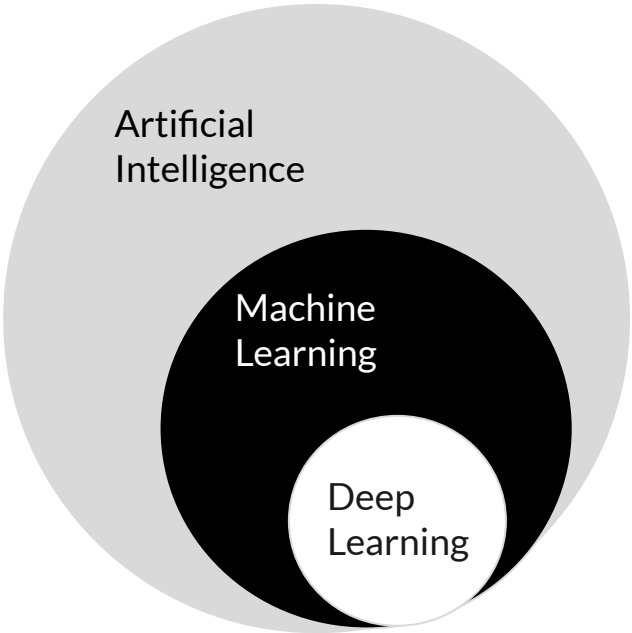
Intro to Deep Learning

Chris Powell
Software Engineer at Hudl

What I'll be talking about today

1. What is 'Deep Learning'?
2. How does a neural network work?
3. What are a few examples of Deep Learning models and why would I use them?
4. What are the downsides to choosing a neural network for my next machine learning project?

Deep Learning is a subset of Machine Learning

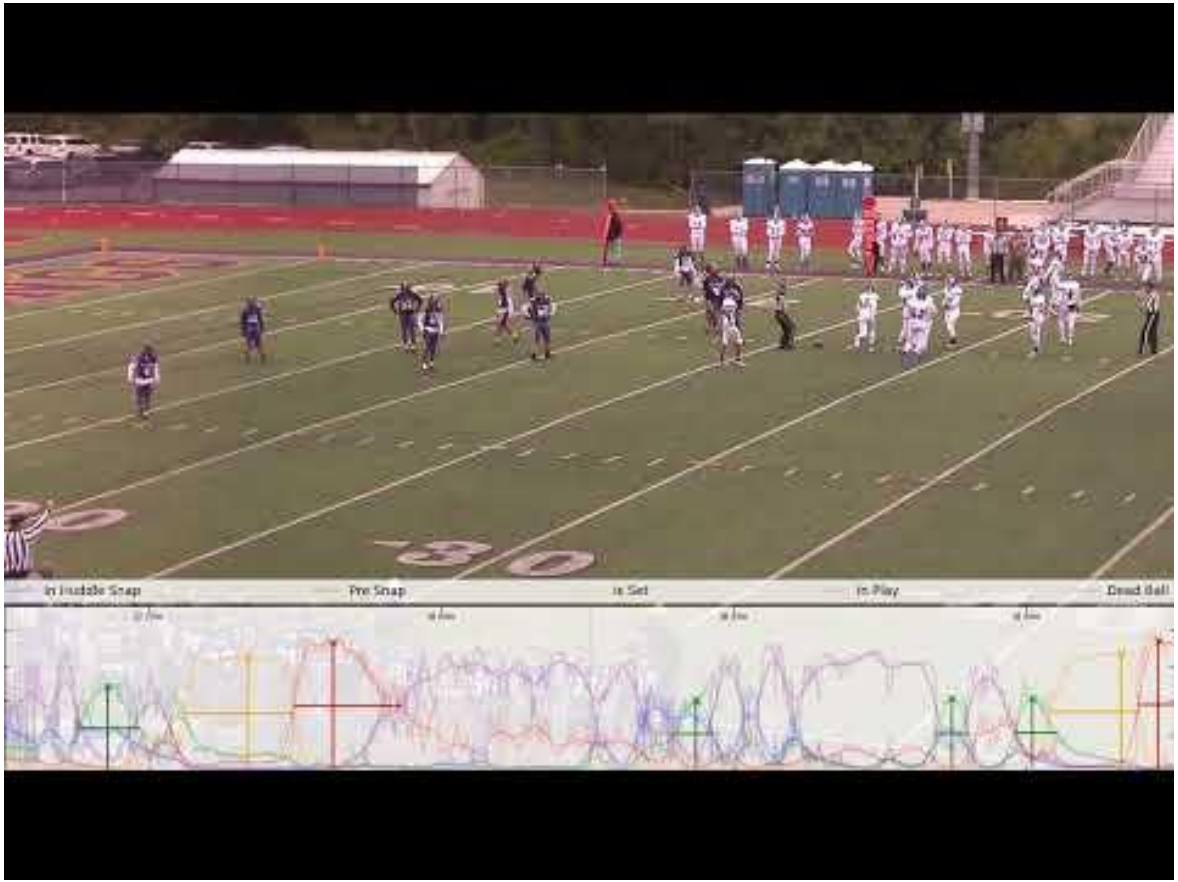


Examples of Deep Learning

“Frame Oracle” →

A series of processes that produce player information, that includes a Deep Learning Convolutional Neural Network.

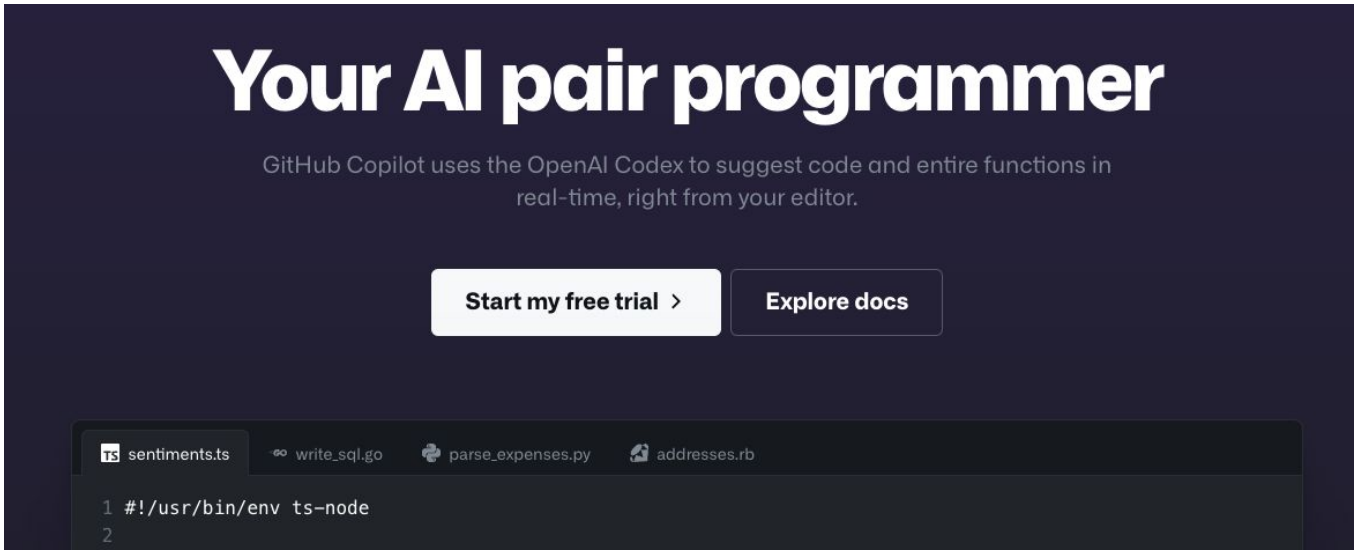
Reviewing the output of Frame Oracle on top of the video that was input to the model.



Predicting the state of American Football games.

Examples of Deep Learning

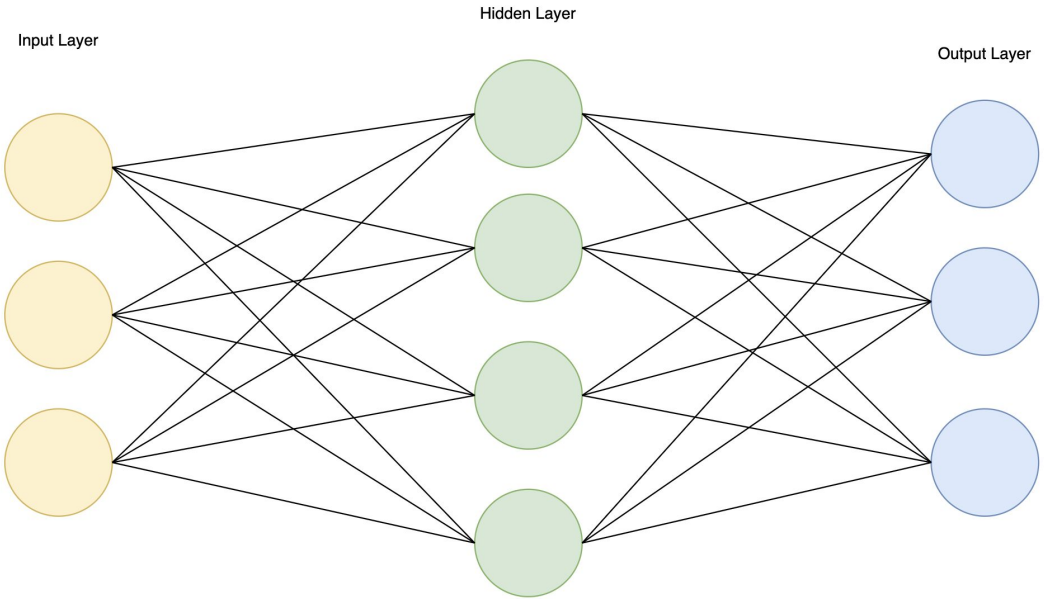
GitHub Copilot



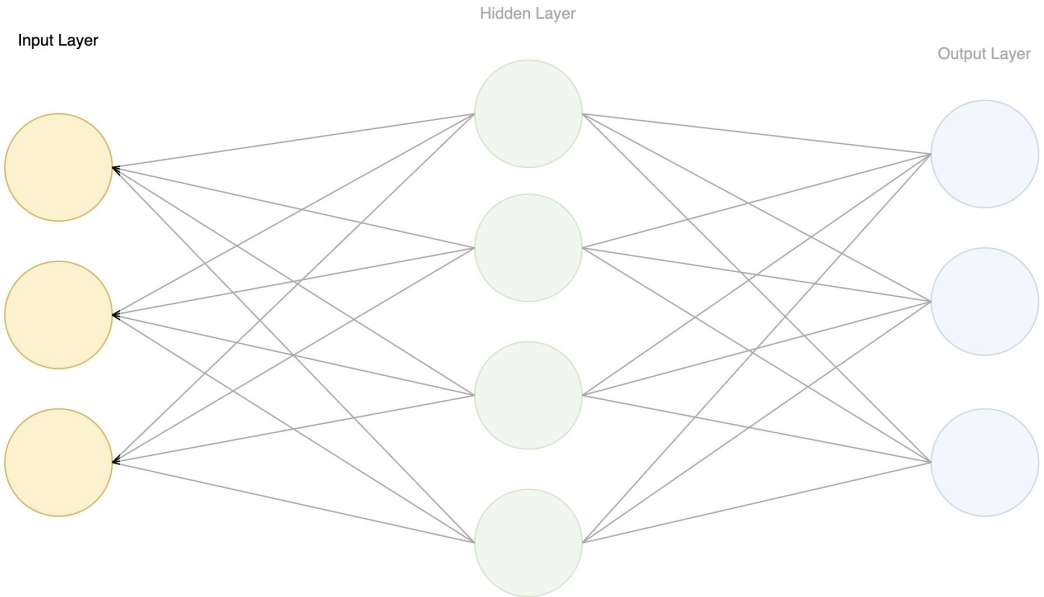
DALL-E created the cover for a recent magazine. DALL-E uses CLIP is an example of a deep learning model.

