

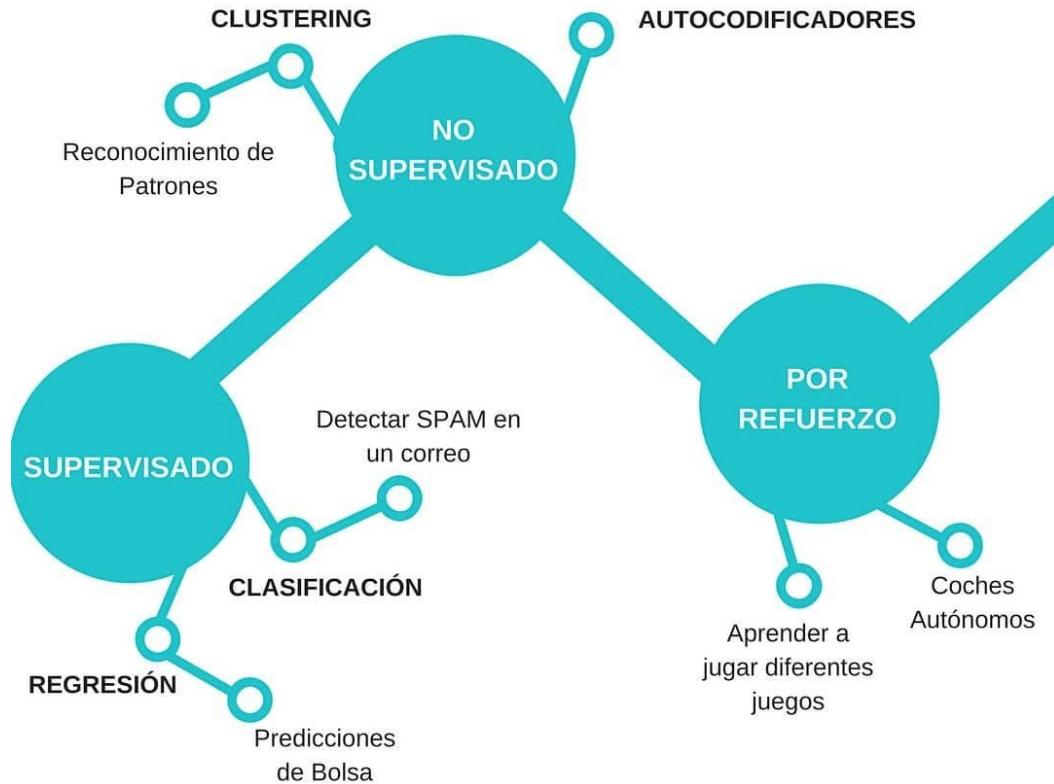
Tipos de Redes Neuronales

<https://twitter.com/osanseviero>

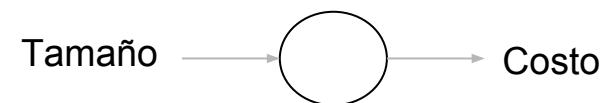
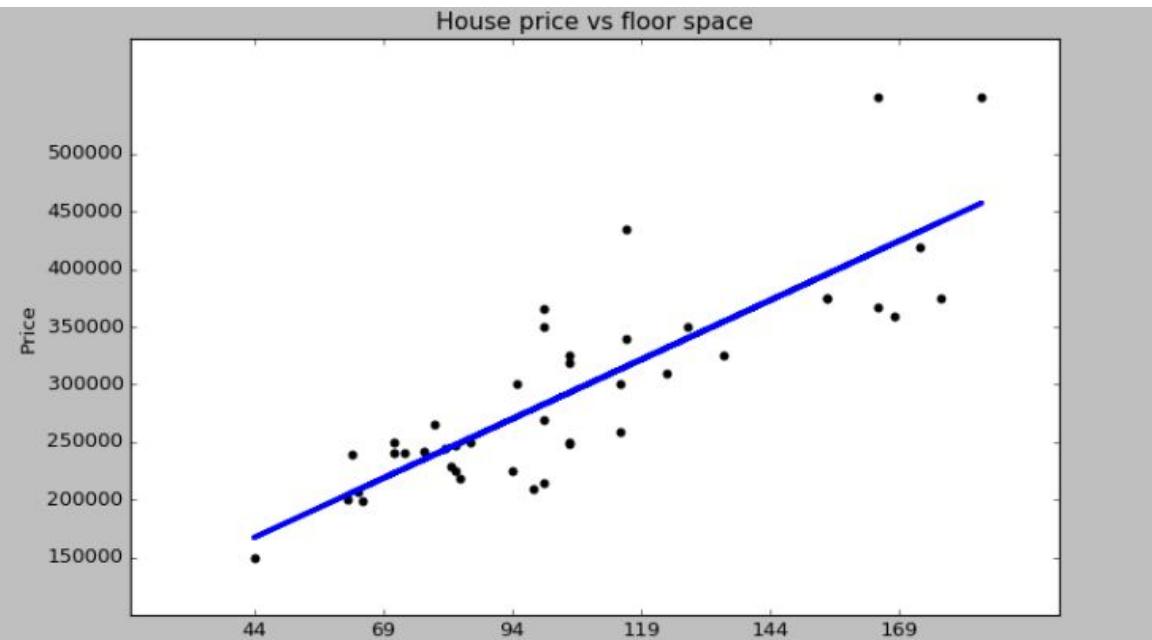
www.linkedin.com/in/omarsanseviero/

