

Juan Carlos Pérez Mesa, Jaime de Pablo Valenciano, Jean-Pierre Levy Manguin
Empleo de redes neuronales de base radial a un modelo econométrico de exportación de tomate
Ciencia Ergo Sum, vol. 14, núm. 1, marzo-junio, 2007, pp. 6-14,
Universidad Autónoma del Estado de México
México

Disponible en: <http://www.redalyc.org/articulo.oa?id=10414102>



Ciencia Ergo Sum,
ISSN (Versión impresa): 1405-0269
ciencia.ergosum@yahoo.com.mx
Universidad Autónoma del Estado de México
México

Empleo de redes neuronales de base radial a un modelo econométrico de exportación de tomate

Juan Carlos Pérez Mesa*, Jaime de Pablo Valenciano** y Jean-Pierre Levy Manguin***

Recepción: 11 de noviembre de 2005
Aceptación: 7 de junio de 2006

*Universidad de Almería.
Departamento de Dirección y Gestión de Empresas.
Cañada de San Urbano s/n 04120 Almería, España.
Correo electrónico: juancarlos@ual.es
**Universidad de Almería.
Departamento de Economía Aplicada.
Cañada de San Urbano s/n 04120 Almería,
España. Correo electrónico: jdepablo@ual.es
***Université du Québec en Outaouais, Québec,
Canada. Correo electrónico:
Jean-Pierre.Levy-Mangin@uqo.ca

Resumen. En este artículo se realiza una comparación entre redes neuronales de base radial y la metodología econométrica tradicional. Ambos sistemas se aplican a una función que explica y predice las exportaciones de tomate de empresas de economía social de la provincia de Almería (España). Los resultados que se obtienen son similares, aunque la fiabilidad del modelo de redes al predecir queda cuestionada. Sin embargo, las redes se configuran como una herramienta fácil de utilizar ya que no están condicionadas por una especificación funcional anterior y una distribución del error.

Palabras clave: función de exportación, modelo econométrico, predicción.

Use of Neural Networks of Radial Base to a Econometric Model of Tomato Exports

Abstract. In this paper a comparison between Radial-Basis Networks and the traditional econometric methodology is realized. Both systems are applied to a function that explains and predicts the tomato exports of one social economy company from the province of Almería Spain. The out come is similar, though the reliability of network model to predict is questioned. Nevertheless, networks are a tool easy to use because they are not determined by a previous functional specification and a distribution of the residuals.

Key words: export function, econometric model, forecast.

Introducción

Este artículo supone una aproximación a las inmensas posibilidades que nos ofrecen los nuevos avances en análisis numérico. La investigación de otras ramas de la ciencia, en este caso, la informática, mediante el desarrollo de una aplicación computacional de Redes Neuronales de Base Radial para el reconocimiento de patrones, nos proporciona una oportunidad de utilización práctica en economía.

La utilización de esta metodología a una función de exportación de tomate, viene motivada por avanzar en el conocimiento del proceso de liberalización que está afectando a las exportaciones de frutas y hortalizas. En concreto, nos hemos centrado en la exportación de tomate español debido a que es uno de los productos más sensibles en cuanto a la competencia de terceros países. En estos momentos a pesar de que España tiene la primacía en las exportaciones en el contexto comunitario; en países concretos, por ejemplo en Francia, Marruecos ha doblegado al potencial exportador español. En este artículo analizamos las exportaciones de las empresas de economía social de Almería (máximas representantes de la venta en destino, es decir, de la negociación directa con clientes fuera de nuestro país), principal provincia exportadora de tomate en España.¹

1. Según datos estadísticos facilitados por la Federación Española de Asociaciones de Productores Exportadores de Frutas, Hortalizas y Plantas Vivas (FEPEX) y para el año 2000, Almería supone el 29% del total exportado en toneladas por España (Murcia es la siguiente provincia en importancia con el 21%).

1. Redes de base radial: metodología

Los primeros trabajos sobre redes neuronales se remontan a los años cincuenta. Estos estudios pretendían mimetizar la estructura del sistema nervioso humano a fin de poder solucionar ciertos problemas que son de difícil resolución mediante una programación algorítmica. Normalmente la referencia histórica más usual es la de Roseblatt (1958) que analizó el dispositivo ‘perceptrón’. Una referencia básica que se recomienda al comenzar el estudio de redes neuronales es McClelland y Rumelhart (1986) o Freeman y Skapura (1991).²

Una red neuronal está compuesta por varios operadores simples –elementos de proceso o nodos– con una pequeña cantidad de memoria. Las unidades se comunican mediante redes (axones) que transportan datos. Grosso modo encontramos seis componentes: 1) unidades de procesamiento (neuronas) que se disponen en capas; 2) un estado de activación que representa, en un instante de tiempo, el nivel de activación de las neuronas; 3) unas salidas de las unidades de activación, que determinan la señal que cada neurona envía sus vecinas; 4) un patrón de conexión (estructura de la red en capas y la intensidad y sentido de las relaciones); 5) una regla de propagación (función de activación); y 6) una regla de aprendizaje.

