



DEPARTAMENTO DE INFORMÁTICA  
UNIVERSIDAD CARLOS III DE MADRID

# Ingeniería Informática

## Redes Neuronales

Febrero 2004.

### Normas generales del examen

- El tiempo para realizar el examen es de **2 horas**
- No se responderá a ninguna pregunta sobre el examen
- Si se sale del aula, no se podrá volver a entrar durante el examen
- No se puede presentar el examen escrito a lápiz

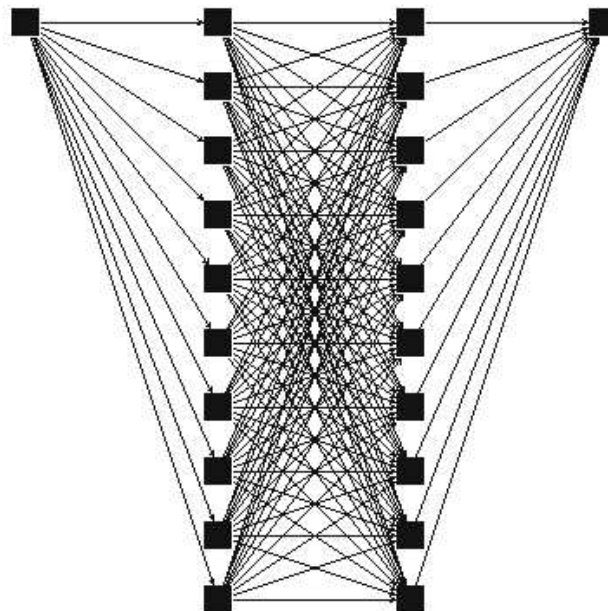
### Problema 1. (3 Puntos)

La empresa “El Corte Francés” encarga a la consultora NERV un estudio sobre las ventas durante el año 2002 para pronosticar las ventas durante el 2003. Para ello les ceden los datos de las ventas totales, en euros, de cada día durante el 2002. En NERV deciden hacer los pronósticos entrenando una red neuronal con estos datos. El modelo de red usado será una red de 4 capas. La capa de entrada tendrá una neurona que recibe el día del año (como un entero entre 1 y 365). La primera y la segunda capas ocultas estarán compuestas por 10 células cada una, y la capa de salida poseerá una única neurona ADALINE en la que mostrará las ventas de ese día en miles de euros.

Las conexiones entre neuronas se harán conectando las de cada capa con todas las de la capa siguiente, de forma que no habrá ningún tipo de recurrencia.

Para el entrenamiento seleccionan los datos de los 300 primeros días del 2002, y para la validación el resto. Finalmente elaboran el informe para “El Corte Francés”.

1. ¿Cuáles piensas que son los puntos débiles del diseño anterior?
2. ¿Crees que la red producirá las salidas deseadas?
3. ¿Qué cambiarías del diseño para que la red fuera más eficaz?



## Problema 2. (3 Puntos)

Se dispone de una serie de observaciones sobre un fenómeno físico. Estas observaciones han sido obtenidas mediante sofisticados aparatos de medida. Se sabe que el fenómeno sigue algún tipo de estructura periódica y simple, pero no se conoce qué estructura es. Los científicos desean determinar dicha estructura para poder realizar predicciones en el futuro. Sin embargo no existe plena confianza en los aparatos de medición. Hay científicos que piensan que las mediciones son exactas y que cualquier modelo que esté de acuerdo con ellas es una alternativa plausible de explicación del fenómeno. Otros científicos, por contra, piensan que el fenómeno no puede ser medido con precisión con la tecnología empleada, y que por lo tanto las señales registradas son únicamente ruido aleatorio.

1. ¿Podrías ayudar a los científicos a decidir qué explicación es la correcta utilizando una red de neuronas?
2. Si es así, ¿de qué forma procederías?

## Problema 3. (4 Puntos)

Realizar una iteración de los patrones de la tabla en una red Adaline en la que: la función de activación es la identidad, las células de la capa de salida tienen umbral, los pesos y umbrales están inicializados a 2 y la tasa de aprendizaje es 0.1.

Patrón	Salida
(1,3,0)	1
(2,4,2)	2
(0,2,1)	2

Hay que indicar cuál es la arquitectura de la red, y los valores de todos los parámetros que intervienen en el aprendizaje, cada vez que se les asigne un nuevo valor.

## Solución al problema 1.

El diseño de red de neuronas de la pregunta tiene varios problemas:

- Solo se introduce una célula de entrada, con lo cual relacionar un único valor numérico, referente al día del año, con las ventas no es muy razonable. El día 127 de un año puede ser domingo y no tener ventas, pero en otro año puede ser lunes y sí tenerlas. Además la información del día no es relevante, tal vez sí el mes, la época del año, el día de la semana.
- Si se toman los primeros 300 días para el entrenamiento y los últimos 65 para la validación, se está introduciendo un sesgo importante que viciará todo el proceso de aprendizaje. Los últimos días del año se corresponden con la campaña de Navidad, en la cual las ventas siguen una estructura absolutamente diferente al resto del año. esto no podría ser extrapolado a partir de la información anterior.

