# REDES NEURONALES

# 2021

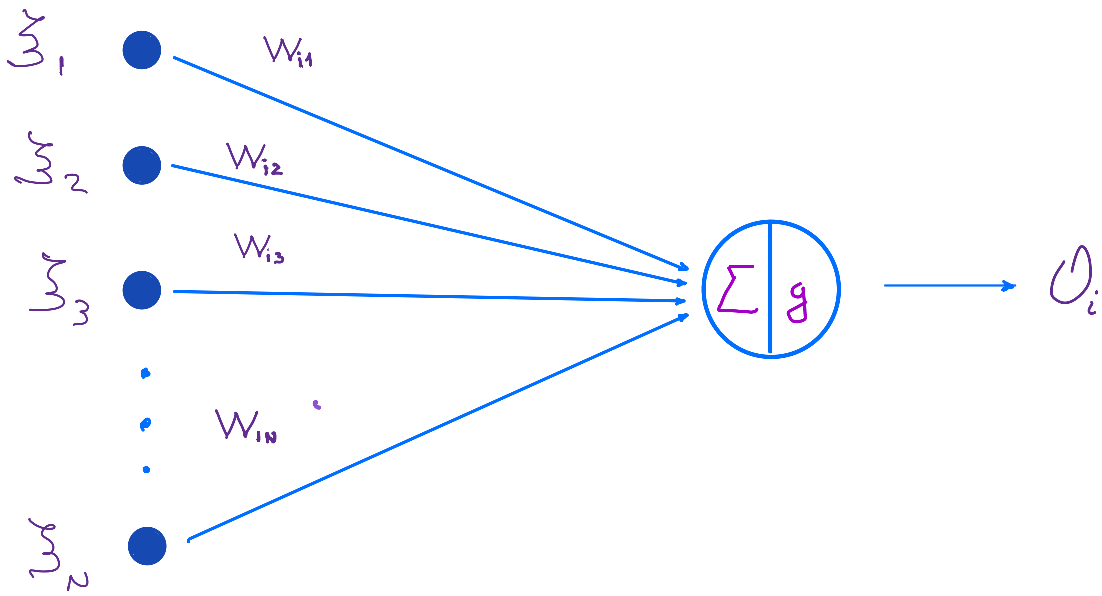
## Clase 16 Parte 1

**Facultad de Matemática, Astronomía, Física y Computación**
**Universidad Nacional de Córdoba**

**Jueves 7 de octubre 2021**

http://www.famaf.unc.edu.ar/~ftamarit/redes2021

https://www.famaf.unc.edu.ar/course/view.php?id=798




Resumiendo la notación:

$\bar{\xi}$ : Vector de entrada (de *N* componentes)

$\bar{\xi}^{\alpha}$ : *p* vectores de entrada del conjunto de entrenamiento (p x N)

$\zeta_i^{\alpha}$ : Salida correcta asociada a $\bar{\xi}^{\alpha}$ (*M* componentes)

$O_i^{\alpha}$ : Salida real asociada a $\bar{\xi}^{\alpha}$ (*M* componentes)

$\mu_i$ : Umbrales de activación de capa de salida (*M* componentes)

$\bar{W}_i$ : Conjunto de acoplamientos sinápticos (*N x M*)

Para el perceptrón simple:

$$\bar{\xi},\ \bar{\xi}^{\alpha},\ \bar{w}_i \in \mathbb{R}^N$$

$$\xi_i^{\alpha},\ O_i^{\alpha},\ \mu_i \in \mathbb{R}$$

## Como deshacernos del molesto umbral

El hecho de que los pesos y el umbral tengan un rol tan diferente en la dinámica de la neurona complica la generación de algoritmos. Pero es una diferencia ficticia. Hay una manera muy simple de equiparar al umbral a una eficacia sináptica. Para ellos imaginamos que tenemos una entrada extra a la que le asignamos el índice k=0 y está fija en el valor -1.

