A walk with deep learning

Victor Rodriguez-Fernandez victor.rfernandez@upm.es
Alejandro Martín alejandro.martin@upm.es

Previously on Stardust LTW-I...

- Use of high levels APIs for deep learning (fastai)
- Transfer learning & fine tuning vision models for image classification
- Classification & Regression on Tabular Data
- Handling imbalanced datasets

https://github.com/stardust-r/LTW-I

Outline

Day 1

1. Introduction
    a. What is deep learning?
    b. Deep learning milestones
    c. Basics on Machine/Deep learning
    d. Deep learning = Machine learning with Neural networks
    e. Areas of success in deep learning
    f. Why is deep learning currently booming?
2. Getting started
    a. Getting a GPU deep learning server
    b. Hands-on 1: Building a simple deep neural network in Pytorch
3. Practical tips & best practices
    a. Experiment tracking
    b. Software development in Jupyter Notebooks with nbdev

Day 2

1. Going deeper into deep learning
    a. Going deeper into deep learning
    b. Convolutional Neural Networks
    c. How does a filter look like?
    d. Generative Adversarial Networks
    e. Building a convolutional neural network with Keras
    f. Natural Language Processing with Deep Learning
    g. Transformers
2. Final discussion
    a. Trending topics and future opportunities
    b. Ethical aspects
    c. Limits of deep learning
    d. Free ML/DL resources

Day 1

A walk with deep learning

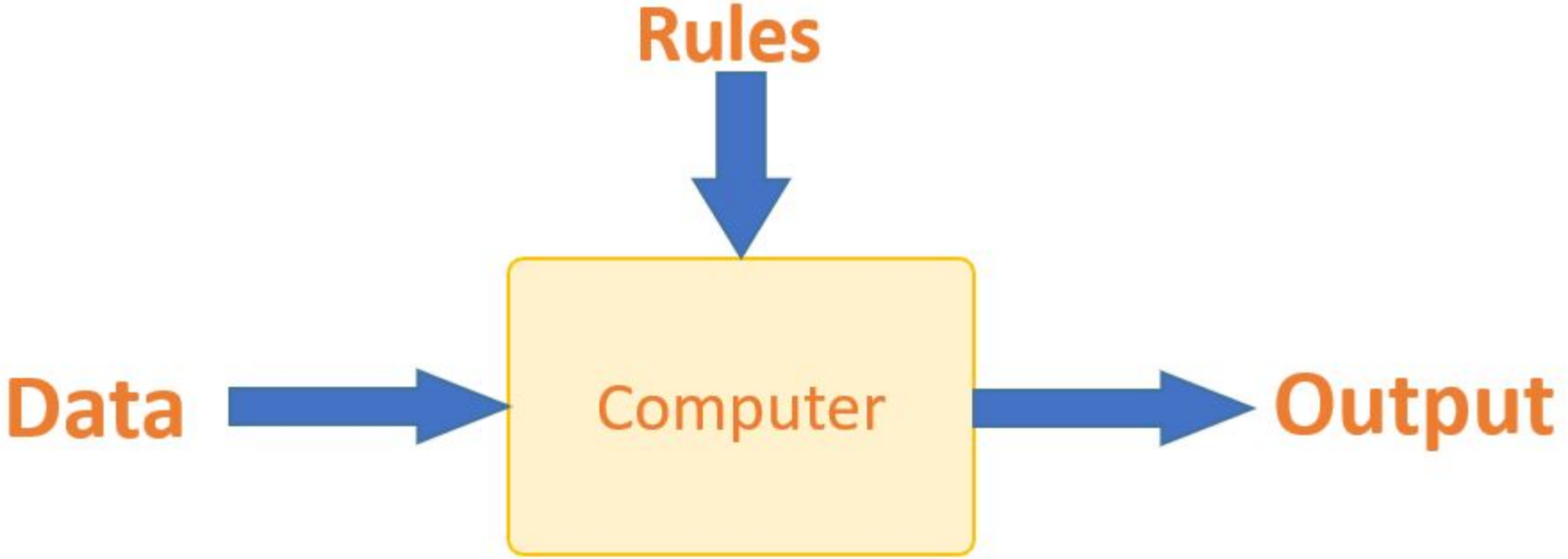
1. Introduction

What is machine learning?

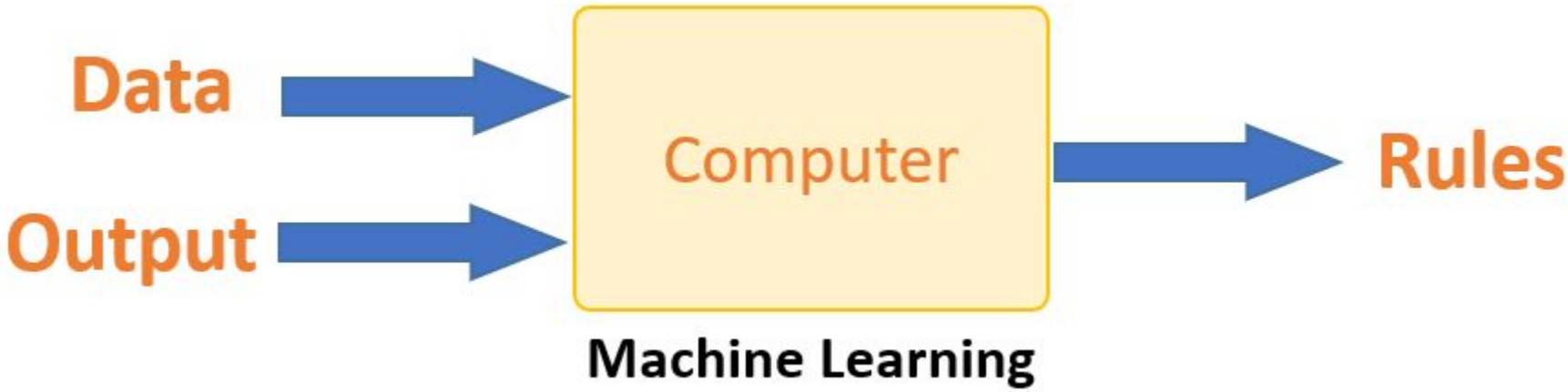
Arthur Samuel (1949)

“Suppose we arrange for some automatic means of testing the effectiveness of any current weight assignment in terms of actual performance and provide a mechanism for altering the weight assignment so as to maximize the performance. We need not go into the details of such a procedure to see that it could be made entirely automatic and to see that a machine so programmed would "learn" from its experience.”

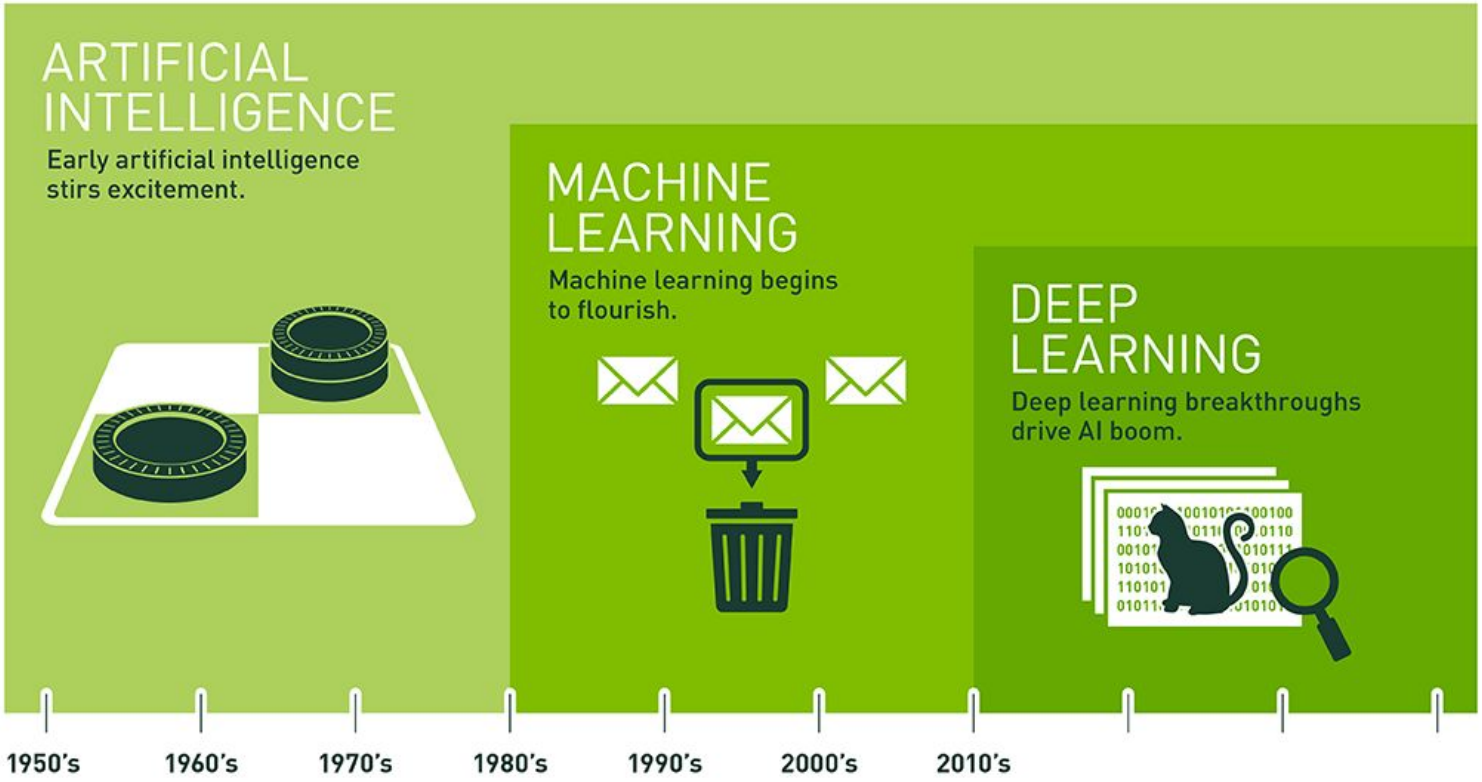
https://www.ibm.com/ibm/history/ibm100/us/en/icons/ibm700series/impacts/



http://www.fictgroup.com/knowledge/17/Differences%20between%20machine%20learning%20and%20software%20engineering



http://www.fictgroup.com/knowledge/17/Differences%20between%20machine%20learning%20and%20software%20engineering



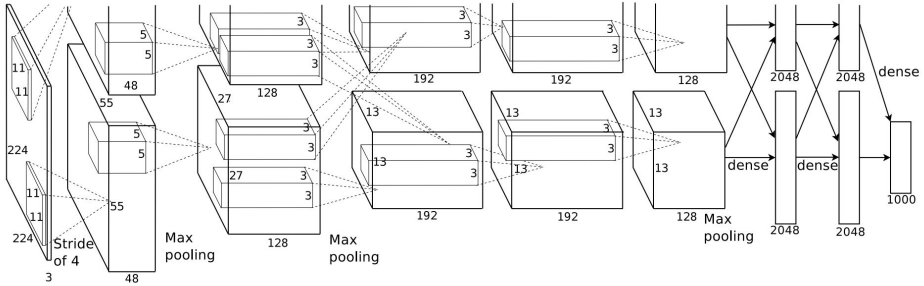
https://docs.paperspace.com/machine-learning/

Since the optimistic beginnings of AI in the 1950s, small subfields of AI - first Machine Learning and then Deep Learning (a subfield of Machine Learning) have made huge milestones.

Deep learning milestones

AlexNet (Krizhevsky et al., 2013)

AlexNet is a classic convolutional neural network architecture. It consists of convolutions, max pooling and dense layers as the basic building blocks. Grouped convolutions are used in order to fit the model across two GPUs.



https://paperswithcode.com/method/alexnet#

AlphaGo (DeepMind, 2016)

https://www.theguardian.com/technology/2016/mar/15/alphago-what-does-google-advanced-software-go-next

