Diplomatura en Ciencia de Datos, Inteligencia Artificial y sus aplicaciones a la economía y los negocios
FCE y FAMAF

MÓDULO 3

UNIDAD 2

Clase 1 Parte 2

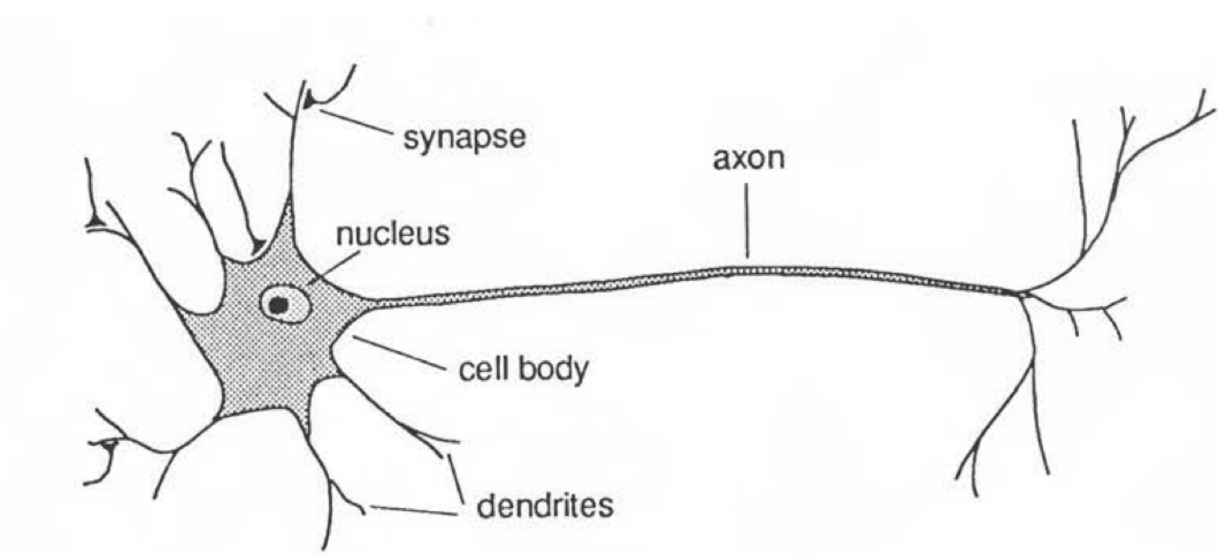
FCE y FAMAF

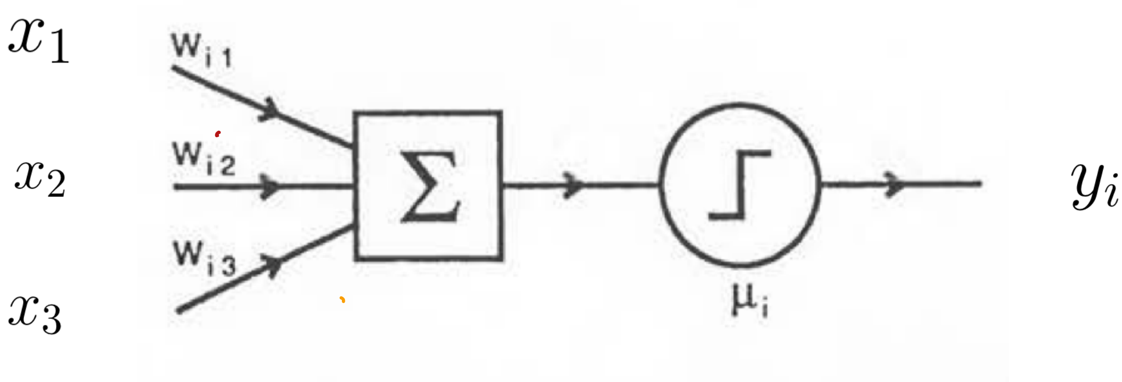
Universidad Nacional de Córdoba
Sábado 23 de octubre 2021

La neurona de McCulloch y Pitts

Comenzamos nuestro camino desde las neuronas artificiales hacia las redes neuronales profundas con la definición de la neurona matemática más simple que capta la esencia del funcionamiento de las neuronas biológicas, la cual se conoce como neurona de McCulloch y Pitts. Veremos que desde su aparición en la literatura hasta hoy, siendo la base de la inteligencia artificial neuronal. Ustedes ya vieron este tema la unidad 1 y ahora lo revistaremos desde una mirada neuronal.

En 1943 el neurólogo y cibernético norteamericano Warren McCulloch junto al matemático especialista en lógica Walter Pitts propusieron este modelo que funciona como unidad de umbral.





Neurona *i*

El gráfico esquematiza el modelo neuronal más simple que conocemos y que permite ensamblar grandes conglomerados de neuronas.