$$h_i - \mu_i = \sum_{k=1}^{N} w_{ik}\xi_k - \mu_i$$

$$= \sum_{k=1}^{N} w_{ik}\xi_k + \mu_i\xi_0$$

$$h_i - \mu_i = \sum_{k=0}^{N} w_{ik}\xi_k$$

si

$$w_{i0} = \mu_i \qquad \xi_0 = -1$$

## Con esto nos olvidamos de darle un tratamiento diferenciado al umbral

A partir de ahora entonces, para nosotros aprender significa encontrar un vector de *N+1* dimensiones que resuelva los datos provistos en el conjunto de entrenamiento.

$$\overline{W} = (w_{i0}, w_{i1}, \ldots, w_{iN}) \in \mathbb{R}^{N+1}$$

o sea, que para cualquier entrada del conjunto de entrenamiento nos de

$$O_i^\mu = g(h_i^\mu) = g\left(\sum_{k=0}^{N} w_{ik} \xi_k^\mu\right) = \zeta_i^\mu$$

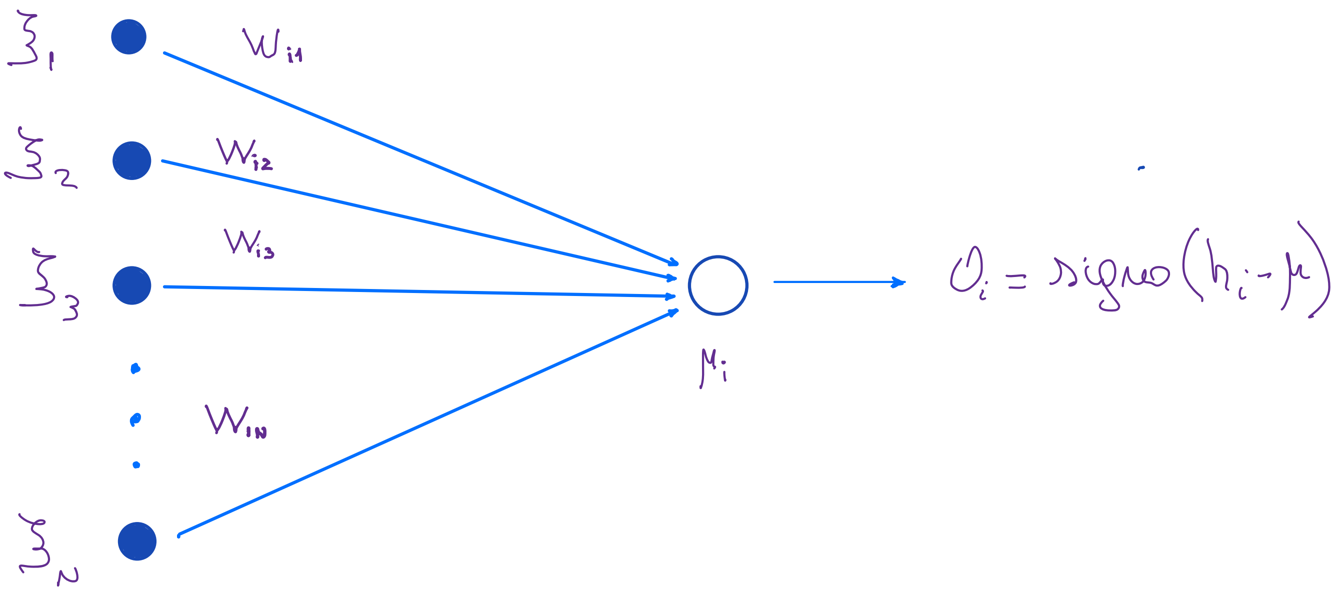
$$O_i^\mu \cong \zeta_i^\mu$$

No siempre podremos aprender perfectamente el conjunto de entrenamiento. Es más, casi nada nunca. Nos conformaremos fijando una tolerancia en el error que comete la red con el conjunto de entrenamiento y dejaremos de buscar mejores soluciones cuando se cumpla que

**Error de entrenamiento** $\leq$ **Tolerancia al error**

Más adelante tendremos que ver qué le exigimos a la red a la hora de generalizar, o sea, de resolver lo que nunca vió, pero lo dejamos para las próximas clases.

