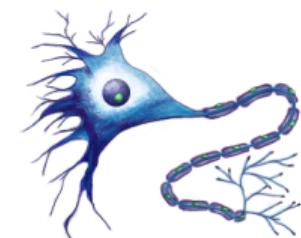


Neurociencias computacionales

1 Neurociencias computacionales

2 Sistema nervioso

3 Cerebro



Neurociencias computacionales

Las *neurociencias computacionales* se interesan en:

- Descripciones y modelos funcionales, biológicamente **realistas**, de neuronas y sistemas neuronales.
- La fisiología y dinámica de dichos elementos.

En ellas se combinan las siguientes disciplinas:

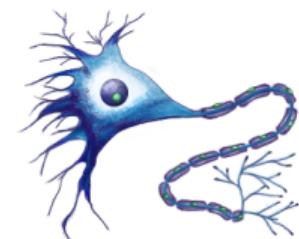
- Biofísica
- Neurociencias
- Matemáticas
- **Ciencias de la computación**
- Ingeniería eléctrica
- Ciencias cognitivas

Tipos de modelos del sistema nervioso

- Existen tres tipos de modelos del sistema nervioso:
 - ① Descriptivos (¿qué hace?)
 - ② Mecanistas (¿cómo lo hace?)
 - ③ Interpretativos (¿por qué o para qué lo hace?)
- Objetivos del modelado: Se modela:
 - Corrientes en membranas
 - proteínas y acoplamiento químico
 - oscilaciones de redes
 - arquitectura topográfica y de columnas
 - aprendizaje
 - memoria

Sistema nervioso

- 1 Neurociencias computacionales
 - 2 Sistema nervioso
 - 3 Cerebro

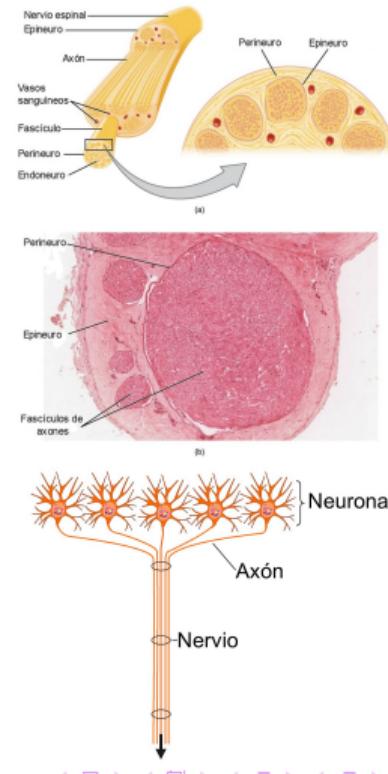


Nervios

- Los nervios son estructuras conductoras de impulsos nerviosos situadas fuera del sistema nervioso central.
- Están formados por un conjunto de axones agrupados, cada uno de los cuales procede de una neurona.
- Pueden ser motores o sensitivos, pero la mayor parte son mixtos y contienen tanto fibras sensitivas como motoras.
- Se originan en la médula espinal (nervios raquídeos) o parten directamente del encéfalo (nervios craneales).

*Principios de Anatomía y Fisiología. Autor:
Tortora-Derrickson.*

Texto y figuras: <https://es.wikipedia.org/wiki/Nervio>



Sistema nervioso periférico

Somático: Nervios conectados a *músculos voluntarios esqueléticos* y *receptores sensoriales*

- ① Aferentes (de entrada)
- ② Eferentes (de salida)

Autónomo: Nervios que se conectan con el corazón, ganglios, etc.

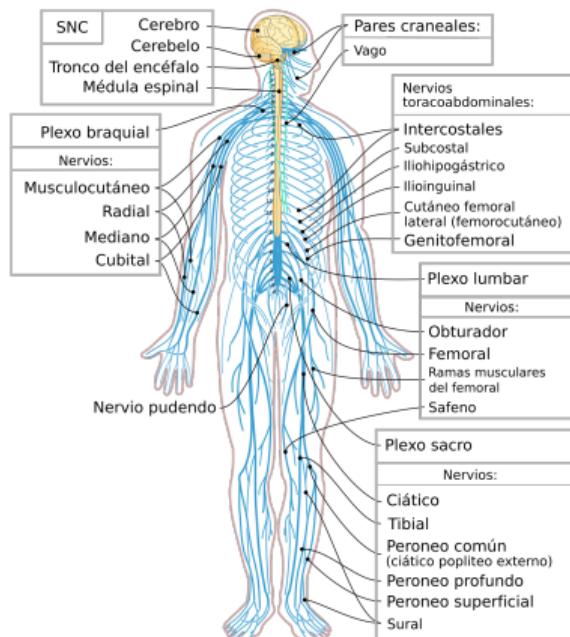


Figura: https://commons.wikimedia.org/wiki/File:Nervous_system_diagram_numbered.svg

Sistema nervioso central

Médula espinal + encéfalo

Médula espinal:

- Ciclos de retroalimentación local (reflejos)
- Señales de control motor descendientes (cerebro → neuronas motoras)
- Axones sensoriales ascendentes (músculos, piel → cerebro)

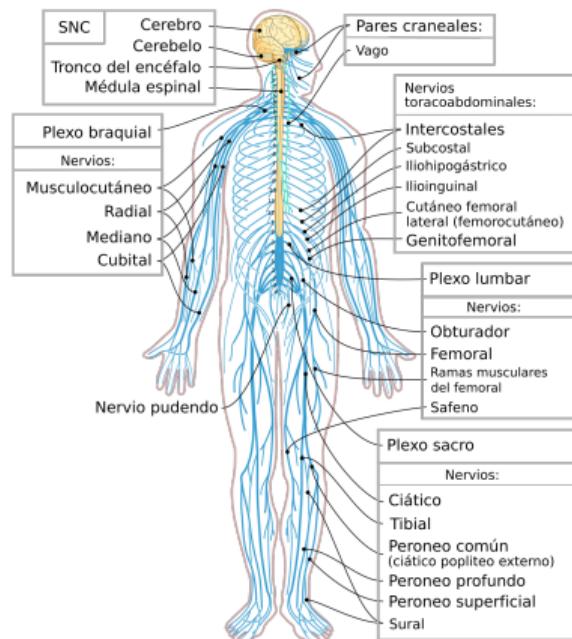


Figura: https://commons.wikimedia.org/wiki/File:Nervous_system_diagram_numbered.svg

Nervios saliendo de la base del cerebro

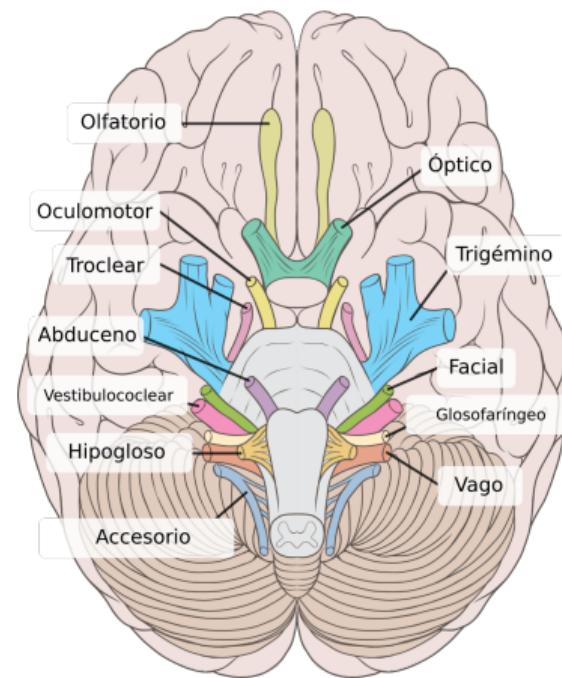
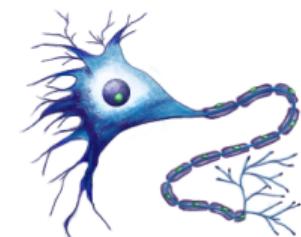


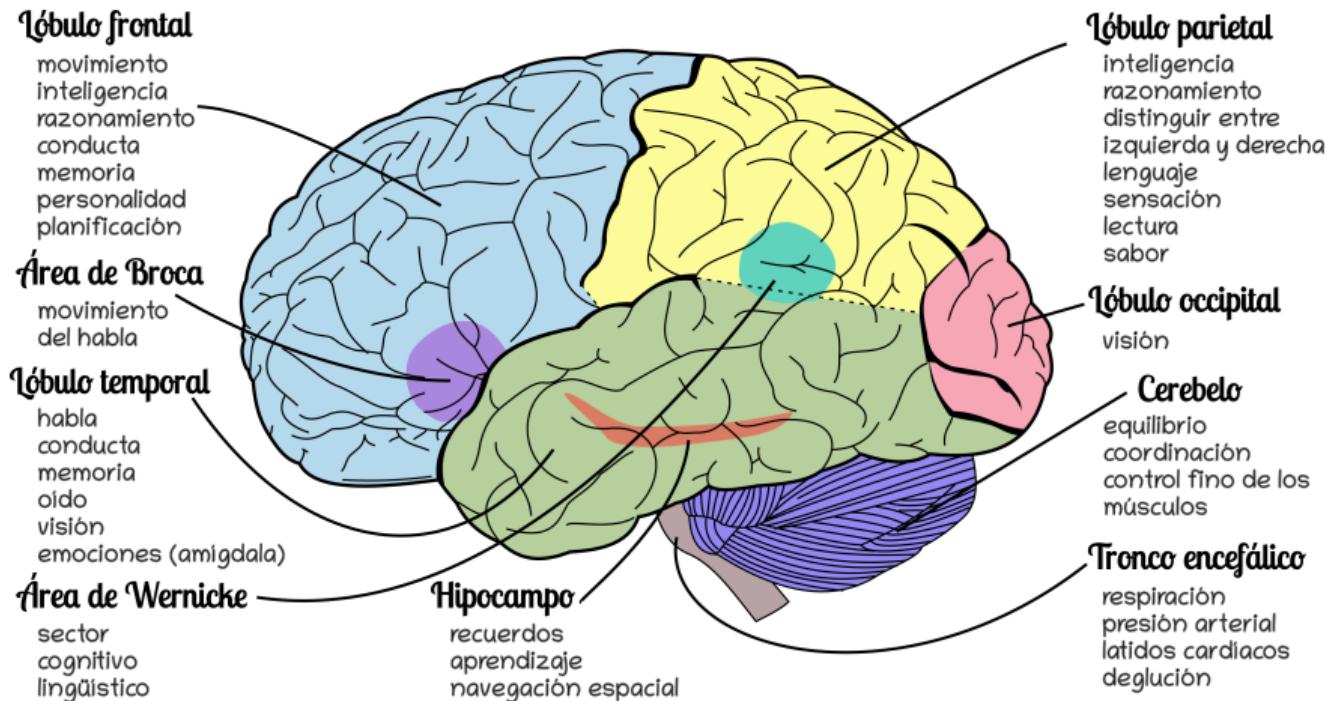
Figura: https://commons.wikimedia.org/wiki/File:Brain_human_normal_inferior_view.svg

Cerebro

- 1 Neurociencias computacionales
- 2 Sistema nervioso
- 3 Cerebro



Regiones del cerebro



Adaptado de: https://commons.wikimedia.org/wiki/File:Brain_diagram_pl.svg

Zonas funcionales del cerebro

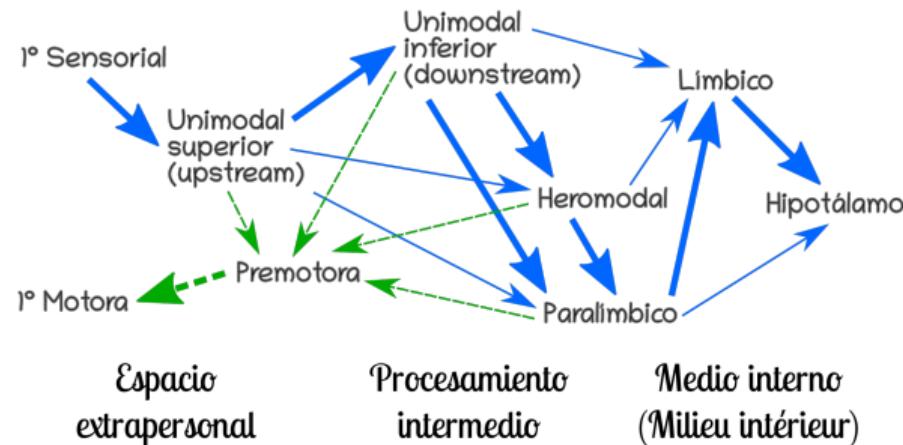


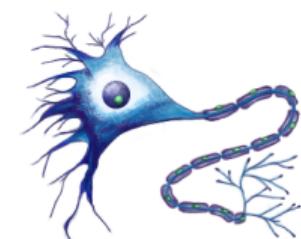
Figura: Las líneas rectas ilustran conexiones monosinápticas en las modalidades visual y auditiva; el grosor representa la cantidad de conexiones. Líneas punteadas ilustran rutas de información motora. Adaptado de Mesulam 1998.

Una versión más sencilla de este tema se puede encontrar en:

<https://psicologiamiente.com/neurociencias/corteza-asociativa>

Introducción

- 1 Introducción
 - 2 Sinapsis química
 - 3 Señal eléctrica
 - 4 Neuroplasticidad



Neuronas de verdad (foto rusa)

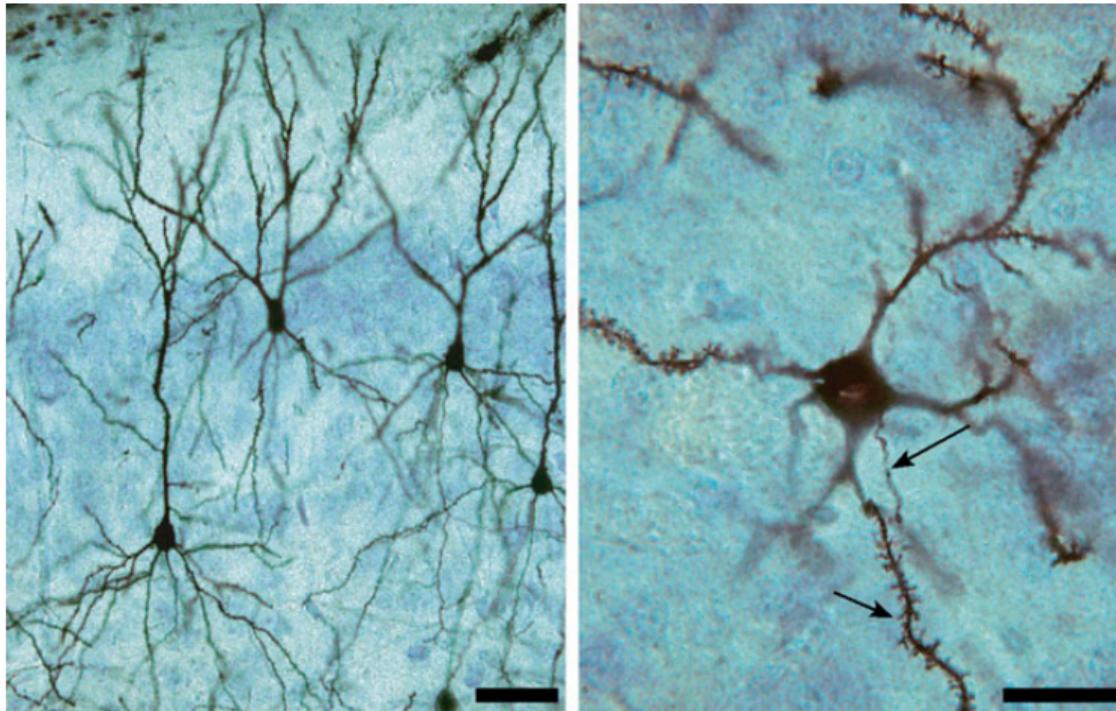


Figura: https://commons.wikimedia.org/wiki/File:Neuron_%D0%BA%D0%BB_%D0%91%D0%B5%D1%86%D0%BD%D0%BE%D0%BC.jpg

Neuronas

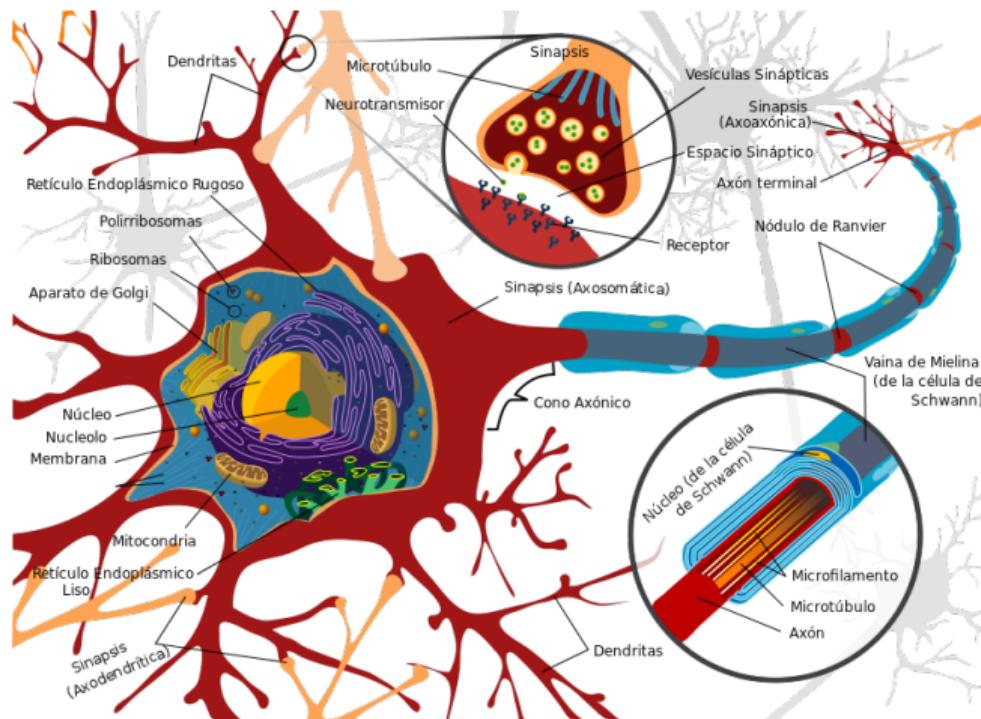


Figura: https://es.wikipedia.org/wiki/Archivo:Complete_neuron_cell_diagram_es.svg

Transmisión de señales y almacenamiento de información

- **Neurotransmisores:** químicos liberados entre neuronas.
 - **Impulsos eléctricos:** potenciales de acción a lo largo del axón.
 - **Plasticidad:** modificación a largo plazo de las conexiones entre neuronas.