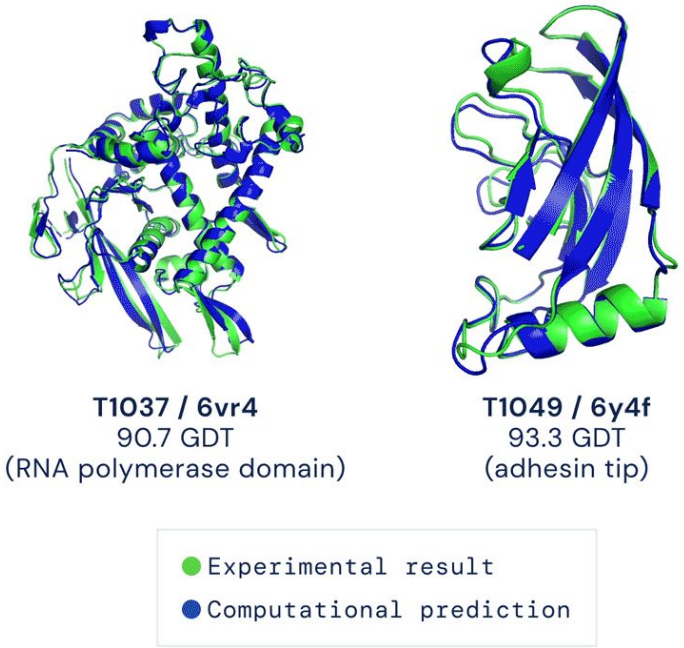
2018 Turing award

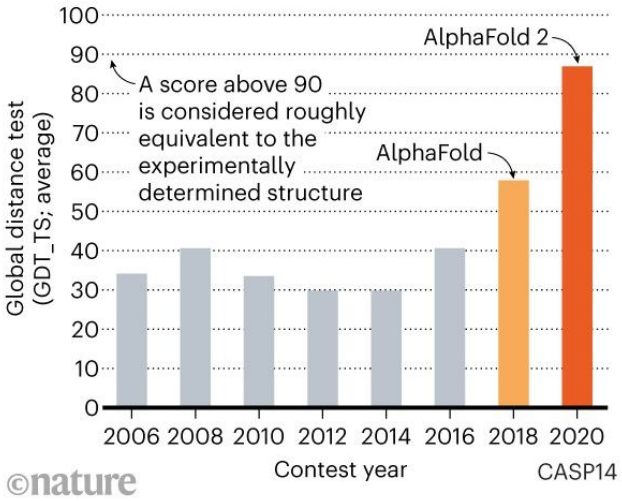
Yoshua Bengio

Geoffrey Hinton

Yann LeCun

AlphaFold2 (DeepMind, 2020)





https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

Speech recognition - Smart speakers

https://www.socialup.it/siri-alexa-google-home-conversano-loop-infinito/

Self driving cars

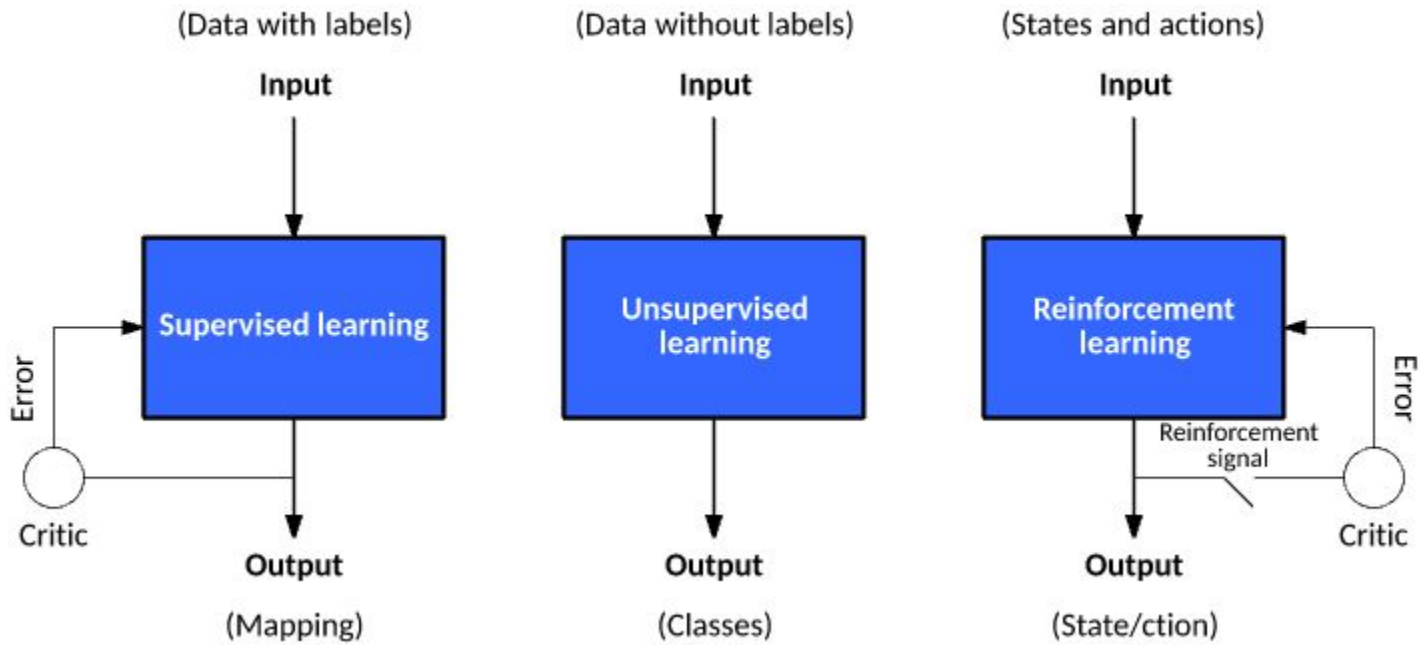
https://medium.com/the-future-is-electric/autonomous-cars-will-increase-urban-road-congestion-91bd05547a93

Basics on Machine/Deep Learning

Jargon

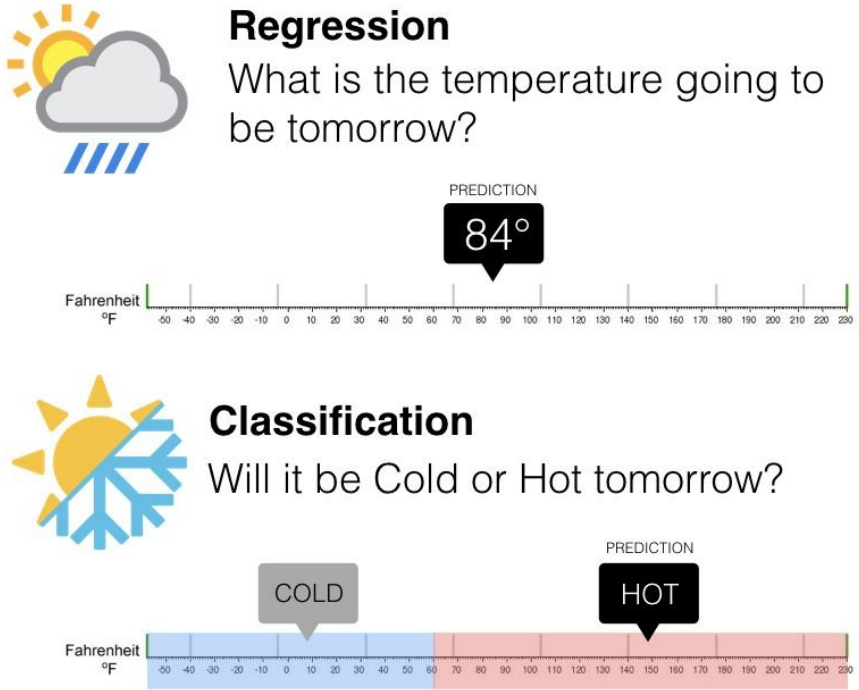
- the functional form of a model -> architecture
- the weights -> parameters
- the results -> predictions
- the measure of performance -> loss
- the dependent variable -> targets, y (labels in the context of classification)

Types of Machine learning



https://www.semanticscholar.org/paper/End-user-Robot-Programming-in-Cobotic-Environments-Liang/f457c9ebceb55b6849057ac93fd7cc8db0953d48

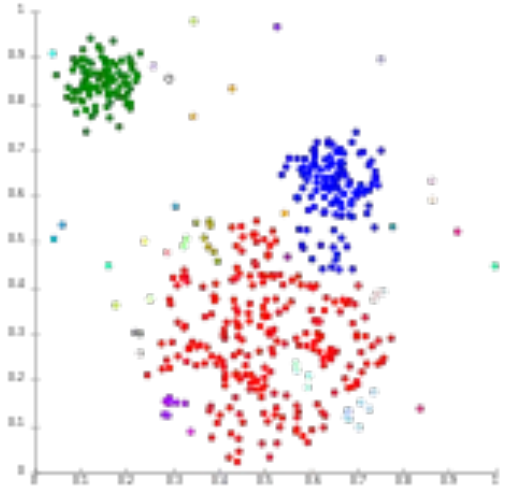
Tasks in supervised learning



https://www.pinterest.com/pin/523332419198760631/

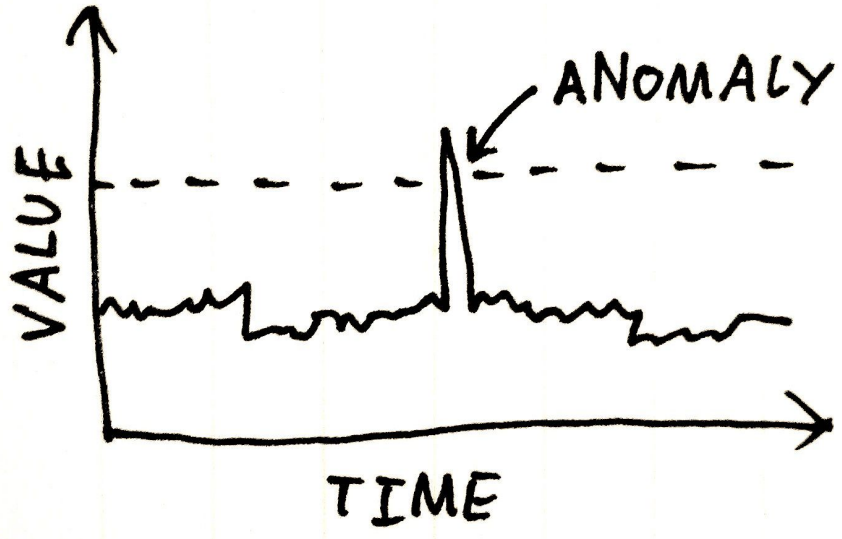
Tasks in unsupervised learning

Clustering



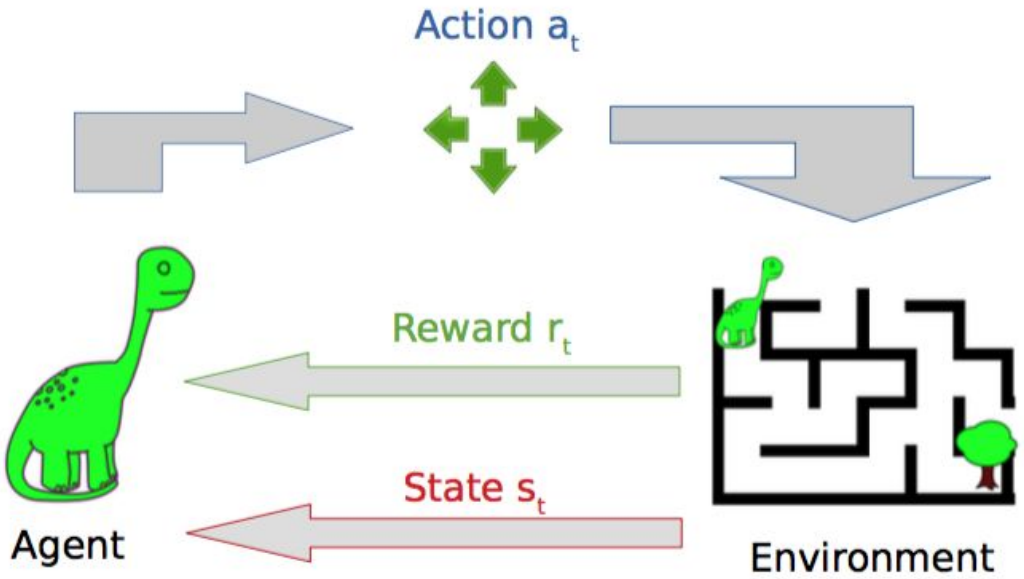
https://en.wikipedia.org/wiki/Cluster_analysis

Anomaly detection



https://github.com/project-anomalia/anomalia/blob/master/time_series.go

Reinforcement learning

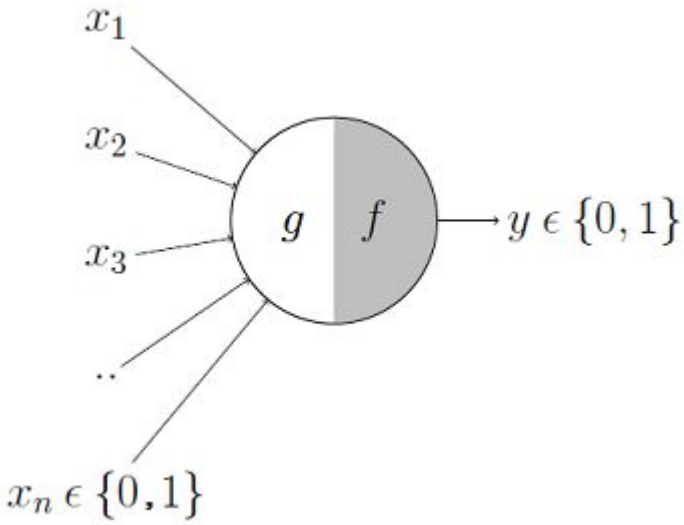


https://twitter.com/hashtag/deepreinforcementlearning

Deep learning = Machine learning with Neural networks

Model of an artificial neuron (1943)

- Binary input data (x)
- Aggregation function (g)
- Activation function (f)



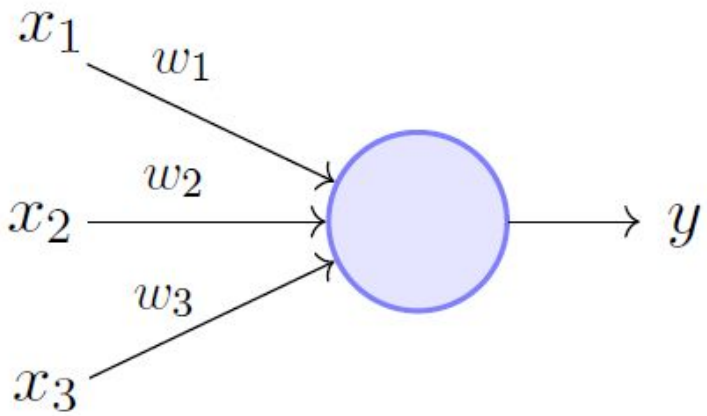
$$y = 1 \quad if \sum_{i=0}^{n} x_i \geq 0$$

$$= 0 \quad if \sum_{i=0}^{n} x_i < 0$$

https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1

The perceptron (1958)

- Inputs can have numeric values, other than 0 and 1
- Inputs have associated **weights**

