



# Aprendizaje Inteligencia Artificial

Cómo enseñamos a las máquinas



# Neuronas artificiales

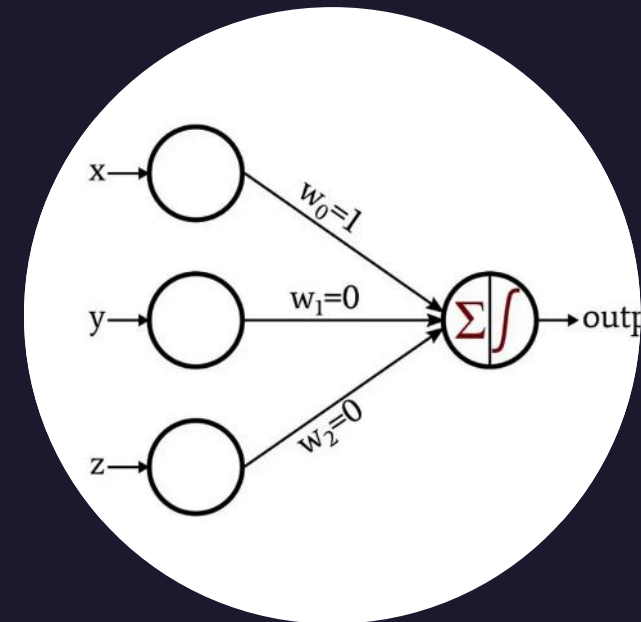
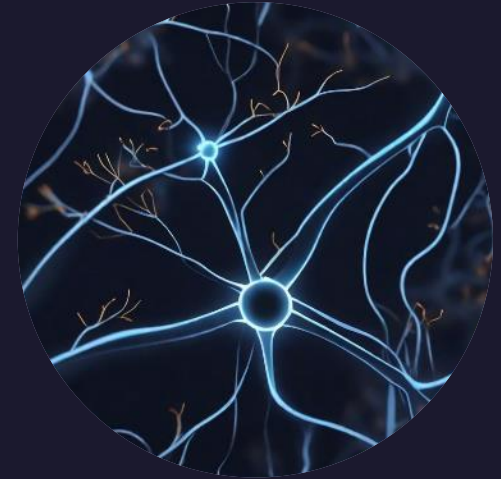
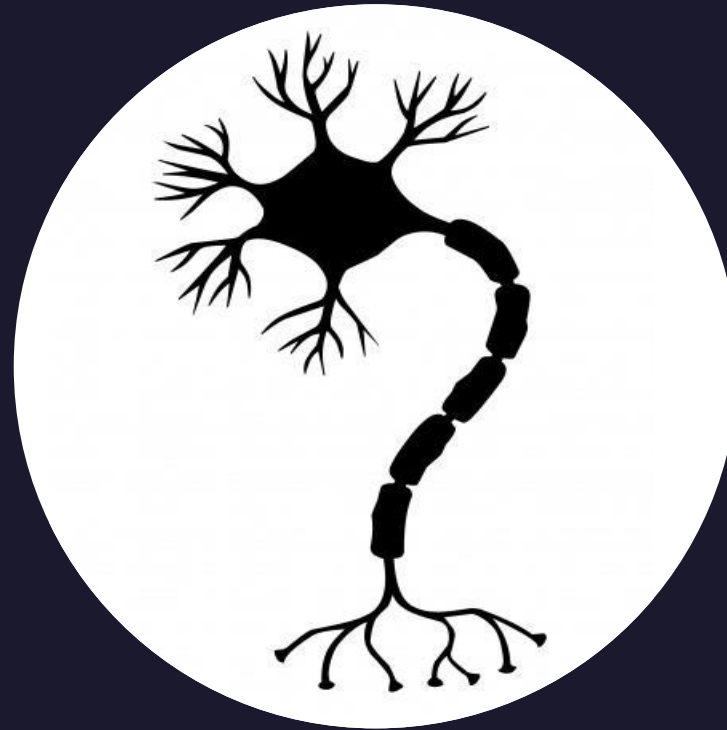
Neurona Biológica: células que transmiten señales eléctricas.

Neurona Artificial: versión simplificada de las neuronas.

Entradas: dendritas que proviene de otras neuronas.

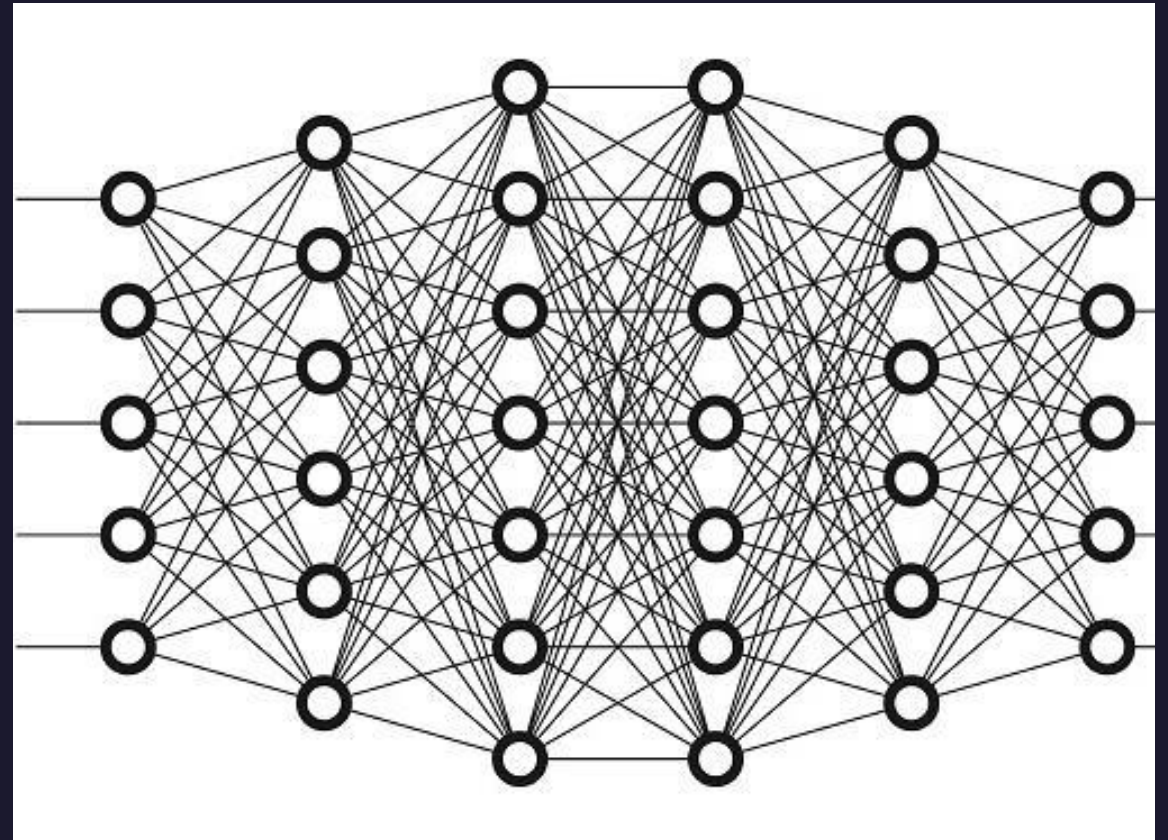
Procesamiento: en función de las entradas se activa o no.

