

# *Redes Neuronales: Sistemas complejos y aplicaciones.*

## Resumen

En este proyecto hacemos una revisión multidisciplinar, revisando desde las primeras ideas y los fundamentos teóricos que dieron origen a las redes neuronales hasta las aplicaciones que podemos encontrar en la actualidad.

Comenzamos examinando los sistemas complejos y su relación con la mecánica estadística para entender los comportamientos emergentes en estructuras interconectadas. A continuación, comparamos las redes neuronales artificiales y biológicas, destacando sus similitudes y diferencias, presentamos la Teoría de Grafos como herramienta matemática para modelar redes neuronales, seguida de un análisis de algunos de los tipos de redes más importantes. Por último, se presentan varias aplicaciones de estas redes en campos como la inteligencia artificial y la medicina, lo que demuestra su impacto en distintas disciplinas.

## Índice

<b>1. Introduction</b>	<b>3</b>
<b>2. Redes neuronales biológicas vs redes neuronales artificiales</b>	<b>4</b>
2.1. Funcionamiento de neuronas biológicas . . . . .	4
2.2. Transmisión de señales y mecanismos de aprendizaje . . . . .	4
2.3. Funcionamiento de neuronas artificiales . . . . .	5
2.3.1. Modelo de McCulloch-Pitts . . . . .	5
2.4. Diferencias clave entre redes neuronales biológicas y artificiales . . . . .	6
2.5. Redes Neuronales Multicapa MLP . . . . .	7
2.6. Redes Neuronales Convolucionales CNN . . . . .	7
2.7. Propiedades de una red compleja . . . . .	8
<b>3. Fundamento Teórico</b>	<b>8</b>
3.1. Teoría de Grafos . . . . .	8
3.1.1. ¿Qué es la teoría de grafos? . . . . .	8
3.1.2. Problema principal . . . . .	9
3.1.3. Conceptos clave . . . . .	9
3.1.4. Métricas estadísticas y estructurales . . . . .	10
3.2. Perceptrón . . . . .	10
<b>4. Modelos de redes</b>	<b>11</b>
4.1. Reed feedforward . . . . .	11
4.2. Redes recurrentes . . . . .	11
4.2.1. Red de Hopfield . . . . .	11
4.3. Redes libres de escala (modelo de Barabási–Albert) . . . . .	12
<b>5. Aplicaciones de las redes neuronales</b>	<b>14</b>
5.1. Feedforward . . . . .	14
5.2. Redes Neuronales Recurrentes (RNN) . . . . .	14
5.3. Hopfield . . . . .	14
<b>6. Red neuronal convolucional para la detección y clasificación del cáncer de mama mediante aprendizaje profundo</b>	<b>15</b>
6.1. Datos y procesamiento . . . . .	15
6.2. Arquitectura . . . . .	16
6.3. Comparación con otros modelos . . . . .	16
6.4. Resultados . . . . .	16

**7. Conclusiones****16**

## 1. Introduction

En las últimas décadas, el estudio de los sistemas complejos ha cobrado gran relevancia en múltiples disciplinas científicas, desde la física y la biología, hasta las ciencias sociales, la economía y la informática. Un sistema complejo se define como aquel que está compuesto por un gran número de elementos interactuantes —como partículas, agentes, nodos, neuronas u organismos— cuya dinámica conjunta da lugar a comportamientos emergentes no triviales. Dichos comportamientos no pueden deducirse directamente a partir del análisis de las partes individuales, pues surgen de interacciones no lineales, acopladas y frecuentemente retroalimentadas. Ejemplos paradigmáticos de estos sistemas incluyen redes neuronales, ecosistemas, mercados financieros, redes sociales y genomas.

Para comprender cómo propiedades globales emergen a partir de componentes locales, es fundamental contar con un marco teórico que permita conectar la dinámica microscópica con el comportamiento macroscópico. En este contexto, la mecánica estadística —una rama de la física originalmente desarrollada para explicar fenómenos termodinámicos— ofrece una poderosa caja de herramientas. Mediante conceptos como micro-estados, distribuciones de probabilidad, entropía, correlaciones y transiciones de fase, la mecánica estadística permite modelar sistemas con un número elevado de grados de libertad, capturando no sólo su comportamiento promedio, sino también sus fluctuaciones y estructuras colectivas.

En el caso de sistemas complejos, el uso de la mecánica estadística se extiende más allá de su dominio tradicional. Por ejemplo, en lugar de considerar átomos o moléculas, se modelan agentes sociales, neuronas, empresas o genes, donde cada micro-estado describe una configuración posible del sistema y a cada uno se le asigna una probabilidad  $P(x)$ . Las propiedades globales del sistema, como orden, consenso o diversidad, se obtienen como valores esperados:

$$\langle A \rangle = \sum_x A(x)P(x)$$

[3] Cuando el número de micro-estados es astronómico, como ocurre en la mayoría de los sistemas complejos reales, es necesario recurrir a distribuciones de probabilidad para modelarlos eficientemente. Un ejemplo es la distribución de Boltzmann:

$$P(x) = \frac{1}{Z} e^{-\beta E(x)},$$

donde  $E(x)$  representa una función de energía efectiva que puede codificar, por ejemplo, disconformidad social, distancia entre preferencias o consumo de recursos, y  $Z$  es la función de partición que garantiza la normalización. Esta estructura permite estudiar el equilibrio estadístico del sistema bajo distintas condiciones.