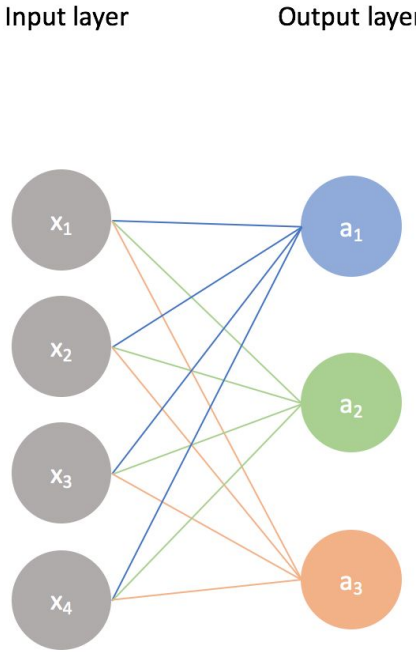


$$y = 1 \quad if \sum_{i=0}^{n} w_i * x_i \geq 0$$

$$= 0 \quad if \sum_{i=0}^{n} w_i * x_i < 0$$

https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1

Matrix multiplication is all you need

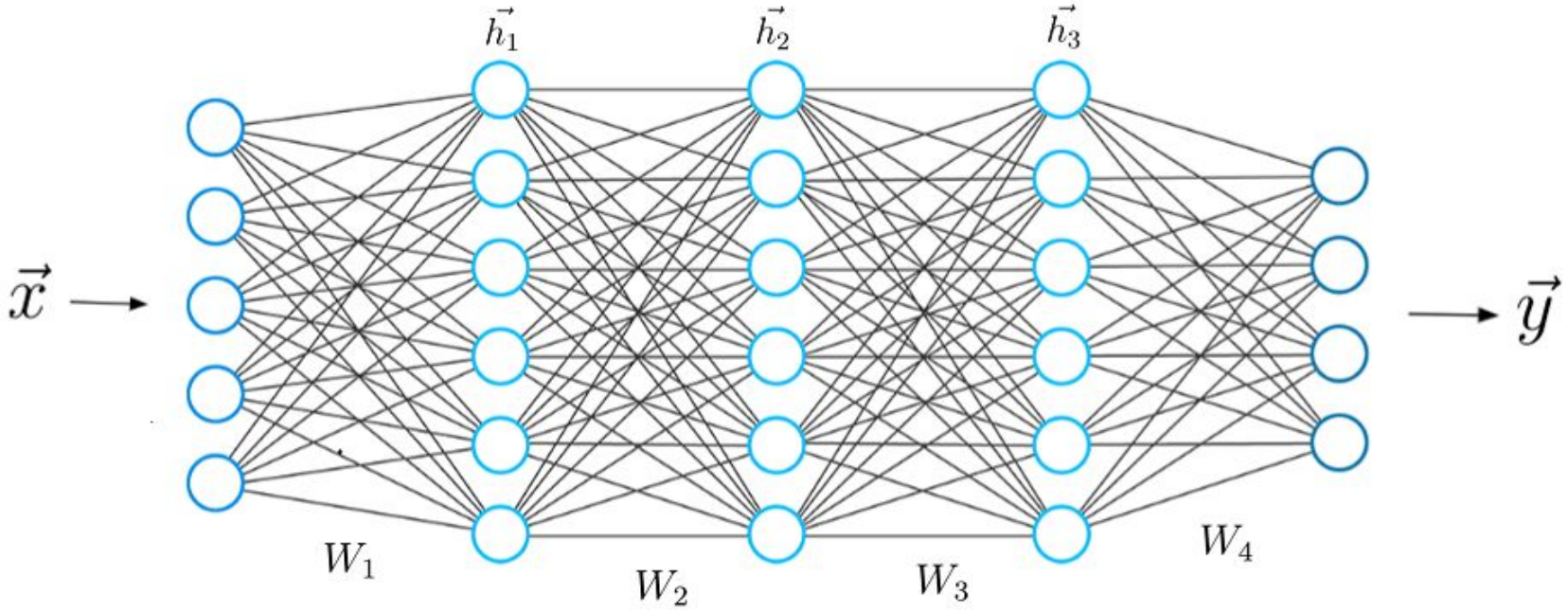


A simple neural network

$$\begin{bmatrix} w_1 & w_2 & w_3 & w_4 \\ w_1 & w_2 & w_3 & w_4 \\ w_1 & w_2 & w_3 & w_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} b \\ b \\ b \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b \\ w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b \\ w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b \end{bmatrix} \xrightarrow{activation} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

https://www.jeremyjordan.me/intro-to-neural-networks/

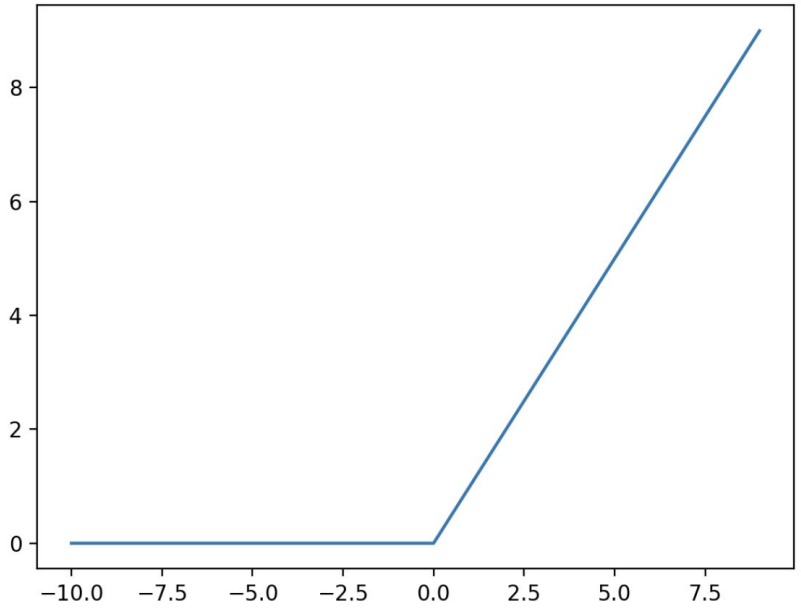
Going deep



https://www.deeplearning-academy.com/p/ai-wiki-machine-learning-vs-deep-learning

Activation functions

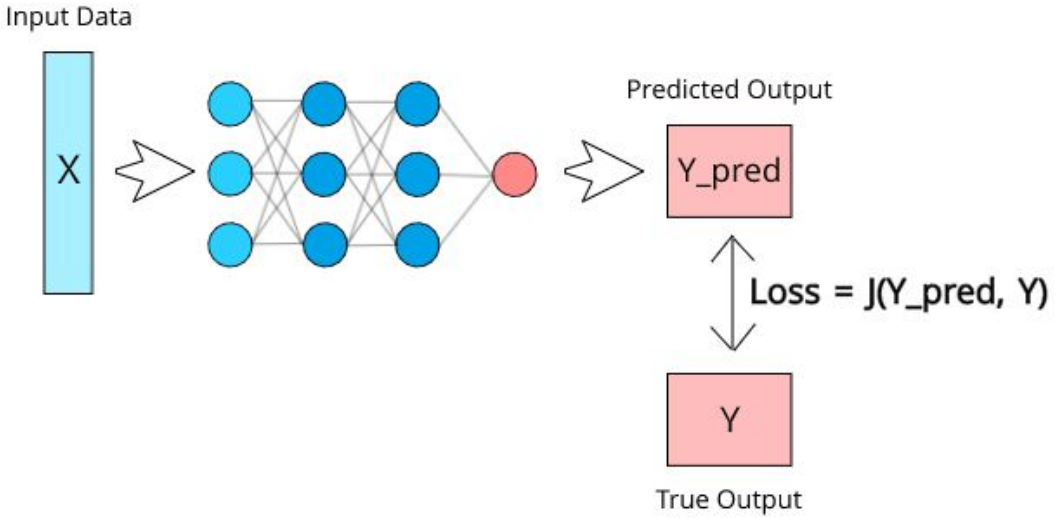
Non linearities applied after the matrix multiplication. The rectified linear unit, or ReLU, has been the most popular in the past decade.



Loss functions

Loss Functions are used to frame the problem to be optimized within deep learning. Most popular ones are:

- Cross Entropy Loss (classification)
- Mean Squared error (regression)

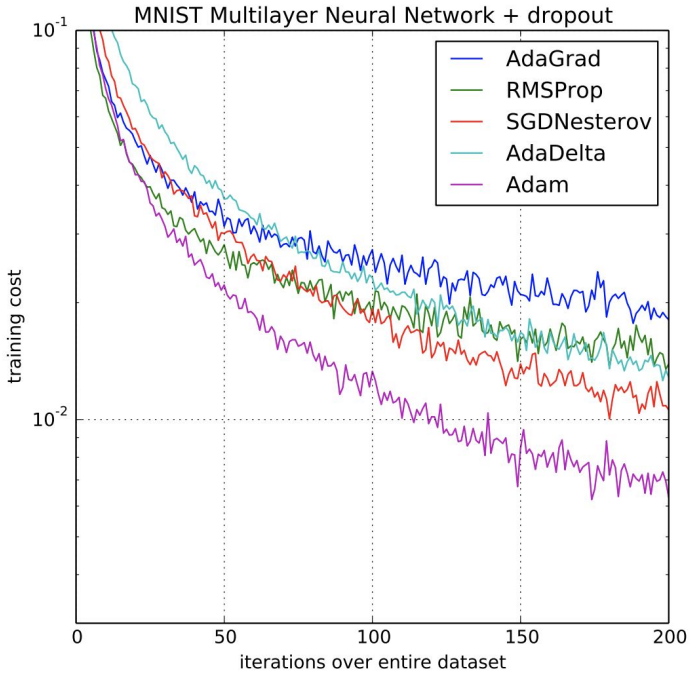


https://deeplearningdemystified.com/article/fdl-3

Stochastic optimization

Used to train neural networks. They iteratively take a “mini-batch” of data, hence ‘stochastic’, and perform gradient descent on the loss function for that batch. Most popular methods are:

- SGD
- Adam



https://paperswithcode.com/method/adam

Areas of success in deep learning: Computer vision

Image classification

Image Classification is one of the most popular tasks in computer vision. It attempts to comprehend an entire image as a whole.

https://paperswithcode.com/task/image-classification

Image segmentation

Semantic segmentation, or image segmentation, is the task of clustering parts of an image together which belong to the same object class.

https://paperswithcode.com/task/semantic-segmentation

Object detection

Object detection is the task of detecting instances of objects of a certain class within an image.

https://paperswithcode.com/task/object-detection

Image generation

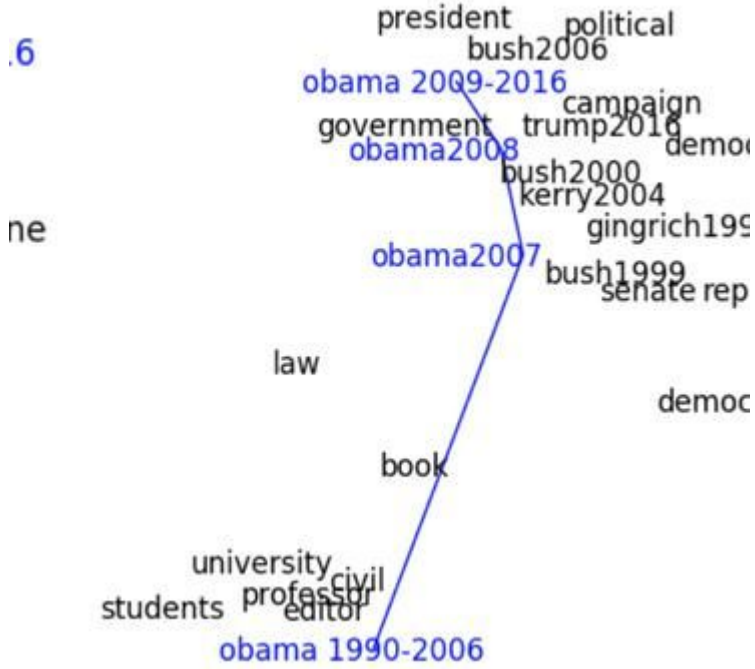
Image generation (synthesis) is the task of generating new images from an existing dataset.

https://paperswithcode.com/task/image-generation

Areas of success in deep learning: Natural Language Processing

Word embeddings

Word embedding is the collective name for a set of language modeling and feature learning techniques in natural language processing (NLP) where words or phrases from the vocabulary are mapped to vectors of real numbers.



https://paperswithcode.com/task/word-embeddings/latest

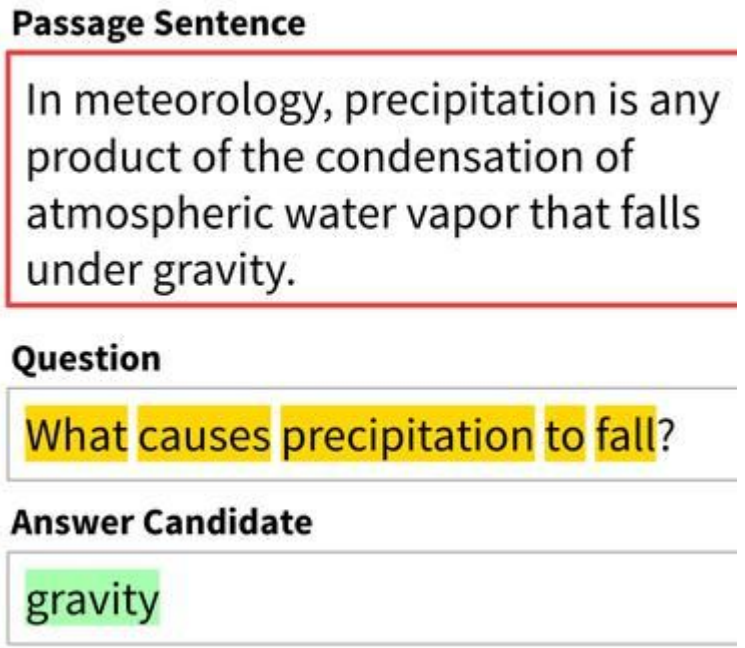
Language modelling

Language modeling is the task of predicting the next word or character in a document.