$x_1,\ x_2,\ x_3$ representan los señales que llegan de las cada neurona pre sinóptica

$W_{i1}\,W_{i2}\,W_{i3}$ representan las eficacias de cada sinápsis que forma la neurona *i* con las neuronas que le envían señales

μ_i representa el umbral de activación de la neurona *i*

El símbolo $\sum$ representa la suma ponderada que realiza la neurona *i* en la región de unión entre el cuerpo y el axón.

$$h_i = W_{i1}x_1 + \ldots + W_{iN}x_N = \sum_{K=1}^{N} W_{ik}x_k = \vec{W}.\vec{X}$$

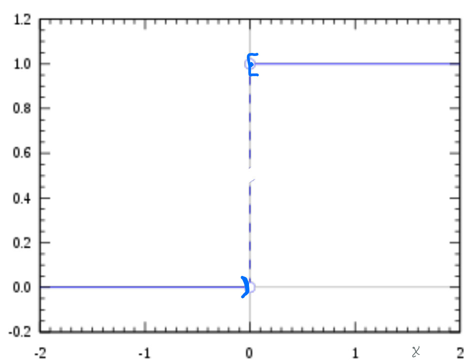
h_i modela el potencial de acción, o sea, la diferencia de potencial entre el interior y el exterior de la neurona

y_i representa la señal que la neurona *i* enviará a través de su axón.

 representa la función de activación Heaviside

Nos interesa ver como la neurona *i* determina su valor de salida y_i en función de los valores de los estados de las tres neuronas pre-sinápticas.

$$y_i = \Theta(h_i - \mu_i)$$



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

Función Heviside

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

$$W_{ij} = \begin{cases} > 0 & \text{si la sinapsis es exhitatoria} \\ = 0 & \text{si la sinapsis está prohibida o ausente} \\ < 0 & \text{si la sinapsis es inhibitoria} \end{cases}$$

Observemos que hemos incluido el tiempo t tanto para describir la neurona i como para las neuronas presinápticas 1, 2 y 3 (que no vemos en el dibujo pero están)

Si tuviéramos N neuronas pre-sinápticas en lugar de solo tres

$$y_i(t+1) = \Theta(h_i(t) - \mu_i) = \Theta\left(\sum_k^N W_{ik}\ x_k\ - \mu_i\right)$$

$$y_i(t+1) = \begin{cases} 1 \ \text{ si } \ \sum_k^N W_{ik} x_k \geq \mu_i \\ \\ 0 \ \text{ si } \ \sum_k^N W_{ik} x_k < \mu_i \end{cases}$$

Entonces, la neurona *i* al tiempo *t* calcula la diferencia de potencial

$h_i(t)$ en su cuerpo debido al input de las neuronas pre sinápticas

como una suma ponderada y compara este valor con el umbral

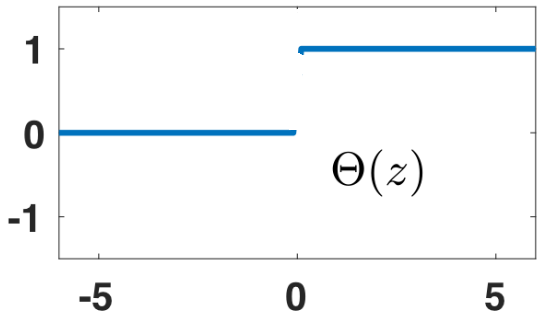
de disparo (resta el segundo del primero). Si es mayor o igual a cero, dispara (toma el valor +1).

Si es menor que cero, se queda en reposo (toma el,valor 0)

A lo largo de la historia de la Inteligencia Artificial, poco ha cambiado el modelado de neuronas desde 1943. Las técnicas más modernas preservan la idea de una suma ponderada y la importancia de considerar los umbrales. Lo que usualmente cambia es la llamada función de activación, que para McCulloch y Pitts era la función Heaviside.

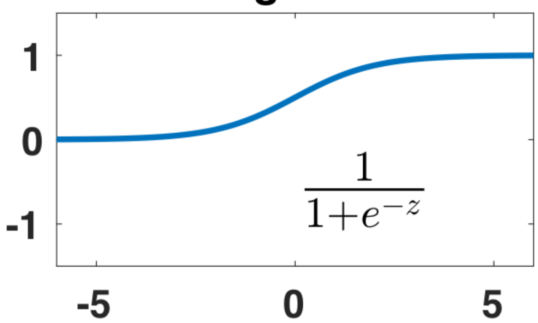
Hoy admitimos otras funciones que preservan la principal propiedad de la Heaviside, la no linealidad.

