

Meta aprendizaje: aprendiendo de pocos ejemplos

Aprendizaje Profundo

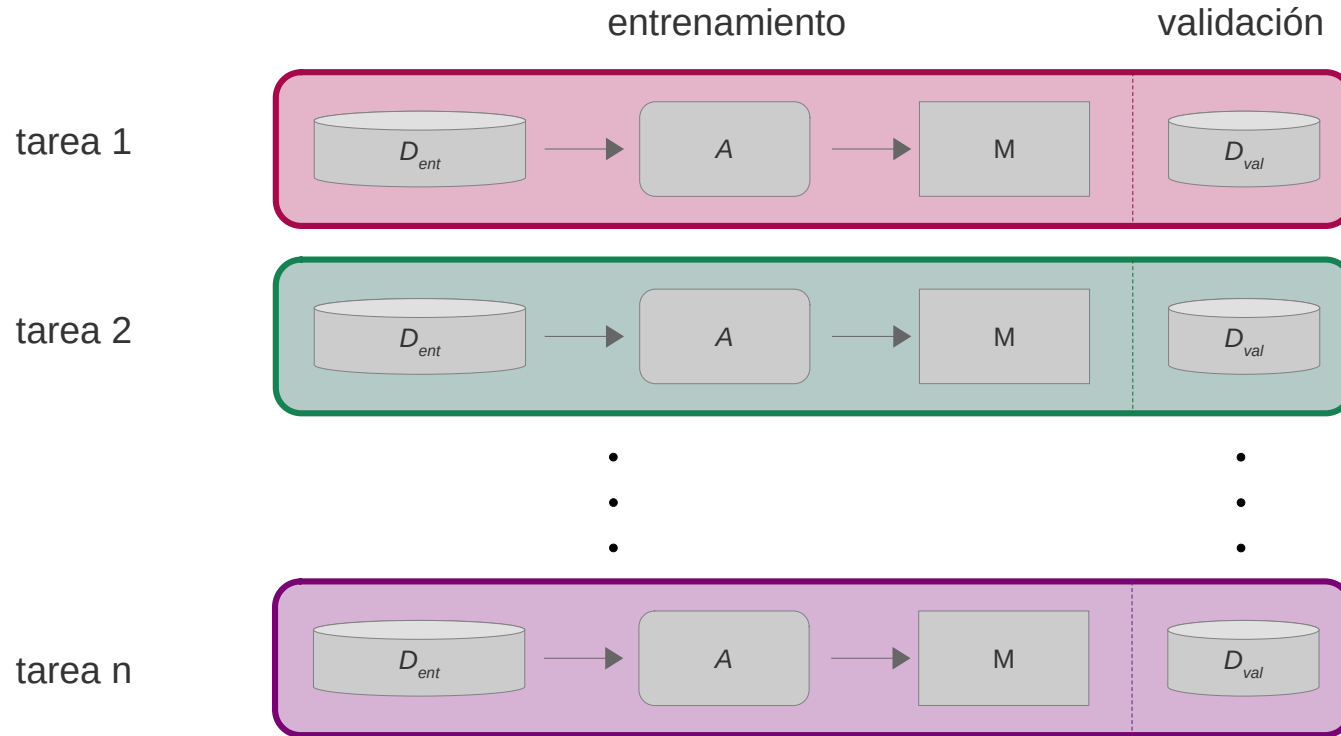
PCIC IIMAS UNAM

Prof. Gibran Fuentes Pineda
Ayud. Bere & Ricardo Montalvo Lezama

Enero 2021

Aprendizaje convencional

- Se entrena un modelo por cada tarea a resolver.



¿Por qué es importante aprender con pocos datos?

- Existen muchos dominios donde los datos etiquetados son escasos.

traducción de lenguas raras

robótica

imágenes médicas



Report: opacidades de aspecto intersticioalveolar parcheadas y bilaterales que predominan en ambos lobulillos inferiores sospechosas de infección por COVID-19 . senos costofrenicos libres .