Nuestro objetivo es ver la capacidad de este tipo de herramientas en contraposición a la *potencia* de los modelos econométricos tradicionales. En este sentido, Caridad y Ceular (2001) utilizan redes neuronales artificiales para comparar los resultados con un modelo lineal de determinación del precio de la vivienda. Otero y Trujillo (1992) establecen comparaciones entre un modelo econométrico tradicional, una modelización ARIMA, y un modelo de redes neuronales, para predecir la tasa de actividad femenina en Andalucía. Richards, Patterson y Van Ispelen (1998) comparan un modelo de redes con otro econométrico de determinación del margen (precio de venta al público-precio en granja) para el tomate fresco en Estados Unidos. El modelo de Perceptrón Multicapa y como algoritmo el *back propagation* ha sido tradicionalmente el tipo de red neuronal más utilizada, por su sencillez, en economía. En nuestro caso utilizaremos otro tipo de red neuronal supervisada, las Redes de Función de Base Radial (*Radial-Basis Function Networks*), también conocidas con las siglas RBF.³ Se utiliza este tipo de red por ser prácticamente desconocida en economía y porque dan un giro en el trabajo de investigación sobre la interpolación tradicional estricta en un espacio multidimensional por lo que en la actualidad son uno de los principales campos en la investigación del análisis numérico

y por tanto su aplicación podría llegar a generalizarse en los próximos años.

El diseño de una red neuronal es un problema de ajuste de una curva o función (aproximación). Desde este punto de vista, el aprendizaje es equivalente a encontrar una superficie en un espacio multidimensional que dé como resultado el ‘mejor ajuste’ para los datos de entrenamiento, siendo medido de una forma estadística. Correspondientemente, la generalización es equivalente a usar esta superficie multidimensional para interpolar los datos del test.

En el contexto de red neuronal, las unidades ocultas suministran un conjunto de funciones que actúan como funciones base para los vectores de entrada cuando son expandidos al espacio de unidades ocultas. Estas funciones se llaman funciones de base radial.

Las funciones de base radial fueron primeramente introducidas en la solución de problemas de interpolación multivariante. Los primeros trabajos se atribuyen a Powell (1987) y ahora es uno de los principales campos en la investigación del análisis numérico.

Bromhead y Lowe (1988) fueron los primeros que exploraron el uso de las funciones de base radial en el diseño de redes neuronales.

Otras contribuciones importantes a la teoría, diseño y aplicación de las redes de función de base radial son los artículos de Moody y Darken (1989), Renals (1989) y Poggio y Girosi (1990). El artículo de Poggio y Girosi enfatizaba el uso de la teoría de la regularización aplicada a esta clase de red neuronal como un método para mejorar la generalización de nuevos datos.⁴ La construcción de las redes de fun-

2. Otra bibliografía, en referencia, al estudio de redes neuronales es: Haykin (1994); Leonar, Kramer y Ungar (1992); Musavi y Liu (1991); Horak (1991); Jetter y Stockler (1991); Thau (1991); Light y Cheney (1991); Pantaleón *et al.* (1993); Werntges (1993) y Corchado *et al.* (2000).
3. Es una modificación respecto del algoritmo de aprendizaje *Back Propagation* (BP). Ambos algoritmos suponen un aprendizaje supervisado, es decir, se intenta ‘mapear’ un conjunto de vectores de entrada y relacionarlos con unos de salida. En general, existe un modelo BP capaz de reproducirse con un algoritmo de base radial, y viceversa. La principal diferencia es que el modelo BP intenta encontrar una única función no lineal (normalmente logística) capaz de representar el problema, sin embargo, las redes de base radial están formadas por un nivel de neuronas con una función radial (en nuestro caso gaussiana) y un nivel integrador (de suma) de la salida del nivel anterior. El entrenamiento de este sistema es un proceso basado en la teoría de la regularización para la aproximación de funciones.
4. Un artículo que profundiza en la dificultad del diseño de redes de base radial es Valls *et al.* (2000).

ción de base radial (RBF), en su forma básica, incluye tres capas totalmente diferentes.

- Capa de entrada: está formada por los nodos fuente (unidades sensoriales).

- Capa intermedia: es una capa oculta de gran dimensión y en la que las unidades (neuronas) que la forman son las funciones base para los datos de entrada.

- Capa de salida: que tiene la responsabilidad en la red de la activación de patrones aplicados en la capa de entrada.

La transformación desde la capa de entrada hasta el espacio de las unidades ocultas se realiza de una forma no lineal, mientras que la transformación desde el espacio de unidades ocultas (capa intermedia) al espacio de salida se realiza de forma lineal.

La justificación matemática para este fundamento podemos encontrarla en el artículo de Cover (1965). En particular, cabe destacar de este artículo que el problema de la clasificación de patrones en un espacio alto-dimensional no lineal es más deseable que sea linealmente separable que en un espacio bajo-dimensional. Por lo tanto, es ésta la razón de hacer alta la dimensión del espacio de unidades ocultas en una red RBF.

A través de un diseño cuidadoso, sin embargo, es posible reducir la dimensión del espacio de unidades ocultas, especialmente si los centros de éstas han sido adaptados.

1.1.El problema de la interpolación

Un punto importante que surge del teorema de Cover en la separabilidad de patrones es que en la resolución del problema de la clasificación de patrones no separables linealmente hay, en general, una ventaja clara que se obtiene acotando el espacio de entrada en un nuevo espacio de mayor dimensión.

Básicamente la acotación no lineal se usa para transformar un problema de clasificación no linealmente separable en otro linealmente separable. De manera similar, podemos usar una acotación no lineal para transformar un problema difícil de filtrado no lineal en otro más fácil que incluya un filtrado lineal.