Además, modelos de espines como el de Ising o Potts han sido adaptados para describir decisiones binarias o multivariadas en sistemas sociales o biológicos. La energía del sistema en el modelo de Ising, por ejemplo, se escribe como:

$$E = -J \sum_{\langle i,j \rangle} s_i s_j,$$

donde  $s_i = \pm 1$  representa una elección individual (por ejemplo, una opinión), y  $J$  determina la fuerza de la interacción (como la tendencia al conformismo en redes sociales).

La evolución temporal de estos sistemas se modela mediante técnicas de dinámica estocástica, como caminatas aleatorias, procesos de Markov o algoritmos de Monte Carlo. Por ejemplo, la evolución de una distribución de probabilidad puede describirse con una ecuación de tipo maestro:

$$P(x_{t+1}) = \sum_{x_t} W(x_{t+1} | x_t) P(x_t),$$

donde  $W(x_{t+1}|x_t)$  es la probabilidad de transición entre estados. Estas herramientas permiten simular cómo la información, las opiniones o los recursos se propagan en el sistema.

[4] Las redes neuronales artificiales, al estar formadas por muchas unidades interconectadas que procesan información de forma distribuida, constituyen un paradigma natural para explorar dinámicas colectivas,

autoorganización, y transiciones de fase. Su comportamiento puede describirse formalmente mediante herramientas de la mecánica estadística, una rama de la física que estudia cómo emergen las propiedades macroscópicas de un sistema a partir de sus interacciones microscópicas, utilizando modelos probabilísticos y conceptos como la energía, la entropía, la función de partición o los estados de equilibrio.

En este contexto, las redes neuronales —especialmente aquellas con arquitectura recurrente o estocástica, como las Restricted Boltzmann Machines y los modelos de Hopfield— pueden ser interpretadas como sistemas de espines generalizados, donde los pesos sinápticos juegan el papel de interacciones, y la dinámica de activación neuronal se asemeja al proceso de relajación térmica de un sistema físico hacia el equilibrio. Esta analogía permite analizar el aprendizaje como una minimización de energía libre, describir configuraciones mediante distribuciones de Boltzmann y estudiar la capacidad de memoria o generalización mediante transiciones de fase y fenómenos críticos.

Además, conceptos fundamentales de la mecánica estadística como la entropía de Shannon, la información mutua o los procesos de Markov son esenciales para cuantificar la complejidad, diversidad y flujo de información en redes neuronales. Gracias a esta intersección entre física y aprendizaje automático, ha sido posible desarrollar modelos más interpretables y eficientes, así como aplicar las redes neuronales como herramientas para analizar datos provenientes de otros sistemas complejos: desde patrones de actividad cerebral hasta dinámicas de opinión en redes sociales o propagación de epidemias.

Este reporte se enfoca en explorar cómo la formulación estadística permite comprender y modelar el comportamiento emergente de redes neuronales, tanto como modelos de sistemas complejos como herramientas para estudiarlos, destacando la profunda conexión con la física estadística.

## 2. Redes neuronales biológicas vs redes neuronales artificiales

### 2.1. Funcionamiento de neuronas biológicas

Del marco de trabajo de la termodinámica se ha aprendido a trabajar con modelos empíricos, es decir modelos que responden a la experiencia fenomenológica, desde esta forma de trabajar, se puede intentar replicar la forma en la que naturaleza se comporta en sistemas como el cerebro, lo cual se tratará de hacer utilizando redes neuronales artificiales. La obra maestra de la naturaleza son las redes neuronales biológicas. El cerebro humano está conformado por billones de neuronas, en las que cada una de ellas se conecta a miles de otras, lo que genera una compleja red para manejar información. Una neurona es una célula nerviosa que tiene la función de recibir, procesar y transmitir información por medio de señales eléctricas y químicas, debido a la excitabilidad eléctrica (capacidad de las neuronas de cambiar su potencial eléctrico y transmitir este cambio a través de su axón) de su membrana plasmática.

Una neurona biológica tiene tres partes principales:

1. Dendritas: Reciben las señales de otras neuronas.
2. Cuerpo celular (soma): Procesan las señales entrantes.
3. Axón: Es la parte que conduce impulsos eléctricos desde el cuerpo celular hacia otras neuronas.

### 2.2. Transmisión de señales y mecanismos de aprendizaje

Las neuronas se comunican mediante impulsos conocidos como potenciales de acción. Cuando una señal alcanza la parte final del axón, se liberan neurotransmisores transfiriendo la señal a la siguiente neurona.

Las neuronas fortalecen o debilitan sus conexiones basándose en la experiencia, la plasticidad sináptica (capacidad del cerebro de modificar la fuerza o eficacia de las conexiones sinápticas entre neuronas) es el equivalente biológico al aprendizaje.

Las redes neuronales biológicas son adaptativas, distribuibles

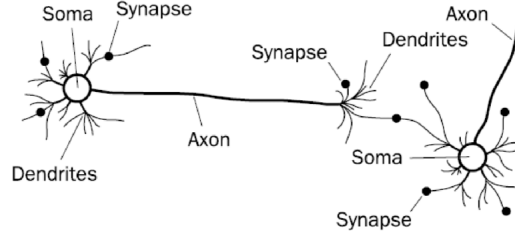


Figura 1: Representación de una neurona

### 2.3. Funcionamiento de neuronas artificiales

Una neurona artificial o perceptrón es la red neuronal más simple de todas ya que sólo contiene una única capa de entrada y un nodo de salida, es un modelo matemático simplificado para tratar de replicar el funcionamiento de una neurona biológica. Sus partes son:

1. Entradas: Reciben señales como lo hacen las dendritas, con la diferencia de que las entradas de las neuronas artificiales son entradas numéricas.
2. Pesos sinápticos: Cada una de las entradas es ponderada, lo cual representa la fuerza de cada conexión.
3. Función de agregación: Las entradas ponderadas se suman.
4. Función de activación: La suma pasa a través de una *función de activación* que determina si la neurona da un valor de salida.
5. Salida: La salida final que pasa a la siguiente capa de neuronas.

