

Autor: Ing. Roberto Piol Puppio

Organismo: Sociedad de Ingeniería de Tasación de Venezuela (SOITAVE)

E-Mail: rpiol@yahoo.com

Area el programa

al que se adscribe: AREA GENERAL: Aplicación de Nuevas Tecnologías en la valoración

Teléfono: 58-212-731.13.19

Título: **Redes Neurales aplicadas al Avalúo Inmobiliario**

Abstract:

El propósito de este trabajo, es investigar la aplicación de la tecnología de redes neurales en la elaboración de Avalúos Inmobiliarios. Ante el declive de los análisis de regresión múltiple frente a la dinámica del mercado inmobiliario, causado principalmente por la discapacidad de los paquetes estadísticos de considerar algo mas que reglas y modelos matemáticos rígidos. Se comparó la habilidad predictiva de una red neural con modelos de regresión múltiple, obteniendo como resultado: En tareas de predicción las redes neurales y los modelos de regresión múltiple generan similares resultados; mientras que para tareas de clasificación las redes neurales rinden mejor.

Palabras Clave: inteligencia artificial, red neural, regresión múltiple, estadística, predicción, clasificación, perceptrón, retropropagación, back-propagation

1.0 Introducción a la Inteligencia Artificial

1.1 La Red Neural Artificial (RNA)

Desde hace años, algunos investigadores han estado creando modelos que simulan la actividad cerebral, en un esfuerzo por producir una forma de inteligencia artificial.

Las RNA están compuestas de un gran número elementos de procesamiento profundamente interconectados (Neuronas) trabajando simultáneamente para la solución de problemas específicos. Las RNA, tal como las personas, aprenden de la experiencia.

En cualquier caso, se trata de una nueva forma de computo, que es capaz de manejar las imprecisiones e incertidumbres que aparecen cuando se trata de resolver problemas relacionados con el mundo real, ofreciendo soluciones precisas y de fácil implementación.

Las RNA están compuestas de muchos elementos sencillos que operan simultáneamente, el diseño de la red está determinado mayormente por las conexiones entre sus elementos, simulando de esta forma las conexiones de las neuronas cerebrales.

1.2 Operación de una Red Neural Artificial:

1.2.1 Los Pesos: Las RNA puede tener factores de peso fijos o adaptables. Las que tienen pesos adaptables emplean leyes de aprendizaje internas para ajustar el valor de la fuerza de interconexión con otras neuronas; de tal manera que los pesos adaptables son esenciales si no se conoce previamente cual deberá ser el valor correcto de los mismos.

Para aquellos software que utilicen rutinas de pesos variables, los mismos serán determinados automáticamente por el programa a partir de la descripción del problema.

Si las neuronas utilizan pesos fijos, entonces su valor o expresión matemática deberá estar previamente definida y será independiente del tipo de datos a procesar por la red.

1.2.2 Las Leyes de Aprendizaje: Son aquellas que determinan como la red ajustará sus pesos utilizando una función del error o algún otro criterio. La ley de aprendizaje o entrenamiento adecuada, será determina la RNA en base a la naturaleza del problema que intenta resolver.

1.2.3 Tipos de Aprendizaje: Existen dos tipo de aprendizaje o entrenamiento: Supervisado y No Supervisado. El primero ocurre cuando se le proporciona a la RNA tanto los datos de entrada como los de salida; de tal manera que la red ajusta sus pesos tratando de minimizar el error de salida calculada. El aprendizaje o entrenamiento no supervisado se presenta cuando a la RNA se le proporcionan únicamente los Datos de entrada, y la red ajusta sus interconexiones basándose únicamente en dichos datos y la salida de la propia red (este tipo de entrenamiento no será utilizado para esta monografía).

1.2.4 Fases de Operación de una RNA:

1.2.4.1 **Entrenamiento o Aprendizaje de la Red:** El usuario proporciona a la red un número preciso de datos de entrada y salida; la red entonces ajusta su pesos de interconexión o sinapsis hasta que la salida de la red se aproxime a la salida correcta.

1.2.4.2 **Recuperación de lo Aprendido:** A la red se le presenta un conjunto de datos de entrada y salida similares, y esta simplemente recalcula la salida correcta. Por lo tanto el conocimiento o capacidad de dar una respuesta de una RNA, se encuentra en la Función de Activación utilizada y en los valores de sus Pesos.

1.2.5 Características de una RNA:

2.2.5.1 **No son algorítmicas:** La gran diferencia del empleo de las redes neuronales en relación con otras aplicaciones como lo es la Regresión Múltiple; radica en que las RNA no son algorítmicas, es decir que no se programan obligándoles a seguir una secuencia predefinida de instrucciones. Las RNA generan ellas mismas sus propias "reglas", para asociar una respuesta a sus entradas. Aprende por ejemplos y de sus propios errores.

2.2.5.2 **Asocian y generalizan sin reglas:** Tal como lo hace el cerebro humano

2.2.5.3 **Requieren de algún tipo de patrón:** Las RNA, no son capaces de reconocer nada que no tenga adjuntamente algún tipo de patrón. Es por esto, que no pueden predecir la lotería ni las carreras de caballos ya que por definición son procesos al azar.

2.2.5.4 **La solución dependerá de la forma de hacer las conexiones:** Así como existen varias formas de conectar neurona biológicas en nuestro cerebro; la solución a un problema variará de acuerdo a la topología de una RNA.

