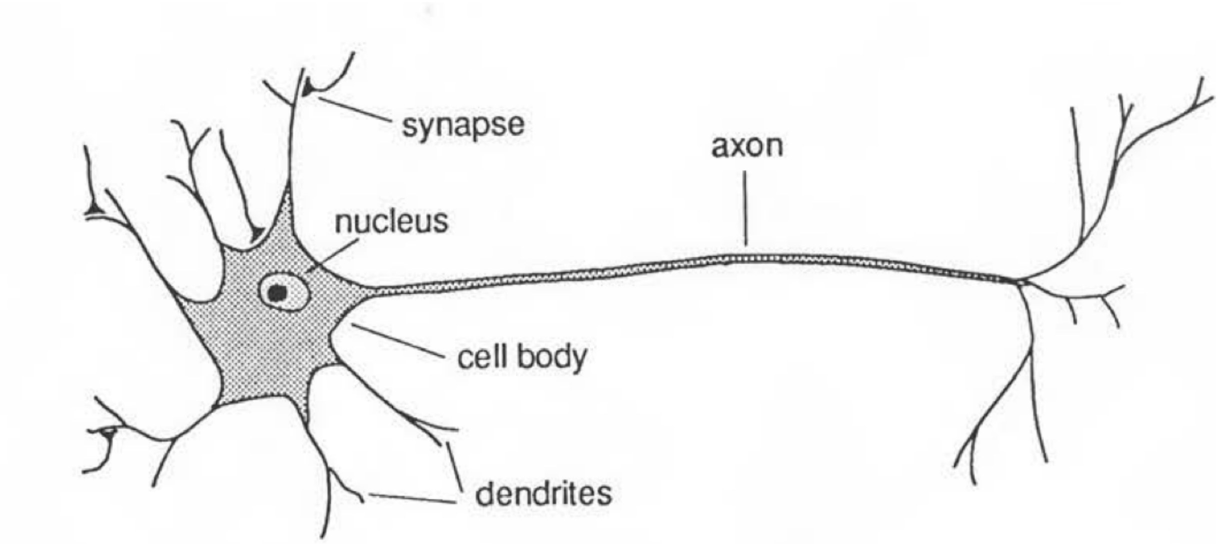
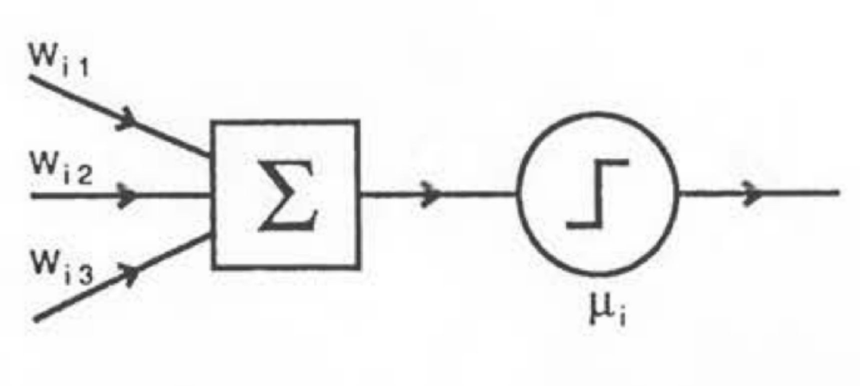
UNA NEURONA ARTIFICIAL

Ya hablamos un poquito de como funciona una neurona biológica



En 1943 el neurólogo y cibernético norteamericano Warren McCulloch, junto al lógico Walter Pitts propusieron un modelo muy simple para emular matemáticamente el funcionamiento de una neurona real. Este modelo de neurona se conoce como McCulloch-Pitts.

El grafico esquematiza el modelo neuronal, que permite ensamblarlas para formar una red neuronal artificial



El cuerpo de la neurona artificial i realiza una suma lineal y ponderada de los estados de las neuronas pre-sinápticas (en este caso 3)

$$h_i = \left(W_{i1} n_i(t) + W_{i2} n_2(t) + W_{i3} n(t) \right)$$

$h_i(t)$ modela el potencial de membrana (entre el interior y el exterior) en el cuerpo de la neurona, el cual surge de la contribución lineal de las neuronas pre-sinápticas.

$n_j(t)$ es el estado de la neurona j

$$n_j(t) = \begin{cases} +1 & \text{si la neurona } j \text{ dispara en tiempo } t \\ 0 & \text{si la neurona } j \text{ no dispara en tiempo } t \end{cases}$$

W_{ij} es la eficiencia sináptica entre la neurona presináptica j y la post-sináptica i.

