

# Aprendizaje por refuerzo y no supervisado en redes neuronales

Luis Ballado

CINVESTAV - UNIDAD TAMAULIPAS

*[luis.ballado@cinvestav.mx](mailto:luis.ballado@cinvestav.mx)*

February 22, 2023

# Contenido

## ① Introducción

## ② Paradigmas de Aprendizaje

Aprendizaje supervisado

Aprendizaje no supervisado ó "clustering"

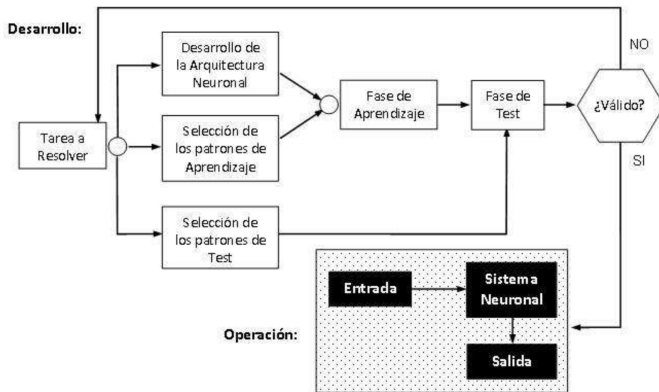
Aprendizaje por refuerzo

# Introducción

Desde la primera mitad del siglo XX se han desarrollado modelos computacionales que han intentado emular el comportamiento del cerebro humano. Aunque se han propuesto una gran cantidad de ellos, todos usan una estructura en red en la cual los nodos o neuronas son procesos numéricos que involucran estados de otros nodos según sus uniones. Una clase de estos modelos computacionales son las Redes de Neuronas Arificiales.

Por su marcada habilidad para obtener resultados de datos complicados e imprecisos, pueden utilizarse para extraer patrones y detectar tramas que son muy difíciles de apreciar por humanos u otras técnicas computacionales.

# Esquema básico del trabajo con RNA



*Las redes neuronales son conjunto de elementos de cálculo simples, usualmente adaptativos, interconectados masivamente en paralelo y con una organización jerárquica que le permite interactuar con algún sistema del mismo modo que lo hace el sistema nervioso biológico.*

— T. Kohonen. *Self-organization and associative memory.* Springer Verlag, New York, 1989....

# El inicio - la neurona artificial

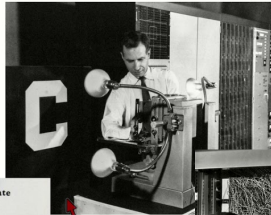
Entre las décadas de 1950 y 1960 el científico Frank Rosenblatt, inspirado en el trabajo de Warren McCulloch y Walter Pitts creó el **Perceptron**, la unidad desde donde nacería y se potenciarían las redes neuronales artificiales.

Un perceptron toma varias entradas binarias  $x_1, x_2$ , etc. para producir una sólo salida binaria. Para calcular la salida, Rosenblatt introduce el concepto de "pesos"  $w_1, w_2$ , etc. que es un número real que expresa la importancia de la respectiva entrada con la respectiva salida. La salida de la neurona será 1 ó 0 si la suma de la multiplicación de pesos por entradas es mayor o menor a un determinado umbral.

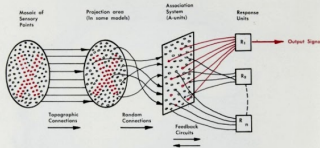
Sus principales usos son decisiones binarias sencillas como compuertas lógicas

# El Perceptrón

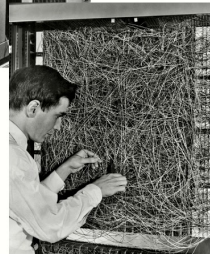
Frank Rosenblatt  
(Cornell Aeronautical Laboratory)  
1957

