Redes Neuronales de Retro propagación

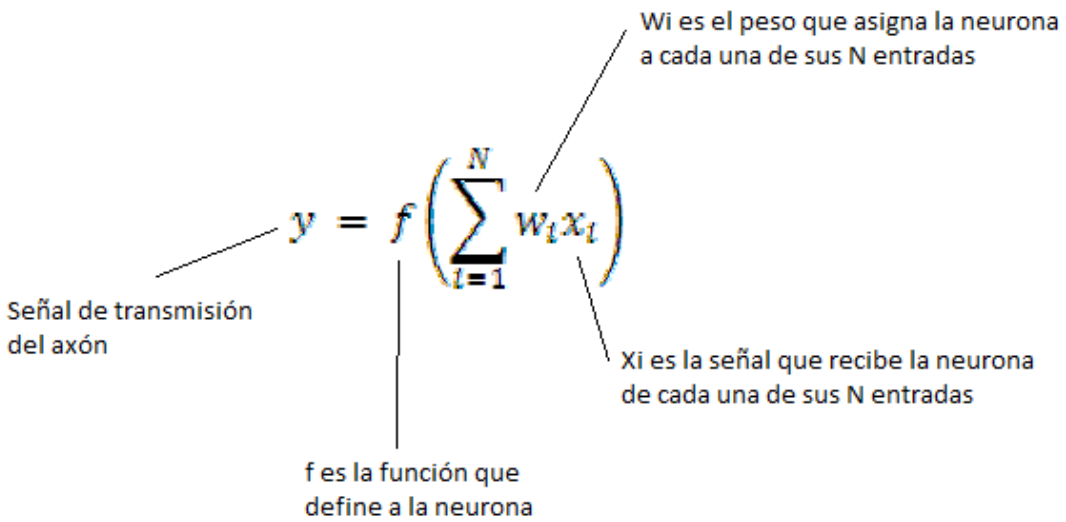
El aprendizaje es un proceso a través del cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el aprendizaje son la destrucción, modificación o creación de conexiones entre las neuronas.

Existen distintos tipos de aprendizajes:

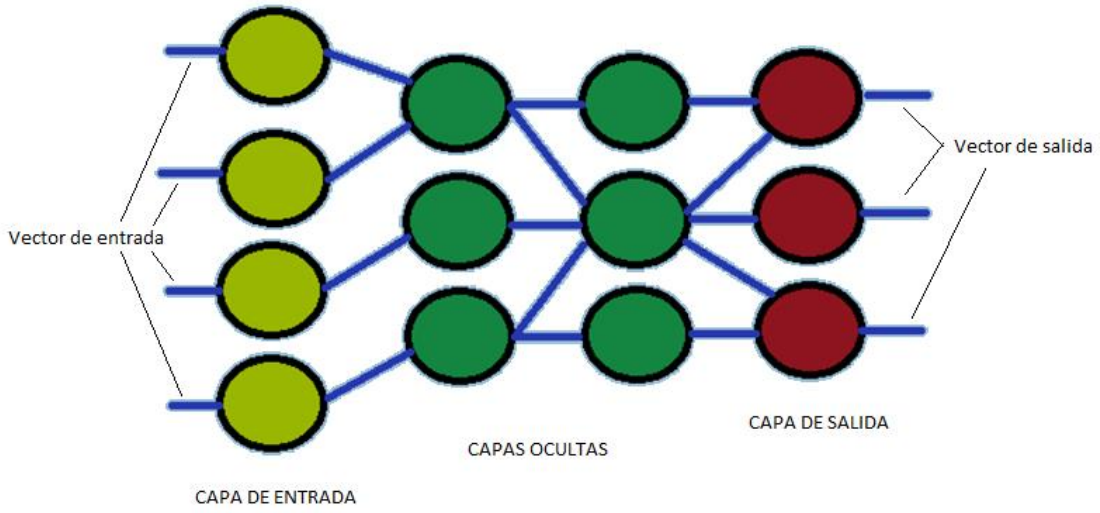
- No Supervisado: a la red solo se le proporciona, datos de entrada y esta debe aprender a categorizarlas (Clustering, SOM)
- Supervisado: a la red se le proporciona un conjunto de datos de entrada y salida conectados. El entrenamiento está controlado por un agente externo que determina la respuesta que debería generar la red a partir de una entrada determinada.

