

RESUMEN REDES NEURONALES

EMILIO YAÑEZ OMAÑA

1. La Neurona Artificial

La neurona artificial es la unidad básica de procesamiento en una red neuronal. Recibe valores de entrada, les aplica un peso asociado (W) y genera un valor de salida tras pasar por una función de activación.

- La operación principal es una **suma ponderada**: $Z = X_1W_1 + X_2W_2 + \dots + X_nW_n + b$ donde b es el término de **sesgo (bias)** que permite ajustar la salida.
- La salida pasa por una **función de activación** (e.g., ReLU, sigmoide o tangente hiperbólica), que introduce no linealidades y permite a las redes aprender patrones complejos.

Cuantas más entradas y funciones no lineales se utilicen, mayor es la capacidad de la red para modelar datos complejos.

2. Capas en la Red

Una red neuronal está formada por varias capas de neuronas, que se agrupan en tres tipos:

1. **Capa de entrada:** Recibe los datos originales del problema.
2. **Capas ocultas:** Procesan y transforman la información usando las neuronas y sus funciones de activación.
3. **Capa de salida:** Entrega el resultado final, como una clasificación o una predicción numérica.

El conocimiento de la red crece con el número de capas y neuronas:

- Una red con una sola capa oculta puede aproximar cualquier función matemática.
- Al sumar más capas, las redes se convierten en **profundas (deep neural networks)**, permitiendo modelar relaciones más complejas.

Para evitar que todas las neuronas colapsen en una salida lineal, se utilizan funciones de activación no lineales como ReLU, seno, coseno, entre otras, que deforman los planos y generan representaciones complejas.

3. Aprendizaje en Redes Neuronales

El aprendizaje consiste en ajustar los pesos (WWW) y sesgos (bbb) para minimizar el error entre la predicción de la red y el resultado real.

- **Perceptrón (1986):** Fue uno de los primeros algoritmos de aprendizaje automático, aunque tenía limitaciones por no ajustar automáticamente los parámetros y corregir errores de forma ineficiente. Esto llevó a la falta de inversión en IA durante el "invierno de la IA".
- **Backpropagation (1986):** Marcó un hito al introducir el **descenso del gradiente** para ajustar parámetros de forma iterativa. Este método calcula cómo cada peso y sesgo afecta el error y los actualiza en consecuencia, desde la capa de salida hacia la capa de entrada.

El aprendizaje iterativo permite a las redes adaptarse a datos cada vez más complejos, lo que ha llevado a su auge en aplicaciones como el reconocimiento de voz, imágenes y texto.

4. Funciones de Activación y Planos Complejos

Las funciones de activación transforman los datos lineales de las neuronas en formas más complejas, permitiendo que la red modele superficies multidimensionales. Ejemplos:

- **ReLU (Rectified Linear Unit):** Deja pasar valores positivos y descarta los negativos, ideal para redes profundas.
- **Sigmoide y tangente hiperbólica:** Suavizan los resultados y son útiles en problemas de clasificación.

Cuantas más neuronas y funciones de activación se agregan, más detallados y precisos son los planos que describen los patrones del problema.