Labels: COVID 19, alveolar pattern, interstitial pattern, pneumonia

Locations: costophrenic angle, lobar, bilateral, lower lobe

DICOM Fields

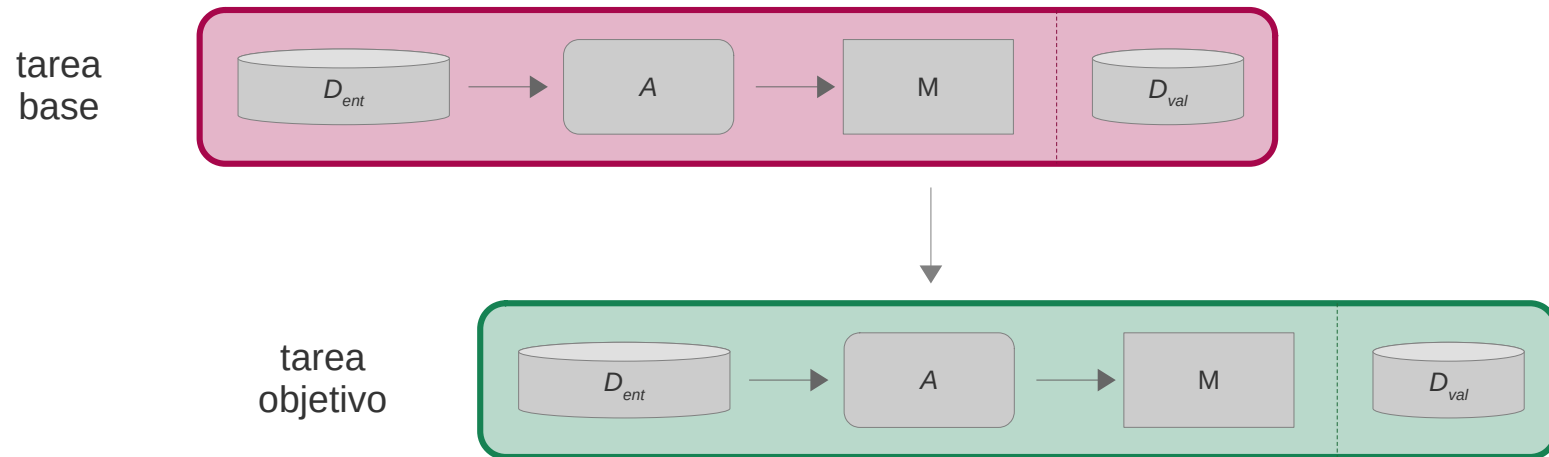
Study Date 20200317
Patient's Sex M
Patient's Birth Date 1986
Modality CR
Manufacturer GE Healthcare
...

Date	Test	Result
17.03.2020	PCR	NEGATIVE
18.03.2020	PCR	NEGATIVE
19.03.2020	IGG	POSITIVE
19.03.2020	IGM	POSITIVE
20.03.2020	PCR	POSITIVE

... ..

Transferencia de conocimiento

- Aprovecha el conocimiento de una tarea base en una tarea objetivo.



Adaptación de dominio

- Aprender un modelo de una distribución origen y se aplica a una distribución diferente.

Dominio origen

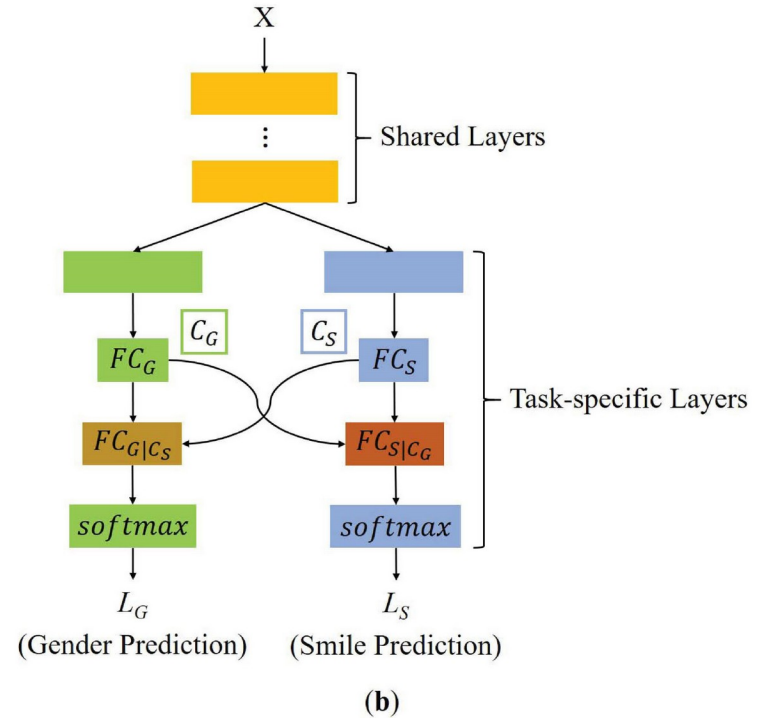
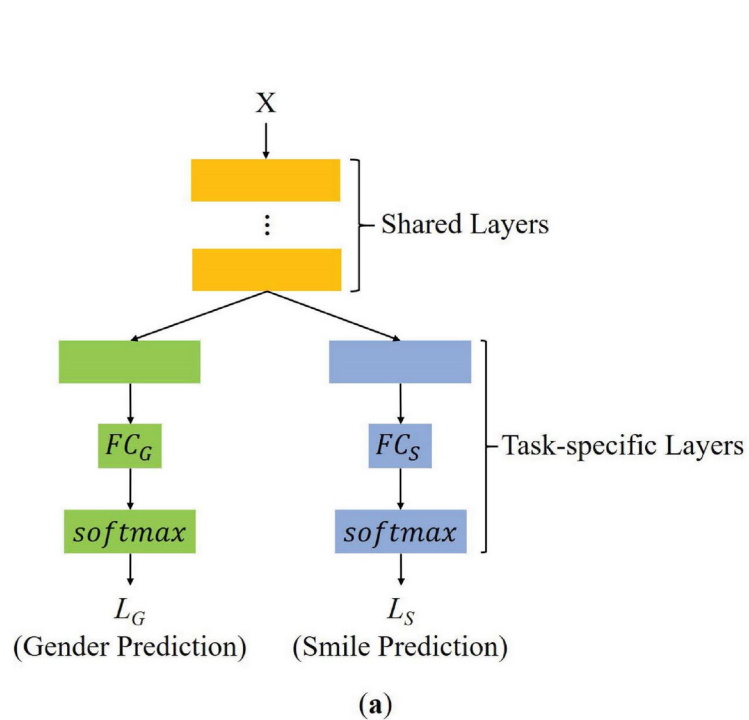


Dominio objetivo



Aprendizaje multitarea

- Aprendizaje simultáneo de varias tareas relacionadas.