https://medium.com/@pierre_guillou/nlp-fastai-language-model-6444819837f0

Question Answering

Question Answering is the task of answering questions (typically reading comprehension questions), but abstaining when presented with a question that cannot be answered based on the provided context



https://paperswithcode.com/task/question-answering

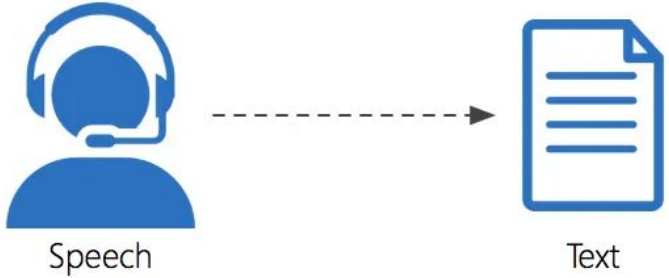
Other areas of success in deep learning

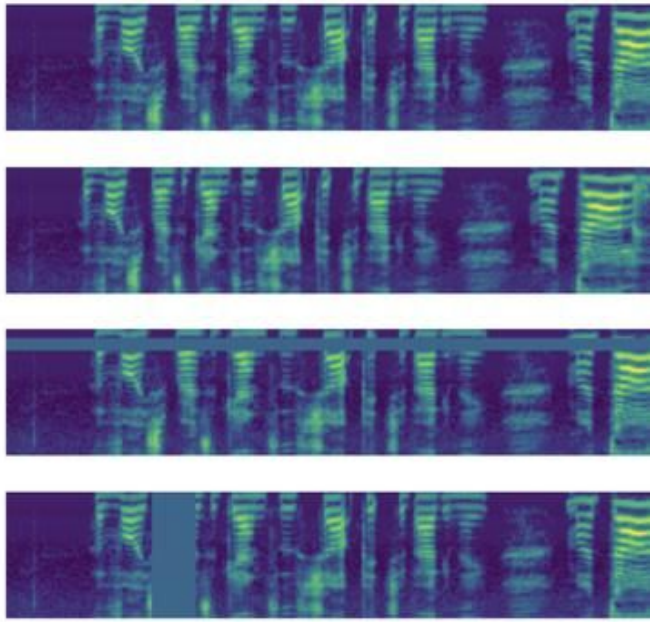
Recommender systems

https://twitter.com/hashtag/rinapiccolo

Speech Recognition

Speech recognition is the task of recognising speech within audio and converting it into text.





https://apkgk.com/com.knowledge.speech_1

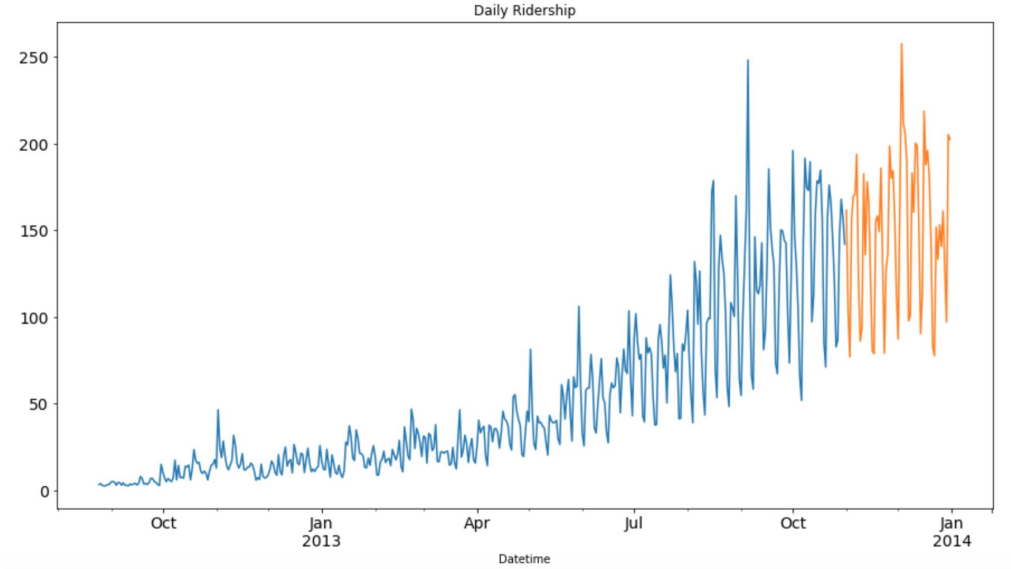
https://paperswithcode.com/task/speech-recognition/codeless?page=7

Speech Synthesis

Speech synthesis is the task of generating speech from some other modality like text, lip movements etc.

Time series forecasting

Time series forecasting is the task of predicting future values of a time series (as well as uncertainty bounds).



https://campus.datacamp.com/courses/introduction-to-portfolio-analysis-in-python/portfolio-optimization?ex=10

Playing games

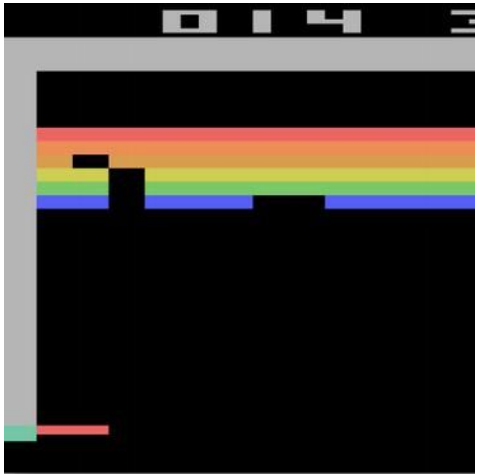
The games task involves training a reinforcement learning agent to play a game with the highest possible score.

https://paperswithcode.com/task/game-of-chess

https://paperswithcode.com/task/game-of-go/codeless

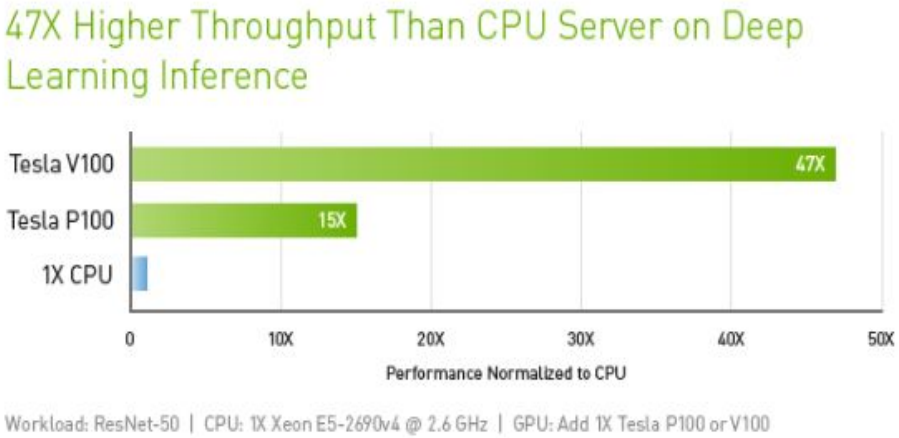
https://paperswithcode.com/task/atari-games

https://paperswithcode.com/task/starcraft-ii



Why is deep learning currently booming?

GPU costs & availability



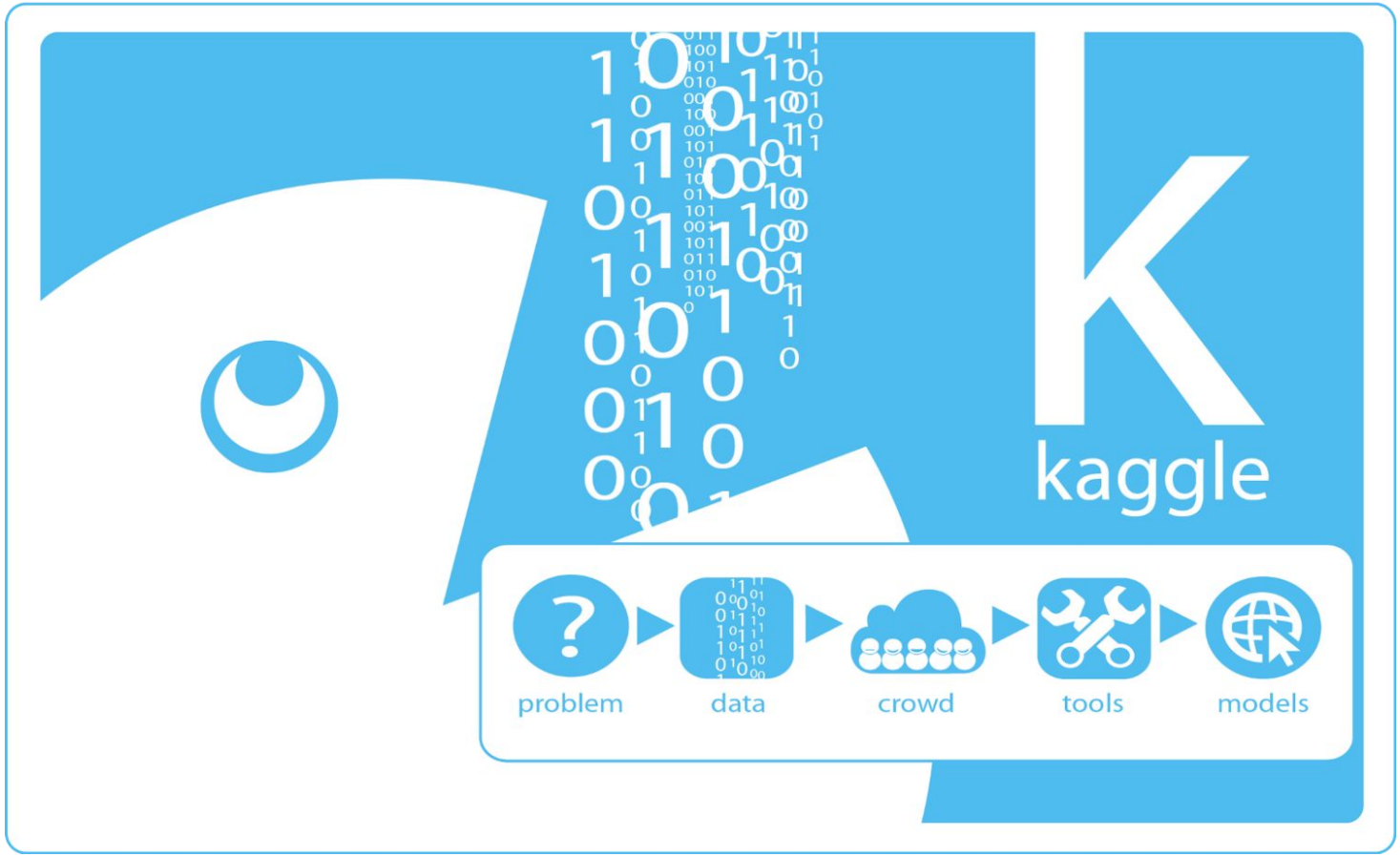
http://nvidia.ihuidian.com/product/detail_17.html

Deep learning frameworks

https://docs.paperspace.com/machine-learning/wiki/comparison-of-ai-frameworks

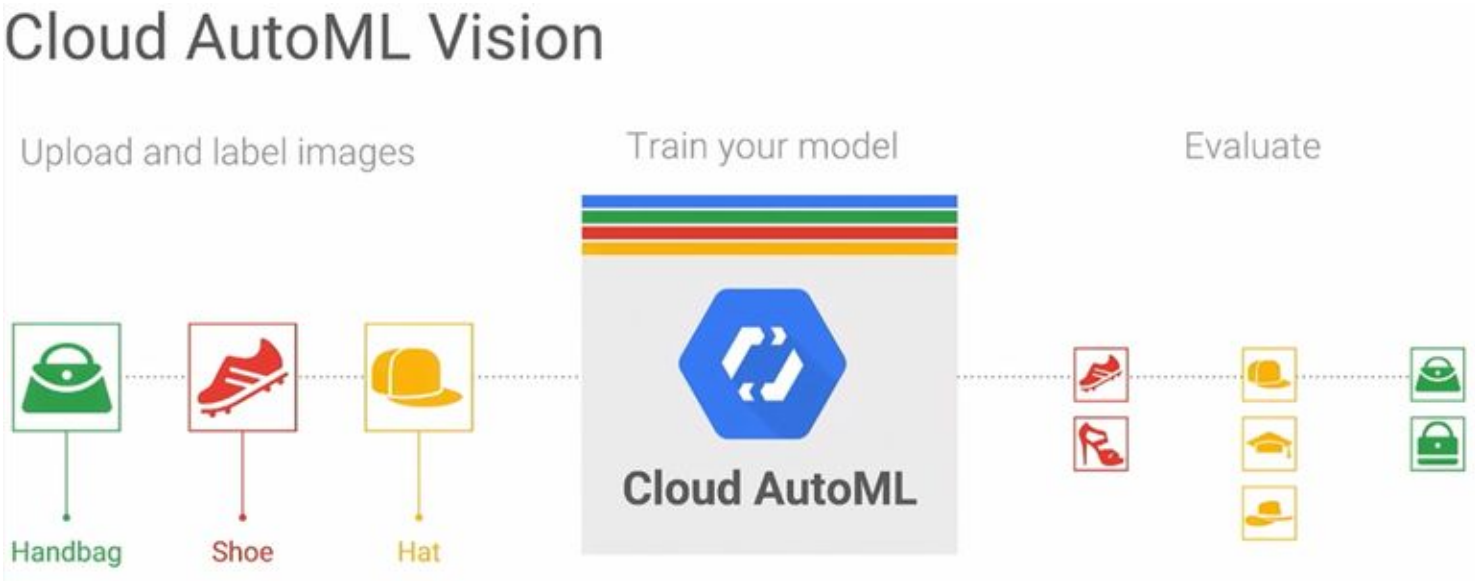
Free learning courses

https://twitter.com/antor/status/1043798393309597696?s=20



https://medium.com/hackernoon/kaggle-kickstarter-of-machine-learning-world-567ba3d39953

AutoML



https://www.zdnet.com/article/google-launches-cloud-automl-an-effort-to-simplify-and-automate-the-grunt-work-behind-ai-and-machine/

