de un total de 961.

Figura 8.18. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.18] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta

denominada sigmoidea-salida identidad (lineal), con tiempo de 60 minutos, utilizando 677 muestras de entrenamiento de la red (*training*), 201 para la prueba o validación, eliminando el modelo 83 observaciones por ser atípicas, de un total de 961.

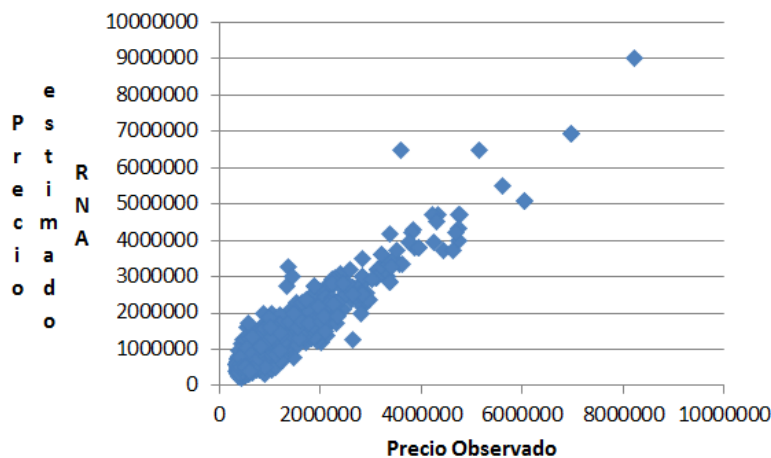
Figura 8.19. Gráfica de precio observado y precio estimado por la RNA

| | | | |
|-----------------|---|-------------------------|-----------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Infraestructura |
| | | 2 | Equiurbano |
| | | 3 | Garage |
| | | 4 | Numbaños2 |
| | | 5 | Garage2 |
| | | 6 | Garage3 |
| | | 7 | Nivelrentabaja |
| | | 8 | Nivelrentaalta |
| | | 9 | Numbaños4 |
| | | 10 | Numbaños4.5 |
| 11 | Numbaños5 | | |
| 12 | Numdormitorio6 | | |
| | 13 | Numniveles3 | |
| | Número de unidades* | 488 | |
| | Método de cambio de escala | Estandarizados | |
| Capas ocultas | Número de capas ocultas | 1 | |
| | Número de unidades en la | 16 | |
| | Función de activación | Sigmoide | |
| Capa de salida | Variables | 1 | Precio |
| | Número de unidades | 1 | |
| | Método de cambio de escala para las dependientes de | Corregidos normalizados | |
| | Función de activación | Tangente hiperbólica | |
| | Función de error | Suma de cuadrados | |

Fuente: Elaboración propia

La figura [8.19] contiene los resultados de las variables o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida tangente hiperbólica, tiempo de entrenamiento de la red 60 minutos, utilizando 677 muestras de entrenamiento de la red (*training*), 201 para la prueba o validación, eliminando el modelo 83 observaciones por ser atípicas, de un total de 961.

Figura 8.20. Gráfica de precio observado y precio estimado por la RNA



Fuente: Elaboración propia

La figura [8.20] contiene la gráfica de resultados de los precios observados y los precios estimados por la RNA con entrada en la red oculta

sigmoidea-salida sigmoidea, con un tiempo de 50 minutos, utilizando 660 muestras de entrenamiento de la red (*training*), 219 para la prueba o validación, eliminando el modelo 82 observaciones por ser atípicas, de un total de 961.

Figura 8.21. Gráfica de precio observado y precio estimado por la RNA

| | | | |
|-----------------|---------------------------------|----------------|-------------------|
| Capa de entrada | Factores | 1 | Terreno |
| | | 2 | construcción |
| | | 3 | Antigüedad |
| | Covariables | 1 | Infraestructura |
| | | 2 | Equiurbano |
| | | 3 | Garage |
| | | 4 | Numbaños2 |
| | | 5 | Garage2 |
| | | 6 | Garage3 |
| | | 7 | Nivelrentabaja |
| | | 8 | Nivelrentaalta |
| | | 9 | Numbaños4 |
| | | 10 | Numbaños4.5 |
| | 11 | Numbaños5 | |
| | 12 | Numdormitorio6 | |
| | 13 | Numniveles3 | |
| | Número de unidades ^a | 480 | |
| | Método de cambio de | | Estandarizados |
| Capas ocultas | Número de capas | 1 | |
| | Número de unidades en | 16 | |
| | Función de activación | | Sigmoide |
| Capa de salida | Variables | 1 | Precio |
| | Número de unidades | 1 | |
| | Método de cambio de | | Normalizada |
| | Función de activación | | Sigmoide |
| | Función de error | | Suma de cuadrados |

Fuente: Elaboración propia

La figura [8.21] contiene los resultados de las variables o factores y covariables en la capa oculta de la red, los precios observados y los precios estimados por la RNA con entrada en la red oculta sigmoidea-salida sigmoidea, tiempo de entrenamiento de la red 50 minutos, utilizando 660 muestras de entrenamiento de la red (*training*), 219 para la prueba o validación, eliminando el modelo 82 observaciones por ser atípicas, de un total de 961.

IX. Bibliografía

Valoración de inmuebles urbanos: Comparativa modelo regresión multivariable versus redes neuronales artificiales para la ciudad de Morelia, Michoacán, México.

D. José Carlos Preciado Carrillo
Dirigido por Dr. D. Francisco Guijarro Martínez

359

9. Bibliografía

Appraisal of real estate. (2002). *El avalúo de bienes raíces*. España: Library of Congress Cataloging in Publication Data.

Apraisal Institute (2002). *El avalúo de bienes raíces*. Chicago: Apraisal Institute.

Acciani, C., Fucilli, V., y Sardaro, R. (2011). Data mining in real estate appraisal: a model tree and multivariate adaptive regression spline approach. Firenze University Press.

Adielsso, S. (2005). Statistical and neural network analysis of pesticide losses to surface water in small agricultural catchments in Sweden. Swedish University of Agricultural Sciences. Department of Soil Sciences. Division of Environmental Physics.

Aguilar Morales, J.E. (2012). Los Artículos Científicos: Clasificación. México: Asociación Oaxaqueña de Psicología A.C.

Aguirre N.C. y León L.D. (2008). Aspectos causales del precio de departamentos nuevos en la comuna de Nuñoa, Santiago de Chile. Chile: Revista URBANO.

Álvarez Cascos Fernández, (1999). Real Decreto 390/1998, de 13 de marzo, que regula las funciones y la estructura orgánica de las Delegaciones de Economía y Hacienda. Madrid: Boletín Oficial del Estado.

Amabile, R. y Rosato, P. (1998). The use of neural networks in the spatial analysis of property values. Minneapolis, Minnesota: Sixth Joint Conference on Food, Agriculture and the Environment.

Amri, S. Z. y Bossomaier, T. R. J. (2005). The Influence of Spatial Attributes on House Price. *Australia: Complexity International* Volume 12.

Anselin, L. (1989). What is special about spatial data? Alternative perspectives on spatial data analysis. California: National Center for Geographic Information and Analysis.

Anas, A. y Kim, I. (1995). General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration. *Journal of Urban Economics* 40, pp. 232-256.

Anton, H. (2011). *Introducción al álgebra lineal*. México: Limusa Wiley.

Anderson, J.A. (2007) *Redes neuronales*. México: Alfaomega Grupo Editor.

Anglin, P.M y Gencay, R. (1995). Semiparametric Estimation of a Hedonic Price Function. Canada: *Journal of Applied Econometrics*, Vol. 11, pp. 633-648.

Anselin, L. (1995). Local Indicators of Spatial Association-LISA. Ohio: *Geographical Analysis*, Vol. 27, No. 2, pp. 95-115.