Neural Networks

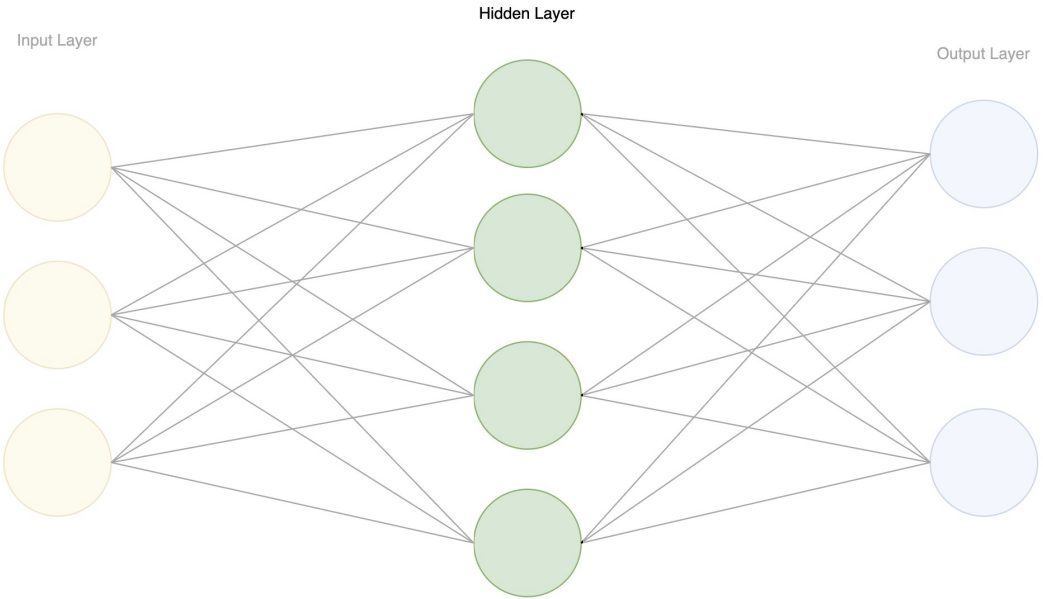
Overview of neural network



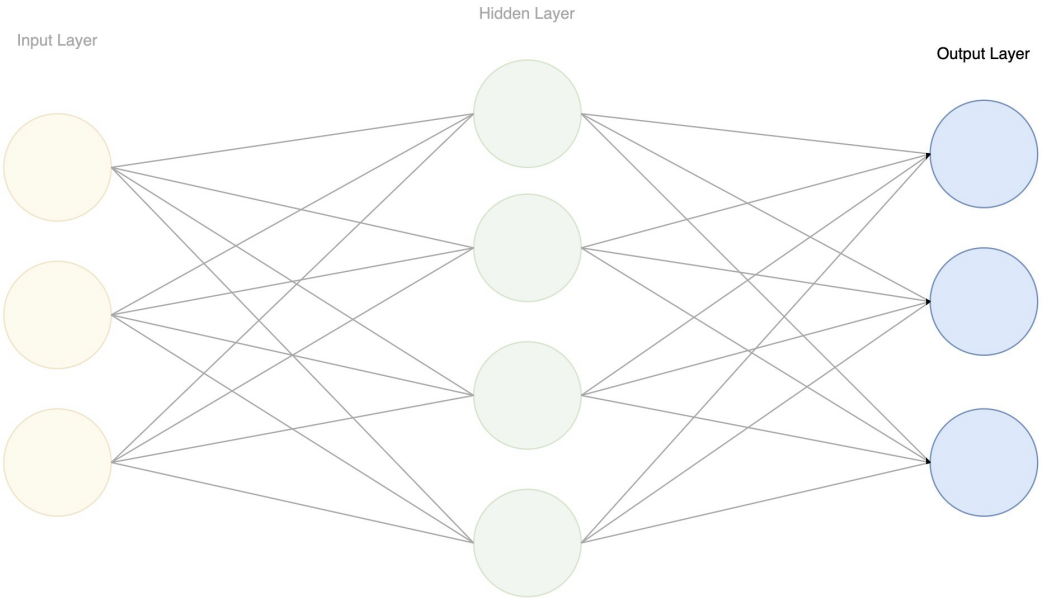
Input Layer = Array of Values



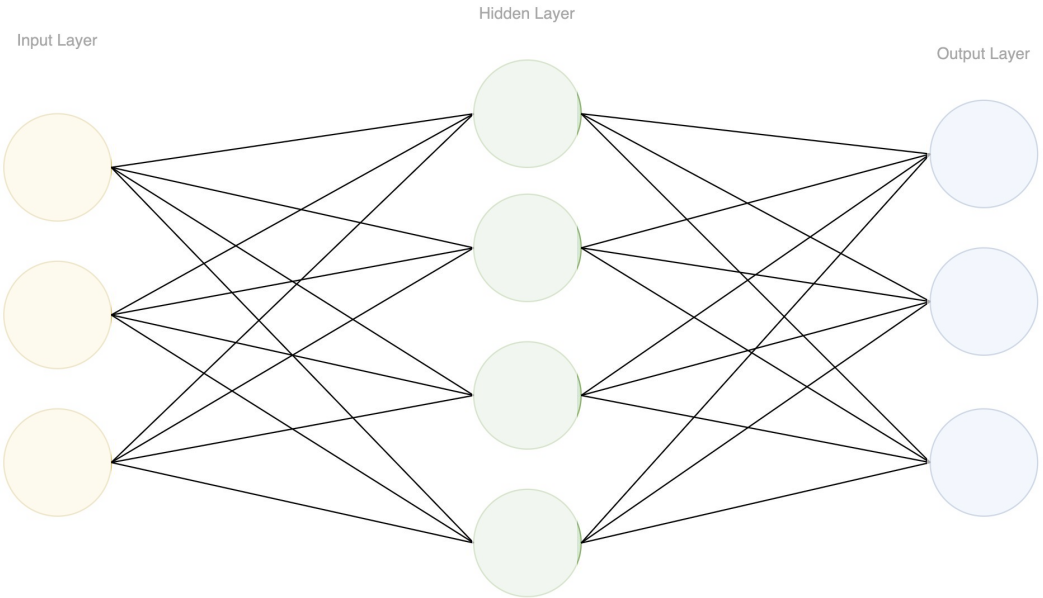
Hidden Layer(s) = where the network learns