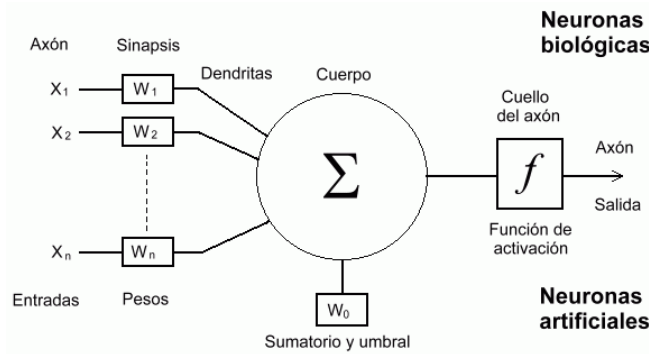


Figura 2: Neurona artificial

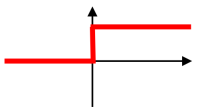
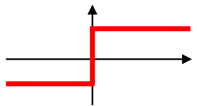
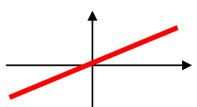
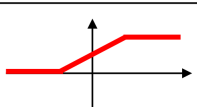
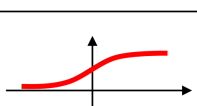
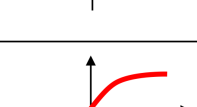
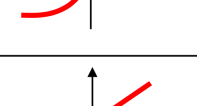

En la fig. 2 se puede ver en la parte inferior las partes que conforman la neurona artificial y en la parte superior a que corresponde esa parte en una neurona biológica.

#### 2.3.1. Modelo de McCulloch-Pitts

El modelo más simple para una neurona artificial es el modelo de McCulloch-Pitts, modelo que se puede observar en la fig. 2, en este modelo las entradas son el estímulo que la neurona artificial recibe del entorno que le rodea, la salida es la respuesta a ese estímulo. En este modelo la neurona se adapta al medio y aprende de él modificando el valor de los pesos sinápticos, por lo cual los mismos serán los parámetros libres que ayudaran a optimizar el aprendizaje de la neurona. La salida neuronal está dada por

$$Y = f\left(\omega_0 + \sum_{i=1}^n \omega_i x_i\right) \quad (1)$$

donde  $\omega_0$  es el umbral,  $\omega_i x_i$  son las entradas  $x_i$  siendo ponderadas por el peso  $\omega_i$ . La función del umbral es la de actuar como un valor constante que permite que la neurona se active o "dispare" incluso si la suma ponderada de las entradas es igual a cero. La función de activación suele elegirse de acuerdo a la tarea realizada por la neurona. En la siguiente tabla se muestran algunas funciones de activación utilizadas para diversos tipos de redes neuronales.

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = \max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^z)$	Multi-layer Neural Networks	

Copyright © Sebastian Raschka 2016  
(<http://sebastianraschka.com>)

Figura 3: Funciones de activación comunes

Como ejemplo se puede decir que una función de activación lineal se utiliza a la salida en redes neuronales que resuelven problemas de regresión. Compuertas del tipo signum o sigmoid se utilizan en redes neuronales que realizan procesos de clasificación binaria.

## 2.4. Diferencias clave entre redes neuronales biológicas y artificiales

Así como una red neuronal biológica es un conjunto de neuronas interconectadas mediante la sinapsis, una red neuronal artificial es un conjunto de neuronas artificiales interconectadas. Por lo que, ahora que ya

se conocen las partes que conforman una neurona artificial y una neurona biológica, se puede resumir en una tabla las principales diferencias una red neuronal biológica y una artificial.

	Redes Biológicas	Redes Artificiales
Componentes	Neuronas (dentrías, axones, sinapsis)	Nodos, pesos sinápticos
Tipo de Señal	Señales eléctricas y químicas	Datos numéricos
Aprendizaje	Plasticidad sináptica	Optimización del peso
Adaptabilidad	Altamente adaptable y auto organizada	Requiere entrenamiento mediante datos
Rapidez	Lenta pero masivamente paralela	Rápida pero menos flexible
Escala	Billones de neuronas	Pocos millones de nodos

Cuadro 1: Diferencias entre redes biológicas y artificiales

## 2.5. Redes Neuronales Multicapa MLP

Las redes neuronales multicapa contienen más de una capa computacional. El perceptrón contiene una capa de entrada y una capa de salida, de las cuales la capa de salida es la única capa que realiza cálculos. La capa de entrada transmite los datos hacia la capa de salida y todos los cálculos son visibles al usuario. Las redes neuronales multicapa contienen múltiples capas computacionales, capas intermedias adicionales entre la entrada y la salida a las cuales se les conoce como *capas ocultas* debido a que los cálculos realizados en ellas no son visibles al usuario. La arquitectura específica de estas redes es de tipo prealimentado ó *feed-forward*, en dicha arquitectura se asume que todos los nodos en cada capa están conectados a aquellos de la siguiente capa. Por lo tanto, la arquitectura de la red neuronal, estará casi definida, una vez que el número de capas y el número/tipo de nodos en cada capa se haya definido. Al número de unidades que tiene cada capa se le conoce como *dimensionalidad* de la capa.