2.3 Topologías de Redes Neurales Artificiales

Existen varias formas de hacer las conexiones en una RNA. Cada tipo sirve para diferentes procesos; algunas de las topologías mas comúnmente usadas son:

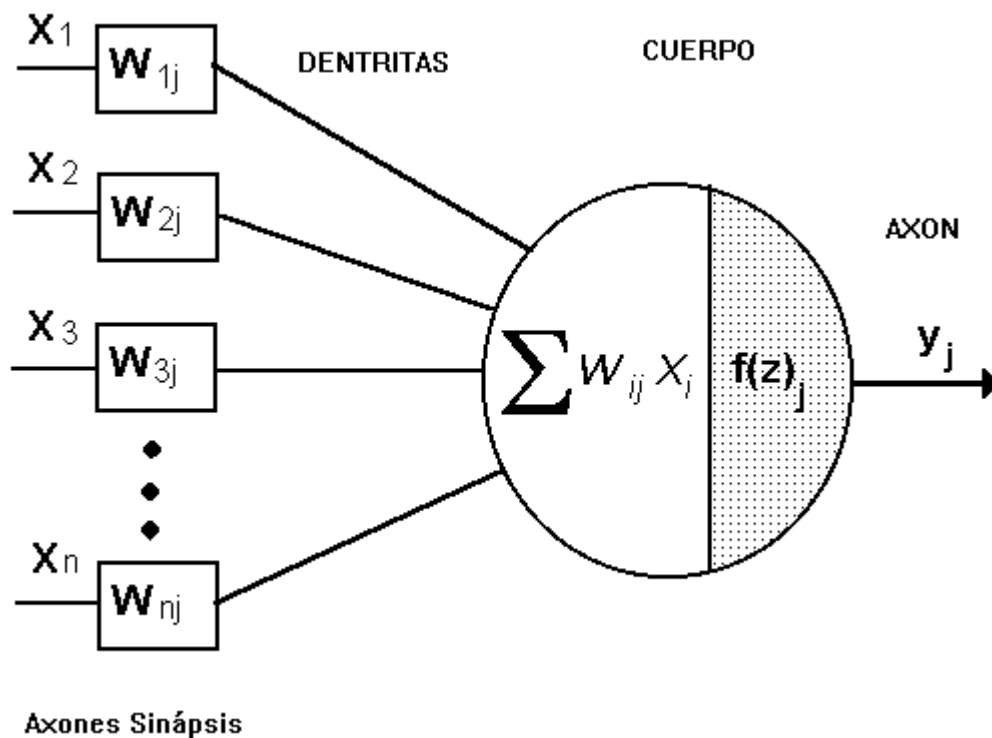
- Perceptrón Multicapa
- Retropropagación (Backpropagation)
- Hopfield
- Kohonen

De ellas, solo se desarrollará en esta monografía las Topologías Perceptrón Multicapa y Retropropagación (Backpropagation), que son las mas comúnmente empleadas en los software de Inteligencia Artificial.

2.3.1 Perceptrón

En la siguiente figura se representa una neurona "artificial", que intenta modelar el comportamiento de una neurona biológica. Aquí el cuerpo de la neurona se representa como un sumador lineal de los estímulos externos $z_j = \sum W_{ij} * X_j$, seguida de una función no lineal $y_j = f(z_j)$.

La función $f(z_j)$ es llamada Función de Activación, y utiliza la sumatoria de estímulos externos ($\sum W_{ij} * X_j$) para determinar la salida de la neurona.



Este modelo se conoce como "Perceptrón de McCulloch-Pitts", y es la base de la mayor parte de las arquitecturas de las RNA.

Las neuronas artificiales, emplean funciones de activación diferentes según la aplicación; algunas veces estas son funciones lineales y la mayoría de las veces no lineales. La eficiencia sináptica se representa por los factores de peso de interconexión W_{ij} , desde la neurona i hasta la neurona j .

Los pesos (W_{ij}) pueden ser positivos (excitación) o negativos (inhibición). Los pesos junto con las Funciones de Activación $f(z_j)$ dictan la operación de la red neuronal.

Si la Sumatoria del producto de cada entrada por el peso correspondiente $\sum W_{ij} * X_j$, es mayor al valor de la función de activación $f(z_j)$, el perceptrón se activará.

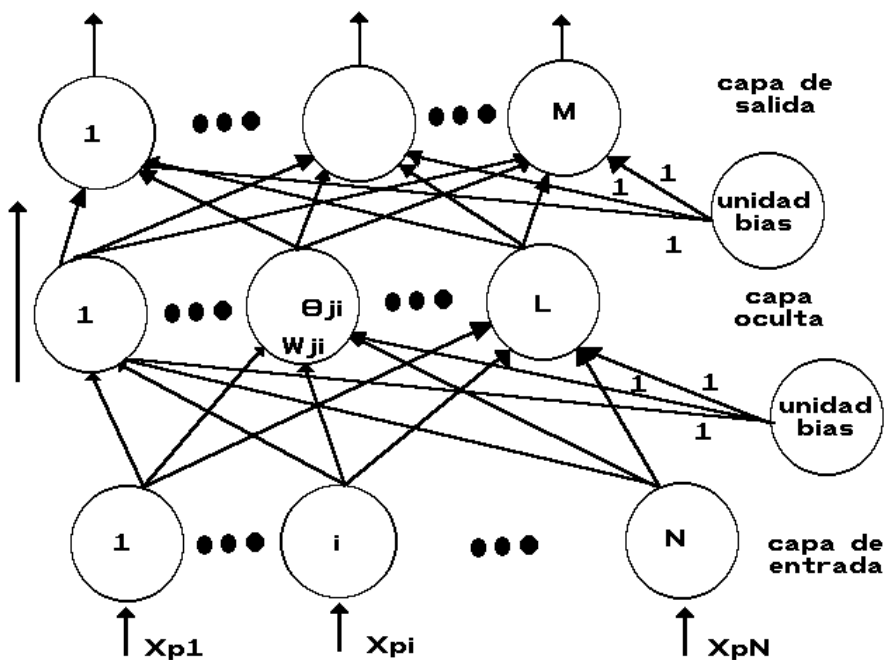
El entrenamiento o aprendizaje de un perceptrón se realiza mediante un incremento diferencial a cada uno de los pesos: Para cada peso W_{ij} se realiza un ajuste dW para cada sinapsis.

El rango de tareas que el Perceptrón puede manejar es mucho mayor que simples decisiones y reconocimiento de patrones.

2.3.2 La Retropropagación

El perceptrón representa un solo elemento de una red neuronal. Si se combinan varios perceptrones en una "capa" y sus correspondientes estímulos de entrada ($W_{ij} * X_{ij}$); se obtendría una red neuronal.