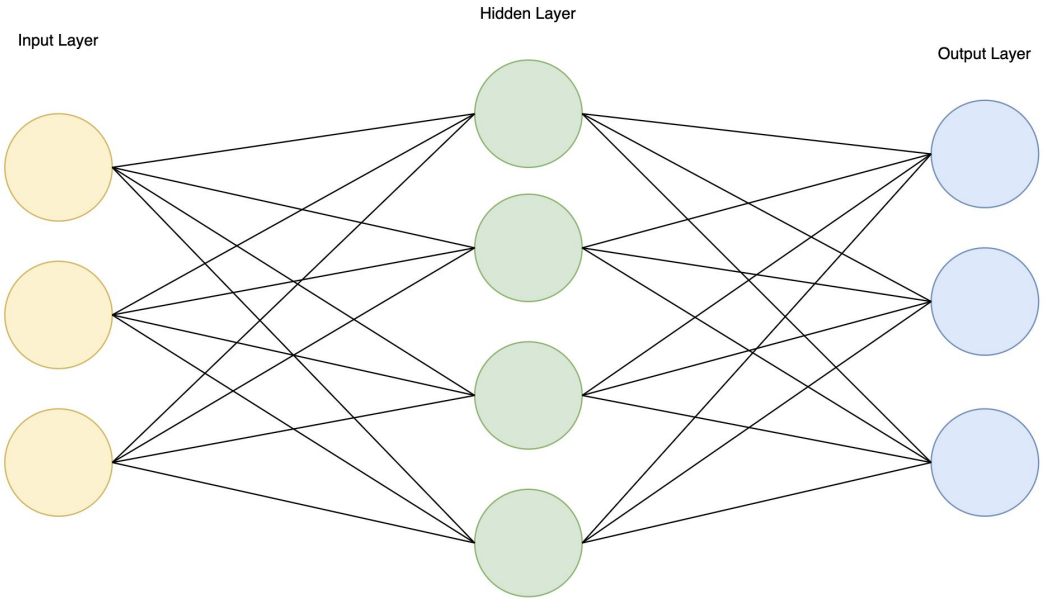
Output Layer = matches the problem type



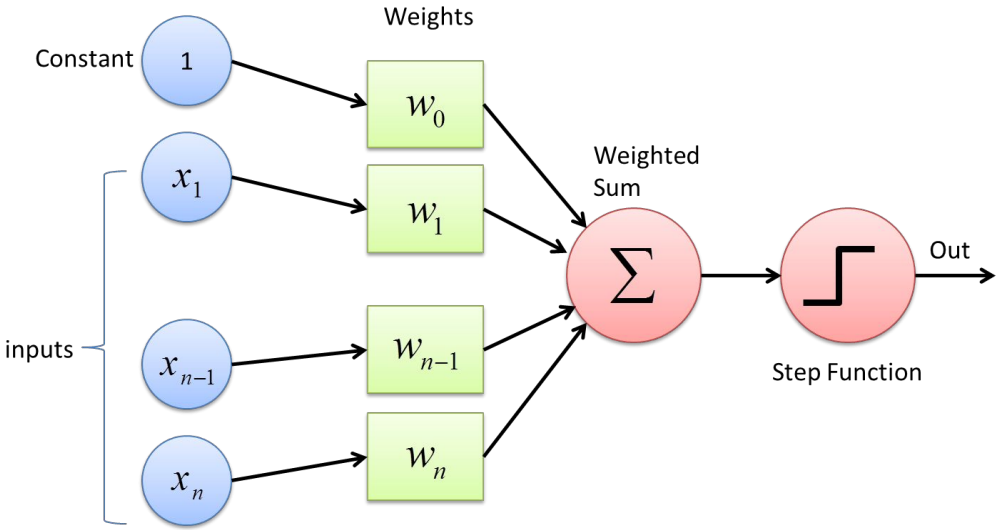
Connections = Weighted Edges



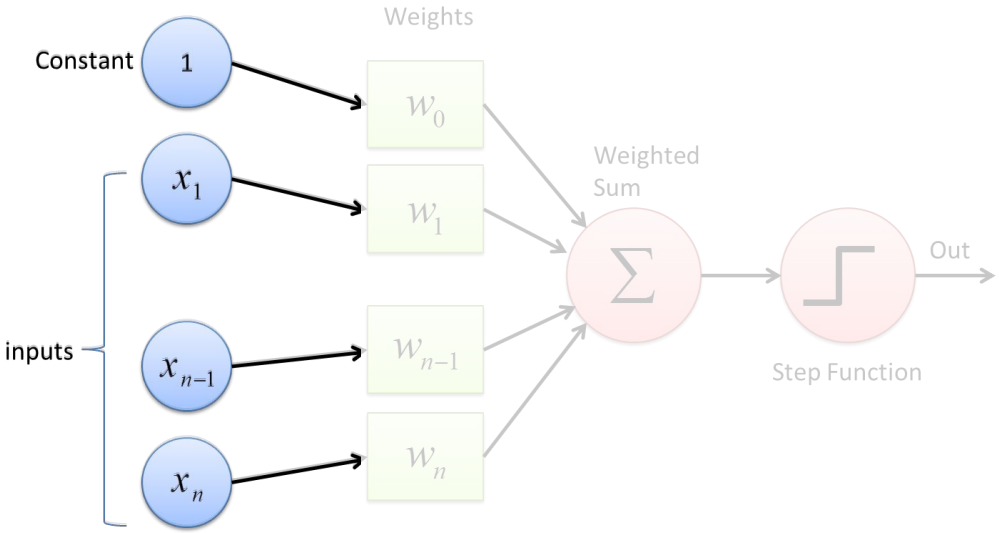
Overview of neural network



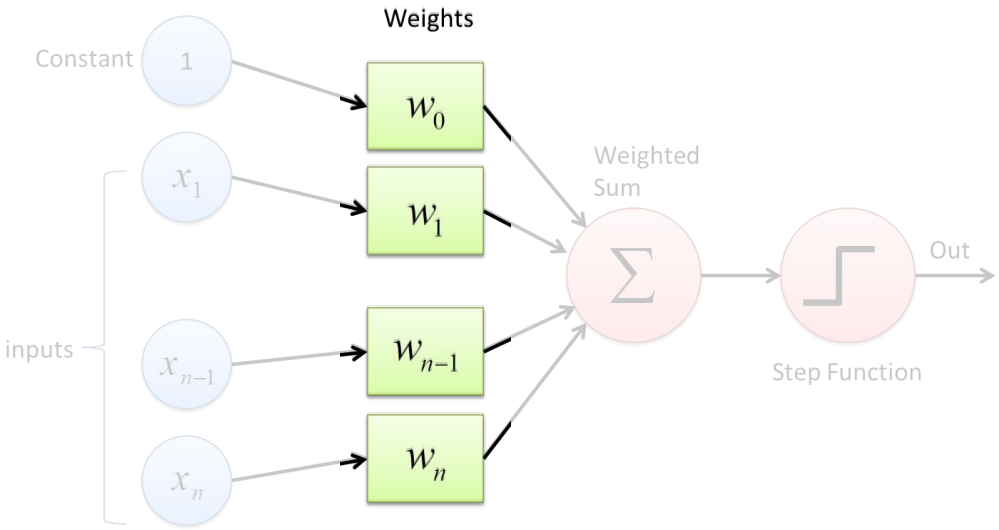
Perceptron: Single Layer Neural Network



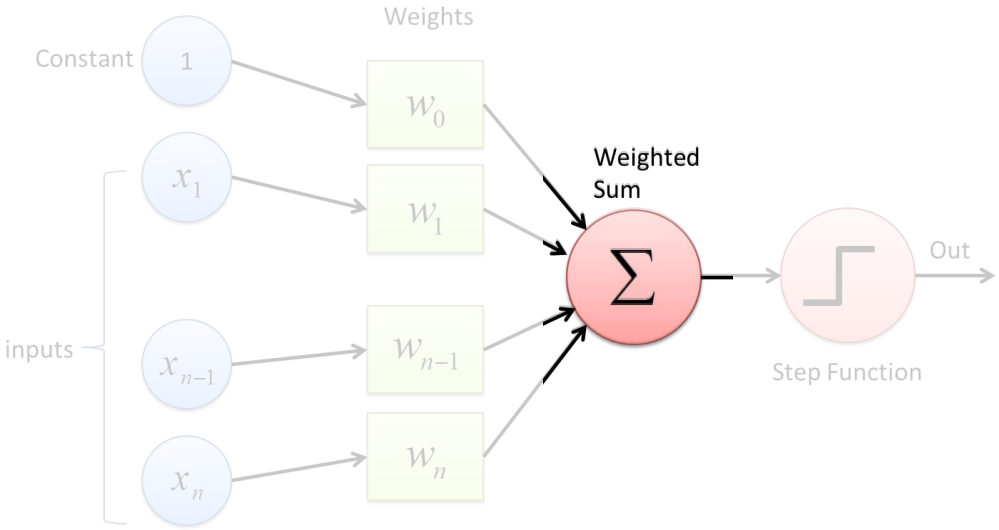
Inputs and Bias



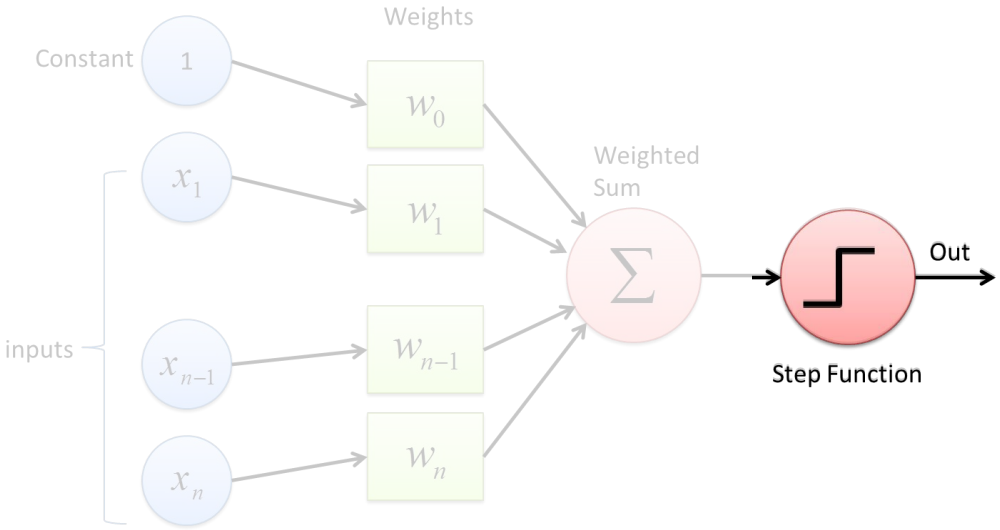
Weights for each input



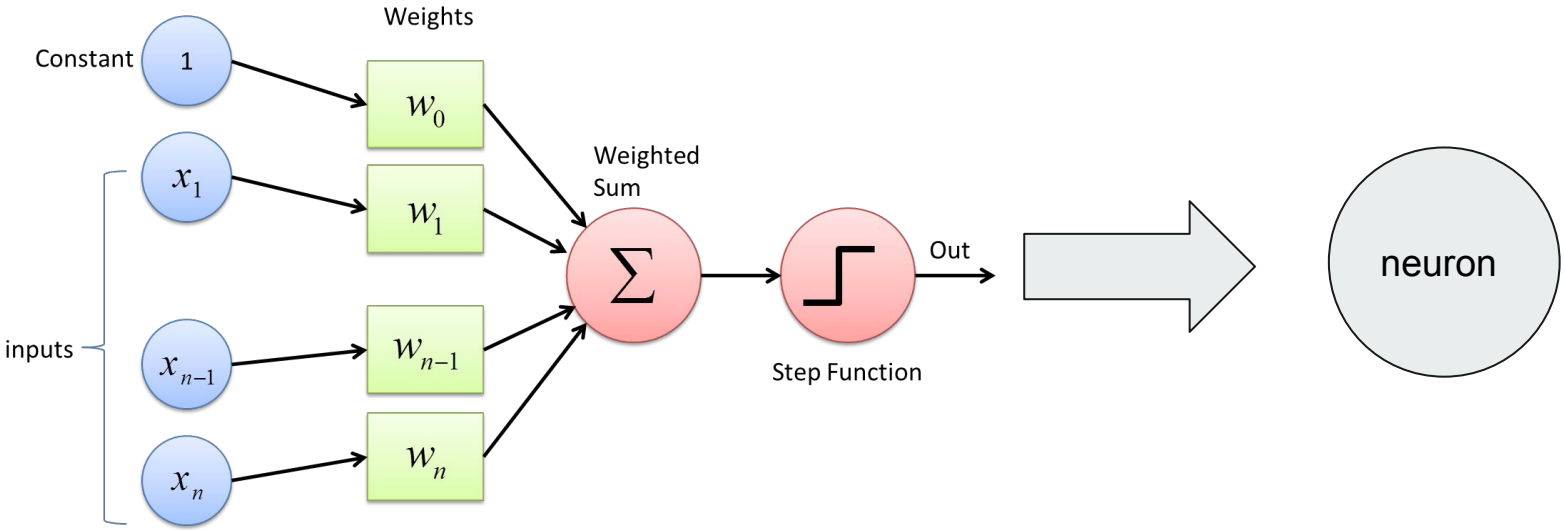
Added together



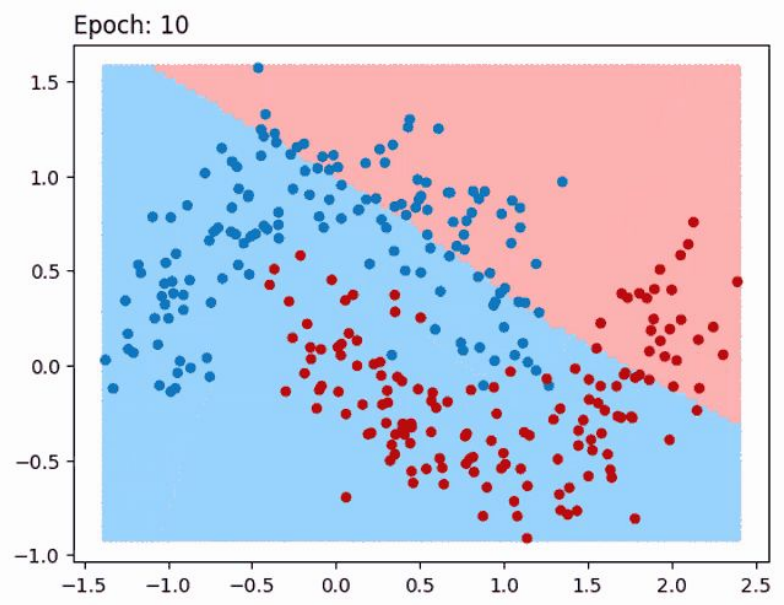
Multiplied by an activation function



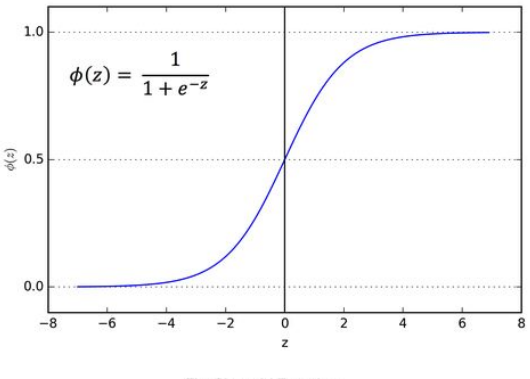
Perceptron: Single Layer Neural Network

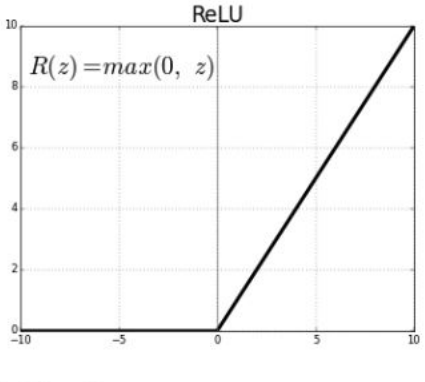


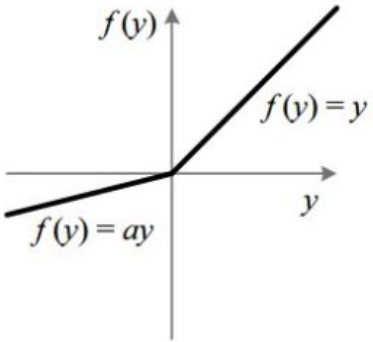
Why use an activation function?



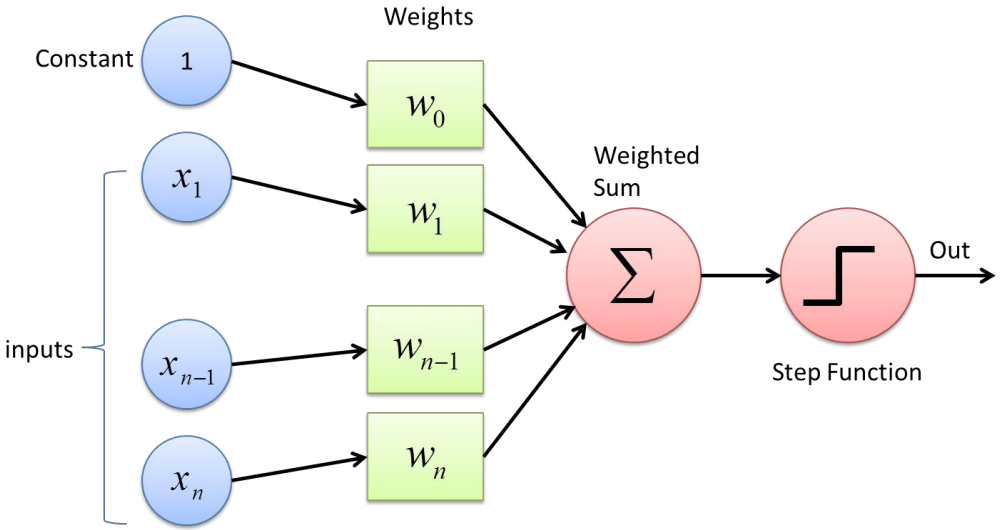
Non-linear Activation Functions



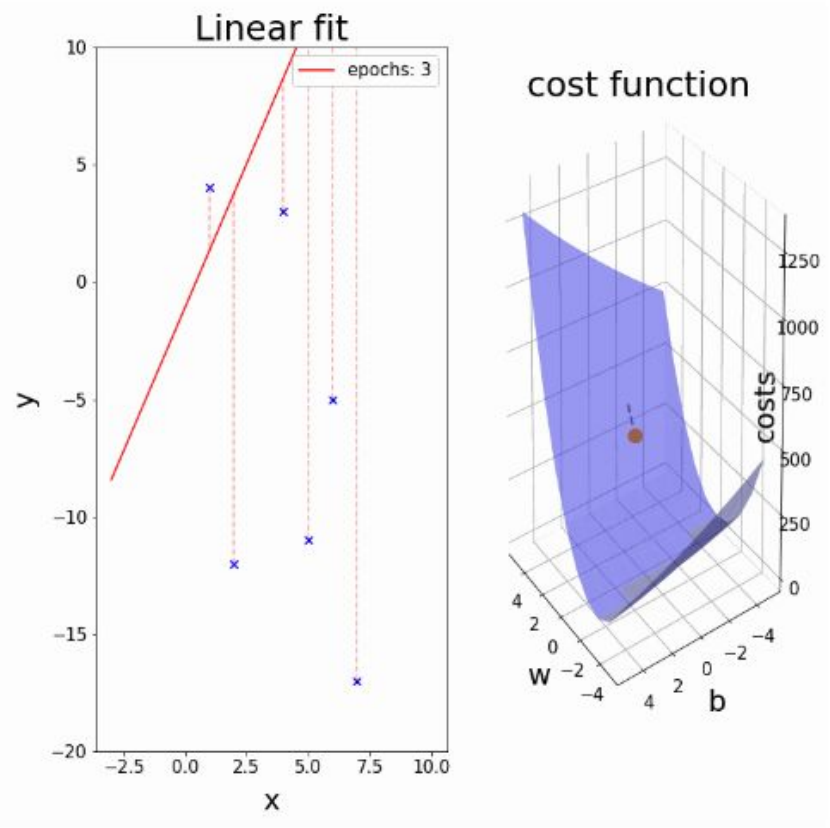




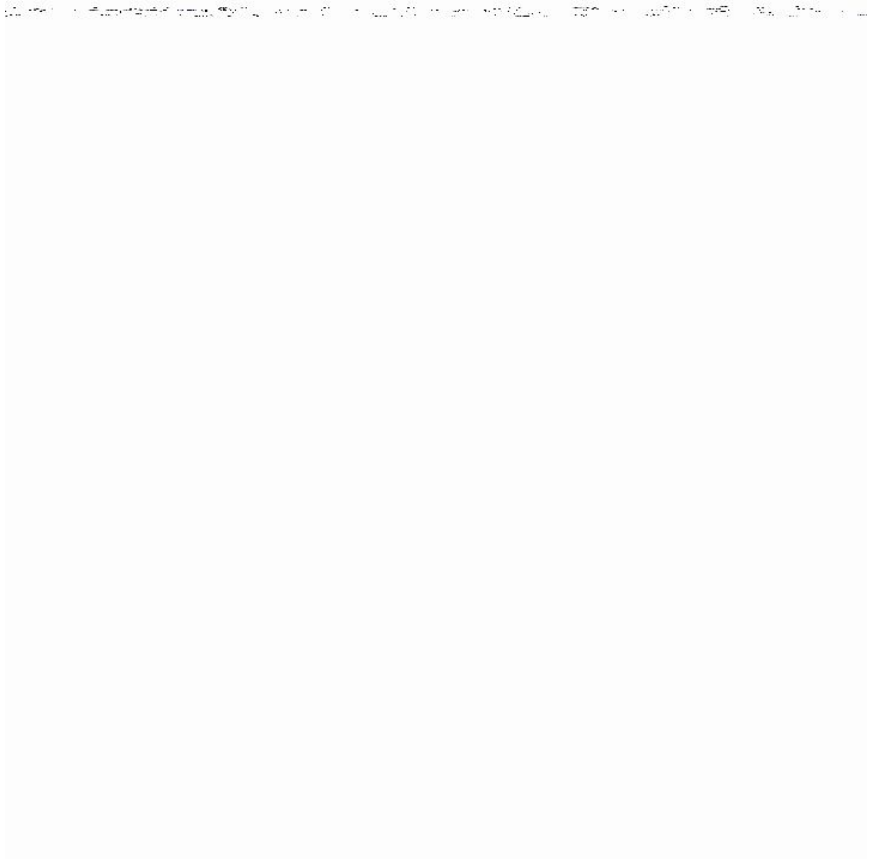
Backpropagation



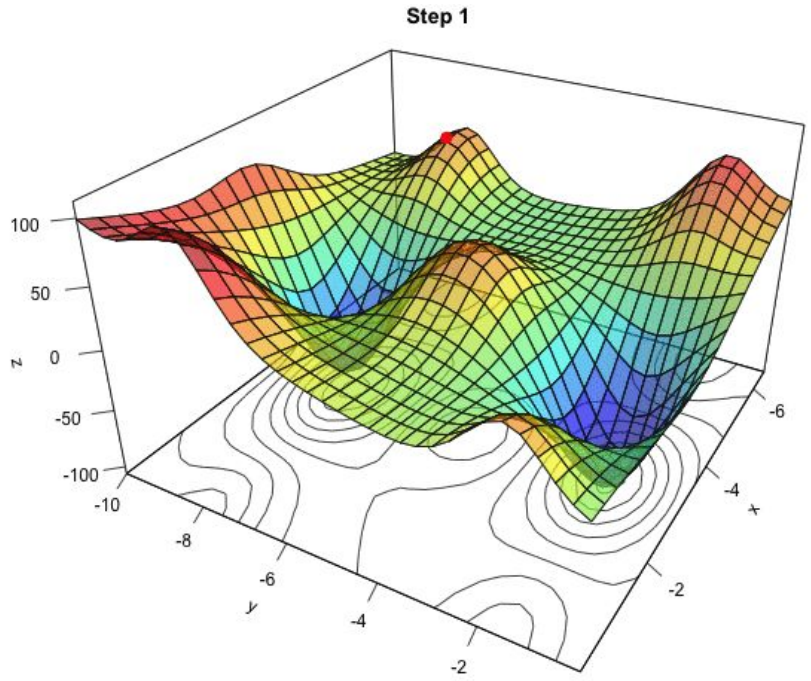
Gradient Descent



Gradient Descent



Gradient Descent



Initializing Weights in Network

Convolutional Neural Networks (CNNs)