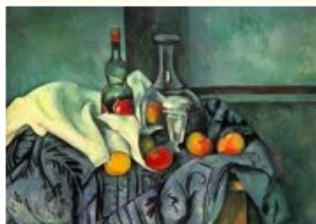
Una intuición de meta aprendizaje

Entrenamiento

Prueba

Braque

Cezanne

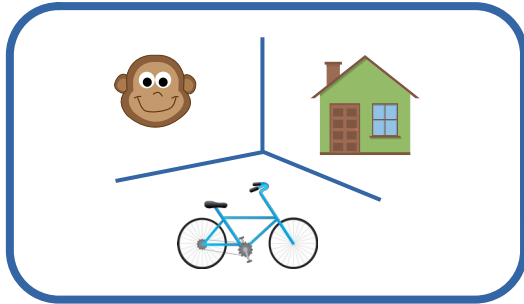


¿Qué artista pintó esta obra?
¿Braque o Cezanne?

- Familia de técnicas enfocadas adaptarse rápidamente a nueva información.

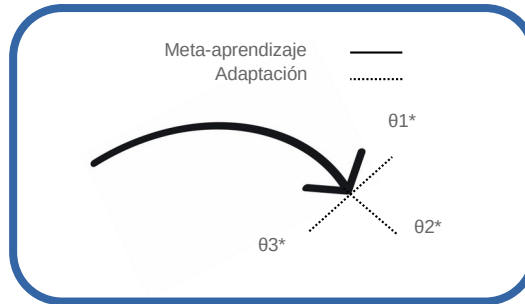
Tipos de meta aprendizaje

Métricas



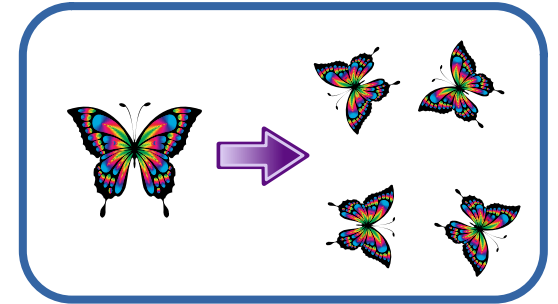
comparación

Optimizadores



optimización

Alucinaciones



aumentado

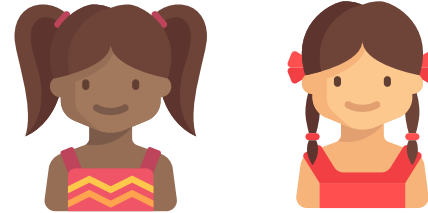
Métricas

Verificación de rostros

- Comparar la imagen del rostro de una persona con otra y verificar si coinciden.