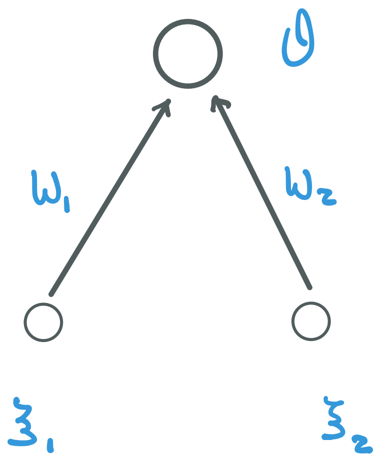
# PERCEPTRON SIMPLE COMO UN CLASIFICADOR




Vamos a analizar la red neuronal más simple que podemos construir para clasificar en dos categorías los valores de entrada. Supongamos que tenemos

$$\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_N)$$

y la salida es binaria, pudiendo tomar valores -1 y +1. Para simplificarlo aún más, supongamos que solo dos entradas  y  .




**PERCEPTRON CON *N=2***

Quitamos aquí el índice i que identifica a la neurona para no complicarnos con los índices en el cálculo que haremos, pero tengamos en mente que estas neuronas simples se pueden ensamblar y que cuando eso ocurre necesitamos preservar el índice.

Suponemos también que las variables de entrada pueden tomar cualquier valor real. Tanto la entrada como las sinapsis vive en el plano. No incluiremos por ahora el umbral.

$$\vec{\xi} \in \mathbb{R}^2 \qquad \vec{\xi} = (\xi_1, \xi_2)$$

$$\vec{w} \in \mathbb{R}^2 \qquad \vec{w} = (w_1, w_2)$$

$$\mu = 0$$

$$O = \text{signo}(h)$$

$$= \text{signo}\left(\sum_{k=1}^{2} w_k \xi_k\right)$$

$$= \text{signo}(\bar{w} \cdot \bar{\xi}) = \begin{cases} +1 \\ -1 \end{cases}$$

$$O \in \{-1, +1\}$$

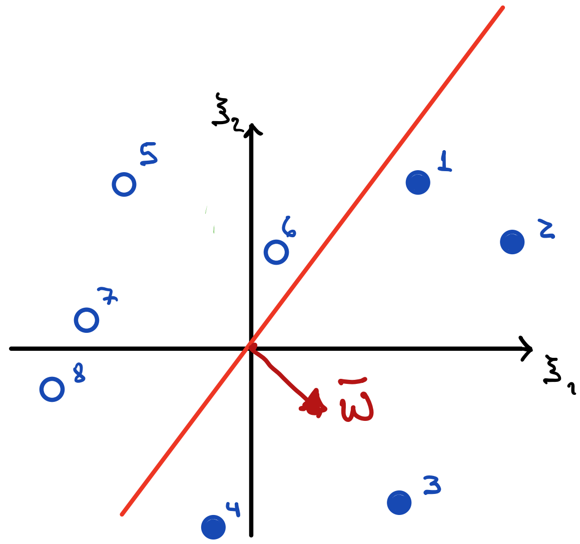
Tenemos relaciones **INPUT-OUTPUT** conocidas, las que serán utilizadas para entrenar. Supongamos que tenemos un conjunto pequeño, de solo **8** elementos en el conjunto de entrenamiento (poquísimas para un caso real).

$$\bar{\xi}^{\mu} \longrightarrow \zeta^{\mu} = \pm 1 \quad \text{deseado o correcto}$$

$$\mu = 1, 2, 3, \ldots \text{ y } 8$$

$$\zeta^{\mu} = \begin{cases} +1 & \text{si } \bar{\xi}^{\mu} \text{ pertenece a la categoría A} \\ -1 & \text{si } \bar{\xi}^{\mu} \text{ pertenece a la categoría B} \end{cases}$$

Voy a representar los puntos en $\mathbb{R}^2$




● A $+1$

○ B $-1$

Cada uno de los 8 puntos representa a uno de los 8 elementos del conjunto de entrenamiento. En este caso en particular 4 elementos pertenecen a la categoría A ( $\mu$ = 1, 2, 3 y 4) y los representamos con círculos azules llenos, en tanto los otros 4 elementos pertenecen a la categoría B ( $\mu$ = 5, 6, 7 y 8) y los representamos con círculos azules vacíos.

