

Módulo 3: Introducción a Modelos de Deep Learning

Duración Estimada: 3 Horas

1. Redes Neuronales Profundas: Fundamentos y Funcionamiento

El **Deep Learning (DL)** es un subcampo del Machine Learning que se basa en las **Redes Neuronales Profundas (DNNs)**. Estas redes están inspiradas en la estructura del cerebro humano.

La Neurona Artificial (Perceptrón)

La unidad básica de una red. Recibe múltiples entradas, las procesa (sumándolas con pesos asignados) y aplica una **función de activación** para decidir si activa y pasa la señal a la siguiente neurona.

La Profundidad (Deep)

Una red es considerada "profunda" cuando tiene **múltiples capas ocultas** entre la capa de entrada y la capa de salida.

Capa	Función Principal
Capa de Entrada	Recibe los datos brutos (ej. píxeles de una imagen, palabras de un texto).
Capas Ocultas	Extraen características y patrones complejos de forma jerárquica. Cuantas más capas, más profundo es el aprendizaje.
Capa de Salida	Genera el resultado final (ej. la clase predicha, la siguiente palabra).

Funcionamiento Jerárquico: En Visión por Computadora (DL), las primeras capas aprenden a detectar bordes simples, las capas intermedias detectan formas y texturas, y las capas finales identifican objetos completos (rostros, animales, etc.). **El modelo aprende las características por sí mismo**, sin intervención humana.

2. Entrenamiento de Modelos: Proceso de Aprendizaje

El entrenamiento es el proceso clave donde una red neuronal profunda ajusta sus miles o millones de **pesos** para mejorar su precisión.

El Ciclo de Entrenamiento:

1. **Propagación Hacia Adelante (*Forward Pass*)**: Los datos de entrenamiento pasan por la red desde la capa de entrada hasta la de salida, generando una **predicción**.
2. **Cálculo del Error (*Loss Function*)**: El modelo compara su predicción con la etiqueta real (la "verdadera respuesta") y calcula el **error** o **pérdida**.
3. **Propagación Hacia Atrás (*Backpropagation*)**: El error se propaga **hacia atrás** a través de la red (de la capa de salida a la de entrada). Este es el mecanismo central del DL.
4. **Optimización (*Gradient Descent*)**: El algoritmo **ajusta los pesos** de cada neurona en función de cuánto contribuyeron al error, buscando el camino más eficiente para **minimizar la pérdida**.

Objetivo: Iterar este proceso millones de veces (épocas) hasta que la red encuentre el conjunto de pesos óptimo que minimice el error de forma consistente.

3. Aplicaciones Prácticas: DL en IA Generativa

El Deep Learning es la base tecnológica de las herramientas generativas más avanzadas, ya que son excelentes para manejar y crear datos no estructurados.

A. Modelos de Lenguaje Grande (LLMs) - Texto 🗣️

- **Arquitectura Clave: Transformer** (basada en el mecanismo de **Atención**).
- **Casos de Uso (ChatGPT, Gemini, etc.):**
 - **Traducción y Resumen:** Los modelos comprenden y reestructuran el lenguaje natural a un nivel que el ML tradicional no puede igualar.
 - **Generación de Código:** El modelo "aprende" las reglas y sintaxis del código a partir de miles de repositorios.

- **Dialogo Coherente:** El mecanismo de *atención* permite al modelo mantener la coherencia a lo largo de largas conversaciones, ponderando la importancia de las palabras anteriores.

B. Modelos de Difusión - Imágenes y Multimedia

- **Arquitectura Clave: Modelos de Difusión** (un tipo avanzado de DL).
 - **Casos de Uso (DALL-E, Midjourney, etc.):**
 - **Creación de Imágenes:** Estos modelos aprenden a generar una imagen añadiendo "ruido" aleatorio y luego invirtiendo el proceso para "des-ruidificar" gradualmente hasta formar la imagen solicitada por el usuario (el *prompt*).
 - **Transferencia de Estilo:** Pueden aplicar el estilo artístico de una pintura famosa a una fotografía personal.
-

4. Deep Learning en Creación de Contenido

El DL se destaca por su capacidad de generar resultados que parecen haber sido creados por humanos.

Generación de Texto

- **El Proceso:** Los LLMs (basados en DL) funcionan prediciendo la **palabra más probable** que sigue a una secuencia de palabras dada.
- **Control Creativo:** Los parámetros como la **temperatura** (que influye en la aleatoriedad de la predicción) permiten al usuario ajustar si quiere un texto predecible y preciso (baja temperatura) o un texto creativo y original (alta temperatura).
- **Ejemplo:**
 - *Input:* "El futuro de la IA se basa en la..."

- *LLM*: "... **capacidad** de los modelos para **aprender** de datos **no** estructurados."

Generación de Imágenes 🎨

- **Entrenamiento:** Los modelos de difusión son entrenados con miles de millones de pares de imágenes y sus descripciones de texto asociadas.
 - **Síntesis:** Cuando un usuario introduce un *prompt* (ej. "Un perro astronauta en Marte, estilo óleo"), el modelo utiliza su comprensión de texto (PLN/DL) y su conocimiento visual (Visión por Computadora/DL) para sintetizar una imagen completamente nueva y única que cumple con ambas condiciones.
-

Ejercicio Práctico (30 min)

Instrucción: Utiliza Gemini, ChatGPT-4, o cualquier otra herramienta de IA Generativa con capacidad para crear imágenes o texto.

Opción A: Generación de Imagen (Si la herramienta lo permite)

- **Prompt de Imagen:** "Crea una imagen de un robot de los años 50 estudiando un libro de *Deep Learning* en una biblioteca futurista, con iluminación *neon*."

Opción B: Generación de Texto (Si la herramienta es solo de texto)

- **Prompt de Texto:** "Explica en 150 palabras, usando una metáfora simple (como un chef o un arqueólogo), el proceso de *Backpropagation* en el entrenamiento de una red neuronal profunda."

Objetivo: Experimentar de primera mano la capacidad de las arquitecturas de Deep Learning para interpretar instrucciones complejas y generar resultados creativos y originales.