

## Sesión 4: Generalización, Sobreajuste y Conceptos Modernos en Redes Neuronales

Curso: Fundamentos de Machine Learning con Redes Neuronales

Sesión 4 (1 hora)

# Introducción

## Objetivos de la sesión

- ▶ Comprender qué es generalización y por qué es importante.
- ▶ Diferenciar entre sobreajuste y subajuste.
- ▶ Conocer técnicas de regularización.
- ▶ Explorar la importancia de los datos y su preparación.
- ▶ Introducir arquitecturas modernas (CNN, RNN, Transformers).
- ▶ Integrar todos los conceptos en un flujo completo de ML.

# Generalización

## Concepto

Generalizar significa que el modelo aprende patrones reales del problema y no simplemente memoriza los datos.

## Analogía

- ▶ Estudiante que memoriza → falla en el examen.
- ▶ Estudiante que comprende → aplica el conocimiento en nuevos casos.

# Sobreajuste y Subajuste

## Sobreajuste (Overfitting)

Ocurre cuando el modelo aprende demasiado bien los datos de entrenamiento, incluso el ruido.

### Señales

- ▶ Pérdida baja en entrenamiento.
- ▶ Pérdida alta en validación.

## Subajuste (Underfitting)

El modelo no tiene suficiente capacidad para aprender el patrón.

### Señales

- ▶ Pérdida alta en entrenamiento y validación.

# Técnicas para Evitar Sobreajuste

## Regularización L2

Añade una penalización a pesos grandes para evitar complejidad excesiva.

## Dropout

Desactiva aleatoriamente neuronas durante entrenamiento para evitar dependencia excesiva.

## Early Stopping

Detiene el entrenamiento cuando la pérdida de validación deja de mejorar.

## Aumentación de Datos

Transformaciones en imágenes, texto o audio para generar variaciones artificiales.

# Importancia de los Datos

## Calidad sobre cantidad

Mejores datos → mejores modelos.

## Procesamiento de datos

### Normalización

Escalar valores a un rango común.

### Estandarización

Ajustar datos a media 0 y desviación estándar 1.

### Balanceo

Evitar clases desbalanceadas para mejorar precisión del modelo.

# Batch Processing y Optimizadores Modernos

## Revisión de batch y mini-batch

Permiten un entrenamiento más eficiente y estable.

## Optimizadores avanzados

### Momentum

Acumula velocidad para acelerar el descenso.

### Adam

Adaptativo, eficiente y el más usado en la práctica moderna.

# Arquitecturas Modernas

## Redes Convolucionales (CNN)

- ▶ Usadas en visión computacional.
- ▶ Detectan bordes, texturas, objetos.

## Redes Recurrentes (RNN, LSTM)

- ▶ Manejan secuencias: texto, audio, series temporales.

## Transformers

- ▶ Arquitectura más exitosa actualmente.
- ▶ Usada en modelos como GPT, BERT, Gemini.
- ▶ Manejan dependencias largas mediante atención.

# Flujo Completo de un Proyecto de Machine Learning

1. Recolección de datos
2. Limpieza y procesamiento
3. División en entrenamiento/validación
4. Selección de modelo
5. Entrenamiento (épocas, batch, optimización)
6. Evaluación
7. Regularización y ajuste final
8. Implementación

## Actividad de Cierre

Reflexionar sobre un ejemplo real e identificar:

- ¿Qué datos se necesitan?
- ¿Qué tipo de modelo usar?
- ¿Dónde podría ocurrir sobreajuste?
- ¿Qué arquitectura moderna sería útil?

## Notas del Presentador

- ▶ Enfatizar que las arquitecturas modernas siguen los mismos principios básicos vistos en sesiones anteriores.
- ▶ Explicar la importancia crítica de los datos y su preparación.
- ▶ Relacionar el contenido con aplicaciones como visión, lenguaje y audio.
- ▶ Conectar con el próximo curso: implementación con Python y librerías (TensorFlow/PyTorch).