Consideremos entonces una red de retropropagación con una capa de entrada, una capa oculta simple y una capa de salida que consiste en una unidad simple. Se ha escogido una unidad de salida simple para simplificar la exposición sin pérdida de generalidad. La red es diseñada para ejecutar una acotación no lineal desde el espacio de entrada hasta el espacio oculto, seguido por una acotación lineal desde el espacio oculto hasta el espacio de salida. Denotamos con p la dimensión del espacio de entrada. Entonces, de manera global, la red representa un plano desde el espacio de entra-

da p -dimensional hasta el espacio de salida unidimensional, el cual denotamos como:

$$s: \mathbb{R}^p \rightarrow \mathbb{R}^1$$

Podríamos pensar en s como el hiperplano (grafo) $\Gamma \subset \mathbb{R}^{p+1}$, donde el plano Γ es un dibujo multidimensional de la salida como función de la entrada. En un caso práctico, el plano Γ es desconocido y los datos de entrenamiento son usualmente contaminados con ruido. Consecuentemente, la fase de entrenamiento y la de generalización del proceso de aprendizaje podrían ser vistas como sigue:

- La fase de entrenamiento constituye la optimización del proceso de ajuste para el plano Γ , basado en puntos de datos conocidos presentados en la red en la forma de ejemplos de patrones de entrada-salida.

- La fase de generalización es sinónima a la interpolación entre los puntos de datos, realizando la interpolación junto con el plano construido, generado por el proceso de ajuste como la aproximación óptima al plano Γ .

De esta manera nos hemos dirigido a la teoría de la interpolación multivariable en un espacio alto-dimensional, la cual tiene una larga historia. El problema de la interpolación, en sentido estricto, puede empezar como sigue:

Cogiendo un conjunto de N puntos diferentes $\{x_i \in \mathbb{R}^p \mid i=1,2,\dots,N\}$ y un correspondiente conjunto de N números reales $\{d_i\}$, encontrar la función F que satisfaga la condición de interpolación:

$$F(x_i) = d_i, \dots \quad (1)$$

Obsérvese que para una interpolación estricta como la especificada aquí, el plano de interpolación (la función F) es construido para pasarlo por todos los puntos de datos de entrenamiento.

La técnica de las funciones de base radial (RBF) consiste en coger una función F que tenga la siguiente forma (Powell, 1987):

$$F(x) = \sum_{i=1}^N w_i \phi(\|x - x_i\|) \quad (2)$$

Donde $\{\phi(\|x - x_i\|) \mid i=1,2,\dots,N\}$ es un conjunto de N funciones ϕ (generalmente no lineales), conocidas como funciones de base radial, y $\|x - x_i\|$ denota la distancia euclídea. Los puntos de datos conocidos $\{x_i\}$ son escogidos como los centros de las funciones de base radial.

Insertando las condiciones de interpolación anteriores, obtenemos el siguiente conjunto de ecuaciones lineales para los coeficientes desconocidos (pesos) $\{w_i\}$.

donde

$$, j, i = 1, 2, \dots, N \quad (4)$$

y definiendo d y w como:

$$w = [w_1, w_2, \dots, w_N]^T \quad (6)$$

Los vectores d y w de longitud N representan el vector de salida deseado y , el vector de pesos lineal, respectivamente.

Denotando con Φ a la matriz ($N \times N$) con elementos que llamaremos matriz de interpolación:

(3) distintos, la matriz de interpolación (7) Φ es positiva y así, de este modo, podemos resolver la ecuación (3) para despejar el vector de pesos w , obteniendo:

Donde Φ^{-1} es la inversa de la matriz de interpolación Φ . Aunque en la teoría se asegura una solución al problema de interpolación estricta, en la práctica no lo podemos resolver cuando la matriz Φ está arbitrariamente cerca de ser singular. En este punto, la teoría de la regularización puede ayudar, y lo hace cambiando la matriz de interpolación Φ por $\Phi + \lambda I$.

2. Un modelo para explicar las exportaciones de tomate de las empresas de economía social de la provincia de Almería

2.1. Metodología

En este apartado intentaremos señalar cuáles son los factores determinantes de la demanda exterior de tomate de las empresas de economía social, que se caracterizan por un sistema de comercialización en destino, diferente al existente en alhóndiga.⁵

Chebil y Briz *op. cit.* (2000) estiman una función de exportación hortícola española, utilizando como regresores: a) variables de competitividad del precio de exportación español en comparación con el de la Unión Europea (UE); b) variables de renta en el lugar de destino y c) una variable de demanda interna (producción + exportaciones importaciones). Tambi (1998) utiliza la variable producción total en su función de exportación de cacao y café, usando, además, una variable de precios en origen y de exportación, debiendo componer una variable adicional de relación entre ambas cotizaciones para evitar la multicolinealidad; otros prefieren utilizar la variación de producción como variable exógena del modelo (Murúa y Araiztegui, 1994), conjuntamente con los precios a nivel nacional e internacional del producto del cual se intentan explicar sus exportaciones (en su caso almendra).

Las variables utilizadas en el modelo son:

XAL_t = Exportación semanal de empresas de economía social de la provincia de Almería para la campaña 1998/99.

Los datos corresponden a la base de datos de COEXPHAL-FAECA. XMA_t = Exportaciones semanales de Marruecos para la campaña 1998/99.