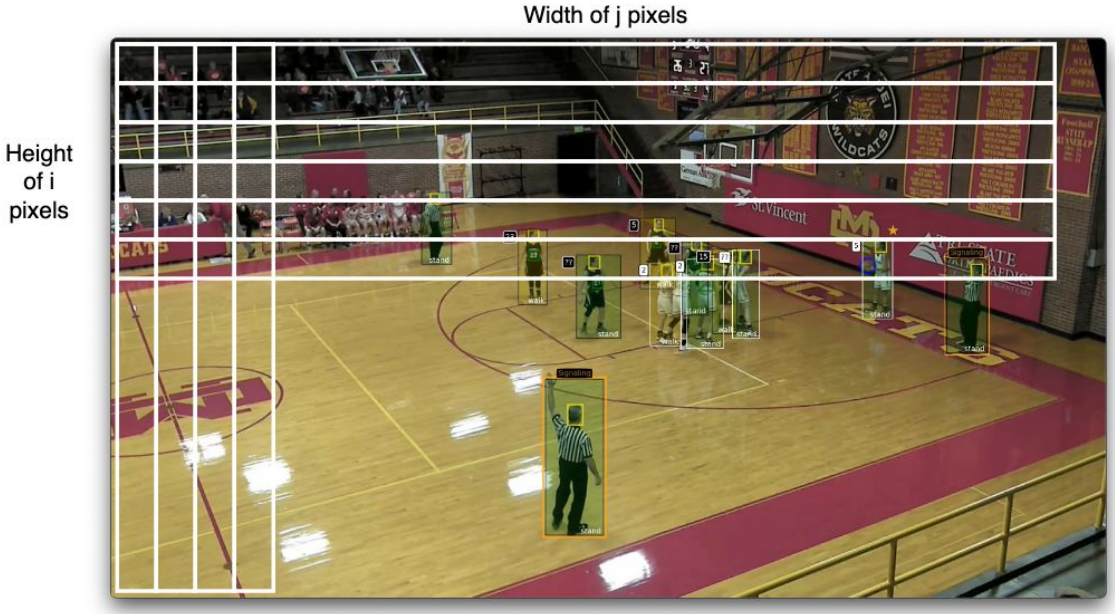
Real World Applications

Frame Oracle

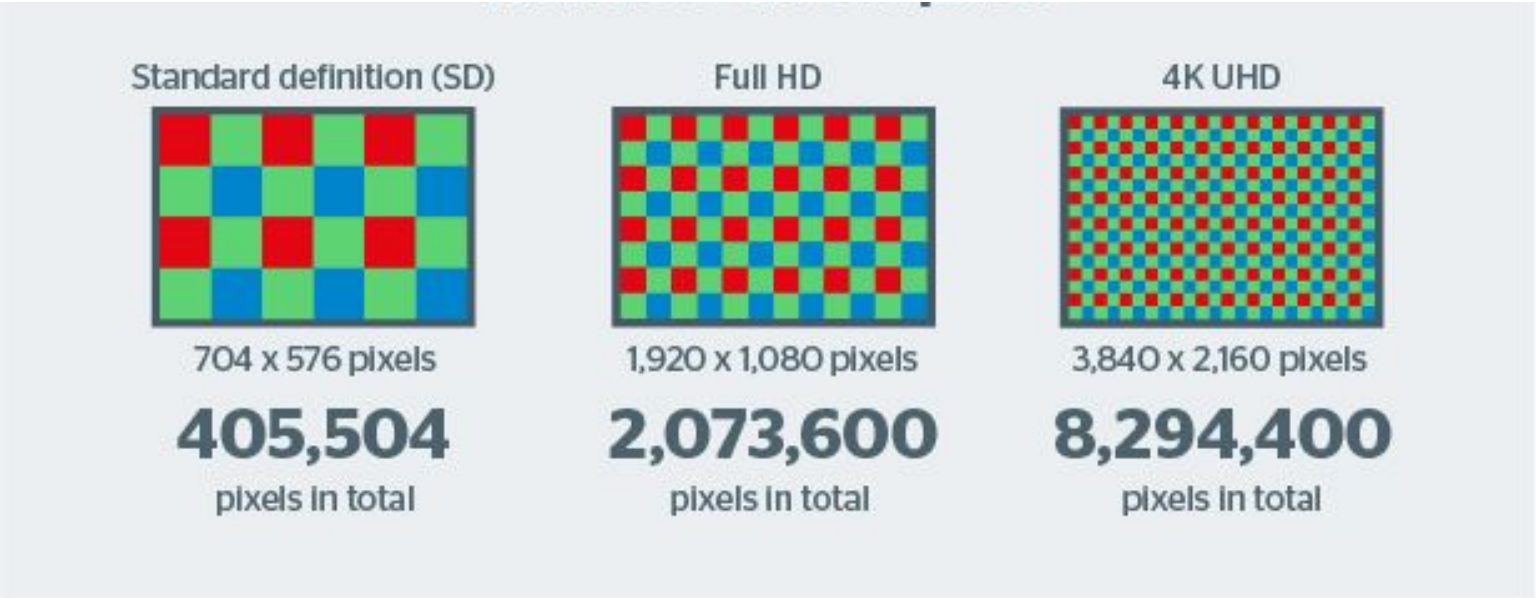
Input: Spatial Relationship

Frame Oracle

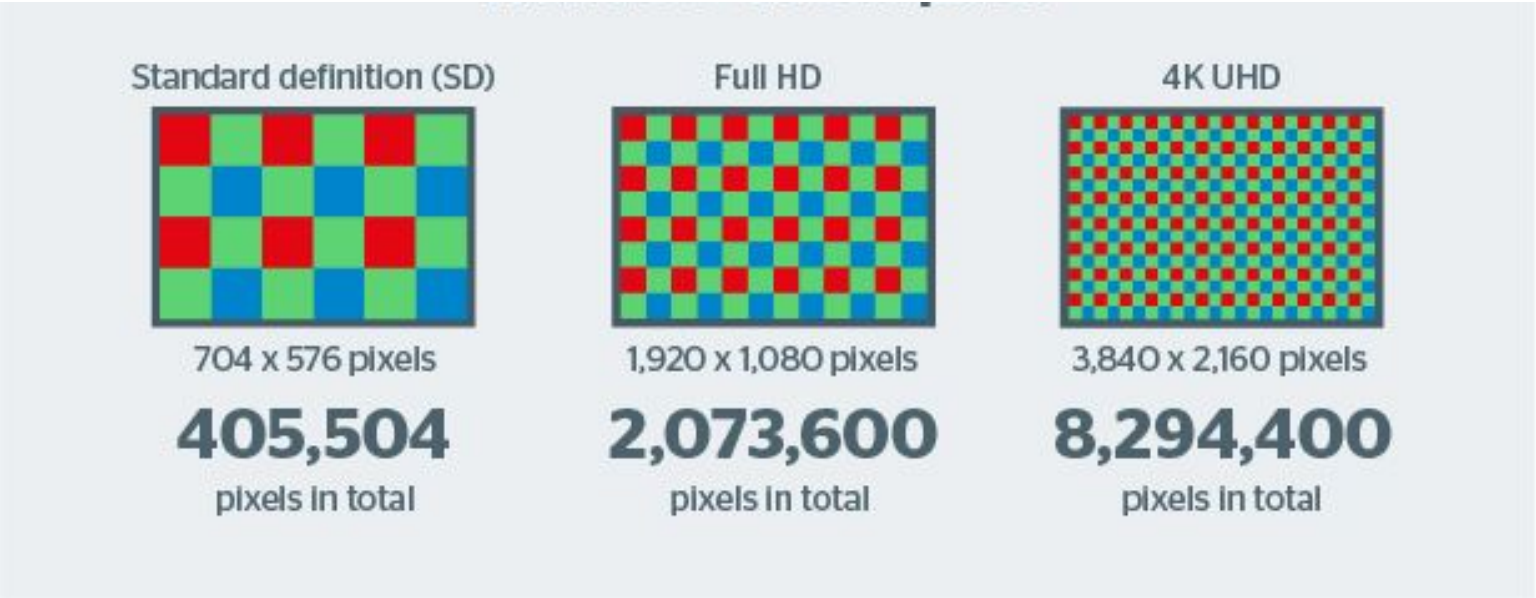
Matrix of (i, j) values that correspond to colors