El primer modelo de red neuronal fue el **Perceptrón**. En este modelo se suman las señales de entrada multiplicadas por unos valores de pesos elegidos aleatoriamente. La entrada se compara contra un patrón preestablecido para determinar la salida de la red. Si en la comparación, la suma de las entradas multiplicadas por los pesos es mayor o igual que el patrón preestablecidos la salida de la red es uno (1), en caso contrario la salida es cero (0). El entrenamiento de la red consiste en un proceso de refuerzo mediante el cual, la salida de las unidades se incrementa o decrementan dependiendo de si aciertan o no a las respuestas correctas.

El perceptrón es un tipo de red de aprendizaje supervisado, es decir necesita conocer los valores esperados para cada una de las entradas presentadas. Una neurona se conecta otras, pero nunca con todas. La señal que recibe cada neurona es la suma de las señales que la pasan todas las que están conectadas con ella; entonces la señal que emite el axón es función de la suma de las entradas



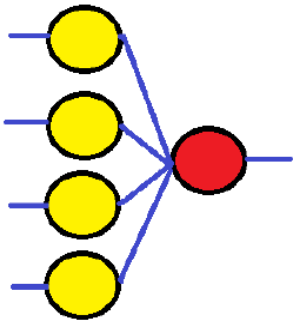
La arquitectura de una red neuronal describe como las salidas de unas neuronas se conectan con las entradas de otras y cuales toman las entradas y proveen salidas fuera de la red.



Del modo que ya describimos, la red servirá para calcular las salidas como función de las entradas. A este modo se lo denomina de transferencia.

Sin embargo la red necesita entrenarse. Para ello se debe asignar pesos W_i para que la red tenga el comportamiento deseado. A este modo se lo denomina de aprendizaje.

La red neuronal más simple para estudiar, tiene una capa de entrada, no tiene capas ocultas y tiene una única salida. A esta red se la denomina “Perceptrón”.



Tener una única salida no significa perder generalidad, ya que se tratará de tener varias veces el mismo esquema y tratarlo en paralelo para llegar a un esquema con varias salidas. Esto es igual a decir que resolvemos el problema de una salida tantas veces separadas como salidas queremos tratar.

Para entrenar al perceptrón, tenemos M casos que asocian un vector de entradas X_i a una salida Y:

$$\left\{\left(X_{1j}, X_{2j}, X_{3j}, \ldots X_{Nj}\right), Y_j \; 1 \leq j \leq M\right\}$$

Se inicializan los valores de los pesos W_i con números aleatorios entre -1 y 1. Para cada uno de los M casos calculamos la salida como función de la entrada obteniendo Z_j

$$z_j = \sum_{i=1}^{N} w_i x_i$$

Si Z_j es igual a Y_j, entonces pasamos al siguiente caso, sino necesitamos cambiar los pesos para entrenar la red.

$$\Delta w_i = \alpha x_{i_j} (y_j - z_j)$$

Luego de haber ajustado los pesos W_i pasamos al siguiente caso hasta agotar los casos disponibles. Con esto el perceptrón queda entrenado.

Un modelo de red con capas ocultas es el algoritmo de retro propagación.

1- Se inicializan los pesos con números aleatorios y pequeños entre -1 y 1.

