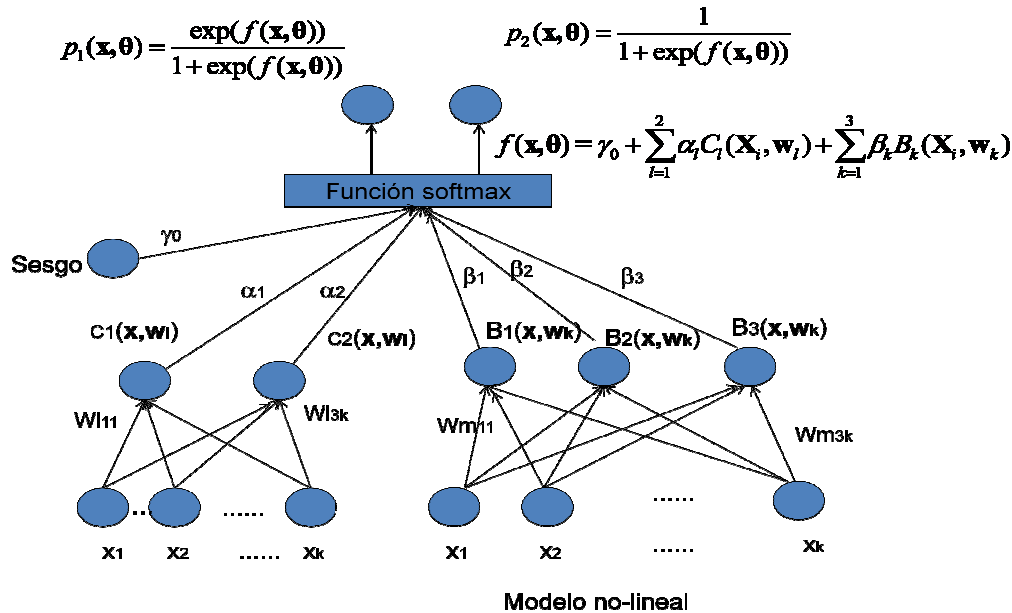


**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

**Ejercicio 1)** Dado el siguiente modelo general de red neuronal, construir uno específico donde haya dos nodos en capa oculta de tipo sigmoidal y dos de tipo unidad producto. Considerar que  $k=2$ .

Con la arquitectura anterior, poner valores a las conexiones entre  $[-10 \text{ y } 10]$  y justificar a que clase pertenecería un patrón cuyo vector de variables independientes es de la forma  $(-2, 2)^T$



La probabilidad de pertenencia a la clase positiva es

$$p_1(\mathbf{x}, \boldsymbol{\theta}) = \frac{\exp(f(\mathbf{x}, \boldsymbol{\theta}))}{1 + \exp(f(\mathbf{x}, \boldsymbol{\theta}))}$$

siendo

$$f(\mathbf{x}, \boldsymbol{\theta}) = \gamma_0 + \sum_{l=1}^2 \alpha_l C_l(\mathbf{X}_i, \mathbf{w}_l) + \sum_{k=1}^3 \beta_k B_k(\mathbf{X}_i, \mathbf{w}_k)$$

**Solución.-** Si, por ejemplo, construimos la función

$$f(\mathbf{x}, \boldsymbol{\theta}) = 3 + 2 \frac{1}{1 + e^{-2+4x_1-2x_2}} - \frac{1}{1 + e^{-3-2x_1-x_2}} + 4x_1^2 x_2^{-1} - 7x_1^3 x_2$$

esta tiene dos nodos ocultos con unidades sigmoideas y dos nodos ocultos con unidades producto. Si sustituimos los valores del patrón  $(-2, 2)$  tenemos que

$$f(\mathbf{x}, \boldsymbol{\theta}) = 3 + 2 \frac{1}{1 + e^{-2-8-4}} - \frac{1}{1 + e^{-3+4-2}} + 4(-2)^2 \frac{1}{2} - 7(-2)^3 2 = 124,27$$

$$P(X=\text{Clase positiva}) = \frac{1}{1 + e^{-124,27}} = 1$$

**Cuestión 2)** Explique brevemente en qué consiste el algoritmo de retropropagación del error en modelos de redes neuronales. Analice el significado de la siguiente regla de actualización de pesos en modelos de redes neuronales.

$$w_k = w_k + \Delta w_k, \text{ donde } \Delta w_k = -\eta \frac{\partial E}{\partial w_k}$$

¿Cómo se obtienen los valores del parámetro  $\eta$ ? ¿Qué significado tiene?