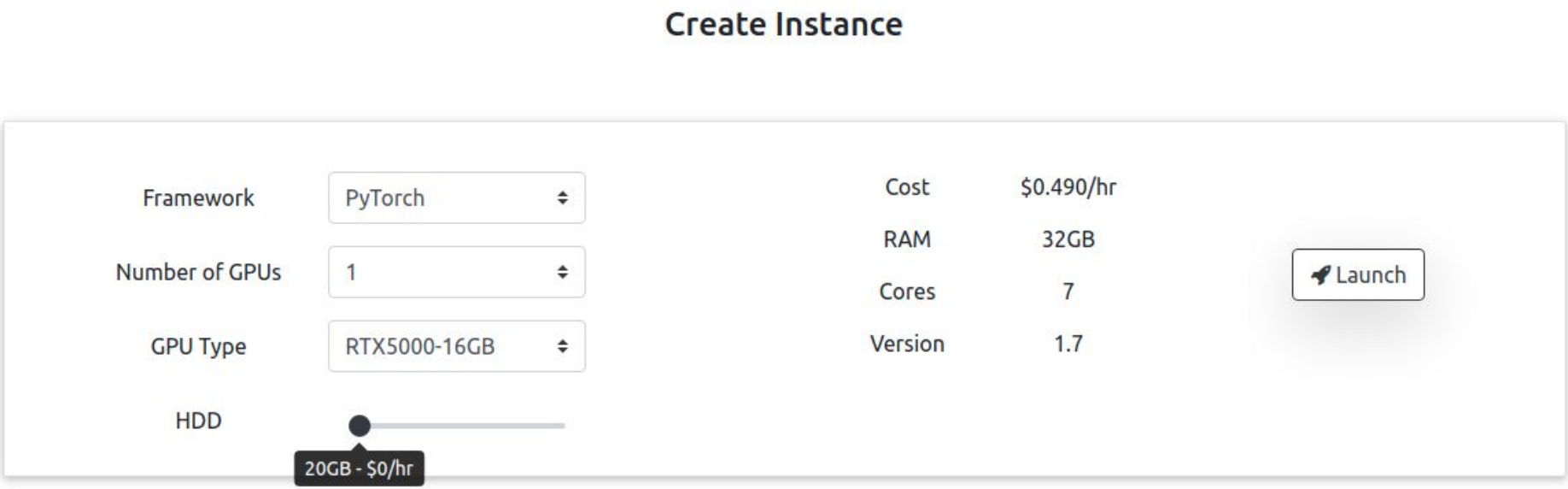
2. Getting Started

Getting a GPU deep learning server

Free online platforms with GPUs

- Google Colab
  - Storage on Google drive account
  - Lots of pre-installed libraries for ML
- Paperspace Gradient
  - Full Jupyter Notebook instance
  - Provides some space to save notebooks and models (5 GB)
  - Only 1 Free Notebook can be run at a time
  - All Notebooks will be set to public and cannot be set to private
- BlazingSQL Notebooks
  - 26 GB free HDD
  - JupyterLab instance

gradient
by *Paperspace*



https://cloud.jarvislabs.ai/

Full remote dedicated servers with GPUs (Datacrunch.io, GoogleCloud, JarvisCloud, ...)

Deep learning 101 with Pytorch and fastai

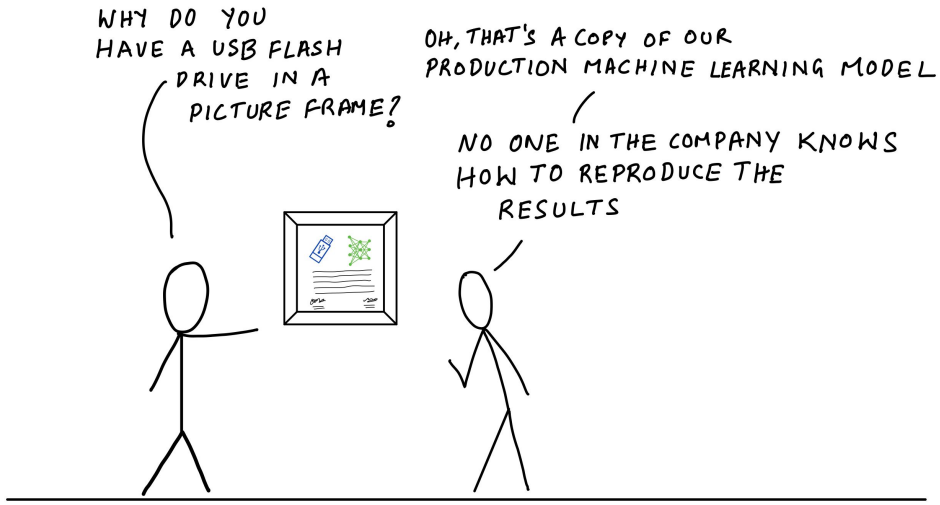
https://github.com/stardust-r/walk-with-deep-learning

3. Practical tips and best practices

Experiment tracking

Who needs Experiment tracking and why?

- Single scientists
  - Coming back to old ideas
  - Compare and visualize runs
  - Hyperparameter tuning
- Teams
  - Share ideas and insights
  - Store experiment metadata
  - Onboarding new members easier



https://towardsdatascience.com/a-quick-guide-to-managing-machine-learning-experiments-af84da6b060b

Tools for ML experiment tracking

See a comparative in: https://neptune.ai/blog/best-ml-experiment-tracking-tools

1. Neptune
2. **Weights & Biases**
3. Comet
4. Sacred
5. MLFlow
6. TensorBoard
7. ...

Weights & Biases (wandb)

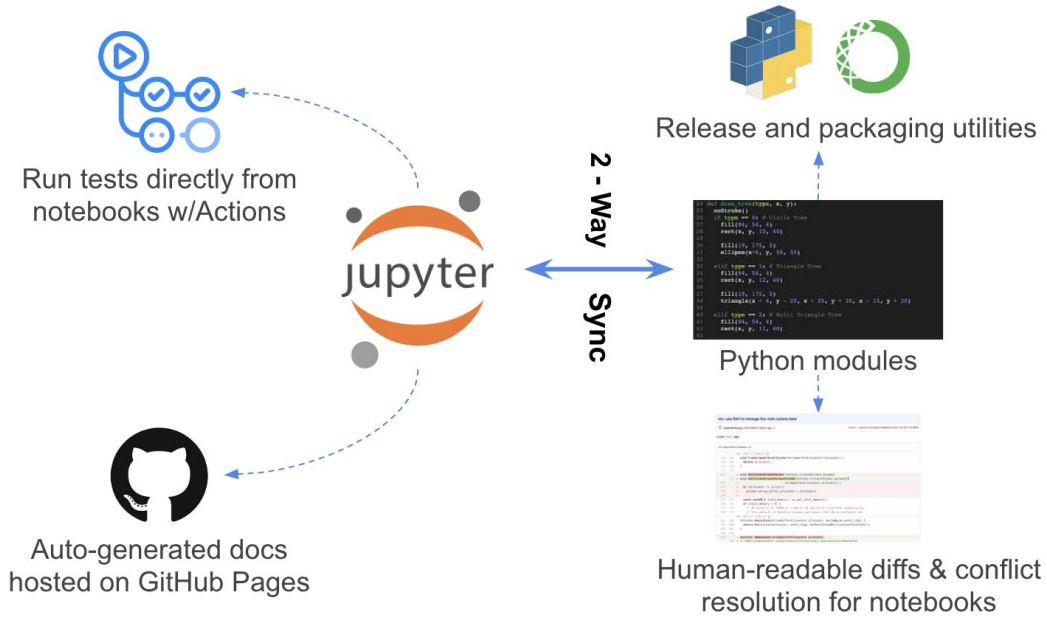
- Created for deep learning experiment tracking
- Easy integration with the most popular ML libraries (Tensorflow, Kerasm Pytorch, fast.ai, scikit-learn, ...)
- Customizable visualisation and reporting tools

Weights & Biases

See wandb project for this talk:
https://wandb.ai/vrodriguezf/TS-III

Software development in Jupyter Notebooks

Nbdev overview



https://github.com/fastai/nbdev

Conflict resolution for notebooks

```
<<<<<<< HEAD
```

In [5]:
```
path = Path.cwd()
path.ls()[0]
```

Out[5]: PosixPath('/home/sgugger/notebooks/01_core.ipynb')

```
=======
```

In [3]:
```
path = Path.cwd()
path.ls()[-1]
```

Out[3]: PosixPath('/home/jphoward/notebooks/01_core.ipynb')

```
>>>>>>> a7ec1b0bfb8e23b05fd0a2e6cafcb41cd0fb1c35
```

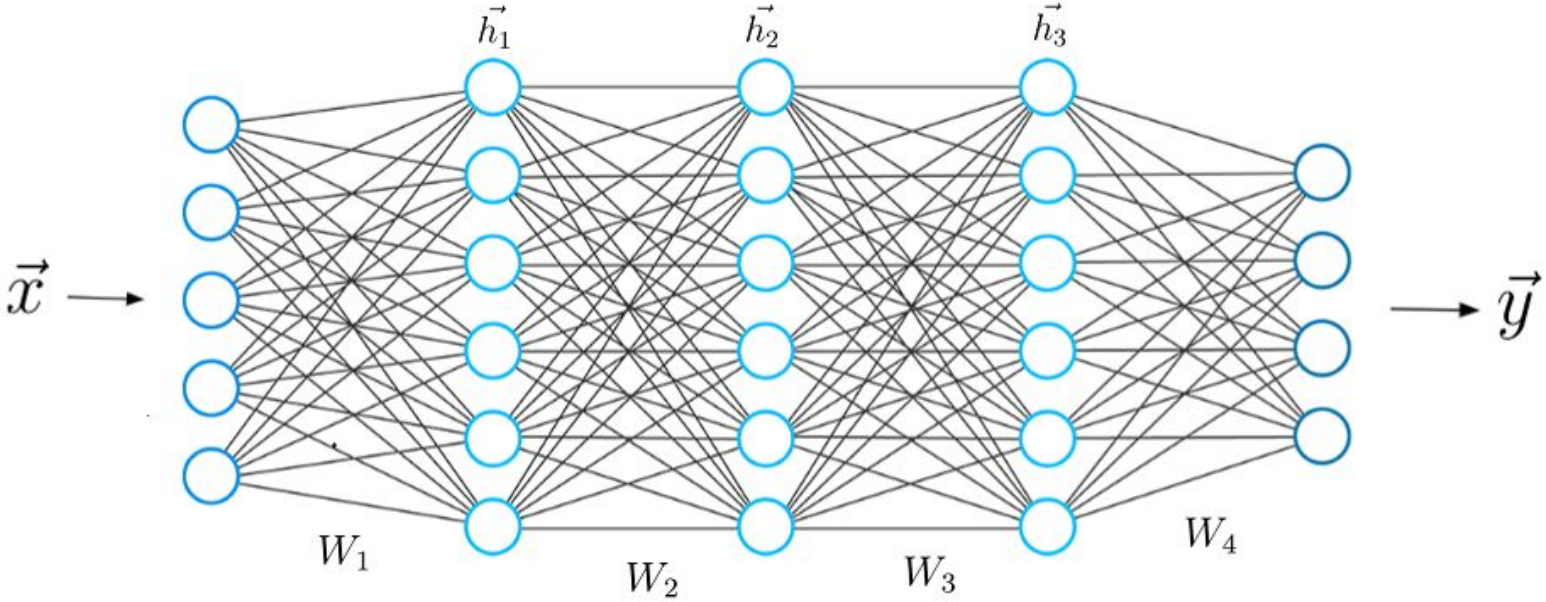
See an example of a nbdev project in: https://github.com/stardust-r/walk-with-deep-learning

Day 2

A walk with deep learning

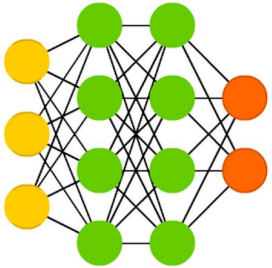
Going deeper into deep learning

Going deeper

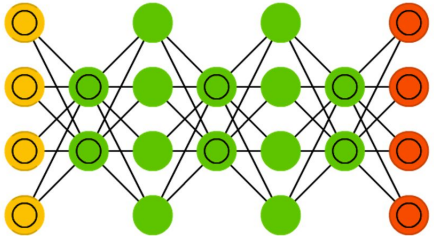


https://www.deeplearning-academy.com/p/ai-wiki-machine-learning-vs-deep-learning

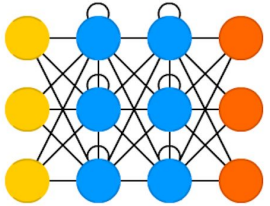
Deep architectures



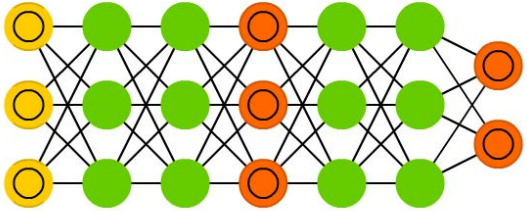
Deep Neural Network
(Fully connected layers)