Aragón Marina, I., Bonnelly Ricart, R., Castañeda Cárdenas, C., Cavallazzi González, M., Herrera Mora, C., Kahane, S., Márquez Valencia, L.F., Martínez Vela, S. y Umaña Salinas M.C. (2011). *Innovación en la investigación de mercados*. Colombia: Alfaomega.

Arévalo Tomé, R. y Chamorro Rivas, J.M. (2008). Geographical heterogeneity in housing. Evidence from

Spain. Facultade de Ciencias Económicas e Empresariais,
Campus As Lagoas-Marcosende, 36310 Vigo.

Arias Ramírez, R. y Sánchez Hernández, L. (2011).
Concentración territorial de las empresas en la región
chorotega: un análisis mediante técnicas de economía
espacial. Ciencias Económicas.

Aragón Torre, A., Calzada Arroyo, J.M., García Güemes,
A. y Pacheco Bonrostro, J. (1996). Aplicaciones de redes
neuronales en economía. *V JORNADAS ASEPUMA*.

Arriaga Navarrete, R. (2012). Ciclos inmobiliarios y precios
de la vivienda: España, EUA y Reino Unido. México:
Universidad Autónoma Metropolitana, Unidad
Azcapotzalco.

Arriaga Navarrete, R. (2009). Efecto riqueza a partir de la
revalorización de los activos inmobiliarios: análisis y
evidencia empírica para el caso español. México:
Universidad Autónoma Metropolitana, Unidad
Azcapotzalco. Red de Revistas Científicas de América
Latina, el Caribe, España y Portugal Sistema de
Información Científica.

Arancibia V.F., Loehnert T.F., y José Soto R.G. (2008). Predicción de signo semanales de las acciones de Falabella, Ripley, Cencosud y D&S con redes neuronales. Santiago de Chile: Universidad de Chile, Facultad de Economía y Negocios, Escuela de Economía y Administración.

Arauzo Azofra, A., Castro Peña, J.L., Delgado Calvo-Flores, M., Navarro Ruíz, M.A. y De la Torre Martínez, J.M. (2000). Predicción de información contable con un modelo híbrido. Comparación con técnicas de regresión clásicas. Granada: Universidad de Granada.

Arnason, O. y Persson, K. (2012). Swedish real estate as a hedge against inflation - with comparison to stocks, bonds and gold. Department of Real Estate and Construction Management. Real Estate Development and Financial Services. Real Estate Management. pp. 2999-3017.

Arya, J.C. y Lardner, R.W. (1992). *Matemáticas aplicadas*. México: Pearson Educación.

Aznar Bellver, J., Guijarro Martínez, F., y Moreno Jiménez, J.M. (2007). *Valoración agraria multicriterio en un entorno con escasa información*. Estudios de Economía Aplicada.

España: Redalyc Sistema de Información Científica.Vol. 25, pp. 551-571.

Aznar Bellver, J. y Guijarro Martínez, F. (2012). *Nuevos métodos de valoración multicriterio*. España: Universidad Politécnica de Valencia.

Aznar Bellver, J. y Guijarro Martínez, F. (2005). *Nuevo método de valoración modelos multicriterio*. España: Universidad Politécnica de Valencia.

Aznar Bellver, J. y López Perales, A.A. (2008). *Valoración inmobiliaria. Métodos y aplicaciones*. España: Universidad Politécnica de Valencia.

Aznar Bellver, J. y Estruch Guitart, A.V. (2012). *Valoración de activos ambientales, teoría y casos*. España: Universidad Politécnica de Valencia.

Aznar Bellver, J., González Mora, R., Guijarro Martínez, F. y López Perales, A.A. (2012). *Valoración inmobiliaria. Métodos y aplicaciones*. España: Universitat Politècnica de València.

Autoridades de la Facultad de Ciencias Económicas y Empresariales (2005). *Boletín matemático*. Instituto de

Investigaciones de Matemática Aplicada. Argentina:
Universidad de Morón.

Baronio, A., Vianco, A. y Rabanal, C. (2012). Una
introducción a la econometría espacial dependencia y
heterogeneidad. Cátedra de Econometría.

Basso, L.R. Servicios Ecosistémicos en Argentina.

Bagdonavicius, A. y Deveikis S. (2006). Modelos de
valoración automatizada en Lituania. Lituania: CT
Catastro.PP. 61-72.

Beamonte San Agustín, M.A. (2008). Análisis estadístico
de modelos hedónicos STAR con efectos de vecindad.
Una aplicación al mercado inmobiliario de Zaragoza.
Zaragoza: Universidad de Zaragoza.

Báez, A., Herrero, L.C., Bedate, A. y Sanz, A. (2012).
Análisis de viabilidad de un proyecto de recuperación del
patrimonio cultural urbano. Valdivia, Chile: Universidad
Austral de Chile. Gestión turística núm. 17, pp. 97-119.

Barthélemy, Q., Gouy-Pailler, C., Isaac, Y., Souloumiac,
A., Larue, A. y Mars, J.I. (2013). Multivariate Temporal

Dictionary Learning for EEG. France: Journal of Neuroscience Methods. pp. 19-28.

Barreras Alconcehl, M. (2010). Matemáticas con Microsoft Excel. México: Alfaomega-Ra-Ma.

Banerjee, A., Vogelstein, J. y Dunson, D. (2013). Parallel inversión of huge covariance matrices.

Bengochea Morancho, A. (2002). A hedonic valuation of urban green areas. España: Landscape and Urban Planning, pp 35-41.

Beck, J., Fralick, J. y Toma, M. (2012). The Housing Price Bubble In A Suburban Georgia Setting: Using The Hedonic Pricing Model In The New South. USA: The Journal of Applied Business Research Volume 28, Number 4, pp. 651-660.

Berenson, M.L., Levine, D.M. y Krehbiel, T.C. (2001). Estadística para administración. México: Pearson-Prentice Hall.

Berenson, M.L., Levine, D.M. (1996). Estadística básica en administración. Conceptos y aplicaciones. México: Pearson Prentice-Hall.

Bonilla, M., Marco, P. y Olmeda, I. (2002). Redes neuronales artificiales: predicción de la volatilidad del tipo de cambio de la peseta. España: Instituto Valenciano de Investigaciones Económicas, S.A.

Borraz, F., Licandro, G. y Ponce, J. (2012). Precios de viviendas: una metodología para evaluar desvíos respecto a sus fundamentos. Uruguay: BCU Banco Central del Uruguay.

Bover, O. y Velilla, P. (2001). Precios hedónicos de la vivienda sin características: el caso de las promociones de viviendas nuevas. España: Banco de España - Servicio de Estudios Económicos núm. 73

Bover, O. y Izquierdo, M. (2003). Quality Adjusted Prices: Hedonic methods and implications for national accounts. España: Investigaciones Económicas, vol. XXVII (2), pp. 199-238.

Box, G.E., Hunter, J.S. y Hunter, W.G. (2008). Estadística para Investigadores. Diseño, innovación y descubrimiento. España: Editorial Reverté.

Bilbao Terol, C. (2001). El otro exceso de gravamen. Un análisis empírico para las políticas de vivienda. España:

Universidad de Zaragoza.Revista de Economía Aplicada, vol. IX, núm. 27, pp.35-41.

Blight, M. y MacFarlane, J. (2002). Property price and valuation modeling (Sydney – 1970 to 2000). Nueva Zelanda: University of Western Sydney.

Brañas Garza, P., Dios Palomares, R. y Escribano Fernández, P. (1997). Casco histórico de Córdoba. Análisis del precio de las características de la vivienda: un enfoque hedónico. Córdoba: I Congreso de Ciencia Regional de Andalucía: Andalucía en el Umbral del siglo XXI, pp. 781-791.