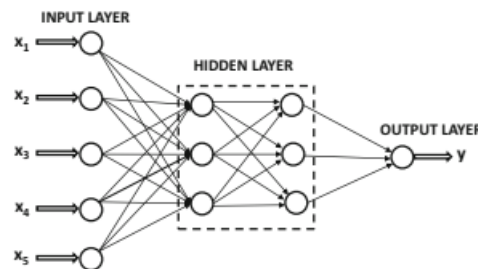


Figura 4: Red Multicapa

Se puede entender por lo tanto que la capa de entrada recibe los datos, la capa oculta procesa los datos y la capa de salida da la respuesta de la red. Este tipo de redes se ponen a prueba con datos estandarizados para ver la capacidad que tiene la red para reproducirlos, optimizando el proceso hasta lograr el resultado esperado, a esto se le conoce como *entrenamiento* de la red.

## 2.6. Redes Neuronales Convolucionales CNN

Las redes neuronales convolucionales son redes inspiradas en redes biológicas y que se usan principalmente en visión computarizada para la clasificación y detección de objetos. La motivación básica para este tipo de redes se obtuvo del entendimiento que Hubel y Weisel lograron acerca de la corteza visual en los gatos, en la cual porciones específicas del campo visual excitan ciertas neuronas.

Para las redes convolucionales, se define una *operación de convolución*, en la cual se usa un filtro para mapear las activaciones de una capa a la siguiente. Una operación de convolución utiliza un filtro tridimensional de pesos con la misma profundidad que la capa en la que se encuentra pero con una medida espacial menor. El producto punto entre todos los pesos en el filtro y cualquier elección de región espacial (del mismo

tamaño del filtro) en una capa define el valor del estado oculto en la siguiente capa. La operación entre el filtro y la región espacial en la capa se realiza en cada posición posible de modo que se pueda definir la siguiente capa.

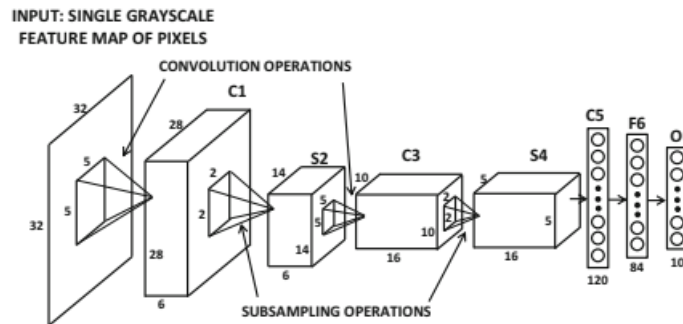


Figura 5: Una de las primeras redes convolucionales

Las redes neuronales convolucionales han sido las redes más exitosas de todos los tipos de redes neuronales. Se utilizan ampliamente para reconocimiento de imágenes, detección de objetos e incluso en procesamiento de texto. El desempeño de estas redes ha superado recientemente a los humanos en la realización de tareas de clasificación de imágenes, por esto pueden llegar a ser bastante útiles como herramienta para el diagnóstico de enfermedades basados en estudios de imagenología clínica.

## 2.7. Propiedades de una red compleja

Es importante conocer e identificar las propiedades que tiene una red compleja como lo son las redes neuronales, desde este punto de vista, la teoría de grafos es una herramienta bastante poderosa para describir y cuantificar las propiedades de redes complejas. Propiedades como la densidad de nodos o el coeficiente de agrupamiento se utilizan en áreas como la neurociencia para relacionar la estructura cerebral con el comportamiento. Por lo tanto en la siguiente sección se iniciará un estudio de las bases teóricas de la teoría de grafos.

## 3. Fundamento Teórico

Se presenta el marco teórico del que se basan las Redes neuronales profundas, se bordan las teorías principales para el desarrollo de estas como lo son la Teoría de Grafos, la Teoría de Redes, algunas estructuras importantes para el modelado de este tipo de redes.

### 3.1. Teoría de Grafos

#### 3.1.1. ¿Qué es la teoría de grafos?

La teoría de grafos es una rama de las matemáticas que estudia las propiedades de sistemas complejos modelados mediante estructuras matemáticas llamadas **grafos**. Estas estructuras se utilizan para representar relaciones entre objetos [2].

Un grafo está compuesto por:

- Nodos: representan los objetos individuales.
- Enlaces: representan relaciones o conexiones entre los objetos.

Esta teoría tiene aplicaciones fundamentales en informática (redes de computadoras), biología (como redes metabólicas o neuronales), y especialmente en inteligencia artificial, se utiliza para modelar redes neuronales.



**Origen histórico.** La teoría de grafos se originó en 1736 con el trabajo de Leonard Euler sobre el problema de los puentes de Königsberg. Este fue el primer problema formalizado usando un grafo y marcó el inicio de esta área.