Salida: Activa su salida que puede llegar a otras neuronas.



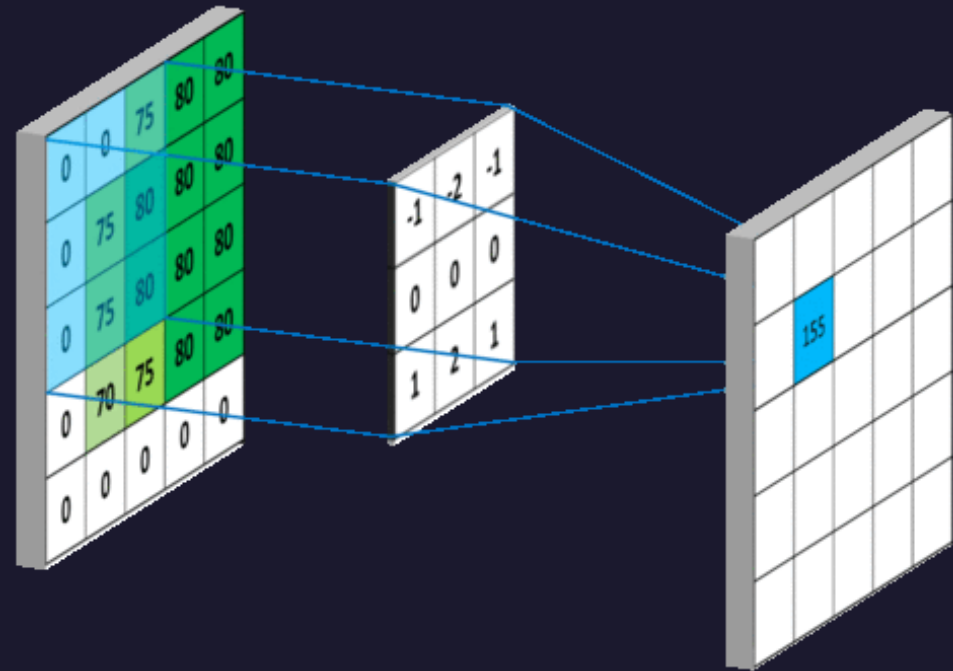
# Deep Learning (aprendizaje profundo)

- Una neurona clasifica binariamente (activa o no).
- No puede resolver problemas complejos.
- Redes de neuronas, se combinan en capas.
- Deep Learning: utilizamos capas, resuelven problemas más complejos.
- Aplicaciones: reconocimiento de patrones, conducción autónoma, traducción, asistentes virtuales...