En la red, se interconectan varias unidades de procesamiento en capas. Las neuronas artificiales de cada capa no se interconectan entre sí; sin embargo: La salida de cada neurona artificial de una capa, proporciona "una entrada" a cada una de las neuronas de la siguiente capa. Es decir: Cada neurona transmitirá su señal de salida a cada neurona de la capa siguiente. La figura siguiente, muestra un ejemplo esquemático de la arquitectura de este tipo de red neuronal.



El algoritmo generalizado de aprendizaje o entrenamiento para una capa intermedia u oculta, es el siguiente: Comienza calculando los Valores de Activación (O) de cada neurona, de acuerdo a la siguiente ecuación:

$$O = f(I_k * W_{kn})$$

Donde:

O	representa la salida (Valor de Activación) de cada neurona de la capa anterior ¹
$f(x)$	es la función de Fermi ²
I_k	corresponde al valor de entrada de cada neurona de esta capa
W_{kn}	expresa el Peso asignado a la conexión de la neurona de la capa anterior con la neurona de esta capa.

Después de que todas las neuronas tienen un valor de activación asociado a un patrón de valores de entrada (O), el algoritmo sigue buscando errores en cada neurona que no sea de la primera capa o de entrada. El término "error" se define como la diferencia del valor de salida estimada por la red y valor de entrada, y es parámetro que indica el grado de entrenamiento o aprendizaje de la RNA.

Los errores encontrados para las neuronas de salidas, deben ser "Retropropagados" a la capa anterior, para que puedan ser asignados a neuronas de las capas intermedias u ocultas, para esto se debe minimizar el error en toda la red.

La función que minimiza el Valor de Activación o Salida $O = f(I_k * W_{kn})$ de cada neurona se denomina "Error de la Señal" (δ_i); y se expresa como la Primera Derivada de dicha función:

$$\delta_i = f' (I_k * W_{kn})$$

Este cálculo se repite para cada capa intermedia u oculta de la red.

¹ Para obtener la Salida (O) de la primera capa de neuronas artificiales, se sustituye $(I_k * W_{kn})$ por los estímulos de entrada $(\sum W_{ij} * X_{ij})$, quedando la ecuación de la forma: $O = f(\sum W_{ij} * X_{ij})$

² La Función de Fermi se expresa como: $f(x) = \frac{1}{1 + e^{(-4*(x-0.5))}}$ y gráficamente se representa como un sigmoide en el espacio (o hiperespacio según sea el caso).

Después de que se ha calculado el error δ_i asociado a cada grupo de neuronas, los pesos se deben actualizar, primero encontrando el incremento de valor de que cada peso debe modificarse (ΔW_{ij}), esto se logra calculando:

$$\Delta W_{ij} = O_i * \delta_j * C$$

Donde:	ΔW_{ij}	Incremento del valor del peso
	O_i	Valor de Activación de la capa anterior
	δ_j	Error de la señal de la capa actual
	C	Constante denominada "Razón de Aprendizaje"

El nuevo peso vendrá dado por la relación:

$$W_{ij}(\text{nuevo}) = W_{ij}(\text{anterior}) + \Delta_{ij}$$

En resumen, para cada corrida la RNA calcula el error en la capa de salida; entonces lo propaga a las capa intermedias u ocultas; inmediatamente la red ajusta los pesos de cada neurona y vuele a recalcular los nuevos Valores de Activación o Salida de cada neurona; repitiendo el procedimiento hasta que la Suma del Cuadrado del Error (SCE), sea muy próximo a cero:

$$\sum (salida_{RNA} - entrada_{Datos})^2 \cong 0$$

3.0 Redes Neurales Artificiales vs. Técnicas de Regresión Múltiple

3.1 Preliminares

3.1.1 El Método Clásico del Mercado

Durante muchos años, la metodología clásica de Aproximación al Mercado (Marketing Approach), fue la principal herramienta del profesional tasador de bienes inmuebles. Esta metodología contemplaba el principio de valuación: "Inmuebles similares se venderán a precios similares" y se basaba en la búsqueda de referenciales comparables o similares al inmueble objeto del avalúo en el mercado inmobiliario.

Hasta aquí no había problemas con el método; el mismo era de fácil comprensión y perfectamente válido. Sin embargo, cuando debido a la escasez o insinceridad de comparables, no era posible obtener una muestra representativa de inmuebles similares, se recurría al proceso de corregir u homologar dichos referenciales, mediante expresiones lógico-matemáticas, a veces

empíricas, a fin de “forzar” al dato referencial, para asemejarlo artificialmente al inmueble objeto del avalúo.

El problema, consiste en la presencia de factores de subjetividad en la determinación del valor de un inmueble, causados por los “criterios” empleados por el profesional tasador al aplicar uno o varios factores de corrección a una serie de referenciales, que evidentemente afectan la exactitud del cálculo del valor del bien.

3.1.2 Las Técnicas de Regresión Múltiple

Con la popularización y accesibilidad de las computadoras personales durante la segunda mitad de la década de los ochenta y la liberación de paquetes estadísticos para las mismas, los profesionales tasadores tuvieron en sus manos la posibilidad de utilizar las técnicas de regresión múltiple como una potente e innovadora herramienta para el cálculo del valor de bienes.

Se consideró entonces a las técnicas de regresión múltiple como “el perfeccionamiento de la metodología del mercado”, ya que los propios referenciales se “autocorrigen” entre sí, sin necesidad de utilizar criterios subjetivos por parte del profesional tasador.

Si bien en teoría, las técnicas de regresión múltiples eran la solución a los problemas de subjetividad que adolecía el método clásico de mercado; para que la regresión múltiple convergiera en un valor real se necesitaba que la serie de referenciales cumplieran una serie de requisitos, siendo el principal de ellos la insinceridad de los valores de los inmuebles declarados en las operaciones de compra-venta en las notarías y registros públicos.