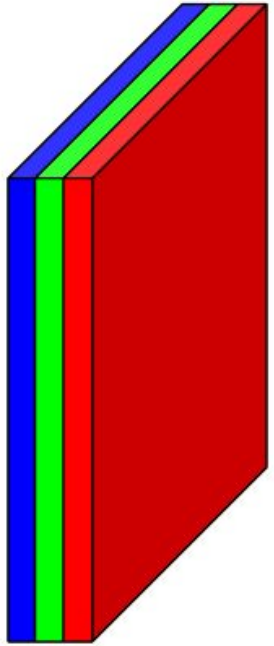
Size of inputs



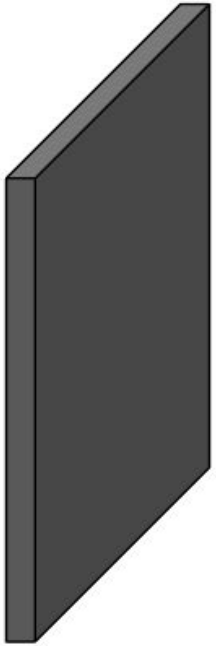
4K UHD² * 100 = 6,872,410,000,000,000



Input: RGB layers

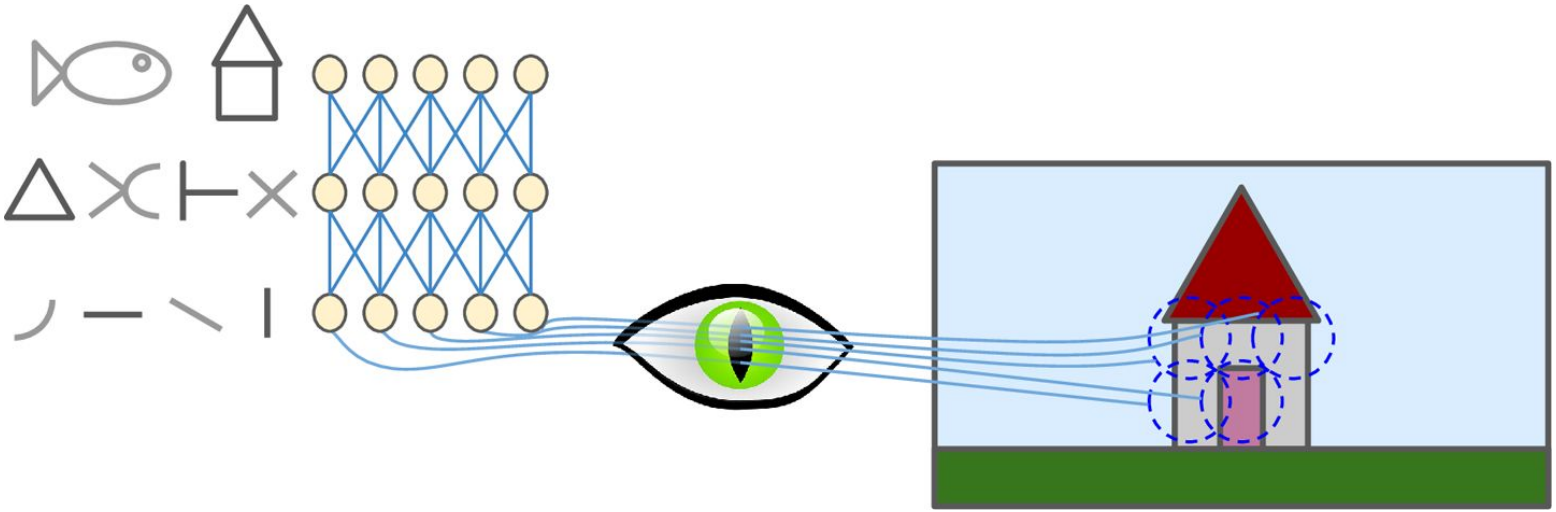


RGB
Image
M x N x 3



Grayscale/Binary
Image
M x N x 1

Breaking Down An Image



Input: Spatial Relationship

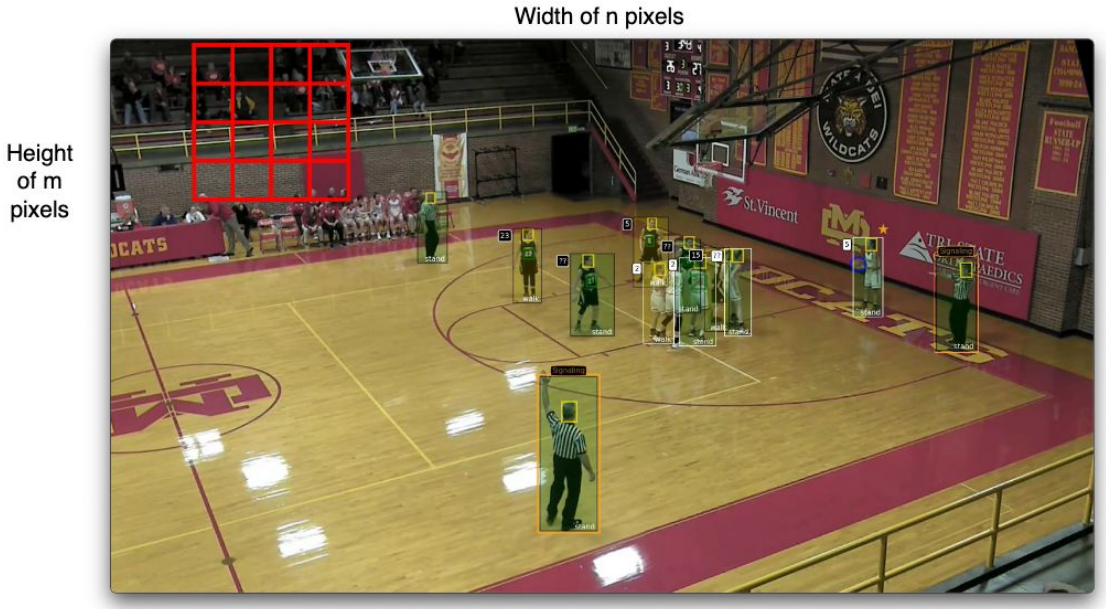
To avoid losing the spatial information, we consider subsections of the image together.

Input: Spatial Relationship

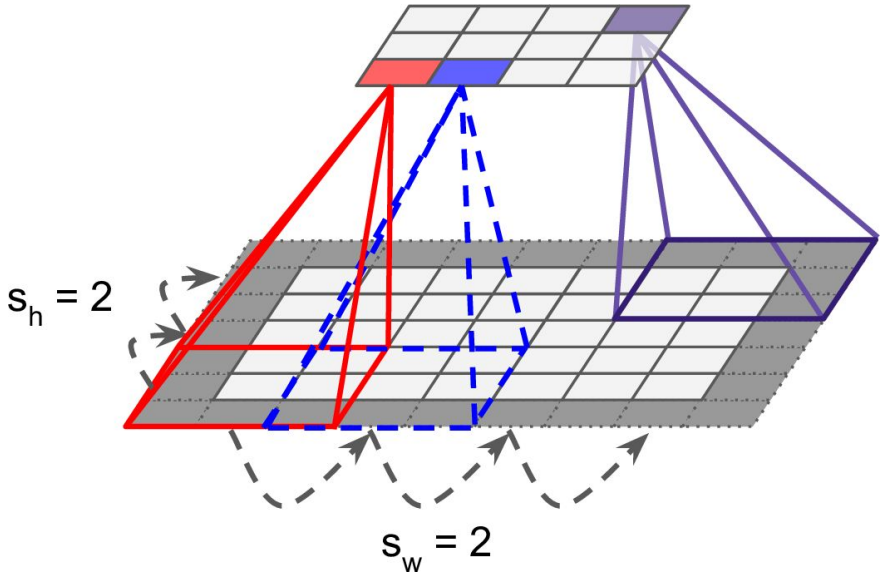
To avoid losing the spatial information, we consider subsections of the image together.

Input: Spatial Relationship

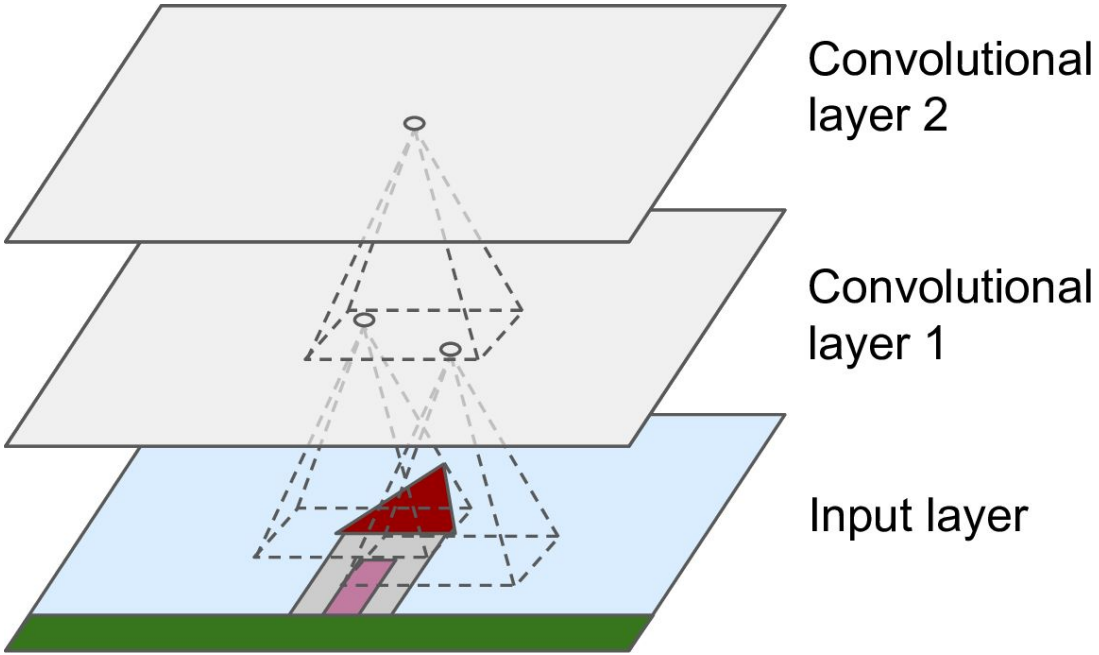
To avoid losing the spatial information, we consider subsections of the image together.



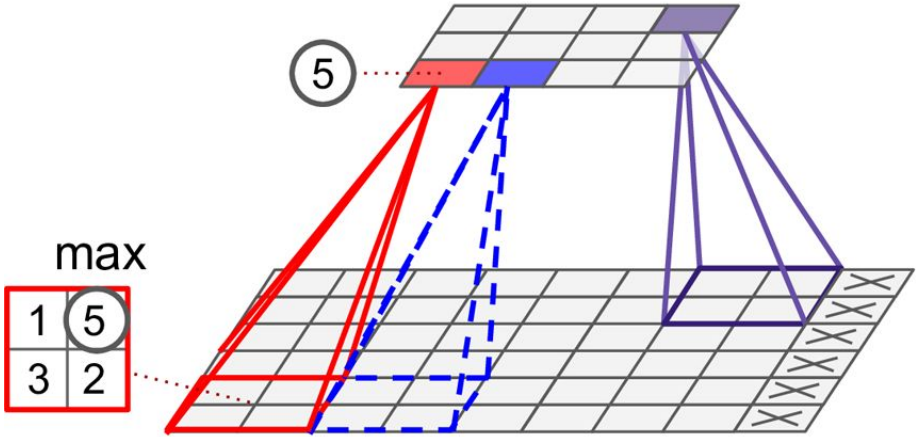
Turning regions into single values

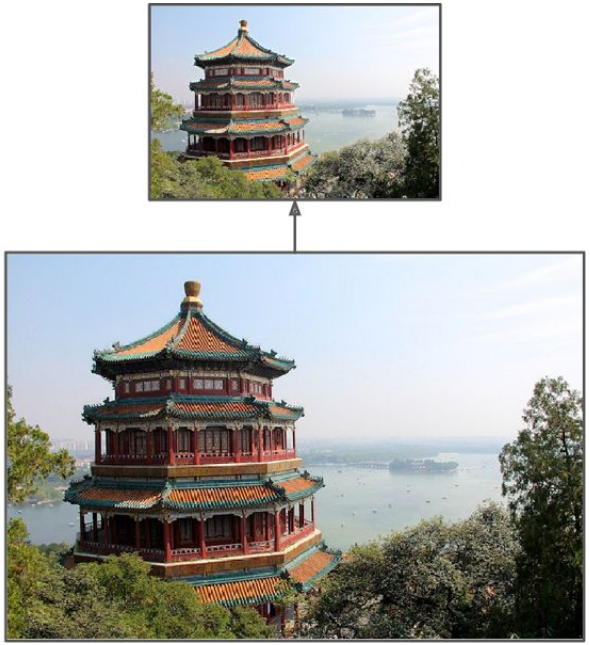


Turning regions into single values

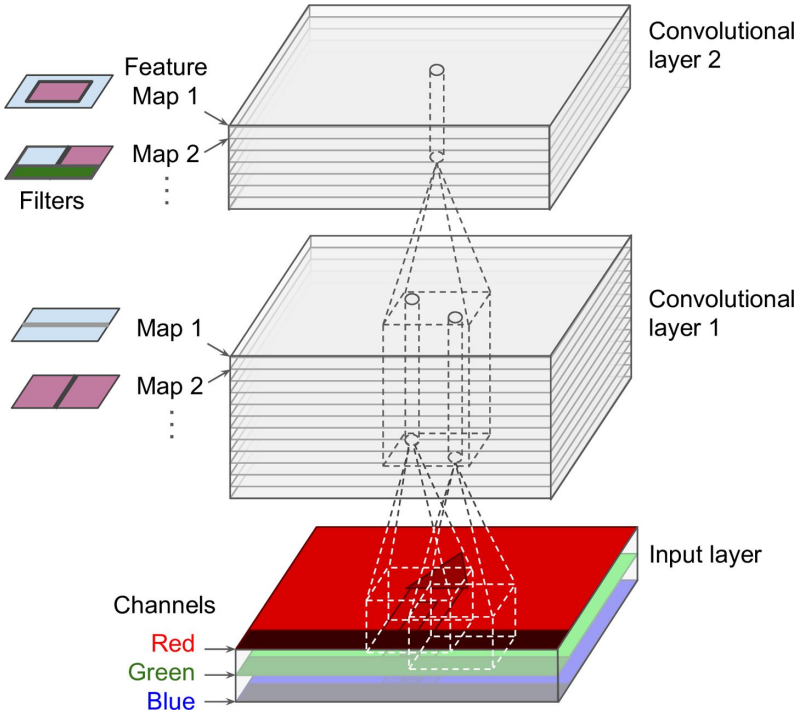


Pooling

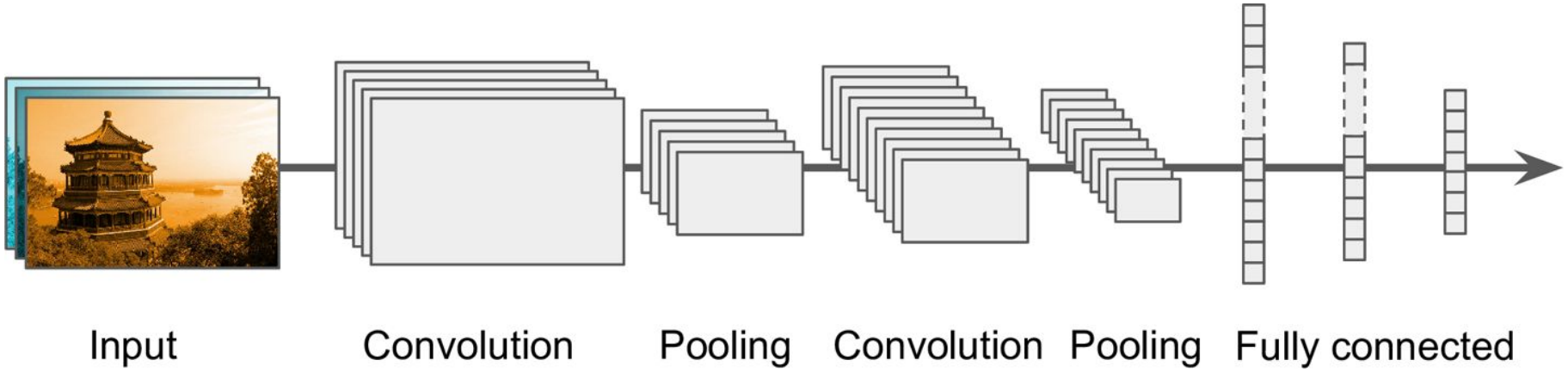




Stacking convolutional layers



So what does a whole CNN look like?