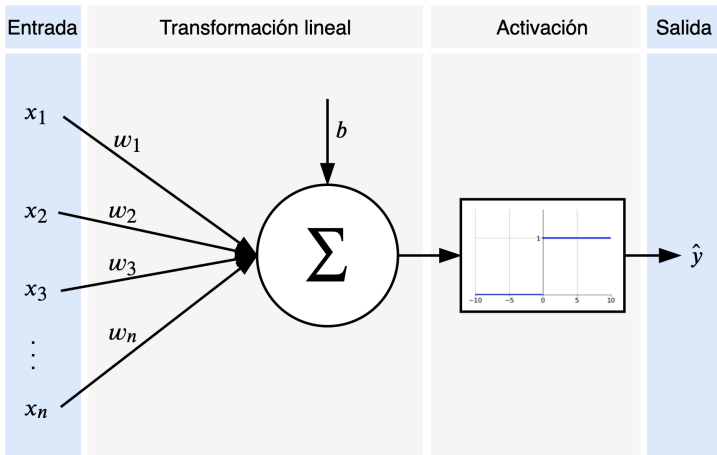


**FIG. 1 — Organization of a biological brain.** (Red areas indicate active cells, responding to the letter X.)



**FIG. 2 — Organization of a perceptron.**



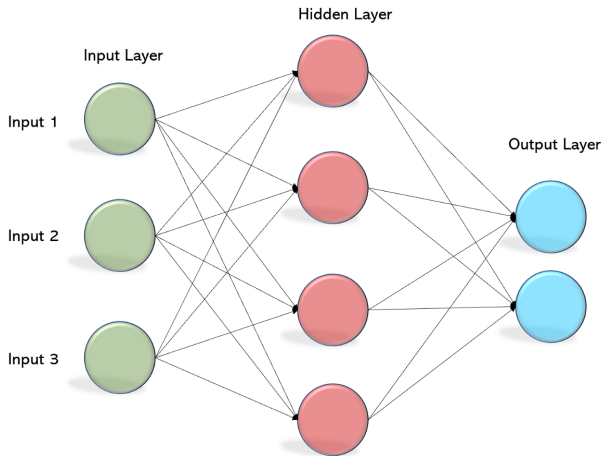




# Multilayer Perceptron - 1965

Es solo una ampliación del modelo que se conoce del perceptron de una sola neurona a más de una. Apareciendo el concepto de capas de entrada, oculta y salida. Pero los valores de entrada/salida binarios.

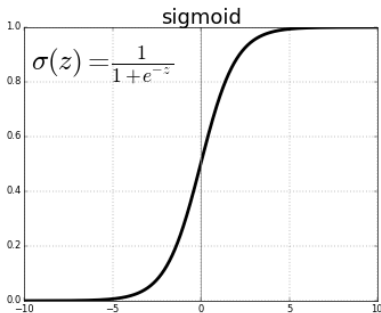
Los valores de los pesos eran asignados manualmente - Entre más perceptrones en las capas, mucho más difícil conseguir los pesos para obtener salidas deseadas.



## Neuronas Sigmoideas - 1980's

Para poder lograr que las redes de neuronas aprendan solas fue necesario introducir un nuevo tipo de neuronas llamadas **Sigmoideas** que son similares al perceptron, pero estas permiten que las entradas puedan tomar valores reales

La función sigmoide definida como  $d(z) = \frac{1}{1+e^{-z}}$  es la primer función de activación



# Backpropagation - 1986

Gracias al algoritmo de backpropagation se hizo posible entrenar redes neuronales de multiples capas de manera supervisada. Al calcular el error obtenido en la entrada e ir propagando hacia las capas anteriores. Se hacen ajustes pequeños (minimizando el costo) en cada iteración para lograr que la red aprenda consiguiendo que la red pueda clasificar las entradas correctamente.

# Convolutional Neural Network - 1989

Son redes multicapa que toman su inspiración del cortex visual de los animales. Es útil en varias aplicaciones, principalmente **procesamiento de imágenes**. La primera CNN fué propuesta por Yann LeCun y estaba enfocada en el reconocimiento de letras manuscritas MNIST <http://yann.lecun.com/exdb/mnist/>

La arquitectura constaba de varias capas que implementaban la extracción de características y luego clasificar.

La imagen se divide en campos receptivos que alimentan una capa convolucional que extrae características de la imagen de entrada. (por ejemplo, detectar lineas verticales, vértices, etc.)

# Long Short Term Memory / Recurrent Neural Network - 1997

Esta arquitectura permite conexiones "hacia atrás" entre las capas. Esto las hace buenas para procesar datos de tipo historicos.

Consisten en unas celdas de memoria que permiten a la red recordar valores por períodos cortos o largos.

Sus aplicaciones se pueden ver en el reconocimiento de voz, de escritura, text to speech por mencionar algunas

# Deep Belief Networks - 2006

Demostraron que utilizar pesos aleatorios al inicializar las redes son una mala idea: por ejemplo al utilizar backpropagation con descenso por gradiente muchas veces se caía en mínimos locales, sin lograr optimizar los pesos. Lo mejor es utilizar una asignación de pesos inteligente mediante un preentrenamiento de las capas de la red "pre-entrenamiento"

Se cree que ese preentrenamiento es una de las causas de la gran mejora en las redes neuronales y permitir el deep learning

# Generative Adversarial Networks - 2014

La idea detrás es de tener dos modelos de redes neuronales compitiendo.