Sin embargo, existen otros problemas tan importantes como el precitado, como lo es la inexactitud de los análisis de regresión múltiple frente a la dinámica del mercado inmobiliario, causado principalmente por la discapacidad de los paquetes estadísticos de considerar algo más que reglas y modelos matemáticos rígidos y la inexactitud en el procesamiento de las variables cualitativas / categoriales³ en los modelos de regresión múltiple lineal.

3.2 Propósito de esta Monografía

El propósito de este trabajo, es investigar la aplicación de la tecnología de redes neurales artificiales, en la elaboración de Avalúos Inmobiliarios con la finalidad de comparar la habilidad predictiva de una RNA con modelos de regresión múltiple; de manera de brindar a los profesionales tasadores una novedosa herramienta basada en la Inteligencia Artificial, para la determinación del valor de los bienes y superar los inconvenientes que plantean los modelos de regresión múltiple.

³ Las variables dicotómicas (tal como “Con Vista al Mar” = 1 ó “Sin Vista al Mar” = 0) son un ejemplo de una variable cualitativa / categorial, que ilustra este problema. Debido a que la variable cualitativa y dicotómica “Vista al Mar” no es lineal, al combinarlas con otras variables cuantitativas (tales como área de construcción, edad del inmueble, etc.) en un modelo de regresión lineal múltiple, será muy poco probable estimar o predecir correctamente la variable “Precio Unitario” en función a dichas variables mixtas.

4.0 Bases Teóricas de esta Monografía

4.1 Redes Neurales vs. Modelos de regresión

4.1.1 Introducción

Los Doctores Alfonso Pitarque, Juan Francisco Roy y Juan Carlos Ruiz, profesores de la Facultad de Psicología de la Universitat de Valencia; investigaron la comparación entre las técnicas de simulación de Redes Neuronales Artificiales con modelos estadísticos sobre tareas de predicción y clasificación⁴.

Como herramientas de Predicción o Clasificación, las RNA han sido conceptualizadas como técnicas estadísticas “no paramétricas” al estar libres del cumplimiento de los supuestos teóricos de la estadística paramétrica, o también se conceptualizan como “técnicas de regresión no lineal”.

El problema surge cuando se encuentran resultados contradictorios o disimilares a la hora de determinar: ¿Qué modelos son más eficientes en la solución de problemas concretos de Predicción o Clasificación?.

Pese a que las RNA son capaces de asociar cualquier patrón de entrada con cualquier patrón de salida, su rendimiento va a depender del ajuste heurístico de numerosos parámetros (pesos, valor de activación, error de la señal, función de activación, retropropagación del error, número de capas, coeficiente de aprendizaje, etc., anteriormente explicadas). Ajustes que no siempre garantizan la solución deseada; además de su estructura de “caja negra” que caracteriza a este tipo de modelos.

4.1.2 Análisis y Conclusión de los Resultados

4.1.2.1 Favorables a las Técnicas de Regresión Múltiple:

- a) Únicamente para tareas de predicción cuantitativas y bajo condiciones idóneas de aplicabilidad; el procedimiento de regresión lineal múltiple obtuvo mejores resultados que las RNA.
- b) Para el resto de los casos en las tareas de predicción cuantitativas, no se encontraron diferencias entre RNA y modelos de regresión

4.1.2.2 Favorables a las RNA:

- a) La gran ventaja del uso de las RNA sobre los modelos estadísticos, consiste en que las RNA pueden admitir como variables de entrada: Conjuntos mixtos de variables cuantitativas y cualitativas.

⁴ Cuando se utilizan variables cuantitativas (numéricas) se define como “Técnicas de Predicción o Estimación”; mientras que cuando se utilizan variables cualitativas o categoriales se define como “Técnicas de Clasificación”.

- b) En tareas de Clasificación, las RNA generan resultados mucho mas exactos que los modelos de regresión logística múltiple. En tareas de Predicción o Estimación las RNA y los modelos de regresión lineal múltiple generan resultados similares

4.1.2.3 Desfavorables a las RNA:

- a) Debido a su naturaleza de “caja negra”, una RNA no da información explícita sobre la importancia relativa de los distintos predictores
- b) El entrenamiento de una RNA es un método de ensayo y error; por lo tanto la calidad de las soluciones dadas por una RNA no puede ser siempre garantizada.

4.2 Comparación entre Redes Neurales y Aproximaciones por Regresión Múltiple: Una aplicación para el Avalúo Residencial en Finlandia

4.2.1 Introducción

Olga Karakozova, M.Sc., en su tesis para alcanzar el grado de Magíster en la Academia Sueca de Economía y Administración de Negocios de fecha Octubre del año 2000; presentó una investigación para determinar los méritos de aplicar las técnicas de RNA para la resolución de problemas relacionados con el Avalúo Inmobiliario Residencial; en comparación con modelos de Regresión Múltiple aplicado a mercado inmobiliario del área metropolitana de Helsinki para el año 1998.

4.2.2 Método Empleado

Para este estudio, Karakozova utilizó seis (6) Modelos de Redes Neurales Artificiales y Cuatro (4) modelos de Regresión Múltiple.

Dichos modelos, fueron aplicados para Tres (3) series de datos: La primera, utilizando la data completa de viviendas en el área metropolitana de Helsinki. Para la segunda serie, se restringió el juego de datos a incluir solo viviendas dentro de la ciudad de Helsinki y La tercera Serie, se restringió aún mas la data, a fin de incluir solo viviendas homogéneas dentro de la ciudad de Helsinki.

4.2.3 Resultados

El estudio arrojó el siguiente resultado: Los Modelos de Redes Neurales Artificiales superaron a los Modelos de Regresión Múltiple, para cada una de las series de datos estudiadas. Sin embargo, se observó solo una muy pequeña diferencia entre las RNA y técnicas de regresión múltiple para el caso de la serie de viviendas homogéneas en Helsinki.