Recurrent Neural Networks

Real World Applications

NEWS

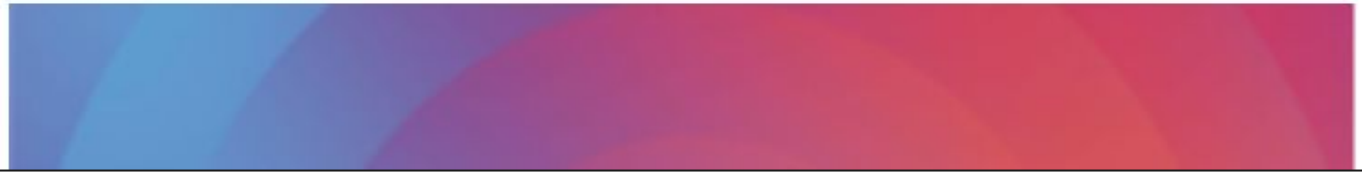
Jul 07, 2020

OpenAI Presents GPT-3, a 175 Billion Parameters Language Model

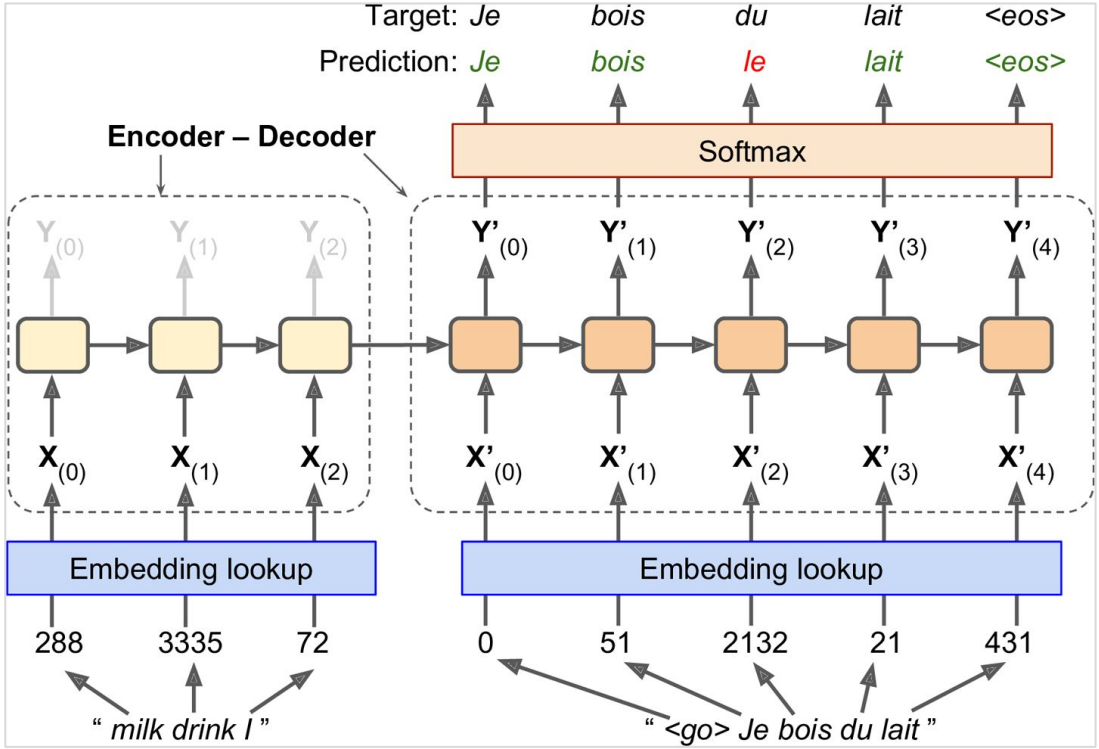
By Nefi Alarcon

Discuss (0) Share 0 Like

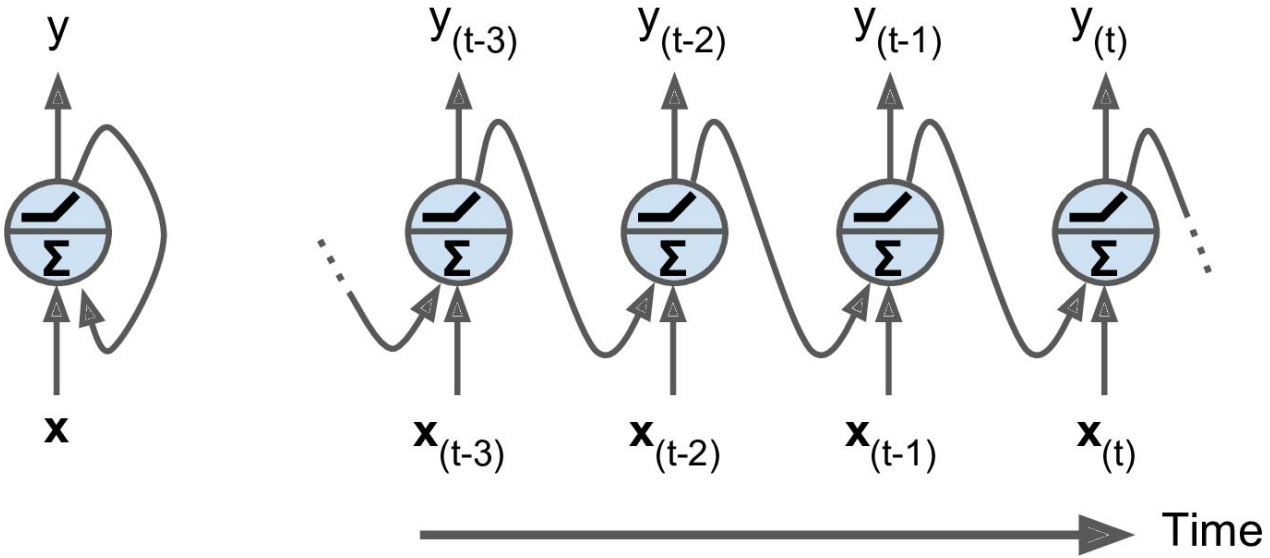
Tags: featured, Machine Learning & Artificial Intelligence, News, Speech & Audio Processing, Supercomputing / Cluster



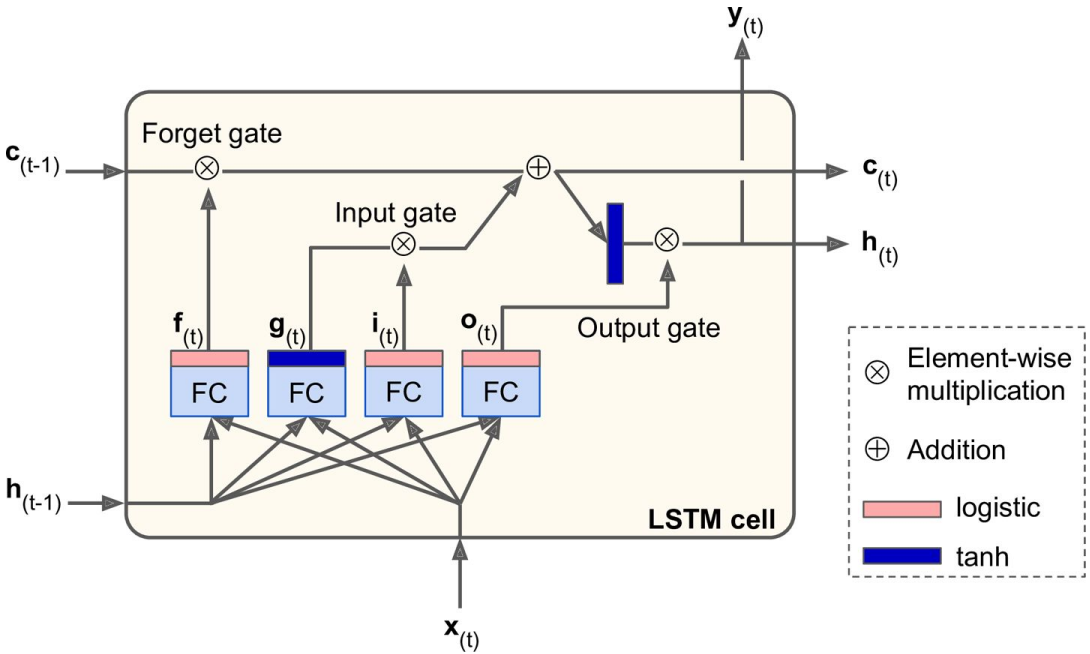
Input


Target: Je bois du lait <eos>
Prediction: Je bois le lait <eos>
Encoder – Decoder
Softmax
Y(0) Y(1) Y(2)
Y'(0) Y'(1) Y'(2) Y'(3) Y'(4)
X(0) X(1) X(2)
X'(0) X'(1) X'(2) X'(3) X'(4)
Embedding lookup
Embedding lookup
288 3335 72
0 51 2132 21 431
" milk drink I "
" <go> Je bois du lait "

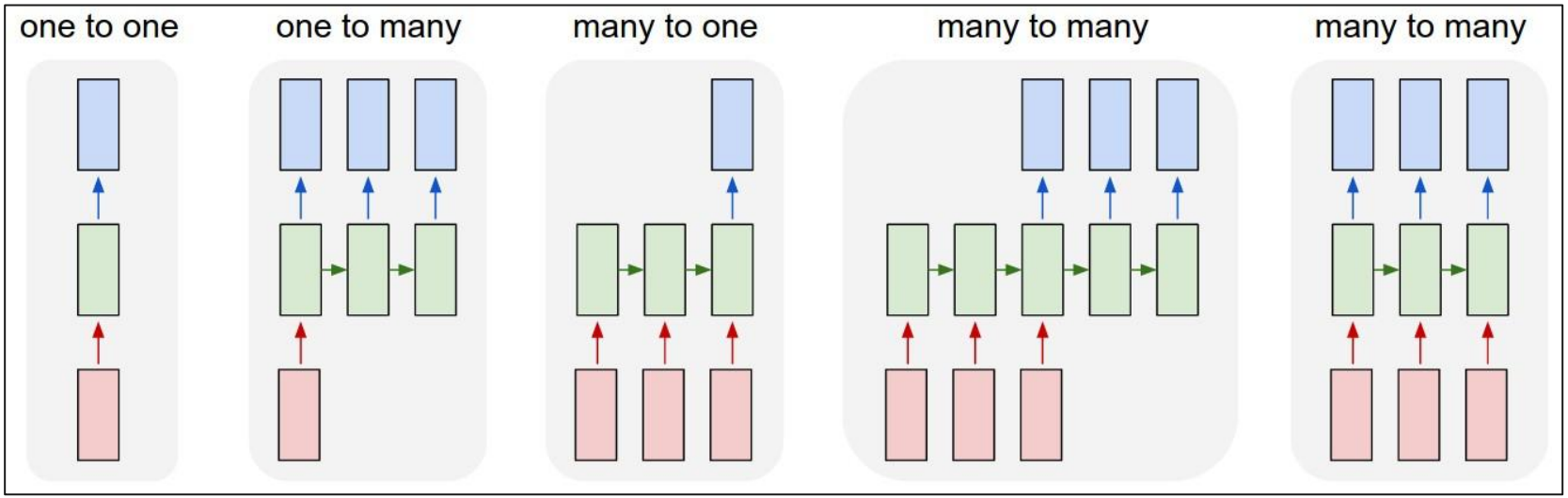
Architecture



Architecture: Long-Short Term Memory Cells



Output



Disadvantages of Neural Networks

Black Box

NEWS

Jul 07, 2020

OpenAI Presents GPT-3, a 175 Billion Parameters Language Model

By Nefi Alarcon

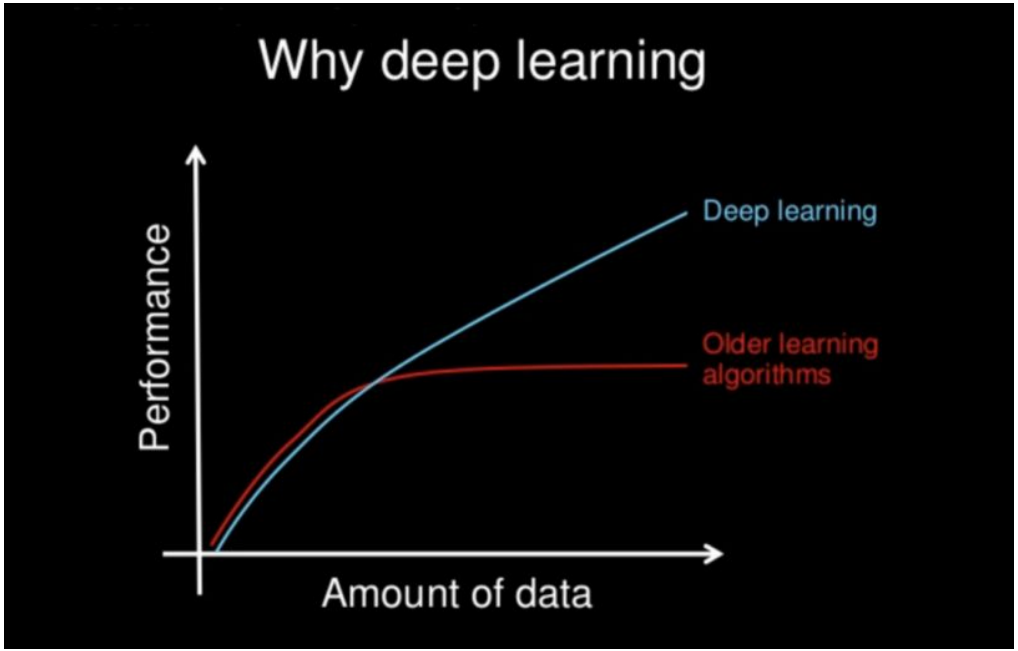
Discuss (0) Share 0 Like

Tags: featured, Machine Learning & Artificial Intelligence, News, Speech & Audio Processing, Supercomputing / Cluster