Brañas Garza, P. y Caridad y Ocerin, J.M. (1997). Análisis de vivienda urbana: modelos por zonas vs modelos completos de la ciudad.

Brewer, B.J. y Elliott, T.M. (2013). Hierarchical Reverberation Mapping. New Zealand: Department of Statistics, The University of Auckland. Vol. 453, pp.1701-1726.

Caballer, V. (2002). Nuevas tendencias en la valoración territorial. España: Universidad Politécnica de Valencia. CT Catastro. pp. 135-145.

Caballer, V. (2008). Valoración Agraria. Teoría y Práctica. Madrid: Ediciones Mundi-Prensa.

Caballer Mellado, V. y Herrerías Herrerías Pleguezuelo, R. (2007). Tasación y Valoración. Situación Actual y Perspectiva de Futuro. España: Estudios de Economía Aplicada, Asociación Internacional de Economía Aplicada España.

Caballer Mellado, V. y De La Poza, E. (2010). La Numismática como objeto de inversión y valoración. España: Estudios de Economía Aplicada, Asociación Internacional de Economía Aplicada España. Vol.28, No. 2, pp.1133-3197.

Caballer Mellado, V. y Guadalajara Olmeda, N. (2005). Incorporación de variables regionales a los modelos econométricos de valoración de la tierra de uso agrícola. Alcalá de Henares: XXXI Reunión Estudios Regionales. Universidad Politécnica de Valencia.

Caballero, R. y Fernández, G.M. (2002). *Toma de decisiones con criterios múltiples*. España: Asepuma.

Cajamarca, R. y Mena, H. (2011). Modelación de series económicas mediante métodos automáticos de regresión difusa. Ecuador: *Analítika, Revista de análisis estadístico*.

Calatrava Leyva, J. y Cañero León R. (2000). Valoración de fincas olivareras de secano mediante métodos econométricos. España: Dpto. de Economía y Ciencias Sociales Agrarias, ETSI Agrónomos.

Canavarró Teixeira, M.C., Caridad y Ocerín, J.M. y Ceular Villamandos, N. (2011). Variáveis Explicativas e a sua importância na formação do preço de um apartamento em Portugal: uma abordagem com redes neurais artificiais. *Finanças*.

Casanovas I Ramon, M. y Santandreu I Gràcia, P. (2009). Metodologías para la valoración de empresas hoteleras. *Revista de Contabilidad y Dirección*. Vol. 8, pp. 153-172.

Carruthers, J.I., Clark, D.E. y Renner, R.N. (2011). The Benefits of Environmental Improvement: Estimates From

Space-time Analysis. College of Business Administration.
Marquette University. pp.173.

Climent, C. J. (2003). Modelos de eficiencia para la tasación de edificios: un caso estudio en la ciudad de Alcoy. 27 Congreso Nacional de Estadística e Investigación Operativa.

Universidad Politécnica de Madrid. Dpto. de Economía, Sociología y Política Agrarias, Escuela Politécnica Universitaria, Universidad de Almería. Cañada de San Urbano.

Castro Hernández, L.M. (2011). Liderazgo de la empresa de intermediación inmobiliaria residencial de segunda mano en España. Departamento de Organización de Empresas Facultad de Ciencias Económicas y Empresariales. Universidad Nacional de Educación a Distancia.

Caridad y Ocerín, J.M., Núñez Tabales, J.M. y Ceular Villamandos, N. (2008). Metodología de precios hedónicos vs. Redes Neuronales Artificiales como alternativas a la valoración de inmuebles. Un caso real. Córdoba: Universidad de Córdoba. CT Catastro.No. 62, pp. 27-42.

Caridad y Ocerín, J.M. y Ceular Villamandos, N. (2004). Determinación de los precios implícitos en bienes inmuebles: Una alternativa a la modelización hedónica. Revista de estudios regionales N° 71 (2004), pp. 85-105.

Caridad y Ocerín, J.M., Brañas Garza, P. y de la Paz, M. (1997). Análisis intraurbano del precio de las características de la vivienda en Córdoba: los barrios menos favorecidos. Córdoba: I Congreso de Ciencia Regional de Andalucía: Andalucía en el Umbral del siglo XXI.

Caridad y Ocerín, J.M., Brañas Garza, P. (1997) El precio de la vivienda urbana. La disyuntiva superficie / ubicación: una ampliación. Córdoba: I Congreso de Ciencia Regional de Andalucía: Andalucía en el Umbral del siglo XXI. Pag. 172.

Cano Guervós, R., Chica Olmo, J. y Hermoso Gutiérrez, J.A. (1999). Metodología para la Zonificación de una ciudad. España: Departamento de Economía aplicada. Universidad de Granada. Estudios de Economía Aplicada, no. 3, pp. 23-49.

Cano Guervós, R. A. y Chica Olmo, J.M. (2004fecha). *Una metodología objetiva para las valoraciones inmobiliarias*. Granada: Departamento de Métodos Cuantitativos para la Economía y la Empresa Universidad de Granada, pp. 122-148.

Cebula, R.J., Foley, M. y Houmes, R. (2011). Property Tax Capitalization *within* a National Historic District versus Property Tax Capitalization *outside* that National Historic District: Another Application of the Tiebout Hypothesis. USA: International Journal of Economics and Finance Vol. 3, núm. 4.

Centro Universitario de Investigación Científica y Tecnológica. (2011). El Investigador. Ecuador: Universidad Técnica del Norte.

Cervelló Royo, R., García, F. y Guijarro, F. (2010). Ranking multicriterio de viviendas: una adaptación del modelo de precio único. España: Estudios de Economía Aplicada. Vol. 28-3, pp.1-22.

Cervelló Royo, R. y Segura García del Río, B. (2011). Efectos de las intervenciones públicas en regeneración urbana: un modelo para evaluar su eficiencia. España:

CIRIEC-España, Revista de Economía Pública, Social y Cooperativa. No. 70, pp. 33-54.

Cohen, J.P. y Coughlin, C.C. (2007). Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices. Federal Reserve Bank of St. Louis. Research Division.

Combris, P., Lecoq, S. y Visser, M. (1997). Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter? USA: *The Economic Journal*, Vol. 107, No. 441. pp. 390-402.

Comité de las Normas Internacionales de Valuación (2005). *Normas internacionales de valuación*. España: Asociación Profesional de Sociedades de Valuación de España (ATASA) y Unión Panamericana de Sociedades de Valuación (UPAV).

Coremberg, A.A. (2000). El precio de la vivienda en Argentina: un análisis econométrico de sus determinantes fundamentales. México: Universidad Autónoma del Estado de México.

Corchado, J.M., Díaz, F., Borrajo, L. y Fernández, F. (2000). *Redes Neuronales Artificiales, un enfoque práctico*. España: Servicio de Publicaciones. Universidad de Vigo.

Chica Olmo, J.M. y Cano Guervós, R.A. (1995). *Aproximación a la variabilidad espacial de las características y del precio de la vivienda. Una aplicación*. España: Departamento de Métodos Cuantitativos para la Economía y la Empresa. Universidad de Granada.

Chica Olmo, J.M. y Cano Guervós, R.A. (2006). *Análisis de la Variación Espacio-Temporal del Precio de la vivienda. El acceso a la información espacial y las nuevas tecnologías geográficas*. pp. 79-92.

Chica Olmo, J.M., Cano Guervós, R y Chica Olmo, M. (2007). *Modelo hedónico espacio-temporal y análisis variográfico del precio de la vivienda*. España: Geo Focus Revista Internacional de Ciencia y tecnología de la Información Geográfica.

Chica Olmo, J.M. (2007). *Prediction of Housing Location Price by a Multivariate Spatial Method: Cokriging*. J R E R.