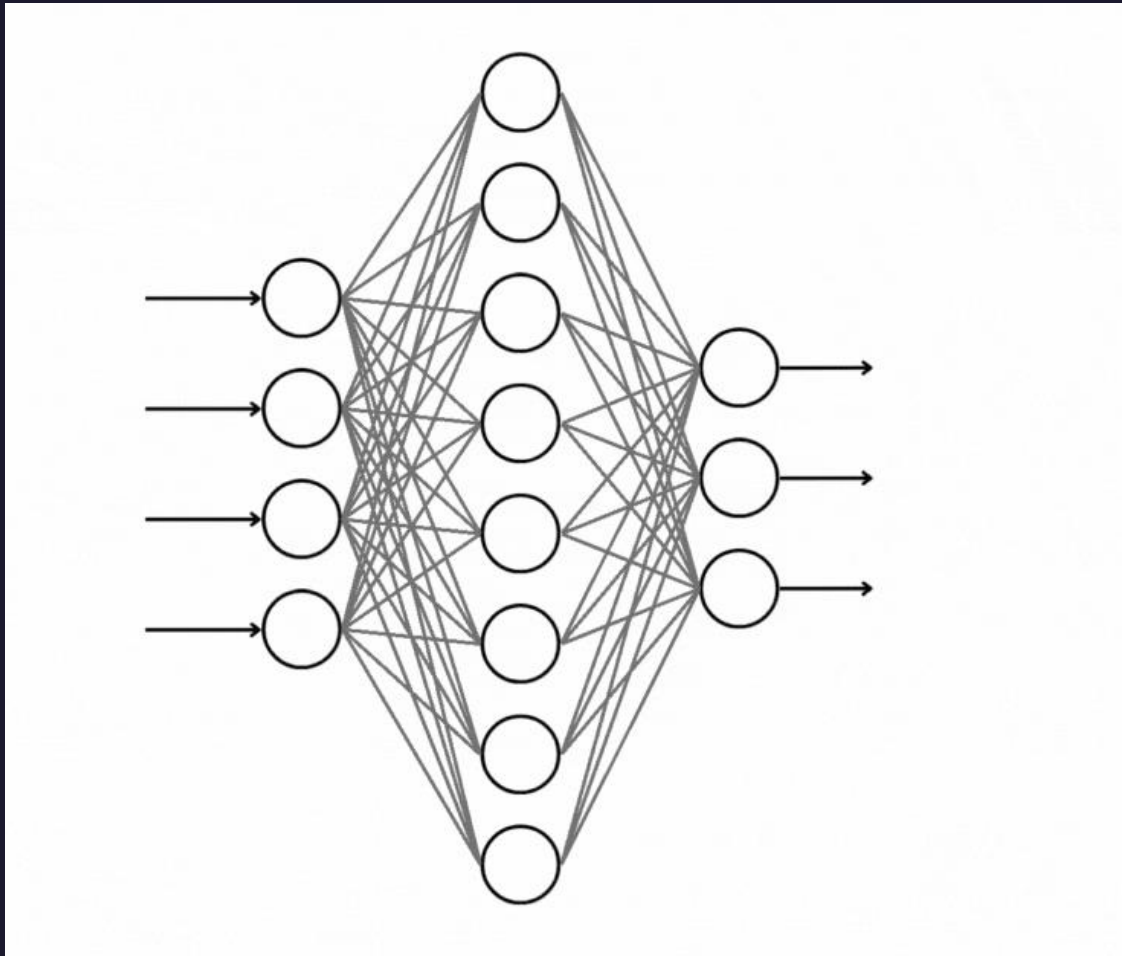
# Tipos de redes neuronales

- Redes Neuronales Artificiales (ANNs)
- Redes Neuronales Convolucionales (CNNs)
- Redes Neuronales Recurrentes (RNNs)
- Redes Neuronales de Memoria a Corto y Largo Plazo (LSTM y GRU)
- Redes Generativas Adversariales (GANs)
- Autoencoders
- Redes de Atención (Transformers)
- Redes Neuronales Híbridas



Red convolucional analizando una imagen

# Fases de funcionamiento



## FASE DE INFERENCIA

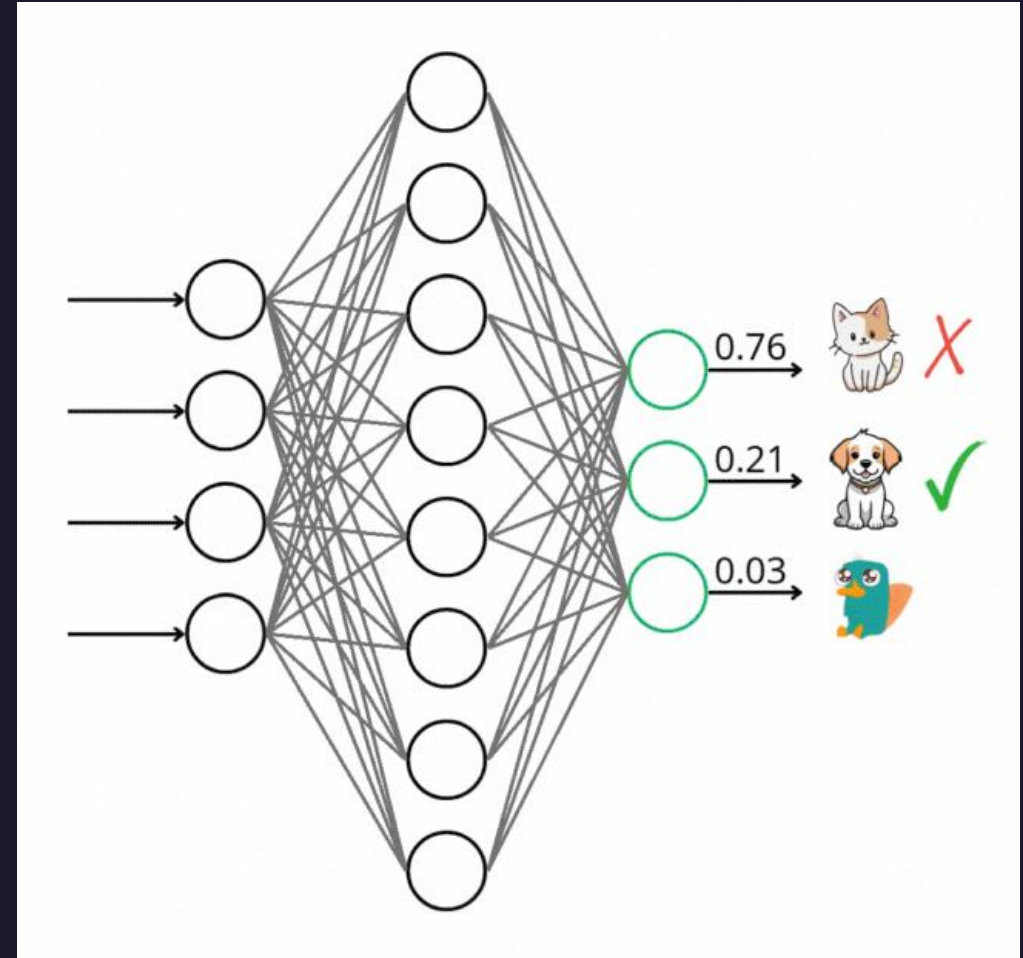
- Desplegado del modelo
- Propagación hacia Adelante
- Obtención de la salida
- Evaluación (si es necesario)

Es posible entrenar parcialmente una red previamente entrenada para tareas similares

