

Redes Neuronales Artificiales y Deep Learning

Percepción Multicapa (MLP), Funciones de Activación y Retropropagación

Una exploración completa de los fundamentos del aprendizaje profundo

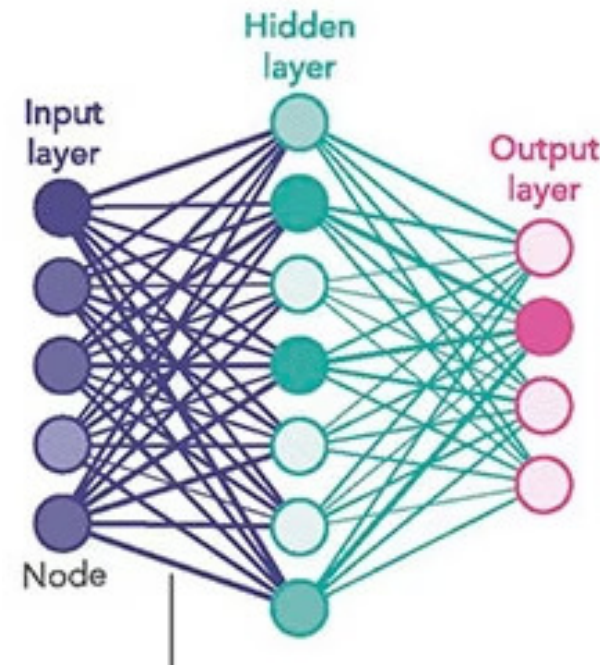
¿Qué es una Red Neuronal Artificial?

Las redes neuronales artificiales están **inspiradas en el cerebro humano**, utilizando neuronas artificiales conectadas en capas que imitan el procesamiento biológico de información.

Estas redes procesan información mediante nodos (neuronas) que reciben señales de entrada, las transforman mediante operaciones matemáticas y transmiten el resultado a las siguientes capas.

Arquitectura básica

- **Capa de entrada:** recibe los datos iniciales
- **Capas ocultas:** procesan y extraen características
- **Capa de salida:** genera la predicción final



Redes Neuronales: Arquitectura y Funcionamiento



Fundamento del Aprendizaje Profundo

Son modelos computacionales inspirados en el cerebro humano, diseñados para reconocer patrones complejos y aprender de grandes volúmenes de datos.



Conexiones y Pesos Sinápticos

Las neuronas artificiales se conectan entre sí mediante 'pesos' que ajustan la fuerza de la señal, imitando las sinapsis biológicas para el aprendizaje.



Propagación de la Información

Los datos fluyen de la capa de entrada a la de salida, transformándose en cada neurona a través de funciones de activación que permiten identificar características complejas.

Funciones de Activación: ¿Por qué son clave?

Las funciones de activación **introducen no linealidad** en la red, permitiéndole aprender patrones complejos y relaciones sofisticadas en los datos.



Sigmoide

$$f(x) = \frac{1}{1+e^{-x}}$$

Rango: 0 a 1

Ideal para clasificación binaria y probabilidades.



ReLU

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

Activa solo valores positivos.



Tanh

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Rango: -1 a 1

Tangente hiperbólica centrada en cero.



Softmax

$$f(x_i) = \frac{e^{x_i}}{\sum e^{x_j}}$$

Usada para clasificación multiclase

¿Cómo aprende una Red Neuronal?



Entrenamiento con datos

La red ajusta sus pesos internos para minimizar el error entre predicciones y valores reales



Proceso iterativo

Múltiples pasadas por los datos mejoran progresivamente la precisión de la red



Retropropagación

Algoritmo clave que distribuye el error hacia atrás para optimizar cada peso

Retropropagación: Ajuste de Pesos Paso a Paso

La **retropropagación** (backpropagation) es el algoritmo fundamental que permite a las redes neuronales aprender de sus errores de manera sistemática.

01

Cálculo del error

Se mide la diferencia entre la salida predicha y el valor real esperado

02

Propagación hacia atrás

El error se propaga desde la capa de salida hacia las capas ocultas anteriores

03

Cálculo de gradientes

Se determina la contribución de cada peso al error total usando derivadas parciales

04

Actualización de pesos

Los pesos se modifican usando gradiente descendente para reducir el error

05

Iteración

El proceso se repite múltiples veces hasta que el error sea mínimo aceptable

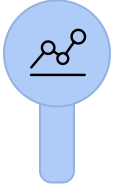
Ejemplo Práctico: Clasificación de Dígitos Manuscritos

Flujo del proceso



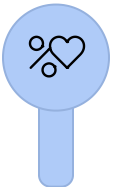
Entrada

Imagen de un dígito de 28×28 píxeles (784 valores numéricos)



Procesamiento MLP

La red procesa la imagen a través de capas ocultas que extraen características



Salida probabilística

Vector de probabilidades para cada dígito (0-9), suma 100%



Clasificación

Reconocimiento automático con precisión superior al 98%

Deep Learning: Redes Neuronales Profundas

El **Deep Learning** representa la evolución del perceptrón multicapa hacia arquitecturas con decenas o incluso cientos de capas ocultas.

Visión por computadora

Reconocimiento facial, detección de objetos, segmentación de imágenes médicas

Procesamiento de lenguaje

Traducción automática, chatbots, análisis de sentimientos, generación de texto

Reconocimiento de voz

Asistentes virtuales, transcripción automática, comandos por voz

Estas redes aprenden **representaciones jerárquicas**: las primeras capas detectan características simples, mientras que las capas profundas identifican conceptos complejos y abstractos.

Retos y Consideraciones

Volumen de datos masivo

Las redes profundas requieren grandes cantidades de datos etiquetados para entrenamiento efectivo. Datos insuficientes pueden llevar a bajo rendimiento.

Costo computacional elevado

El entrenamiento demanda GPUs o TPUs especializadas. Tiempos de entrenamiento pueden extenderse de horas a días o semanas.

Riesgo de sobreajuste

Sin técnicas de regularización (dropout, normalización), la red puede memorizar en lugar de generalizar patrones.

Diseño de arquitectura

Elegir el número de capas, neuronas por capa y funciones de activación adecuadas requiere experiencia y experimentación.

❏ **Nota importante:** El equilibrio entre complejidad del modelo y disponibilidad de recursos es crucial para el éxito del proyecto.