Chen, S. F. y Goodman, J. (1998). *An Empirical Study of Smoothing Techniques for Language Modeling*.

Massachusetts. Center for Research in Computing Technology Harvard University.

Christensen, H. B. (2012). *Estadística paso a paso*. México: Trillas.

Dasso, A. y Funes, A. (2013). Desarrollo de Modelos de Evaluación Usando Operadores de una Lógica Continua. San Luis Argentina: XV Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional de San Luis.

De Amato, M. (2007) Comparing Rough Set Theory with Multiple Regression Analysis as Automated Valuation Methodologies. Italia: Faculty of Engineering, Technical University of Bari, Politecnico di Bari, Italy. Vol. 10 no. 2, pp. 42-65.

De Vicente Lama, M., Molina Sánchez, H., y Ramírez Sobrino, J.N. (2013). Inversiones inmobiliarias: la elección contable valor razonable vs. Coste en los grupos cotizados españoles. Córdoba: Etea. Centro Adscrito Universidad de Córdoba. NIC 40.

Des Rosiers, F., Thériault, M. y Villeneuve, P.Y. (1999). Sorting out Access and Neighborhood Factors in Hedonic Price Modelling: An Application to the Quebec City Metropolitan Area. Canada: Faculté des sciences de l'administration.

Donovan, S. y Munro, I. (2013). Impact of urban form on transport and economic outcomes. New Zealand: NZ Transport Agency. Reporr 513, pp. 74.

Domínguez Domínguez, J. y Domínguez López, J.A. (2006). Estadística y Probabilidad. El mundo de los datos y el azar. México: Oxford University Press.

Duncan, D.T., White, K., Aldstadt, J., Castro, M.C., Whalen, J. y Williams, D.R. (2012). Space, race, and poverty: Spatial inequalities in walkable neighborhood amenities? Alemania: *Demographic research* volume 26, article 17, pp. 409-448.

Dumitru, E., Szegedy, C., Toshev, A. y Dragomir, A. (2013). Scalable Object Detection using Deep Neural Networks.

Dubin, R.A. (1992). Spatial autocorrelation and neighborhood quality. USA: Regional Science and Urban Economics 433452. North-Holland.

Enríquez Martínez, M.A., Franco Maass, S., Osorio García, M., Ramírez de la O, I.L, y Nava Bernal, G. (2010). *Evaluación multicriterio de los recursos turísticos del Parque Estatal Sierra de Nanchititla*, Estado de México. México: Universidad Autónoma del Estado de México. Núm.8, pp. 7-35.

Emil Panduro, T. y Lausted Veie, K. (2013). Classification and valuation of urban green spaces – A hedonic house price valuation. Arbejdsrapport - De Økonomiske Råds Sekretariat.

Escalante Vázquez, E.J. (2013). Seis-Sigma. *Metodología y Técnicas*. México: Limusa.

Facultad de Ciencias Administrativas y Económicas. (2012). *Estudios Gerenciales*. Cali- Colombia: Universidad Icesi, Colombia.

Fedea Sociedad Abierta. (2008). *La Crisis de la Economía Española: Lecciones y Propuestas*. España: FEDEA.

Fernández Durán, L., Llorca Ponce, A., Valero Cubas, S., Botti Navarro, V.J. (2012). Incidencia de la localización en el precio de la vivienda a través de un modelo de red neuronal artificial. Una aplicación a la ciudad de Valencia. España: Departamento de Organización de Empresas. Universidad Politécnica de Valencia. CT catastro.

Fernández Avilés, G. y Larraz, B. (2008). Aportaciones de los métodos geo-cuantitativos a la estimación de precios de vivienda. Universidad de Castilla-La Mancha. Cim. Economía, pp.217-242.

Fitch Osuna, J.M. y Garcia Almirall, P. (2008). La incidencia de las externalidades ambientales en la formación espacial de valores inmobiliarios: el caso de la región metropolitana de Barcelona. ACE Arquitectura Ciudad y Entorno.

Figueroa, B.E. y Lever, D.G. (1992). Determinantes del precio de mercado de los terrenos en el área urbana de Santiago. Cuadernos de Economía. Año 29, No. 89, pp.99-113.

Fotheringham, A.S., Brunson, C. y Charlton, M. (2002). Geographically Weighted Regression. England: John Wiley & Sons Ltd.

Francke, M. y Van de Minne, A. (2013). Land, Structure & Depreciation. Amsterdam: University of Amsterdam and Ortec Finance.

Friedman, F. (2013). Apartment Price Models for the Glenwood Riverfront Development. Oregon Undergraduate Research Journal.Vol.5.

Flores, S.L. y Flores, J.O. (2008). Evaluación del mercado inmobiliario con fines de inversión. Caso: casco urbano del municipio barinas, período 2001-2005. Venezuela: Universidad Central de Venezuela. Vol. XIV, núm.1, pp. 221-253.

Flores Xolocotzi, R., González-Guillén, M.J. y De los Santos-Posadas, H.M. (2010). Valoración económica del servicio recreativo del parque Hundido de la Ciudad de México. México: Región y Sociedad / Vol. XXII / Núm. 47.

Gallego Mora-Esperanza, J. (2004). La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid. Madrid: CT catastro.

Gallego Mora-Esperanza, J. (2008). Modelos de valoración automatizada. Madrid: Jefe de Área Regional de Inspección Gerencia Regional del Catastro de Madrid. CT catastro.

Gao, X., Asami, Y. y Jo Chung, C. (2002). An Empirical Evaluation of Hedonic Regression Models. Japon: Center for Spatial Information Science The University of Tokyo. Núm.46.

Gautier, E. y Tsybakov, A.B. (2013). Pivotal estimation in high-dimensional regression via linear programming. France: CREST (ENSAE).

García L. de Meneses, T. (2003). Desarrollo de la valoración catastral de fincas rústicas. Aplicación a la Comunidad Foral de Navarra. Departamento Gestión de Empresas Universidad Pública de Navarra. CT.

García García, C., Herrerías Velasco, J. y García Pérez, J. (2007). Los procesos de *Elicitación* en el Método de las

dos funciones de distribución: Un caso práctico. España: Estudios de Economía aplicada, pp. 245-274.

García Almirall, M.P. (2004). Un SIG para la evaluación ambiental de la ciudad. Centro de Política del Suelo y Valoraciones, Universidad Politécnica de Catalunya.

García Oré, C. (2011). Estadística Descriptiva y Probabilidades para Ingenieros: Lima. Macro Empresa Editora.

García Serrano, A. (2013). Inteligencia Artificial, Fundamentos, práctica y aplicaciones. México: Alfaomega.

García I. De Meneses, T. (2000). Un modelo analógico para la valoración catastral. España: Estudios Agrosociales y Pesqueros, núm 186, 2000, pp. 105-127.

García Meca, E., Martínez Conesa, I. y Laffarga Briones, J. (2009). Normas Internacionales de valoración: la opinión de los profesionales en España International Valuation Standards: the view from Spanish practitioners. Revista española de financiación y contabilidad. Vol. XXXVII, núm. 143, pp. 479-504.

García Pozo, A. (2006). Una aproximación a la aplicación de la metodología hedónica: especial referencia al caso del mercado de la vivienda. España: Facultad de Ciencias Económicas de la Universidad de Málaga. Unesco 531299.

García Pozo, A. (2007). Determinantes del precio de la vivienda usada en Málaga: Una aplicación de la metodología hedónica. España: revista de estudios regionales nº 82, I.S.S.N.: 0213-7585 (2008), pp. 135-158.

Garrigues Carnicer, C. y Durán Boo, I. (2005). Ministerio de Economía y hacienda; Incides de 1994-2004. Madrid: Ministerio de Economía y Hacienda Secretaría General Técnica. ISSN.1138-3488.