4.2.4 Conclusión de los Resultados

- a) En tareas de predicción las RNA supera las técnicas de Regresión Múltiple por pequeño margen

- b) Para series de datos heterogéneos, las RNA superan a las Técnicas de Regresión Múltiple
- c) Las RNA satisfacen los criterios necesarios para las técnicas de avalúos masivos de vivienda.
- d) Las RNA, mas que un sustituto a los análisis econométricos, son un complemento a los mismos

4.3 Análisis del Efecto de la Edad en el Valor de las Viviendas, a través del Análisis de Redes Neurales Artificiales.

4.3.1 Introducción

El profesor A. Quang Do, del Departamento de Finanzas de la Escuela de Administración de Negocios de la Universidad Estatal de San Diego y el profesor G. Grudnitski del Departamento de Contaduría de la Escuela de Administración de Negocios de la Universidad Estatal de San Diego en Diciembre de año 1992; presentaron un estudio donde demostraron que el uso de RNA permite superar los problemas relacionados con el uso de las técnicas de regresión múltiple, como lo son la multicolinealidad, la heterosedasticidad, etc.

4.3.2 Descripción del Problema

El estudio de los Profesores Quang y Grudnitski, se basó examinar la relación que existe entre la edad de una propiedad (vivienda) y su valor de mercado.

Usando técnicas estadísticas de regresión múltiple, se determinó que la edad de un inmueble es inversamente proporcional a su valor durante toda la vida útil del mismo.

Sin embargo, también se demostró que el método utilizado (regresión múltiple) generaba errores significantes. Esta inexactitud era magnificada por problemas inherentes a las técnicas de regresión múltiples tales como interacción entre las variables, variables de comportamiento no-lineal, problemas de multicolinealidad entre las variables independientes, problemas de heterosedasticidad etc.

Las RNA, tienen la particularidad de adaptarse muy bien a la solución de problemas relacionados con el reconocimiento de esquemas, clasificación de variables cualitativas o categoriales; y las mismas se corresponden a técnicas no lineales. Por lo tanto, las RNA poseen la capacidad de proveer una aproximación muy exacta a una amplia clase de funciones no-lineales.

La razón del éxito de las RNA sobre las técnicas de regresión, consiste en la forma de cómo su función de activación es especificada. Mientras que en las técnicas de regresión múltiple, su Función de Activación está pre-especificado independientemente de la data; Las RNA, autodeterminan su Función de Activación “afinando” los parámetros que permiten el mejor ajuste a la data.

4.3.1 Método Empleado

Se tomó una muestra de 242 viviendas unifamiliares (casas), durante el período Enero – Septiembre de 1991, en el suroeste de la ciudad de San Diego (California, EE.UU.), teniendo cuidado de que dichos datos perteneciera a vecindarios similares. Se seleccionaron las siguientes Variables Independientes:

- Edad del inmueble
- Nro. de Habitaciones
- Nro. de baños (1/4 de baño = 1)
- Area de Construcción
- Nro. de Puestos de Estacionamiento
- Nro. de Estaciones de Bomberos en el Vecindario
- Nro. de pisos
- Area del terreno

4.3.2 Resultados y Conclusiones

- a) Se concluyó que la “Edad” de un inmueble es inversamente proporcional a su “Valor”, únicamente durante los primeros 16 a 20 años de su vida útil.
- b) Habiendo transcurrido ese período de tiempo, las variables “Edad” y “Valor”, pasan a relacionarse en forma directamente proporcional.
- c) Demostrando de esta manera que, el uso de técnicas de RNA corrigió el problema de heterosedasticidad presente en el modelo de regresión lineal, que erróneamente determinaba que las variables “Edad” y “Valor” eran inversamente proporcionales durante toda la vida útil de una vivienda.

5.0 Ejemplo Comparativo entre técnicas de Regresión Múltiple y Redes Neurales Artificiales para la determinación del Precio Unitario de apartamentos en la ciudad de Pampatar (Venezuela).

5.1 Selección de la Data

Los datos utilizados para este ejemplo de comparación entre las técnicas de RNA y la regresión múltiple, se corresponde a una muestra de Noventa y un (91) operaciones de compra-venta de apartamentos en la ciudad de Pampatar protocolizados durante el Cuarto Trimestre del año 2001.

Pampatar, es una ciudad ubicada en la Isla de Margarita, actualmente Pampatar está físicamente unida a la ciudad de Porlamar (Principal ciudad de la Isla). La condición de “Puerto

Libre” de la Isla de Margarita, unido a las bellezas naturales de esa porción de tierra caribeña; es el destino turístico preferido tanto de los habitantes de Venezuela como de turistas extranjeros.

Pampatar, está sesgado en dos tipos principales de viviendas: Unifamiliares, destinadas principalmente a personas que viven en la Isla y Multifamiliares (Apartamentos y town-houses), destinadas principalmente a servir como viviendas vacacionales o segundas viviendas.

5.2 Características de las Serie:

5.2.1 Descripción de la data para la aplicación de técnicas de Regresión Múltiple

REGRESIÓN MÚLTIPLE

NUMERO DE DATOS:	91
VARIABLE DEPENDIENTE:	1
VARIABLES INDEPENDIENTES:	7

DESCRIPCION DE LAS VARIABLES:		
VARIABLE	CLASE	TIPO
FECHA	INDEPENDIENTE	CUANTITATIVA
AREA	INDEPENDIENTE	CUANTITATIVA
HAB	INDEPENDIENTE	CUANTITATIVA
BAÑOS	INDEPENDIENTE	CUANTITATIVA
EDAD (1)	INDEPENDIENTE	CUANTITATIVA
VENTA (2)	INDEPENDIENTE	CATEGORIAL
VISTA (3)	INDEPENDIENTE	CATEGORIAL
PU (4)	DEPENDIENTE	CUANTITATIVA

Notas Explicativas a la Regresión Múltiple:

- (1) La Variable Independiente EDAD, viene definida como la fecha de protocolización del Documento de Condominio del edificio donde se ubica el referencial.
- (2) La Variable Independiente VENTA, viene definida como un juego de datos dicotómicos de la forma VENTA = 0 se corresponde a la venta de un apartamento usado. VENTA = 1 se corresponde a la venta de un apartamento nuevo.
- (3) La Variable Independiente VISTA, viene definida como un juego de datos dicotómicos de la forma VISTA = 0 se corresponde a la venta de un apartamento sin vista al mar. VISTA = 1 se corresponde a la venta de un apartamento con vista al mar.
- (4) La Variable Dependiente PU, viene definida como el Precio Unitario del referencial y es el resultado del cociente entre el Precio de Venta y su Area.