W_1	W_2	W_3
-0,52	**0,39**	**0,23**

2- Se dispone de un set de datos de entrenamiento, como, por ejemplo:

ID	Estado (X_1)	Humedad (X_2)	Viento (X_3)	Juega Tenis (Y)
1	1	2	2	**0**
2	3	1	2	**0**
3	1	2	1	**0**
4	1	2	1	**0**
5	3	2	2	**1**
6	2	2	2	**1**
7	2	1	2	**1**
8	1	1	2	**1**
9	3	2	1	**1**
10	2	2	1	**1**
11	3	1	1	**1**
12	3	1	1	**1**
13	2	1	1	**1**
14	1	1	1	**1**

3- Luego se recorren los M casos para realizar la retro propagación desde la entrada hacia la salida para calcular Z_j:

$$z_j = \sum_{i=1}^{N} w_i x_i$$

Por ejemplo: 1 * - 0.52 = **- 052**

Una vez obtenidos los productos, se suman los valores de cada caso para obtener Z_j

Estado (W_1X_1)	Humedad (W_2X_2)	Viento (W_3X_3)	$Z_j = \sum W_iX_i$
-0,52	0,78	0,46	0,72
-1,99	0,10	-0,12	-2,01
0,54	1,01	0,74	2,29
0,08	-0,83	0,29	-0,46
0,52	-0,46	0,76	0,82
0,57	-0,31	0,90	1,16
0,44	-0,22	0,78	1,00
0,22	-0,22	0,78	0,78
0,80	-0,35	0,48	0,92
0,62	-0,29	0,49	0,83
1,14	-0,08	0,53	1,59
0,07	-0,19	0,41	0,29
0,90	-0,05	0,55	1,40
0,29	-0,13	0,47	0,63

4- Como la red Perceptrón, solo puede resolver problemas cuyas salidas estén clasificadas en dos categorías diferentes, se asignará a todos aquellos valores que sean mayor o iguales a 0.50 el valor 1 y los que sean menores el valor 0.

Estado (W_1X_1)	Humedad (W_2X_2)	Viento (W_3X_3)	$Z_j = \sum W_iX_i$	Juega Tenis
-0,52	0,78	0,46	0,72	1
-1,99	0,10	-0,12	-2,01	0
0,54	1,01	0,74	2,29	1
0,08	-0,83	0,29	-0,46	0
0,52	-0,46	0,76	0,82	1
0,57	-0,31	0,90	1,16	1
0,44	-0,22	0,78	1,00	1
0,22	-0,22	0,78	0,78	1
0,80	-0,35	0,48	0,92	1
0,62	-0,29	0,49	0,83	1
1,14	-0,08	0,53	1,59	1
0,07	-0,19	0,41	0,29	0
0,90	-0,05	0,55	1,40	1
0,29	-0,13	0,47	0,63	1

5- Para determinar si los pesos permitieron determinar el resultado esperado, se debe calcular el error, restando al Resultado Esperado el Resultado Obtenido, y ese valor se debe restar a los pesos utilizados en la retropropagación para ser utilizados en la próxima.

Entrenamiento

w1	w2	w3	Error
0,00	**0,04**	**0,61**	-1
1,00	1,04	1,61	-1
2,00	2,04	2,61	-1
3,00	3,04	3,61	-1
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0
4,00	4,04	4,61	0

6- Por otra parte, para saber si la red quedó entrenada, se debe verificar el porcentaje de aciertos obtenido. Para ello se debe comparar el resultado obtenido Z_j versus el resultado real Y_1; y se suman la cantidad de aciertos. Si los resultados son aceptables, en cuanto a porcentajes obtenidos, se usan los pesos para realizar reconocimientos sobre nuevos casos.

Juega Tenis (Y)	Suma (Zj)	Juega Tenis	¿Acertó?
0	0,72	1	No
0	-2,01	0	Si
0	2,29	1	No
0	-0,46	0	Si
1	0,82	1	Si
1	1,16	1	Si
1	1,00	1	Si
1	0,78	1	Si
1	0,92	1	Si
1	0,83	1	Si
1	1,59	1	Si

1	0,29	0	No
1	1,40	1	Si
1	0,63	1	Si
			79%

Nota: Ver ejercicio completo en planilla de cálculo adjunto, donde se puede observar cómo se entrena una red de tipo perceptrón. Para llevar a cabo dicho entrenamiento se aplica la fórmula que abajo se describe.