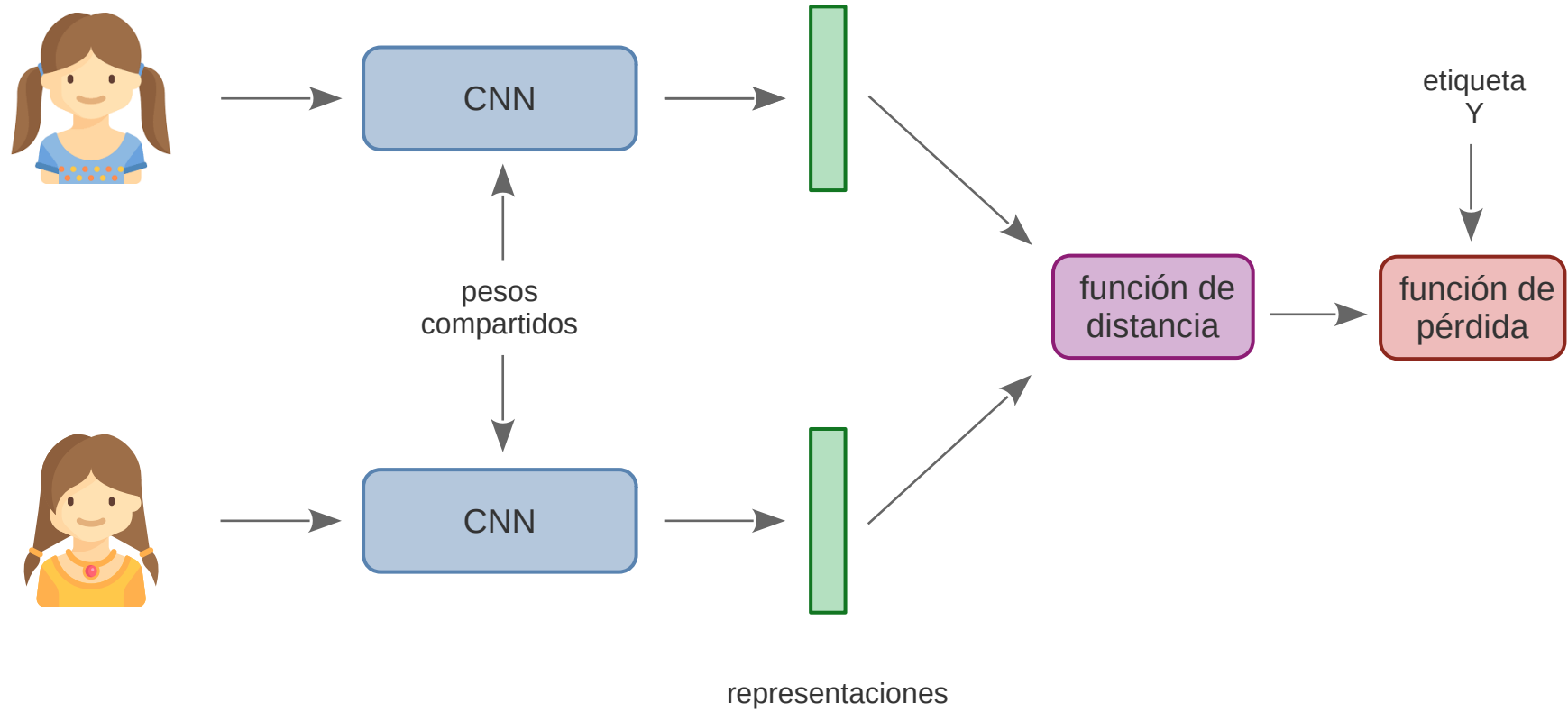


ejemplo positivo

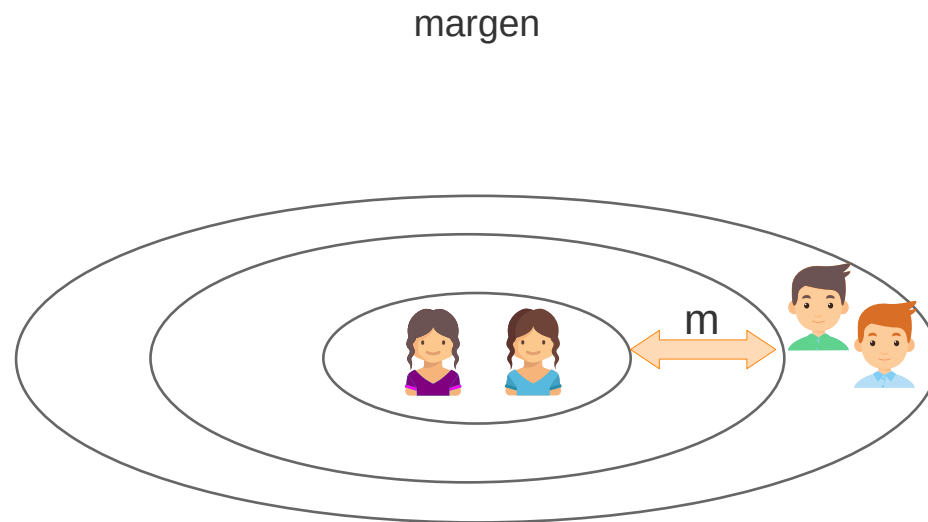
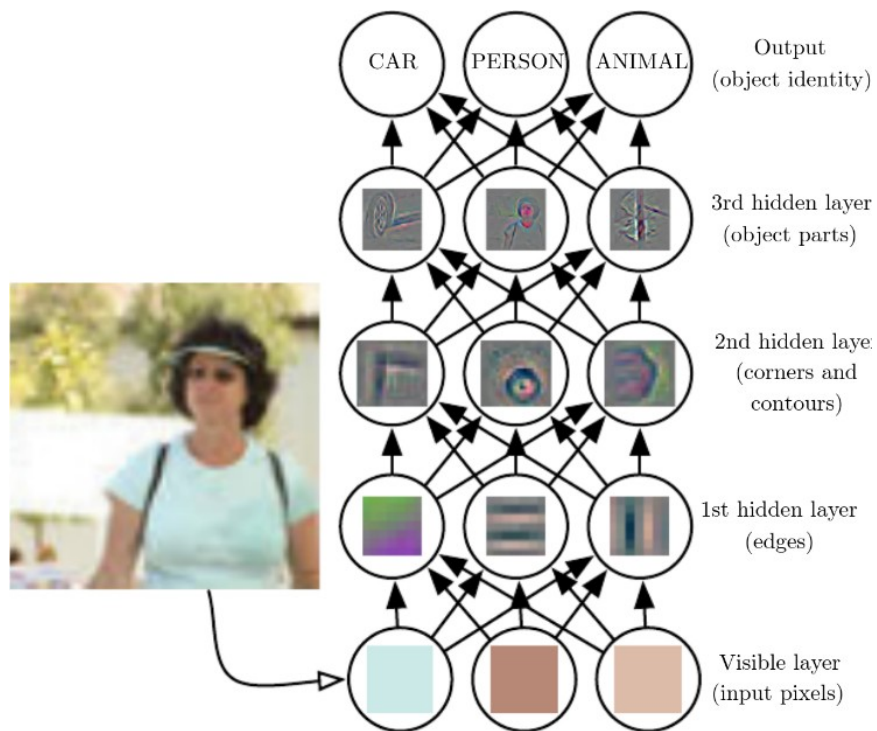