# Fases de funcionamiento

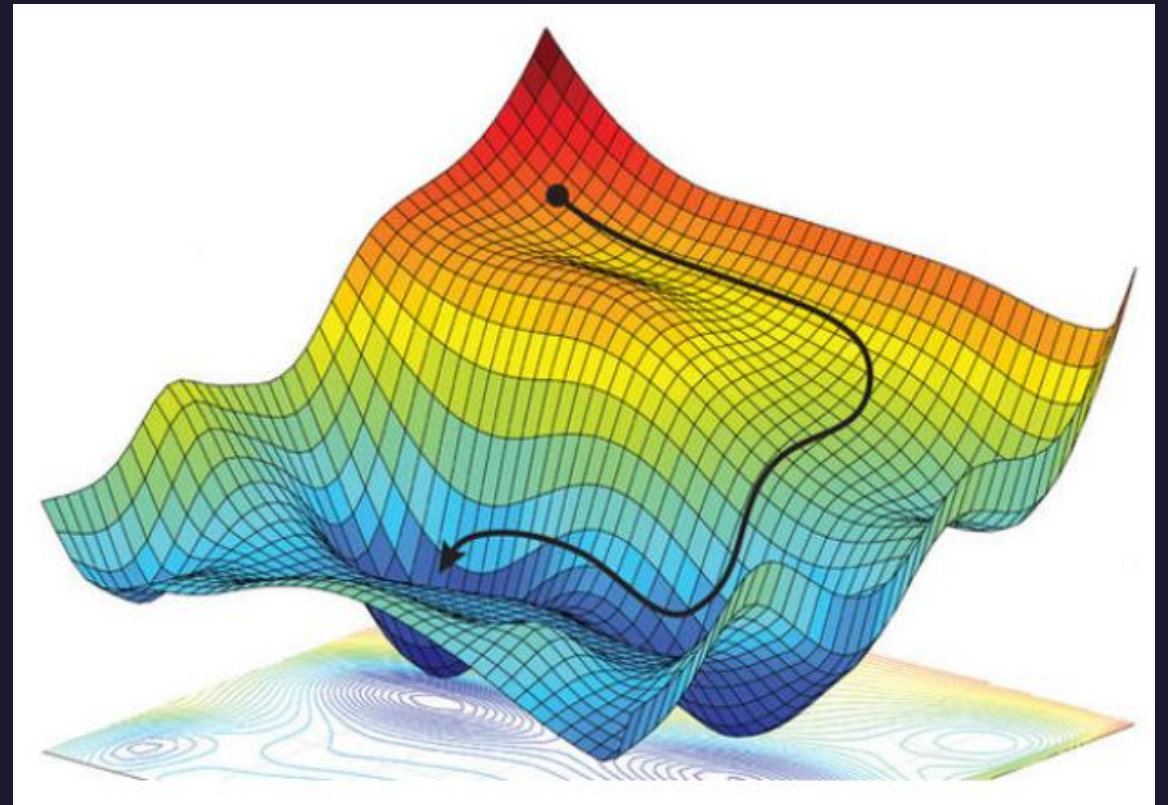
## FASE DE ENTRENAMIENTO

- Recopilar un conjunto de datos amplio de entrenamiento (dataset)
- Construcción del modelo e inicialización
- Los datos se pasan por la red hasta la salida
- Cálculo de error (función de pérdida)
- Ajuste de la red (retropropagación) desde la salida a la entrada
- Repetir hasta converger u obtener rendimiento satisfactorio

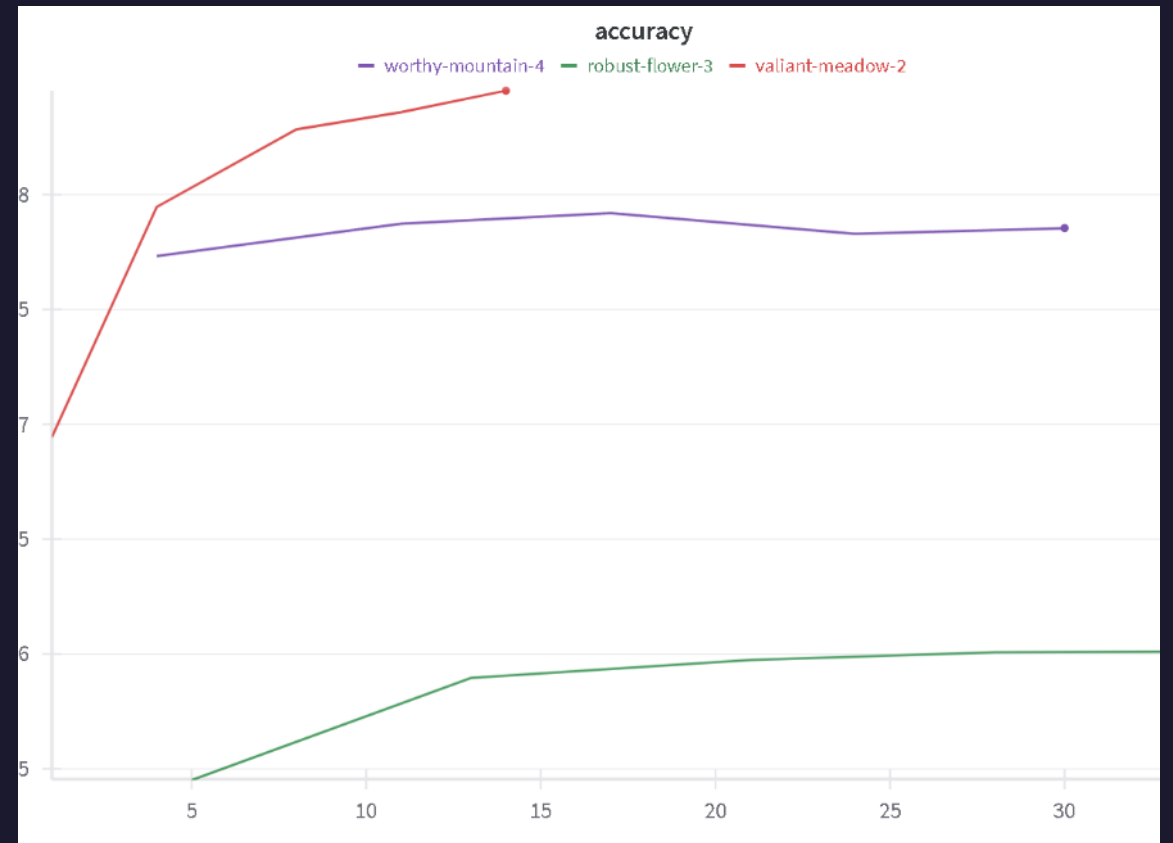


# Aspectos del aprendizaje

- Tamaño del lote (batch size)
- Épocas (epochs)
- Tasa aprendizaje
- Perdida (loss)



# Aspectos del aprendizaje



# Teachable Machine

Interfaz Intuitiva: No requiere experiencia en programación.

Permite entrenar modelos con imágenes, sonidos y poses corporales.