César Hervás Martínez  
Pedro A. Gutiérrez Peña

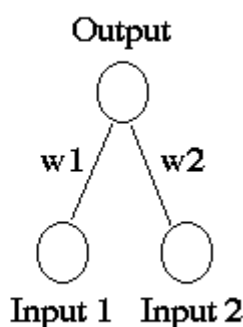
**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

**Ejercicio 3.-** Considere una red neuronal simple compuesta de dos entradas conectadas a una unidad de salida única (Ver Figura). La salida de la red se determina mediante el cálculo de una suma ponderada de sus dos entradas,  $I_1$  e  $I_2$ , y la comparación de este valor con un umbral. Si la entrada de red (net) es mayor que el umbral, la salida es 1, de lo contrario es 0. Matemáticamente, podemos resumir el cálculo realizado por la unidad de salida como sigue:

$$\text{net} = w_1 I_1 + w_2 I_2$$

si  $\text{net} > \theta$  entonces  $o = 1$ , en otro caso  $o = 0$ .

Input 1	Input 2	Output
1	1	1
1	0	0
0	1	0
0	0	0



Encuentre un umbral y unos valores para los coeficientes del modelo de forma tal que se pueda implementar la función AND.

Si ahora tenemos que el patrón (0.0) es de la clase 1 (implementar la función XOR). ¿Es suficiente con el esquema de la Figura 1 o hay que utilizar un modelo diferente? Encuentre un modelo de red neuronal que implemente dicha función.

**Cuestión 4)** La siguiente ecuación que tipo de modelo representa? Explique lo que significan la función, variables y coeficientes que aparecen en el mismo. Construya el grafo del modelo ¿Cuál es su misión?

$$f(\mathbf{x}) = \gamma_0 + \sum_{j=1}^k \alpha_j x_j + \sum_{m=1}^3 \beta_m B_m(\mathbf{x}_i, \mathbf{w}_m),$$

siendo  $B_m(\mathbf{x}_i, \mathbf{w}_m) = \prod_{i=1}^p x_i^{w_{m,i}}$

**Ejercicio 5)** Analice la siguiente ecuación propuesta como función objetivo, E, a minimizar en un modelo de red neuronal de regresión con multivariantes dependientes indicando lo que significa cada sumando y cada variable. ¿Cómo se calcularía el valor  $\hat{y}_{lk}$ ? ¿Para qué se pone el segundo sumando de la función objetivo? ¿Qué ocurre si  $w_{ij} \ll 1$ , para todo  $i$  y  $j$ ? ¿Y si  $w_{ij} \gg 1$ , para todo  $i$  y  $j$ ?

**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

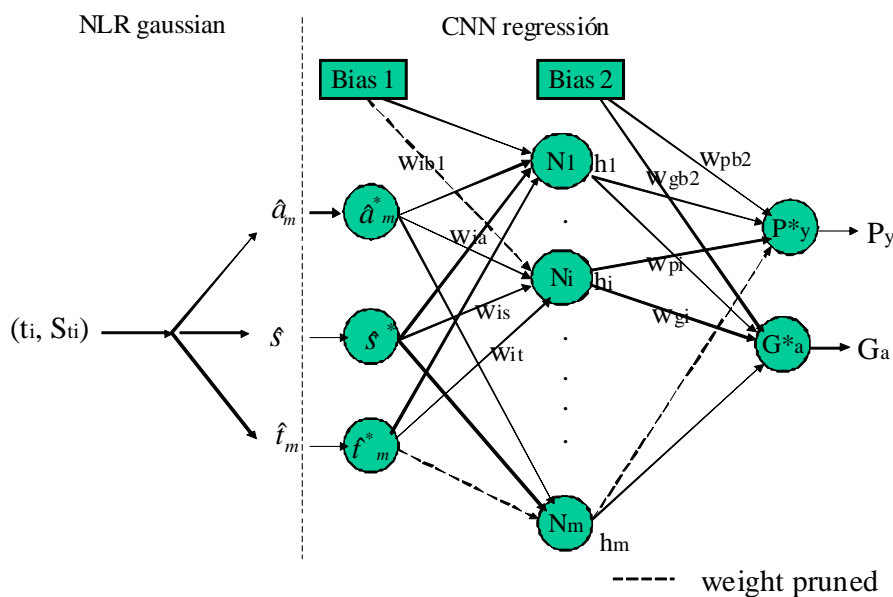
$$E = (1-\lambda) \frac{\sum_{l=1}^n \sum_{k=1}^K (y_{lk} - \hat{y}_{lk})^2}{nK} + \lambda \sum_{ij} \frac{w_{ij}^2}{1 + w_{ij}^2}$$

donde la suma sobre el subíndice l, implica suma sobre todos los patrones del conjunto de entrenamiento, y la suma sobre el subíndice k, implica suma sobre todos los nodos de salida, y donde  $\lambda$  es un parámetro determinado por el investigador.

**Cuestión 6)** ¿Que se entiende por sobre-entrenamiento de una red neuronal? ¿De qué manera se detecta? ¿Cómo se puede evitar? ¿Qué ventajas e inconvenientes tiene considerar un conjunto de validación? Poner un ejemplo en un problema de regresión.