ejemplo negativo

Red Siamesa



Representaciones



Pérdida constructiva

$$J(x_1, x_2, y) = (y d^2) + (1 - y) (\max(0, m - d)^2)$$

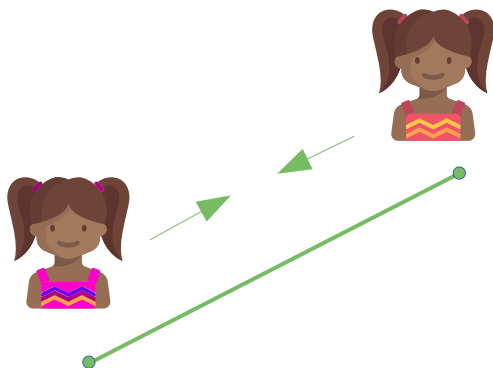
d : distancia euclídeana entre representaciones $d = \sqrt{(x_1 - x_2)^2}$

y : etiqueta de las parejas (1 similares, 0 distintas)

m : margen

Similares

$y = 1$

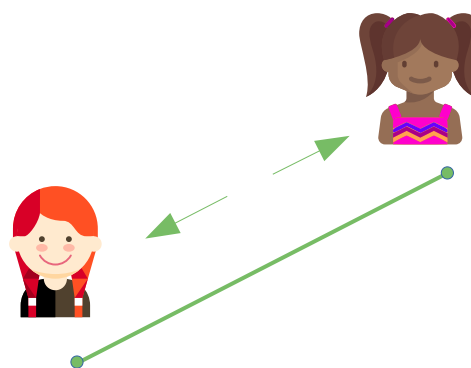


$$J(x_1, x_2, y) = d^2$$

Para minimizar J
se minimiza d

Distintos

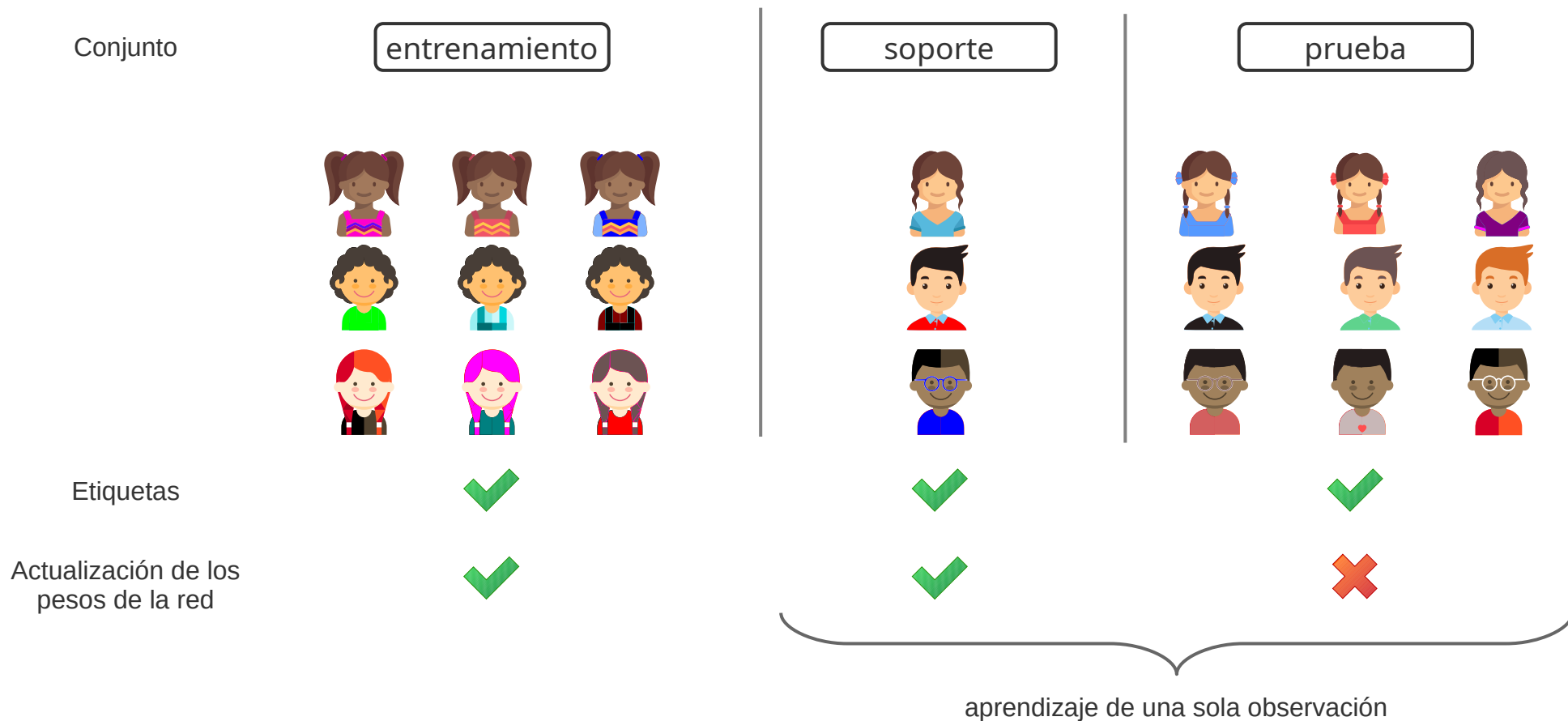