Redes Neuronales

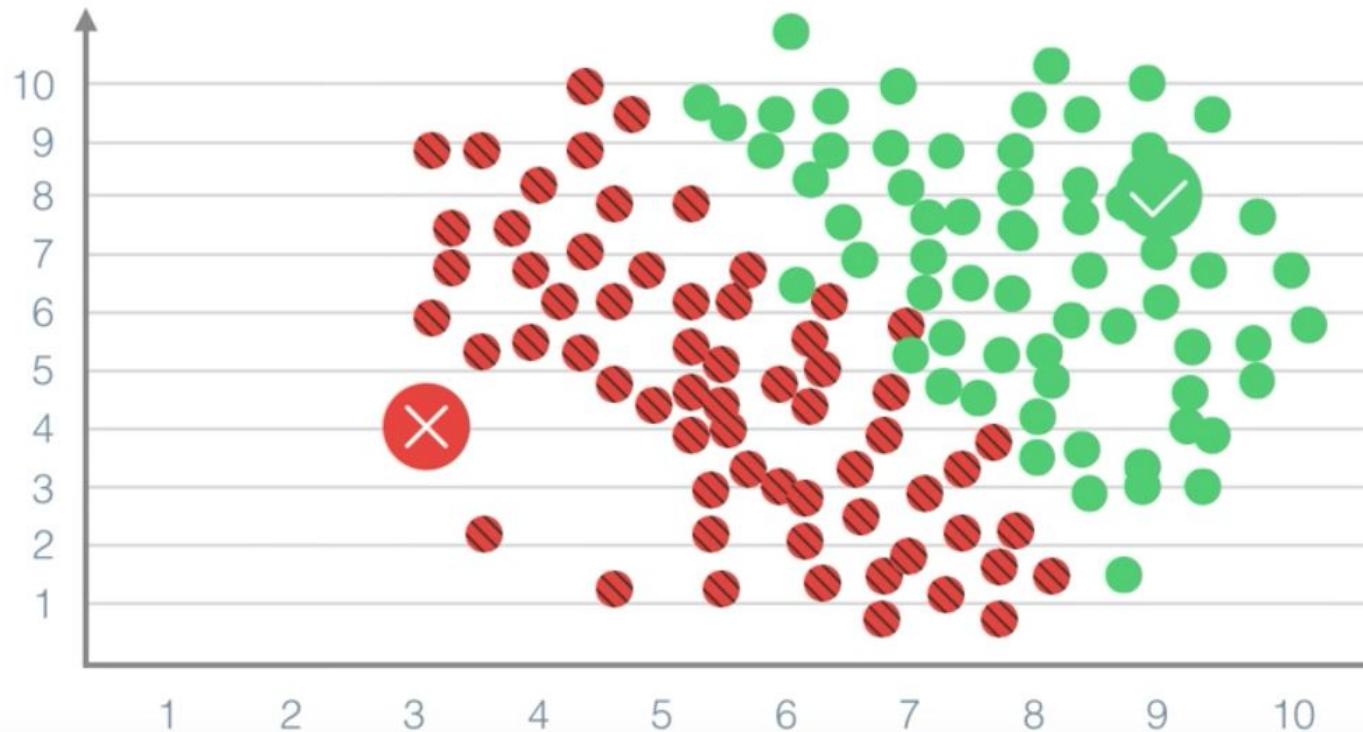
TIPOS DE MACHINE LEARNING



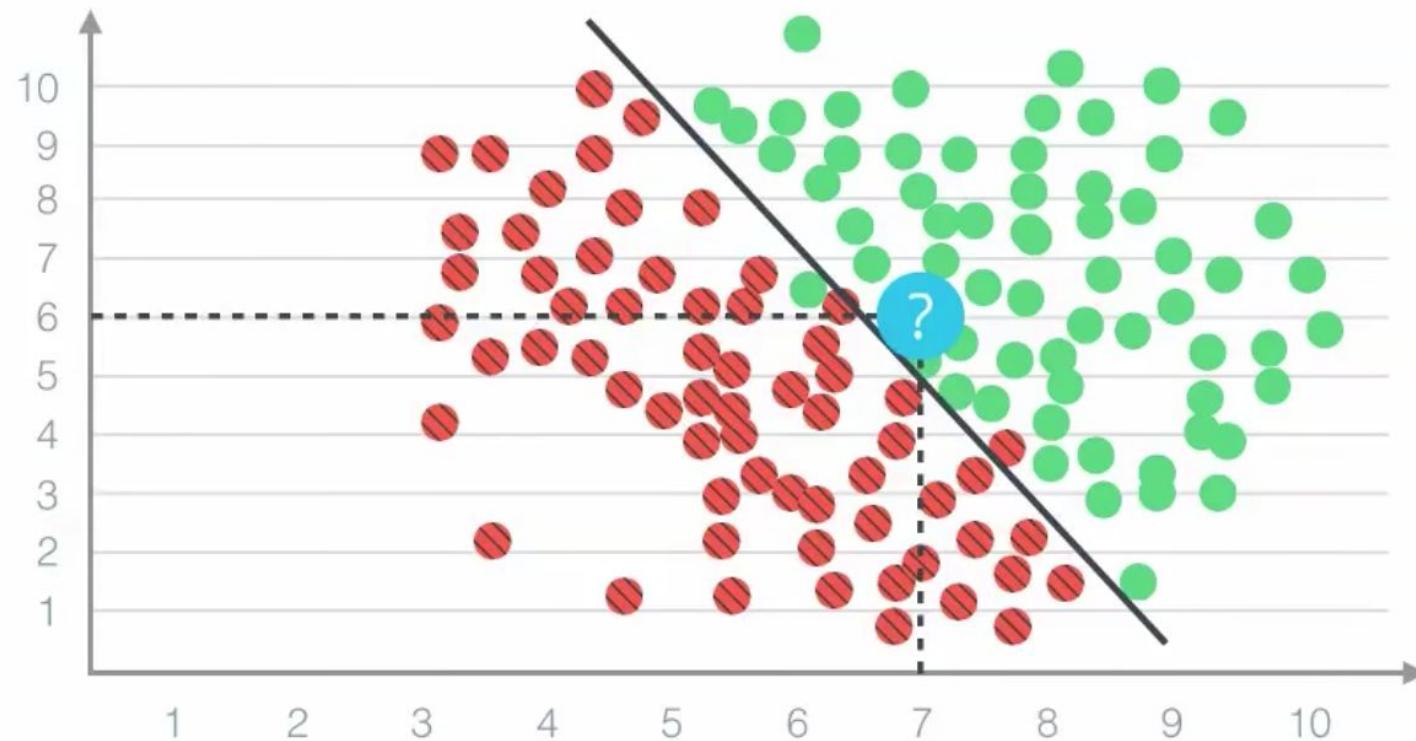
Predicción de costos de casa



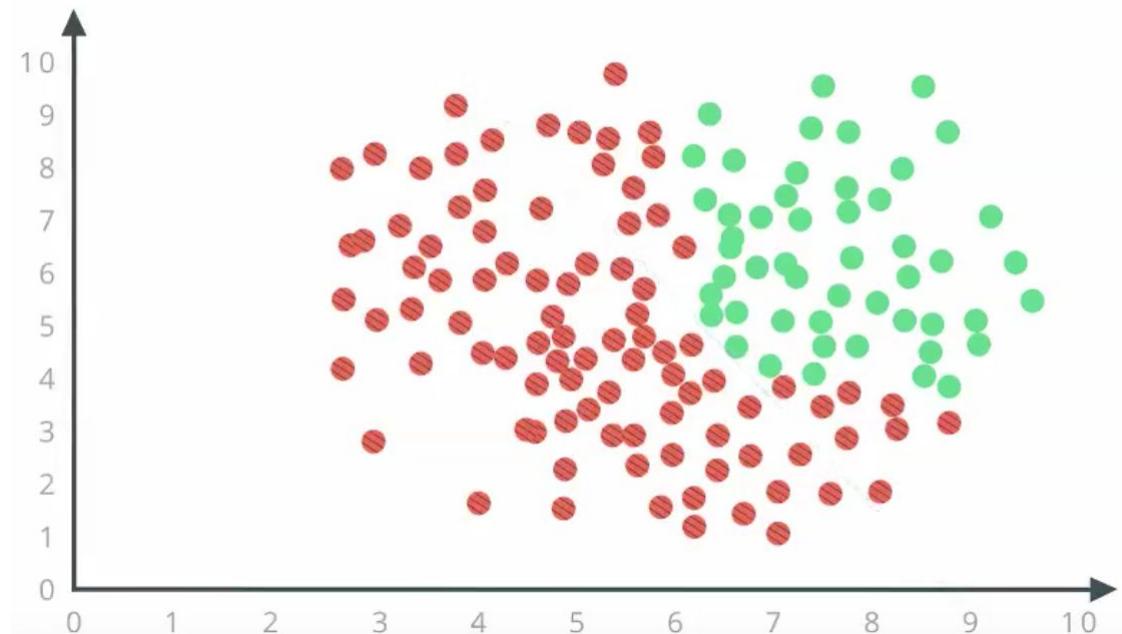
Regresión Logística



Regresión Logistica

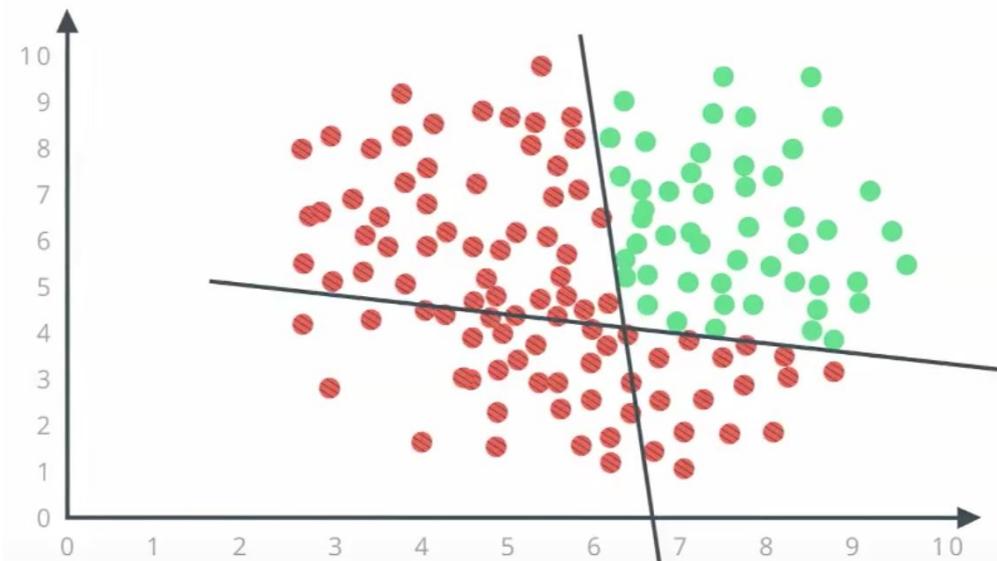


Redes Neuronales



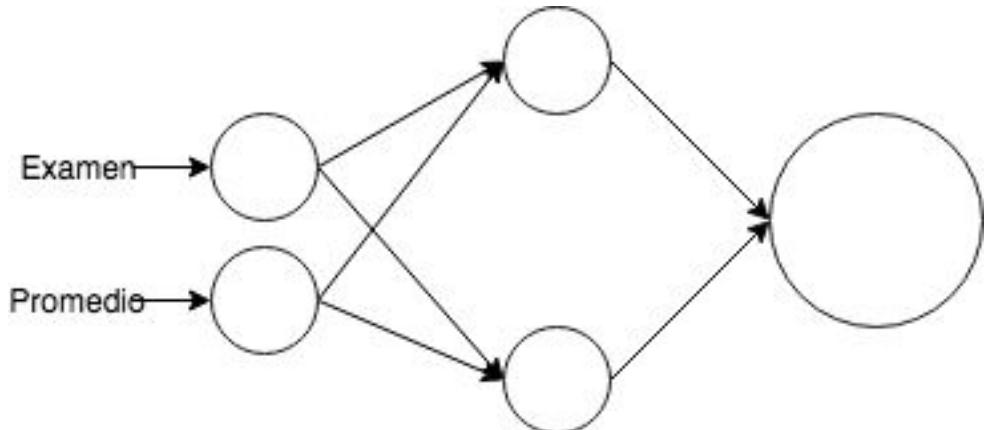
Redes Neuronales

- ¿El punto está arriba de la primera línea?
- ¿El punto está a la derecha de la otra línea?
- ¿La respuesta a las otras dos preguntas fue sí?

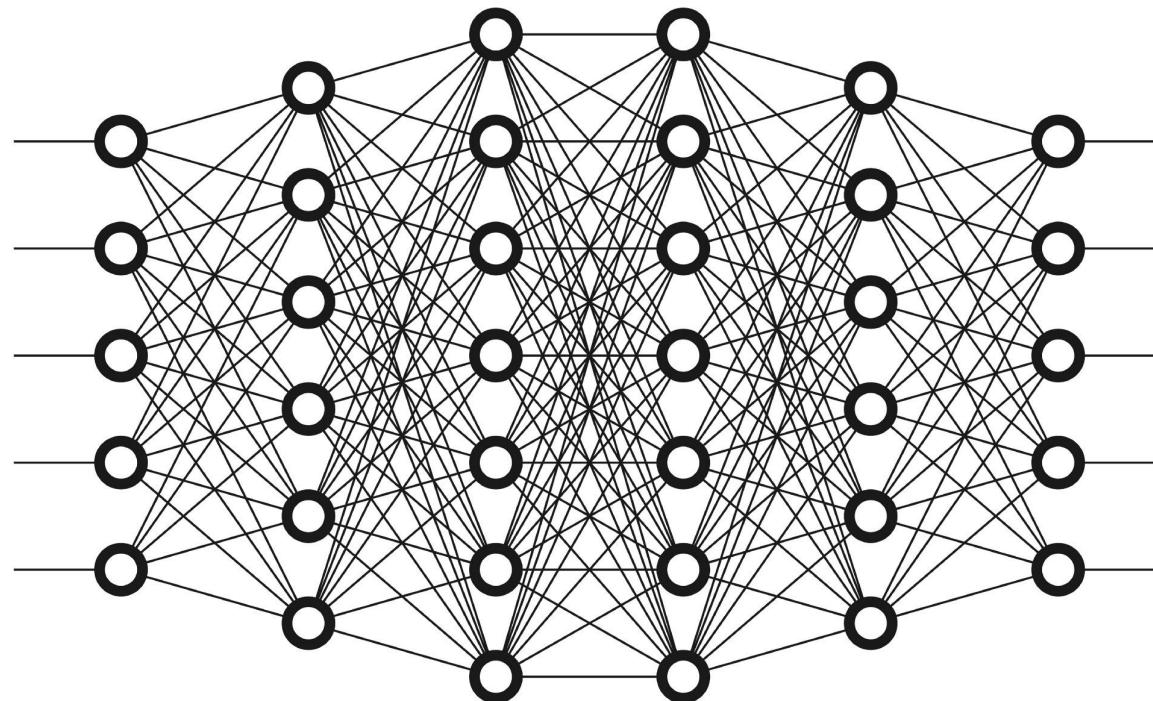


Redes Neuronales

- ¿El punto está arriba de la primera línea?
- ¿El punto está a la derecha de la otra línea?
- ¿La respuesta a las otras dos preguntas fue sí?



Deep Learning



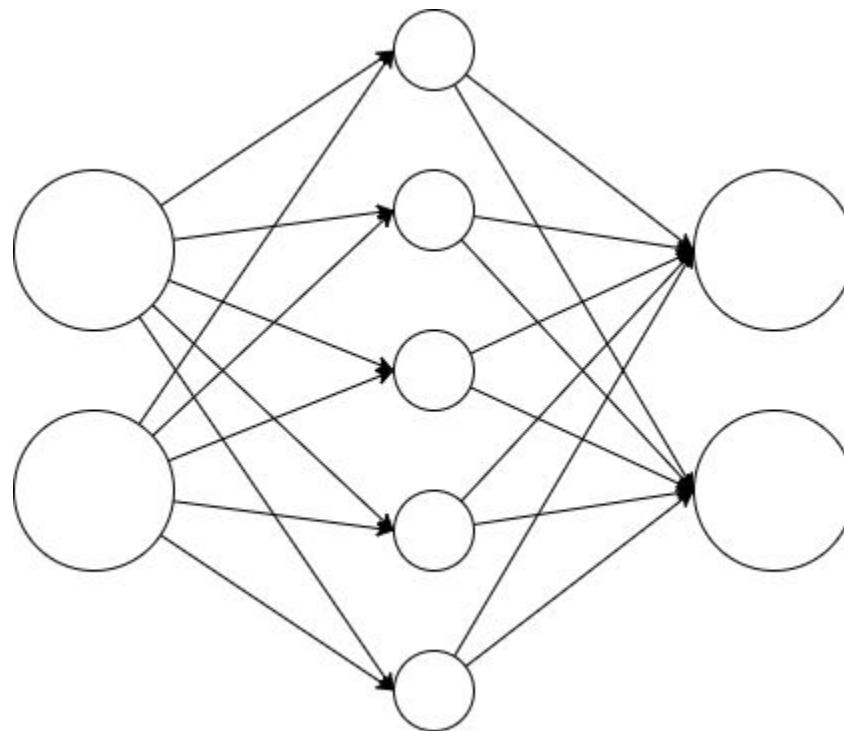
Información de entrenamiento (training set)