**Ejercicio 7)** Dada la siguiente tabla sitúe los coeficientes del modelo en las conexiones del grafo del modelo de red. ¿Qué tipo de modelo de red es? ¿Qué estructura tiene? ¿Cuál es el valor asociado al patrón de entrada (2,5, 3,2, 0,7) sabiendo que los mínimos y máximos de las variables  $a_m$ ,  $s$  y  $t_m$  para los patrones de entrenamiento son [1-4] [2-4] [0-1] y de las variables  $[P]$  y  $[Q]$ , [1-4] [2-5]?

$\hat{a}_m^*$ ,  $\hat{s}^*$  y  $\hat{t}_m^*$  son valores normalizados, entre 0,1 y 0,9, de las variables de entrada y  $[P]^*$  y  $[GA]^*$  valores normalizados, también entre 0,1 y 0,9, de las concentraciones de los ácidos Pyrogallol y Gálico que forman parte de los nodos de salida.



Procedimiento de estimación de los parámetros del modelo

$$[P]^* = 1.25h_1 - 0.35h_2 + 0.09h_4 \quad [GA]^* = -0.19 - 0.25h_1 + 1.74h_2 - 0.12h_3 - 0.34h_4$$

$$h_1 = \frac{1}{1 + \exp(2.45 - 2.18 \hat{a}_m^* - 1.82 \hat{s}^* - 0.93 \hat{t}_m^*)}$$

$$h_2 = \frac{1}{1 + \exp(0.59 - 3.6 \hat{a}_m^* + 2.07 \hat{s}^*)} \quad h_3 = \frac{1}{1 + \exp(0.79)}$$

César Hervás Martínez  
Pedro A. Gutiérrez Peña

**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

**Ejercicio 8.-**

Se inscribieron los pacientes incluidos en la cohorte HERACLES. Esta cohorte prospectiva observacional incluyó pacientes infectados por el VIH con infección crónica activa por VHC en seguimiento en 19 centros de referencia en Andalucía. Se han utilizado diferentes modelos de redes neuronales de funciones de base para la clasificación automática de pacientes tratados o no tratados. Los pacientes que iniciaron la terapia se clasificaron como aquellos que i) cumplieron con los criterios, o ii) no cumplieron los criterios, según cumplieran o no los criterios para el tratamiento del VHC establecidos en el plan estratégico.

Las 17 características de los pacientes se pueden ver en la tabla Anexo. La variable a predecir es si recibió el paciente, o no, la terapia para la infección por hepatitis C, VHC. Las variables de entrada están normalizadas entre 1 y 2, mientras que la salida está entre 0 y 1.

a) Analice la salida adjunta del mejor modelo de red RBF utilizando el software NNEP. Escriba la ecuación del modelo para la base de datos VIH-VHC, determinando el nº de nodos de la capa oculta, así como el nº de enlaces. ¿El software hace selección de características? ¿El mejor modelo clasifica de forma similar ambas clases? ¿Es un modelo linealizable? ¿Es un modelo interpretable? ¿Cuáles son las variables que más incidencia tienen en el modelo de clasificación binaria?

Generation 50 - Methodology EP

-----  
-4.25 \* ( x4^1.21 \* x9^-0.45 \* x11^-0.90)

+4.32 \* (1)

Train CCR: 75.60

Train Confussion Matrix

-----  
Pred.    0        1        CCR  
Target  
0        1276    171       0.88  
1        364    382       0.51

Test CCR: 77.24

Test Confussion Matrix

-----  
Predicted        0        1        CCR  
Target  
0        444    61       0.88  
1        109    133       0.55

**Solución.-**

a) La ecuación del modelo de red PUNN es

$$f(\mathbf{x}, \boldsymbol{\theta}) = 4.32 - 4.25 \frac{(x_4^*)^{1.21}}{(x_9^*)^{0.45} (x_{11}^*)^{0.90}}$$

El nº de nodos en capa oculta es 1 y el nº de conexiones 5. El software hace selección de características dado que sólo aparecen 3 de las 17 características iniciales ( $x_4$ ,  $x_9$  y  $x_{11}$ )

**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

X4=Reciente categoría PWID (Persons who inject drugs) , X9= Experiencia con tratamiento previo con Peg-IFN/RBV (la combinación de interferón pegilado (P-IFN) y ribavirina (RIB)) y X11=grado de fibrosis del hígado.

El mejor modelo en test tiene una sensibilidad para la clase 1 de 0.88 y para la clase 2 de 0.55, luego clasifica mejor la clase codificada como 0, esto es la clase de pacientes Tratados; donde la clase positiva es la de No Tratados.