Es completamente gratuita y funciona directamente desde el navegador.

Entrenamiento rápido y en tiempo real.

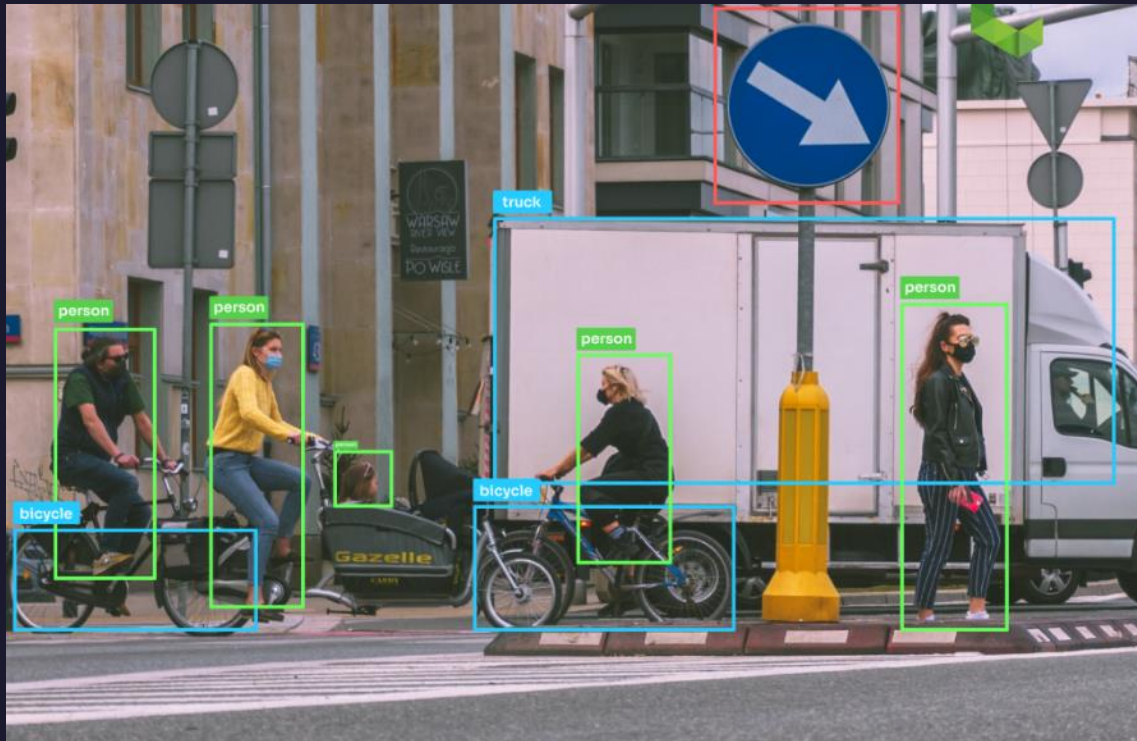
<https://teachablemachine.withgoogle.com/>  
<https://dancingwithai.media.mit.edu/>

The screenshot displays the Teachable Machine web interface. On the left, three training classes are visible: 'Manzana' (Apple), 'Platano' (Banana), and 'Melón' (Melon). Each class has a 'Webcam' and 'Subir' (Upload) button, followed by a row of 20 image samples. A 'Preparación' (Preparation) panel is open, showing 'Modelo preparado' (Model prepared) and 'Avanzado' (Advanced) options. On the right, a 'Vista previa' (Preview) section shows a green apple with a bounding box. Below it, a 'Salida' (Output) section shows progress bars for each class: 'Manz...' is at 100%, 'Platano' is at 0%, and 'Melón' is at 0%.