### 3.1.2. Problema principal

El objetivo principal de la teoría de grafos es representar, analizar y resolver relaciones entre pares de elementos. Algunos problemas clásicos incluyen:

- **Camino más corto:** ¿Cuál es la ruta más eficiente entre dos nodos?
- **Conectividad:** ¿Qué nodos están conectados entre sí directa o indirectamente?
- **Flujo en redes:** ¿Cuál es la máxima cantidad de flujo (información, energía, etc.) que se puede transmitir de un nodo a otro?
  - Se modela asignando capacidades a los enlaces y utilizando algoritmos como Ford-Fulkerson.
- **Detección de ciclos y árboles:**
  - **Ciclo:** Secuencia de enlaces que comienza y termina en el mismo nodo.
  - **Árbol:** Grafo conexo sin ciclos, que conecta todos los nodos con el mínimo número de enlaces ( $n - 1$  para  $n$  nodos).

### 3.1.3. Conceptos clave

Concepto	Descripción
<b>Grafo dirigido</b>	Los enlaces tienen una dirección (de un nodo a otro).
<b>Grafo no dirigido</b>	Los enlaces no tienen dirección, la relación es mutua.
<b>Camino</b>	Secuencia de nodos conectados por enlaces.
<b>Ciclo</b>	Camino cerrado donde el nodo inicial y final coinciden.
<b>Grado (degree)</b>	Número de enlaces conectados a un nodo (grado de entrada y salida si es dirigido).
<b>Conectividad</b>	Existencia de caminos entre nodos en el grafo.
<b>Subgrafo</b>	Cualquier subconjunto de nodos y enlaces que forma un grafo.
<b>Árbol</b>	Grafo conexo sin ciclos. Si tiene $n$ nodos, tiene $n - 1$ enlaces.

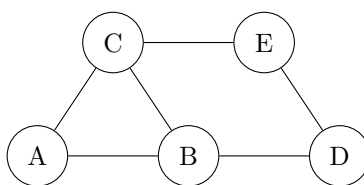


Figura 6: Ejemplo de grafo no dirigido. Contiene un ciclo entre A, B y C.

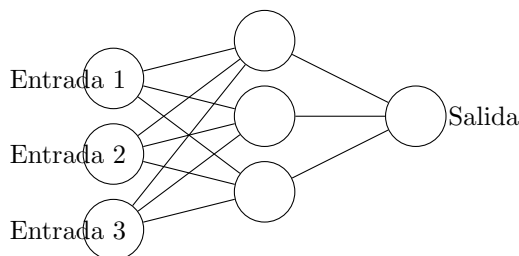


Figura 7: Red neuronal representada como grafo dirigido.

### 3.1.4. Métricas estadísticas y estructurales

Estas métricas permiten analizar cuantitativamente las propiedades del grafo. Son esenciales para entender su estructura global y la importancia de sus componentes.

Métrica	Significado
<b>Grado medio</b> ( $\langle k \rangle$ )	Promedio de conexiones por nodo.
<b>Distribución de grados</b>	Probabilidad de que un nodo tenga cierto número de conexiones.
<b>Centralidad</b>	Mide la importancia de un nodo. Tipos: <ul style="list-style-type: none"> <li>▪ <b>Degree centrality:</b> número de conexiones.</li> <li>▪ <b>Closeness:</b> inverso de la distancia promedio al resto.</li> <li>▪ <b>Betweenness:</b> cuántos caminos más cortos pasan por el nodo.</li> </ul>
<b>Diámetro</b>	Máxima distancia (camino más corto) entre dos nodos cualesquiera.
<b>Longitud promedio del camino</b>	Promedio de distancias más cortas entre todos los pares de nodos.
<b>Densidad</b>	Relación entre el número de enlaces existentes y el número máximo posible.
<b>Coefficiente de agrupamiento</b>	Mide cuán conectados están los vecinos de un nodo (transitividad local).

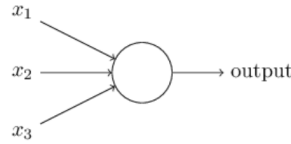
## 3.2. Perceptrón

Un perceptrón es un tipo de neurona artificial que toma varias entradas binarias,  $x_1, x_2, \dots$ , y produce una única salida binaria:

$$\text{salida} = \begin{cases} 0 & \text{si } \sum w_i x_i \leq \text{umbral} \\ 1 & \text{si } \sum w_i x_i \geq \text{umbral} \end{cases} \quad (2)$$

donde se introducen los pesos,  $w_1, w_2, \dots$  números reales que expresan la importancia de las respectivas entradas para la salida. La salida de la neurona (0 o 1), se determina en función de si la suma ponderada  $\sum_j w_j x_j$  es menor o mayor que un valor umbral. Al igual que las ponderaciones, el umbral es un número real que constituye un parámetro de la neurona.

La representación gráfica de un perceptrón es



Simplificando la descripción de los perceptrones. La condición  $\sum_j w_j x_j > \text{umbral}$  es compleja, y podemos hacer dos cambios de notación para simplificarla. El primer cambio consiste en escribir  $\sum_j w_j x_j$  como un producto escalar,  $w \cdot x = \sum_j w_j x_j$ , donde  $w$  y  $x$  son vectores cuyos componentes son los pesos y las entradas, respectivamente. El segundo cambio consiste en desplazar el umbral al otro lado de la desigualdad y reemplazarlo por lo que se conoce como el sesgo (bias) del perceptrón,  $b = -\text{umbral}$ . Utilizando el sesgo en lugar del umbral, la regla del perceptrón se puede reescribir:

$$\text{salida} = \begin{cases} 0 & \text{si } w \cdot x + b \leq 0 \\ 1 & \text{si } w \cdot x + b \geq 0 \end{cases} \quad (3)$$