5.2.2 Descripción de la data para la aplicación de técnicas de Redes Neuronales Artificiales (RNA)

RED NEURAL ARTIFICIAL

NUMERO DE DATOS:	91
VARIABLE DE SALIDA:	1
VARIABLES DE ENTRADA:	9

DESCRIPCION DE LAS VARIABLES:		
VARIABLE	CLASE	TIPO
FECHA	INDEPENDIENTE	CUANTITATIVA
AREA	INDEPENDIENTE	CUANTITATIVA
HAB	INDEPENDIENTE	CUANTITATIVA
BAÑOS	INDEPENDIENTE	CUANTITATIVA
EDAD	INDEPENDIENTE	CUANTITATIVA
PRIM (5)	INDEPENDIENTE	DICOTOMICA
SEC (6)	INDEPENDIENTE	DICOTOMICA
CON_VIS (7)	INDEPENDIENTE	DICOTOMICA
SIN_VIS (8)	INDEPENDIENTE	DICOTOMICA
PU	DEPENDIENTE	CUANTITATIVA

Notas Explicativas a la RNA:

- (5) La Variable Independiente PRIM, viene definida como una variable dicotómica de la forma PRIM = 0 se corresponde a la venta de un apartamento usado. PRIM = 1 se corresponde a la venta de un apartamento nuevo.
- (6) La Variable Independiente SEC, viene definida como una variable dicotómica de la forma SEC = 1 se corresponde a la venta de un apartamento usado. SEC = 0 se corresponde a la venta de un apartamento nuevo.
- (7) La Variable Independiente CON_VIS, viene definida como una variable dicotómica de la forma CON_VIS = 1 se corresponde a la venta de un apartamento con vista al mar. CON_VIS = 0 se corresponde a la venta de un apartamento sin vista al mar.
- (8) La Variable Independiente SIN_VIS, viene definida como una variable dicotómica de la forma SIN_VIS = 1 se corresponde a la venta de un apartamento sin vista al mar. SIN_VIS = 0 se corresponde a la venta de un apartamento con vista al mar.

5.3 Aplicación de la Técnica de Regresión Múltiple

Para el cálculo del modelo lineal que mejor se ajusta a la serie de datos (Apartamentos en Pampatar), se utilizó la Hoja de Cálculo Microsoft Excel.⁵

⁵ Microsoft Excel forma parte del paquete Microsoft Office 2000 y se corresponde a un paquete de uso general, que posee un completo análisis estadístico así como otras funciones.

5.3.1 Salida del software (Correlación Lineal Múltiple):

202,477.377711	236,109.889352	18.923058	67,939.886667	-24,566.171778	-1,494.834647	-0.988009	-177,121.560286
50,780.004823	45,148.691693	8.533792	36,376.117672	28,070.473549	902.841868	1.387923	292,945.352578
0.754409	139,558.944842	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
36.422931	83.000000	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
4.965789E+12	1.616566E+12	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A

Modelo de Regresión Lineal Múltiple:

$$Y = -177,121.560286 - 0.988009 * X1 - 1,494.834647 * X2 - 24,566.171778 + 67,939.886667 * X4 + 18.923058 * X5 + 236,109.889352 * X6 + 204,477.377711 * X7$$

Variables:

Fecha	Area	Hab	Baños	Edad	Venta	Vista	PU
X1	X2	X3	X4	X5	X6	X7	Y

Coefficiente de Determinación:	0.754409
Estadístico F:	36.422931
Grados de Libertad:	83
Variables Independientes:	7
F de Prueba (Fo):	> 2.36
Error Estándar de la Correlación:	139,558.944842 [Bs/M2]
SCR:	4.965789 E +12 [Bs/M2]
SCE:	1.616566 E +12 [Bs/M2]

5.3.2 Análisis de la Matriz de Correlación:

	CTE	VISTA	FECHA	BANOS	EDAD	HAB	VENTA	AREA
CTE	1	-0-	-0-	-0-	-0-	-0-	-0-	-0-
VISTA	-0-	1	-0-	-0-	-0-	-0-	-0-	-0-
FECHA	-0-	-0.092	1	-0-	-0-	-0-	-0-	-0-
BANOS	-0-	-0.173	0.53	1	-0-	-0-	-0-	-0-
EDAD	-0-	-0.35	-0.67	-0.094	1	-0-	-0-	-0-
HAB	-0-	0.357	0.031	-0.263	0.147	1	-0-	-0-
VENTA	-0-	-0.618	0.067	0.89	-0.372	-0.204	1	-0-
AREA	-0-	0.087	-0.124	-0.549	-0.109	-0.408	0.79	1

Observaciones a la matriz:

- Ninguno de los Coeficientes de Correlación de las Variables Independientes indica una correlación Fuerte (> 0.75).
- No existen problemas de Multicolinealidad entre las Variables Independientes.

5.3.3 Análisis de los Residuos:

Datos =	91
Probabilidad	0.010989011
Sigma =	139,558.94
k =	-2.290744305
k*Sigma =	(319,693.86)
Valores Atípicos =	0

5.3.4 Determinación del Precio Unitario de un Inmueble Patrón

Se determinó el Precio Unitario (PU) de un "Inmueble Patrón", representativo de la serie de 91 Referenciales correspondientes a apartamentos en la ciudad de Pampatar:

Fecha del "Avalúo"	X1 =	11-Abr-2002
Area del Apartamento	X2 =	80.97 [M2]
Habitaciones	X3 =	2
Baños	X4 =	2
Fecha del Documento de Condominio	X5 =	06-Oct-1996
Tipo de Venta	X6 =	0 (Apart. Usado)
Vista al Mar	X7 =	0 (Sin Vista al Mar)
Precio Unitario	Y =	420,503.52 [Bs/M2]

5.4 Aplicación de la técnica de Red Neural Artificial (RNA)

5.4.1 Descripción del software empleado

Para el uso de las técnicas de inteligencia artificial, se utilizó el software "Ainet". Este programa, se basa en Redes Artificiales Neuronales con una topología de Red Perceptrón Multicapa con Retropropagación de error.