XFE_t = Exportaciones del resto de asociaciones pertenecientes a FEPEX (Murcia, Alicante, Las Palmas y Tenerife) para la campaña 1998/99.

PX_t = Precios en origen del tomate: corresponde a la cotización media en Alhóndiga para la campaña 1998/99.

Hay una clase de funciones de base radial que cumplen la siguiente propiedad (Light, 1992):

Si x_1, x_2, \dots, x_N son puntos distintos en \mathbb{R}^d . Entonces la matriz de interpolación de ($N \times N$) elementos (Φ) cuyos elementos son $\phi(x_i, x_j)$ es positiva.

Este teorema es más poderoso que el resultado previo atribuido a Micchelli (1986) donde concluía que la matriz de interpolación es no-singular, es decir, que la matriz de interpolación tiene inversa.

El teorema de Light (1992) se aplica, entre otros, a los siguientes casos:

- Multicuadrados inversos

$$\phi(r) = \frac{1}{(r^2 + c^2)^{1/2}} \text{ para cualquier } c > 0, \text{ y } r \geq 0$$

- Funciones Gausianas

$$\phi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \text{ para cualquier } \sigma > 0, \text{ y } r \geq 0$$

No obstante, algunas investigaciones teóricas y resultados prácticos, parecen mostrar que el tipo de función no lineal no es crucial para el rendimiento de la red de función de base radial.

Si retornamos a la implicación del teorema de Light, podemos decir que preservando que los puntos de datos sean

5. Una explicación de esta fórmula de comercio de frutas y hortalizas mediante subasta puede verse en De Pablo y Pérez Mesa (2002).

PAL_t = Producción semanal de las empresas almeriense. Datos COEXPHAL-FAECA. Campaña 1998/99.

El estudio abarca, según observaciones,⁶ desde la semana 41 de 1998 hasta la semana 20 de 1999. Trabajamos con datos transformados en logaritmos, de esta manera los coeficientes que acompañan a las variables serán elasticidades. El modelo propuesto⁷ es el siguiente:

$$XAL_t = a_1 + b_1 XMA_t + b_2 XFE_t + b_3 PX_t + b_4 PAL_t + e_t \quad (8)$$

El análisis de todas las variables que conforman ambos modelos (tabla 1) muestra que no son estacionarias; sin embargo, los residuos de los dos modelos sí lo son (tabla 2 y 3), esto nos indica de la presencia de cointegración. Las exportaciones de Marruecos no resultan significativas (al 5%) para explicar las exportaciones almerienses, al igual que las exportaciones del resto de provincias españolas. Se construye por este motivo un modelo que explica las exportaciones almerienses en función de su producción y los precios en origen⁸ (tabla 3): la elasticidad precio y produc-

Tabla 1. Prueba de Dickey-Full aumentada (ADF) con un rezago.

Variables	(1)	(2)	(3)	$\Delta(1)$	$\Delta(2)$	$\Delta(3)$
XAL_t	-0.600	-1.315	-1.197	-4.210	-3.893	-3.679
XMA_t	-2.967	-1.904	-0.750	-4.745	-4.788	-4.444
PAL_t	-1.472	-1.654	-1.577	-4.651	-3.276	-3.207
XFE_t	-0.426	-1.718	-0.087	-4.424	3.638	-3.341
PX_t	-1.754	-0.150	-0.321	-5.252	4.490	-4.812

(1) Con término independiente y tendencia. Valor Mackinnon al 5% del test ADF para rechazar la hipótesis de raíz unitaria = 3.561.

(2) Con término independiente. Valor Mackinnon al 5% del test ADF para rechazar la hipótesis de raíz unitaria = 2.959.

(3) Sin término independiente ni tendencia. Valor Mackinnon al 5% del test ADF para rechazar la hipótesis de raíz unitaria = 1.951.

- La serie de Marruecos provenía del Ministerio de Agricultura Marroquí. Sin embargo ésta dejó de ser facilitada en cuanto las autoridades de Marruecos descubrieron que se utilizaba para el control del contingente de exportación a la UE, por parte de las asociaciones de empresas exportadoras españolas.
- Si se considera que alguna de las variables independientes está medida con error, es aplicable la estimación con variables instrumentales.
- Sobre esta especificación se construye un modelo con MCE (ver anexo). El MCE es significativo y muy próximo a uno, con signo negativo, lo que nos indica que cada semana se corrige 99% de las desviaciones producidas en el periodo anterior. El modelo propuesto sirve además como test de verificación de la relación de cointegración examinando la significación del coeficiente que acompaña a la corrección del error (Kremers *et al.*, 1992).
- Para la utilización de redes neuronales de base radial se ha empleado una aplicación desarrollada para MATLAB, por el Departamento de Lenguajes y Sistemas de la Universidad de Almería.

Tabla 2. Relación 1 a largo plazo. Variable dependiente

	Coeficiente	t-student	Probabilidad
C	5.256	2.687	0.012
PX_t	-0.276	-2.260	0.014
PAL_t	0.894	13.411	0.000
XFE_t	-0.257	-1.650	0.110
XMA_t	0.054	1.102	0.285
R^2 Aj.	0.968		
DW	2.009		
F	230.980		
Prob	0.000		
ADF (1)	-3.420		
ADF (2)	-3.365		
ADF (3)	-3.498		

(1) Con término independiente y tendencia. Valor Mackinnon al 5% de la prueba ADF para rechazar la hipótesis de raíz unitaria = 3.561.