Tipos de neuronas

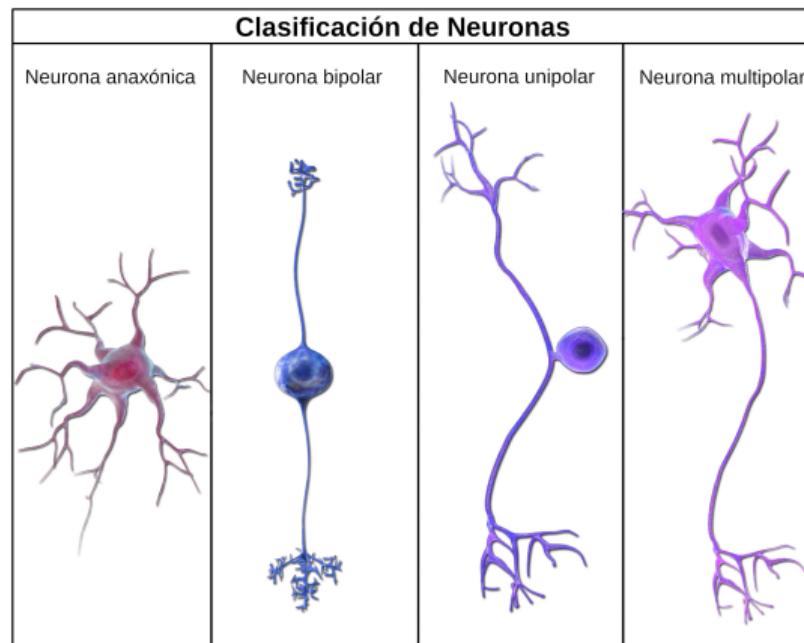
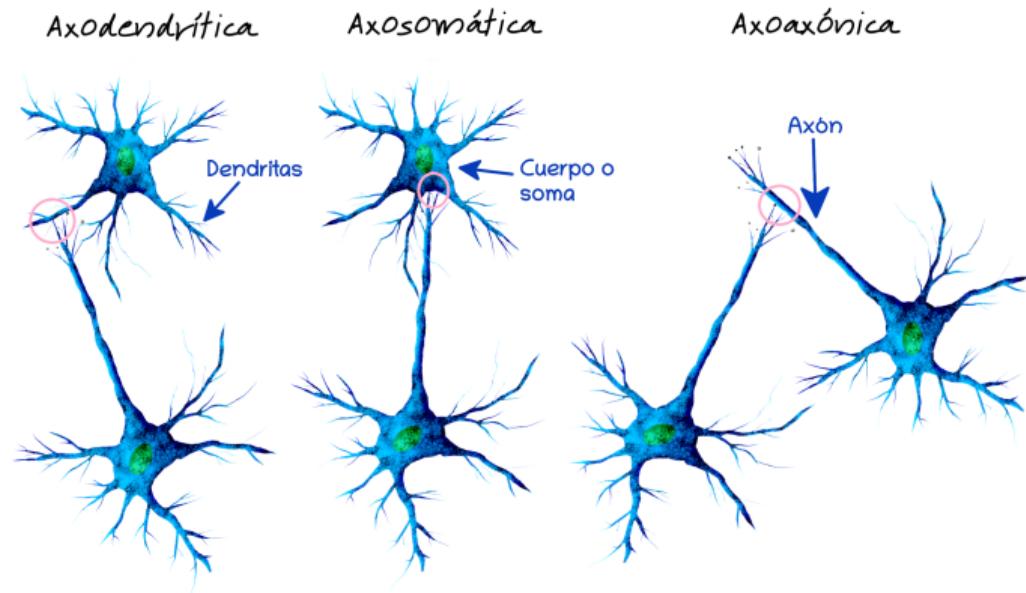


Figura: https://commons.wikimedia.org/wiki/File:Neuron_Classification.png

Sinapsis

- Un axón puede hacer contacto con:
 - Las dendritas de otra neurona,
 - su cuerpo (soma)
 - su axón.

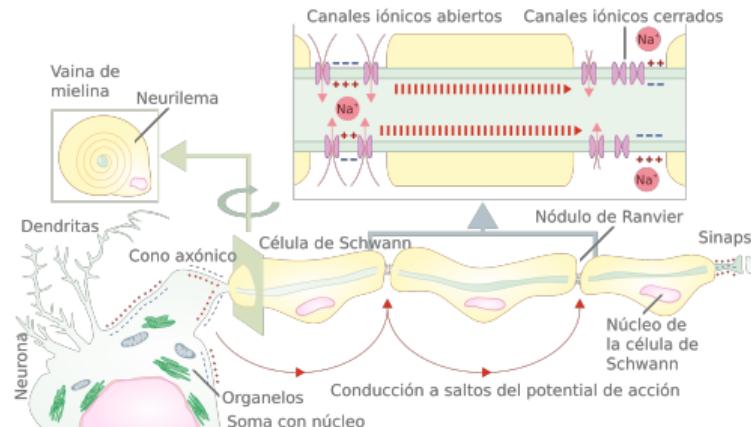


- El cerebro humano tiene alrededor de 86 mil millones de neuronas. [Universitam](#)
- Se piensa que neuronas en el cerebelo pueden llegar tener hasta 200,000 conexiones con otras neuronas. Kerr 2016

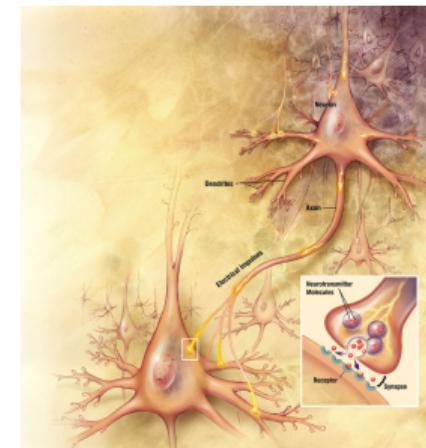
Sinapsis

Funcionalmente hay dos mecanismos de transmisión sináptica:

- 1 Eléctrica, a través de la *brecha sináptica*.
- 2 Química, mediante *neurotransmisores*.



(a) Corriente iónica por el axón (efecto de corto plazo)



(b) Cambio en la permeabilidad de la membrana

Figuras: https://commons.wikimedia.org/wiki/File:Propagation_of_action_potential_along_myelinated_nerve_fiber_en.svg,

https://commons.wikimedia.org/wiki/File:Neurons_big1.jpg

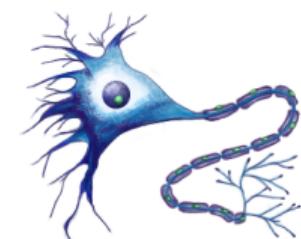
Sinapsis química

1 Introducción

2 Sinapsis química

3 Señal eléctrica

4 Neuroplasticidad



Sinapsis química

Hay dos tipo de sinapsis químicas:

- ① Excitatoria. Despolariza la membrana postsináptica.
- ② Inhibitoria. Hiperpolariza la membrana postsináptica.

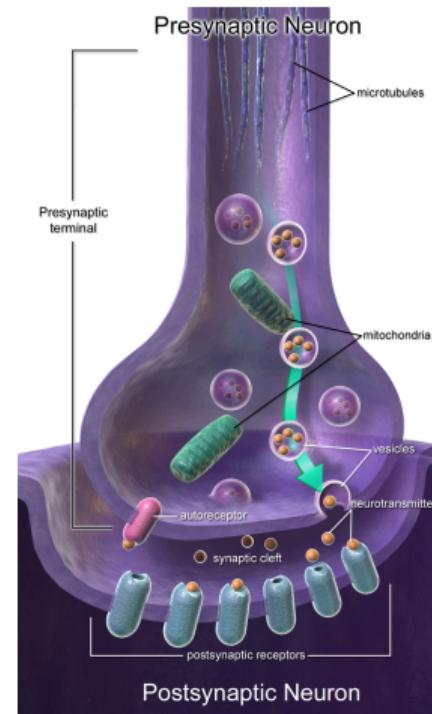
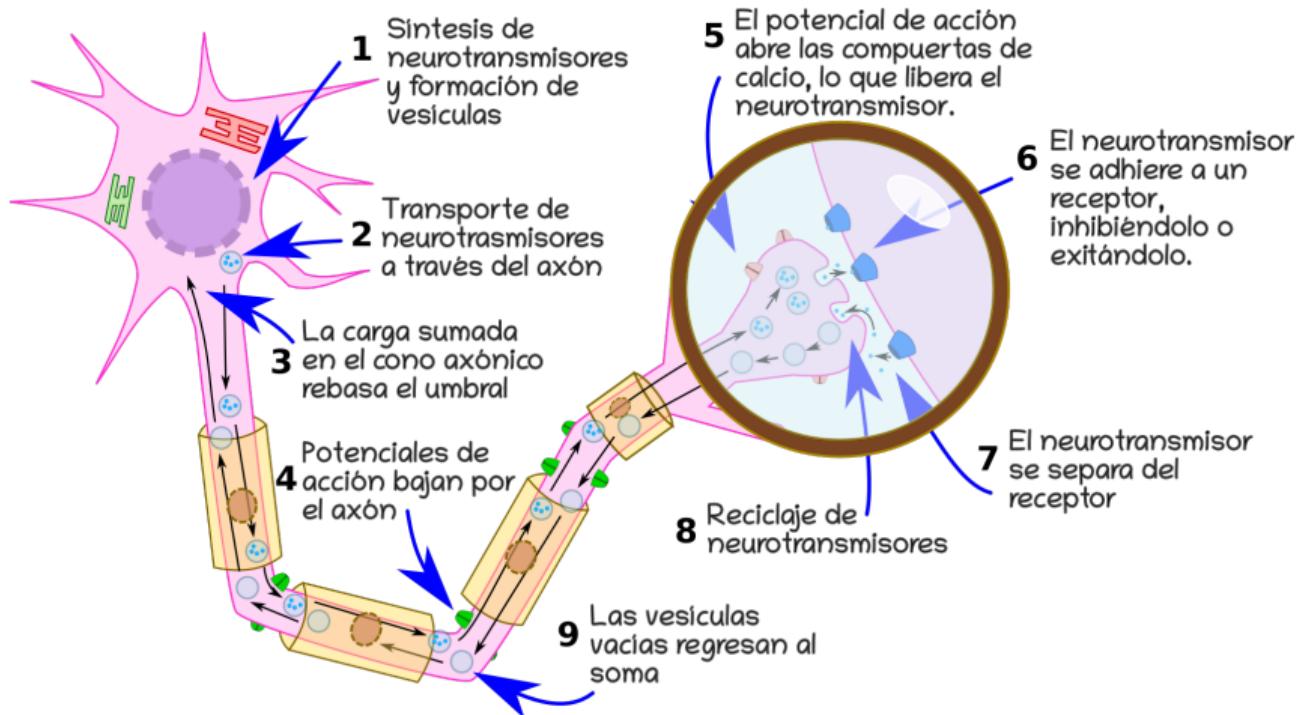


Figura: https://commons.wikimedia.org/wiki/File:Synapse_Presynaptic_Neuron.png

Neurotransmisión



Ejemplos de neurotransmisores

Cada neurotransmisor tiene diversos efectos, pero se mencionan los más icónicos de cada uno.

Serotonina Hormona de la felicidad.

Dopamina Adicción, placer y aprendizaje.

Oxitocina Parto, lactancia, aprendizaje, empatía, fidelidad, regulación del estrés^[1].

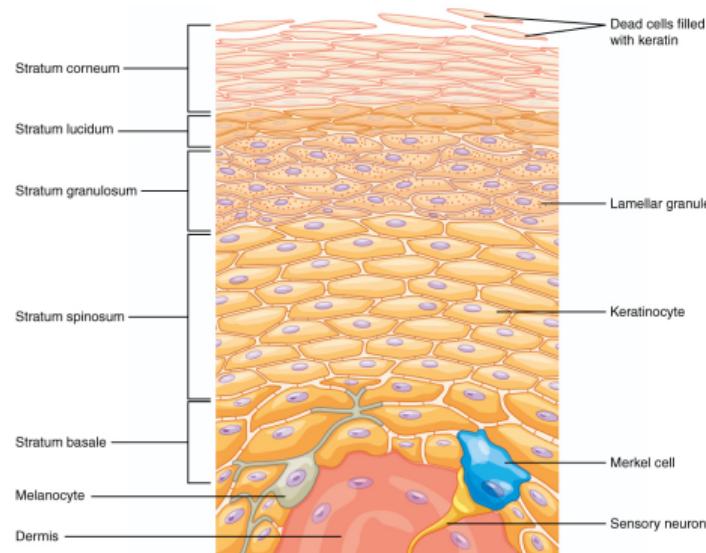
Endorfinas Euforia.

Adrenalina Supervivencia, mecanismos de alerta.

[1] <https://www.abc.es/familia-padres-hijos/20150418/abci-oxitocina-parto-feliz-201504161753.html>

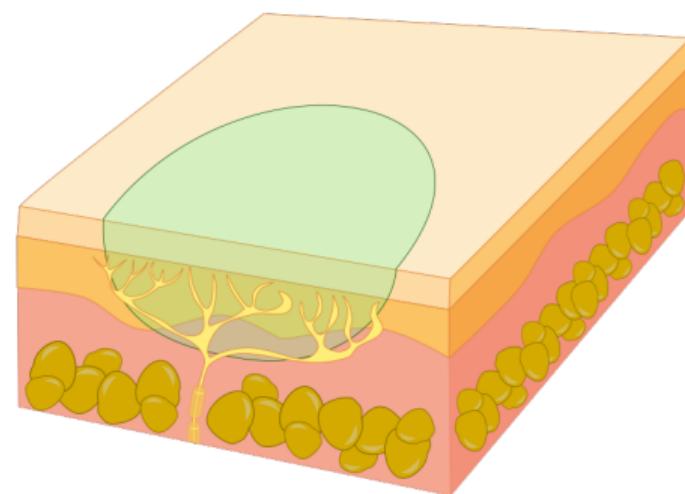
Campos receptores

Propiedades específicas de un estímulo sensorial que generan una respuesta fuerte de la célula.



(a) Capas de la epidermis

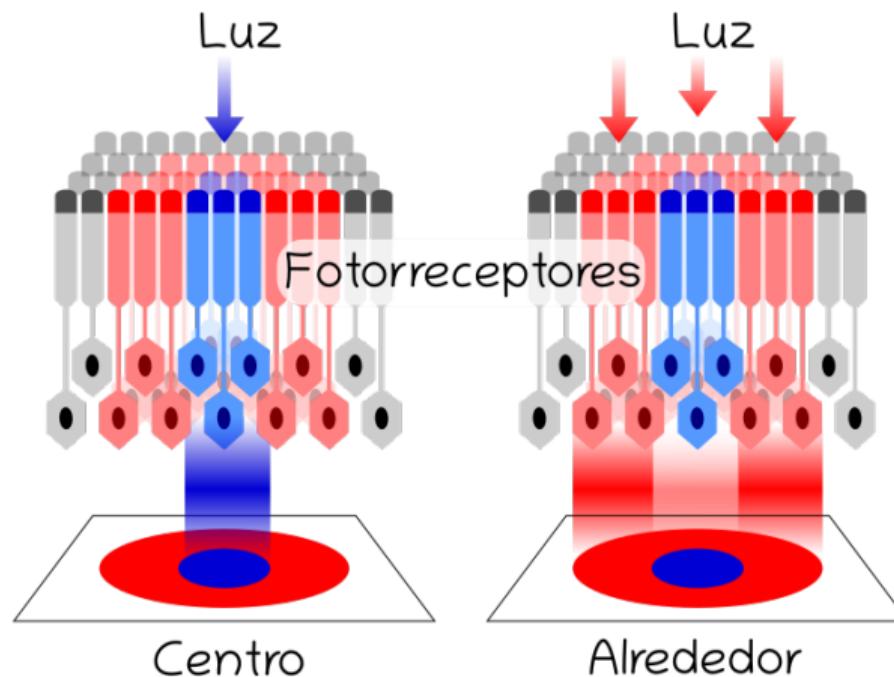
Figura izquierda: https://commons.wikimedia.org/wiki/File:502_Layers_of_epidermis.jpg



(b) Campos receptores

Neuronas sensitivas

Campos receptores en el ojo.