Esta aplicación, es uno de las mas sencillas herramientas para la resolución de problemas que necesiten el uso de Redes Neuronales Artificiales (RNA) para alcanzar un resultado.

El algoritmo utilizado por "Ainet", no requiere una fase de entrenamiento o aprendizaje de la RNA; y el resultado se genera inmediatamente obteniendo una gran velocidad de solución a este tipo de problemas.

El algoritmo de “Ainet”, tampoco requiere especificar los “Pesos” iniciales; simplemente un coeficiente denominado “Coeficiente de Penalidad”, controla la variabilidad de la distribución de los pesos. Así como tampoco requiere la especificación del número de capas de perceptrones intermedias u ocultas, ni especificación de las conexiones entre las neuronas⁶.

Adicionalmente, la interfase de usuario es muy simple de operar, los datos entran y salen con un simple “copiar y pegar”, muy similar a una hoja de cálculo.

Sin embargo, el software “Ainet” adolece de estimadores estadísticos de uso común como lo es el Coeficiente de Determinación. Debido a su algoritmo, mas parecido a un circuito de audio; “Ainet” genera como indicador principal de la bondad del ajuste, el estadístico RMS (Raíz Cuadrada del Promedio del Error). Por lo tanto, para poder comparar la RNA con la Regresión Lineal Múltiple, los estadísticos Residuo, Coeficiente de Determinación, SCE, SCR, SCT y F, se calcularon manualmente a través de la “Predicción” (Y Calculado) de cada referencial, calculada por el programa.

5.4.2 Característica de la Red Neural:

Tipo:	MLBP ⁷
Variables de Entrada (9):	FECHA AREA HAB BANOS EDAD PRIM SEC CON_VIS SIN_VIS
Variables de Salida (1) :	PU

5.4.2 Predicción del Precio Unitario (PU) de un “Inmueble Patrón” a través de la RNA

El software “Ainet”, después de correr la serie correspondiente a 91 apartamentos en la ciudad de Pampatar, utilizando 9 variables de entrada (5 Cuantitativas y 4 Dicotómicas) y una (1) variable de salida (PU) en una Red Perceptrón Multicapa con Retropropagación del Error; predijo como valor unitario del “Apartamento Patrón”:

⁶ Al introducir la data, el software calcula y optimiza automáticamente tanto el número de capas ocultas como las conexiones entre las “neuronas artificiales” necesarias para llegar a un resultado.

⁷ Perceptrón Multicapa con Retropropagación del Error.

Fecha del "Avalúo"	FECHA	Entrada #1 =	11-Abr-2002	Cuantitativa
Area del Apartamento	AREA	Entrada #2 =	80.97 [M2]	Cuantitativa
Habitaciones	HAB	Entrada #3 =	2	Cuantitativa
Baños	BANOS	Entrada #4 =	2	Cuantitativa
Fecha del Documento de Condominio	EDAD	Entrada #5 =	06-Oct-1996	Cuantitativa
Tipo de Venta	PRIM	Entrada #6 =	1	Dicotómica
	SEC	Entrada #7 =	0	Dicotómica
Vista al Mar	CON_VIS	Entrada #8 =	1	Dicotómica
	SIN_VIS	Entrada #9 =	0	Dicotómica
Precio Unitario	PU	Salida #1=	492,849.94 [Bs/M2]	Cuantitativa

5.4.3 Análisis de los Estadísticos de Control⁸

SCR=	6.16633E+12
SCE=	4.39901E+11
SCT=	6.58236E+12
Coeficiente de determinación=	0.936797335
k =	9
n =	91
Grados de libertad=	81
Fo =	< 2.24
F =	126.1578518

⁸ Calculados en forma manual a través de la hoja de cálculo MS-Excel

6.0 Conclusiones y Recomendaciones

6.1 Comparación de las Soluciones entre las Dos (2) Técnicas

Al comparar los resultados de aplicar las técnicas de Regresión Múltiple Lineal⁹ y Red Neural Artificial (RNA)¹⁰, sobre la misma data, se obtiene el siguiente resultado:

ESTADISTICO	REGRESION MULTIPLE	RNA
PU [Bs/M2]	420,503.52	492,849.94
VALOR [Bs]	34,048,170.01	39,906,059.64
COEF.DETERMINACION	0.754409182	0.936797335
F (FISCHER a=0.95)	36.42293119	126.1578518
Fo (F prueba a=0.95)	< 2.36	< 2.24
GRADOS DE LIBERTAD	83	81
VAR.INDEPENDINTES	7	9
VAR.DEPENDINTES	1	1

Donde resalta el hecho que las técnicas de RNA explican mejor el comportamiento del fenómeno: “Valor Unitario de Apartamentos en la ciudad de Pampatar”.

6.2 Conclusiones

a) Las técnicas de RNA, explican mucho mejor el comportamiento de fenómenos estadísticos que las técnicas de Regresión Múltiple, en el caso de comportamientos No – Lineales¹¹ de una serie de datos.

b) Se comprueban los resultados de los Doctores Alfonso Pitarque, Juan Francisco Roy y Juan Carlos Ruiz, profesores de la Facultad de Psicología de la Universitat de Valencia, en cuanto a:

- i. La gran ventaja del uso de las RNA sobre los modelos estadísticos, consiste en que las RNA pueden admitir como variables de entrada: Conjuntos mixtos de variables cuantitativas y cualitativas.