Hagamos algunas cuentas simple:

$$O^\mu = \zeta^\mu \qquad \text{(lo deseado)}$$

$$\text{signo}\left(\sum_{k=1}^{2} w_k \xi_k^\mu\right) = \zeta^\mu$$

$$\text{signo}\left(\vec{w} \cdot \vec{\xi}^\mu\right) = \zeta^\mu$$

$$\zeta^{\mu}\,\text{signo}\left(\bar{w}\cdot\bar{\xi}^{\mu}\right) = \zeta^{\mu}\zeta^{\mu} = +1 \quad \text{pues} \quad \zeta^{\mu} = \pm 1$$

$$\text{signo}\left(\zeta^{\mu}\right)\text{signo}\left(\bar{w}\cdot\bar{\xi}^{\mu}\right) = +1$$

$$\text{signo}\left(\left(\bar{w}\cdot\bar{\xi}^{\mu}\right)\zeta^{\mu}\right) = 1$$

$$\left(\bar{w}\cdot\bar{\xi}^{\mu}\right)\zeta^{\mu}$$

$$\bar{w}\cdot\left(\bar{\xi}^{\mu}\zeta^{\mu}\right)$$

$$\text{signo}\left(\bar{w}\cdot\left(\bar{\xi}^{\mu}\zeta^{\mu}\right)\right) = 1$$

$$\text{signo}\left(\bar{w}\cdot\bar{X}^{\mu}\right) = 1$$

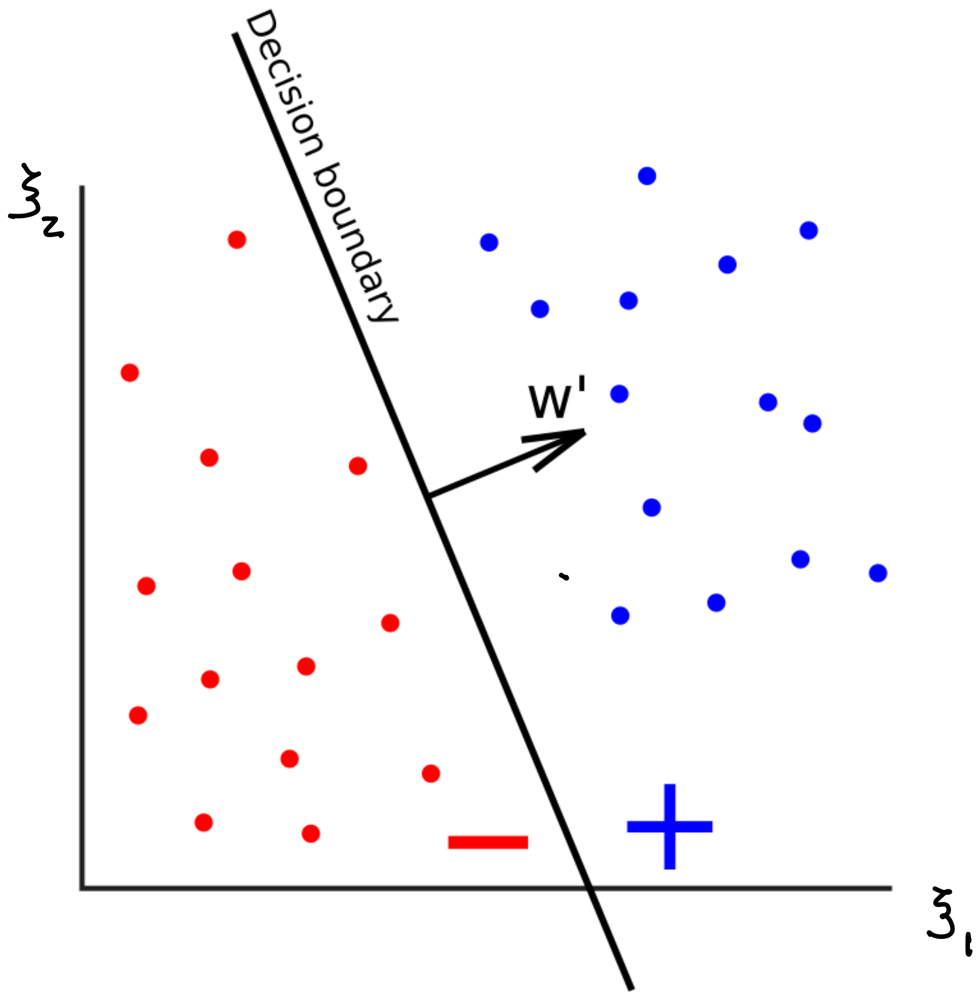
$$\boxed{\bar{w}\cdot\bar{X}^{\mu} > 0} \qquad \forall \mu$$