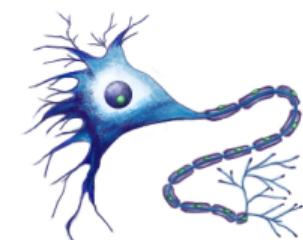
Señal eléctrica

1 Introducción

2 Sinapsis química

3 Señal eléctrica

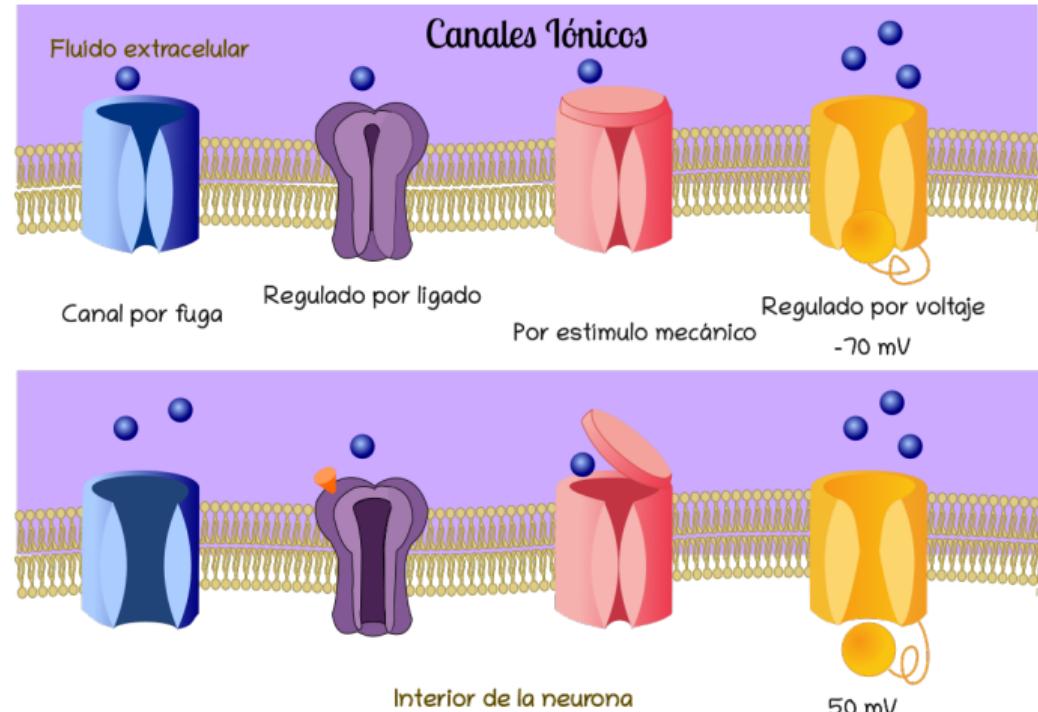
4 Neuroplasticidad



Diferentes tipos de compuertas iónicas

Ejemplo de iones

- K^+
- Na^+
- Cl^-



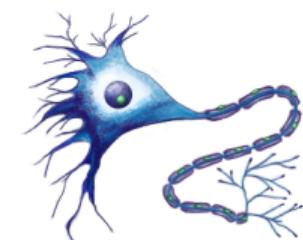
Neuroplasticidad

1 Introducción

2 Sinapsis química

3 Señal eléctrica

4 Neuroplasticidad



Neuroplasticidad

- La *neuroplasticidad* es un mecanismo de aprendizaje del cerebro en el cual:
 - cuando las neuronas se activan simultáneamente con frecuencia la conexión entre ellas se fortalece,

Este mecanismo constituye la principal inspiración para el diseño de las redes neuronales artificiales.

Unión sináptica eléctrica

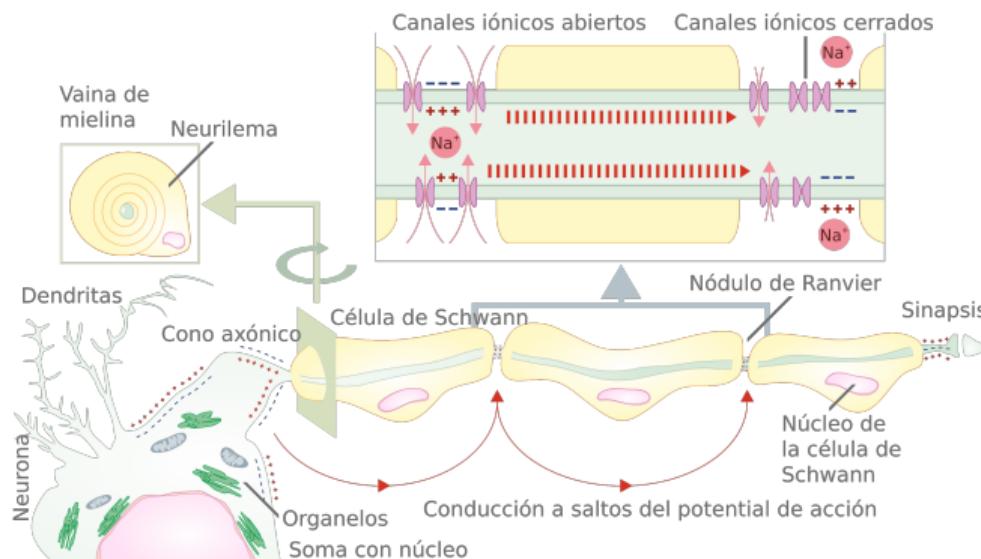


Figura: https://commons.wikimedia.org/wiki/File:Propagation_of_action_potential_along_myelinated_nerve_fiber_en.svg

Figura: Corriente iónica (efecto sincronizado de corto plazo)

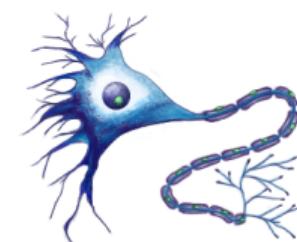
Modelo de Hodgkin Huxley

1 Modelo de Hodgkin Huxley

2 Dinámica del voltaje

3 Simulación

4 Información codificada en las dendritas



Axón gigante de calamar

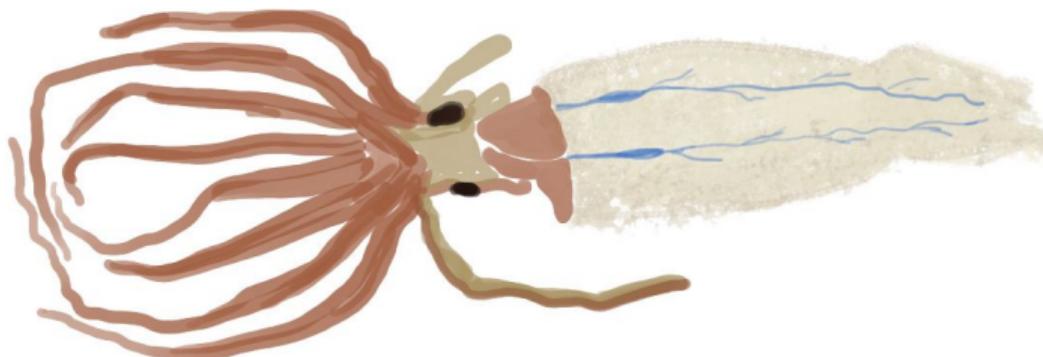


Figura: El calamar puede medir hasta 4m de largo. Este axón tiene aproximadamente 0.5mm de diámetro.

Modelo de Hodgkin Huxley

El modelo fue propuesto originalmente en 1952, aunque tomó varios años comprender los mecanismos detrás del comportamiento descrito.

- Estudiaron las conductancias de Na^+ y K^+ .

Modelo de Hodgkin Huxley

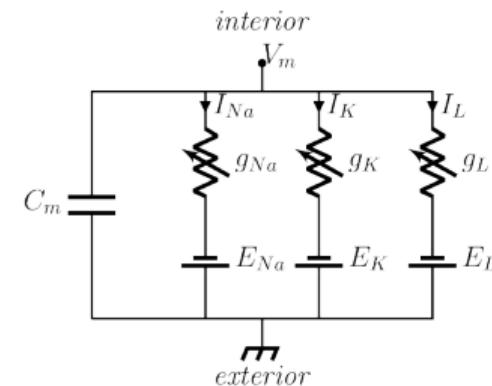
Se presentan dos Canales Controlados por Voltaje (*Voltage Gated Channel (VGC)*)
Wells s.f.:

Na^+ Un canal transitorio (con compuerta para inactivación).

K^+ Un canal persistente.

Adicionalmente un canales para intercambio aleatorio de otros iones (ej: Cl^-).

L Un canal aleatorio (*leaky*).



Repasso de conceptos de electricidad

Potenciales eléctricos E ó V. resultan de la separación de cargas opuestas. Se mide en mV.

Corriente I. Movimiento de cargas. Se mide en μA .

Resistencia R. Medida de la oposición al movimiento de las partículas cargadas.

Conductancia g. Inverso de la resistencia, es decir, facilidad de transmisión de las partículas cargadas.

Capacitancia o capacidad eléctrica C. Cantidad de energía eléctrica almacenada en un capacitor para una diferencia de potencial eléctrico dada.

Potenciales de Nerst o de reposo

Son los potenciales a los cuales el flujo neto de iones a través de los canales abiertos es cero: E_{Na} , E_K y, agrupando otros iones como el Cl^- , E_L .

$$\begin{array}{ll} E_{Na} & 50\text{mV} \\ E_{Ca} & 150\text{mV} \\ E_K & -80\text{mV} \\ E_{Cl} & -60\text{mV} \end{array}$$

La membrana de una neurona es modelada como un elemento de un circuito con **capacitancia** C_m y **potencial** V , los cuales están regidos por las ecuaciones siguientes:

$$C_m \frac{dV}{dt} = -g_{Na}m^3h(V - E_{Na}) - g_Kn^4(V - E_K) - g_L(V - E_L) + I_{ext} \quad (1)$$

$$\frac{1}{\gamma(T)} \frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n = \frac{n^\infty(V) - n(t)}{\tau_n(V)} \quad (2)$$

$$\frac{1}{\gamma(T)} \frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m = \frac{m^\infty(V) - m(t)}{\tau_m(V)} \quad (3)$$

$$\frac{1}{\gamma(T)} \frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h = \frac{h^\infty(V) - h(t)}{\tau_h(V)} \quad (4)$$

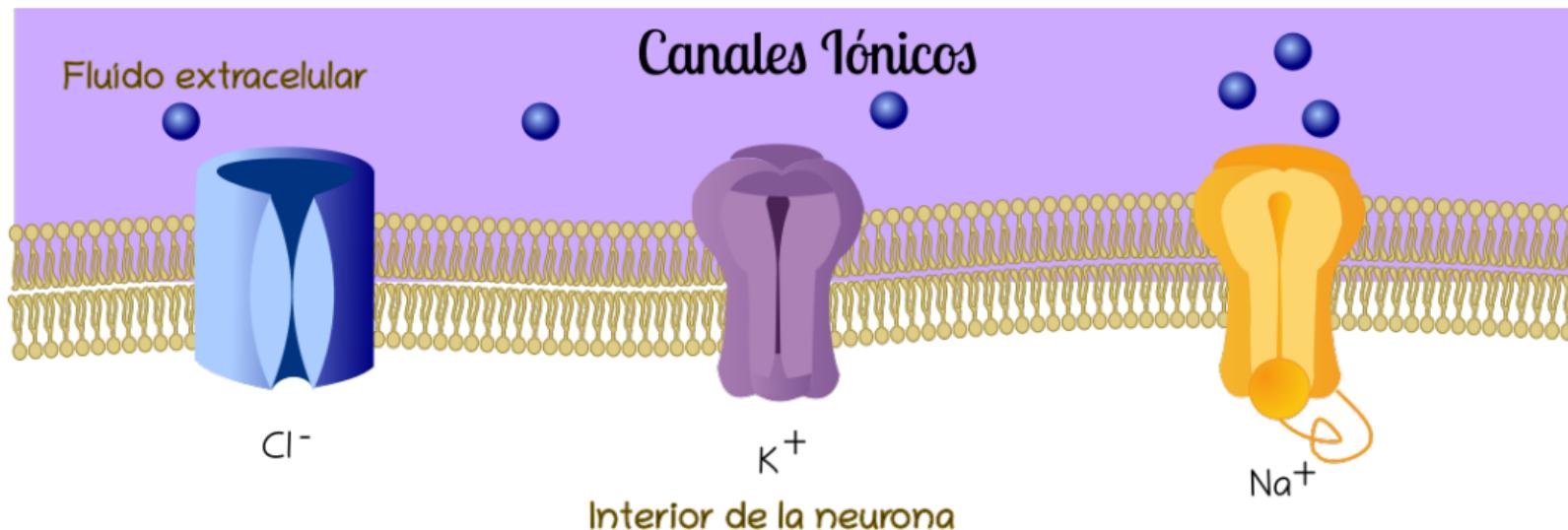
con los potenciales de Nernst E_{Na} , E_K , E_L .

Temas

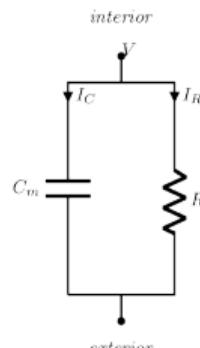
1 Modelo de Hodgkin Huxley

- Bi-capa de lípidos
- Compuertas iónicas

Bi-capa de lípidos



Bi-capa de lípidos



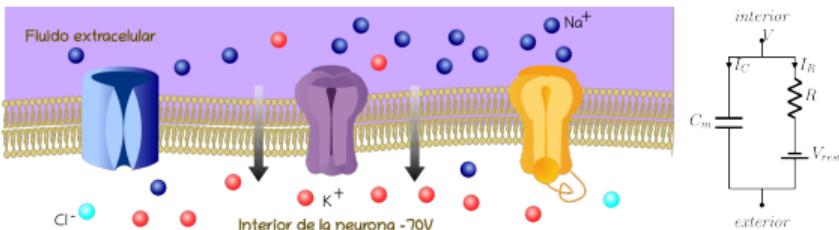
$$I_C + I_R - I_{ext} = 0 \quad (5)$$

$$C \frac{dV}{dt} + \frac{V}{R} - I_{ext} = 0 \quad (6)$$

$$C \frac{dV}{dt} = -\frac{V}{R} + I_{ext} \quad (7)$$



- Una pequeña cantidad de carga pasa entre ambas capas.
- La carga es almacenada por un breve periodo, comportándose como un *capacitor* $C = \frac{Q}{V}$.
- Ofrece cierta *resistencia* al paso de corriente $V = RI$.



$$C \frac{dV}{dt} = -\frac{(V - V_{rest})}{R} + I_{ext} \quad (8)$$

- Reacomodar como:

$$\tau \frac{dV}{dt} = -V + V_\infty \quad (9)$$

- ¿Cuánto vale la constante de tiempo τ ?
- ¿Cuál es el estado estacionario $\frac{dV}{dt} = 0$? ¿Cuánto vale V_∞ ?

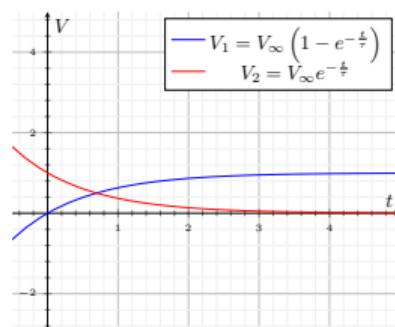
Solución para un pulso escalón

- Demostrar que las soluciones para los segmentos de un pulso escalón son:

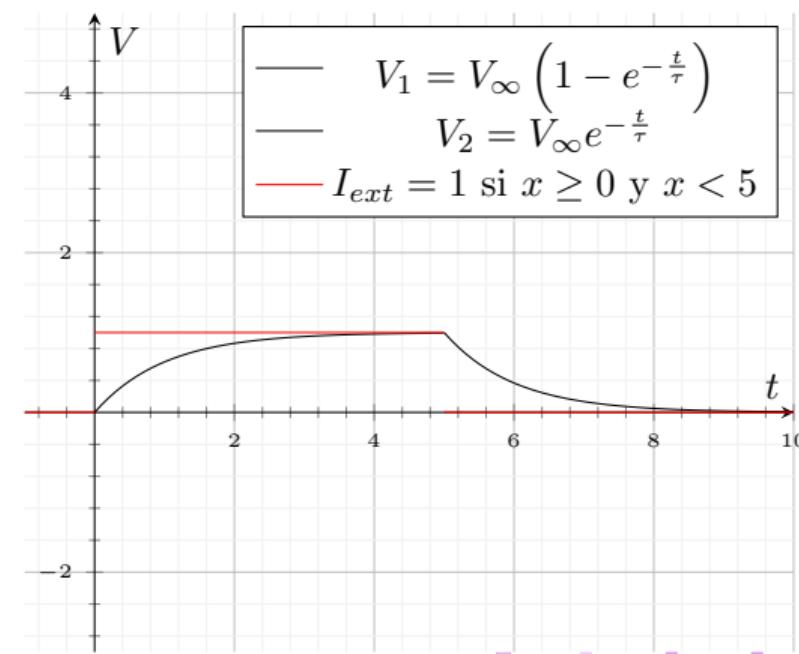
$$V = V_\infty \left(1 - e^{-\frac{t}{\tau}}\right) \quad \text{si } I_{ext} > 0 \quad (10)$$

$$V = V_\infty e^{-\frac{t}{\tau}} \quad \text{si } I_{ext} = 0 \quad (11)$$

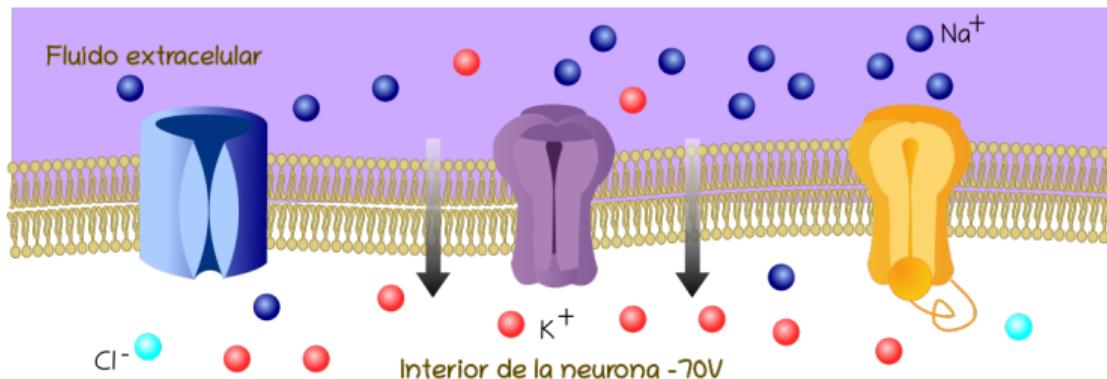
Soluciones para el pulso



Soluciones para el pulso



Diferencia de potencial afuera y adentro

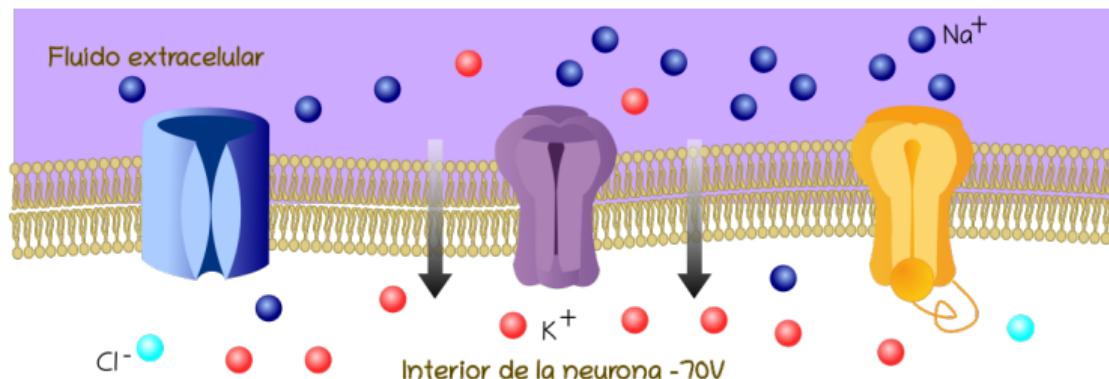


- Esto provoca el efecto de una pila.
- Potenciales de Nernst E_{Na} , E_{K} , E_{L} dados por:

$$E = \frac{k_B T}{zq} \ln \frac{[\text{adentro}]}{[\text{afuera}]} \quad (12)$$

con k_B la constante de Boltzman, q es la carga del ion y z , el número de iones.

Polarización



La membrana se puede encontrar:

Polarizada en su estado de reposo con $V < 0$ ($V \approx -70\text{mV}$).

Despolarizada cuando $V \geq 0$.

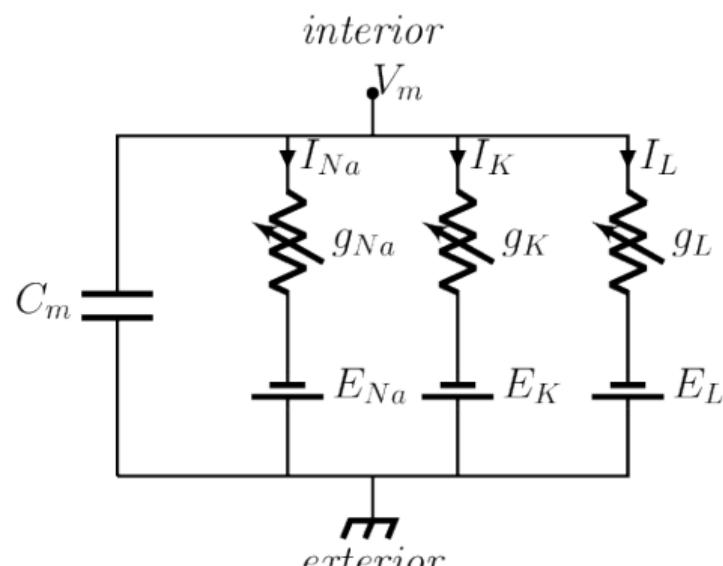
Hiperpolarizada cuando la diferencia de potencial incrementa su magnitud $V \ll 0$

Temas

1 Modelo de Hodgkin Huxley

- Bi-capa de lípidos
- Compuertas iónicas

Modelo de Hodgkin Huxley



$$C_m \frac{dV}{dt} = -g_{Na} m^3 h (V - E_{Na}) - g_K n^4 (V - E_K) - g_L (V - E_L) + I_{ext} \quad (1)$$

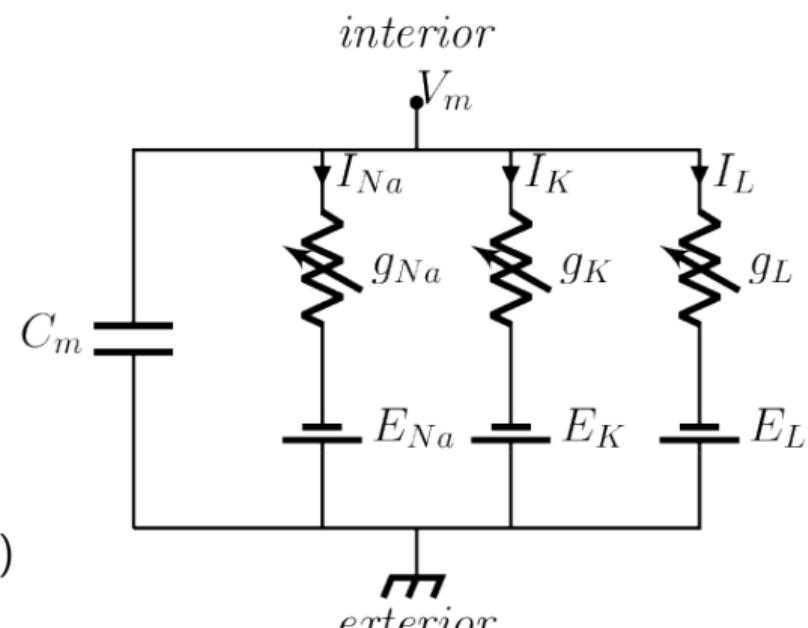
Conductancias

$$G = \frac{1}{R}$$

$$G_{Na} = g_{Na} m^3 h$$

$$G_K = g_K n^4$$

$$C_m \frac{dV}{dt} = - G_{Na}(V - E_{Na}) - G_K(V - E_K) - g_L(V - E_L) + I_{ext} \quad (1)$$



Canal persistente Tiene un sólo tipo de compuerta y dos estados posibles:

Activado

Desactivado

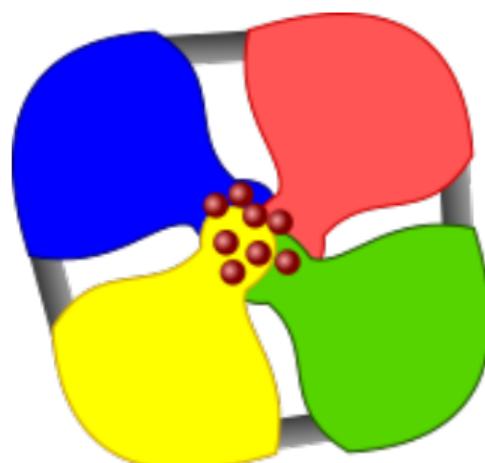
Canal transitorio Tiene compuertas de *activación* e *inactivación*, y tres estados.

Desactivada Compuerta de activación cerrada, inactivación abierta.

Activada Ambas compuertas abiertas.

Inactivada Compuerta de inactivación cerrada.

Modelo estadístico de conductancia de los canales



Compuerta para el Potasio (K)

n Probabilidad de que cada compuerta esté abierta.

$(1 - n)$ Probabilidad de que cada compuerta esté cerrada.

$P_K \sim n^4$ Probabilidad de que el canal de K se abra.

Cada compuerta se abre o cierra con una probabilidad independiente de las otras puertas, pero dependiente del voltaje V de la membrana.

$\gamma(T)$ Escala temporal dependiente de la temperatura.

$\alpha_n(V)$ Probabilidad de que una compuerta transite de cerrada a abierta.

$\beta_n(V)$ Probabilidad de que una compuerta transite de abierta a cerrada.

$$\frac{1}{\gamma(T)} \frac{dn}{dt} = \alpha_n(V)(1 - n) - \beta_n(V)n = \frac{n^\infty(V) - n(t)}{\tau_n(V)} \quad (2)$$

$$\tau_n(V) = \frac{1}{\alpha_n(V) + \beta_n(V)} \quad (13)$$

$$n^\infty(V) = \frac{\alpha_n(V)}{\alpha_n(V) + \beta_n(V)} \quad (14)$$

$n^\infty(V)$ probabilidad de compuerta abierta en el equilibrio cuando $t \rightarrow \infty$.

$\tau_n(V)$ Tiempo que toma llegar al equilibrio.

Canales controlados por voltaje para el Sodio (Na)

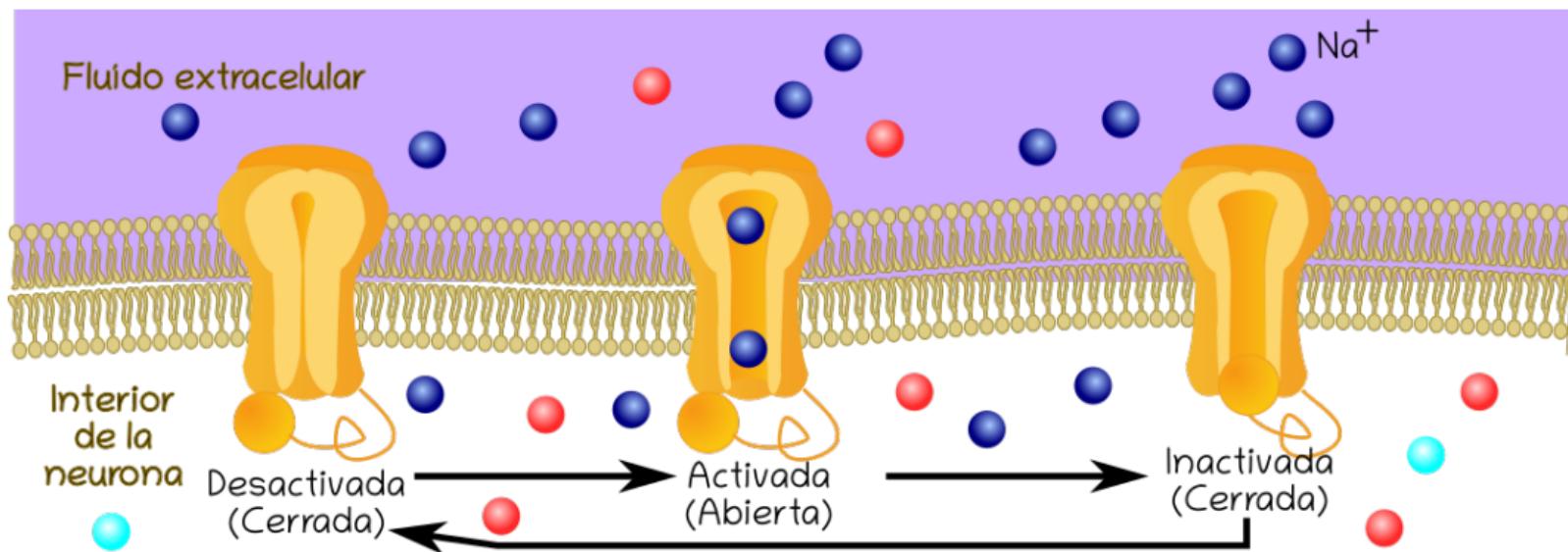


Figura: Vista de los estados desactivado, activado e inactivado.

m Probabilidad de que cada compuerta esté abierta.

$(1 - m)$ Probabilidad de que cada compuerta esté cerrada.

h Probabilidad de que la puerta de activación no esté cerrada.

$$P_{Na} \sim m^3 h$$

$$\frac{1}{\gamma(T)} \frac{dm}{dt} = \alpha_m(V)(1 - m) - \beta_m(V)m = \frac{m^\infty(V) - m(t)}{\tau_m(V)} \quad (3)$$

$$\frac{1}{\gamma(T)} \frac{dh}{dt} = \alpha_h(V)(1 - h) - \beta_h(V)h = \frac{h^\infty(V) - h(t)}{\tau_h(V)} \quad (4)$$

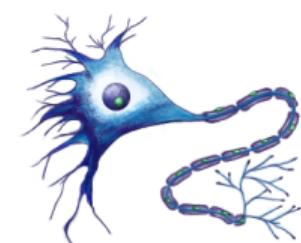
Dinámica del voltaje

1 Modelo de Hodgkin Huxley

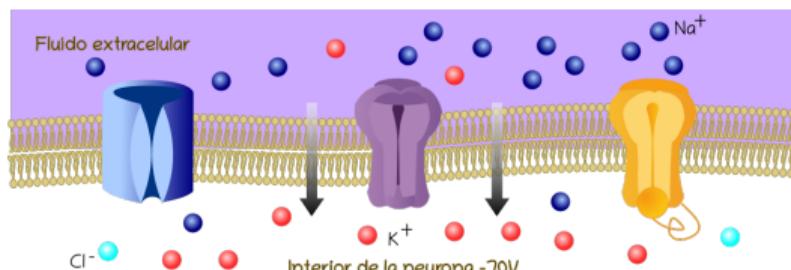
2 Dinámica del voltaje

3 Simulación

4 Información codificada en las dendritas



Polarización



La membrana se puede encontrar:

Polarizada en su estado de reposo con $V < 0$ ($V \approx -70\text{mV}$).

Despolarizada cuando $V \geq 0$.

La despolarización puede ser provocada por corrientes positivas externas.
(Iones positivos entran a la célula).

Hiperpolarizada cuando la diferencia de potencial incrementa su magnitud $V \ll 0$

La hiperpolarización puede ser provocada por corrientes positivas desde el interior. (Iones positivos salen de la célula).

Dinámica del voltaje

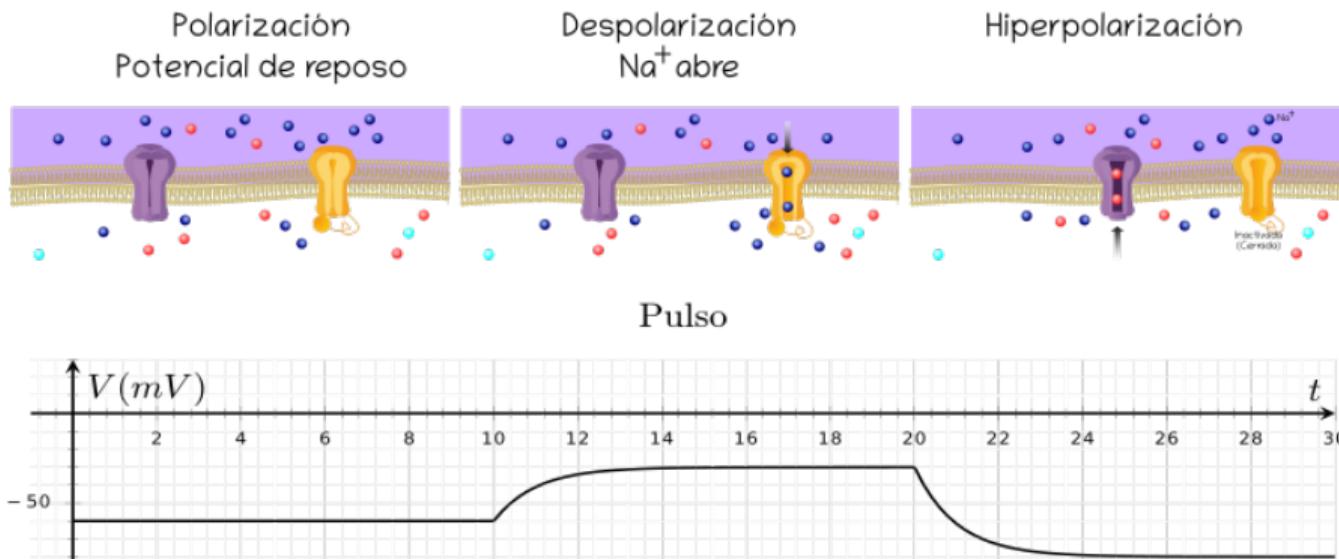


Figura: Cuando se rebasa el voltaje umbral, los canales de Na^+ y K^+ interactúan para producir una rápida despolarización de la membrana, para luego hiperpolarizarla.