<https://amozo.es/ViTArt/>

# Otras tareas aprendizaje con imágenes

## DETECCIÓN OBJETOS



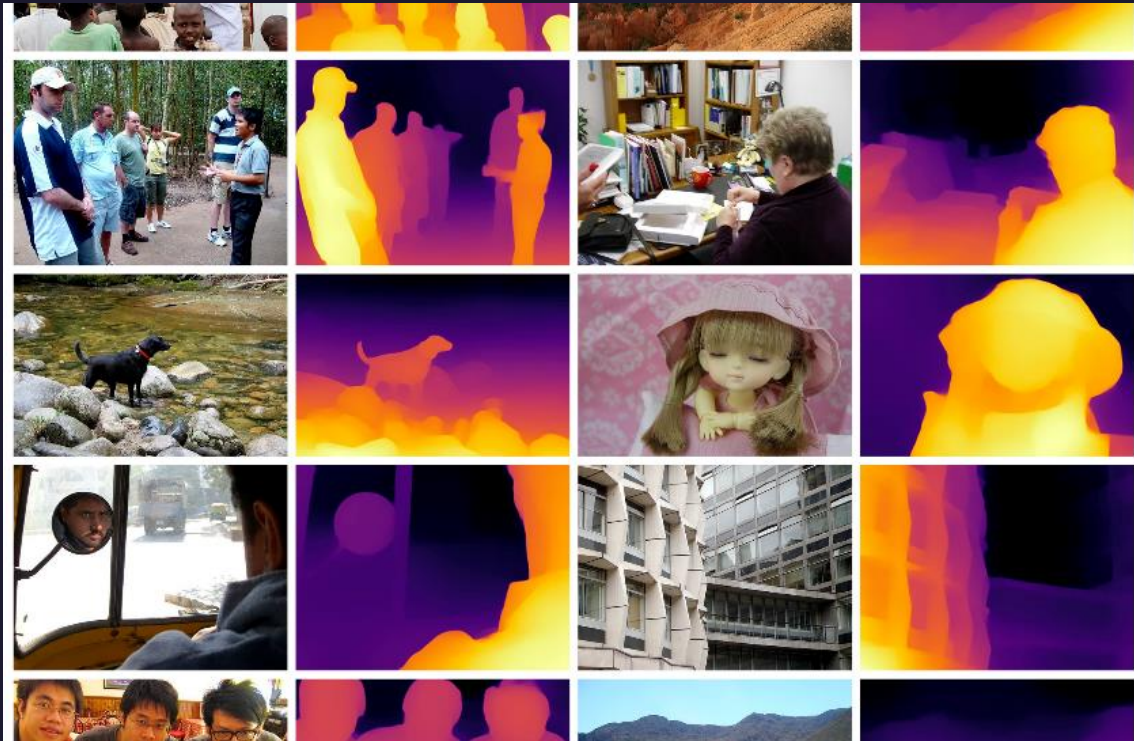
## SEGMENTACIÓN OBJETOS



<https://docs.ultralytics.com/>

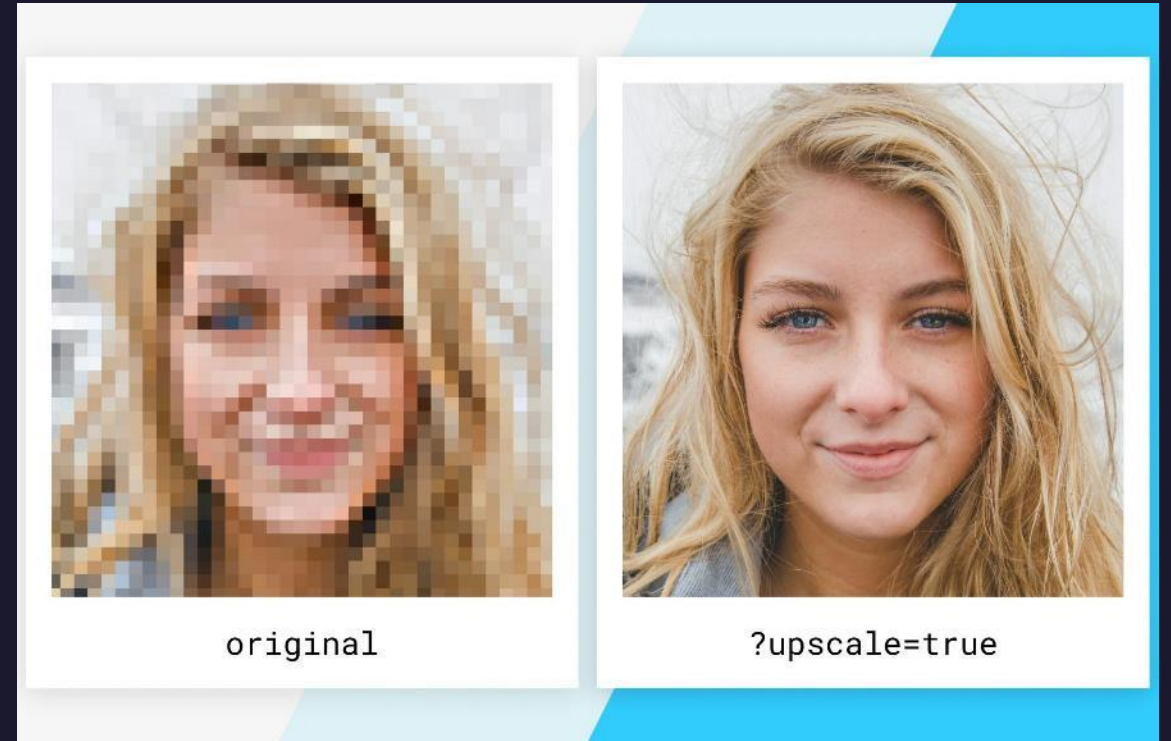
# Otras tareas aprendizaje con imágenes

## MAPAS PROFUNDIDAD



[https://pytorch.org/hub/intelisl\\_midass\\_v2/](https://pytorch.org/hub/intelisl_midass_v2/)