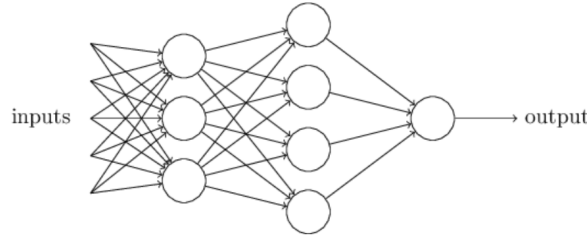
Los perceptrones se pueden usar para calcular las funciones lógicas elementales que solemos considerar como computación subyacente, como AND, OR y NAND. Por ejemplo, supongamos que tenemos un perceptrón con dos entradas, cada una con un peso de  $-2$  y un sesgo general de  $3$ , entonces vemos que la entrada

00 produce la salida 1, ya que  $(-2) \times 0 + (-2) \times 0 + 3 = 3$  es positivo. Cálculos similares muestran que las entradas 01 y 10 producen la salida 1. Pero la entrada 11 produce la salida 0, ya que  $(-2) \times 1 + (-2) \times 1 + 3 = -1$  es negativo. Por lo tanto, nuestro perceptrón implementa una puerta NAND.

## 4. Modelos de redes

### 4.1. Reed feedforward

Una red de propagación hacia adelante (también conocida como red neuronal feedforward) es un tipo de red neuronal artificial donde la información fluye solo en una dirección, desde las capas de entrada a las capas de salida, sin conexiones cíclicas. Tomando como ejemplo. Lo que llamaremos la primera capa



de perceptrones, toma tres decisiones muy simples, sopesando la evidencia de entrada. ¿Qué pasa con los perceptrones de la segunda capa? Cada uno de estos perceptrones toma una decisión sopesando los resultados de la primera capa de toma de decisiones. De esta manera, un perceptrón de la segunda capa puede tomar una decisión a un nivel más complejo y abstracto que los perceptrones de la primera capa. Y el perceptrón de la tercera capa puede tomar decisiones aún más complejas. De esta manera, una red de perceptrones de múltiples capas puede participar en una toma de decisiones sofisticada.

El ejemplo de NAND muestra que podemos usar perceptrones para calcular funciones lógicas simples. De hecho, podemos usar redes de perceptrones para calcular cualquier función lógica. Esto se debe a que la puerta NAND es universal para la computación; es decir, podemos construir cualquier computación a partir de puertas NAND.

### 4.2. Redes recurrentes

Las redes neuronales recurrentes (RNNs, por sus siglas en inglés) son un tipo de red neuronal profunda que se diseñaron para procesar datos secuenciales, es decir, datos que tienen un orden o secuencia. A diferencia de las redes neuronales tradicionales (que procesan datos de forma independiente en cada capa), las RNNs tienen una estructura que permite que la información persista a través del tiempo, lo que les permite recordar información pasada y usarla para predecir o clasificar información futura.

#### 4.2.1. Red de Hopfield

Una red de Hopfield es un tipo de red neuronal recurrente que se utiliza para el almacenamiento y recuperación de patrones, funcionando como una forma de memoria asociativa. Fue propuesta por John Hopfield en 1982.

La red de Hopfield es formalmente análoga al modelo de Ising de la física estadística. En el modelo de Ising. Los "espines"  $s_i$  pueden estar en estado  $+1$  o  $-1$ . Hay interacciones entre espines vecinas, caracterizadas por acoplamientos  $J_{ij}$ .

En la red de Hopfield, los pesos  $w_{ij}$  juegan un papel similar a los acoplamientos  $J_{ij}$ , y el estado de la red corresponde a una configuración de espines. La analogía es tan directa que se puede definir una función de energía donde  $\theta_i$  es el umbral (bias) de activación de la neurona  $i$ .

$$E(\mathbf{s}) = -\frac{1}{2} \sum_{i,j=1}^N w_{ij} s_i s_j + \sum_{i=1}^N \theta_i s_i \quad (4)$$

La red de Hopfield funciona como un sistema de **memoria asociativa**, capaz de recuperar patrones completos a partir de versiones incompletas o ruidosas. La decisión que toma la red corresponde al patrón almacenado más cercano (en términos de similitud) al estado inicial que se le proporciona.

En el estado inicial se proporciona un patrón de entrada  $\mathbf{s}^{(0)}$  que puede estar incompleto o contaminado por ruido cada neurona  $i$  calcula su entrada neta

$$h_i = \sum_{j=1}^N w_{ij} s_j - \theta_i$$

y actualiza su estado mediante la regla:

$$s_i(t+1) = \text{sgn}(h_i)$$

La actualización puede realizarse: Tras varias actualizaciones, el sistema alcanza un estado estable (mínimo de energía). Ese estado final es la “decisión” de la red: el patrón memorizado que mejor coincide con la entrada inicial.

Existen diversas reglas de aprendizaje que pueden utilizarse para almacenar información en la memoria de la red de Hopfield. Es deseable que una regla de aprendizaje tenga las dos propiedades siguientes

- **Local** : una regla de aprendizaje es local si cada peso se actualiza utilizando la información disponible para las neuronas en cada lado de la conexión asociada con ese peso en particular.
- **Incremental** : Se pueden aprender nuevos patrones sin usar información de los patrones antiguos que también se usaron para el entrenamiento. Es decir, cuando se usa un nuevo patrón para el entrenamiento, los nuevos valores de los pesos dependen únicamente de los valores antiguos y del nuevo patrón.

