Cuestión 1.- Analice la siguiente ecuación propuesta como función objetivo, E, a minimizar en un modelo de red neuronal de regresión con múltiples variables 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_{jik}<<1, para todo i, j y k? ¿Y si w_{jik} >>1, para todo i, j y k?

$$E = (1\text{-}\lambda)\frac{\sum_{i=1}^{n}\sum_{k=1}^{K}(y_{ik} - \hat{y}_{ik})^2}{nK} + \lambda\sum_{jik}\frac{w_{jik}^2}{1+w_{jik}^2}$$

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

Cuestión 2.- ¿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.

Cuestión 3.- Las siguientes ecuaciones 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? ¿Cuál es la regla de decisión para asignar un nuevo patrón?

$$f(\mathbf{x},\boldsymbol{\theta}) = \gamma_0 + \sum_{j=1}^{k}\alpha_j x_j + \sum_{m=1}^{3}\beta_m B_m(\mathbf{x_i},\mathbf{w}_m),\ P(\mathbf{x}\in C_1) = \frac{e^{f(\mathbf{x},\boldsymbol{\theta})}}{1+e^{f(\mathbf{x},\boldsymbol{\theta})}}$$

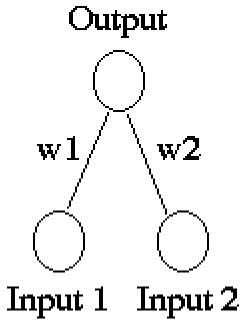
siendo $B_m(\mathbf{x}_i,\mathbf{w}_m) = \prod_{i=1}^{p} x_i^{w_{m,i}}$

Ejercicio 1.- 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:

net = $w_1I_1 + w_2I_2$

si net > θ entonces o = 1, en otro caso o = 0.

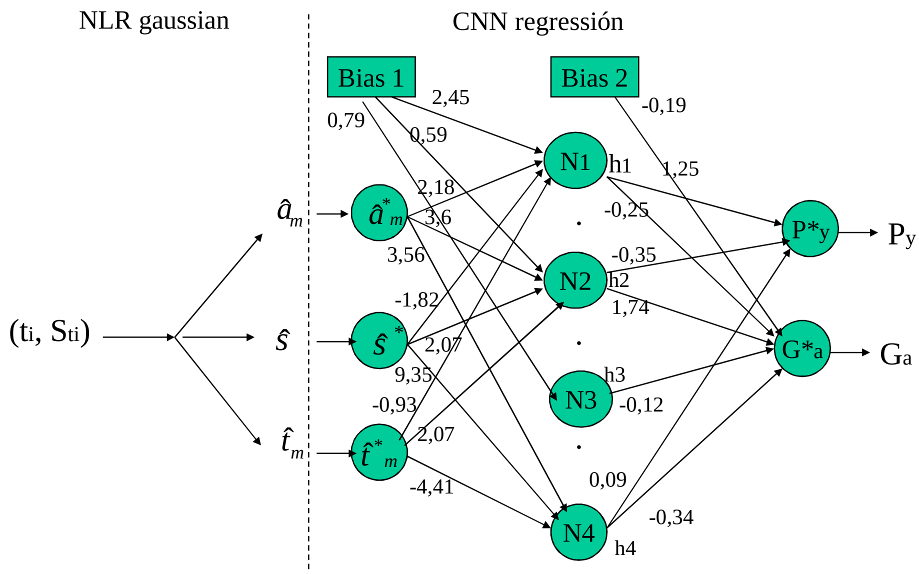
Input 1	Input 2	Output
1	1	1
1	0	0
0	1	0
0	0	0



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 (implementando 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.

Ejercicio 2.- Dado el siguiente grafo sitúe los coeficientes de los modelos en las conexiones 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 am, s y tm para los patrones de entrenamiento son [1-4] [2-4] [0-1] y de las variables [P] y [Q] , [1-4] [2-5]?
$\widehat{a}_m^*, \widehat{s}^* \, y \, \widehat{t}_m^*$ 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.



$$[P]^* = 1.25h_1 - 0.35h_2 + 0.09h_4$$

Grafo del modelo de red neuronal con los parámetros del modelo

$$[GA]^* = -0.19 - 0.25h_1 + 1.74h_2 - 0.12h_3 - 0.34h_4$$

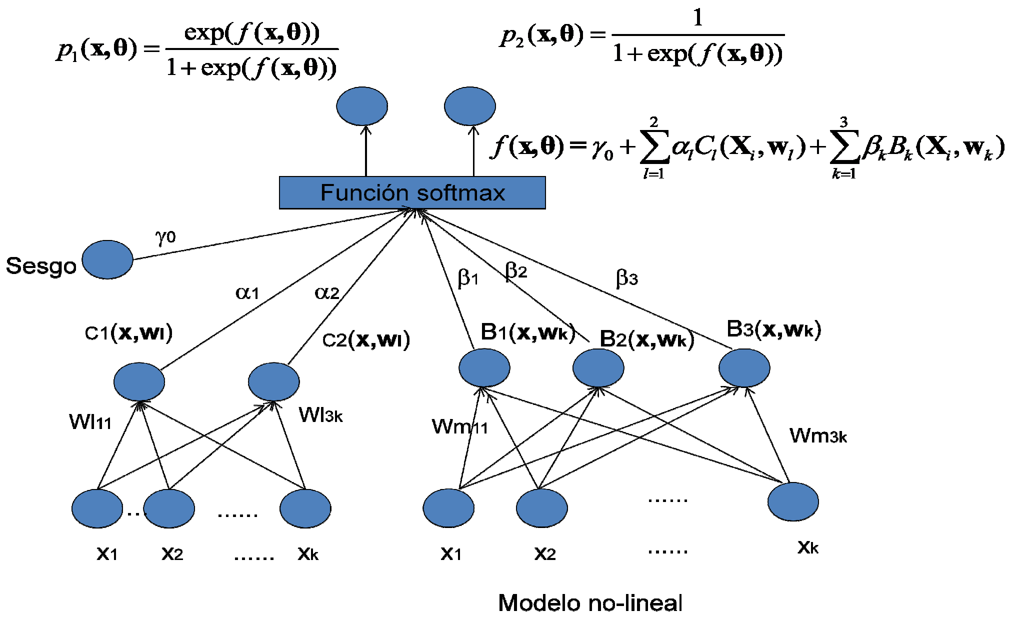
$$h_1 = \frac{1}{1 + \exp(2.45 - 2.18\,\widehat{a}_m^* - 1.82\,\widehat{s}^* - 0.93\,\widehat{t}_m^*)}$$

$$h_2 = \frac{1}{1 + \exp(0.59 - 3.6\,\widehat{a}_m^* + 2.07\,\widehat{s}^*)}$$

$$h_3 = \frac{1}{1 + \exp(0.79)}$$

$$h_4 = \frac{1}{1 + \exp(3.56\,\widehat{a}_m^* - 9.35\,\widehat{s}^* - 4.41\,\widehat{t}_m^*)}$$

Ejercicio 3.- 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 [-1 y 1] 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

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