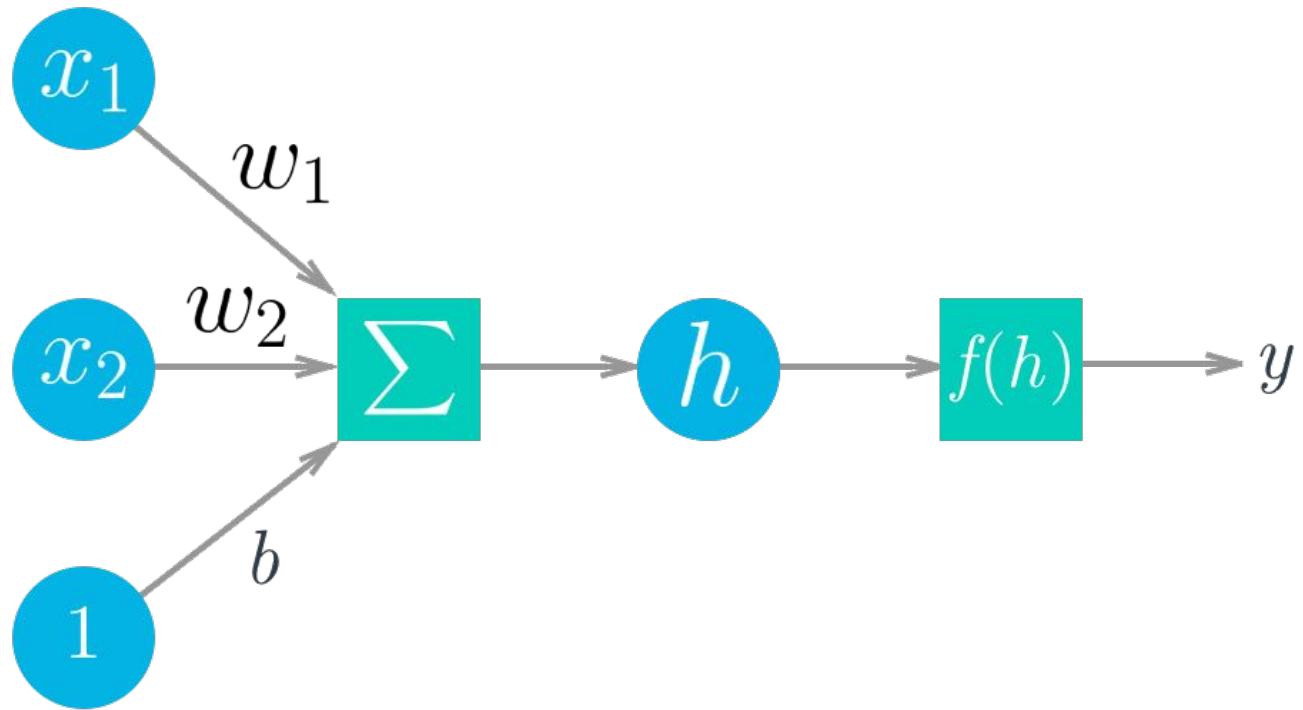
Promedio Colegio	Examen Admisión	Aceptado
9.8	8.9	1
8.7	9.2	1
6.3	7.2	0
7.5	9.3	1
9.8	7.2	0
8.8	8.4	1
8.3	8.9	1
9.5	9.4	1
7.3	6	0

Input Layer

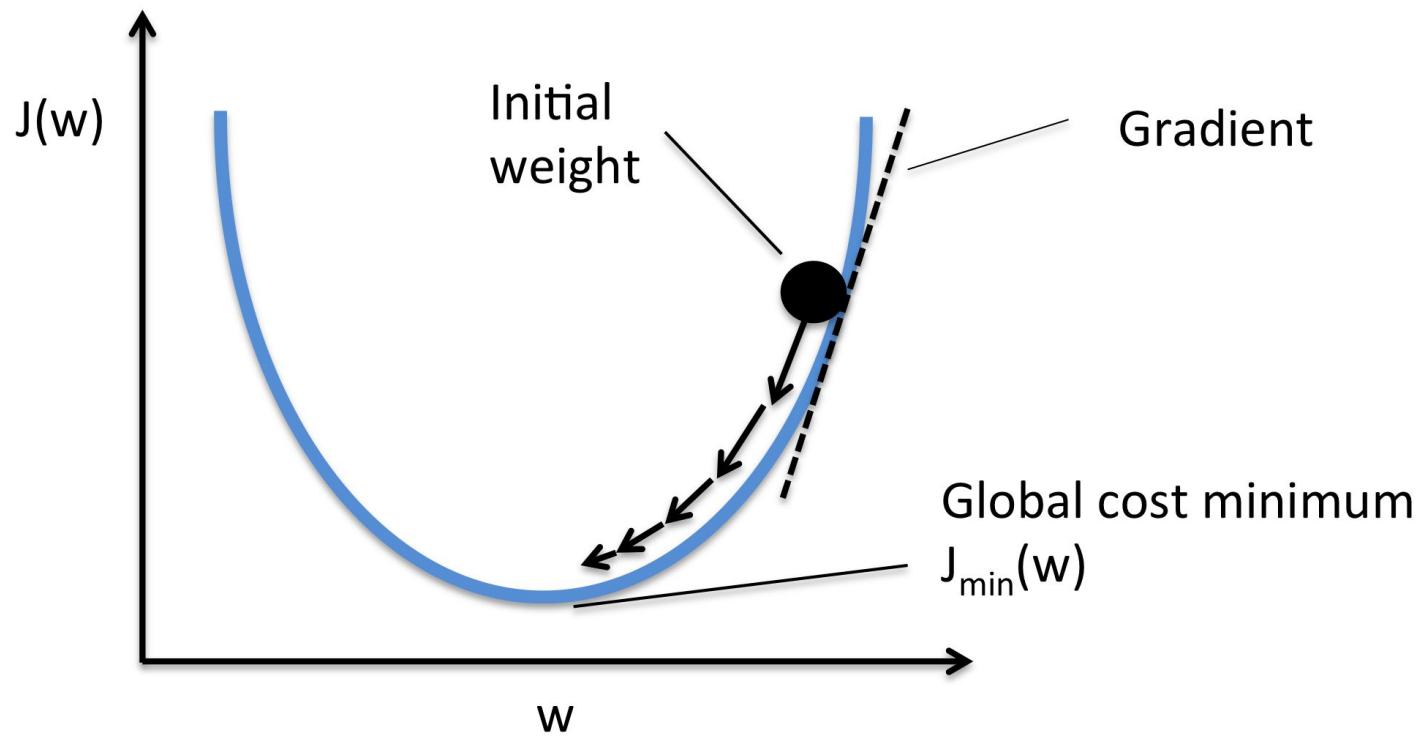
Hidden Layers

Output
Layer

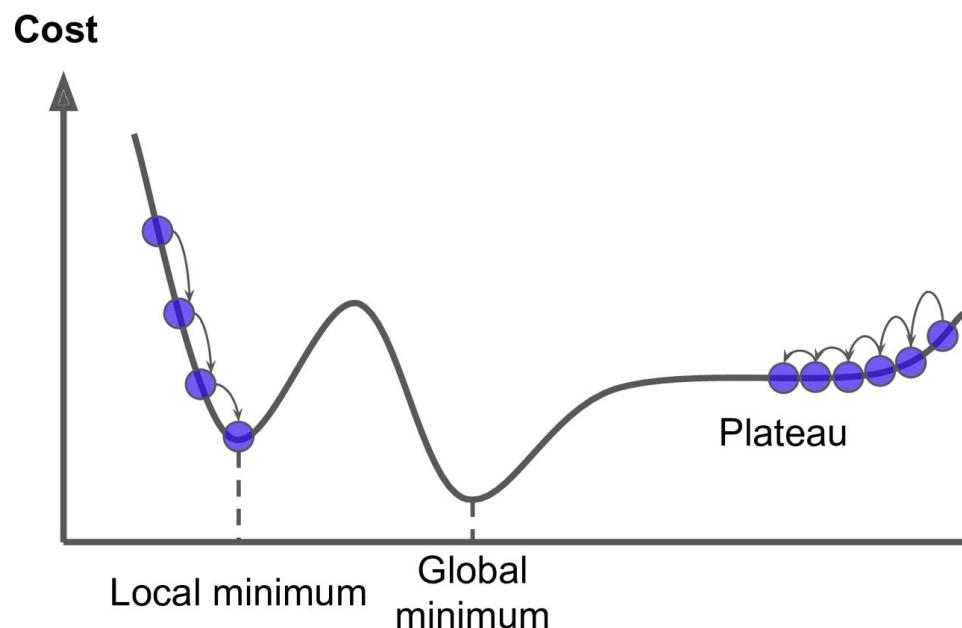




Gradient Descent



Gradient Descent



Convolutional Neural Networks

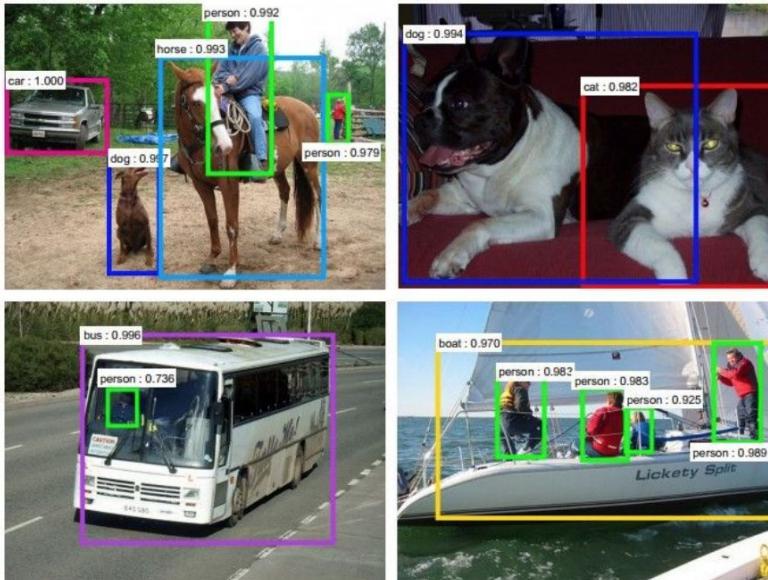
YOLO - You Only Look Once

- Estado del arte
- Sistema de detección de objetos



Detección y segmentación de objetos

Detection

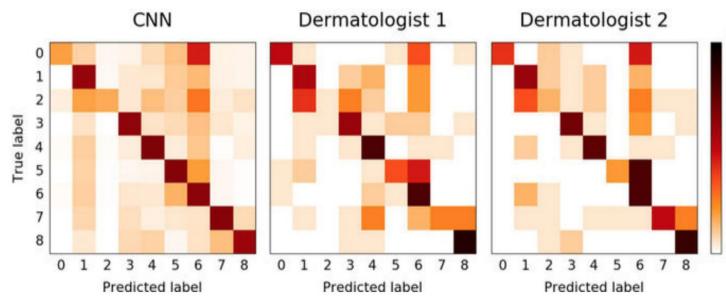
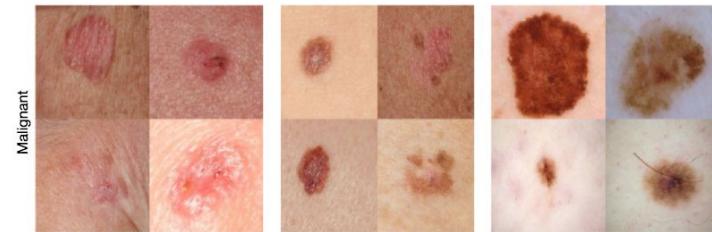
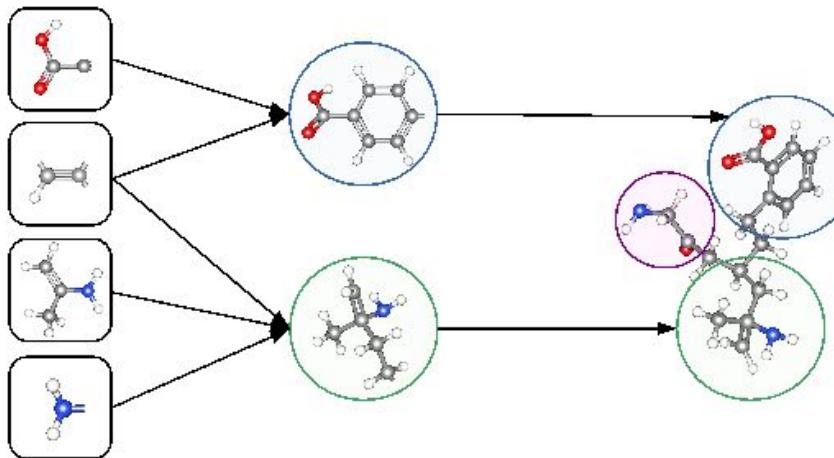


Figures copyright Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, 2015. Reproduced with permission.

[*Faster R-CNN: Ren, He, Girshick, Sun 2015*]



CNNs en medicina



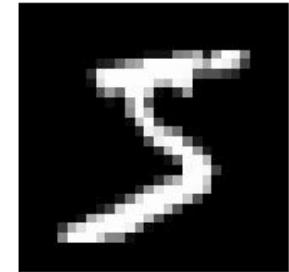
Retos de Visión por Computadora

- Visión es una de las tareas más complicadas
 - Para resolución baja hay $2^{524,288}$ imágenes posibles.
 - Y sólo existen 10^{24} estrellas en el universo.
- Variación en los objetos
 - Diferentes escalas, puntos de vista, iluminación, etc.
- Variación semántica
 - Variación en la misma clase
 - Superposición de clases.