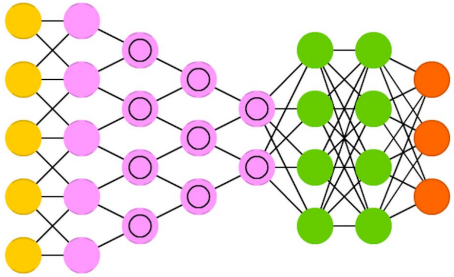
Deep Belief Network



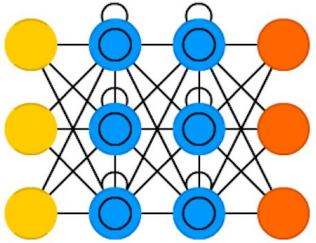
Recurrent Neural Network



GANs



Convolutional Neural Network

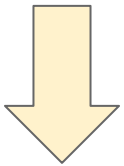


LSTMs

https://www.deeplearning-academy.com/p/ai-wiki-machine-learning-vs-deep-learning

Going deeper

An overview of neural networks architectures

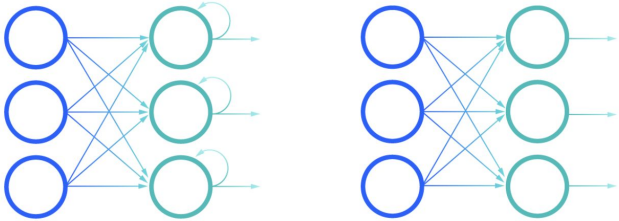


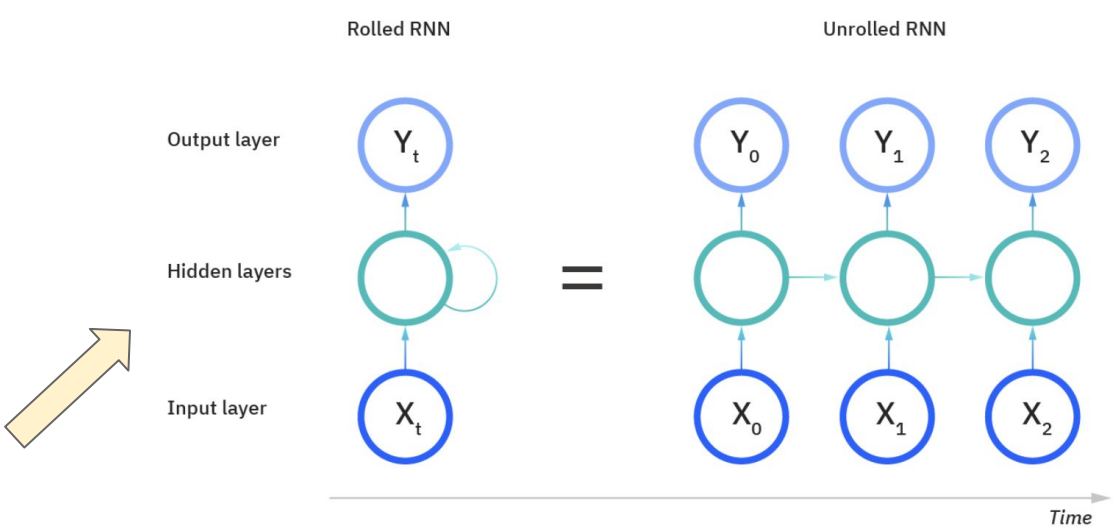
https://www.asimovinstitute.org/neural-network-zoo/

Recurrent Neural Network

In temporal series, sequential or ordinal data, we need memory!

RNN save information from previous states

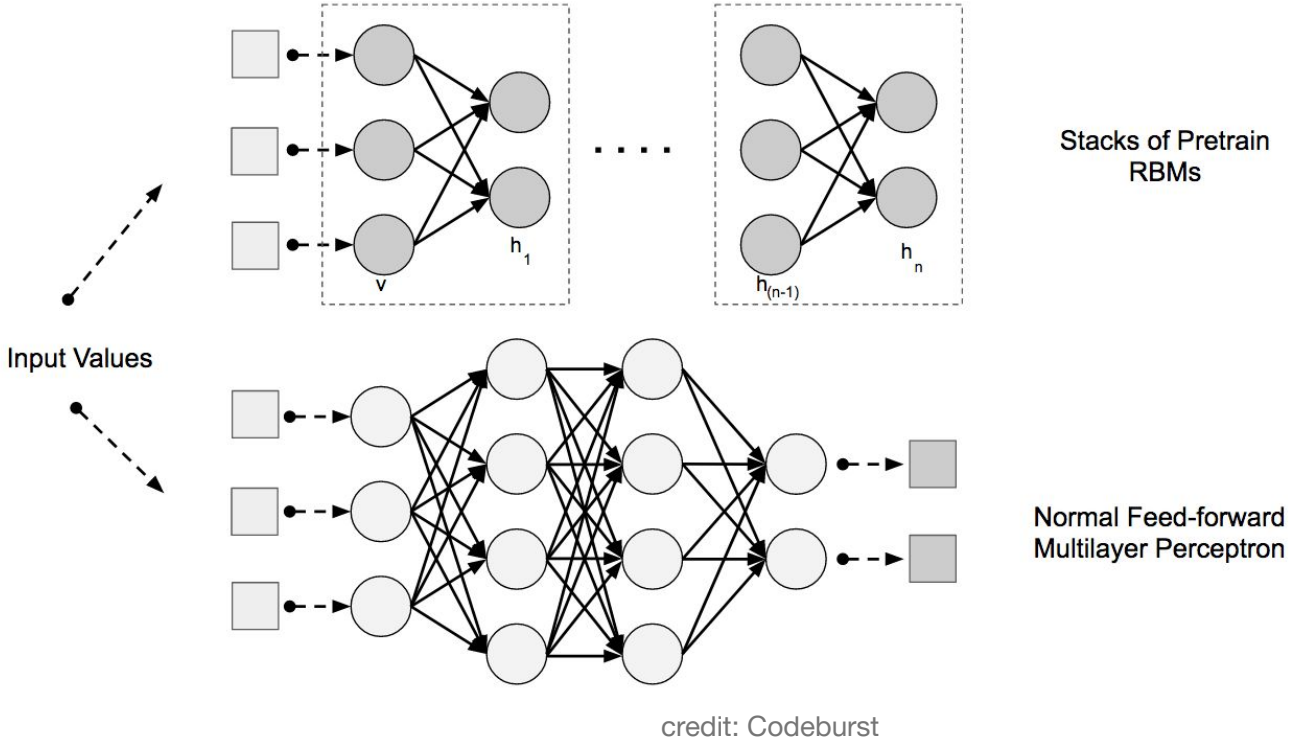




https://www.ibm.com/cloud/learn/recurrent-neural-networks

Deep Belief Networks

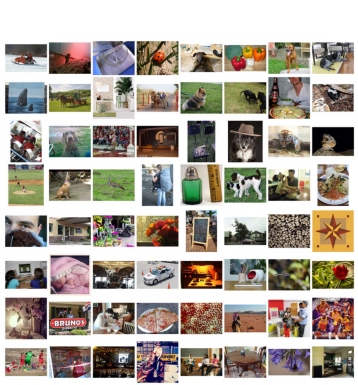
DBNs can be described as a stack of Restricted Boltzmann Machines (RBMs)

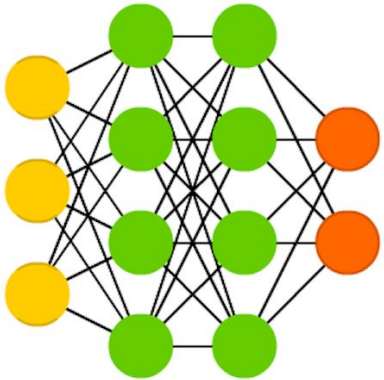


credit: Codeburst

Addressing image classification tasks

DNNs infer knowledge from complex non-linear relationships between the inputs

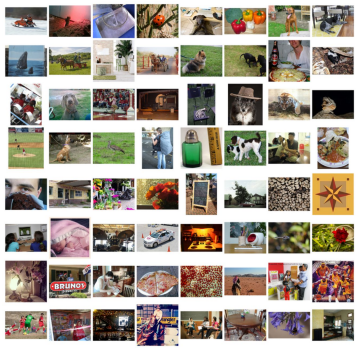




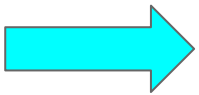
But they do not infer spatial relationships

https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

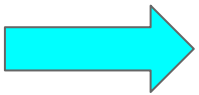
Addressing image classification tasks

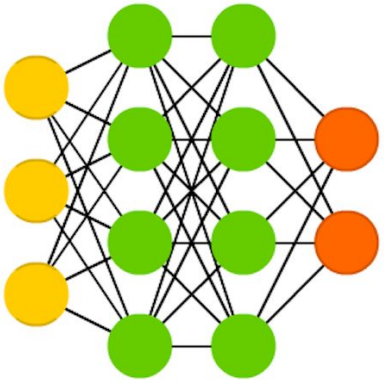


Input data



Feature extraction (by hand in classical Machine Learning)



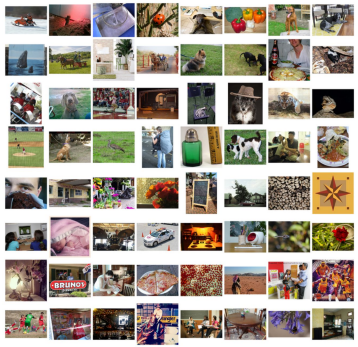


Knowledge inference (i.e. image classification)

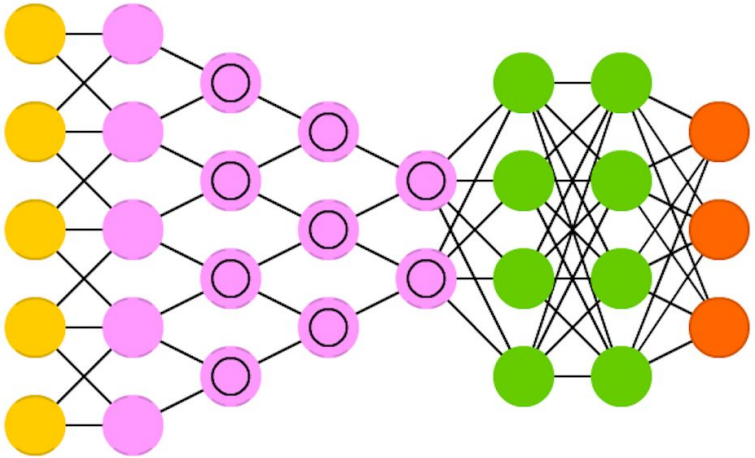
https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

Convolutional Neural Networks

Convolutional Neural Networks



Input data



Feature extraction
+
Knowledge inference

https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

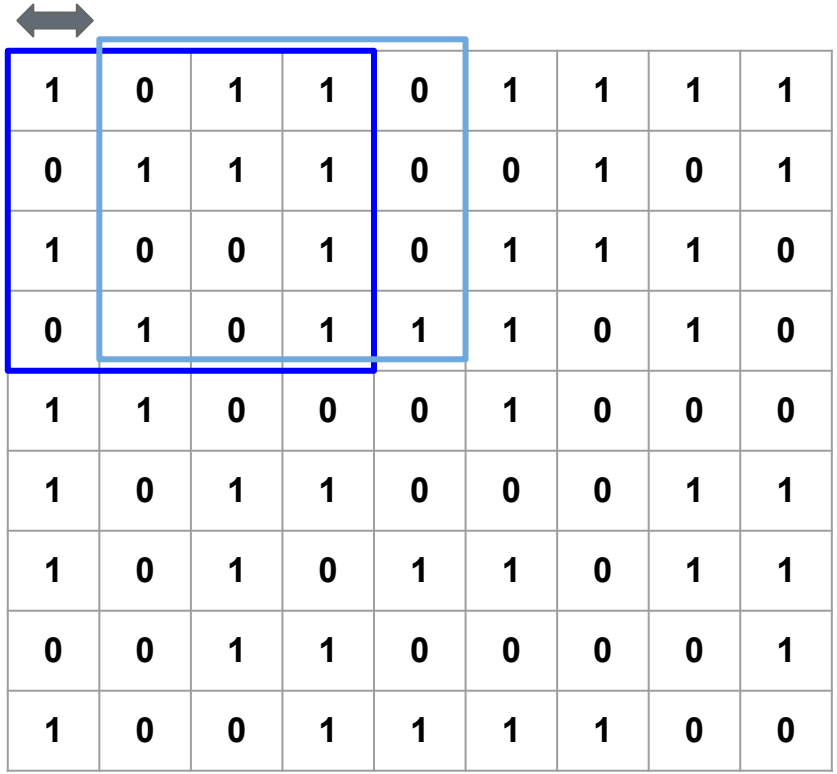
Convolutional Neural Networks

Also called shift invariant networks

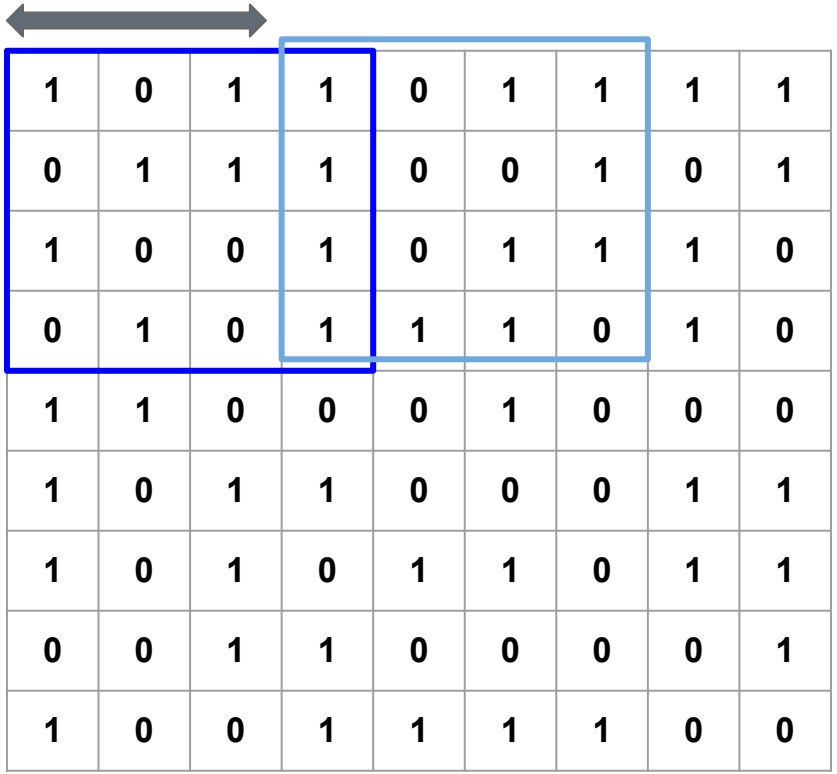
They support translation, rotation or size variances

https://stats.stackexchange.com/questions/208936/what-is-translation-invariance-in-computer-vision-and-convolutional-neural-netwo