(2) Con término independiente. Valor Mackinnon al 5% de la prueba ADF para rechazar la hipótesis de raíz unitaria = 2.959.

(3) Sin término independiente ni tendencia. Valor Mackinnon al 5% de la prueba ADF para rechazar la hipótesis de raíz unitaria = 1.951.

ción de las exportaciones almerienses son inferiores a la unidad. Los ajustes, medidos por el R^2 , son en los dos casos superiores al 95%.

2.2. Aplicación de redes neuronales de base radial

Con el modelo econométrico (8), aplicamos la metodología de redes neuronales de base radial.⁹ Vamos a probar la inclusión de las variables, que mediante análisis econométrico tradicional no resultaban significativas. El modelo de redes neuronales, puede encontrar relaciones que la metodología tradicional estadística soslaya. De esta forma probaremos dos alternativas: a) una modelización con las variables que nos aconseja la metodología econométrica tradicional (tabla 3), b) otra alternativa que incluye todas ellas (tabla 2).

La construcción de la red neuronal asociada al trabajo es una red de funciones de base radial que se construye de forma dinámica, por lo que dependerá del tamaño de la muestra de entrenamiento y de las condiciones (aleatorias) iniciales.

Se han ensayado 50 modelos en total con la misma arquitectura, el primer grupo (25 modelos) utiliza 20 datos como

Tabla 3. Modelo 2 a largo plazo. Variable dependiente XAL_t (OLS).

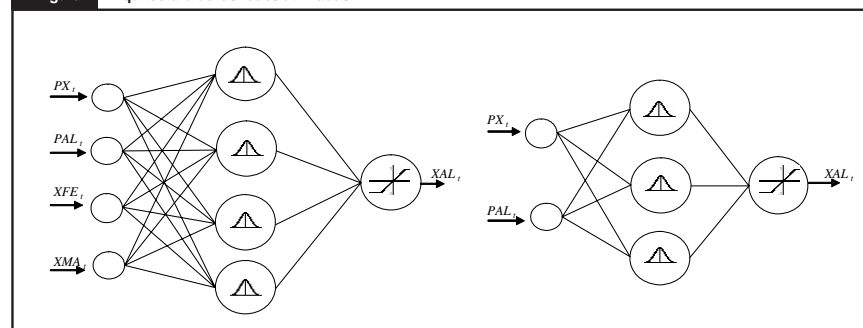
	Coeficiente	t-student	Probabilidad
C	2.621	2.097	0.045
PX_t	-0.266	-2.559	0.016
PAL_t	0.845	15.466	0.000
R^2 Aj.	0.966		
DW	1.772		
F	443.124		
Prob	0.000		
ADF (1)	-3.177		
ADF (2)	-3.128		
ADF (3)	-3.196		

patrones de entrada escogidos de forma aleatoria-. Posteriormente se realiza una estimación con la muestra completa de 32 puntos. Con la red entrenada se realizan las predicciones sobre la serie completa utilizada en el entrenamiento.

Hemos incluido cuatro neuronas en la capa intermedia en el modelo con inclusión de todas las variables y tres neuronas con el modelo de variables significativas (figura 1). El procedimiento que seguimos para escoger el número de neuronas fue el de prueba y error.¹⁰ Se ha utilizado un entrenamiento supervisado.¹¹ La función de activación de la capa oculta ha sido gaussiana. La razón de aprendizaje se ha fijado en 0.01.

Los 12 puntos restantes del primer grupo (para ambos modelos –tablas 4 y 5–) se utilizan como muestra intermedia de validación, obteniendo en su predicción un R^2 medio

Figura 1. Arquitectura de las redes utilizadas.



del 25% y 65% respectivamente, lo que nos muestra un proceso de generalización correcto para el caso de la inclusión de variables significativas y no significativas, es decir, no existe sobreaprendizaje;¹² para el primer caso parece existir un proceso de generalización incorrecto. Se observa, además, cómo los resultados del modelo de redes mejoran con la dimensión de la muestra de aprendizaje (al pasar de 20 a 32 datos). La comparación con el modelo econométrico se realiza con la red entrenada con la totalidad de patrones

Tabla 4. Modelización con variables estadísticamente significativas.

Entrenamiento	R^2
Grupo de datos 1:20	0.4237
Grupo de datos 2:32	0.9394
MCO	0.9688
Prueba sobre residuos del modelo final de redes	
Prueba ADF	Ljung-Box
(1)	2.870
(2)	3.037
(3)	2.961
	Q(1)
	Q(4)
	Q(10)
	0.327
	0.292
	0.382

En redes R^2 = medio de 25 modelos por grupo de datos.

Tabla 5. Modelización con variables estadísticamente significativas y no significativas.

Entrenamiento	R^2		
Grupo de datos 1:20	0.6706		
Grupo de datos 2:32	0.9954		
MCO	0.9704		
Prueba sobre residuos del modelo final de redes			
Prueba ADF	Ljung-Box	Probabilidad	
(1)	5.241	Q(1)	0.128
(2)	5.150	Q(4)	0.112
(3)	5.334	Q(10)	0.505
En redes $R^2 =$ medio de 25 modelos por grupo de datos.			

En redes R^2 = medio de 25 modelos por grupo de datos.