Uno llamado Generador, toma inicialmente "datos basura" como entrada y genera muestras.

El otro modelo, llamado Discriminador que recibe a la vez muestras del Generador y del conjunto de entrenamiento y deberá ser capaz de diferenciar entre las dos fuentes.

Sus aplicaciones principales son la generación de imágenes realistas, pero también la mejora de imágenes ya existentes, o generar textos en imágenes o hasta el desarrollo de moléculas para industria farmacéutica.





# ¿Qué es el aprendizaje?

Norbert Wiener en su libro "Dios y el Golem" define aprendizaje como:

## Definition

Un sistema organizado puede definirse como aquel que transforma un cierto mensaje de entrada en uno de salida, de acuerdo con algún principio de transformación. Si tal principio está sujeto a cierto criterio de validez de funcionamiento, y si el método de transformación se ajusta a fin de que tienda a mejorar el funcionamiento del sistema de acuerdo con ese criterio, se dice que el sistema aprende"



# Aprendizaje supervisado

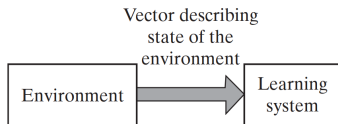
Surge en enseñarle a estos algoritmos cual es el resultado que quieres obtener para un determinado valor tras mostrarles muchos ejemplos si se dan las condiciones, el algoritmo será capaz de dar un resultado correcto incluso cuando le muestres valores que no ha visto antes.

Se denomina "clasificación"

- Los datos deben estar etiquetados, requiere supervisión humana.
- Se usan datos etiquetados para entrenamiento y se predicen individuos fuera de la muestra de entrenamiento

A partir de nuestros datos etiquetados entrenar un algoritmo que aprenda en relación a sus atributos.

# Aprendizaje no supervisado



No existe alguien que enseñe o evalúe el proceso de aprendizaje, la red desarrolla la habilidad de formar representaciones internas para abstraer características de la entrada y crear nuevas clases automáticamente.

- Los datos no están etiquetados
- El algoritmo busca por su cuenta grupos o cluster

Nuestro algoritmo es ciego y sólo conoce los atributos y en base a esas características crear grupos

Se hace uso de reglas de aprendizaje competitivas  
"winner-takes-all"

- Capa de entrada recibe la información disponible
- Capa competitiva consiste de neuronas que compiten entre ellas (en base a la regla de aprendizaje) para la esperanza de responder, la neurona con el mejor acierto gana y se activará, dejando a las demás neuronas apagadas.

# Aprendizaje por refuerzo ¿Fuerza Bruta?

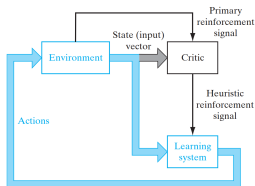
No tenemos una "etiqueta de salida", por lo que no es de tipo supervisado y si bien estos algoritmos aprenden por sí mismos, tampoco son de tipo no supervisado, en donde se intenta clasificar grupos teniendo en cuenta alguna distancia entre muestras.

El Aprendizaje por refuerzo, intentará hacer aprender a la máquina basándose en un esquema de "premios y castigos" - en un entorno en donde hay que tomar acciones y que está afectando por múltiples variables que cambian con el tiempo.

En modelos de Aprendizaje Supervisado (o no supervisado) se intenta minimizar la función de coste,.. es decir reducir el error.

En cambio en Aprendizaje Reforzado intenta "maximizar la recompensa" con la continua interacción con el ambiente para minimizar el index scalar de desempeño.

# Componentes del RL



- Agente - será nuestro modelo que queremos entrenar y que aprenda a tomar decisiones
- Ambiente - será el entorno en donde interactúa y se mueve el agente. Contiene las limitaciones y reglas posibles en cada momento
- Acción - posibles acciones que puede tomar en un momento determinado el agente
- Recompensas (ó castigos) - podremos obtener un premio ó penalización que oriente al agente en si lo está haciendo bien o mal.

# Q-Learning, el algoritmo más usado

El objetivo es entrenar nuestro modelo a través de simulaciones de manera que las decisiones que vaya tomando nuestro agente obtengan "la mayor recompensa"

- Políticas - Tabla que indicará al modelo como actuar en cada estado
- Acciones - las diversas acciones que puede hacer el agente
- Recompensas - si sumamos ó restamos puntaje con la acción tomada
- Comportamiento greedy - Tener grandes recompensas inmediatas o ir explorando y valorando las riquezas a largo plazo