Puede tomar cualquier valor real

$$W_{ij} \begin{cases} > 0 & \text{sinapsis exhitatoria} \\ = 0 & \text{sinapsis ausente} \\ < 0 & \text{sinapsis inhibitoria} \end{cases}$$

Nos interesa entender como se determina el estado de la neurona post-sináptica i en el tiempo $t+1$

$$n_i(t+1) = \Theta\left(h_i(t) - \mu_i\right)$$

donde $\Theta(x)$ es la función Heaviside o escalón

$$\Theta(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$$

Entonces:

$$n_i(t+1) = \Theta\left(h_i(t) - \mu_i\right)$$

$$= \Theta\left(\sum_{j=1}^{3} w_{ij} n_j(t) - \mu_i\right)$$

$$= \begin{cases} 0 & \text{si } \sum_{j=1}^{3} w_{ij} n_j(t) < 0 \\ 1 & \text{si } \sum_{j=1}^{3} w_{ij} n_j(t) \geq 0 \end{cases}$$

Entonces, nuestra neurona artificial de McCulloch-Pitts calcula la diferencia de potencial en el tiempo t en el cuerpo celular, $h_i(t)$, y la compara con

μ_i : umbral de activación de la neurona i

Si la diferencia de potencial sobrepasa el umbral de activación, la neurona i estará en el estado $+1$, o sea, disparará.

$$h_i(t) \geq \mu_i \Longrightarrow n_i(t+1) = +1.$$

Si la diferencia de potencial no sobrepasa el umbral de activación, la neurona i estará en el estado 0, o sea, disparará.

$$h_i(t) < \mu_i \implies n_i(t+1) = 0 .$$

A lo largo del curso mantendremos la estructura de esta neurona artificial. A lo sumo cambiaremos la función θ, o incluso la complejizaremos definiendo varias funciones por cada neurona.

Con esto podemos definir una RED NEURONAL como un conjunto de N neuronas artificiales. En el tiempo t, la red neuronal está definida por la palabra entera (en términos computacionales)

$$\vec{n}(t) = (n_1(t), n_2(t), n_3(t), \ldots, n_N(t))$$

Pero la red está definida además por la matriz de eficiencias sinápticas

$$\mathbf{W} = \begin{vmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & & & \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{vmatrix}$$

que nos dice quién se conecta con quién, y con que eficacia y con que caracter (si exhitatorio o inhibitorio). Y tambien se define por un vector de umbrales de activación, uno por neurona:

$$\bar{\mu} = (\mu_1, \mu_2, \mu_3, \ldots, \mu_N)$$

En general (salvo casos contados), tendremos que APRENDER los elementos de $\mathbf{W}$ y de $\bar{\mu}$, de forma que asocie correctamente configuraciones iniciales $\bar{n}(0)$ con estados estacionarios $\lim_{t\to\infty} \bar{n}(t)$.

Cuando nos referimos a APRENDER, queremos decir que, partiendo de valores aleatorios de $\mathbf{W}$ y $\vec{\mu}$ vamos variándolos lentamente

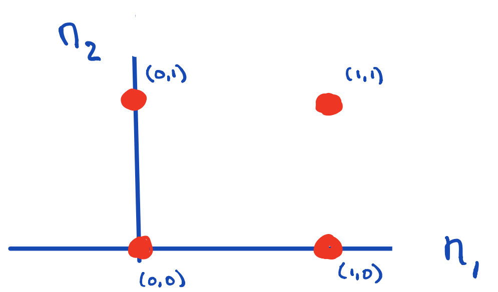
$$w_{ij}^{\text{nuevo}} \longrightarrow w_{ij}^{\text{anterior}} + \Delta w_{ij}$$

$$\mu_i^{\text{nuevo}} \longrightarrow \mu_i^{\text{anterior}} + \Delta \mu_i$$

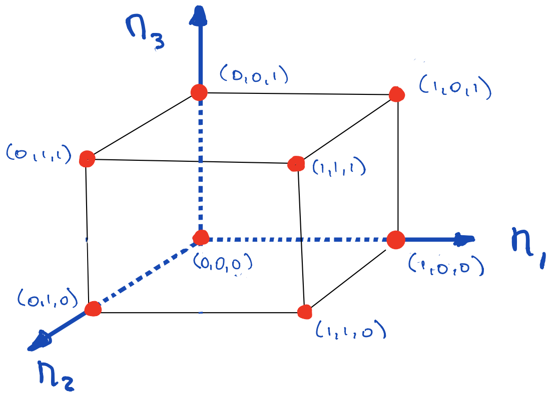
hasta alcanzar un conjunto de valores que cumplen con la premisa de asociar correctamente los estados iniciales con los finales.

Una red de neuronas de McCulloch-Pitts "viven" en un hipercubo de lado 1

$N = 1$



$N = 2$



$N = 3$

Dado que la red tiene N neuronas de 2 estados, el número total de estados posibles es:

$$2^N$$

Si tenemos $N = 4$ hay 16 estado posibles

Miremos una evolución posible

$$\vec{n}(0) = (1, 0, 0, 1)$$

$$\vec{n}(1) = (0, 0, 0, 1)$$

$$\vec{n}(2) = (1, 0, 1, 1)$$

$$\vec{n}(3) = (1, 0, 1, 1)$$

estado estacionario

Si en dos tiempos sucesivos está en el mismo estado, se quedará para siempre en ese estado, pues el sistema es determinista.

La red asoció el estado inicial

input $\qquad\qquad$ output

$$(1, 0, 0, 1) \longrightarrow (1, 0, 1, 1)$$