(32 casos): en general se observa, en las tablas 4 y 5, cómo los resultados, al comparar redes y modelo tradicional, son similares. También se puede ver (tabla 5) cómo la inclusión de todas las variables (significativas y no significativas) mejora el ajuste mediante redes.¹³ El análisis de la estacionariedad y la autocorrelación de los residuos, en el modelo de redes, nos puede servir para contrastar la ‘mejorabilidad’ del modelo ya que se puede deber a la existencia de comportamientos no modelizados. En este sentido, las dos versiones muestran autocorrelación

10. Normalmente el número de neuronas de la capa oculta se determina de forma artesanal dependiendo de la experiencia del propio investigador (con carencia de esta experiencia el método de prueba y error es el más utilizado); otro algoritmo típico es el desarrollado por Platt (1991) que consiste en ir añadiendo neuronas a la capa oculta siempre que disminuya el error.

11. En este tipo de entrenamiento, tanto los centros como los pesos, las desviaciones, son calculados en la propia fase de entrenamiento. El entrenamiento de este tipo de aprendizaje se realiza mediante cálculos iterativos basados en el método del gradiente descendente.

12. Uno de los problemas fundamentales que puede aparecer cuando se entrena una red neuronal es que se produzca un sobreaprendizaje, es decir, la red es capaz de aprender de forma precisa los patrones con los que se entrenó, pero cuando se le presentan otros que no tuvieron que ver en el entrenamiento (patrones de validación) devuelve valores alejados de la realidad. Normalmente el error de validación debe ser similar al de entrenamiento, de esta forma se puede ver que la red está llevando a cabo un proceso de generalización correcto. El problema del sobreaprendizaje suele venir provocado porque los patrones de aprendizaje son muy diferentes a los utilizados para validar.

13. La inclusión de variables no relevantes en un modelo econométrico provoca que los estimadores tengan mayores errores estándar: lo que podría provocar errores de especificación (rechazar variables relevantes) que afecten a los resultados de la predicción.

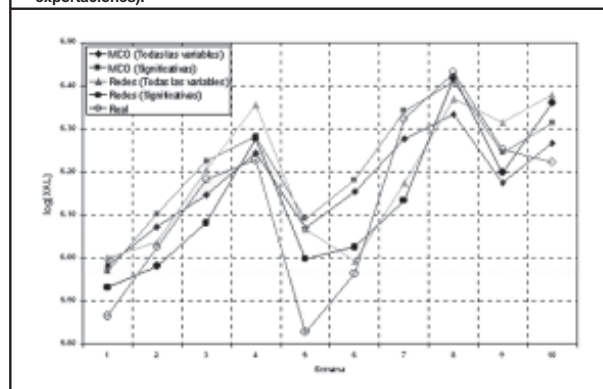
en los residuos, aunque ésta parece mejorar con la inclusión de las variables estadísticamente no significativas. El análisis de los residuos nos sirve como punto adicional de valoración de las redes neuronales.

Al analizar los resultados de la predicción (tabla 6), vemos cómo con la inclusión de las variables estadísticamente significativas, el modelo de red proporciona resultados en la predicción muy similares a los obtenidos al introducir también las variables no significativas, estando por debajo del nivel que los MCO: el mejor modelo de redes alcanza un R^2 de 73% frente al 85% de los MCO. Para completar la información que nos proporcionan los coeficientes de determinación se utiliza la U de Theil como medida del error de predicción (Theil, 1961), su cálculo muestra resultados mejores para los modelos neuronales, sin embargo, para ambos tipos de modelización la U está muy cercana a cero por lo que las diferencias son mínimas.

Tabla 6. Resultados de la predicción con diez puntos adicionales (Semana 41 a 50 de la Campaña 1999/2000).

Resultados de predicción	R^2	Theil's U^{14}
MCO (todas las variables)	0.805	0.019
MCO (variables significativas)	0.858	0.020
Redes (todas las variables)	0.731	0.017
Redes (variables significativas)	0.702	0.017

Gráfica 1. Predicciones sobre los 10 puntos adicionales de la serie de exportaciones).



Fuente: Elaboración propia

14. Donde $U = \left(\sum (\hat{y} - y)^2 / \sum y^2 \right)^{1/2}$, si $U=0$ los pronósticos serán perfectos.

15. Al hablar de mejor modelo nos referimos a la existencia de un R^2 más elevado al realizar el entrenamiento con 32 patrones.

16. Para la campaña 2005/2006 el contingente base se sitúa en 185 000 toneladas, cantidad a la que debemos sumar 38 000 toneladas de contingente adicional.

17. Esta situación parece estar cambiando en campañas más recientes: Almería está incrementando sus exportaciones hacia el mercado alemán entrando en competencia directa con el resto de provincias españolas.

Muchas veces, a la hora de predecir, suele ser más importante la capacidad del modelo para captar los cambios de dirección, que los resultados numéricos (Henriksson y Merton, 1981): si éste fuera nuestro interés vemos (gráfico 1) cómo los dos mejores modelos en predicción (variables significativas – con MCO– e inclusión de todas las variables –con redes–) muestran resultados algo distintos: los MCO consiguen predecir los tres cambios de dirección en la serie, mientras que el mejor¹⁵ modelo de redes falla al predecir la subida de la semana sexta de la serie de predicción. Ambos tipos de modelización predicen un aumento de las exportaciones en la semana décima, cuando en realidad se produce un descenso.