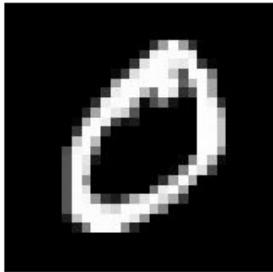


Dígitos Escritos a Mano (MNIST)

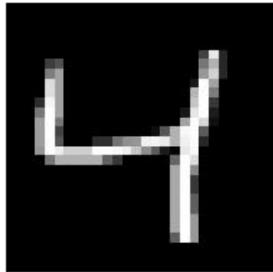
5



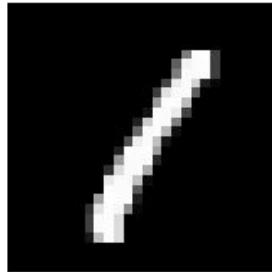
0



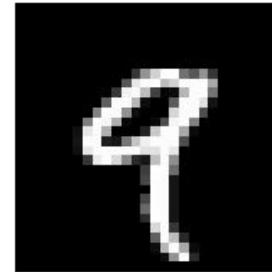
4



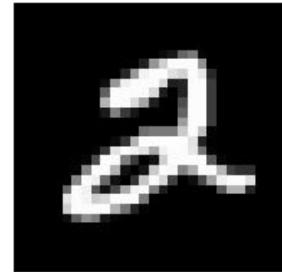
1



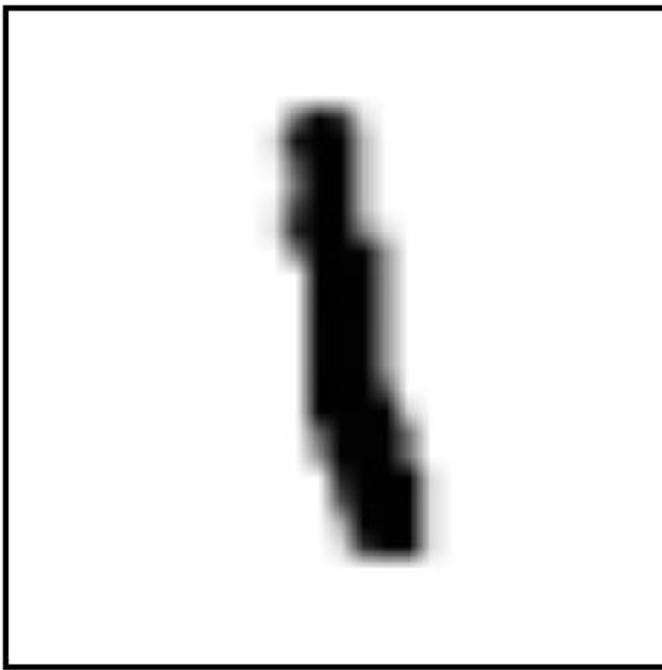
9



2



Features



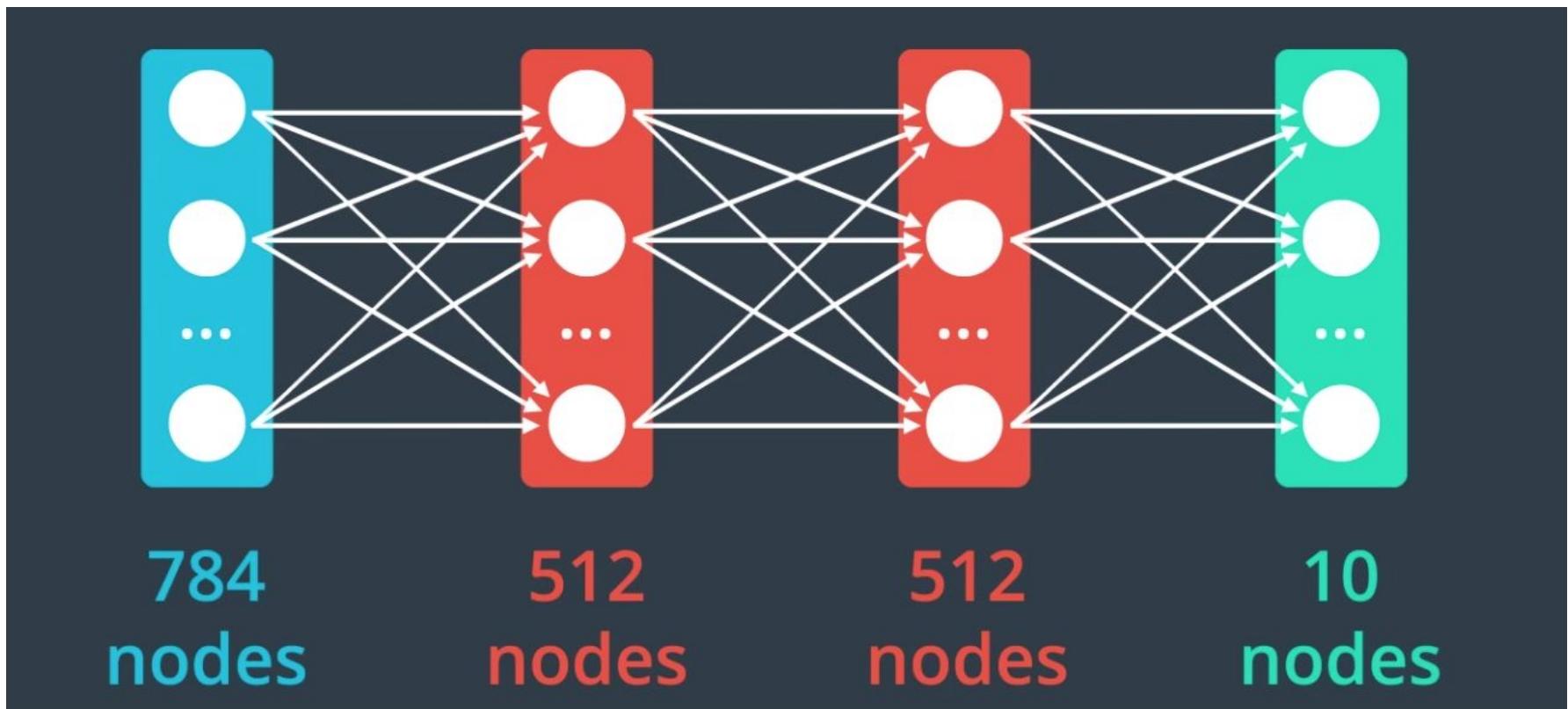
2

Con FFNN

- Se hace algo llamado flattening
- Es decir, convertimos la matriz en un gran vector

```
model = Sequential()
model.add(Flatten(input_shape=x_train.shape[1:]))
model.add(Dense(512, activation='relu'))
model.add(Dropout(0.2))
model.add(Dense(512, activation='relu'))
model.add(Dropout(0.2))
model.add(Dense(10, activation='softmax'))
```

Con FFNN

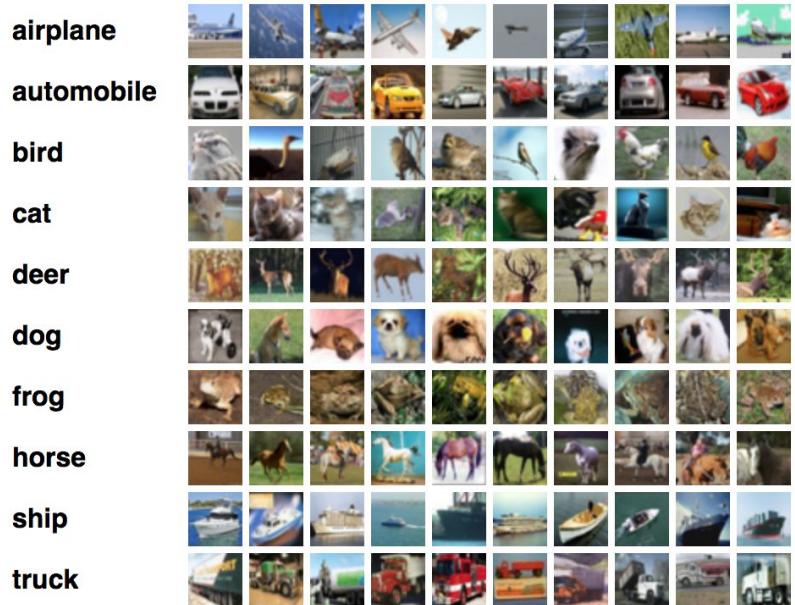


¿Por qué funciona mal?

- Un FFNN puede tener 98% de precisión con MNIST, pero un CNN llega a 99.79%.

¿Qué otros problemas hay además de MNIST?

- CIFAR 10 (96.53% 2015)
 - 10 clases con 6,000 imágenes de color de 32x32 cada una



¿Qué otros problemas hay además de MNIST?

- CIFAR 10 (96.53% 2015)
 - 10 clases con 6,000 imágenes de color de 32x32 cada una
- CIFAR 100 (75.72% 2015)
 - 100 clases con 600 imágenes cada una
 - 20 superclases con 5 clases cada una

Superclass
aquatic mammals
fish
flowers
food containers
fruit and vegetables
household electrical devices
household furniture
insects
large carnivores
large man-made outdoor things
large natural outdoor scenes
large omnivores and herbivores
medium-sized mammals
non-insect invertebrates
people
reptiles
small mammals
trees
vehicles 1
vehicles 2

Classes
beaver, dolphin, otter, seal, whale
aquarium fish, flatfish, ray, shark, trout
orchids, poppies, roses, sunflowers, tulips
bottles, bowls, cans, cups, plates
apples, mushrooms, oranges, pears, sweet peppers
clock, computer keyboard, lamp, telephone, television
bed, chair, couch, table, wardrobe
bee, beetle, butterfly, caterpillar, cockroach
bear, leopard, lion, tiger, wolf
bridge, castle, house, road, skyscraper
cloud, forest, mountain, plain, sea
camel, cattle, chimpanzee, elephant, kangaroo
fox, porcupine, possum, raccoon, skunk
crab, lobster, snail, spider, worm
baby, boy, girl, man, woman
crocodile, dinosaur, lizard, snake, turtle
hamster, mouse, rabbit, shrew, squirrel
maple, oak, palm, pine, willow
bicycle, bus, motorcycle, pickup truck, train
lawn-mower, rocket, streetcar, tank, tractor

¿Qué otros problemas hay además de MNIST?

- CIFAR 10 (96.53% 2015)
 - 10 clases con 6,000 imágenes de color de 32x32 cada una
- CIFAR 100 (75.72% 2015)
 - 100 clases con 600 imágenes cada una
 - 20 superclases con 5 clases cada una
- STL 10 (74.33% 2015)
 - Similar a CIFAR 10: imágenes de 96x96
 - 500 imágenes con sus labels
 - 100,000 imágenes sin labels - aprendizaje no supervisado
- SVHN (98.31% 2016)
 - Números en ambientes reales
 - Dígitos de casas de Street View

¿Por qué funciona mal?