### 4.3. Redes libres de escala (modelo de Barabási–Albert)

Las redes libres de escala son un tipo de red compleja caracterizada porque la distribución de grados de sus nodos sigue una ley de potencias, es decir, la probabilidad de que un nodo tenga  $k$  conexiones es proporcional a  $k^{-\gamma}$ , donde  $\gamma$  suele estar entre 2 y 3:

$$P(k) \sim k^{-\gamma}$$

Esto implica que la mayoría de los nodos tienen pocos enlaces, pero unos pocos (*hubs*) tienen muchos. Esta característica contrasta con redes aleatorias como las de Erdős–Rényi, donde la distribución de grados es de tipo Poisson.

#### Modelo de Barabási–Albert

Propuesto en 1999 por Albert-László Barabási y Réka Albert [1], el modelo BA simula el crecimiento de una red con las siguientes reglas:

- **Crecimiento**: La red comienza con un pequeño conjunto de nodos. En cada paso del tiempo, se agrega un nuevo nodo con un número fijo de enlaces ( $m$ ) que se conectan a nodos ya existentes.
- **Conectividad preferencial**: Los nuevos nodos tienen mayor probabilidad de conectarse a nodos que ya poseen un alto número de conexiones. Formalmente, la probabilidad de conectar con un nodo  $i$  es proporcional a su grado  $k_i$ :

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

Este mecanismo genera automáticamente una red con distribución de grados en forma de ley de potencias.

### Ventajas del modelo Barabási–Albert respecto a otros modelos

Comparado con modelos clásicos como el modelo de Erdős–Rényi (red aleatoria) o las redes regulares (como cuadrículas o anillos), el modelo BA presenta varias ventajas al modelar sistemas reales:

- **Realismo estructural:** La ley de potencias en la distribución de grados reproduce fielmente estructuras observadas en redes reales como Internet, redes sociales, redes biológicas y de transporte.
- **Presencia de hubs:** La generación natural de nodos altamente conectados permite modelar nodos críticos en sistemas reales (por ejemplo, aeropuertos grandes, personas influyentes, servidores principales, etc.).
- **Robustez topológica:** Las redes BA son resistentes a fallos aleatorios de nodos, lo que las hace útiles para modelar sistemas tolerantes a fallos (como redes distribuidas o neuronales).
- **Pequeño mundo emergente:** Aunque el modelo no impone explícitamente una estructura de “mundo pequeño”, suele generar redes con baja longitud de camino promedio, lo que favorece una rápida propagación de información.
- **Simplicidad y escalabilidad:** Su implementación es simple y eficiente computacionalmente, lo que lo hace práctico para simular redes grandes.
- **Aplicabilidad en aprendizaje automático:** Las redes con hubs ofrecen estructuras jerárquicas que pueden ser aprovechadas por modelos como GNNs para mejorar la propagación y agregación de información.

### Propiedades adicionales de las redes libres de escala

- **Emergencia de hubs:** Pocos nodos muy conectados, esenciales para la conectividad global.
- **Robustez estructural:** Resistentes a la eliminación aleatoria de nodos, pero vulnerables a ataques dirigidos a hubs.
- **Pequeño diámetro:** La mayoría de los nodos pueden alcanzarse con pocas conexiones.
- **Alta heterogeneidad:** Distribución desigual del grado de los nodos.

### Ejemplos en el mundo real

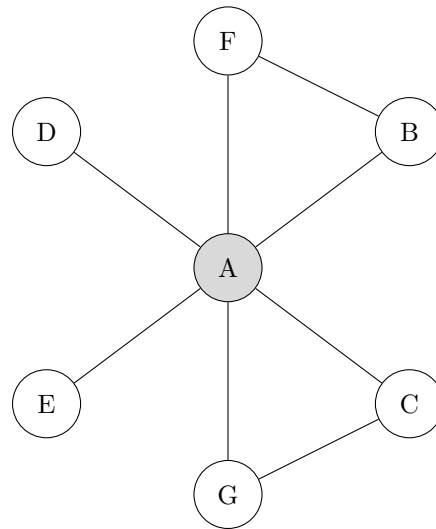
- Internet (a nivel de routers o dominios).
- Redes sociales (algunos usuarios tienen millones de conexiones).
- Redes de transporte aéreo.
- Redes de citas académicas.
- Redes biológicas como redes metabólicas.

### Aplicación a redes neuronales y GNN

En el contexto de redes neuronales, el análisis estructural de la arquitectura puede beneficiarse del modelo de redes libres de escala:

- Las **Graph Neural Networks (GNN)** son capaces de operar sobre grafos arbitrarios, lo que las hace ideales para trabajar con redes complejas como las libres de escala.
- Las GNN pueden explotar la estructura jerárquica de los hubs para mejorar la eficiencia en la propagación de información.
- Redes neuronales profundas presentan propiedades de conectividad similares a las redes libres de escala.

### Visualización del modelo BA



En este grafo, el nodo A es un hub: tiene muchas más conexiones que el resto, lo cual es característico del modelo.

Las redes libres de escala representan una clase fundamental de redes complejas que se encuentran en numerosos sistemas naturales y artificiales. El modelo de Barabási–Albert provee una base simple para modelar este tipo de estructuras. En el análisis de arquitecturas neuronales y el diseño de algoritmos de aprendizaje automático sobre grafos, este modelo resulta esencial.