Conclusiones

Se analizan en este apartado las conclusiones más relevantes en cuanto a la interpretación general de modelo de exportaciones de tomate para las empresas de economía social. Los precios en origen tienen influencia en las exportaciones; sin embargo, la elasticidad es menor que uno, lo que nos dice que los cambios de los precios producirán una variación menos que proporcional en las exportaciones provinciales de las empresas de comercialización en destino.

Las exportaciones de tomate marroquí no parecen influir en la exportación de tomate de la provincia. Esto puede deberse a que el límite de exportación a la UE, situado en la campaña 1998/1999 en 168,757 toneladas (el contingente de tomate entre los meses de octubre a marzo se situaba para la campaña 1998/99 en 150,676 tm), era un volumen asimilable por el mercado en destino;¹⁶ lo que no quiere decir que con futuros contingentes la situación no pueda variar.

Las exportaciones del resto de asociaciones FEPEX no resultan significativas, esto es explicable si tenemos en cuenta que en España existe un cierto reparto de los mercados y por tanto no existe una competencia sustitutiva (De Pablo, Pérez y García, 2004).¹⁷ Los cambios en la producción almeriense producen alteraciones en las exportaciones, en nuestro caso, menos que proporcionales: se aprecia preferencia al mercado interior frente al exterior que se puede deber a la dificultad de colocar 'género' en los mercados de exportación en los momentos de producción no esperada.

Un inconveniente general de la red es que no se posee, en sentido estricto, un modelo sobre el que realizar interpretaciones económicas de los coeficientes, por lo que en principio, las redes neuronales se pueden destinar, más fácilmente, a procesos de predicción que de interpretación; por este motivo, en este apartado, evaluamos las redes neuronales en comparación con la modelización tradicional en su vertiente de predicción.

Los resultados de los modelos desarrollados en este artículo muestran que no parecen existir problemas en las redes al incluir variables estadísticamente no significativas (los resultados son prácticamente idénticos): la inclusión de éstas en modelos de MCO incrementa el coeficiente de correlación y además puede producir problemas en la especificación del modelo y, por ende, afectar la predicción. Este problema no existiría con la modelización de redes neuronales pues no están condicionadas por una especificación funcional anterior y una distribución del error, y son estimadores muy robustos bajo condiciones extremas de no-normalidad (Richards, Patterson y Van Ispelen, 1998).

Los resultados de los modelos de redes son similares a los de MCO (utilizando 32 patrones de entrenamiento, y como medida de bondad el R^2); sin embargo, al realizar

predicciones con datos fuera de la serie original, éstos son inferiores que los obtenidos mediante modelización econométrica. Al realizar un análisis de los residuos de los modelos de redes vemos que existe autocorrelación, lo que sirve de aviso de una ‘potencialidad’ no explotada en el modelo.

A modo de coda, aunque a primera vista ambas fórmulas de modelización son similares, los modelos de redes muestran condicionantes que cuestionan su fiabilidad, como son: la existencia de sobreentrenamiento (generalización incorrecta: inferiores resultados en la predicción) e incluso patrones no aleatorios en los residuos. Por tanto, en el ejemplo desarrollado, no existirían evidencias concluyentes para seleccionar las redes neuronales frente a la metodología tradicional.

Bibliografía

- Bromeard, D. S. y D. Lowe (1988). “Multivariable Function Interpolation and Adaptive Networks”, *Complex System*. 21.
- Caridad, J. M. y N. Ceular (2001). “Un análisis del mercado de vivienda a través de redes neuronales artificiales”, *Estudios de Economía Aplicada*. Núm. 18.
- Chebil, A. y J. Briz (1999). “Escenario competitivo del sector hortícola español”, *Rev. Distribución y Consumo*. Núm. 52.
- Corchado J. M.; Díaz, F.; Borrajo L.; Fernández, F. (2000). *Redes neuronales artificiales. Un enfoque práctico*, Servicio de publicaciones de la Universidad de Vigo.
- Cover, T.M. (1965). “Geometrical and statistical properties of system of linear inequalities with applications in pattern recognition”. *IEEE Transactions on Electronic Computers*. 14.
- De Pablo, J. y Pérez Mesa, J.C. (2002). “Las Alhóndigas, pasado, presente y futuro”. *Rev. Distribución y Consumo*. Núm. 66.
- De Pablo, J.; J. C. Pérez-Mesa y R. García (2004). *Caracterización de la comercialización hortofrutícola española: una aplicación a la producción intensiva*. Thomson-Civitas, Madrid.
- Freeman, J. A. y D. Skapura (1991). *Neural Networks. Algorithms, Applications, and Programming Technique*. Addison-Wesley Publishing Company, Inc. Massachusetts.
- Haykin, S. (1994). *Neural Networks*. IEEE press.
- Henriksson, R.D. y R. C. Merton (1981). “On Market Timing and Investment Performance, II. Statistical Procedures for Evaluating Forecasting Skills”. *Journal of Business*. 54.
- Horak, D.T. (1991). “Recursive Radial Basis Functions for Multivariable Function Approximation”. Allied-Signal Aerosp. Co., Columbia, MD, USA. Proceedings of the 1991 American Control Conference (IEEE Cat. No. 91CH2939-7). *American Autom. Control Council*, Evanston, IL, USA. Vol.1, pp.25-27.
- Jetter, K. y J. Stockler (1991). “Algorithms for Cardinal Interpolation Using Box Splines and Radial Basis Functions”. *Numerische Mathematik*. Vol.60, Núm.1. FB Math., Duisburg Univ., Germany.
- Kremers, J., Ericsson N. y Dolado, J. (1992). “The Power of Cointegration Test”. *Oxford Bulletin of Economics and Statistics*. Vol. 54.
- Leonard, J.A.; M. A. Kramer y L. H. Ungar (1992). “Using Radial Basis Functions to Approximate a Function and its Error Bounds”. *Dept. of Chem. Eng., MIT, Cambridge, MA, USA. IEEE Transactions on Neural Networks*. Vol.3, Núm. 4; July, pp 624-627.
- Light, W. A. (1992). “Some Aspects of Radial Basis Function Approximation”, en S.P. Singh (edit.). *Approximation Theory, Spline Functions and Applications*. NATO. ASI Series. Vol. 256, Kluwer Acad., Boston.
- Light W.A. y E. W. Cheney (1991). “Interpolation by piecewise-linear radial basis functions. II”, *Journal of Approximation Theory*. Vol.64, núm. 1. Dept. of Math., Lancaster Univ., UK.
- McClelland, J. y Rumelhart (1986). *Parallel Distributed Processing*. Vol. I y 2. MIT Press. Cambridge, MA.
- Micchelli, C.A. (1986). “Interpolation of scattered data: distance matrices and conditionally positive definite functions”. *Constructive Approximation*, 2.