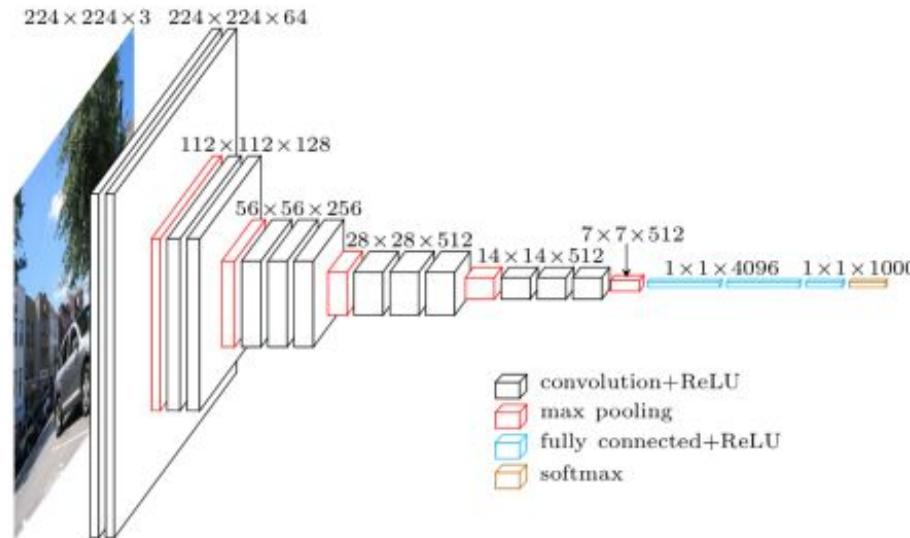
- Un FFNN puede tener 98% de precisión con MNIST, pero un CNN llega a 99.79%.
- Sólo usamos redes completamente conectadas
 - Millones de parámetros, por lo cual es muy lento
- Perdemos información dimensional por hacer flattening

Convolutional Neural Networks (CNN)

- Entiende información espacial
- No aplana imágenes
- Es muy similar en términos generales de un FFNN

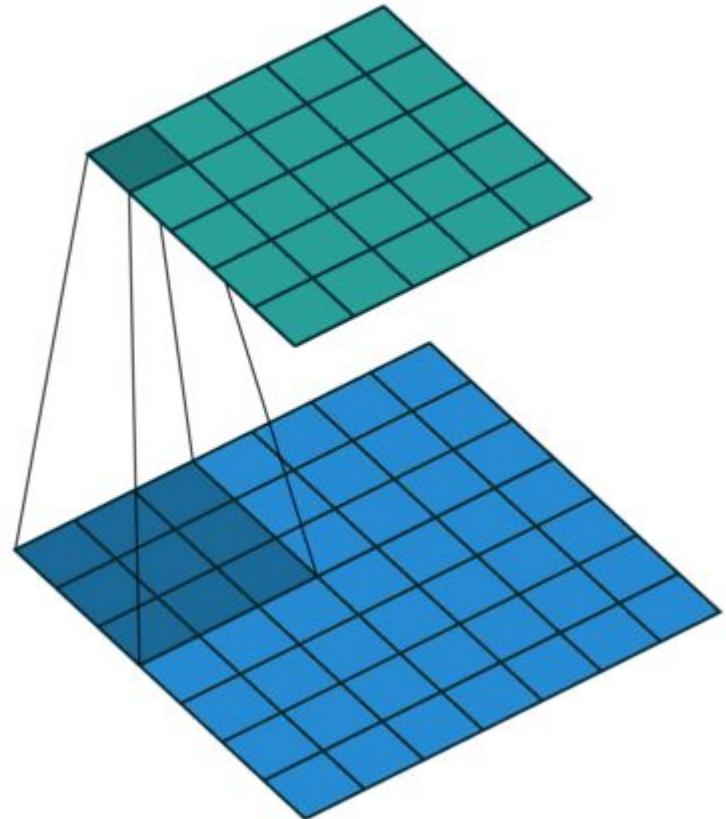
Convolutional Neural Network

- Vamos reduciendo dimensiones pero agregando canales hasta tener un gran vector



CNN

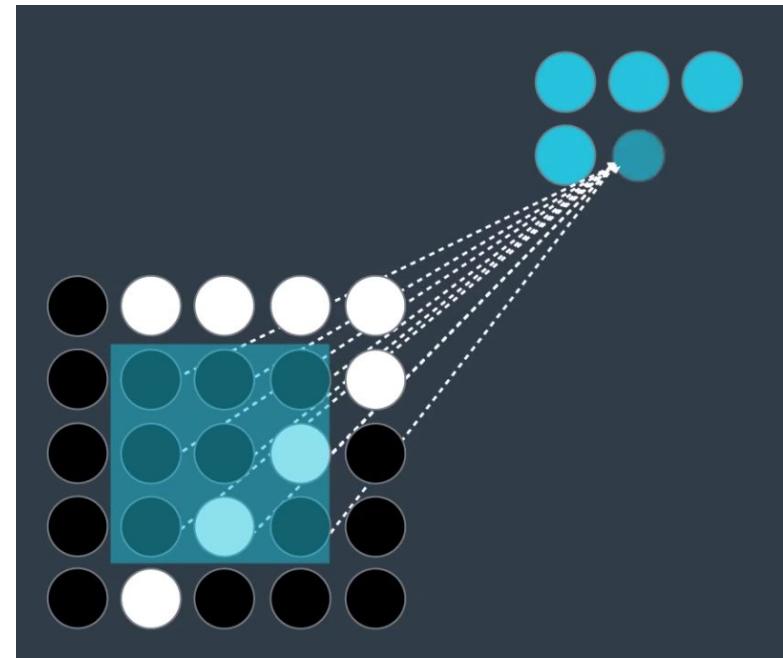
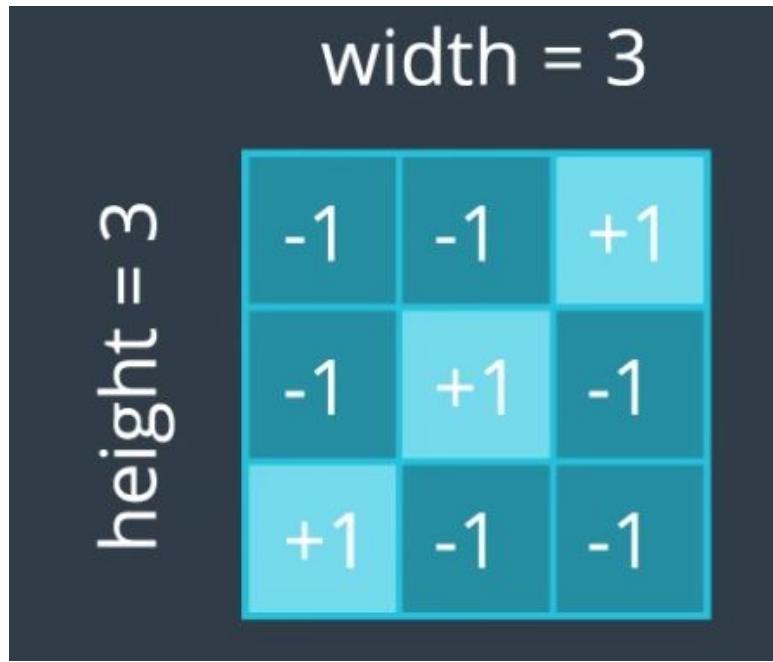
CNN



CNN



Filtro





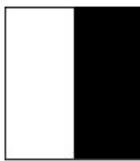
Filter 1



Activation Map for Filter 1



Filter 2



Activation Map for Filter 2



Filter 3



Activation Map for Filter 3



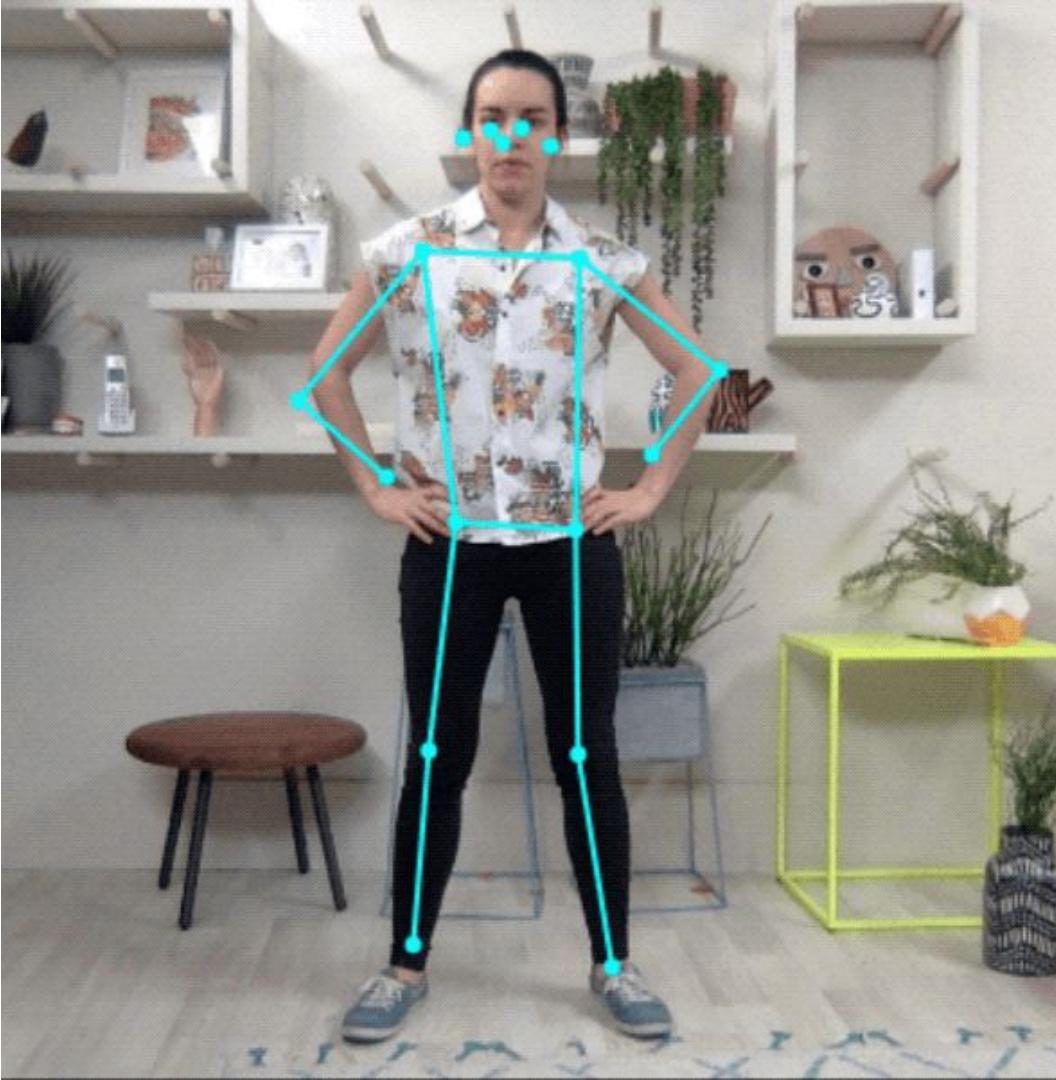
Filter 4



Activation Map for Filter 4



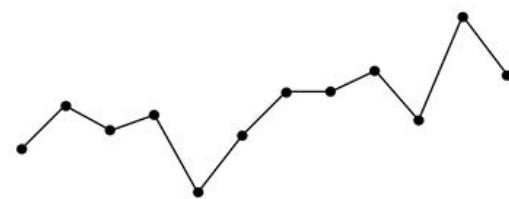
PoseNet



Recurrent Neural Networks

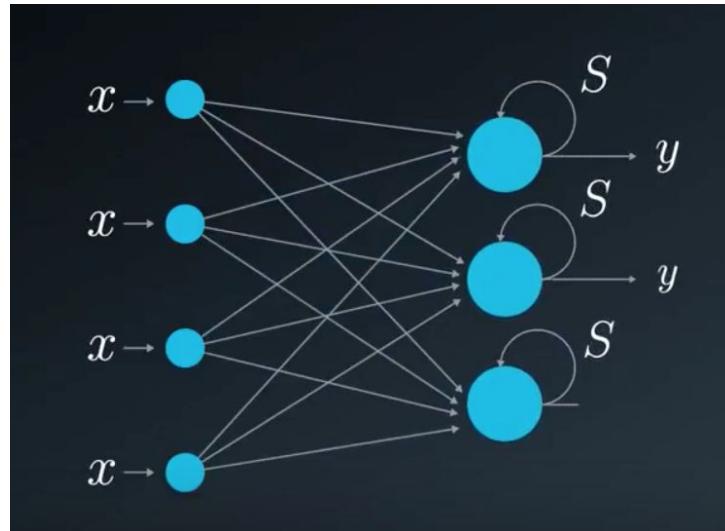
RNN

- Tienen memoria
- Dependencia sobre el tiempo
- No sólo usamos el input actual, sino el input previo
- Ejemplos:
 - Completar secuencias
 - Distinguir acciones de videos



RNN

- Recurrent: Ocurre repetidamente
- Hacemos la misma tarea en la secuencia de entrada múltiples veces