Ghosh, J. (2007). Macroeconomía y políticas de crecimiento. Nueva York: Departamento de asuntos económicos y sociales (ONU DESA).

Guijarro Martínez, F. y Moya Clemente, I. (2008). Un procedimiento eficiente para la selección de variables en modelos factoriales de valoración de empresas. España:

Universidad Politécnica de Valencia. Revista Europea de Dirección y Economía de la Empresa. Vol. 18, núm. 3, pp. 83-98.

Guijarro Martínez, F. (2013). Guijarro Martínez, F. Estadística aplicada a la valoración modelos multivariantes. España: Universidad Politécnica de Valencia.

Gobierno de Chile Ministerio de Obras Públicas. (2013). Análisis e implementación de un modelo de planificación para la macro zona central etapa i: análisis y evaluación del sistema de transporte de las provincias sector sur poniente región metropolitana. Desarrollo y Calibración de Modelos. Chile: Ministerio de obras públicas-dirección de planeamiento.

González Tamez, C.A. (2008). Valoración económica del ruido como componente ambiental en la formación del precio del mercado inmobiliario de la vivienda. *El caso de la Ciudad de Barcelona*. España: Universitat Politècnica de Catalunya.

González Manteiga, W. (1988). Una perspectiva general con nuevos resultados de la aplicación de la estimación no

paramétrica a la regresión lineal. Santiago de Compostela: Departamento de Estadística e Investigación Operativa Facultad de Matemáticas. Vol. 30, núm. 118, pp. 141-179.

González Alonso, F., Cuevas Gonzalo, J.M., Arbiol Bertrán, R. y Baulies Bochaca, X. (1997). Utilización Diacrónica de encuestas de campo e Imágenes de Satélite en la Estimación de superficies cultivadas por el Método de Regresión. España: VII Congreso Nacional de Teledetección, pp. 27-30.

Gómez Baggethun, E. y Barton, D.N. (2012). Classifying and valuing ecosystem services for urban planning. Barcelona: Ecological Economics. Elsevier. pp. 235-245.

Grossman, S.I. (1996). *Álgebra lineal*. México: Mc Graw Hill.

Hair, J.F., Anderson, R.E., Tatham, R.L. y Black, W.C. (1999). Análisis Multivariante. España: Pearson- Prentice Hall.

Hansen, J.L. y Benson, E.D. (2013). The value of a water view: variability over 25 years in a coastal housing market.

Western Washington. The Coastal Business Journal.
Volume 12, Number 1.

Haliloğlu Kahraman, Z.E. (2013). Dimensions of housing satisfaction: a case study based on perceptions of rural migrants living in dikmen. Turkey: Department of City and Regional Planning. pp.1-27.

Hernández Santoyo, A., León Sánchez, M.A., y Casas Vilardell, M. (2004). Valoración económico - Ambiental de los recursos forestales basada en técnicas de decisión multicriterio. Estudio de caso: parque nacional viñales, pinar del río. Cuba: Universidad de Pinar del Río.

Hernández Sampieri, R., Fernández collado, C. y Baptista Lucio, P. (2010). Metodología de la Investigación. México: Mc Graw Hill.

Hilera, J.R. y Martínez, V.J. (1995). Redes Neuronales Artificiales, Fundamentos, Modelos y aplicaciones: Madrid: Serie Paradigma Ra-Ma.

Humarán Nahed, I. y Roca Cladera, J. (2010). Hacia una medida integrada del factor de localización en la valoración residencial: el caso de Mazatlán. México.

España: Universidad Autónoma de Sinaloa. Universidad Politécnica de Cataluña.

Huanga, Z., Chena, H., Hsua, C-J., Chenb, W-H. y Wuc, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. Elsevier B.V.pp.543-558.

InterMoney, S.A. (1999). Fondos de inversión inmobiliaria y precios de inmuebles. Fondos de Inversión Inmobiliaria y Precios de Inmuebles.

Isasi Viñuela, P. y Galván León, I.M. (2004). Redes de Neuronas Artificiales, Un enfoque práctico. España: Pearson- Prentice Hall.

Jylanki, P., Nummenmaa, A. y Vehtari, A. (2013). Expectation Propagation for Neural Networks with Sparsity-promoting Priors. Finland: Department of Biomedical Engineering and Computational Science. Athinoula A. Martinos Center for Biomedical Imaging Massachusetts General Hospital. Department of Biomedical Engineering and Computational Science. pp.1-39.

Jory Rubilar, M.E. (2007). Predicción de las variaciones de costos para proyectos de construcción utilizando redes neuronales. Santiago de Chile: Universidad de Chile facultad de ciencias físicas y matemáticas departamento de ingeniería civil.

Kallus, N. (2010). Regression-robust designs of controlled Experiments. Massachusetts Institute of Technology. The National Science Foundation.

Ko, J. y Jason Cao, X. (2013). The Impact of Hiawatha Light Rail on Commercial and Industrial Property Values in Minneapolis. Minnesota: Journal of Public Transportation, Vol. 16, No. 1.

Komarova, V. (2009). Valuing Environmental Impact of Air Pollution in Moscow with Hedonic Prices. University of Navarra, Pamplona. España: World Academy of Science, Engineering and Technology. Vol. 3, no. 9.

Kiel, K.A. y Williams, M. (2006). The impact of Superfund sites on local property values: Are all sites the same?. USA: Journal of Urban Economics 61 (2007). pp. 170–192. Elsevier Inc.

Kroll, Ph.D., C.A. y Cray, J.D., A.F. (2010). Hedonic Valuation of Residential Resource Efficiency Variables A Review of the Literature The Center for Resource Efficient Communities University of California, Berkeley. California: The Center for Resource Efficient Communities University of California, Berkeley. pp.1-53.

Kryvobokov, M. y Wilhelmsson, M. (2007). Analysing location attributes with a hedonic model for apartment prices in donetsk, ukraine. International Journal of Strategic Property Management (2007) 11, pp. 157–178.

Laterra, P., Jobbagy E.G. y Paruelo, J.M. (2013). Valoración de servicios Ecosistématicos. Conceptos, herramientas y aplicaciones para el ordenamiento territorial. Instituto Nacional de Tecnología Agropecuaria.

Lei Zhang, X. y Wu, J. (2013). Denoising deep neural networks based voice activity detection. China: Multimedia Signal and Intelligent Information Processing Laboratory, Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing.

Lisi, G. (2012). On the functional form of the hedonic price function: a matching-theoretic model and empirical evidence. *The International Real Estate Review*.

Lind, D.A., Mason, R.D. y Marchal, W.G. (2000). *Estadística. Para administración y economía*. México: McGraw Hill-Irwin McGraw-Hill.

Lind, D.A., Marchal, W.G. y Wathen, S.A. (2012). *Estadística aplicada a los negocios y la economía*. New York: McGraw Hill.

Limsombunchai, V. (2004). House Price Prediction: Hedonic Price Model vs. Artificial Neural Network. New Zealand: Commerce Division, Lincoln University.

López Andi6n, M.C. (2002). Modelos econom6tricos del mercado de la vivienda en las regiones espa6olas. Universidad de Santiago de Compostela, Working Paper Series Economic Development.

L6pez, F.A. (2013). Sistemas de informaci6n geogr6fica y econometr6a espacial en la tasaci6n de inmuebles urbanos. Proyecto Piloto en la Ciudad de Cartagena. Cartagena:

Departamento de Métodos Cuantitativos e Informáticos
Universidad Politécnica de Cartagena.

Long, F., Páez, A. y Farber, S. (2007). Spatial Effects in Hedonic Price Estimation: A Case Study in the City Of Toronto. McMaster University. CSpA Centre for Spatial Analysis.