## 5. Aplicaciones de las redes neuronales

### 5.1. Feedforward

- Estas redes son útiles para tareas en las que hay que aprender una relación entre entradas y salidas, típicamente:
- Clasificación de imágenes: Reconocimiento de productos, detección de defectos en imágenes.
- Predicción de series temporales: Pronósticos de ventas, análisis del mercado bursátil, gestión energética.
- Detección de fraudes: Identificación de transacciones fraudulentas en tarjetas de crédito.

### 5.2. Redes Neuronales Recurrentes (RNN)

Permiten procesar secuencias de datos, ya que la información se retroalimenta a través de bucles en la red.

Se usan comúnmente para problemas ordinarios o temporales como, por ejemplo, la traducción de idiomas, el procesamiento del lenguaje natural (NLP), el reconocimiento de voz y los subtítulos de imágenes; se incorporan a aplicaciones populares como Siri y Google Translate.

Algunas aplicaciones típicas: Traducción automática: Traducción de textos de un idioma a otro.

Reconocimiento de voz: Conversión de voz a texto.

Análisis de sentimientos: Detección de emociones en textos o publicaciones en redes sociales.

### 5.3. Hopfield

La Red de Hopfield es una red recurrente, es decir, existe realimentación entre las neuronas. De esta forma, al introducir un patrón de entrada, la información se propaga hacia adelante y hacia atrás, produciéndose una dinámica. En algún momento, la evolución se detendrá en algún estado estable. En otros casos, es posible que la red no se detenga nunca.

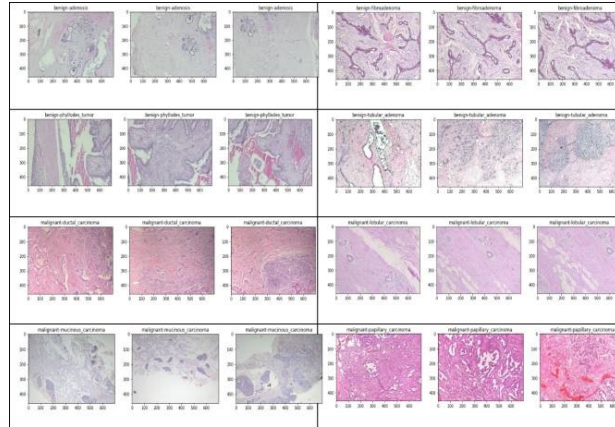


Figura 8:

Las aplicaciones típicas de las Redes Hopfield incluyen el reconocimiento de patrones, la memoria direccionable por contenido y la optimización .

## 6. Red neuronal convolucional para la detección y clasificación del cáncer de mama mediante aprendizaje profundo

El estudio presenta una nueva arquitectura de CNN llamada BCCNN, diseñada para detectar y clasificar imágenes de cáncer de mama en ocho clases distintas (presumiblemente diferentes tipos o estados de tumores). El objetivo es mejorar la precisión y versatilidad en el diagnóstico, y demostrar que una CNN con mecanismos de regulación de aprendizaje puede alcanzar resultados clínicamente útiles.

### 6.1. Datos y procesamiento

Orígenes del dataset: Obtienen imágenes de mamografías a partir de bases públicas en Kaggle (como CBIS-DDSM, MiniDDSM y otras) . Estas bases incluyen imágenes originales junto con etiquetas clínicas (benigno, maligno, subcategorías).

Preprocesamiento: Normalización del tamaño y contraste. Aumento de datos: rotaciones, espejos y escalados para ampliar el set y evitar sobreajuste.

#### Clases de tumores identificadas:

- BA (adenosis benigno)
- BF (fibroadenoma benigno)
- BPT (tumor benigno de phyllodes)
- BTA (adenoma tubular benigno)
- MDC (carcinoma ductal maligno)
- MLC (carcinoma lobular maligno)
- MMC (carcinoma mucinoso maligno)
- MPC (carcinoma papilar maligno)

## 6.2. Arquitectura

- Entrada  
Imagen de RM (resonancia) con distintas magnificaciones (40x–400x).
- Capas Convolucionales Varias capas CNN que extraen características visuales progresivas (texturas, formas).  
No se especifican filtros exactos, pero siguen un patrón típico de extracción jerárquica.
- Pooling  
Capas de max pooling para reducir la información espacial y concentrar lo relevante.
- Flatten  
Convierte los mapas de características en un vector 1D.
- Capas Fully Connected  
Una o más capas densas que combinan las características para clasificación.
- Capa de salida (Softmax)  
8 neuronas (una por clase), con función softmax para obtener la probabilidad de cada tumor.

## 6.3. Comparación con otros modelos

El artículo también evalúa y compara BCCNN con seis modelos pre-entrenados en ImageNet: Xception, InceptionV3, VGG16, MobileNet, ResNet50.

Realizan un total de 30 experimentos, combinando los seis modelos con las cinco magnificaciones de imágenes (40x–400x)

## 6.4. Resultados

F1-score del BCCNN: 98,28 %.

Comparado con los demás:

ResNet50: 98,14 %

VGG16: 97,67 %

Xception: 97,54 %

MobileNet: 93,98 %

Esto muestra que su arquitectura personalizada logró resultados ligeramente mejores que modelos grandes preentrenados .

## 7. Conclusiones



## Referencias

- [1] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439): 509–512, 1999.
- [2] Béla Bollobás. *Modern Graph Theory*. Springer, 1998.
- [3] Kerson Huang. *Statistical Mechanics*. Wiley, 2nd edition, 1987.
- [4] Melanie Mitchell. *Complexity: A Guided Tour*. Oxford University Press, 2009.