Perceptron



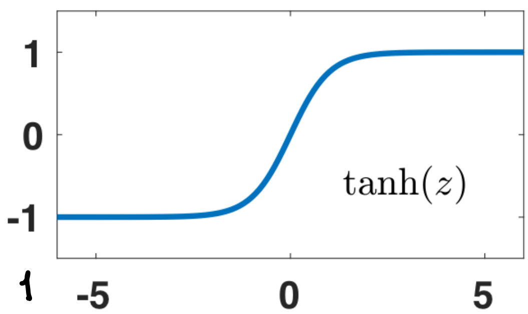
$$y = 0 \text{ o } 1$$

Sigmoid



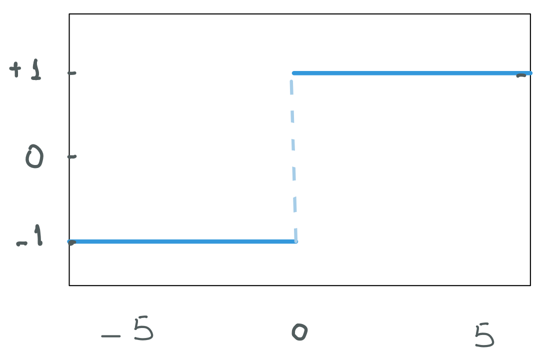
$$0 < y < 1$$

Tanh

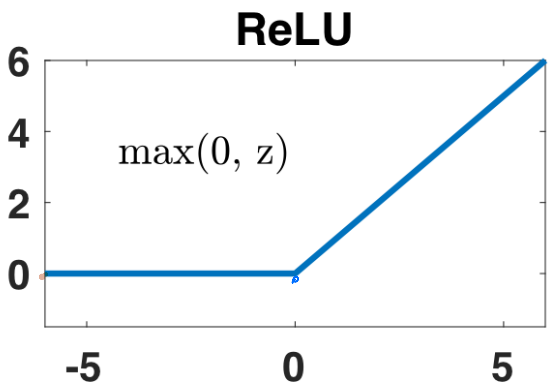


$$-1 < y < 1$$

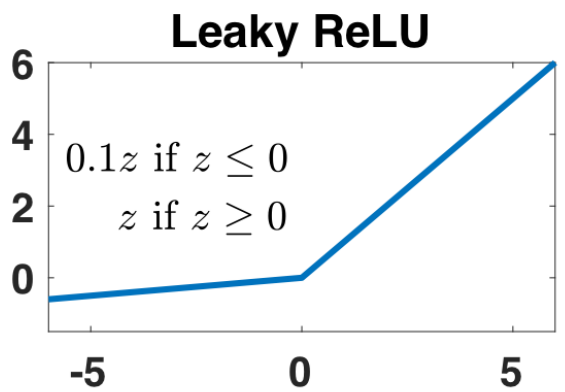
Signo



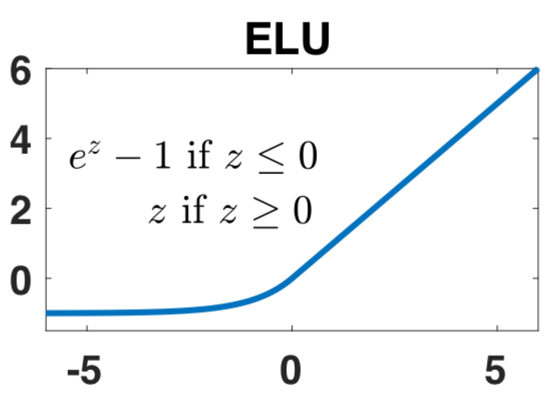
$$y = -1 \text{ o } +1$$



$$0 \leq y < \infty$$



$$-\infty \leq y < \infty$$



$$-1 \leq y < \infty$$

Una RED NEURONAL es un ensamblado de muchas neuronas artificiales.

$$\vec{y}(t) = (y_1(t), \ldots, y_N(t))$$

La red está definida no solo por la ***N*** neuronas sino también por

- Las ***NxN*** eficacias sinápticas, determinadas por medio de una matriz cuadrada.
- Los ***N*** umbrales pensados como un vector.

APRENDER**, en términos de Inteligencia Artificial, es asignar los valores de las *NxN* eficacias sinápticas y los *N* umbrales de forma tal que la red asigne, dinámicamente, una salida correcta ante una entrada dada.**

Cuando nos referimos a APRENDER queremos decir que, partiendo de valores aleatorios de las sinápsis y los umbrales vamos variándolo lentamente hasta alcanzar valores que le permiten a la red hacer las conexiones entrada salida correctas.

$$W_{ik}^{n} = W_{ik}^{v} + \Delta W_{ij}$$

$$\mu_i^n = \mu_i^v + \Delta \mu_i$$

La actualización puede ser

- En línea: se aplica back propagation después de cada ejemplo.
- En batch: se aplica back propagation después de mostrar todo el conjunto de entrenamiento.