Lützkendorf, T. (2010). *Inmuebles sostenibles – ¿sueño o tendencia?* Vol. 62, 517, pp. 5-22.

Lu, B., Harris, P., Charlton, M. y Brunsdon, C. (2013). The GWmodel R package: Further Topics for Exploring Spatial Heterogeneity using Geographically Weighted Models. Ireland: National Centre for Geocomputation, National University of Ireland Maynooth, Maynooth.

Masry, E. (1997). Multivariate regression estimation: Local polynomial fitting for time series. California: Department of Electrical and Computer Engineering. Elsevier Science B.V.

Martínez Blasco I. (1998). Modelos Econométricos aplicados a la Valoración de Bienes Inmuebles Rústicos. pp. 181-212.

Martínez Estudillo, F. J., Hervás-Martínez, C., Torres Jiménez, M. y Martínez Estudillo, A. C. (2007). Modelo no lineal basado en redes neuronales de unidades producto para clasificación. Una aplicación a la determinación del riesgo en tarjetas de crédito. España: Universidad Pablo de Olavide Sevilla. Revista de Métodos Cuantitativos para la economía y la empresa. pp. 40-62.

Martínez Pagés, J. y Ángel Maza, L. (2003). Análisis del precio de la vivienda en España. España: Banco de España, Madrid. Documento de trabajo 0307.

Martín del Brío, B. y Sanz Molina, A. (2007). Redes Neuronales y Sistemas Borrosos. México: Alfaomega-Rama.

Martín Martín, Q. y Santana, Y. (2007). Aplicación de las Redes Neuronales Artificiales a la Regresión. Madrid: Editorial La Muralla.

Marqués, F. (2011). Microsoft Excel 2010 en Profundidad. México: Alfaomega.

Marmolejo Duarte, C. y González Tamez, C.A. (2009). El impacto del ruido sobre la formación espacial de los valores inmobiliarios: un análisis para el mercado residencial de Barcelona. Barcelona: Doctorando del Programa en Gestión y Valoración Urbana y Arquitectónica. UPC.

Matthew S.Y., Yu, j. y Jin, L. (2012). Detecting Bubbles in Hong Kong Residential Property Market. Singapore Management University Institutional Knowledge at Singapore Management University.

Masoud Bakhtyari, K., Pirasteh, S., Pradhan, B., Rodzi Mahmud. A., Azmin Sulaiman, W.N. y Moradi, A. (2011). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Springer-Verlag.

Massiani, J. (2013). How to Value the Benefits of a Recreational Area? A Cost-Benefit Analysis of the Conversion of a Brownfield to a Public Beach in Muggia (Italy). Università cà Foscari di Venezia. pp. 86-102.

Malpezzi, S. (2002). Hedonic Pricing Models: A Selective and Applied Review. Wisconsin: The Center for Urban Land Economics Research.

Mermelstein, D.A. (2006). Mortgage defaults, macroeconomics, and institutional arrangements: Beyond the standard Credit Scoring. Buenos Aires: MPRA Munich Personal RePEc Archive.

Mendieta, J.C. y Perdomo, J.A. (2007). Especificación y estimación de un modelo de precios hedónico espacial para evaluar el impacto de transmilenio sobre el valor de la propiedad en Bogotá. Colombia: CEDE Universidad de los Andes.

Meloni, O. y Ruiz Nuñez, F. (2002). El precio de los terrenos y el valor de sus atributos. Un enfoque de precios hedónicos. Argentina: XXXII Reunión Anual de la Asociación Argentina de Economía Política llevada a cabo en Mendoza.

Meelun, G., Whittal, J. y Evans, K. (2011). Analysis of spatial heterogeneity in modelling high-priced residential properties of cape town using neural networks. Geomatics Division, School of Architecture, Planning and Geomatics.

Montero Lorenzo, J.M. (2007). Estadística Descriptiva. España: Internacional Thomson Editores.

Michaels, R.G. y Smith, V.K.(1990). Market Segmentation and Valuing Amenities with Hedonic Models: The Case of Hazardous Waste Sites. Academic Press, Inc. Journal of urban economics. pp. 223-242.

Modugno, L., Cagnone, S. y Giannerini, S. (2013). The Analysis of Tribal Art Prices: a Multilevel Model with Autoregressive Components. Italy: Department of Statistical Sciences, University of Bologna.

Morales, I. y González, M.I. (2003). Comparación de las técnicas de análisis de variancia y regresión lineal múltiple: aplicación a un experimento de almacenamiento de mango. Costa Rica: Universidad de Costa Rica.

Morales Zurita, L. y Arias Arbeláez, F.A. (2005). La calidad de la vivienda en Bogotá: Enfoque de precios hedónicos de hogares y de agregados espaciales. Colombia: Universidad del Valle. Revista de Sociedad y Economía, núm. 9, pp. 47-80.

Montañés, A. y Clemente J. (2005). Un análisis empírico del precio de la vivienda en Aragón y su relación con la renta. Zaragoza: Departamento de Análisis Económico Universidad de Zaragoza. pp.1-58.

Montero Lorenzo, J.M. y Larraz Iribas, B. (2006). Estimación espacial del precio de la vivienda mediante métodos de Krigeado(*). España: Estadística española.

Montero Lorenzo, J.M. (2004). El precio medio del metro cuadrado de la vivienda libre: Una aproximación metodológica desde la perspectiva de la Geoestadística. España: Red de Revistas Científicas de América Latina, el Caribe, España y Portugal Sistema de Información Científica.

Montúfar, G., Rauh, J. y Ay, N. (2013). Maximal Information Divergence from Statistical Models defined by Neural Networks. USA: University Park, Max Planck Institute for Mathematics in the Sciences, Santa Fe Institute.

Molina, Isabel. (2000). La valoración catastral de los inmuebles singulares. *Dirección General del Catastro. CT*

Muñoz Gielen, D. (2009). El justiprecio expropiatorio en los países bajos: ACE: Architecture, City and Environment. Año III, núm. 9, pp. 51-56.

Navarro Pastor, J.B. y Losilla Vidal, J.M. (2000). Análisis de datos faltantes mediante redes neuronales artificiales. Barcelona: Universidad Autónoma de Barcelona. Vol. 2, núm. 3, pp. 503-510.

Navarro Galera, A. y Pérez López, M.C. (2009). La repercusión del valor razonable de los bienes inmuebles en la utilidad de los estados financieros: una nota de investigación. España: Revista de Contabilidad-Spanish Accounting Review Vol. 12 - N° 1. pp. 141-162.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers. USA: European Economic Review 39 1697-1716.

Nikolaos, K., Dimitra, V. y Agapi, X. (2011). Real estate values and environment: a case study on the effect of the environment on residential real estate values. Grecia: International Journal of Academic Research Vol. 3. No. 1. January, 2011, Part III.

Núñez Cerda, F. y Schovelin Surhoff, R. (2002). Modelo de precio de suelo urbano en el gran concepción. Concepción – Chile: Universidad del Bío-Bío.

Núñez Cerda, F. y Schovelin Surhoff, R. (2002). Estimación de un Modelo Hedónico para conjuntos de viviendas nuevas. Concepción – Chile: Universidad del Bío-Bío.

Núñez Tabales, J.M., Caridad y Ocerin, J.M. y Ceular Villamandos, N. (2009). Propuestas metodológicas para valoración de inmuebles urbanos. Granada: Editorial Comares.

Núñez Tabales, J.M., Caridad y Ocerin, J.M., Ceular Villamandos N., y Fuentes Jiménez, A.M. (2009). Estimación del precio de la vivienda mediante redes neuronales artificiales (rna) en diferentes marcos temporales. Córdoba: Dpto. Estadística, Econometría, I.O. y Org. Empresas. Universidad de Córdoba. Vol. 11, pp. 79-111.

