

Transfer learning con Hugging Face

Reúsa modelos preentrenados; entrena lo justo.

CONCEPTOS CLAVE

- Tres niveles: **feature extraction** (congelado, 100 ej.), **fine-tuning parcial** (1 000 ej.), **total** (10 000+ ej.).
- Hugging Face Hub: 475 000 modelos preentrenados. Antes de entrenar, busca.
- Licencias relevantes para uso comercial: Apache 2.0 y MIT.
- Learning rate para fine-tuning: $\sim 2 \cdot 10^{-5}$, dos órdenes por debajo del entrenamiento desde cero.

EJEMPLO A MANO

ResNet-50 con cabeza nueva de 10 clases. El backbone tiene 25,5 M parámetros, todos congelados (`requires_grad = False`). La cabeza añadida es una `Linear(2048 → 10)`, con

$$\text{params entrenables} = 2048 \cdot 10 + 10 = \mathbf{20\,490}.$$

Caben en cualquier portátil sin GPU. Tiempo de entrenamiento con 800 imágenes y 10 épocas: ~ 75 segundos en CPU. Coste cloud: 0 \$.

CONCEPTO → APLICACIÓN CON MODELOS DEL HUB

Modelo	Aplicación	Licencia
ViT-base	Catálogo retail genérico.	Apache 2.0
ResNet-50	Control de calidad en planta.	Apache 2.0
CLIP-large	Búsqueda imagen-texto similar.	MIT
DiT-base	OCR de documentos y facturas.	MIT
SegFormer-b0	Segmentación de imágenes satélite.	Apache 2.0

Idea clave del Día 05. El Hub de Hugging Face es el lugar donde alguien ha pagado ya **72 \$** entrenando ResNet-50 en ImageNet; tu trabajo es reusarlo, no repetirlo. La receta exprés en 6 pasos (Hub → cabeza → congelar → `TrainingArguments` → `Trainer` → evaluar) cubre el 90 % de los casos.

Para llevarte: comprueba siempre la licencia antes de uso comercial.