The convolution operation: **stride**



Stride = 1



Stride = 3

The convolution operation: **stride**

If a filter size / stride combination does not match the size of the image

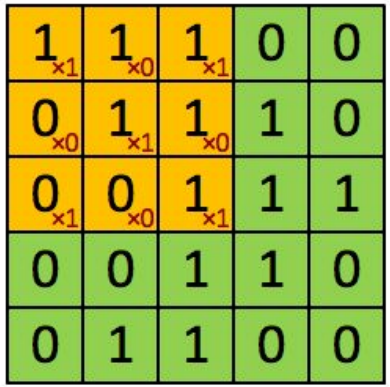
- **Padding**
- Reduce filter size
- Drop pixels

Convolutional Neural Networks

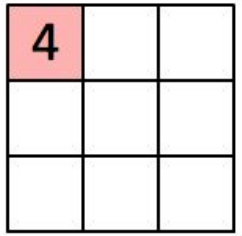
The convolution operation

Filter or kernel:

1	0	1
0	1	0
1	0	1



Image

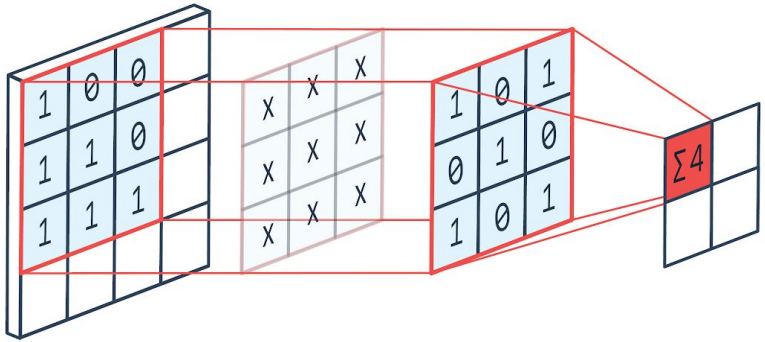


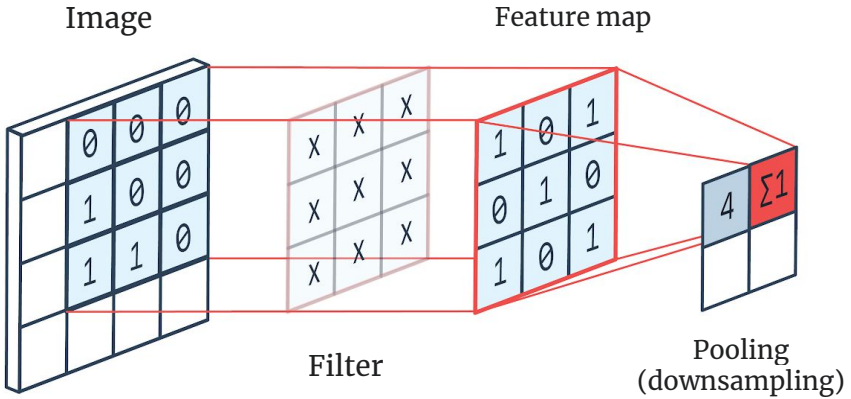
Convolved Feature

https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Convolutional Neural Networks

Convolution + Pooling

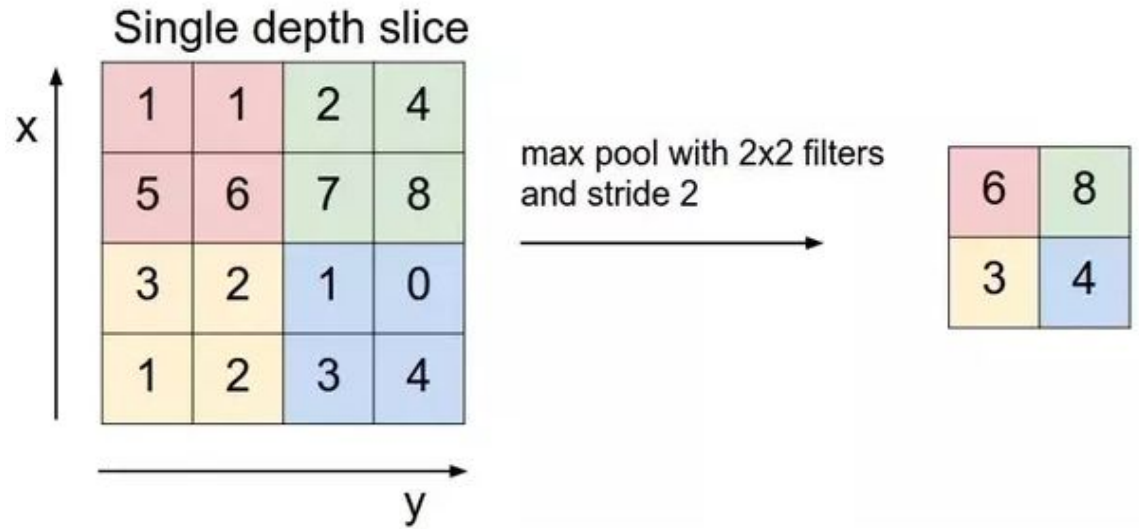




https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/blocks/2d-convolution-block

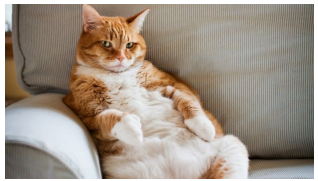
Convolutional Neural Networks

MaxPooling



https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

Architecture of a CNN



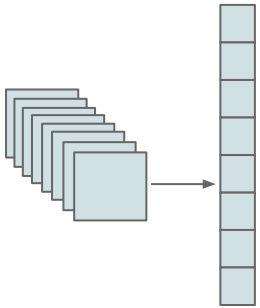
Input image

n times

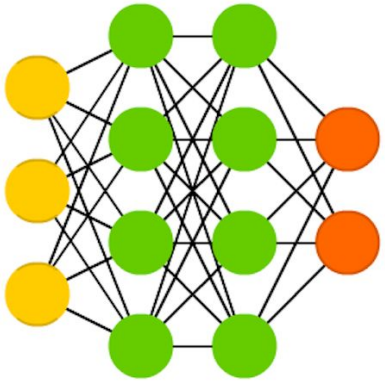
Convolution

Activation

Pooling



Flatten layer

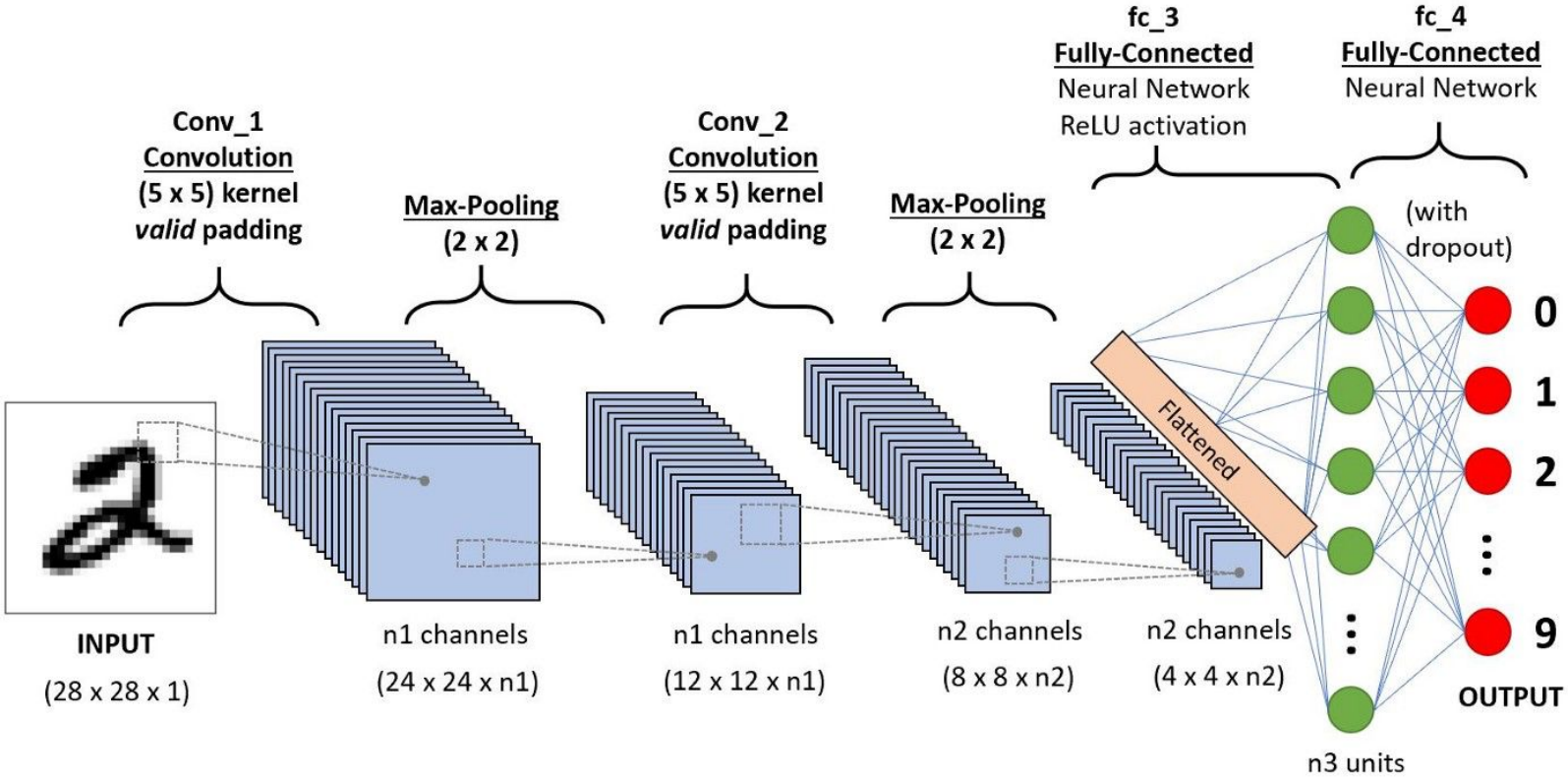


DNN

Feature extraction

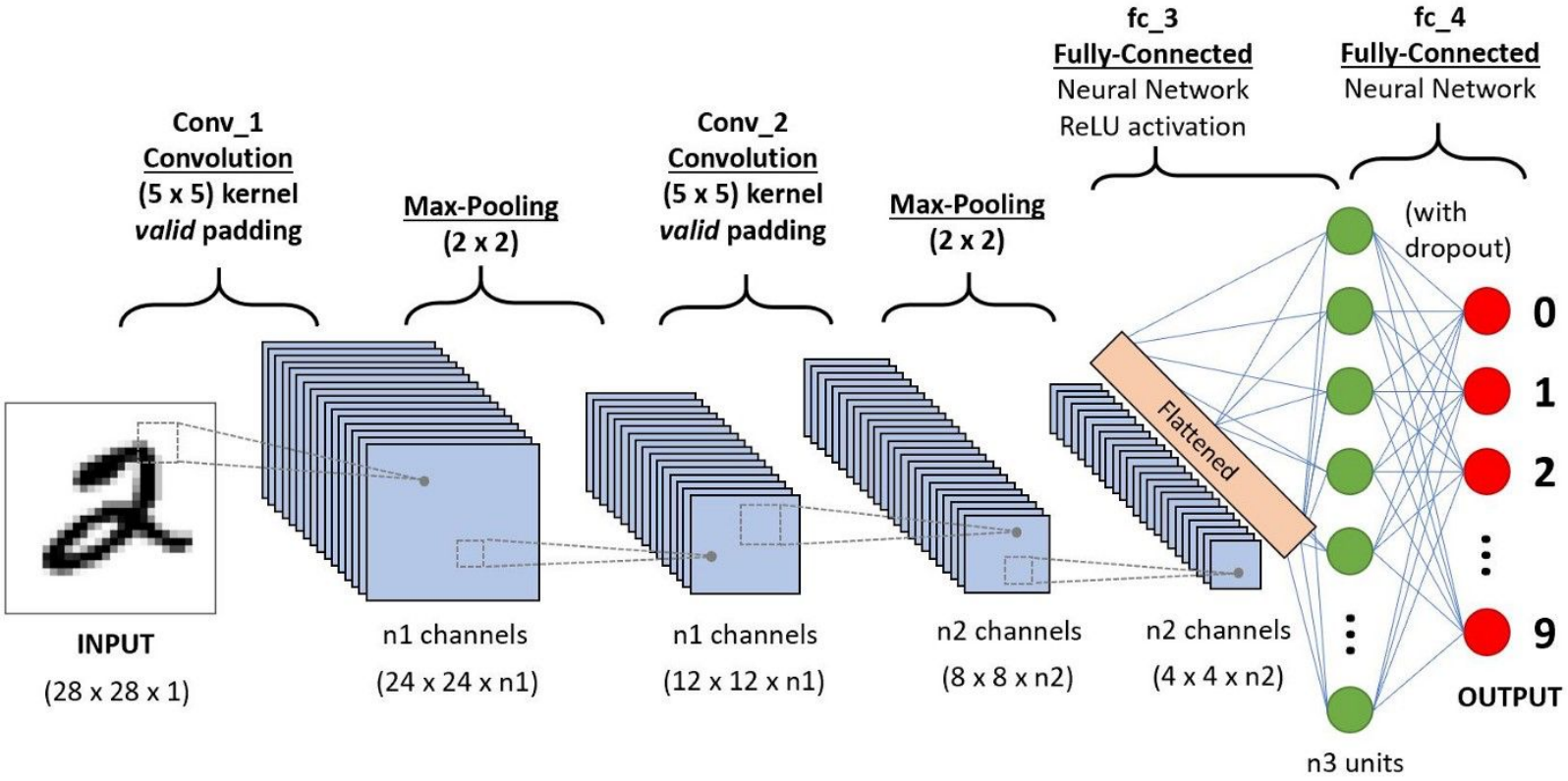
Knowledge inference

Architecture of a CNN



https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Architecture of a CNN

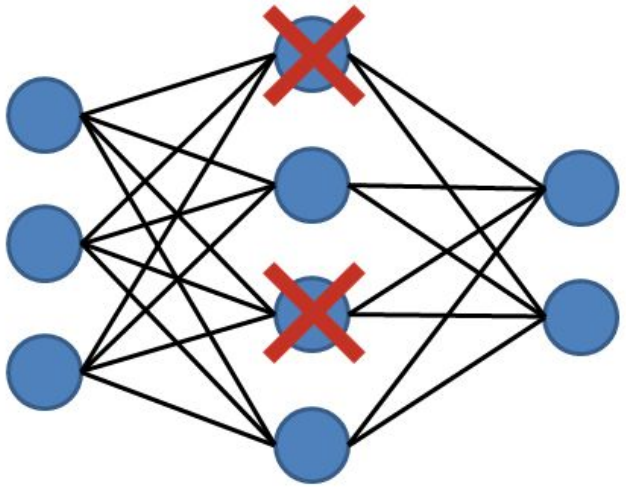


High level to lower level features

Dropout layer

Preventing overfitting

A set of random units are set to 0. This improves the performance of every unit and avoids “isolated units” with wrong knowledge