Núñez Tabales, J.M. (2007). Mercados Inmobiliarios: Modelización de los precios. Córdoba: Departamento de Estadística, Econometría, I. O. y Organización de Empresas.

Núñez Tabales, J.M., Caridad y Ocerin, J.M., y Rey Carmona, F.J. (2013). Redes neuronales artificiales para la predicción de precios inmobiliarios. España: Universidad Pablo de Olavide Sevilla. Revista de Métodos Cuantitativos para la economía y la empresa. pp. 29-44.

Núñez Tabales, J.M., Ceular Villamandos N. y Millán Vázquez de la Torre, G. (2007). Aproximación a la Valoración Inmobiliaria mediante la metodología de precios Hedónicos (MPH). Ayala Calvo, J.C. y grupo de investigación FEDRA. Conocimiento, innovación y emprendedores: Camino al futuro.

Nicholls, S. y Crompton, J.L. (2005). The Impact of Greenways on Property Values: Evidence from Austin, Texas. Texas: *National Recreation and Park Association*. Vol. 37, No. 3, pp. 321-341.

Oduwole, H. K. y Eze, H. T. (2013). A Hedonic Pricing Model on Factors that Influence Residential Apartment

Rent in Abuja Satellite Towns. Nigeria: Mathematical Theory and Modeling Vol.3, No.12.

Ortiz González Rafael (2003). Formulación y Evaluación de Proyectos Inmobiliarios para valuadores.

Oviedo Pro, J.L., Caparrós Gass, A. y Campos Palacín, P. (sin fecha). Valoración económica y ambiental de los bosques: los servicios recreativos. España: Instituto de Políticas y Bienes Públicos Centro de Ciencias Humanas y Sociales Consejo Superior de Investigaciones Científicas.

Organización del Centro de Desarrollo de la OCDE. (2013). Perspectivas económicas de América Latina 2014 Logística y Competitividad para el Desarrollo. OCDE/CEPAL/CAF 2013.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T. y French, N. (2003). Practice Briefing. Real estate appraisal: a review of valuation methods. Grecia: Journal of Property Investment & Finance Vol. 21 No. 4, 2003. pp. 383-401.

Pardini, A. (2007). A perspective on the valorization of agro-silvo-pastoral systems in the Mediterranean Basin

Pastos y Forrajes. Cuba: Estación Experimental de Pastos y Forrajes "Indio Hatuey". Vol.30, núm. 1, pp. 77-105.

Pardew, J.B., Shane, R.L. y Yanagida, F.J. (1986). Structural Hedonic Prices of Land Parcels in Transition from Agriculture in a Western Community. *Western Journal of Agricultural Economics*, 11(1). pp. 50-57.

Parnreiter, C. (2011). Formación de la ciudad global, economía inmobiliaria y transnacionalización de espacios urbanos. El caso de Ciudad de México. Chile: Pontificia Universidad Católica de Chile. Red de Revistas Científicas de América Latina, el Caribe, España y Portugal Sistema de Información Científica. Vol. 38, núm. 111, pp. 5-24.

Paytona, S., Lindseyb, G., Wilsonc, J., Ottensmanna, J.R. y Mand, J. (2008). Valuing the benefits of the urban forest: a spatial hedonic approach. Londres: *Journal of Environmental Planning and Management* Vol. 51, No. 6, November 2008, pp. 717–736.

Pérez Delgado, M.L. (2003) Aplicaciones de las redes neuronales artificiales a la estadística. Madrid: Editorial La Muralla.

Pitarque, A., Roy, J.F. y Ruiz, J.C. (1998). Redes neurales vs modelos estadísticos: Simulaciones sobre tareas de predicción y clasificación. España: *Psicológica* (1998) 19, pp. 387-400.

Pi-ying, L. (2011). Analysis of the Mass Appraisal Model by Using Artificial Neural Network in Kaohsiung City. Taiwan, China: *Journal of Modern Accounting and Auditing*, Vol. 7, No. 10, pp. 1080-1089.

Pradeep Chaudhry, M.P. Sharma, G. Singh y Arun Bansal. (2013). Valuation of Urban Environmental Amenities in Developing Countries: A Case Study from Chandigarh, India. India: *Global Journal of Science Frontier Research Environment & Earth Science*. Vol. 13.

Proïa, F. (2013). Testing for residual correlation of any order in the autoregressive process.

Plata Rocha, W. (2010). Descripción, Análisis y Simulación del Crecimiento Urbano Mediante Tecnologías de la Información Geográfica. El Caso de la Comunidad de Madrid. Alcalá de Henares: Departamento de Geografía Programa de Doctorado en Cartografía, Sistemas de Información Geográfica y Teledetección.

Pulido Calvo, I. y Manuela Portela, M. (2007). Aproximaciones Neuronales Univariantes para la predicción de Caudales Diarios en Cuencas Portuguesas. Fundación para el Fomento de la Ingeniería del agua. Vol. 14, Núm. 2.

Puchstein, R. y Preuß, P. (2013). Testing for stationarity in multivariate locally stationary processes. Germany: Ruhr-Universität Bochum Fakultät für Mathematik.

Ponce Cruz, P. (2010). Inteligencia Artificial. Con aplicaciones a la ingeniería. México: Alfaomega.

Póczos, B., Rinaldo, A., Singh, A. y Wasserman, L. (2013). Distribution-Free Distribution Regression. Pittsburgh, PA USA: Carnegie Mellon University. pp. 507-515.

Quezada Lucio, N. (2012). Estadística con SPSS20, Investigación y manejo de datos, Manipulación de casos y variables, Informes, estadística descriptiva y tablas, Pruebas no paramétricas y Gráficos. Lima – Perú: Empresa Editora Macro.

Quispe Villafuerte, A. (2012). Una aplicación del modelo de precios hedónicos al mercado de viviendas en Lima Metropolitana. Sociedad de Economía y Derecho UPC. Vol. 9, núm.36.

Rangel Aquino, A.M. (2008). Análisis sobre un Rango de Variación Aceptable en factores de homologación, en la Elaboración de Avalúos de inmuebles en el centro histórico de la Ciudad de Campeche. México: Cámara Mexicana de la Industria de la Construcción.

Ramos, A., Sánchez, P., Ferrer, J.M., Barquín, J., y Linares, P. (2010). Modelos Matemáticos de Optimización. Madrid: Escuela Técnica Superior de Ingeniería Departamento de Organización Industrial.

Razi, M.A. y Athappilly, K. (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. Expert Systems with Applications Elsevier. pp. 65-74.

Renigier Biłozor, M. y Wiśniewski, R. (sin fecha). The impact of macroeconomic factors on residential property prices indices in Europe. XLI Incontro di Studio del Ce.S.E.T. pp. 149-166.

Render, B., Stair, R.M. y Hanna, M.E. (2012). Métodos cuantitativos para los negocios. México: Always Learning-Pearson.

Reyes Bueno, F.R. (2012). Valoración de Tierras en Vilcabamba (Ecuador): Diseño de un modelo de Valoración catastral Rural. Lugo: Universidad de Santiago de Compostela Escuela Politécnica Superior. Departamento de Ingeniería Agroforestal.

Rios Saavedra, O.A, y Borrero, O. (2010). Curso de econometría aplicada a los avalúos técnicas modernas de valuación de inmuebles. Colombia: Lonja de Propiedad Raíz de Cali y Valle del Cauca.

Rivas, D., Maldonado, J.L., Borges, R., y Colmenares, G. (2007). Aplicación de los algoritmos genéticos para estimar los parámetros en un modelo de regresión de Cox. Colombia: Universidad de Los Andes. Economía XXXI, pp. 57-74.

