

# Convolutional Neural Networks (CNNs)

De bloques convolucionales a modelos completos

Cesar Garcia

2025

- En la sesión anterior vimos **qué es** una convolución
- Hoy veremos **cómo se combinan** para resolver un problema real
- Construiremos una CNN completa para **clasificación**
- Compararemos CNN vs MLP

*¿Qué piezas faltan para pasar de filtros a un modelo usable?*

# Arquitectura típica de una CNN

## Bloques

Una CNN suele componerse de:

- Bloques convolucionales:
  - Conv  $\rightarrow$  ReLU  $\rightarrow$  Pool
- Capas finales:
  - Flatten  $\rightarrow$  Linear  $\rightarrow$  Output

La profundidad crea **jerarquía de representaciones**.

*¿Por qué no usamos solo una capa convolucional?*

## Patrón recurrente

Un bloque típico:

- Conv2d (extrae patrones)
- ReLU (no linealidad)
- Pooling (reducción e invariancia)

Se repite varias veces.

*¿Qué efecto tiene apilar varios bloques similares?*

## Flatten

Después de convoluciones:

- tenemos tensores ( $C \times H \times W$ )
- las capas densas esperan vectores

Flatten convierte:

$$(C, H, W) \rightarrow (C \cdot H \cdot W)$$

*¿Qué información espacial se pierde al hacer flatten?*

## Salida

Para clasificación:

- última capa lineal produce logits
- softmax convierte logits en probabilidades

Ejemplo MNIST:

- 10 neuronas de salida
- una por dígito

*¿Por qué usamos softmax solo al final?*

# CNN vs MLP (repaso)

## Diferencias clave

CNN: \* pesos compartidos \* menos parámetros \* inductive bias espacial

MLP: \* conexiones globales \* muchos parámetros \* aprende desde cero la estructura

*¿En qué tipo de datos una CNN NO sería buena idea?*

## Similaridades

Entrenar una CNN es igual que entrenar un MLP:

- forward
- loss
- backward
- optimizer

La diferencia está en la **arquitectura**, no en el algoritmo.

*¿Qué partes del pipeline no cambian entre MLP y CNN?*



## Pitfalls

- CNN demasiado grande → overfitting
- Pooling excesivo → pérdida de detalle
- Muy pocas capas → baja capacidad

Arquitectura es **tradeoffs**.

*¿Cómo detectarías que una CNN es demasiado compleja?*

## Arquitectura importa

Las CNN funcionan porque:

- imponen estructura correcta
- reducen el espacio de hipótesis
- aprenden representaciones jerárquicas

***No solo entrenamos mejor: buscamos el modelo adecuado.***

*¿Qué ganamos al introducir inductive bias?*