Capítulo 2

Redes neuronales

2.1. Introducción a las redes neuronales

2.1.1. Primeros modelos

En el año 1943 el neurólogo Warren McCulloch y el lógico Walter Pitts publicaron un artículo titulado *Cálculo lógico de ideas inherentes en la actividad nerviosa* en el que plantean el primer modelo matemático que intentaba imitar el comportamiento de una neurona natural.

Para describir el modelo matemático que representa el funcionamiento de una neurona artificial utilizaremos la siguiente notación:

- $X = (x_i)_{i=1,\dots,p}$: vector de entrada
- y: valor de salida
- $W = (w_i)_{i=1,\dots,p}$: vector de pesos
- θ: umbral
- f: función de activación

Así, para obtener el vector de salida se comienza calculando el potencial sináptico que viene dado por la combinación lineal del vector de entrada y el vector de pesos. Estos sirven para aumentar o disminuir los valores de la señal de entrada y por tanto tenemos

$$x_1 w_1 + \cdots + x_p w_p.$$

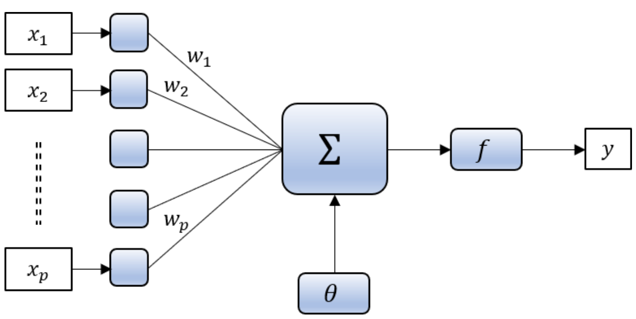

Figura 2.1: Esquema de funcionamiento de una neurona.

Después se resta el umbral, que sirve para limitar el potencial sináptico, y se aplica la función de activación f obteniendo así

$$y = f\left(\sum_{i=1}^{p} x_i w_i - \theta\right).$$

La función de activación utilizada en el modelo de McCulloch y Pitts es la función signo pero en otros modelos neuronales se pueden utilizar la función umbral, la función logística, la función sigmoidea o la tangente hiperbólica.

Basándose en este modelo e implementando el concepto de aprendizaje, Rosenblantt desarrolló un modelo conocido como perceptrón simple. En este se van realizando iteraciones, y, en cada una de ellas, se actualiza el vector de pesos para minimizar el error utilizando la técnica de descenso de gradiente.

Para explicar este método de aprendizaje utilizamos la misma notación que en la neurona explicada anteriormente pero añadimos el superíndice n para indicar el número de la iteración. Además, también añadimos d que representa el valor de la salida esperada y $\eta^{(n)}, \eta^{(n)} \in (0,1]$ que es la tasa de aprendizaje en cada iteración.

Del mismo modo que antes obtenemos la salida de la neurona en la iteración $n-1$

$$y^{(n-1)} = f\left(\sum_{i=1}^{p} x_i^{(n-1)} w_i^{(n-1)} - \theta^{(n-1)}\right)$$

y queremos minimizar el error cuadrático medio

$$E^{(n-1)} = \frac{1}{2}(d - y^{(n-1)})^2.$$

Para ello se utiliza la fórmula de actualización del peso utilizando la tasa de aprendizaje y, derivando, llegamos a la siguiente expresión

$$w_i^{(n)} = w_i^{(n-1)} - \eta^{(n-1)} \frac{\partial E^{(n-1)}}{\partial w_i^{(n)}}.$$

Derivando y utilizando la regla de la cadena llegamos a la expresión final

$$w_i^{(n)} = w_i^{(n-1)} + \eta^{(n-1)}(d - y^{(n-1)}) x_i^{(n-1)}.$$

Este modelo de perceptrón simple tiene una limitación dado que no puede resolver problemas que no sean linealmente separables.

2.1.2. Funcionamiento y clasificación

Debido a esto, surgió la necesidad de crear métodos más complejos que sí fuesen capaces de resolverlos. Es así como nacen las redes neuronales artificiales, que, además de simular el comportamiento de las neuronas, emulan las conexiones entre ellas.

Estos sistemas se organizan en tres tipos de capas que contienen tipos diferentes de neuronas:

- Neuronas de entrada: son las que reciben los datos iniciales y forman la capa de entrada.
- Neuronas de salida: se agrupan en la capa de salida y reciben la información que se extrae como datos finales.
- Neuronas ocultas: forman la capa o capas ocultas. Reciben información de la capa anterior (o bien de entrada o bien otra capa oculta) y emiten una señal a la capa posterior (o bien de salida o bien otra capa oculta). Por tanto, no podemos ni introducir ni extraer información directamente de ellas.

Además, podemos diferenciar distintos tipos de redes neuronales según su número de capas, sus tipos de conexiones o el grado de conexiones.

- Según el número de capas:
  - Redes monocapa: solo tienen la capa de entrada y la de salida, es decir, no tienen ninguna capa oculta.
  - Redes multicapa: tienen al menos una capa oculta entre la capa de entrada y la de salida.

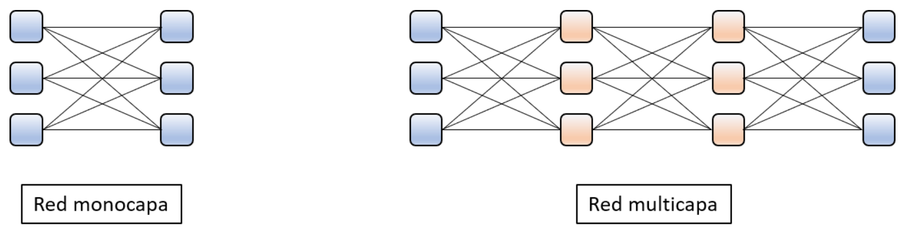

Figura 2.2: Redes neuronales según el número de capas.

- Según el número de conexiones:
  - Redes parcialmente conectadas: algunas neuronas no están conectadas con todas las neuronas de las capas adyacentes.
  - Redes totalmente conectadas: todas las neuronas de cada una de las capas están conectadas con todas las neuronas de cada una de las capas adyacentes.

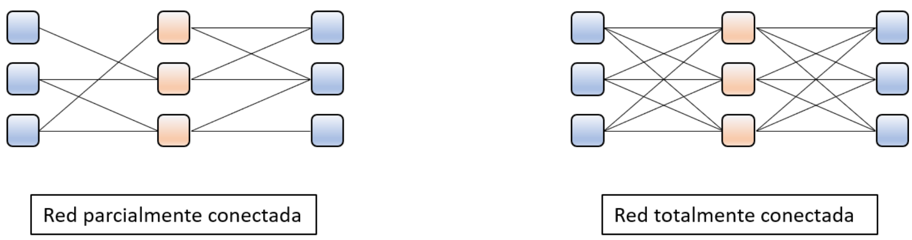

Figura 2.3: Redes neuronales según el número de conexiones.

- Según el tipo de las conexiones
  - Redes no recurrentes o redes hacia delante: la señal siempre viaja hacia delante, es decir, pasa de una capa a la siguiente si que exista retroalimentación.
  - Redes recurrentes: introducen el concepto de memoria, ya que existen conexiones que permiten que la salida de una neurona sea la entrada de una neurona de la capa anterior (o la misma). Esto hace que se incluya la variable del tiempo, ya que la red puede “recordar” entradas y salidas de la misma neurona en instantes anteriores.