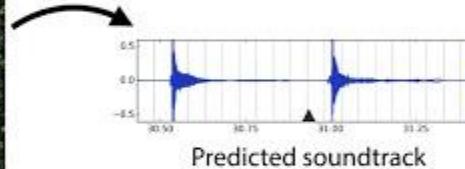
RNN

- FNN (1980s)
 - Sin dependencias temporales
- Las neuronas biológicas son recurrentes
- Time Delay Neural Network (TDDN) (1989)
 - Limitado a un poco tiempo
- Elman Networks - RNN simples (1990)
- Estos tienen problema de vanishing gradient
 - Los features cada vez tenían menor impacto en la red (derivadas más pequeñas)
- Long Short-Term Memory (LSTM) (1990s)
 - Resuelve este problema
- Variaciones
 - GRU - Gated Recurrent Units

Aplicaciones - Agregando Sonido



Silent video



Predicted soundtrack

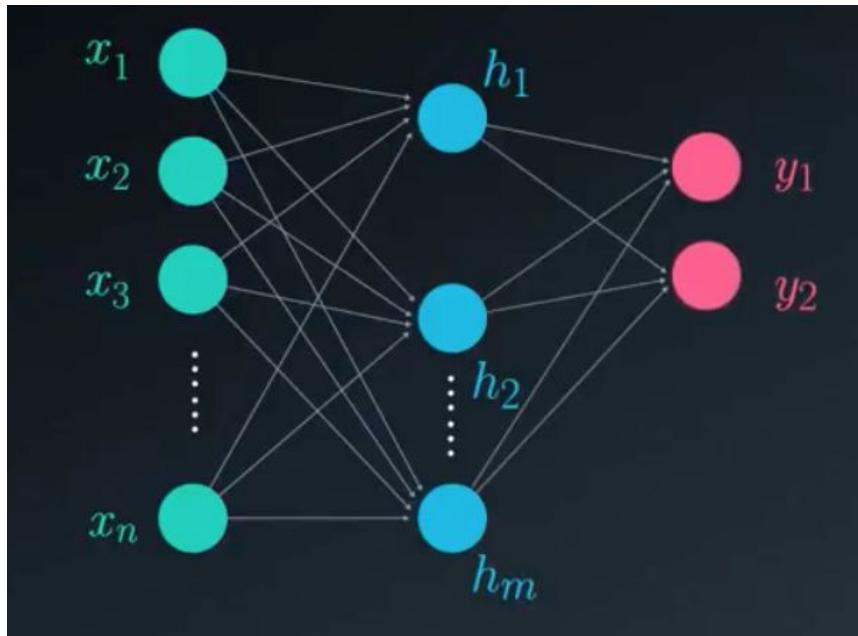
Aplicaciones - Generar Texto

dogs are way better than cats

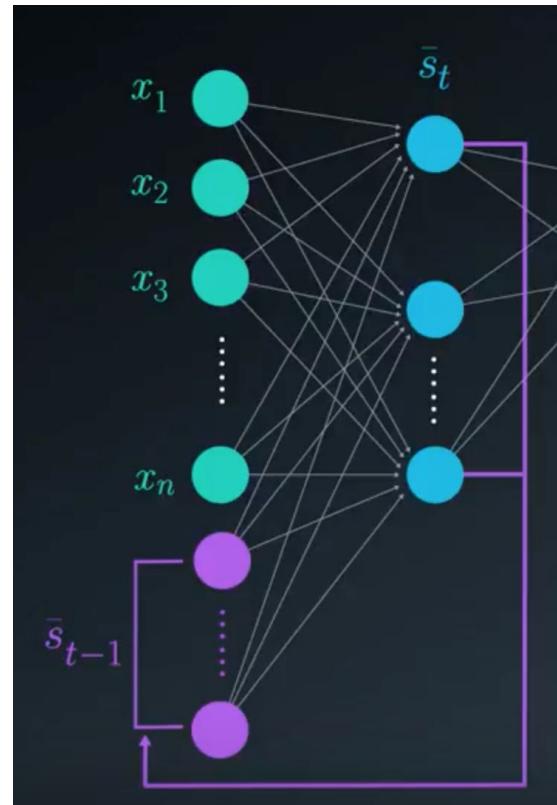
Aplicaciones

- Amazon Lex
 - Voz a texto
 - Speech Recognition (Siri, Google Assistant, etc.)
- Facebook
 - Modelado de lenguajes
- Netflix
 - Sistemas de recomendación
- Predicciones de series
 - Tráfico (Waze)
 - Acciones de bolsas y movimientos
- Natural Language Processing (NLP)
 - Traducción
 - Responder preguntas
 - Chatbots
- Y más

FNN

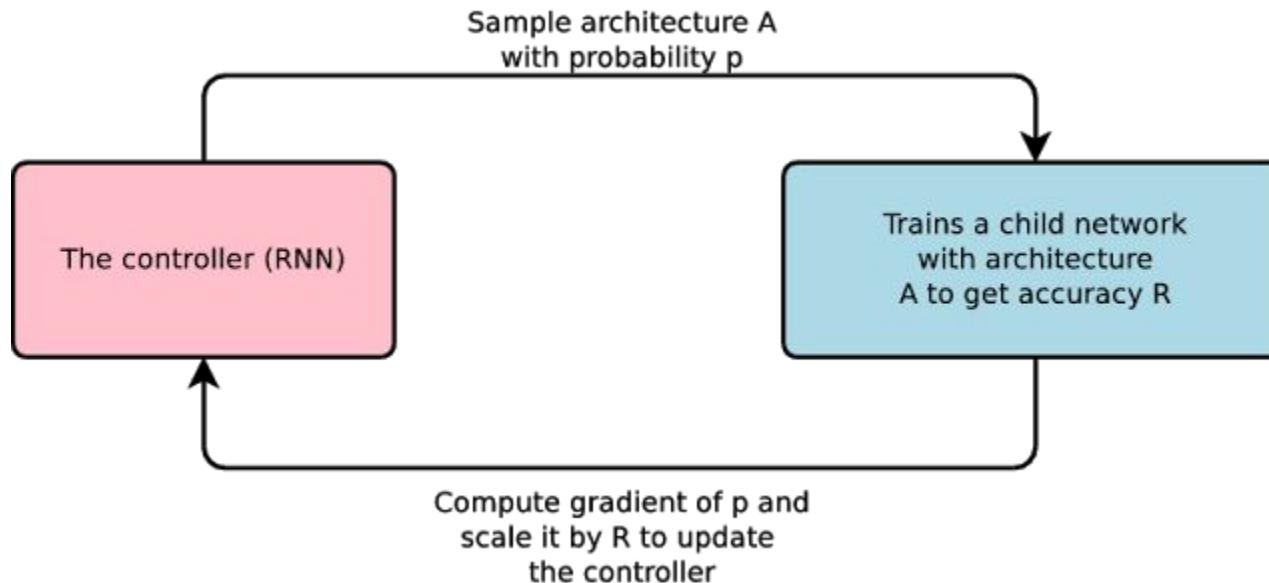


RNN

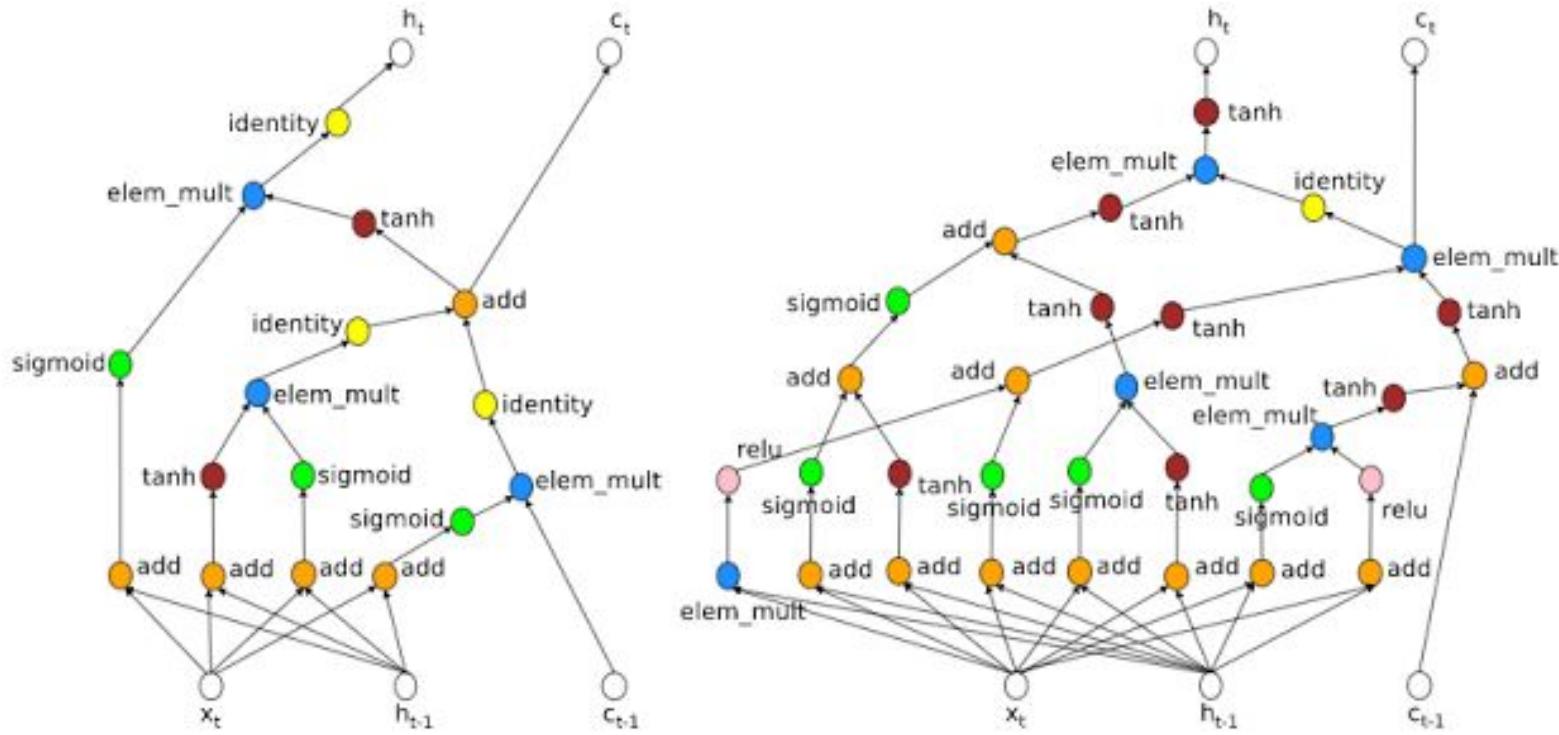


AutoML

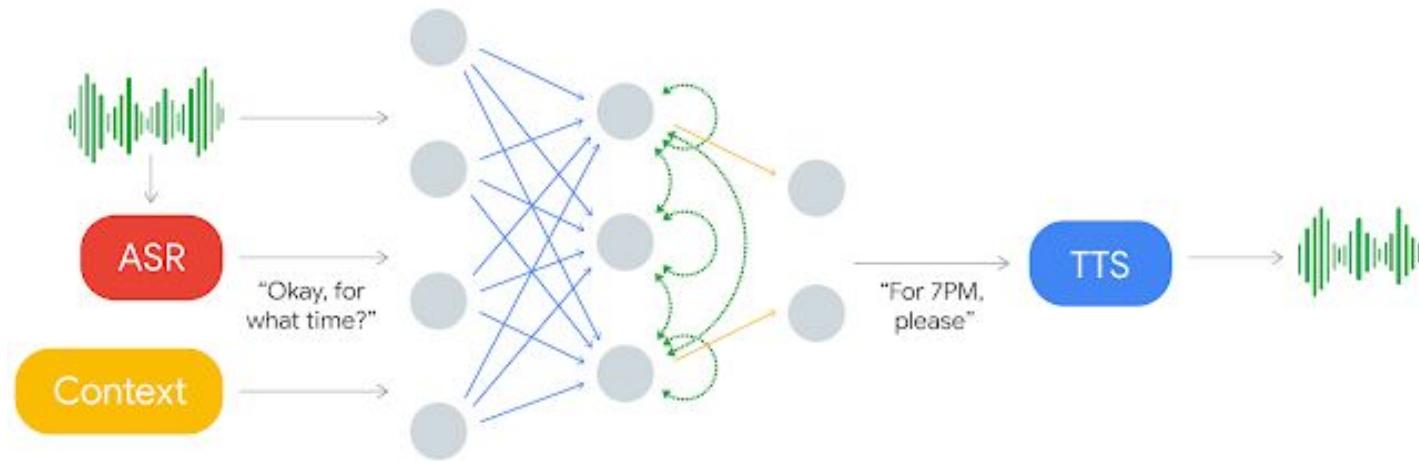
- Un modelo de 10 capas tiene 10^{10} redes candidatas.
- Diseñar la arquitectura manualmente puede ser doloroso.