α y β medidas experimentalmente

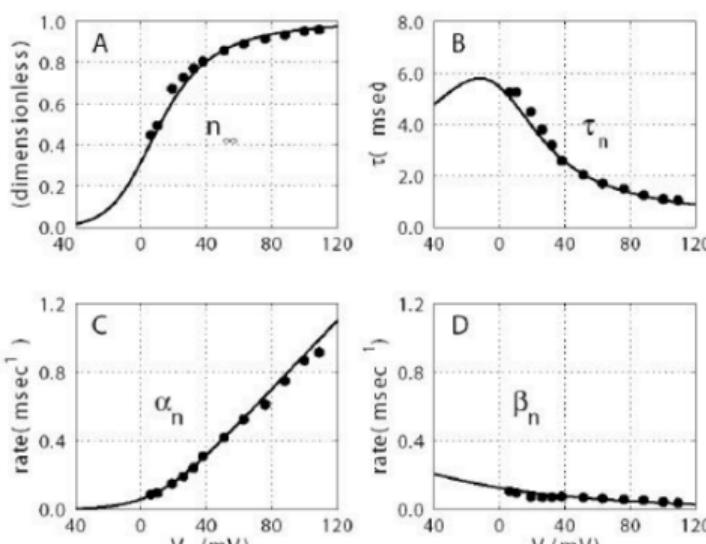


Figura: Medición experimental de los parámetros y ajuste manual de curvas. Imagen de Nelson 2004

Curvas paramétricas ajustadas

$$\alpha_n(V) = \frac{0.01(10 - V)}{e^{\left(\frac{10-V}{10}\right)} - 1} \quad \beta_n(V) = 0.125e^{-\frac{V}{80}} \quad (15)$$

$$\alpha_m(V) = \frac{0.1(25 - V)}{e^{\left(\frac{25-V}{10}\right)} - 1} \quad \beta_m(V) = 4e^{-\frac{V}{18}} \quad (16)$$

$$\alpha_h(V) = 0.07e^{-\frac{V}{20}} \quad \beta_h(V) = \frac{1}{e^{\left(\frac{30-V}{10}\right)} + 1} \quad (17)$$

Activación e inactivación de los canales

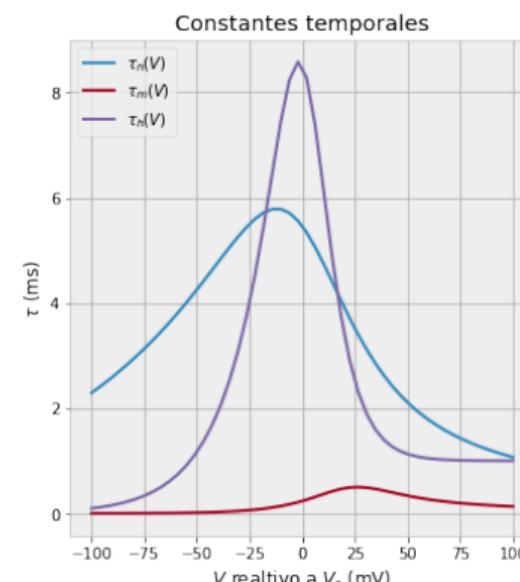
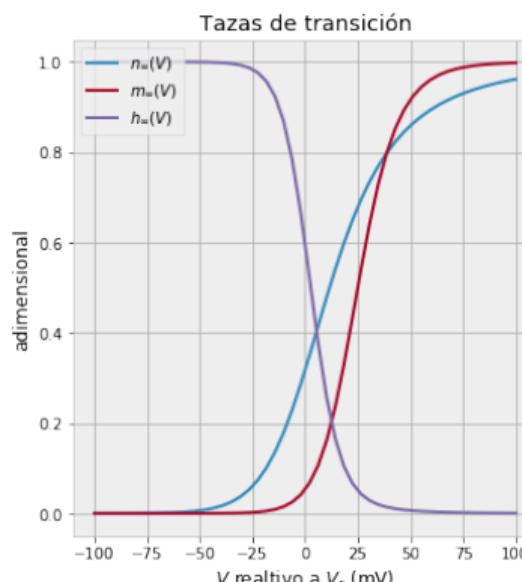
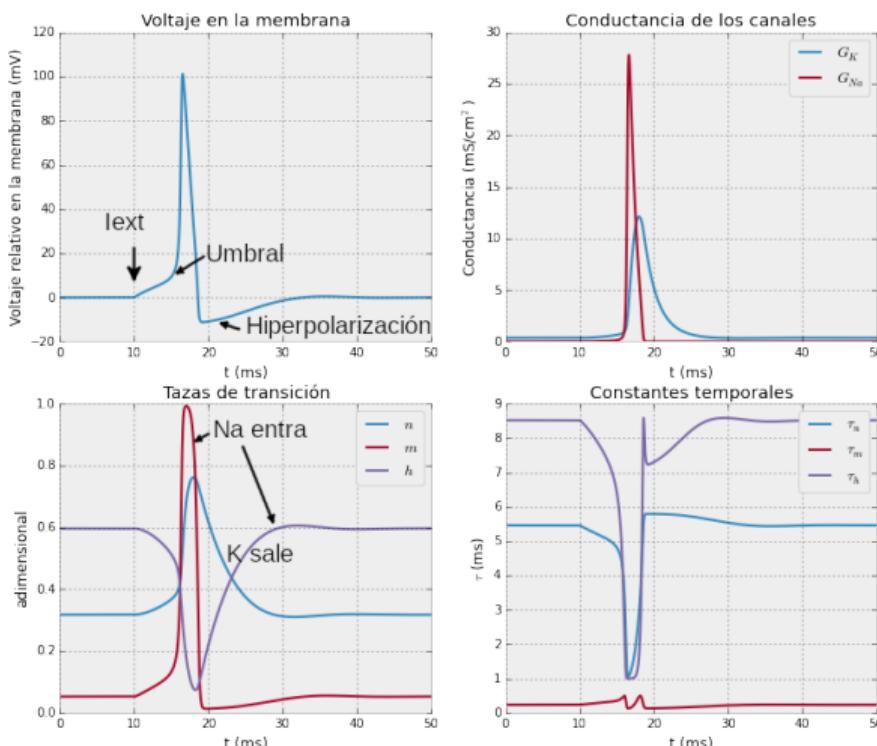


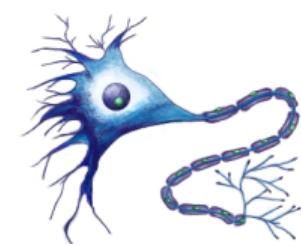
Figura: Las compuertas de Na^+ se abren primero, luego las de K^+ y esto provoca que se inactiven las de Na^+ .

Disparo



Simulación

- 1 Modelo de Hodgkin Huxley
- 2 Dinámica del voltaje
- 3 Simulación
- 4 Información codificada en las dendritas



Algoritmo de integración de Euler I

Algoritmo 1 Algoritmo de integración de Euler [Wells pp51].

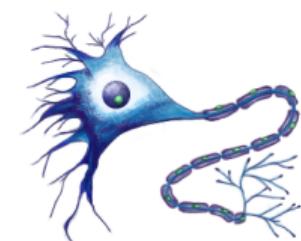
- 1: **function** INTEGRADISPARO($T, \Delta T, V_0, I_{ext}(t)$)
- 2: Inicializar arreglos de longitud T :
 $T[], V[], n[], m[], h[], GNa[], GK[], \tau_{nN}, \tau_{mM}, \tau_h \leftarrow$
 arreglo[numeroDePasos]
- 3: $V[0] \leftarrow V_0$
- 4: **for** $t = 0$ a $t = T$ cada Δt **do**
- 5: Calcular $\alpha_n, \beta_n, \alpha_m, \beta_m, \alpha_h, \beta_h$ utilizando $V(t)$.
- 6: Calcular las τ_x y x^∞ a partir de las anteriores.
- 7: Calcular las probabilidades de las compuertas n, m, h , utilizando las ecuaciones en diferencias en su forma matricial $\Pi(t + \Delta t) = A_\pi \Pi(t) + B_\pi$.

Algoritmo de integración de Euler II

- 8: Calcular $G_{Na} = g_{Na}m^3h$ y $G_K = g_Kn^4$.
- 9: Almacenar los resultados de este paso en los arreglos
 $T[], V[], n[], m[], h[], GNa[], GK[], tauN[], tauM[], tauH[]$.
- 10: $I_{ext} \leftarrow IEXT(t)$
- 11: Calcular $V_m(t + \Delta t)$
- 12: Devolver los arreglos $T[], V[], n[], m[], h[], GNa[], GK[], tauN[], tauM[], tauH[]$
con los resultados para los tiempos $[0, T]$.

Información codificada en las dendritas

- 1 Modelo de Hodgkin Huxley
- 2 Dinámica del voltaje
- 3 Simulación
- 4 Información codificada en las dendritas



Disparos de una neurona

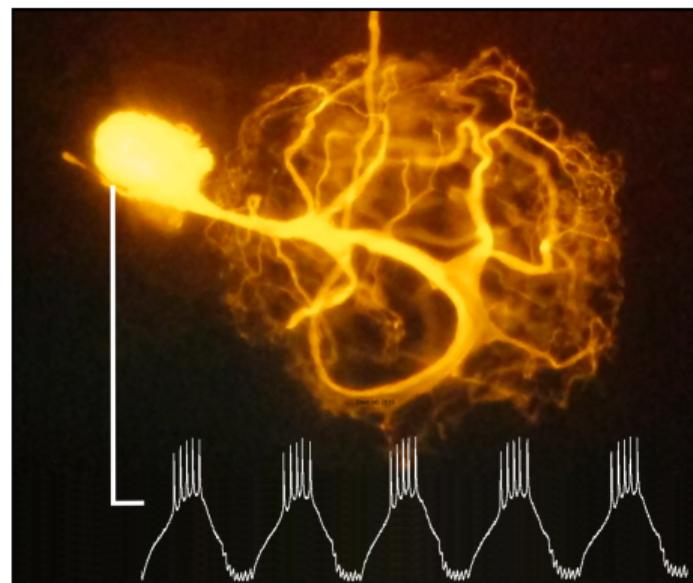


Figura: Tinción fluorescente de una neurona de cangrejo. También se muestran las oscilaciones rítmicas del potencial en la membrana.

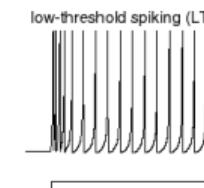
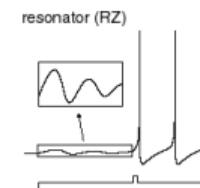
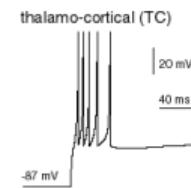
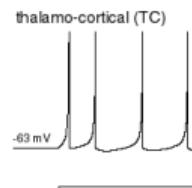
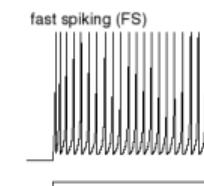
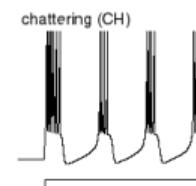
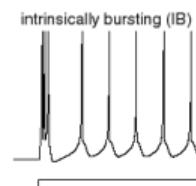
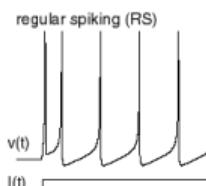
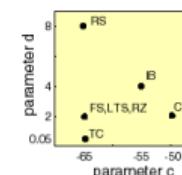
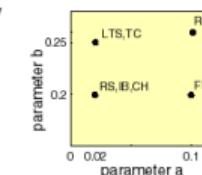
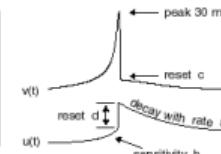
https://commons.wikimedia.org/wiki/File:PD_neuron_staining_and_recording.png

Diferentes patrones de disparo (simulados)

$$v' = 0.04v^2 + 5v + 140 - u + l$$

$$u' = a(bv - u)$$

if $v = 30 \text{ mV}$,
then $v \leftarrow c$, $u \leftarrow u + d$



<http://www.izhikevich.org/publications/spikes.htm>

Localización del sonido por disparidad

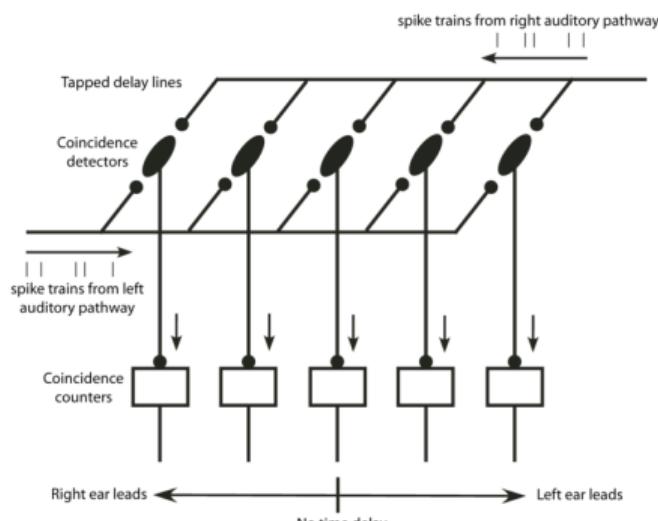


Figura: El sistema auditivo registra y analiza pequeñas diferencias en el tiempo de llegada de los sonidos a los dos oídos para estimar la dirección desde la cual el sonido es emitido.

Fuente: http://www.scholarpedia.org/article/Jeffress_model

Aprendizaje de máquina

- 1 Aprendizaje de máquina
- 2 Espacio de hipótesis
- 3 Conjuntos de datos para el entrenamiento
- 4 Tipos de aprendizaje

Aprendizaje de máquina

El *aprendizaje de máquina* es el campo de estudio que dota a las computadoras de la habilidad de aprender sin haber sido programadas explícitamente.

Definición

Se dice que un programa de computadora aprende de la experiencia E, con respecto a una tarea T, y una medida de desempeño D si su desempeño en T, como lo mide D, mejora con la experiencia E.^a

^aMitchell 1997

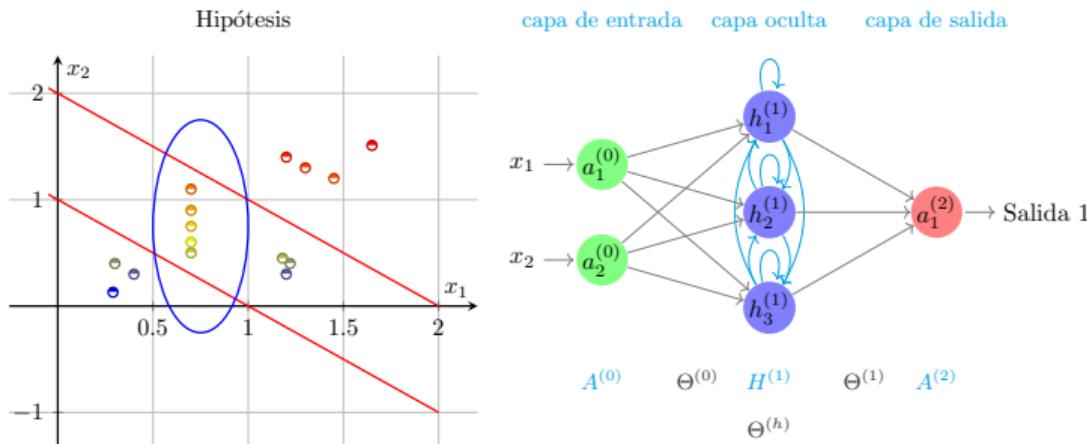
Espacio de hipótesis

- 1 Aprendizaje de máquina
- 2 Espacio de hipótesis
- 3 Conjuntos de datos para el entrenamiento
- 4 Tipos de aprendizaje

Hipótesis

- El aprendizaje se realiza con respecto a una familia de *hipótesis*.
- La forma de este espacio está dada por la familia funciones que podría mapear los datos entrada en datos de salida.
- El aprendizaje consiste en seleccionar una hipótesis de ese espacio (una función), de tal manera que, dado un dato de entrada, sea posible predecir alguna característica de salida deseada.

- En el caso de las redes neuronales, el espacio de hipótesis está determinado por la arquitectura de la red.



Conjuntos de datos para el entrenamiento

- 1 Aprendizaje de máquina
- 2 Espacio de hipótesis
- 3 Conjuntos de datos para el entrenamiento
- 4 Tipos de aprendizaje

Clasificación de los conjuntos de datos

Para entrenar un algoritmo de aprendizaje se requieren datos de entrenamiento, separados en tres conjuntos:

Entrenamiento Datos con los cuales se ajustan los parámetros de la hipótesis.

Validación Datos utilizados para ajustar los parámetros del algoritmo de entrenamiento, que puedan afectar qué hipótesis es seleccionada.

Prueba Datos utilizados para evaluar la posibilidad de que la hipótesis aprendida generalice^[1] a datos no vistos anteriormente.

^[1]Es decir, que sea válida también para datos nuevos.

Tipos de aprendizaje

- 1 Aprendizaje de máquina
- 2 Espacio de hipótesis
- 3 Conjuntos de datos para el entrenamiento
- 4 Tipos de aprendizaje

Aprendizaje supervisado

Decimos que el aprendizaje es *supervisado* si para cada ejemplo x entre los datos de entrenamiento X se conoce la respuesta correcta y .