$y = 0$



$$J(x_1, x_2, y) = \max(0, m - d)^2$$

Para minimizar J
se maximiza d

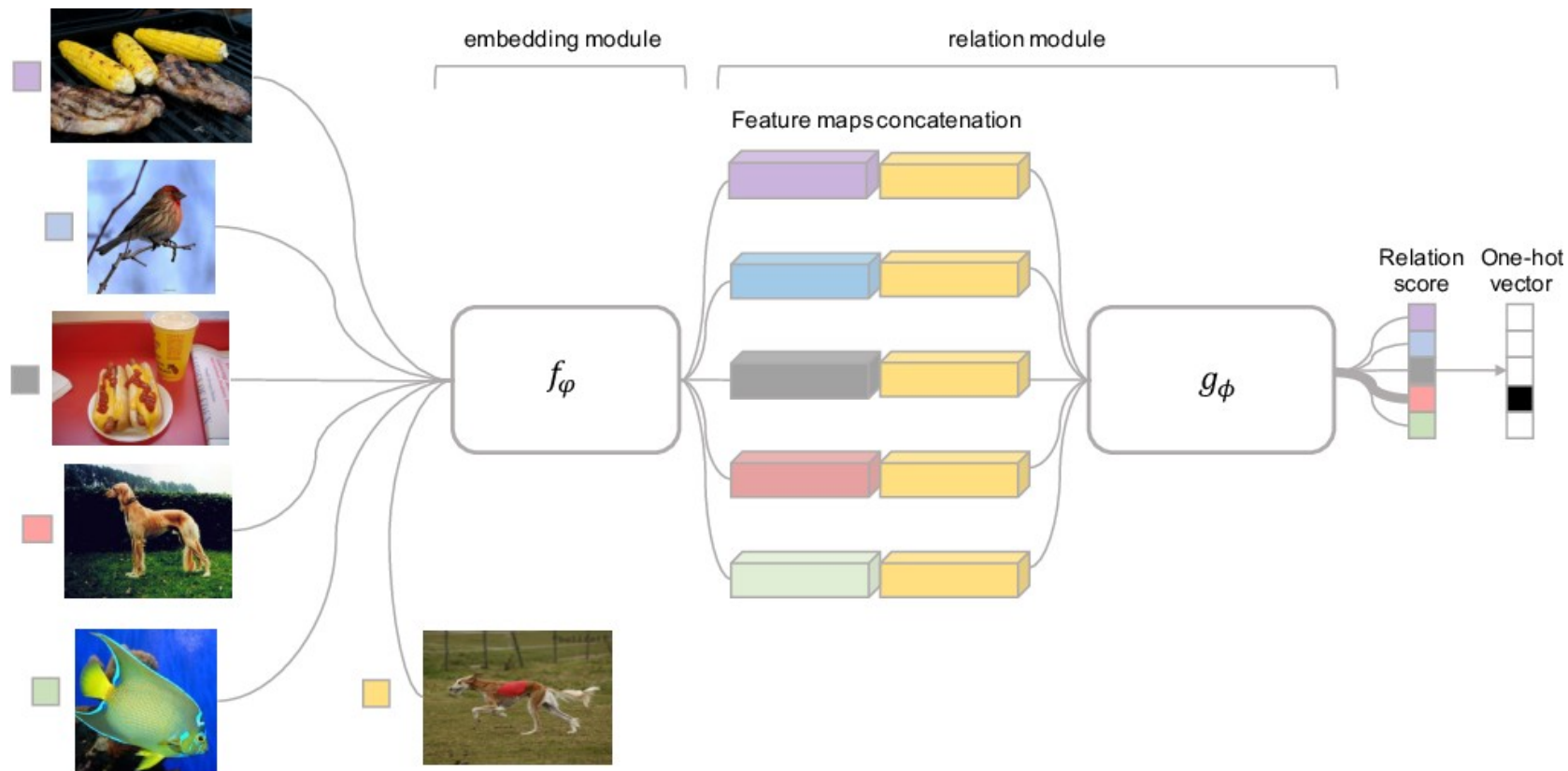
Aprendizaje de una observación





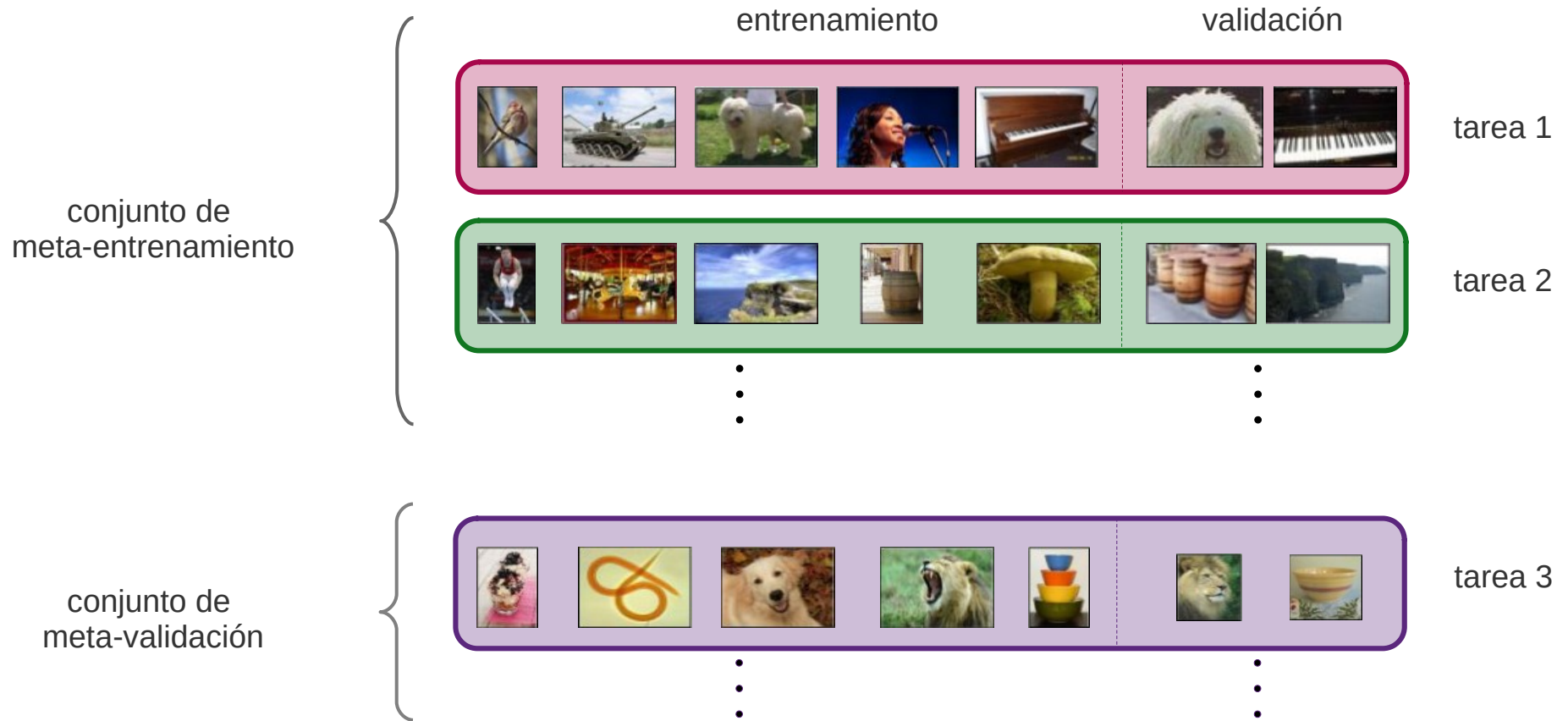
¡tiempo de programar!
7b_mtl_siamese.ipynb