AutoML



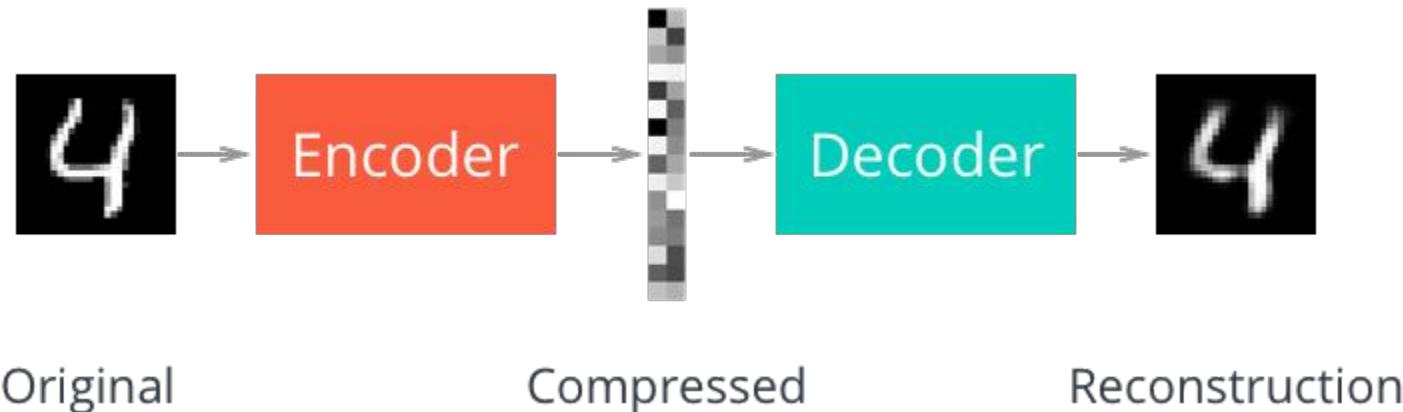
Duplex



Autoencoders

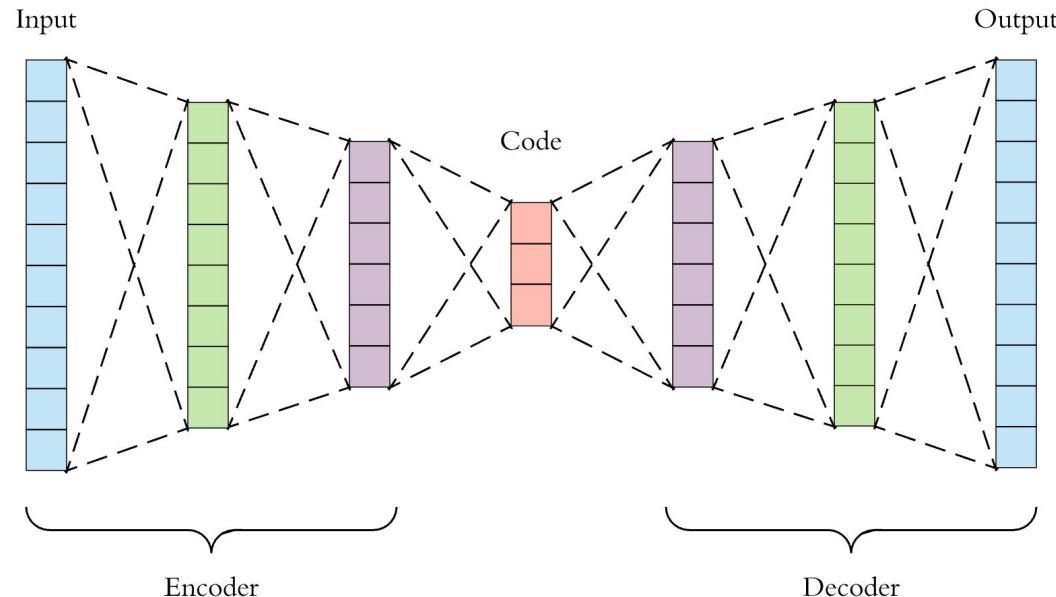
Autoencoders (2008, 2009, 2011, 2013)

- Modelos usados para comprimir la información
- Se crea una modelo fino que permite representar la información de manera más sencilla.

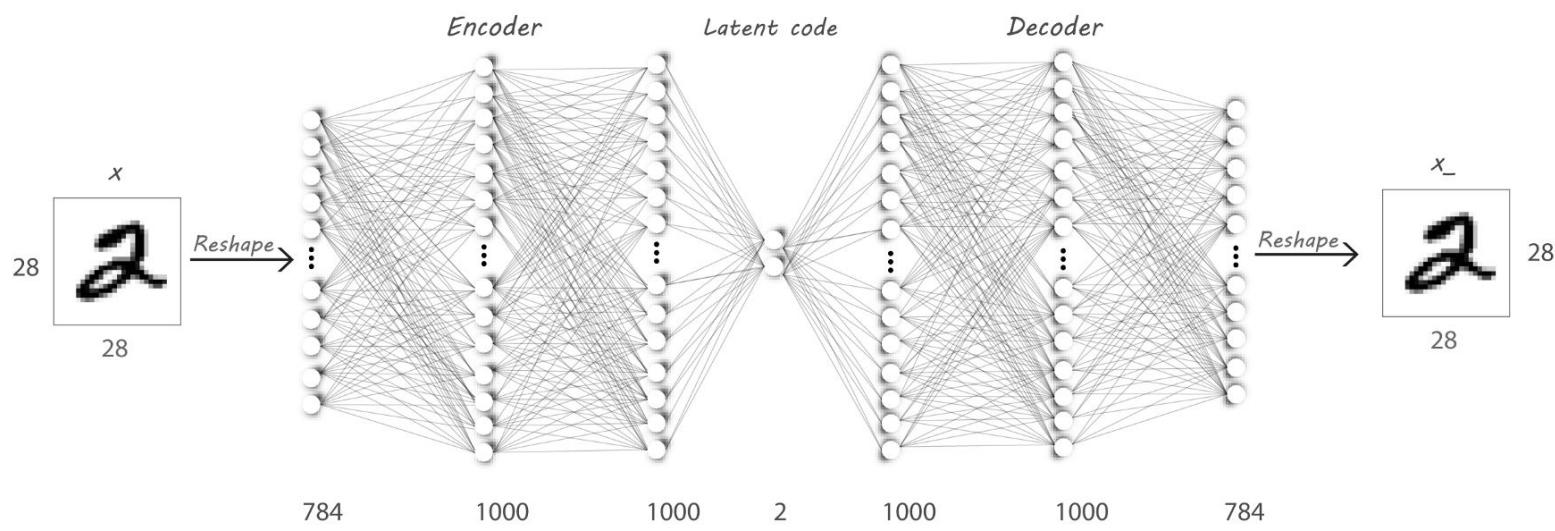


Autoencoders

- Deep compression - se encarga el modelo, no lo definimos nosotros
- Tenemos **encoder** y **decoder**, ambos son redes neuronales



Autoencoders



Generative Adversarial Networks

Generative Adversarial Network

- Propuesta por Ian Goodfellow en el 2014
- Pueden imaginar/generar nuevas cosas

StackGAN

- Recibe una descripción y genera imágenes en alta resolución
- El modelo que las genera nunca ha visto las fotos

StackGAN
Stage-I
64x64
images

This bird is white with some black on its head and wings, and has a long orange beak

This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments



StackGAN
Stage-II
256x256
images



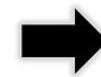
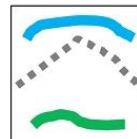
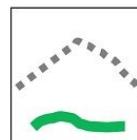
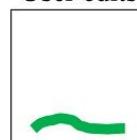
Vanilla GAN
256x256
images



iGAN



User edits



Generated images



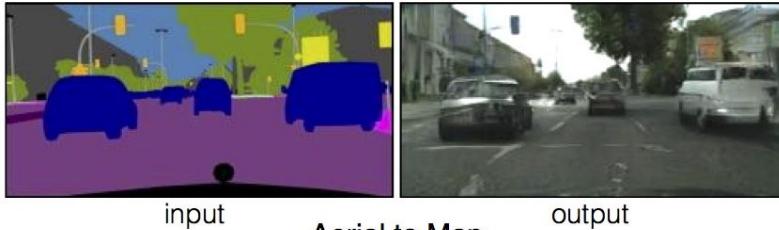
— Color

— Sketch

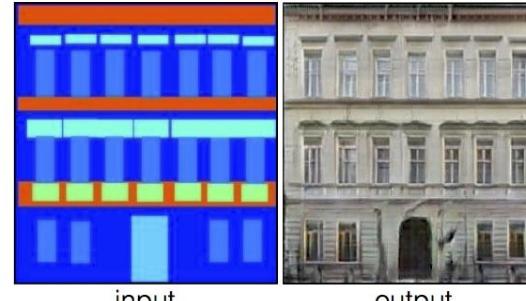
pix2pix (<https://affinelayer.com/pixsrv/>)