Black Box

Training Time/Computation

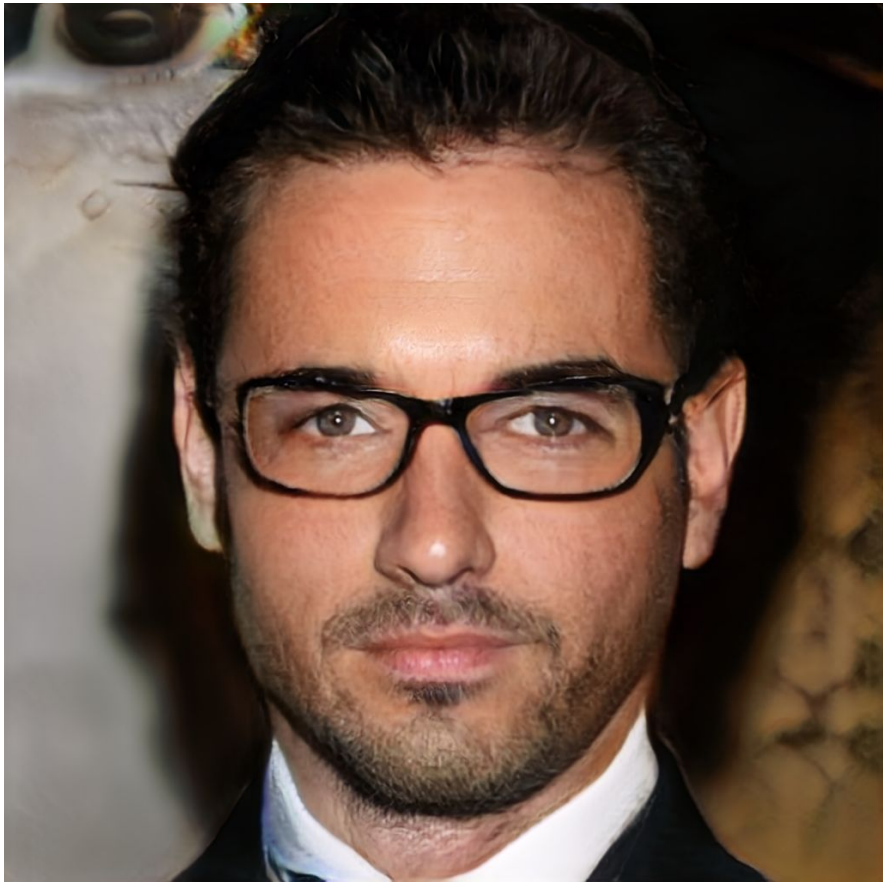
Neural Networks need a lot of data



Thanks for listening to my presentation

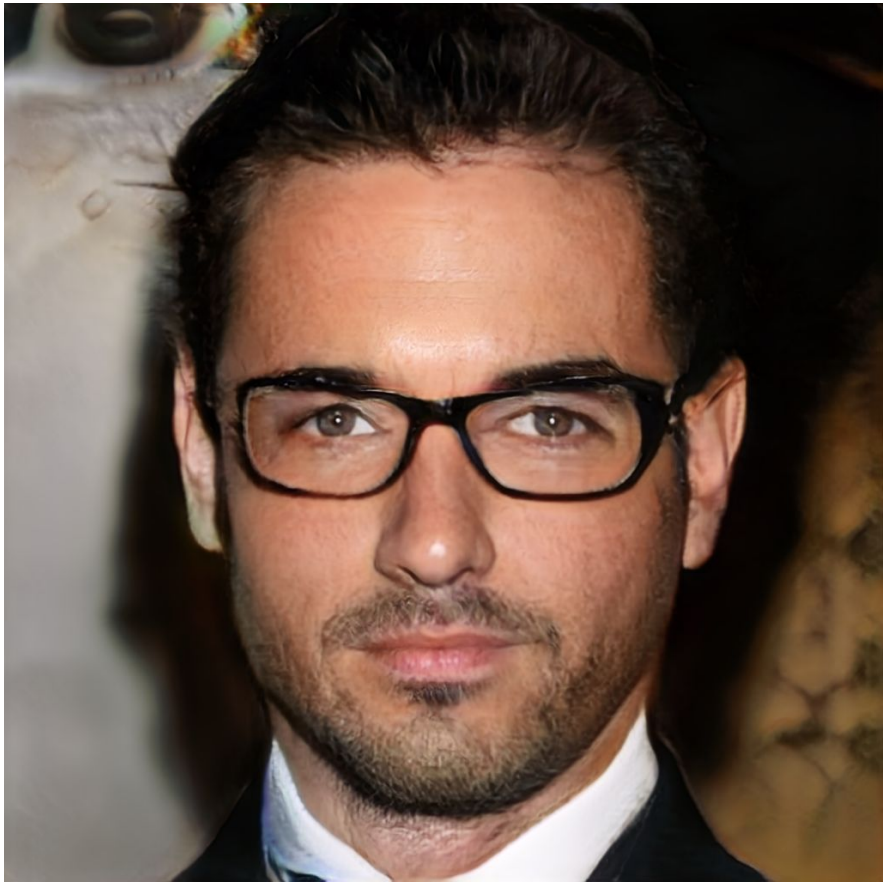
Additional topics and citations after this slide

Generative Adversarial Networks (GANs)



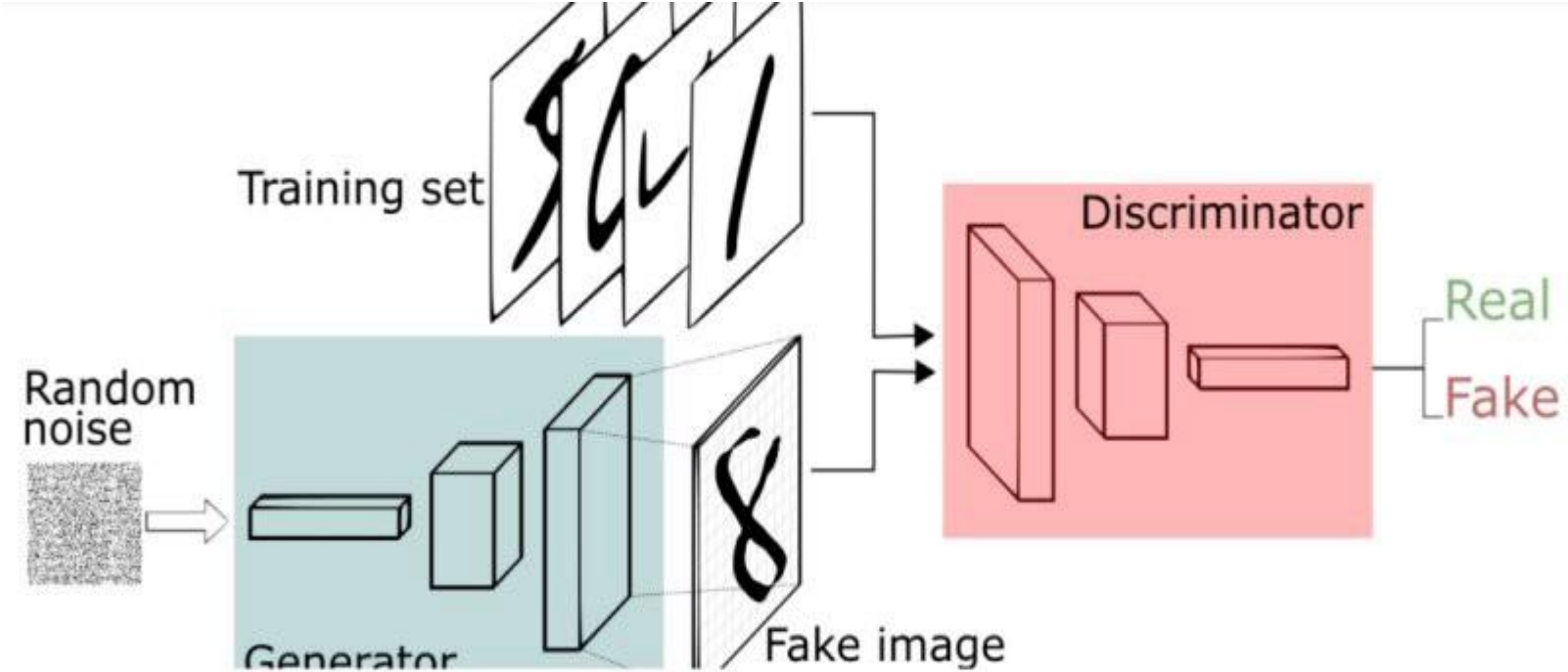
Do you know this celebrity?

Do you know this celebrity?



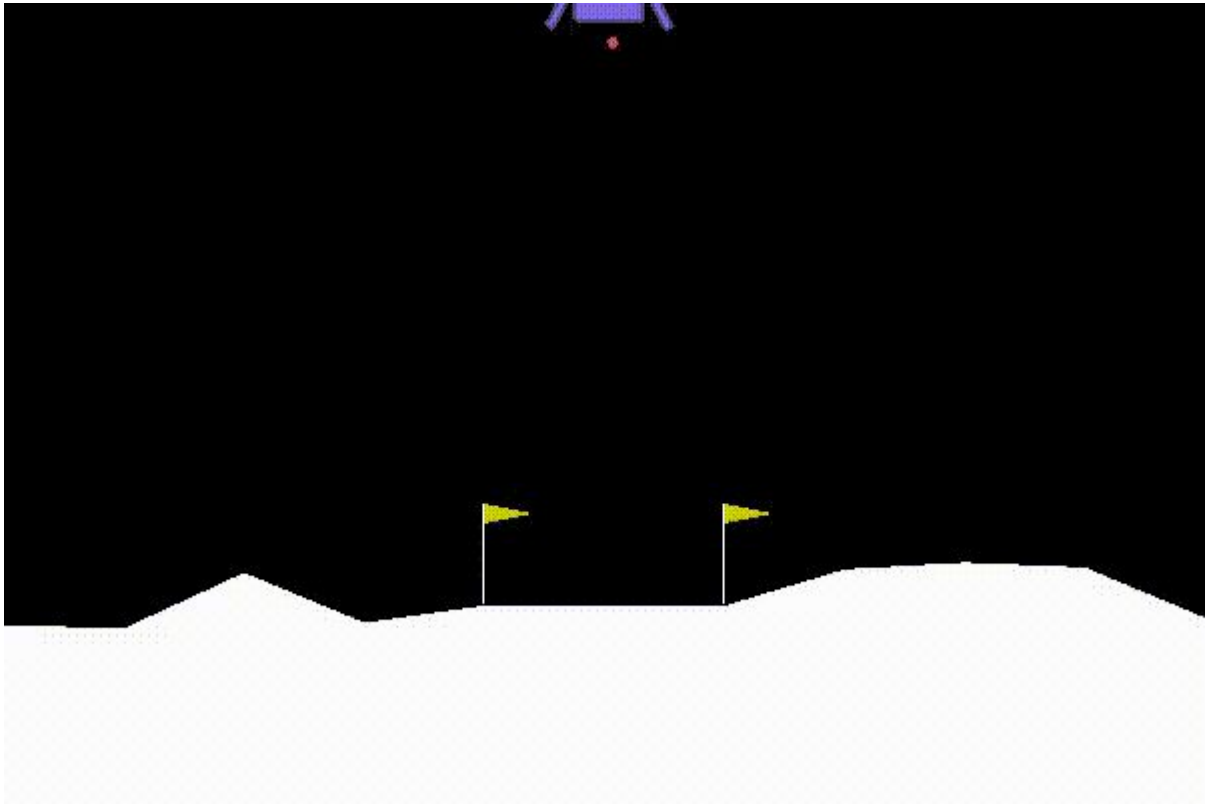
His backdrop is very odd. The glasses don’t seem to connect to his right ear.

Why is she wearing two different earrings? Does the skin around her left eye look older than the right?