⁹ Para caso de la Regresión Múltiple, las Variables “Venta” y “Vista”: Se consideran “Categoriales” mas no “Dicotómicas” (aún y cuando los datos de entrada de esta variable sean 0 y 1).

¹⁰ La diferencia entre el Número de Variables Independientes de la Regresión Múltiple y RNA, de debe al formato de entrada de las variables “Venta” y “Vista”. Los software de RNA consideran a cada variable como una “neurona artificial” de entrada; por lo tanto exigen la diferenciación de las Variable Dicotómicas “Venta” (en: PRIM – SEC) y “Vista” (en: CON_VIS – SIN_VIS) para poder enterarlas en el sistema.⁹¹

¹¹ El solo hecho de existir una mezcla de variables cuantitativas y dicotómicas en una serie de datos, obligatoriamente indica un comportamiento No Lineal del modelo. Las variables dicotómicas no son

lineales ya que las mismas son de la forma: $P(Y = 1) = \frac{e^{a+bX}}{1+e^{a+bX}}$

ii. En tareas de Clasificación, las RNA generan resultados mucho mas exactos que los modelos de regresión logística múltiple.

c) Se comprueban los resultados de Olga Karakozova, M.Sc. de la Academia Sueca de Economía y Administración de Negocios; en cuanto a: "...Para series de datos heterogéneos, las RNA superan a las Técnicas de Regresión Múltiple..."

d) Se comprueban los resultados de los profesores A. Quang Do y G. Grudnitski de la Escuela de Administración de Negocios de la Universidad Estatal de San Diego; en cuanto a: "...El uso de RNA permite superar los problemas relacionados con el uso de las técnicas de regresión múltiple, como lo son la multicolinealidad, la heterosedasticidad, etc..."

e) El uso de la RNA es una poderosa herramienta para el avalúo de bienes muebles e inmuebles, ya que permite obtener resultados coherentes para series donde las técnicas de regresión múltiple no son capaces de converger a un resultado.

f) Debido a la estructura de "Caja Negra", de las técnicas de Redes Neurales Artificiales (RNA). No se puede conocer la descripción del modelo matemático que explica el comportamiento de una serie de datos. Solo se podrá obtener los resultados (Valores Calculados o Predecidos por el software).

g) Se analizaron los softwares de RNA siguientes:

- i. Ainet (Turbajeva 42 SI-3000 Celje. Eslovenia. www.ainet-sp.si)
- ii. BrainMaker Neural Networks (California Scientific Software 10024 Newtown Rd. Nevada City. California 95959. EE.UU. www.calsci.com)
- iii. Phytia – The Neural Network Designer (Runtime Software, EE.UU.)
- iv. Easy NN (S. Wolstenholme, Cheshire, U.K. www.easynn.com)
- v. Pathfinder: Add-in de MS-Excel. Neural Network System (Z Solutions, Atlanta, EE.UU. www.zsolutions.com)

Aunque, todos difieren en su interfase, mas o menos generan la misma información. Algunos son mas complicados que otros en su manejo, funcionamiento e interpretación de los resultados. Pero todos coinciden en el hecho de, que es engorroso adaptarlos como una herramienta sencilla de análisis para ser usados por tasadores sin conocimientos básicos sobre Inteligencia Artificial y RNA.

16-Abr-2002

Bibliografía

- ADAIR A. y MCGREAL W. (1988) "Computer asissted valuation of residential property". Revista: The Real Estate Appraisall and Analyst. Edición Invierno 1988. pp. 18-21. EE.UU.
- BORST R. (1991) "Artificial Neural Networks: The Next Modeling / Calibration technology for the assessment comunity". Revista: Property Tax Journal. Vol 10. pp. 69-94. EE.UU.
- CHERKASSK V., FRIEDMAN J. y WECHSLER I. (1994) "From statistics to neural network". Springer-Velag. Berlín.
- CROALL L. y COHEN P. (1983) "Industrial applications of neural networks". Springer-Velag. New York.
- HECHT-NIELSEN R. (1990) "Neurocomputing". Addison-Wesley. California.
- HERTA J., KROGH A. y PALMER R. (1991) "Introduction from the theory of neural computation". Addison-Wesley. California.
- HAYNES J. y TAN C. (¿?) "An artificial neural network real estate price simulator". School of Information Technology. Bond University. Austria.
- HILERA J. y MARTINEZ V. (1995) "Redes neuronales artificiales: fundamentos modelos y aplicaciones". Ra-ma. Madrid.
- KARAKOZOVA O. (2000) "A comparison between neural network and multiple regression approaches: An application to residential valuation in findland". Tesis de Maestría de: Swedish School of Economics and Business Administration. Suecia.
- QUANG A. y GRUDNISKI G. (1992) "A neural network approach to residential property appraisal". Revista: The Real Estate Appraisal. Vol 58. Nro. 3 Diciembre 1992. pp: 38-45. EE.UU.
- QUANG A. y GRUDNISKI G. (1993) "A neural network analysis of the effect of age on housing values". Revista: The Journal of Real Estate Research. Vol 8. Nro. 2 Primavera de 1993. pp: 253-264. EE.UU.
- PIOL R. (1989-2002) "Métodos estadísticos aplicados a la valuación de bienes inmuebles". SOITAVE. Caracas.
- PITARQUE A., ROY J. y RUIZ J (1998) "Redes neurales vs. Modelos estadísticos: Simulaciones sobre tareas de predicción y clasificación". Univesitat de Valencia. España.
- SMITH M. (1993) "Neural network for statistical modelling". Van-Nostrand-Reinhold. New York.
- VARIOS AUTORES (2000) "Redes neurales artificiales". TREC Internet. Acceso: <http://electronica.com.mx/neural>
- VON DER BECKE C. (1998) "Glosario de Carlos Von der Becke". Acceso: <http://www.argenet.co.ar/~vo/rumellhrt.html>
- WALKER R. (2000) "Neural network fundamental Appraisal". Acceso: http://www.walkrich.com/value_investing/aprsl.htm
- WASERMAN P. (1989) "Neural computing: theory and practice". Van-Nostrand-Reinhold. New York.