Ejercicios
Tema 6: Redes Neuronales

Gonzalo A. Aranda Corral

1. Verdadero o Falso. Si tomamos la función *identidad* como función de activación, entonces el perceptrón simple es capaz de aprender la función booleana binaria $f_{AND}(f_{AND}(1,1) = 1, f_{AND}(0,1) = 0, f_{AND}(1,0) = 0, f_{AND}(0,0) = 0)$.

   Si la respuesta es Verdadero tienes que dar un perceptrón que caracterice a f_{AND}. Si la respuesta es Falso, tienes que dar las razones de por qué no puede existir tal perceptrón.

   ¿Verdadero o falso?:....

2. Verdadero o Falso. El perceptrón simple es capaz de representar la fórmula lógica $(A \vee B) \wedge \neg C$ usando la función umbral como función de activación.

$$f(x) = \begin{cases} 1 & si\ x > 0 \\ 0 & si\ x \leq 0 \end{cases}$$

   Si la respuesta es Verdadero tienes que dar los pesos del perceptrón. Si la respuesta es Falso, tienes que dar las razones de por qué no puede existir tal perceptrón.

3. Considera un perceptrón con función de activación sigmoide y pesos $\vec{w'} = (0.1, 0.2, 0.3)$ y un conjunto de entrenamiento $D = \{E_1, E_2\}$ con $E_1 = \langle(0.5, 0.5), 0.3\rangle$ y $E_2 = \langle(0.4, 0.7), 0.2\rangle$. [Nota: Para hacer los cálculos tomaremos los cuatro primeros decimales.]

   (a) Calcula el error cuadrático del perceptrón con pesos $\vec{w'}$ sobre el conjunto de entrenamiento D.

   (b) Calcula en valor del gradiente de la función anterior $\nabla E(\vec{w})$ para el vector de pesos $\vec{w}'$.

4. Consideremos un perceptrón con pesos $\vec{w} = (w0, w1, w2)$, $w0 = w1 = w2 = 0.8$, la función sigmoide como función de activación y un conjunto de entrenamiento $D = \{E_1, E_2\}$ con $E_1 = \langle(0, 1), 1\rangle$ y $E_2 = \langle(1, 0), 0\rangle$.

   (a) Calcular el error cuadrático $E(\vec{w})$ que se comete al considerar este perceptrón sobre ese conjunto de entrenamiento.

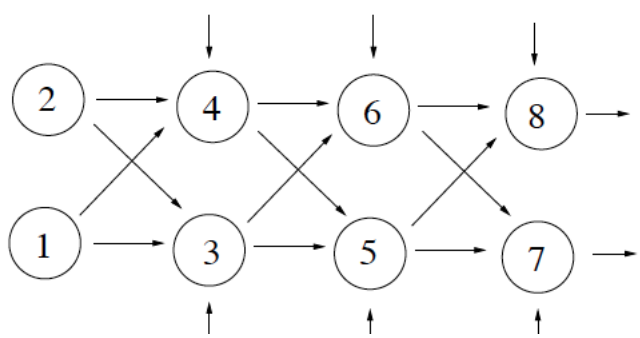
(b) Devolver el valor del peso w_2 tras una actualización usando el algoritmo de entrenamiento del perceptrón por **Descenso por el Gradiente**, con un factor de aprendizaje $\eta = 0.2$. Sólo hay que dar los cálculos necesarios para esa actualización.

5. Aplica el Algoritmo de Entrenamiento de **Descenso por el Gradiente** hasta obtener la primera actualización del peso w_2 al problema de aprendizaje del perceptrón con $\eta = 0.1$, la función sigmoide como función de activación y el conjunto de aprendizaje

	x_1	x_2	x_3	y
e_1	0.7	0.2	0.1	0.3
e_2	0.3	0.5	0.2	0.8
e_3	0.1	0.1	0.8	0.6

Tomar como pesos iniciales $w_0 = 0.1$, $w_1 = 0.1$, $w_2 = 0.1$ y $w_3 = 0.1$.

6. Consideremos la siguiente red neuronal, un factor de aprendizaje $\eta = 0.1$ y la función sigmoide como función de activación. Representaremos por w_{ij} el peso asociado a la sinapsis desde la neurona i a la neurona j y por w_{0j} el peso asociado a la entrada virtual de la neurona j. Consideraremos todos los pesos iniciales de la red neuronal valen 1.0.



¿Cuánto vale el peso w_{36} tras dar un paso del algoritmo de retropropagación asociado al ejemplo $\langle(0,5,0,5),(0,8,0,8)\rangle$? Detallar todos los pasos del algoritmo.

[Nota: Para hacer los cálculos tomaremos 4 decimales.]

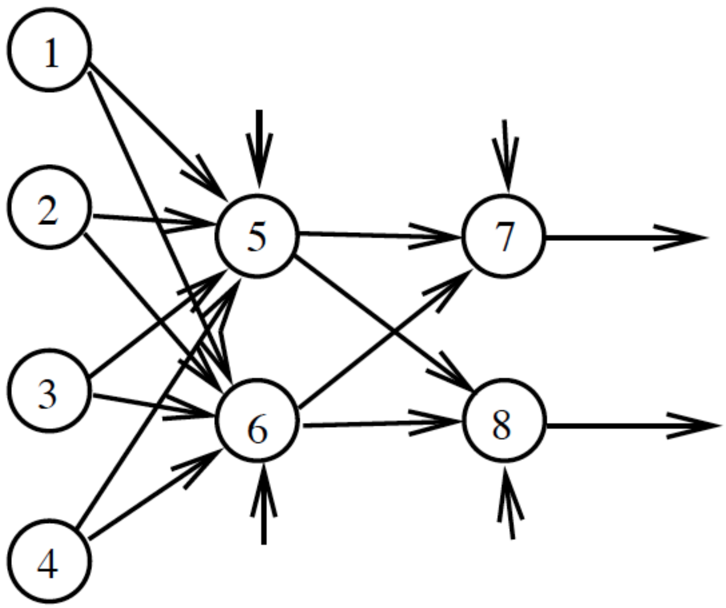
7. Sea f una función de $R \times R$ en $\{-1, 1\}$. Consideremos el problema de aprender f mediante un perceptrón simple bipolar, para ello se tiene el siguiente conjunto de entrenamiento:

	Entradas	Salida
E_1	$(2, 0)$	1
E_2	$(0, 0)$	-1
E_3	$(2, 2)$	1
E_4	$(0, 1)$	-1
E_5	$(1, 1)$	1
E_6	$(1, 2)$	-1

Aplicar el algoritmo de entrenamiento del perceptrón simple bipolar con el conjunto de entrenamiento anterior, considerando los ejemplos en el mismo orden en que aparecen, hasta que se clasifiquen correctamente todos los ejemplos. Tomar 0 como valor inicial para los pesos y 0,1 como factor de aprendizaje.

Con los pesos aprendidos, ¿qué salida se obtiene para las siguientes entradas: (0, 2), (1, 0) y (2, 1)?

8. Sea una red neuronal con la siguiente estructura en la que se usa el sigmoide como función de activación:



Supongamos dado un ejemplo (x_1, x_2, x_3, x_4) con salida esperada $(y7, y8)$.

Supongamos también que ya hemos calculado la salida a_i en cada unidad $i = 1, ..., 8$. Según el algoritmo de retropropagación: cuáles son las fórmulas para calcular los errores δ_8, δ_7 y δ_6 respectivamente? ¿y las fórmulas para actualizar los pesos $w_{6,7}$ y $w_{6,8}$, respectivamente?