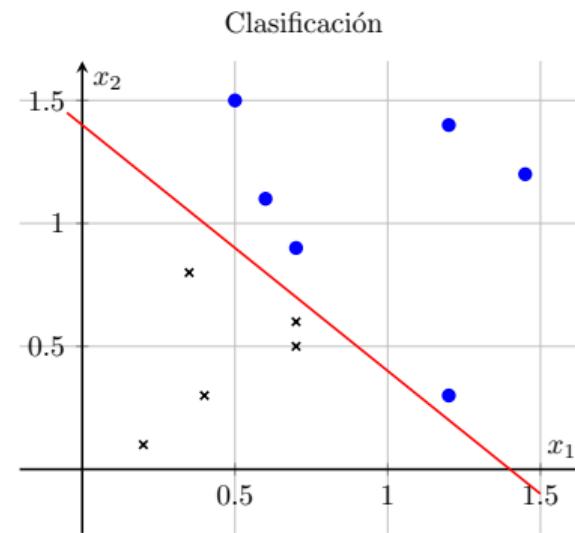
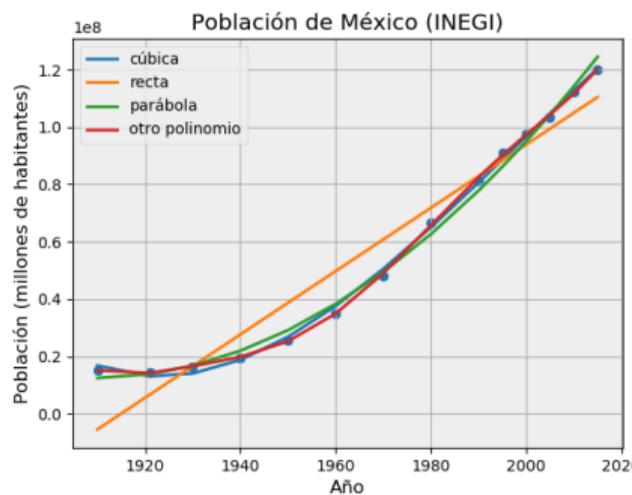


Figura: Ejemplos de aprendizaje supervisado. Izquierda: *Regresión*. Derecha *Clasificación*

Regresión

Un modelo de *regresión* busca predecir valores de salida **continuos**.

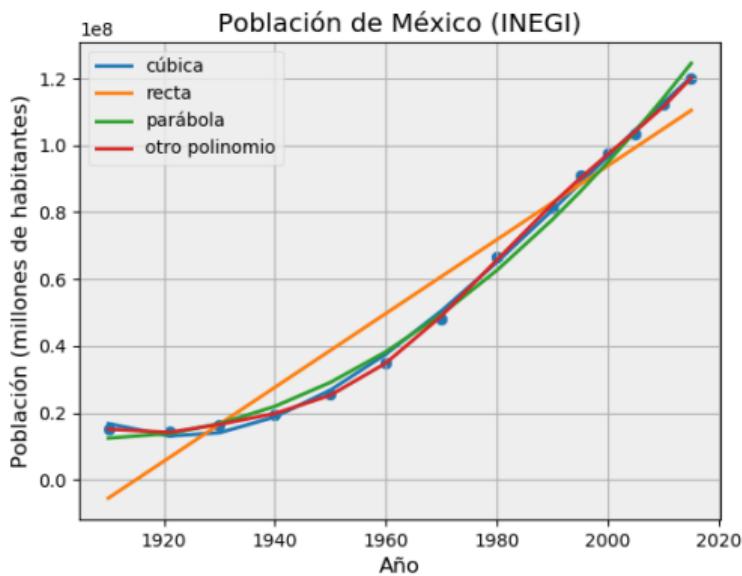


Figura: Función de *Regresión*: Dada la entrada $x = (\text{año})$, ¿cuántos habitantes tendría México?.

Clasificación

En un problema de *clasificación* se desea predecir una salida discreta.

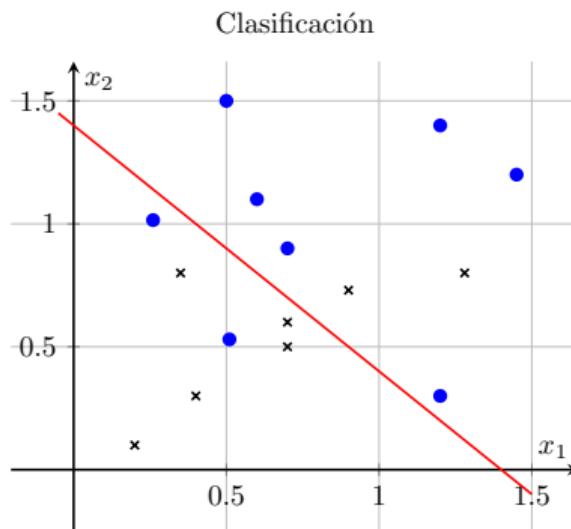


Figura: Problema de *Clasificación*: Dadas las entradas $x = (\text{edad}, \text{tamaño_del_tumor})$, ¿el cáncer es maligno o benigno?

Aprendizaje no supervisado

- No se tienen valores *correctos* o *incorrectos*
- El objetivo del aprendizaje no supervisado es descubrir estructura en los datos.

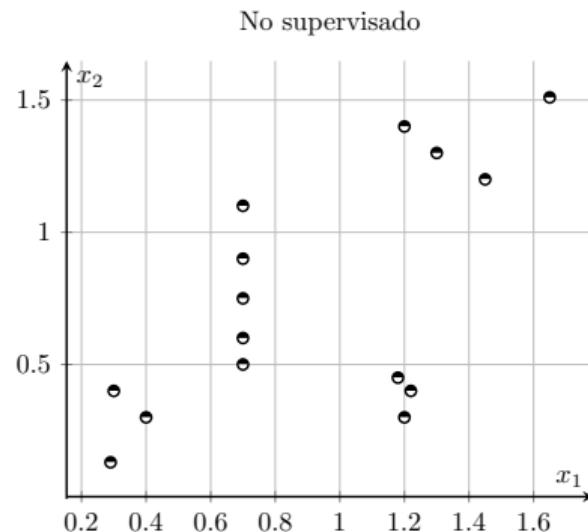
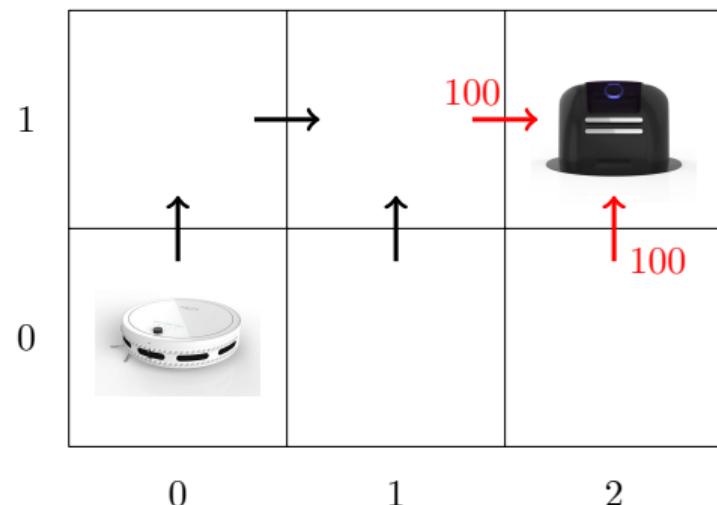


Figura: Aprendizaje no supervisado ¿qué estructura tienen los datos?

Aprendizaje por refuerzo

Para cada estado se desea aprender una *política* que indique la acción que maximizará la *recompensa* recibida en el menor tiempo posible.

Política π	
(0,0)	↑
(1,0)	↑
(2,0)	↑
(0,1)	→
(1,1)	→



Perceptrón

1 Perceptrón

2 Funciones de activación

3 Funciones de error

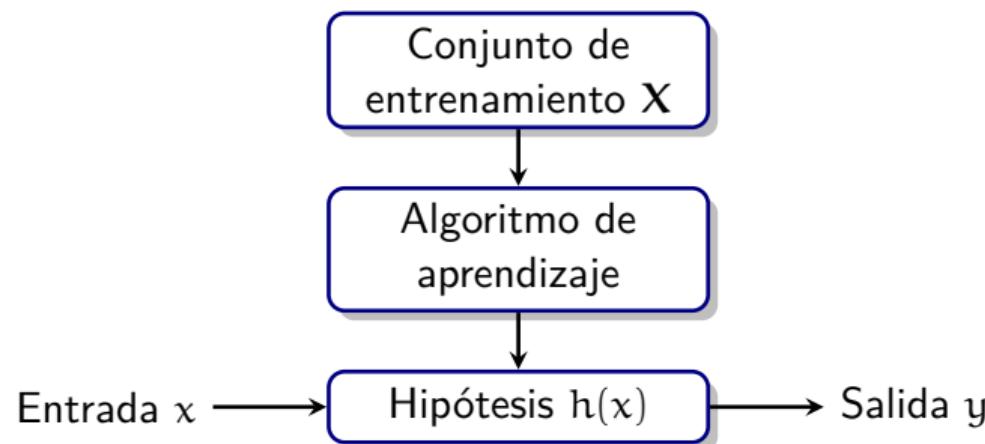
Temas

1 Perceptrón

- Clasificación lineal
 - Compuertas lógicas con neuronas

Problema

- En un problema de clasificación se pretende, dadas las características de un ejemplar, identificar la clase a la cual pertenece.



$$h(x) : \mathbb{R}^n \rightarrow \{e | e \text{ es una etiqueta de clase}\}$$

Clasificación con una sola frontera

- En el caso más sencillo, se trata de separar los datos de entrada en dos clases.

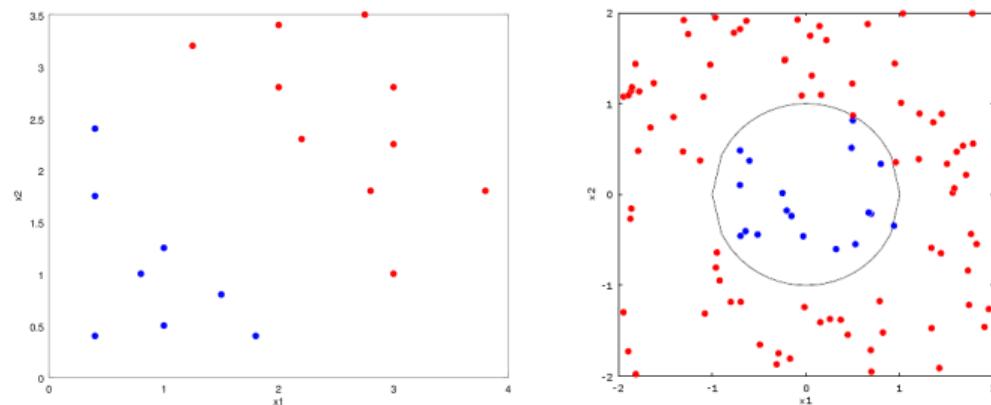


Figura: Problema de clasificación, ilustrado en el plano. Izq: frontera lineal, se dice que los datos son *linealmente separables*. Der: frontera circular.

Perceptrón (Rosenblatt 1958)

- Un *perceptrón* es una función matemática inspirada en el funcionamiento de una neurona.
- Sirve para obtener fronteras lineales entre datos multivariados mediante un algoritmo de aprendizaje.
- Fueron popularizados por el psicólogo Frank Rosenblatt a inicios de los 1960's, en un libro llamado *Principios de neurodinámica*, donde presentó varios modelos de perceptrones.
- Los perceptrones cayeron en desuso porque Minsky y Papert demostraron sus limitaciones en un libro llamado *Perceptrones*.

Perceptrón

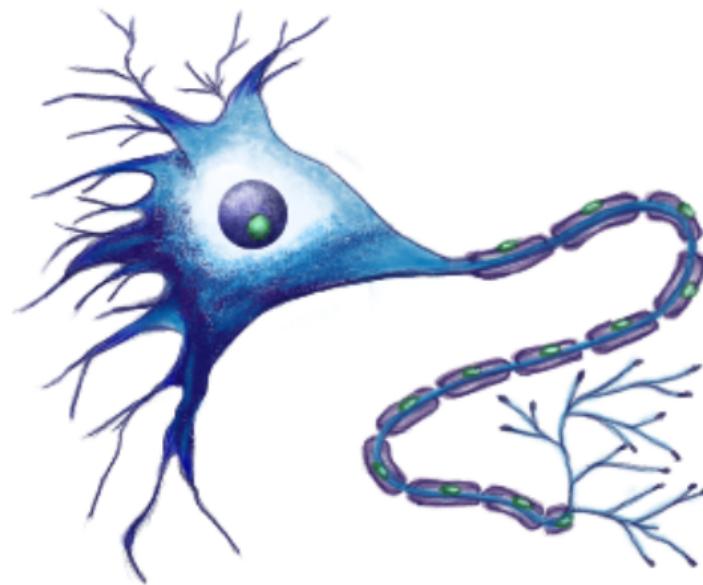


Figura: Neurona.

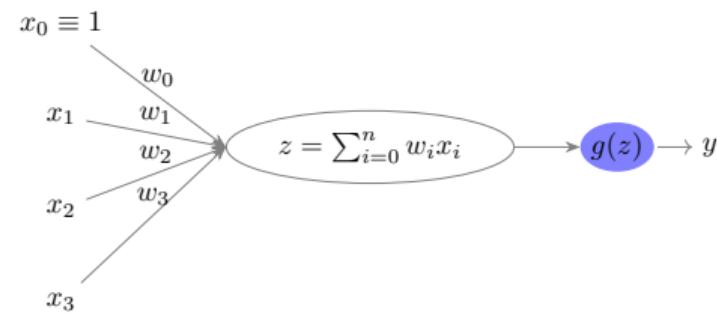
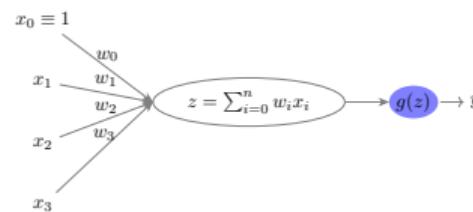


Figura: Perceptrón.

Donde:

- x_0 es el *sesgo* (*bias*),
- g es la *función de activación*,
- w_i son los *pesos* de las conexiones de entrada.

Función de activación



Donde g , la *función de activación*, puede ser alguna de:

- Función escalón

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

- ReLU

$$f(x) = \max(0, x)$$

- Logística

$$g = \frac{1}{1 + e^{-z}}$$

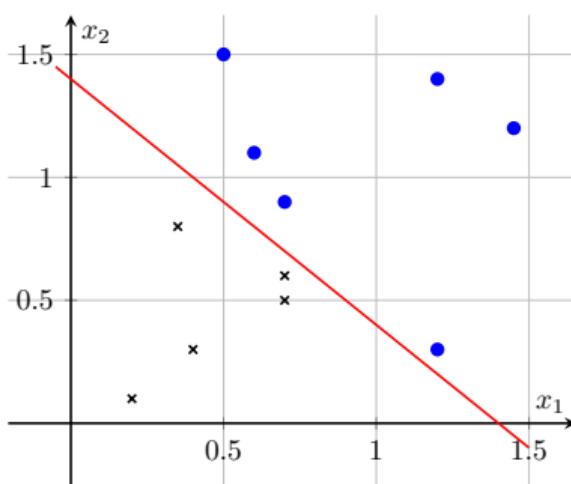
- Tangente hiperbólica

$$\tanh = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Interpretación

El perceptrón modela una neurona que dispara o no dependiendo de los valores de sus señales de entrada.

Clasificación



$$\begin{aligned}x_0 &\equiv 1 \\w_0 &= -1.4 \\w_1 &= 1 \\w_2 &= 1 \\x_1 & \quad \quad \quad g(\sum_{i=0}^n w_i x_i) \rightarrow y \\x_2\end{aligned}$$

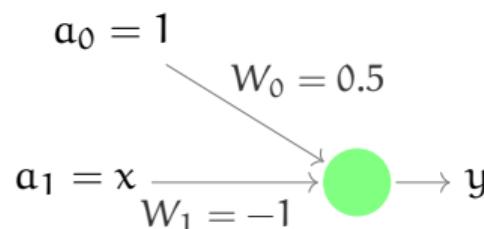
$$y = \frac{1}{1 + e^{-(1.4 + x_1 + x_2)}}$$

Temas

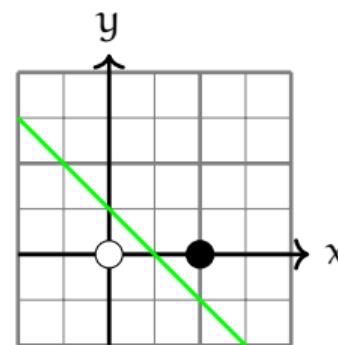
1 Perceptrón

- Clasificación lineal
- Compuertas lógicas con neuronas

NOT



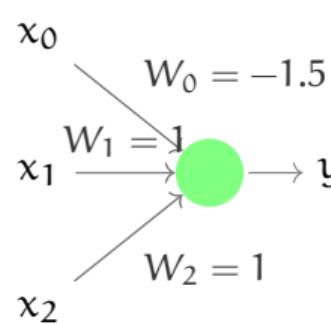
NOT



$$\text{Salida} = g(0.5 - x)$$

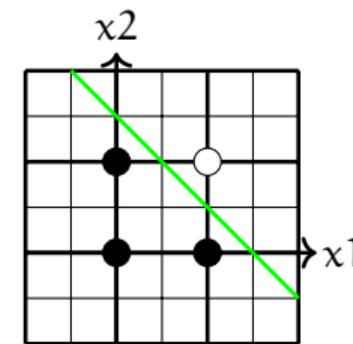
x	z	y
0	0.5	1
1	-0.5	0

AND



$$\text{Salida} = g(-1.5 + x_1 + x_2)$$

x_1	x_2	z	h_{umbral}
0	0	-1.5	0
0	1	-0.5	0
1	0	-0.5	0
1	1	0.5	1

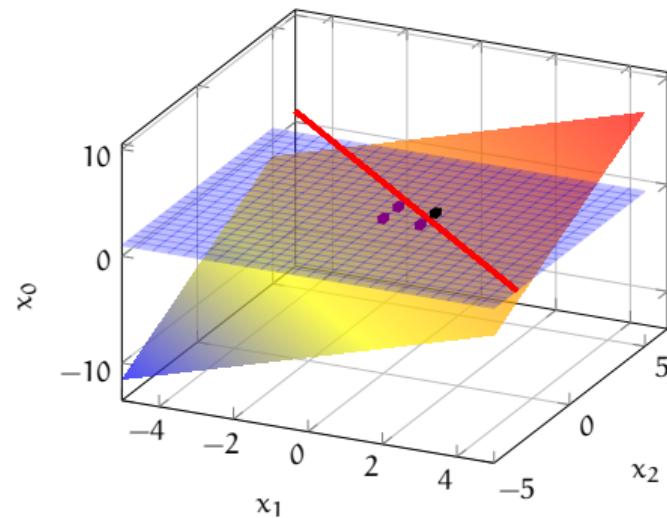


$$\text{Salida} = g(-15 + 10x_1 + 10x_2)$$

x_1	x_2	z	h_{sigmoide}
0	0	-15	0.0000003
0	1	-5	0.0067
1	0	-5	0.0067
1	1	5	0.993

Linealidad

Para que un perceptrón en \mathbb{R}^2 represente una función lineal, debemos considerar el espacio tridimensional \mathbb{R}^3 .