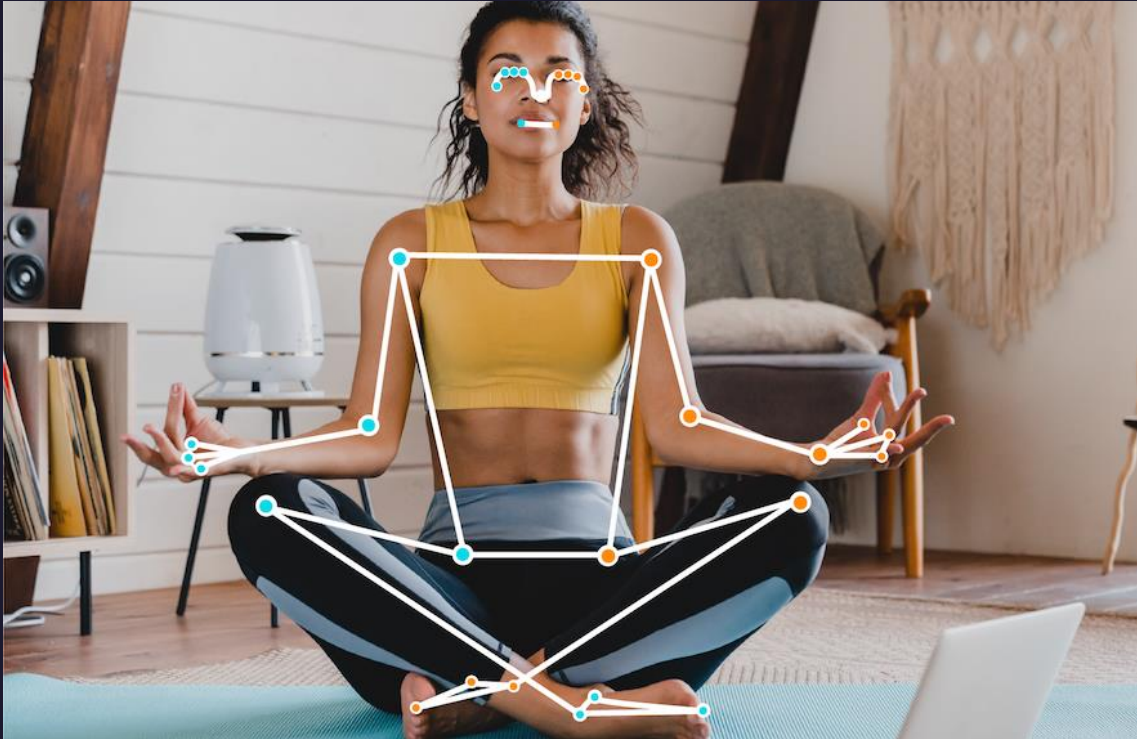
## SUPER RE-ESCALADO



<https://magnific.ai/>

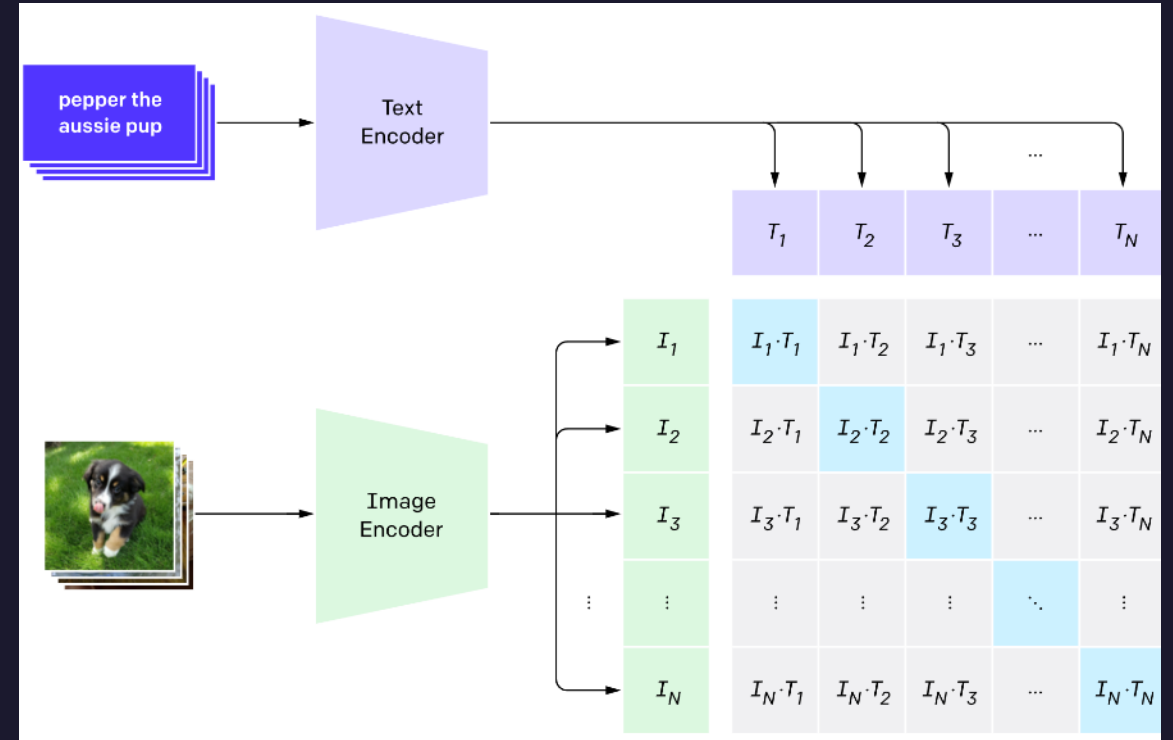
# Otras

## DETECCIÓN DE POSES



<https://semiconductor.withgoogle.com/posenet>

## CONTENIDO SEMÁNTICO DE IMÁGENES



<https://amozo.es/PicassoAlg/>

# Tipos aprendizajes

Supervisado: proporcionamos respuesta

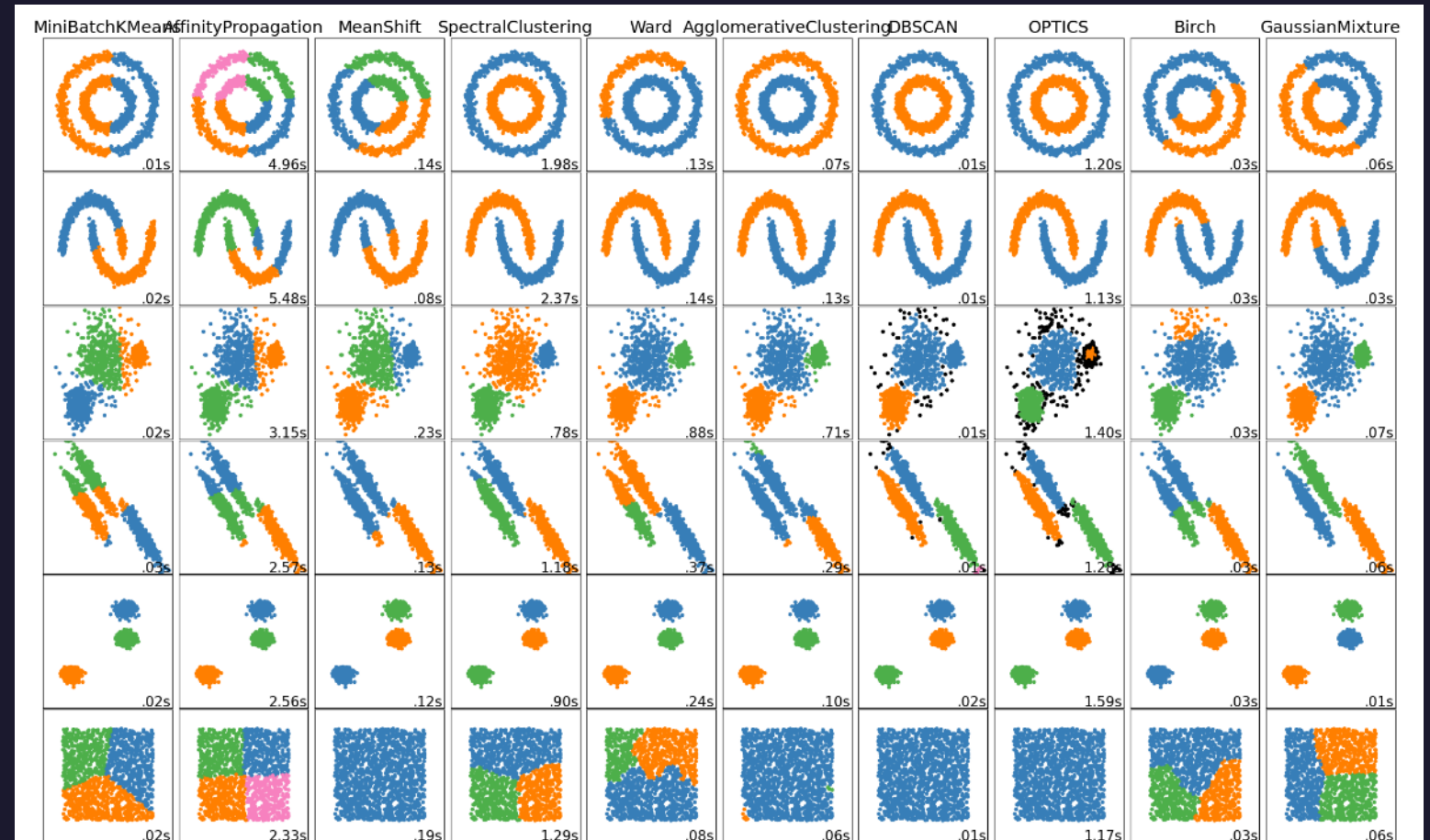
Clasificación

No supervisado

Clustering

Autosupervisado

GPT



<https://amozo.es/clusters3D/>

# Problemas aprendizaje

## SOBREAJUSTE (OVERFITTING)

- Aprende demasiado bien los detalles
- Un ligero cambio produce error
- Bien en entrenamiento
- Mal inferencia
- Ejemplo:
  - Entrenamiento con manzanas rojas
  - En inferencia no reconoce una manzana verde

## SUBAJUSTE (UNDERFITTING)

- No capta los detalles
- Posible modelo simple
- Posible pocos datos de entrenamiento
- Mal en entrenamiento e inferencia
- Ejemplo:
  - Entrenamiento con una imagen de pera y manzana
  - En inferencia no reconoce correctamente



# Problemas aprendizaje

## DATOS INSUFICIENTES O DESEQUILIBRADOS

- Una o varias clases se identifican bien
- Una o varias clases no se identifican bien
- Ejemplo:
  - 1000 imágenes de perros
  - 10 imágenes de gatos

## DATOS CON RUIDO O ERRÓNEOS

- Datos con errores
- Datos incompletos
- Datos irrelevantes
- Pueden confundir al modelo



# Datasets

## OPEN IMAGES DATASET V7 DE GOOGLE

- 9 millones de imágenes
- etiquetas a nivel de imagen
  - 61.404.966 etiquetas que abarcan 20.638 clases
- cuadros delimitadores (bounding boxes)
  - 15.851.536 cuadros en 600 clases
- máscaras de segmentación
  - 2.785.498 segmentaciones en 350 clases

[Google Cloud Storage](#)