- Traduce imagen a imagen en el mismo dominio

Labels to Street Scene



Labels to Facade



BW to Color



Aerial to Map



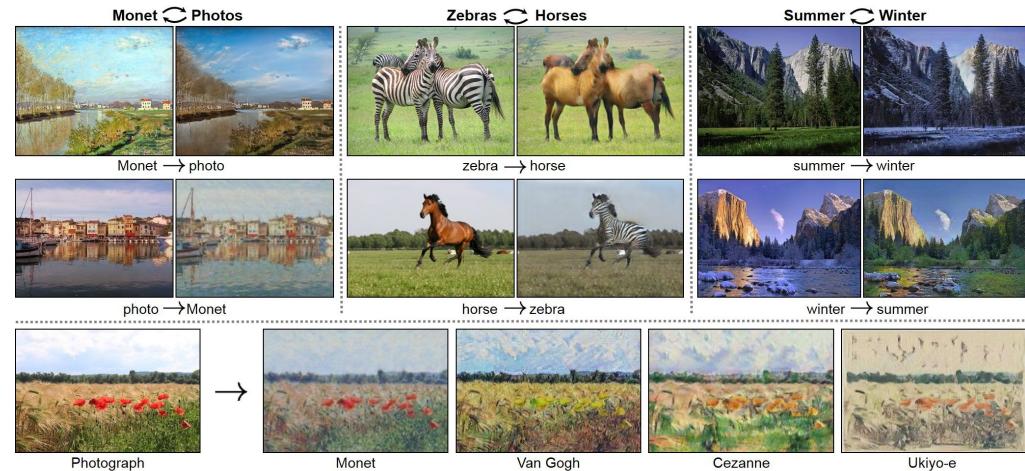
Day to Night



Edges to Photo



CycleGAN - imagen a imagen



CycleGAN - Cambio de temporada



winter Yosemite → summer Yosemite

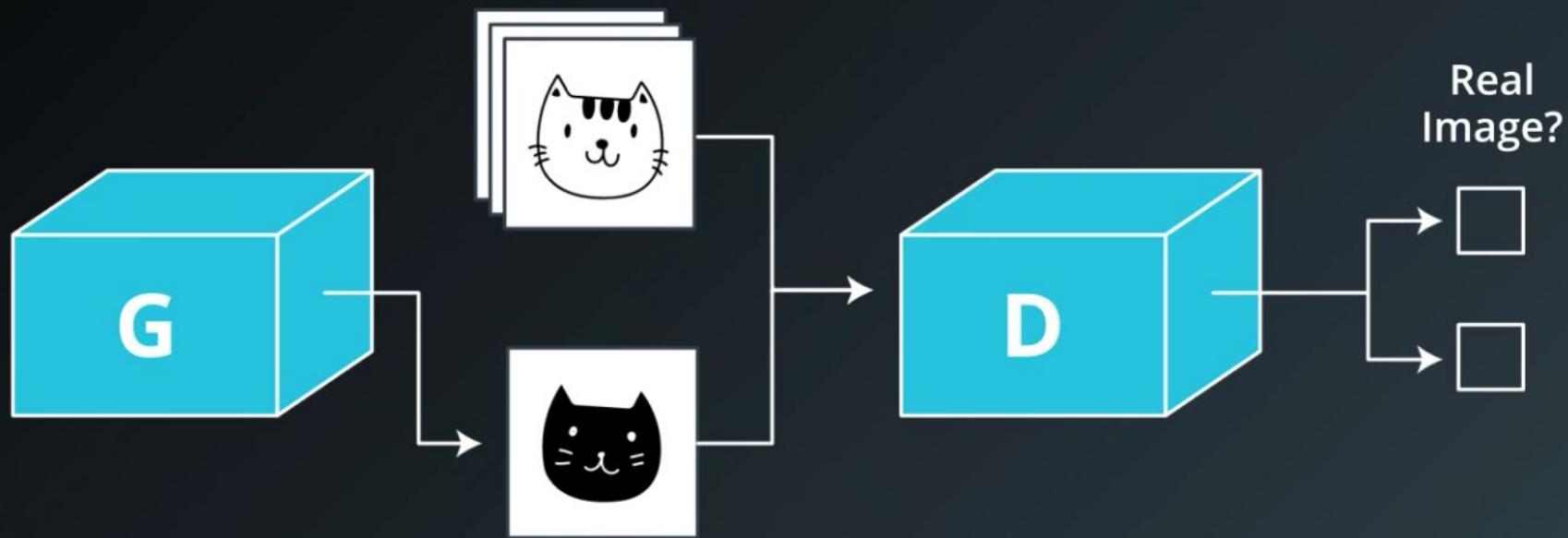


summer Yosemite → winter Yosemite

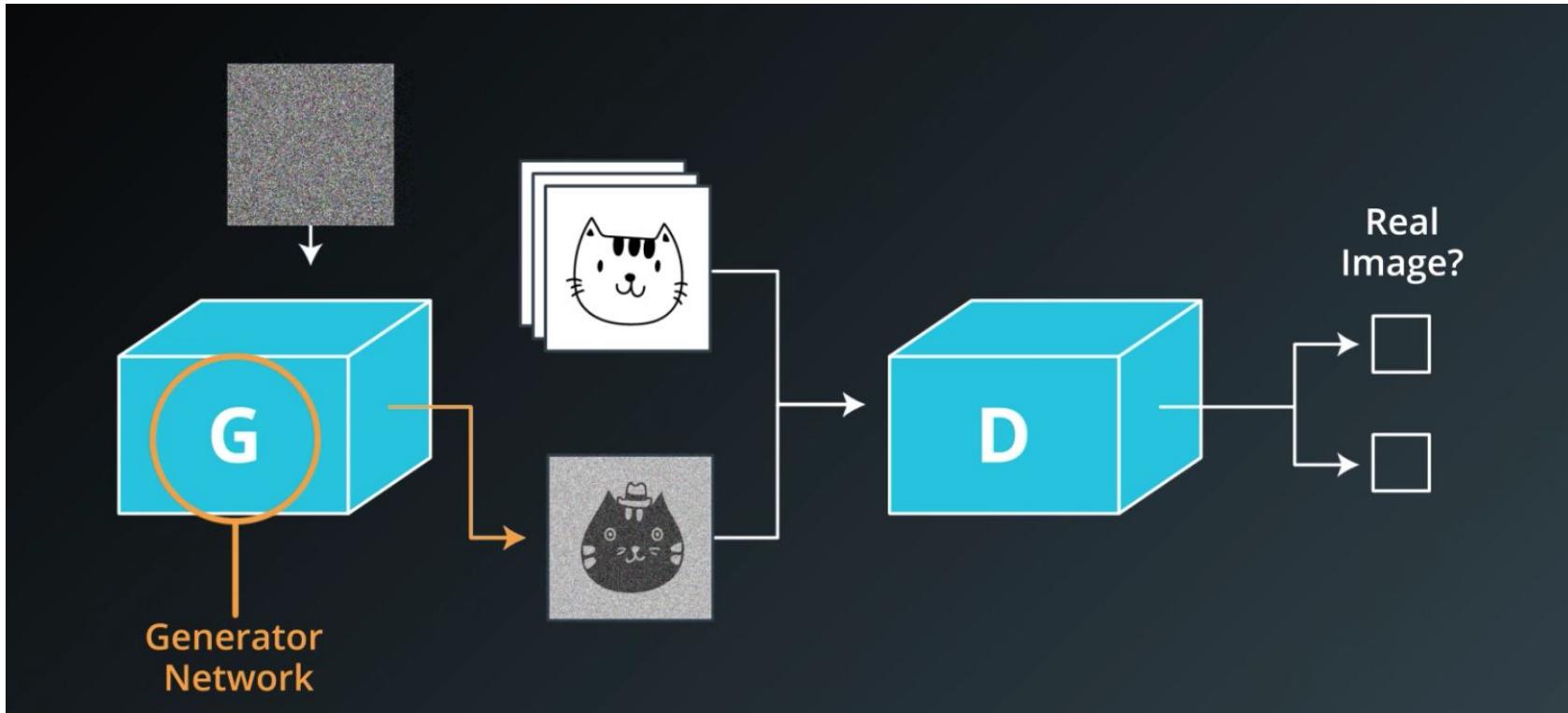
¿Cómo funciona? <https://arxiv.org/pdf/1406.2661.pdf>

- Ya vimos cómo generar con RNNs - generaban una palabra a la vez
- Podemos hacer lo mismo con imágenes - generamos un pixel a la vez
- ¿Cómo generamos una oración o una imagen completa de una vez?

¿Cómo funciona? <https://arxiv.org/pdf/1406.2661.pdf>

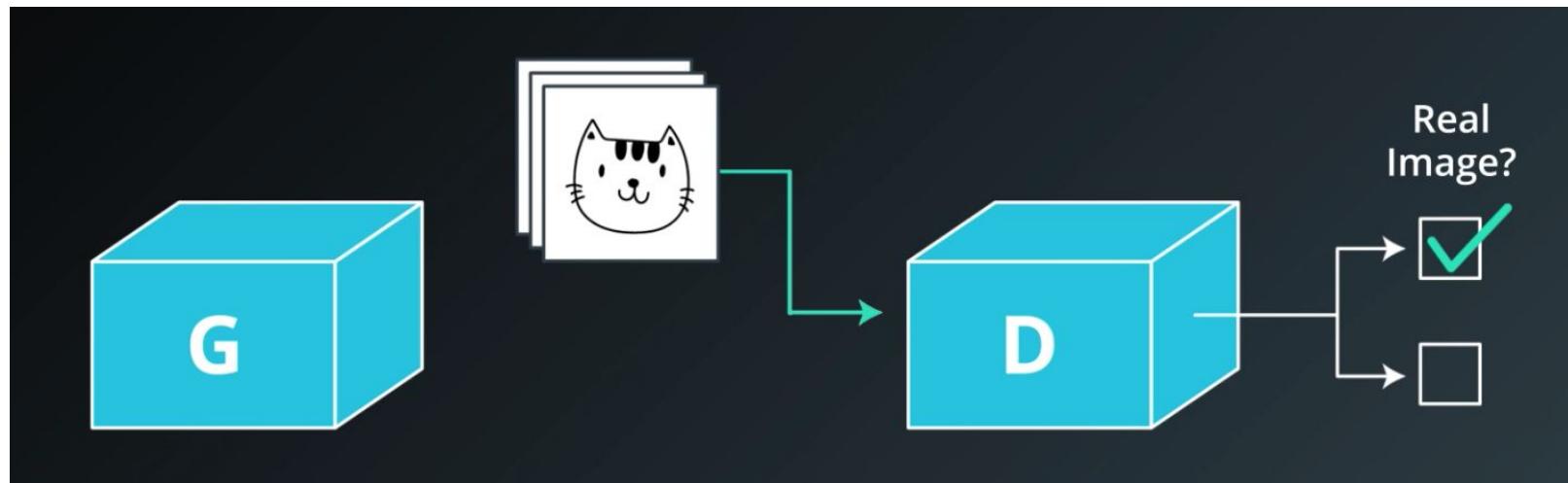


¿Cómo funciona? <https://arxiv.org/pdf/1406.2661.pdf>



¿Cómo funciona? <https://arxiv.org/pdf/1406.2661.pdf>

- No es supervisado - no hay labels
- Vemos muchas imágenes y pedimos que genere imágenes de la misma distribución
- Discriminator dirige al generator



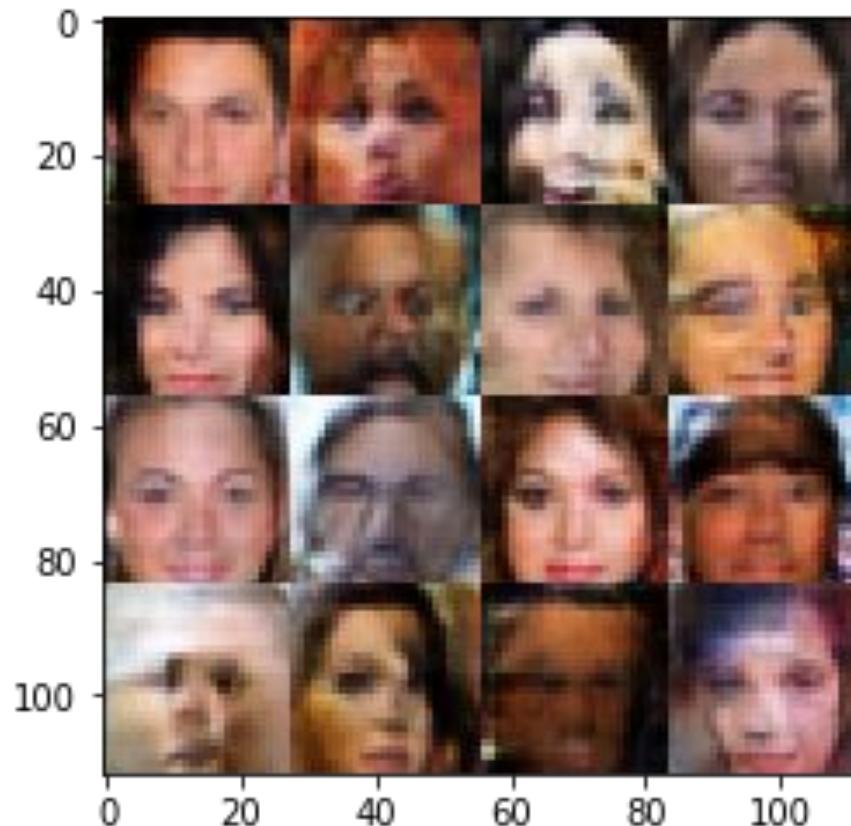
GAN

- Discriminador - intenta asignar un 1 si la imagen es real y 0 si no es real
- Generador - intenta que las imágenes generadas sean consideradas reales - intenta engañar al discriminador
- Generador tiene que hacer imágenes realistas

Generación de caras de Anime



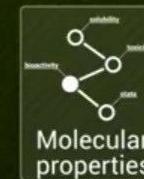
Generación de caras reales



Drug Database



discriminator



Disease

Generative Adversarial Networks



INSILICO MEDICINE

Siguientes pasos

- Intro teórica a DL: <https://www.deeplearning.ai/>
- Curso práctico de DL: <http://www.fast.ai/>
- DL para coches autónomos: <https://selfdrivingcars.mit.edu/>
- DL para NLP: <http://cs224d.stanford.edu/>
- DL para CV: <http://cs231n.stanford.edu/>
- TensorFlow
 - [https://www.tensorflow.org/get started/](https://www.tensorflow.org/get_started/)
 - <https://www.tensorflow.org/tutorials/>
 - https://www.tensorflow.org/programmers_guide/
- Proyectos reales

osanseviero@gmail.com