Una alternativa más eficaz podría ser diseñar una red en la que introdujeran los datos de ventas también en la entrada, como una serie temporal. El tamaño de la red dependerá del tamaño de la ventana temporal que se quiera elegir. Debería ser mayor o igual a 7, ya que habitualmente los patrones de ventas se repiten semanalmente. La salida podría ser la misma que en la pregunta, pero los conjuntos de datos deben generarse de otra manera. Una posibilidad es la de elegir días del año de forma aleatoria para evitar introducir cualquier tipo de sesgo.

## Solución al problema 2.

La manera más sencilla de averiguar si los datos son arbitrarios o no es la de intentar resolver el problema mediante una red de neuronas. Si los datos son aleatorios, la red no podrá producir salidas estables por muchos ciclos de aprendizaje que se impongan. Así pues si después de un gran número de ejecuciones de la red el error cometido en aprendizaje oscila de forma arbitraria, se podrá afirmar que los datos que se manejan se corresponden con ruido aleatorio. Esto también podría producirse si la red no fuera capaz de recoger la estructura de los datos, pero en el texto de la pregunta se afirma que se conoce que el fenómeno objeto de estudio es periódico y tiene una estructura sencilla, con lo cual es muy probable que una red backpropagation fuera capaz de determinarlo.

## Solución al problema 3.

La salida del ADALINE se podía haber considerado de dos formas: una

$$y = \sum_{i=1}^3 w_i X_i + u \quad (1)$$

o también como

$$y = \sum_{i=1}^3 w_i X_i + u \quad (2)$$

Con cualquiera de los dos modelos se ha dado como bueno el ejercicio, aunque los resultados finales serán distintos. A continuación se dará el resultado para el modelo 1.

Los patrones son:

Patrón	Salida
(1,3,0)	1
(2,4,2)	2
(0,2,1)	2

Los pesos variarán según la ecuación

$$\Delta w_i = \alpha(d(p) - y(p))X_i(p)$$

donde  $X_i(p)$  es la coordenada  $i$  del patrón  $p$ ,  $d(p)$  es la salida deseada para el patrón  $p$ ,  $y(p)$  es la salida de la neurona para el patrón  $p$ , y  $\alpha$  es la tasa de aprendizaje, con valor 0.1

La variación del umbral será

$$\Delta u = \alpha(d(p) - y(p))$$

si hemos usado la ecuación 1, y

$$\Delta u = -\alpha(d(p) - y(p))$$

si hemos usado la ecuación 2.

Por tanto, para cada patrón (en el orden establecido) tendremos que:

$$w_i = w_i + \alpha(d(p) - y(p))X_i(p)$$

y

$$u = u + \alpha(d(p) - y(p))$$

(recordemos que vamos a resolver el problema usando el modelo 1).

Para el primer patrón, la salida de la neurona será:

$$y = w_1X_1 + w_2X_2 + w_3X_3 + u = 2 \cdot 1 + 2 \cdot 3 + 2 \cdot 0 + 2 = 10$$

Dado que la salida deseada era 1,  $(d - y) = -9$ , y por tanto  $\alpha(d - y) = -0,9$ . Asignamos pues este valor a una constante  $a = -0,9$ . Por tanto, tras la primar iteración, los pesos deberán tomar los valores:

$$w_1 = w_1 + aX_1 = 2 - (0,9 \cdot 1) = 1,1$$

$$w_2 = w_2 + aX_2 = 2 - (0,9 \cdot 3) = -0,7$$

$$w_3 = w_3 + aX_3 = 2 - (0,9 \cdot 0) = 2$$

y finalmente el umbral

$$u = u + a = 2 - 0,9 = 1,1$$

La salida de la red para el segundo patrón será

$$y = 1,1 \cdot 2 - 0,7 \cdot 4 + 2 \cdot 2 + 1,1 = 4,5$$

Dado que la salida deseada para este patrón es 2, la diferencia será  $(d - y) = 2 - 4,5 = -2,5$ . A continuación tomamos la constante  $b = \alpha(d - y) = -0,25$ . Los pesos y el umbral cambiarán siguiendo las ecuaciones:

$$w_1 = w_1 + bX_1 = 1,1 - (0,25 \cdot 2) = 0,6$$

$$w_2 = w_2 + bX_2 = -0,7 - (0,25 \cdot 4) = -1,7$$

$$w_3 = w_3 + bX_3 = 2 - (0,25 \cdot 2) = 1,5$$

$$u = u + b = 1,1 - 0,25 = 0,85$$

La salida de la red para el tercer patrón será

$$y = 0,6 \cdot 0 - 1,7 \cdot 2 + 1,5 \cdot 1 + 0,85 = -1,05$$

Dado que la salida deseada para este patrón es 2, la diferencia será  $(d - y) = 2 + 1,05 = 3,05$ . A continuación tomamos la constante  $c = \alpha(d - y) = 0,305$ . Finalmente los pesos y el umbral cambiarán siguiendo las ecuaciones:

$$w_1 = w_1 + cX_1 = 0,6 + (0,305 \cdot 0) = 0,6$$

$$w_2 = w_2 + cX_2 = -1,7 + (0,305 \cdot 2) = -1,09$$

$$w_3 = w_3 + cX_3 = 1,5 + (0,305 \cdot 1) = 1,805$$

$$u = u + c = 0,85 + 0,305 = 1,155$$