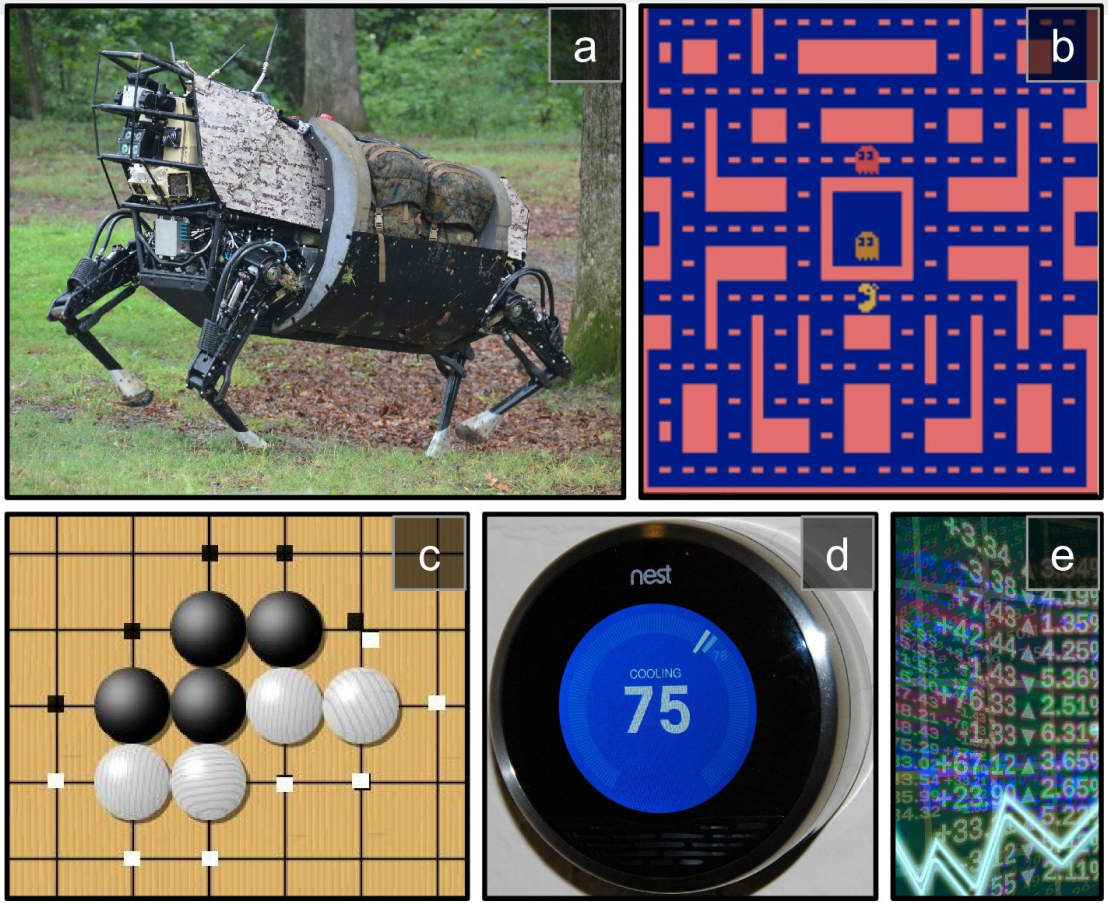


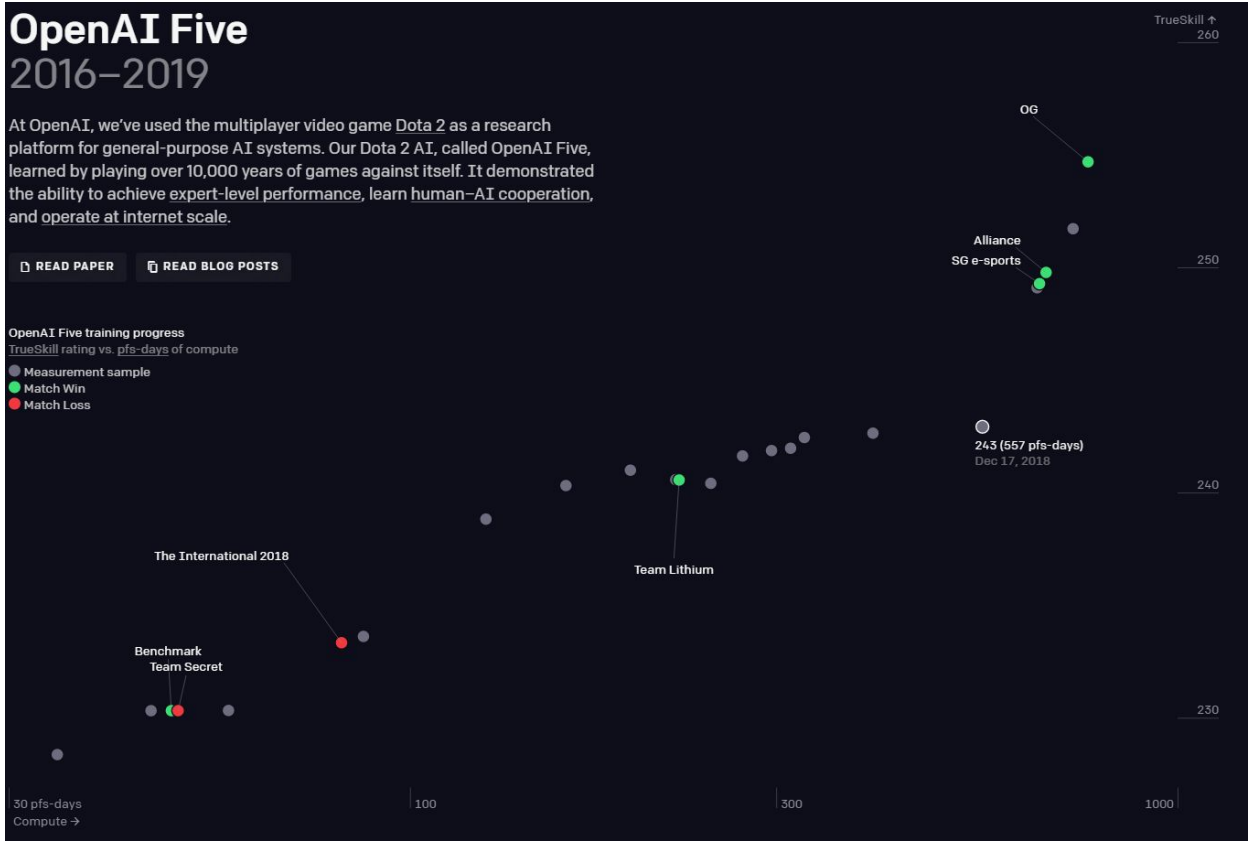
Architecture for a GAN

Reinforcement Learning



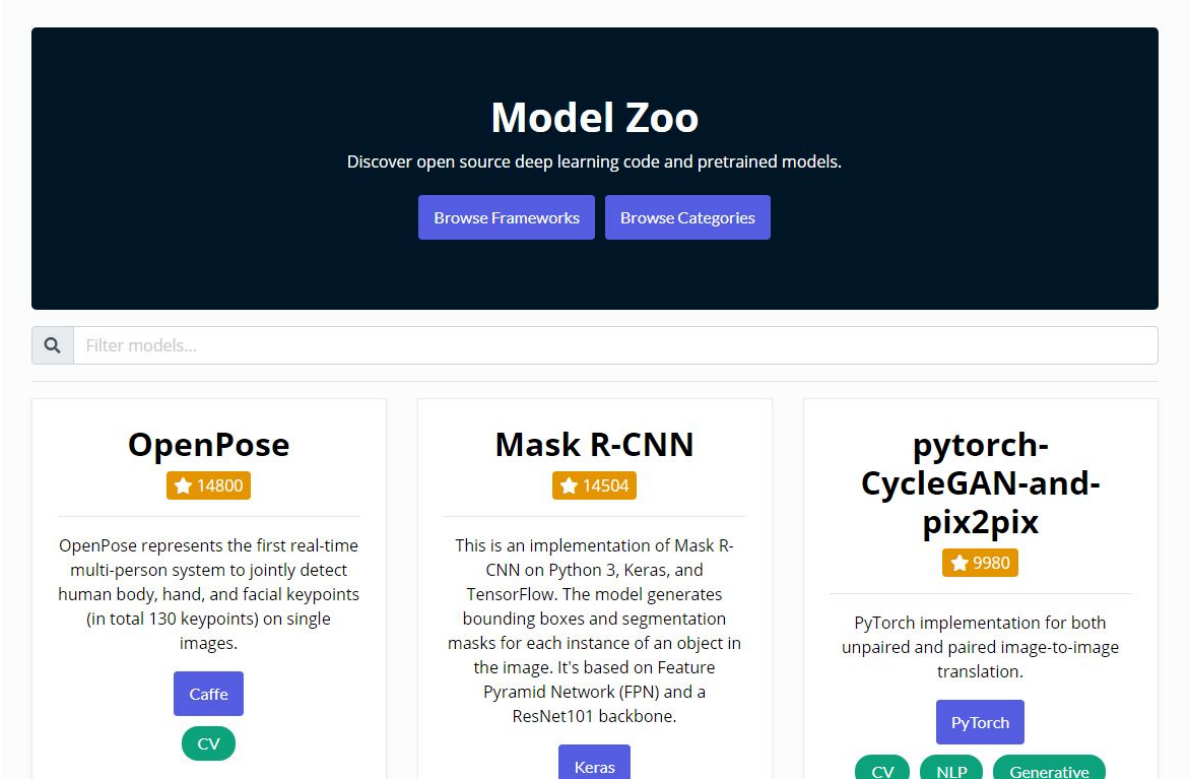
Examples



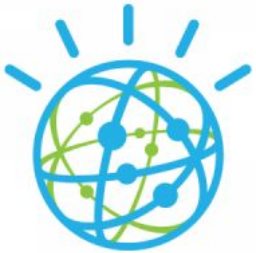

OpenAI Five
2016–2019
At OpenAI, we've used the multiplayer video game Dota 2 as a research platform for general-purpose AI systems. Our Dota 2 AI, called OpenAI Five, learned by playing over 10,000 years of games against itself. It demonstrated the ability to achieve expert-level performance, learn human–AI cooperation, and operate at internet scale.
READ PAPER
READ BLOG POSTS
OpenAI Five training progress
TrueSkill rating vs. pfs-days of compute
Measurement sample
Match Win
Match Loss
TrueSkill ↑
260
250
240
230
OG
Alliance
SG e-sports
243 (557 pfs-days)
Dec 17, 2018
Team Lithium
The International 2018
Benchmark
Team Secret
30 pfs-days
100
300
1000
Compute →

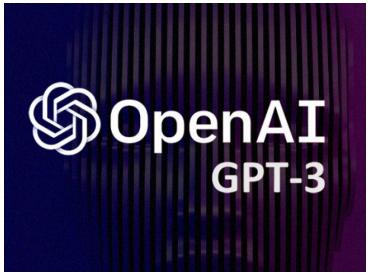
Model Zoos and Transfer Learning

Model Zoo?

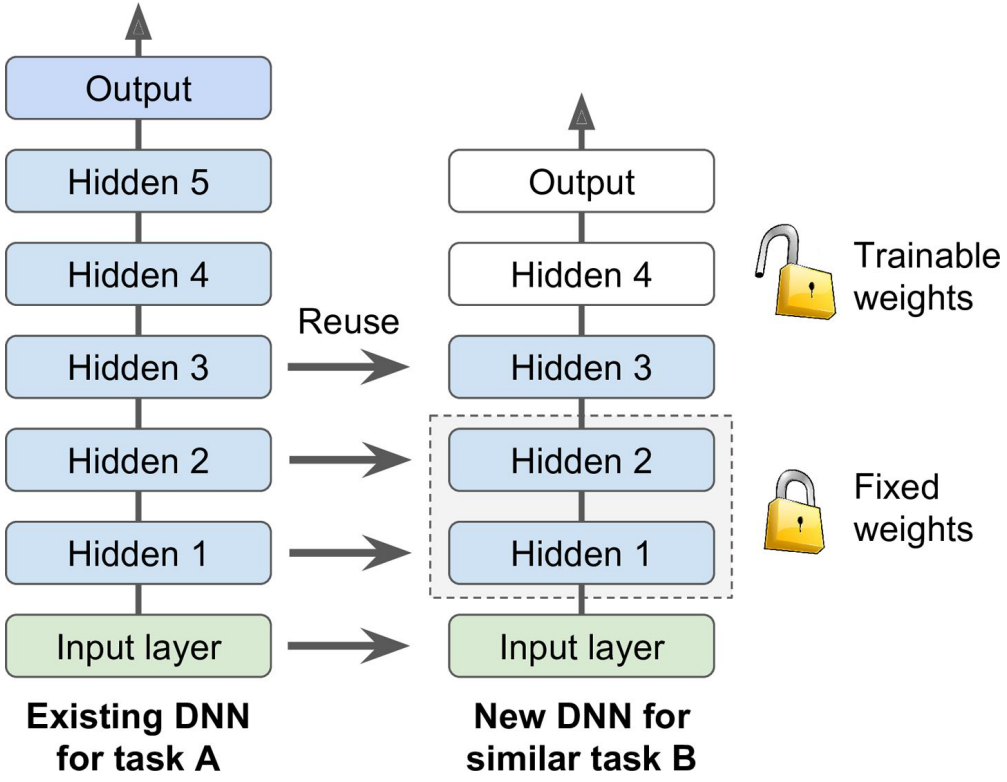

Model Zoo
Discover open source deep learning code and pretrained models.
Browse Frameworks
Browse Categories
Filter models...
OpenPose
14800
OpenPose represents the first real-time multi-person system to jointly detect human body, hand, and facial keypoints (in total 130 keypoints) on single images.
Caffe
CV
Mask R-CNN
14504
This is an implementation of Mask R-CNN on Python 3, Keras, and TensorFlow. The model generates bounding boxes and segmentation masks for each instance of an object in the image. It's based on Feature Pyramid Network (FPN) and a ResNet101 backbone.
Keras
pytorch-CycleGAN-and-pix2pix
9980
PyTorch implementation for both unpaired and paired image-to-image translation.
PyTorch
CV
NLP
Generative

Alternatively, use an API

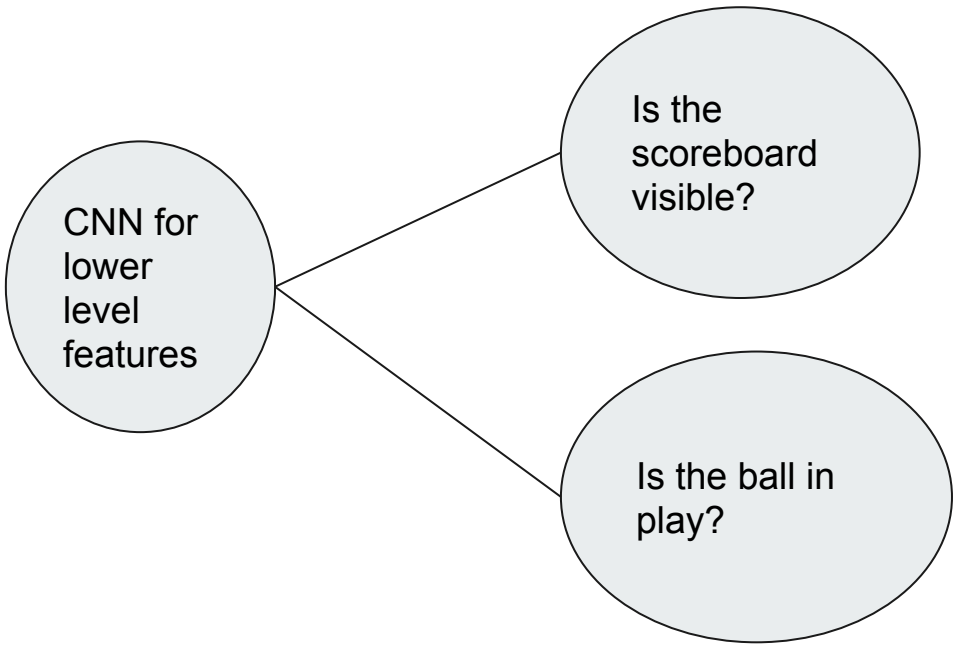




Transfer Learning



Multi-task Models


CNN for lower level features
Is the scoreboard visible?
Is the ball in play?

Why Deep Learning Now?

Big Data

GPU

Continued advancement in neural networks