- Moody, J. y C. Darken (1989). "Fast Learning in Networks of Locally Tuned Processing Units". *Neural Computation*. Vol. 1.
- Murúa J. y A. Araiztegui (1994). "Exportaciones de almendra: análisis de calendarios y demanda". *Rev. Española de Economía Agraria*. Núm. 170.
- Musavi, M.T. y W. J. Liu (1991). *Best Match Algorithm with Radial Basis Functions*. Dept. of Comput. Eng., Maine Univ., Orono, ME, USA. *IJCNN-91-Seattle: International Joint Conference on Neural Networks* (Cat. No.91CH3049-4). IEEE. Vol. 2. New York, NY, USA.
- Otero, J. M. y F. Trujillo (1992). "Predicción multivariante y multiperiodo mediante una red neuronal: estudio de un caso". *Estudios de Economía Aplicada. VI Reunión anual de ASEPELT-ESPAÑA*.
- Pantaleón, C.J.; Díaz de María, F.; Figueras Vidal A. R. (1993). "On training RBF networks". *Sixth International Conference Neural Networks and their Industrial and Cognitive Applications*. NEURO-NIMES 93 Conference Proceedings and Exhibition Catalog. EC2 Dept. of Electron., ETSI Telecom-Univ. Cantabria, Santander, Spain.. Nanterre, France.
- Platt, J. (1991). "A Resource-allocating Networks for Function Implementation". *Neural Computation*, 3.
- Poggio, T.; Girosi, F. (1990). "Networks for Approximation and Learning". *Proc. IEEE*, 78.
- Powell, M. J. D. (1987). "Radial basic functions for multivariable interpolation, a review". Mason and Cox (eds.), *Algorithms for Approximation*. Oxford. Clarendon Press.
- Renals, D. (1989). "Radial basics function network for speech pattern classification". *Electronic Letters*, 25, pp. 437-439.
- Richards, T.; Patterson, P.; Van Ispelen, P. (1998). "Modeling Fresh Tomato Marketing Margins: Econometrics and Neural Networks". *Agricultural and Resource Economics*. December.
- Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain". *Psychological Review*. Núm. 65.
- Serrano C. (2000). "Las redes neuronales artificiales", *5campus.org, Sistemas Informativos Contables*. <<http://www.5campus.org/leccion/redes>> (26/02/2002).
- Tambi, E. (1999). "Co-integration and Error Correction Modelling of Agricultural Export Supply in Cameroon". *Journal of Agricultural Economics*, 20.
- Thau R.S. (1991). "Application of Generalized Radial Basis Functions to the problem of object recognition", *Proceedings of the SPIE-The International Society for Optical Engineering*. Vol.1 MIT, Artificial Intelligence Lab., Cambridge, MA, USA.
- Theil, H. (1961). *Economic Forecasts and Policy*. Amsterdam, North-Holland.
- Valls, J. M.; Molina, J.M.; Galván, I. M. (2000). "Sistema Multiagente para el diseño de Redes de Neuronas de Base Radial Óptimas". *Revista Iberoamericana de Inteligencia Artificial*. Núm. 10.
- Werntges, H.W. (1993). "Partitions of unity improve neural function approximators", *IEEE International Conference on Neural Networks* (Cat. No.93CH3274-8). Dept. of Biophys., Dusseldorf Univ., Germany. York, NY, USA.

Sociedad Mexicana de Física · Gobierno del Estado de México · Universidad Autónoma del Estado de México

Convocan al

1er Concurso de Talentos en Física

Estado de México 2007

Podrán participar los alumnos inscritos y con asistencia regular en el sistema de Educación Básica del Estado de México de 2do. y 3er. grado de secundaria. Será por regiones el 28 de abril del 2007.

Mayores informes: Comisión Organizadora. Facultad de Ciencias, UAEM. Unidad Académica El Cerrillo Piedras Blancas, Km 15 Carretera Toluca-Atzacomulco, entronque Tlachaloya, Toluca, Estado de México. CP. 50200.
Dra. María Guadalupe Frías Palos: Teléfono: (722) 2965556 ext 112. Correo electrónico: mgfriaspalos@yahoo.com.mx