Entrenamiento

- En ocasiones, no es posible visualizar los datos de entrada, especialmente si son multidimensionales.
- El perceptrón incluye un algoritmo de entrenamiento que permite encontrar los pesos adecuados para realizar la predicción correcta.

Seleccionar casos de entrenamiento utilizando cualquier política que garantice que cada caso de entrenamiento continuará siendo elegido.

- ① Si la salida es correcta, dejar los pesos como están.
- ② Si la salida incorrectamente indica 0, sumar el vector de entrada al vector de pesos.
- ③ Si la salida incorrectamente indica 1, restar el vector de entrada al vector de salida.

Está garantizado que este procedimiento encontrará un conjunto de pesos que obtienen la respuesta correcta para todos los casos de entrenamiento, *si existen*. Esto dependerá de que se hayan elegido las características correctas.

Uso de los perceptrones

En reconocimiento de patrones estadísticos tradicionalmente se siguen tres pasos:

- ① Se convierten los vectores de entrada en vectores de activación, que representan características. Estas características se eligen a mano.
- ② Se aprende con qué peso contribuye cada una de las características, y la suma pesada de las entradas determina un valor numérico.
- ③ Si esta cantidad se encuentra por encima de un valor numérico, se dice que el vector de entrada corresponde con la clase objetivo.

Funciones de activación

1 Perceptrón

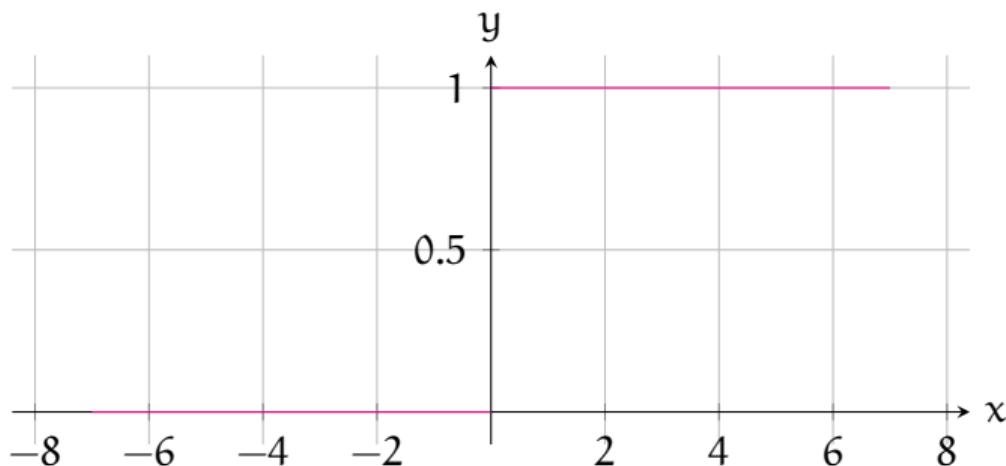
2 Funciones de activación

3 Funciones de error

Función escalón

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

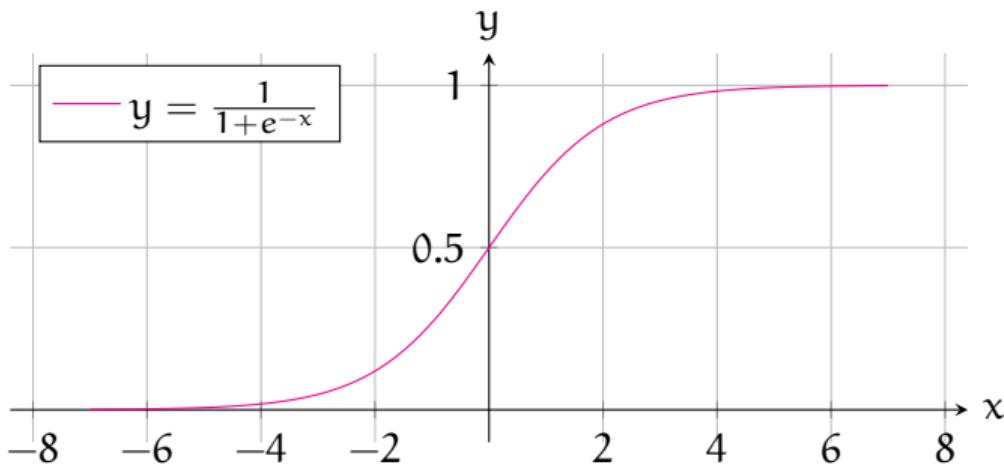
Rectificadora o función rampa



Función logística

$$y = \frac{1}{1 + e^{-x}} \quad (2)$$

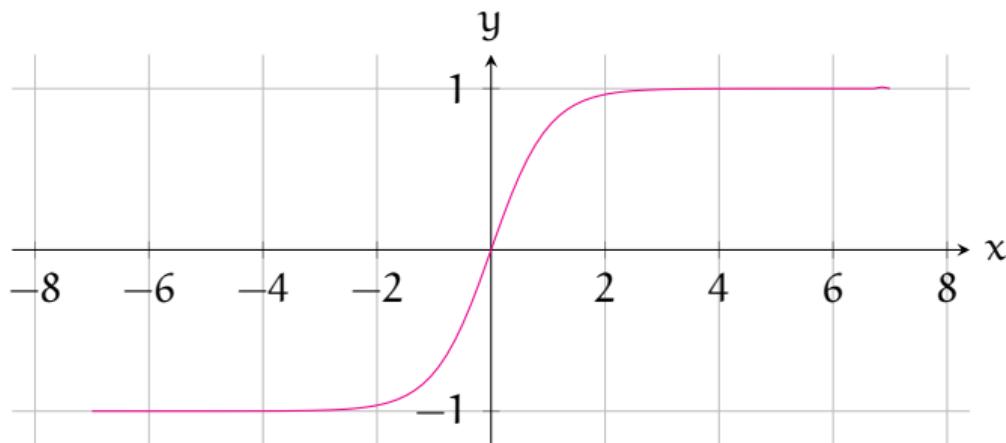
Función logística



Tangente hiperbólica

$$y = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

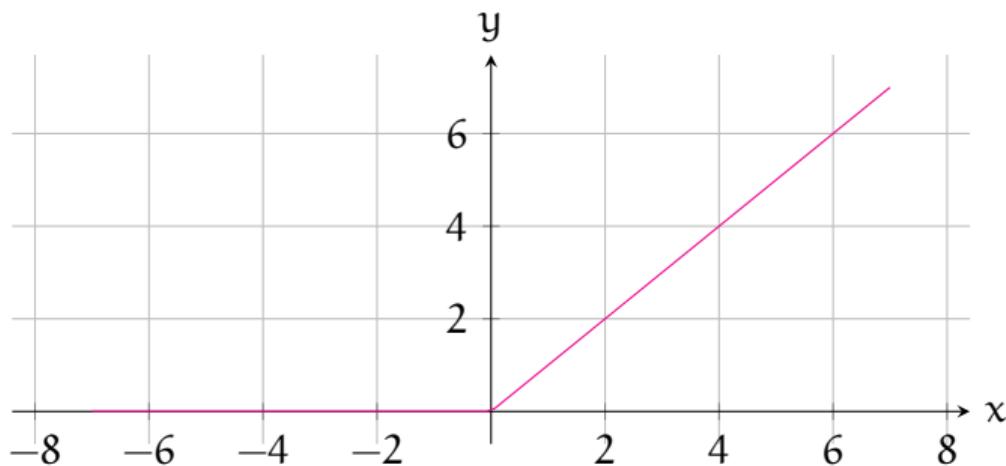
Tangente hiperbólica



Rectificadora lineal

$$y = \max(0, x) \quad (4)$$

Rectificadora o función rampa



Funciones de error

1 Perceptrón

2 Funciones de activación

3 Funciones de error

Funciones de error

- Para determinar si el perceptrón está clasificando correctamente los datos o no, se definen funciones de error.
- Para entrenar un perceptrón basta con utilizar cualquier método de optimización de funciones para encontrar los parámetros θ que minimizan la función de error.
- Sea $y^{(m)}$ el valor de salida deseado para el m -ésimo ejemplar de entrenamiento para un perceptrón con pesos Θ , es decir, la clasificación correcta 0 ó 1.
- Sea $a^{(m)}$ el valor de salida del perceptrón.
- Existen varias funciones de error, pero se presentan a continuación las dos más utilizadas.

Diferencias al cuadrado

La función de error *diferencias al cuadrado* se define como:

$$L(\Theta) = \frac{1}{2m} \sum_{m=0}^{M-1} (y^{(m)} - a^{(m)})^2 \quad (5)$$

Históricamente fue la primera en popularizarse, pero no es la más adecuada para problemas de clasificación.

Entropía cruzada

La función de error *entropía cruzada* se define como:

$$L(\Theta) = -\frac{1}{m} \sum_{m=0}^{M-1} \left(y^{(m)} \log(a^{(m)}) + (1 - y^{(m)}) \log(1 - a^{(m)}) \right) \quad (6)$$

Esta función tiene un mejor comportamiento ya que define una superficie más suave en el espacio de parámetros para problemas de clasificación.

Funciones no separables linealmente

1 Funciones no separables linealmente

2 Modelo computacional de una red neuronal

Temas

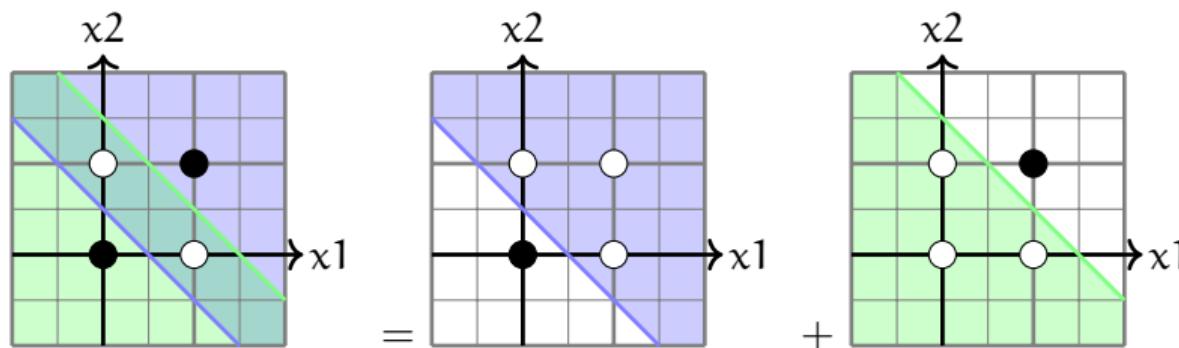
1 Funciones no separables linealmente

- XOR
- Alimentación hacia adelante

XOR (Minsky & Papert, 1969 → Rumerhart et al. 1986)

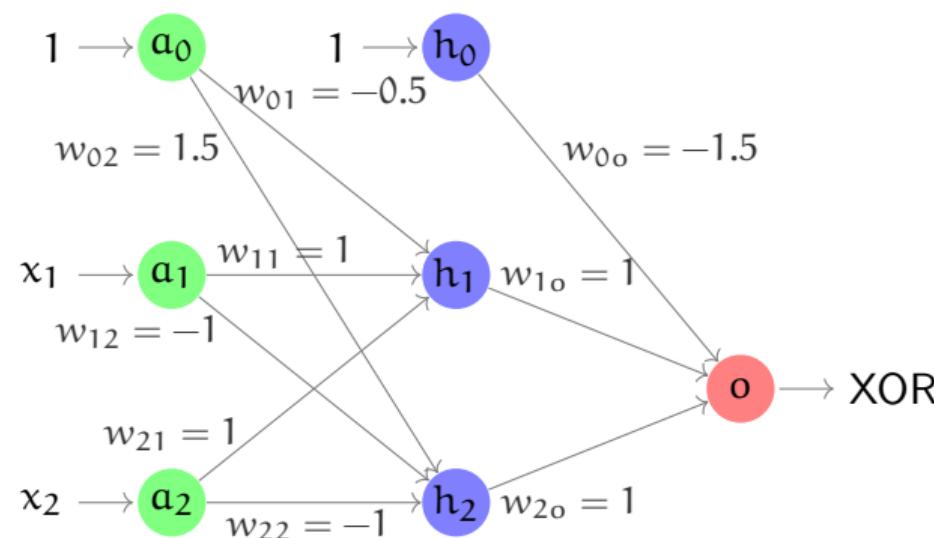
- XOR es una compuerta lógica con valores no separables linealmente en el plano.
- Se puede escribir como:

$$\text{XOR} = (x_1 \vee x_2) \wedge \neg(x_1 \wedge x_2) = (x_1 \vee x_2) \wedge (x_1 \text{ NAND } x_2) \quad (1)$$



XOR

capa de entrada capa oculta capa de salida

Con $W_{\langle \text{origen} \rangle \langle \text{destino} \rangle}$.

Temas

1 Funciones no separables linealmente

- XOR
- Alimentación hacia adelante

XOR evaluación de la neurona

Para la capa oculta:

$$z_j = g \left(\sum_i w_{ij} a_i \right) \quad h_j = \frac{1}{1 + e^{-z}}$$

En el ejemplo:

$$h_0 \equiv 1$$

$$h_1 = g(-0.5a_0 + 1a_1 + 1a_2)$$

$$h_2 = g(1.5a_0 + -1a_1 + -1a_2)$$

Para la capa de salida:

$$z_o = g \left(\sum_j w_{jo} h_j \right) \quad o = \frac{1}{1 + e^{-z_o}}$$
$$o = g(-1.5h_0 + 1w_{1o} + 1w_{2o})$$

XOR evaluación de la neurona (ejemplo convención 1)

$$\mathbf{A} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{W}^{(1)} = \begin{bmatrix} w_{01} & w_{11} & w_{21} \\ w_{02} & w_{12} & w_{22} \end{bmatrix} = \begin{bmatrix} -0.5 & 1 & 1 \\ 1.5 & -1 & -1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{H} &= g(\mathbf{W}^{(1)}\mathbf{A}) \\ &= \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = g \left(\begin{bmatrix} -0.5 & 1 & 1 \\ 1.5 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) = g \left(\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

XOR evaluación de la neurona (ejemplo convención 2)

$$A = [a_0 \quad a_1 \quad a_2] = [1 \quad 0 \quad 1]$$

$$W^{(1)} = \begin{bmatrix} w_{01} & w_{02} \\ w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} -0.5 & 1.5 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}$$

$$H = g(AW^{(1)})$$

$$[h_1 \quad h_2] = g \left([1 \quad 0 \quad 1] \begin{bmatrix} -0.5 & 1.5 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \right) = g([0.5 \quad 0.5]) = [1 \quad 1]$$

$$H' = [h_0 \ h_1 \ h_2] = [1 \ 1 \ 1] \quad W^{(2)} = \begin{bmatrix} w_{0o} \\ w_{1o} \\ w_{2o} \end{bmatrix} = \begin{bmatrix} -1.5 \\ 1 \\ 1 \end{bmatrix}$$

$$O = [o] = g \left([1 \ 1 \ 1] \begin{bmatrix} -1.5 \\ 1 \\ 1 \end{bmatrix} \right) = g([0.5]) = [1]$$

Evaluando sobre varias entradas

¿Qué pasa ahora si queremos evaluar la red sobre varias entradas?

$$X = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{con sesgos} \rightarrow \quad X' = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Veamos sólo la primera capa:

$$H = g(W^{(1)}X'^T) = g\left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -0.5 & 1.5 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}\right) = g\left(\begin{bmatrix} -0.5 & 1.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \\ 1.5 & -0.5 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Entonces, H^T contiene las activaciones para todas las entradas.

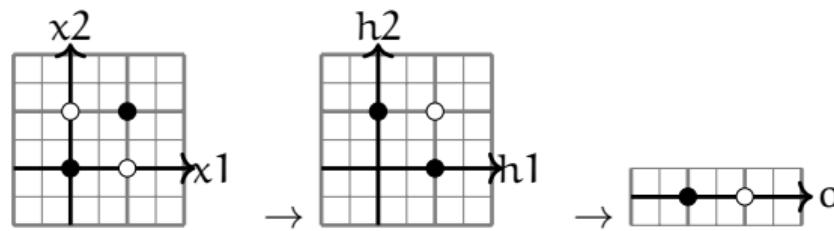
Neuronas como mapeos entre espacios

Observemos la tabla de entradas para la compuerta XOR.