Red relational

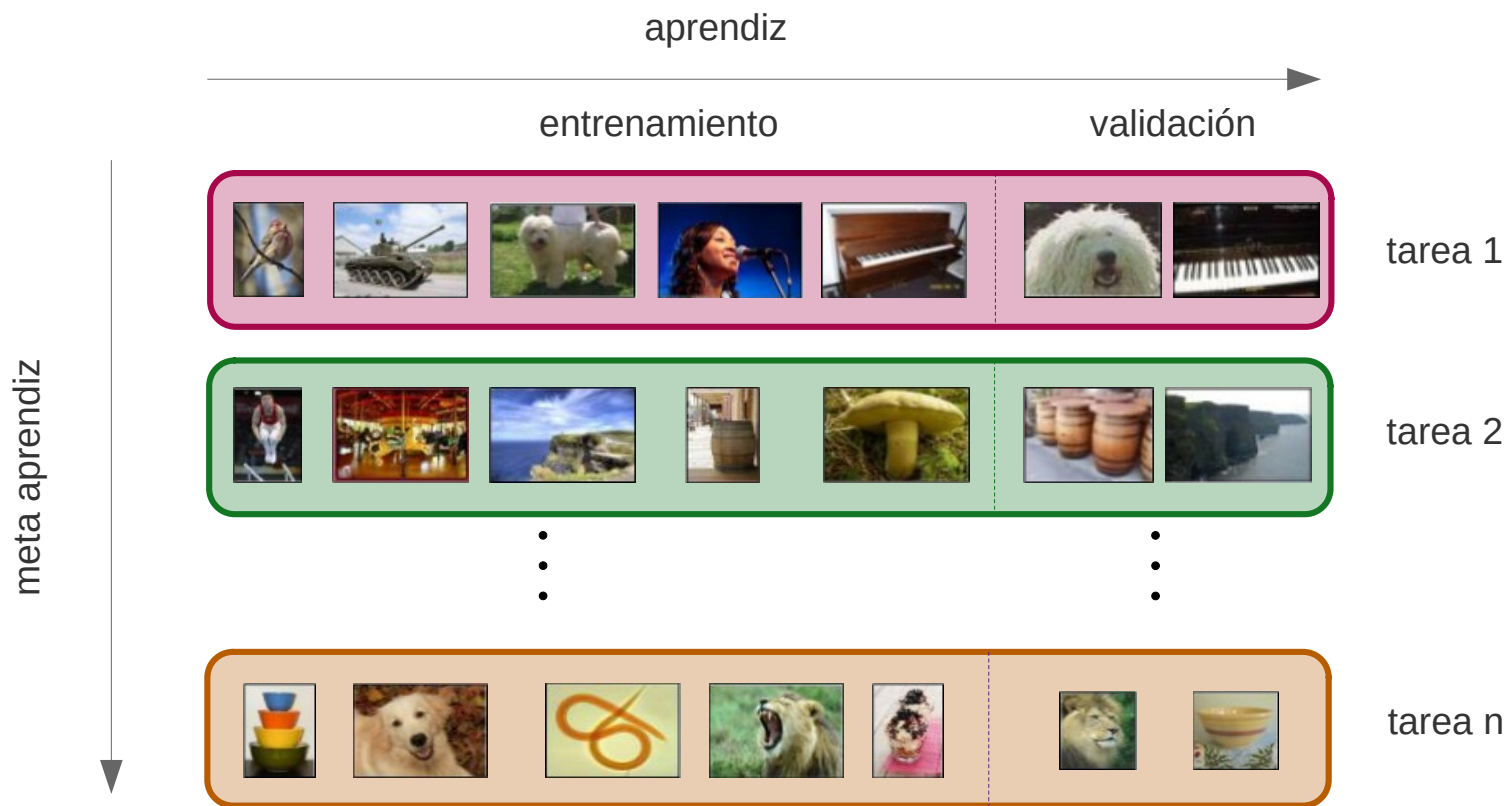


Optimizador

Clasificación con pocas observaciones con episodios



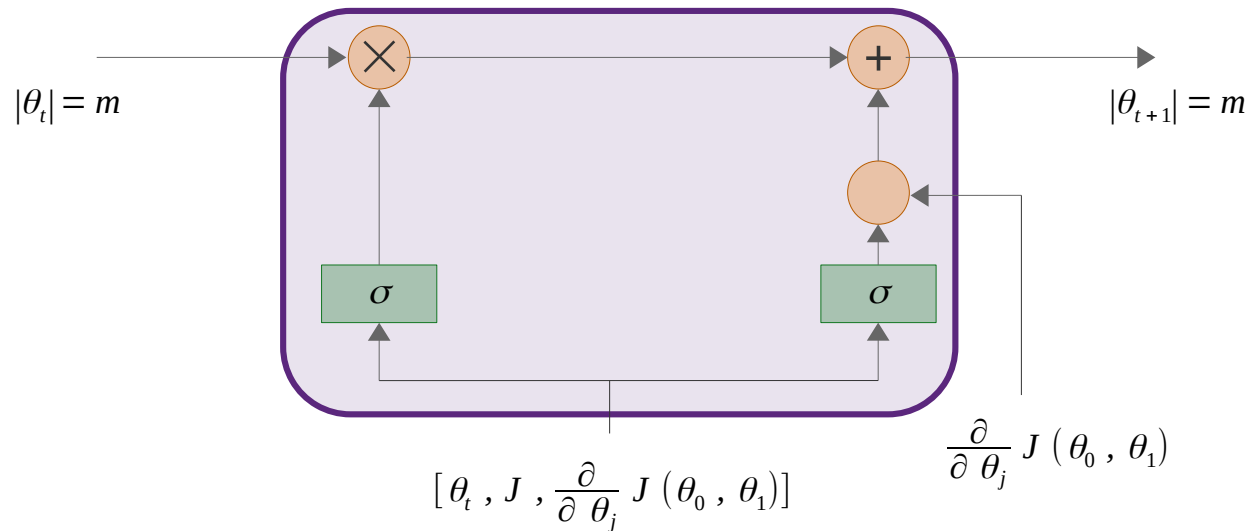
Meta optimizador



Meta optimizador vs SGD

$$\theta = 1 \theta_{t-1} + \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \quad \text{SGD}$$

$$\theta = f \theta_{t-1} + i \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \quad \text{meta-aprendiz}$$





¡Gracias!