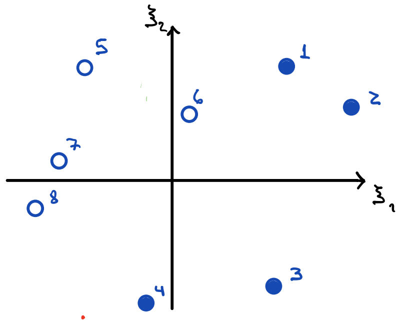
donde $\bar{X}^{\mu} = \bar{\xi}^{\mu}\zeta^{\mu}$

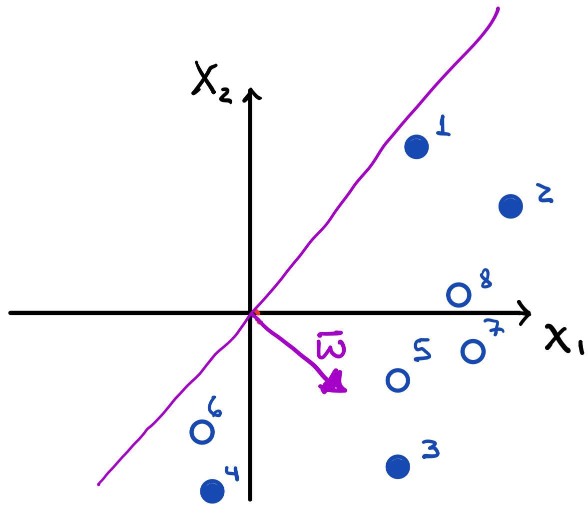
Notemos que:

si $\zeta^{\mu} = +1$ entonces $\bar{X}^{\mu} = \bar{\xi}^{\mu}$

si $\zeta^{\mu} = -1$ entonces $\bar{X}^{\mu} = -\xi$

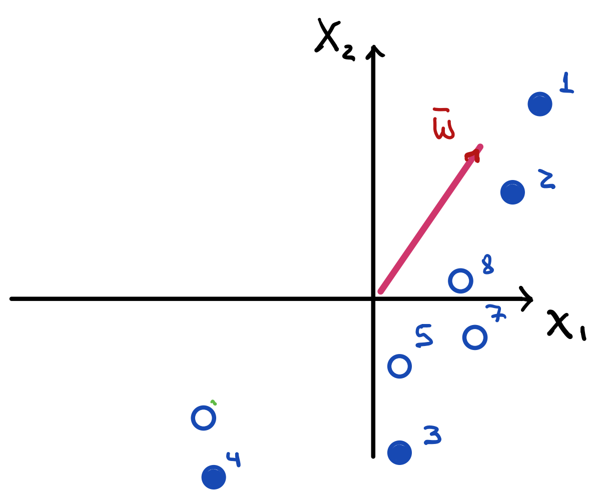

Decision boundary
w'
−
+
ξ2
ξ1





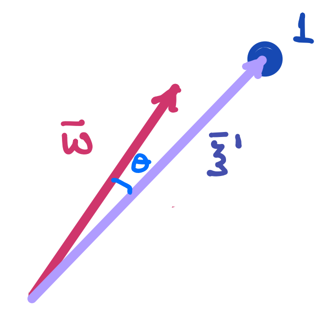

Sabemos que $\bar{w} \cdot \bar{x}^{\mu} > 0$ Implica que la proyección de $\overline{W}$ sobre cada uno de los elementos del conjunto de entrenamiento. O sea, de TODOS LOS EJEMPLOS.

Como vimos, el $\overline{W}$ que buscamos vive en el plano



¿Resuelve este $\bar{W}$ los elementos del conjunto de entrenamiento?

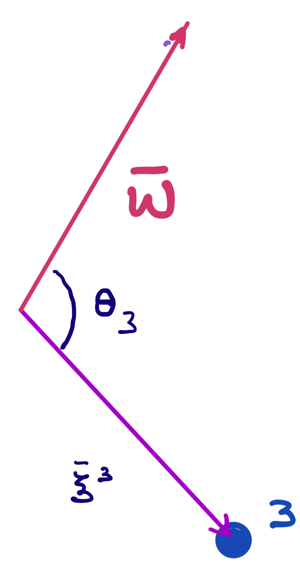
Para el elemento **1** funciona bien:




$$\bar{w} \cdot \bar{\xi}^1 = |\bar{w}| \cdot |\bar{\xi}^1| \cdot \cos(\theta) > 0$$

pues $0 < \theta < \frac{\pi}{2}$ (agudo).

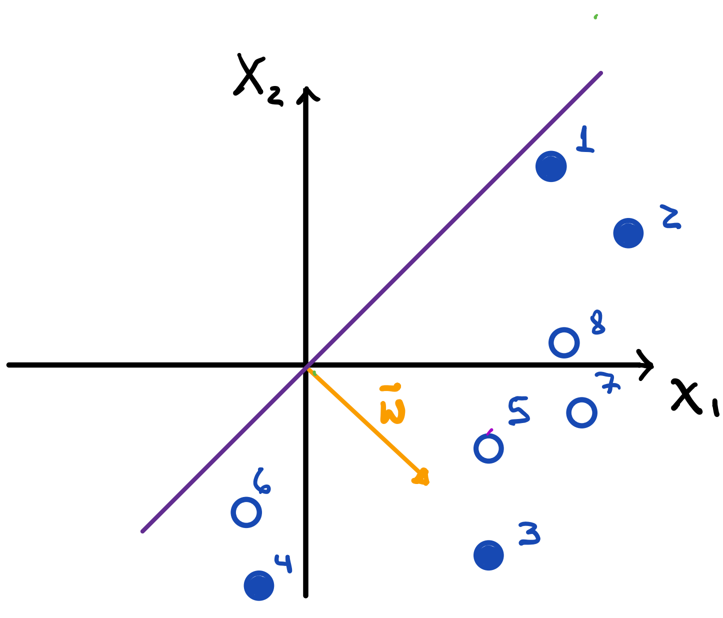
Eso miremos qué pasa con el elemento **3**:



$$\bar{w} \cdot \bar{\xi}^3 = |\bar{w}| \cdot |\bar{\xi}^3| \cos(\theta_3) < 0$$

pues $\frac{\pi}{2} < \theta_3 < \pi$ (obtuso)

O sea, debemos encontrar un $\vec{W}$ que forme un ángulo agudo con los ocho elementos del conjunto de entrenamiento en su representación $\bar{X}$. En este caso, como en muchos casos, existen infinitos vectores W que resuelven el problema.




Este vector $\vec{W}$ resuelve bien todos los ejemplos. O sea, con este conjunto de sinapsis aprendemos a resolver el problema de aprendizaje. Resta saber cuán bien resuelve los casos que no ha visto.