## COCO (COMMON OBJECTS IN CONTEXT)

- 123,287 imágenes.
- 80 categorías que abarcan desde personas y animales hasta vehículos y objetos domésticos
- 886.284 segmentaciones

<https://cocodataset.org/#explore>



# GPT (Generative Pre-trained Transformer)

- Modelo de lenguaje basado en la arquitectura Transformer.
- Entrenado con datos de Internet.
- Modelo autoregresivo: genera texto de modo secuencial, predice la siguiente palabra en función de las anteriores (método de atención).
- El coste de entrenar GPT-4 fue más de 100 millones de dólares.
- GPT-1 (2018): 117 Millones de parámetros
- GPT-2 (2019): 1500 Millones de parámetros
- GPT-3 (2020): 175.000 Millones de parámetros
- GPT-4 (2023): ¿100 Billones de parámetros?

El que no corre ...

Salta (0.04)

Nada (0.01)

Vuela (0.75)

Camina (0.20)



# DALL·E 2 (3)

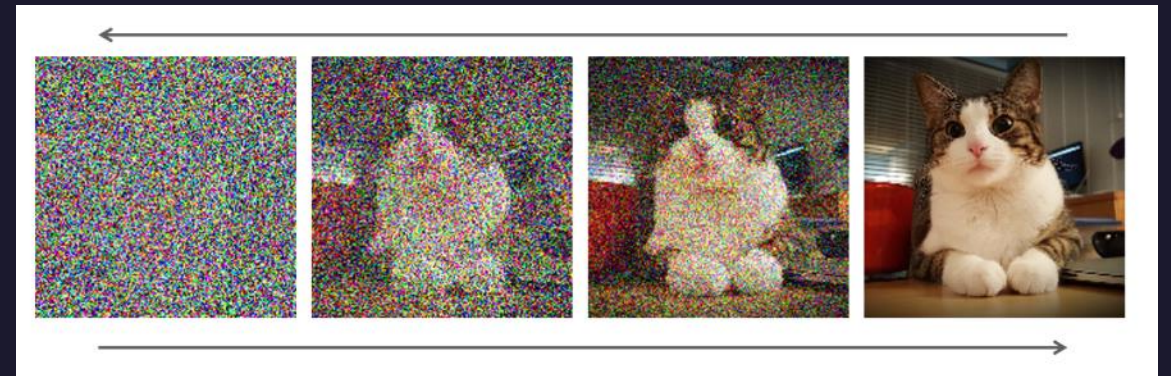
- Generación de imágenes a partir de un texto.
- Outpainting: genera contenido alrededor de una imagen previa.
- Inpainting: cambia parte del contenido de una imagen previa.
- Variaciones: toma una imagen como partida.



Ejemplos extraídos de <https://openai.com/dall-e-2>

# DALL·E 2 (3)

- Modelo de difusión.
- Parte de una imagen de ruido, sucesivamente elimina el ruido hasta llegar a la imagen generada.
- La generación se condiciona con el prompt, otra imagen, inpainting, mapas de profundidad, poses, etc.
- Alternativa a las redes GAN (redes generativas adversarias)





<http://amozo.es/ia-fancine>

Presentación

Dataset imágenes fruta

Github



# ¿Qué nos espera en el futuro?

Alejandro Mozo Quesada

Transversal Skills in Applied Artificial Intelligence (tsaii.eu)



UNIVERSIDAD  
DE MÁLAGA