Rosas Cabal, J. (2013). Diccionario Matemáticas. México: Editores Mexicanos Unidos.

Rossini, P. (1999). Application of Artificial Neural Networks to the Valuation of Residential Property. Nueva Zelanda: Third Annual Pacific-Rim Real Estate Society Conference.

Rodríguez Rodríguez, A.M. (2005). Por qué es necesario la Sinergia entre el Campo de la Tasación Inmobiliaria y la Contabilidad. Revista Científica de Vol. IXI N° 1.

Rojas López, M.D., Bran Rueda, N.Y. y Rincón López, C. (2013). Modelos de financiación de vivienda. casos: México, Chile y Colombia. Colombia: Revista Ingenierías Universidad de Medellín, vol. 12, No. 22. pp. 97 – 108.

Romero, C. (1993). Teoría de la decisión multicriterio: Conceptos, técnicas y aplicaciones. Madrid: Alianza Universidad Textos.

Ruá Aguilar, M.J. (2011). Método de valoración de viviendas desde la perspectiva medioambiental y análisis de costes. Valencia: Universidad Politécnica de Valencia, Departamento de Economía y Ciencias Sociales.

Russel, S. y Norvig, P. (2003). Inteligencia Artificial. Un enfoque Moderno. Madrid: Pearson Prentice-Hall. pp. 1-1240.

Saiful Islama, K. y Asamib, Y. (2011). Addressing Structural Instability in Housing Market Segmentation of the Used Houses of Tokyo, Japan. International Conference: Spatial Thinking and Geographic Information Science 2011. Elsevier. pp. 93-109.

Sagner, A. (2009). Determinantes del precio de viviendas en Chile. Chile: Central Bank of Chile Working Papers no. 549.

Saginer, J., Simons, R. y Throupe, R. (2012). Emerald Article: A meta-analysis of the effect of environmental contamination on non-residential real estate values. Journal of Property Investment & Finance, Vol. 29 Iss: 4 pp. 460 - 478.

Sánchez Camperos, E.N., y Alanis García, A.Y. (2006). Redes Neuronales. Conceptos fundamentales y aplicaciones a control automático. España: Pearson-Prentice Hall.

Saptutyingsih, E. (2013). Impact of air pollution on property values: a hedonic price study. *Jurnal Ekonomi Pembangunan Volume 14, Nomor 1, Juni 2013, hlm. pp. 52-65.*

Selím,S.(2008). Determinants of house prices in turkey: a hedonic regression model. Celal Bayar Üniversitesi, İ.İ.B.F., İşletme Bölümü. pp. 65-76.

Seong-Hoon, Ch., J.M. Bowker, y William M. Park. (2006). Measuring the Contribution of Water and Green Space Amenities to Housing Values: An Application and Comparison of Spatially-weighted Hedonic Models. Tennessee: The American Agricultural Economics Association Annual Meeting.

Segura, B., Garcia Portillo, R. y Vidal, F. (1997). Modelos econométricos de valoración. aplicación a la valoración fiscal. Valencia: Invest. Agr.: Prod. Prot. Veg. Vol. 13 (1-2), 1998.

Schfer, M., Schoenauer, T.,Wolff, C., Hartmann, G., Klarb, H. y Rckert, U. (2001). Simulation of Spiking Neural Networks - Architectures and Implementations. Alemania: Elsevier.

Silva, C.A., Cruz Trejos, E.A. y Medina, P.D. (2010). Valoración de inmuebles compaginando teoría y práctica. Colombia: Universidad Tecnológica de Pereira.

Simpson, D.P., Turner, I.W, Strickland, C.M. y Pettitt, A.N. (2013). Scalable iterative methods for sampling from massive Gaussian random vectors. Australia: Department of Mathematical Sciences, Norwegian University of Science and Technology, Mathematical Sciences School, Queensland University of Technology School of Economics, Australian Business School.

Schovelin Surhoff, R. (2004). Modelos para Maximizar Ingresos de Proyectos de Viviendas en el Gran Concepción. Concepción, Chile: Universidad del Bío-Bío. pp. 69-80.

Shachmurove, Y. (2005). Business Applications of Emulative Neural Networks. Departments of Economics, The City College of the City University of New York and, The University of Pennsylvania, SMC Premier Holdings.

Spiegel, M.R., Schiller, J.J., Srinivasan, R.A. y Stephens, M. (2009). Probabilidad y Estadística. México: Mc Graw Hill.

Suriatini Ismail, Ph.D. (2006). Spatial Autocorrelation and Real Estate Studies: A Literature Review. Johor, Malaysia: Department of Property Management, Faculty of Geoinformation Science and Engineering, Universiti Teknologi Malaysia.

Takeuchi, T. y Sekido, H. (2013). An Approximate Approach to E-optimal Designs for Weighted Polynomial Regression by Using Tchebycheff Systems and Orthogonal Polynomials. Graduate School of Informatics, Kyoto University.

Tránchez Martín, J.M. (2000). Características de la vivienda determinante de su valor de mercado. Una aproximación utilizando el modelo de precios hedónicos. Zaragoza: U.N.E.D Departamento de economía aplicada e Historia Económica.

Varin, C., Cattelan, M. y Firth, D. (2013). Statistical Modelling of Citation Exchange Among Statistics Journals.

Italy: Università Ca' Foscari, Università di Padova, University of Warwick. pp. 135-150.

Valderrey Sanz, P. (2010). SPSS 17 Extracción del conocimiento a partir del análisis de datos. España: Alfaomega-Ra-Ma.

Vergés Escuín, R. (2002). El mercado residencial español 1987-2006. Expectativas de vivienda y suelo. España: *Revista de Economía y Finanzas de Castilla y León.* pp. 73-85.

Vergés Escuín, R. (2012). Estadísticas de vivienda libre en España. 1987-2011. Series operativas de compraventas y precios. España: Estadística Española Volumen 54, número 179 / 2012, pp. 357-420.

Ventolo, W.L. y Williams, M.R. (1997). Técnicas del Avalúo inmobiliario. Chicago: Real Estate Education Company.

Vidal Domínguez, M.J. y Vinuesa Angulo, J. (2012). Valoración Crítica de análisis de correlación aplicada al estudio de la oferta del suelo en estructuras de mercado

de carácter rural. Madrid: Universidad Autónoma de Madrid.

Voorman, A., Shojaie, A. y Witten, D. (2012). Graph Estimation with Joint Additive Models. Department of Biostatistics, University of Washington. Biometrika Trust. pp. 1-25.

Xu, B. (2007). An hedonic analysis of southwestern louisiana wetland prices using gis. The Department of Agricultural Economics & Agribusiness.

Yusof Hussain, M., Abdul Rahman, R., Mohamed Husain, F.N., Lyndon, N. y Nabila Ibrahim, N. (2012). Housing Bubbles Assessment 2005-2010 : Experiences in Klang Valley, Malaysia. Advances in Natural and Applied Sciences, 6(1). pp. 33-41.

Walpole, R.E., Myers, R.H., Myers, s.L. y Ye, K. (2007). Probabilidad & Estadística. Para Ingeniería & Ciencias. México: Pearson-Prentice Hall.

WEN Hai-zhen, JIA Sheng-hua y GUO Xiao-yu. (2005). Hedonic price analysis of urban housing: An empirical

research on Hangzhou, China: Journal of Zhejiang University SCIENCE. pp. 907-914.

Worzala, E., Lenk, M. y Silva, A. (1995). An Exploration of Neural Networks and Its Application to Real Estate Valuation. Colorado: The Journal of Real Estate Research.

Zurada, J.M., Levitan, A.S. y Guan, J. (2006). Non-Conventional Approaches To Property Value Assessment. Louisville: Journal of Applied Business Research.