Si, es un modelo linealizable dado que si tomamos logaritmos tenemos que

$$\log(4.32 - f(\mathbf{x}, \boldsymbol{\theta})) = \log(4.25) + 1.21\log(x_4^*) - 0.45\log(x_9^*) - 0.90\log(x_{11}^*)$$

de esta forma linealizamos el modelo y lo hacemos interpretable. La variable que más incide en el modelo es la  $x_4^*$ , seguida de la  $x_{11}^*$  y por último de la  $x_9^*$ . Las variables se han transformado en la forma

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + 1$$

b) Las siguientes tablas son algunas de las salidas de 50 ejecuciones. La primera, son los resultados del conjunto de test, mientras que la segunda, son los resultados para el conjunto de entrenamiento. ¿Estamos sobreentrenando? ¿Con que iteración nos quedaríamos? Si

**Test**

Iteración	Generación	CCR	Nº de enlaces
29	50	76,57	8
30	50	77,11	8
31	50	76,44	10
32	50	77,11	7
33	50	77,24	6
Media		76,78	10,58
Desviación típica		0,63	3,52

**Entrenamiento**

Iteración	Generación	CCR	Nº de enlaces
29	50	75,15	8
30	50	75,74	8
31	50	75,51	10
32	50	75,74	7
33	50	75,60	6
Media		75,50	10,58
Desviación típica		0,38	3,52

**Solución.-** No estamos sobreentrenando dado que los resultados de CCR medio en test son superiores a los resultados medios de CCR en entrenamiento. Nos quedaríamos con la iteración 33 dado que tiene el CCR mayor y menor nº de conexiones

c) Dados los resultados obtenidos con modelos de unidades producto (PUNN), sigmoides (SUNN) y RBF (RBFNN). ¿Con que modelo se quedaría? Justifique la respuesta

Modelo Arquitectura (17-1,2,4-1)	Media ± Desviación típica				Mejor modelo				
	CCR	MS	AUC	#con	CCR	MS	AUC	#con	Variables. Independientes
PUNN	0.767 ± 0.004	0.559 ±0.012	0.794 ±0.003	9.46 ± 1.57	0.771	0.579	0.801	11	X2,X4,X5,X9, X11,X13,X14,X17
SUNN	0.762 ±	0.563 ±	0.793 ±	14.10 ±	0.771	0.545	0.799	19	X2,X3,X4,X6,X7,

**GRADO EN INGENIERÍA INFORMÁTICA  
(IMC) CURSO 2021-2022  
EJERCICIOS Y CUESTIONES DE REDES NEURONALES**

	0.006	0.014	0.002	1.63					X9,X10,X11,X13
RBFNN	0.768 ±0.004	0.550 ± 0.008	0.795 ± 0.005	7.36 ± 1.83	0.776	0.550	0.802	8	X3,X4,X8, X11,X13,X16

**Solución.-**

El mejor modelo en CCR , en AUC y en nº de conexiones es el obtenido con la tipología RBF; en cambio el mejor modelo en MS es el obtenido con la tipología PUNN, por tanto este modelo es el que mejor clasifica la clase minoritaria, esto es, la clase 1, los No Tratados.

**APÉNDICE**

**Table 1.** Characteristics of the Spanish HERACLES cohort (NCT02511496).

Description	Variable	Values	Occurrences	Percentage
Met the Spanish criteria plan	X <sub>1</sub>	No	1287	43.77%
		Yes	1653	56.22%
PWID Category	X <sub>2</sub>	Lifetime PWID	2169	73.78%
	X <sub>3</sub>	OST PWID	339	11.53%
	X <sub>4</sub>	Recent PWID	47	1.60%
	X <sub>5</sub>	Never PWID	385	13.10%
Presented major psychiatric disorders	X <sub>6</sub>	No	2886	98.16%
		Yes	54	1.84%
Been in jail	X <sub>7</sub>	No	2823	96.02%
		Yes	117	3.98%
Previous treatment experience	X <sub>8</sub>	Naïve to therapy	2053	69.83%
	X <sub>9</sub>	With Peg-IFN/RBV	725	24.66%
	X <sub>10</sub>	With DAAs/Peg-IFN/RBV	162	5.51%
Liver fibrosis	X <sub>11</sub>	Stage F0-F1	898	30.54%
		Stage F2	475	16.16%
		Stage F3	787	26.77%
		Stage F4	780	26.53%
Gender	X <sub>12</sub>	Female	491	16.70%
		Male	2449	83.30%
Age	X <sub>13</sub>	Continuous variable	min	18
			max	76
			mean	48.95
HCV genotype	X <sub>14</sub>	Genotype 1	1741	59.22%
	X <sub>15</sub>	Genotype 2	27	0.92%
	X <sub>16</sub>	Genotype 3	484	16.46%
	X <sub>17</sub>	Genotype 4	688	23.40%
Received therapy for HCV infection	Output	No	988	33.60%
		Yes	1952	66.40%