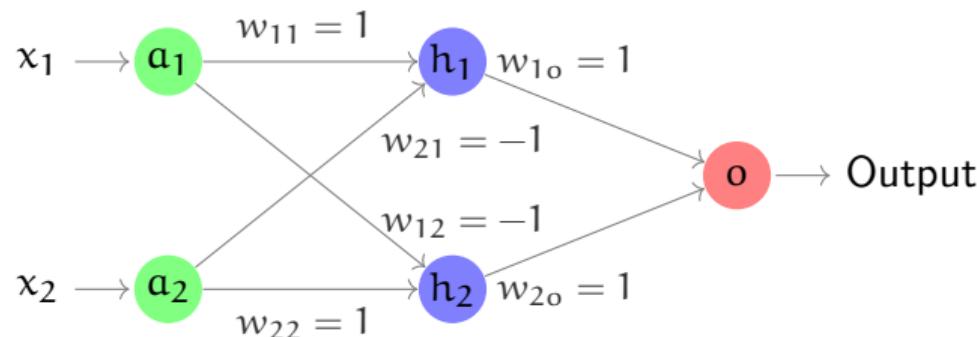
x_1	x_2	h_1	h_2	o
0	0	0	1	0
0	1	1	1	1
1	0	1	1	1
1	1	1	0	0

Podemos interpretar cada capa de la neurona como una función que transforma espacios de varias dimensiones. En este caso:

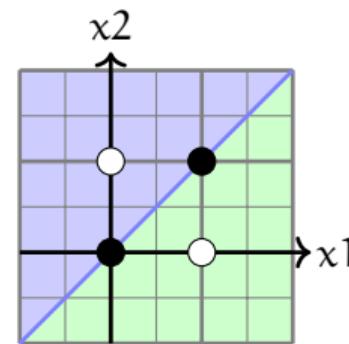
$$\mathbb{R}^2 \rightarrow \mathbb{R}^2 \rightarrow \mathbb{R} \quad (2)$$



XOR (Utilizando la función umbral)



x_1	x_2	z_1	h_1	z_2	h_2	z	o
0	0	0	0	0	0	0	0
0	1	-1	0	1	1	1	1
1	0	1	1	-1	0	1	1
1	1	0	0	0	0	0	0

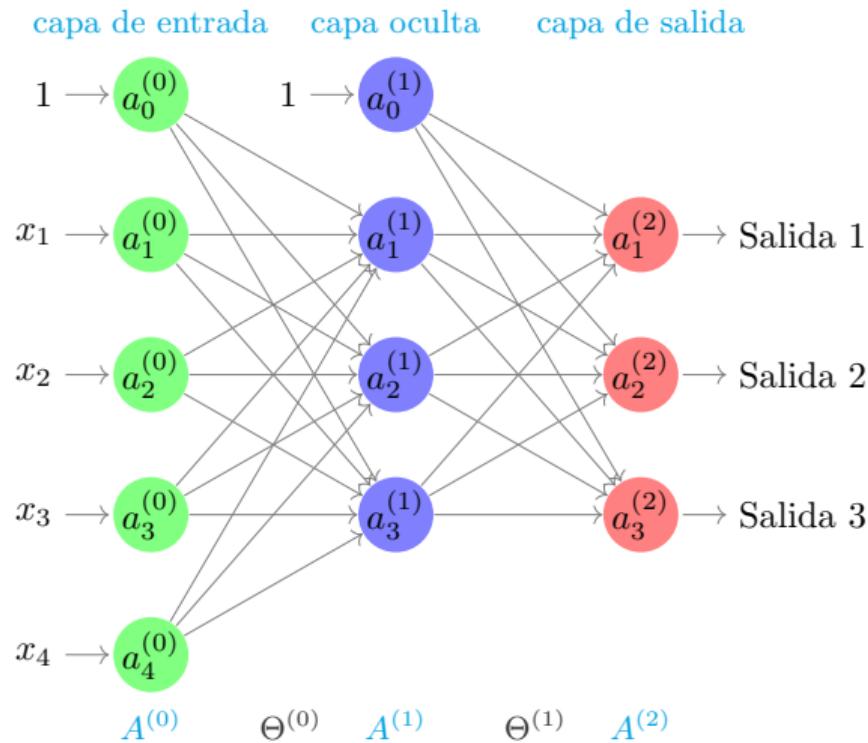


Modelo computacional de una red neuronal

1 Funciones no separables linealmente

2 Modelo computacional de una red neuronal

Red neuronal



One-hot encoding

Por ejemplo:

$$\text{Salida 1} = \text{coche} = [1 \ 0 \ 0] = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (3)$$

$$\text{Salida 2} = \text{casa} = [0 \ 1 \ 0] = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (4)$$

$$\text{Salida 3} = \text{vac\u00e1} = [0 \ 0 \ 1] = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (5)$$

Valor de una neurona (propagación hacia adelante)

$$a_j^{(l+1)} = g \left(\sum_i w_{ij}^{(l)} a_i^{(l)} \right) \quad (6)$$

$$g = \frac{1}{1 + e^{-x}} \quad (7)$$

(8)

Para una sola salida:

$$\begin{bmatrix} a_0^{(l+1)} \\ a_1^{(l+1)} \\ \dots \\ a_{n'}^{(l+1)} \end{bmatrix} = g \left(\begin{bmatrix} w_{01}^{(l)} & \dots & w_{n1}^{(l)} \\ \dots & & \dots \\ w_{0n'}^{(l)} & \dots & w_{nn'}^{(l)} \end{bmatrix} \begin{bmatrix} a_0^{(l)} \\ a_1^{(l)} \\ \dots \\ a_n^{(l)} \end{bmatrix} \right) \quad (9)$$

Para varios ejemplares en paralelo:

$$A^{(l+1)} = g(A^{(l)}W) \quad (10)$$

$$\begin{bmatrix} a_0^{(l+1)} & a_1^{(l+1)} & \dots & a_{n'}^{(l+1)} \end{bmatrix} = g \left(\begin{bmatrix} a_0^{(l)} & a_1^{(l)} & \dots & a_n^{(l)} \end{bmatrix} \begin{bmatrix} w_{01}^{(l)} & \dots & w_{0n'}^{(l)} \\ \vdots & & \vdots \\ w_{n1}^{(l)} & \dots & w_{nn'}^{(l)} \end{bmatrix} \right) \quad (11)$$

Teorema

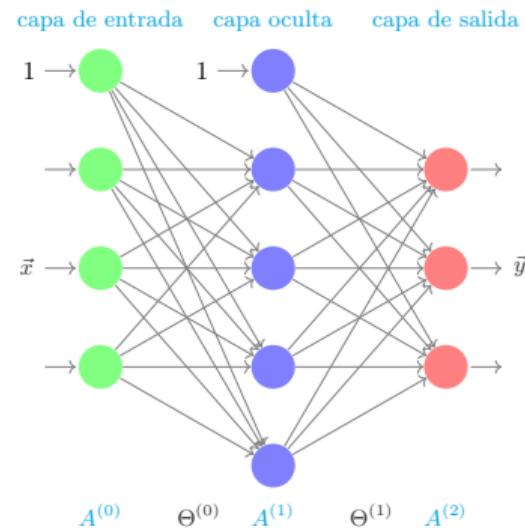
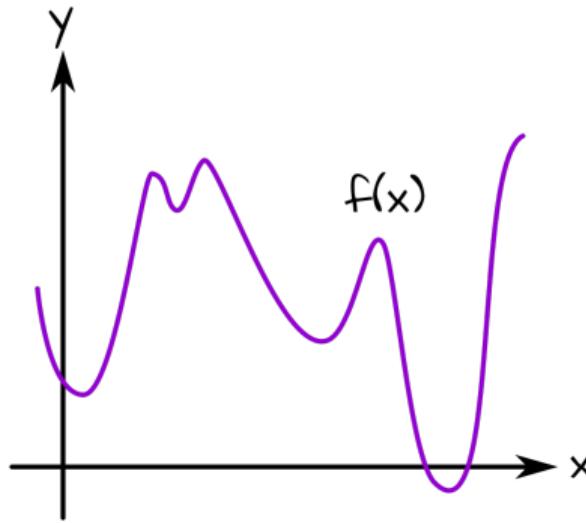
1 Teorema

2 Elementos

Aproximador universal

Definición

Dada una función continua cualquiera $f(\vec{x})$, existe una red neuronal de tres capas tal que para cada entrada x la salida es $\hat{f}(\vec{x})$ con $|\hat{f}(\vec{x}) - f(x)| < \varepsilon$



Funciones $\mathbb{R}^{s_0} \rightarrow \mathbb{R}^{s_L}$

- El número de componentes s_0 en el vector de **entrada** \vec{x} es la dimensión de los argumentos de la función.
- El número de componentes s_L en el vector de **salida** \vec{y} es la dimensión de la salida de la función.

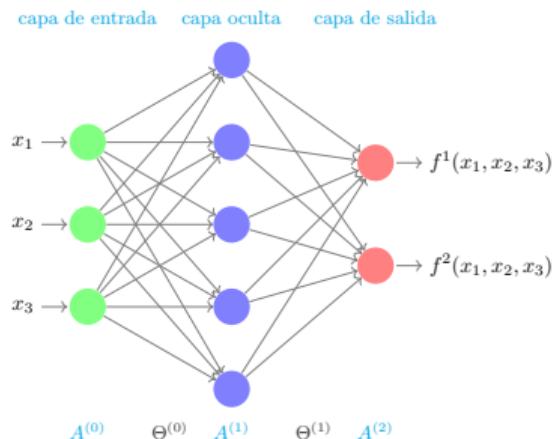


Figura: Ejemplo con $s_0 = 3$ y $s_L = 2$, sin sesgos

Observaciones

- Para incrementar la precisión de la aproximación se requiere incrementar el número de neuronas s_1 de la capa oculta.

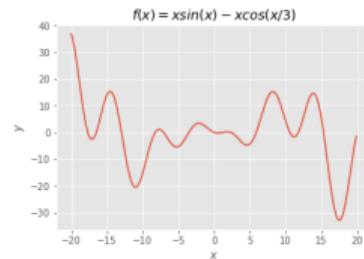


Figura: Función analítica

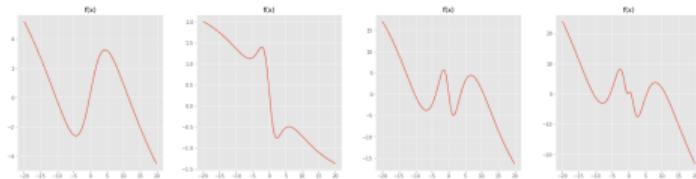


Figura: Aproximaciones en entrenamiento para 2, 5, 50 y 500 neuronas en la capa oculta.

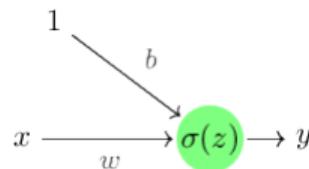
Elementos

1 Teorema

2 Elementos

1 perceptrón

- Observemos el comportamiento de un solo perceptrón con las funciones:



$$\sigma(z) = \frac{1}{1 + e^{-z}}$$
$$z = wx + b$$

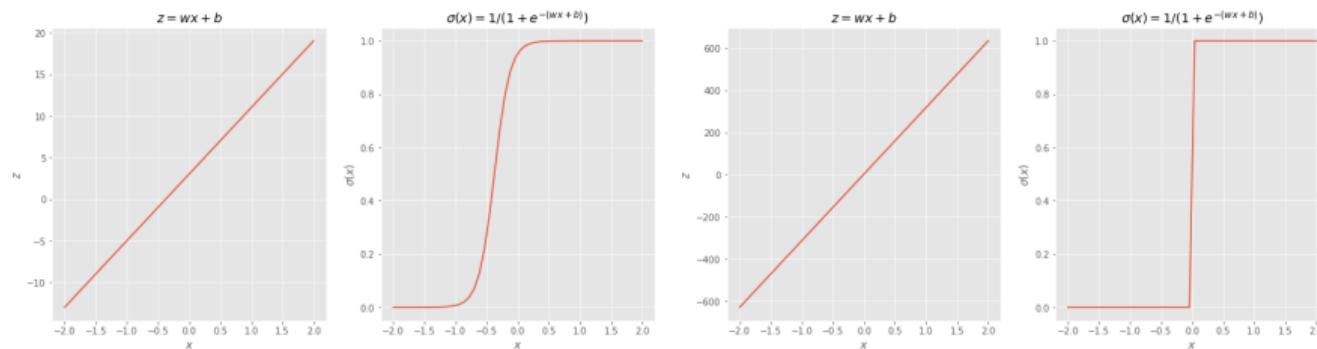


Figura: Izquierda: $w = 8, b = 3$ Derecha: $w = 316, b = 3$

- El valor del sesgo b indica la posición de la frontera entre $f(x) \sim 0$ y $f(x) \sim 1$.
- Conforme w se vuelve muy grande (> 500), la frontera tiende a ser vertical, para valores más pequeños la curva se vuelve más ancha.
- *La posición de la frontera es **directamente proporcional** a b e **inversamente proporcional** a w .* Nielsen 2019

$$s = -\frac{b}{w}$$

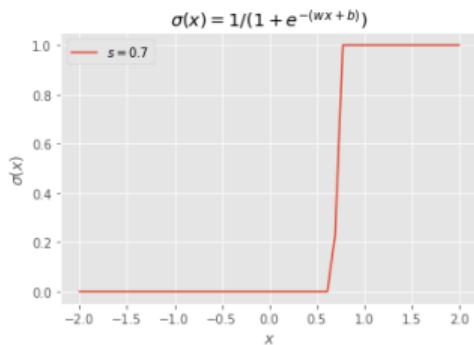


Figura: Frontera fija con $w = 200, s = 0.7$

2 perceptrones

- Utilizando aproximaciones a funciones escalón en una red con dos perceptrones.

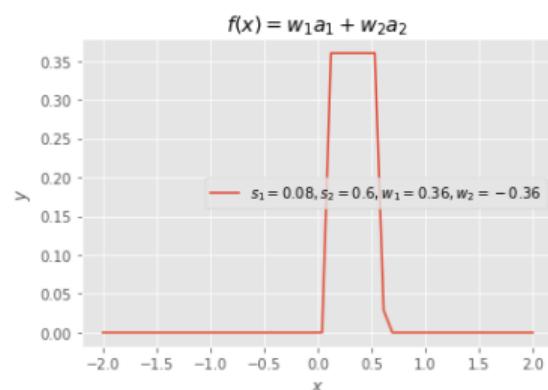
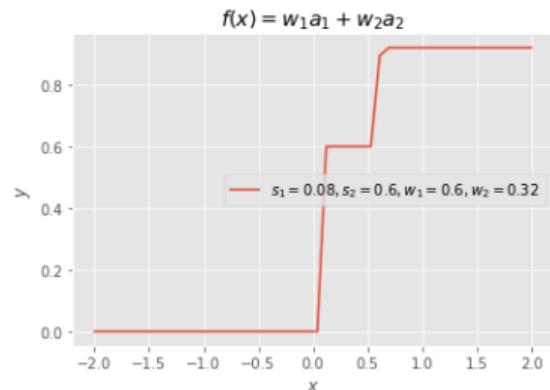
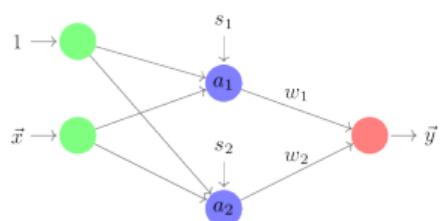


Figura: **Centro:** Los escalones de ambos perceptrones se suman. **Derecha:** Si las contribuciones de ambos son iguales y opuestas $w_1 = -w_2$, obtenemos un bloque cuyo ancho es $s_2 - s_1$ y su altura es $h = w$.

Aproximando la función

- Se pueden utilizar entonces pares de perceptrones, diseñados intencionalmente para aproximar cualquier función utilizando secuencias de barras hasta la precisión deseada, dada por el ancho de las barras.
- Según la figura siguiente ¿cuántos pares más y en qué posiciones podrías ponerlos para aproximar mejor la función?

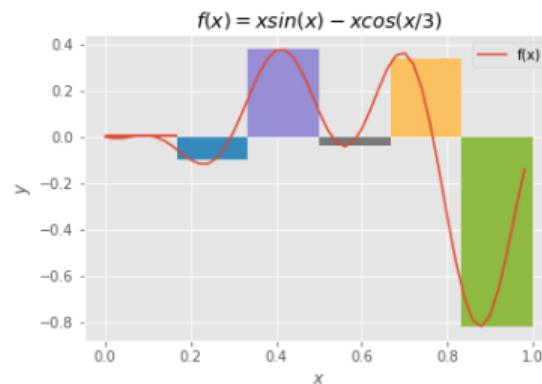


Figura: 6 pares de neuronas equidistantes.

- Un argumento semejante se puede utilizar para construir aproximaciones mediante bloques en más dimensiones.

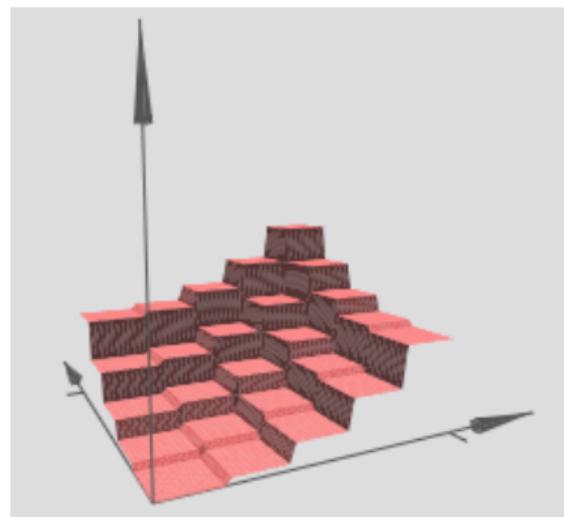


Figura: Construcción para funciones $\mathbb{R}^2 \rightarrow \mathbb{R}$. Tomado de Nielsen 2019