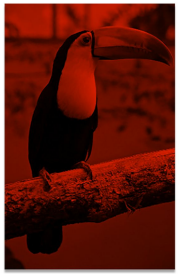


Channels

Grayscale image
1 channel

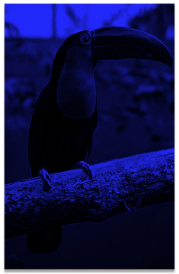
Data with size = $w * h$

Original Image



Red Channel

Green Channel



Blue Channel

Color image
3 channel

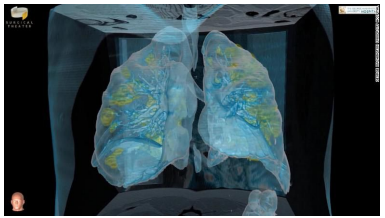
Data with size = $w * h * c$

Conv1D, Conv2D, Conv3D, ...

CNNs are typically applied to 2D data, but can be applied to other number of dimensions

1D Audio

2D Image



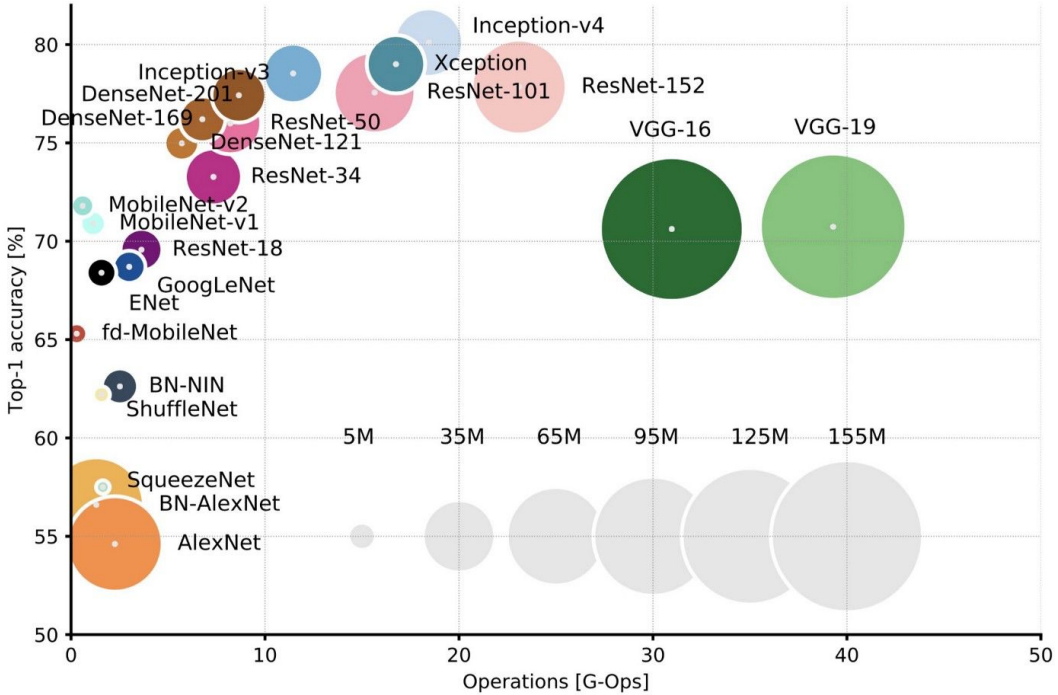
3D Image

4D spatio-temporal data

Weaknesses of CNNs

Is it a duck or a rabbit?

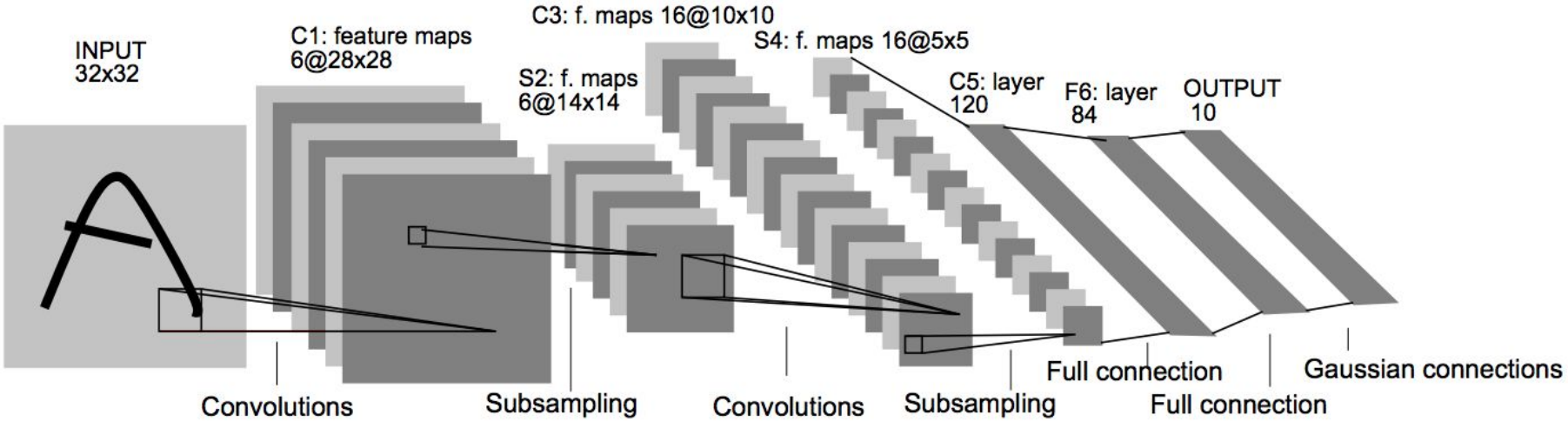
CNN architectures



https://towardsdatascience.com/neural-network-architectures-156e5bad51ba

CNN architectures

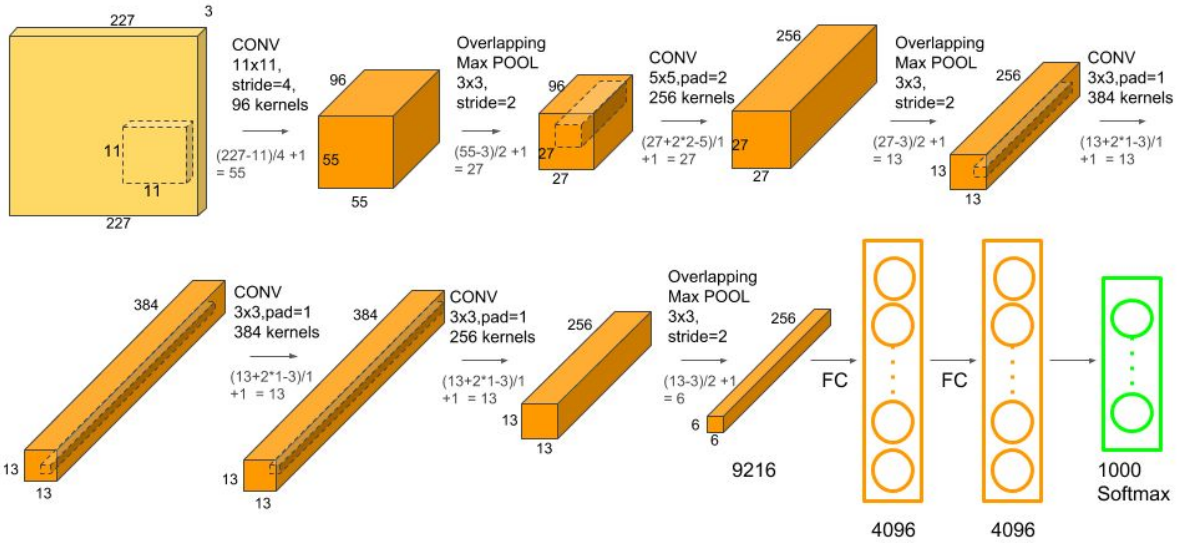
LeNet-5 (1998)



https://www.jeremyjordan.me/convnet-architectures/

CNN architectures

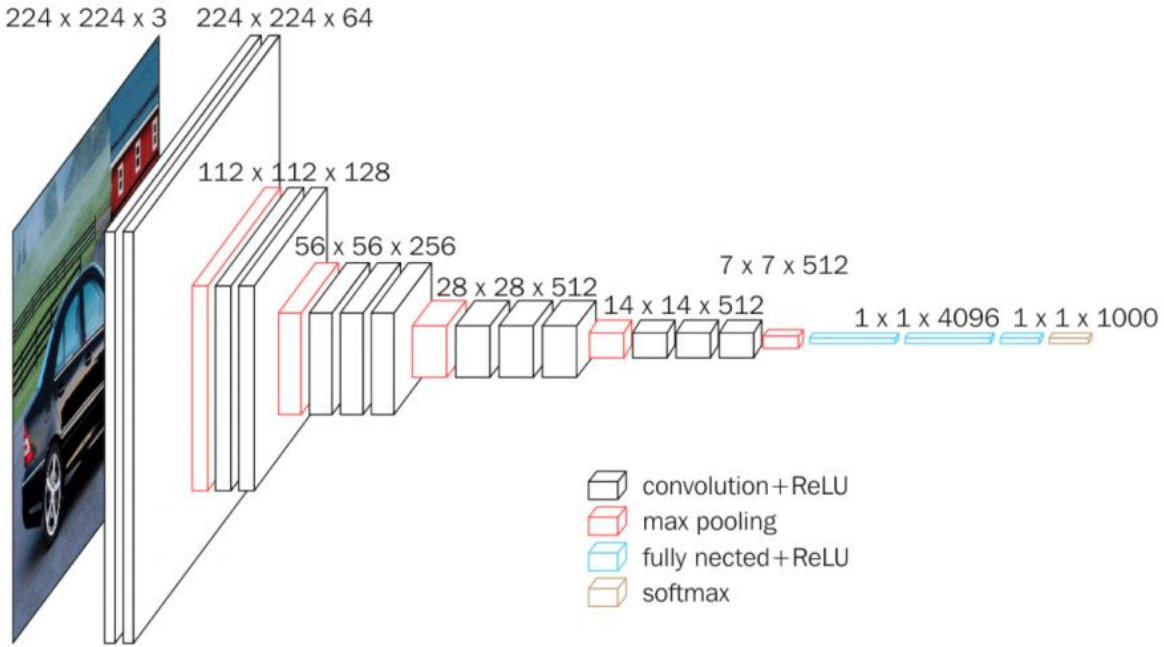
AlexNet (2012)



https://neurohive.io/wp-content/uploads/2018/10/AlexNet-1.png

CNN architectures

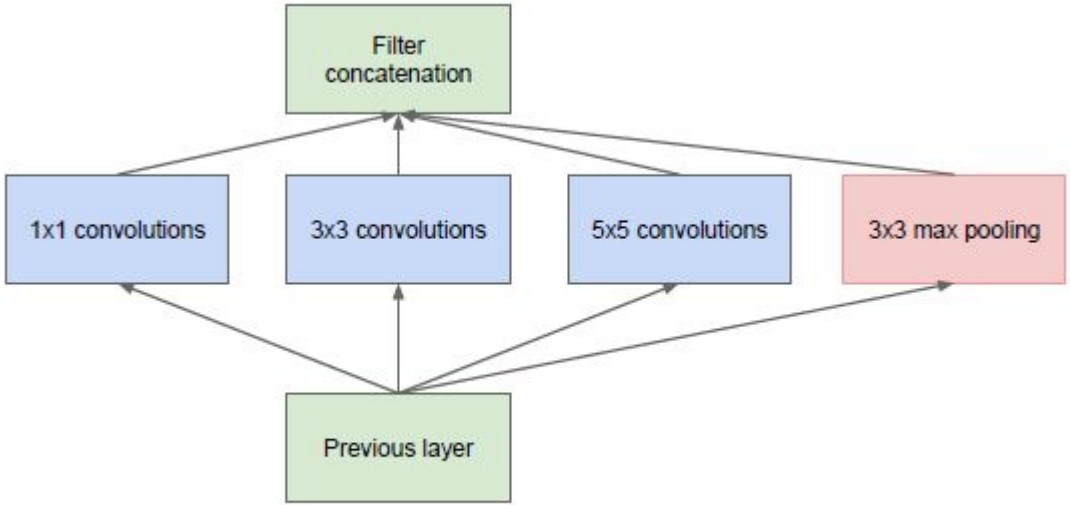
VGG-16 (2012)



https://neurohive.io/wp-content/uploads/2018/11/vgg16-1-e1542731207177.png

CNN architectures

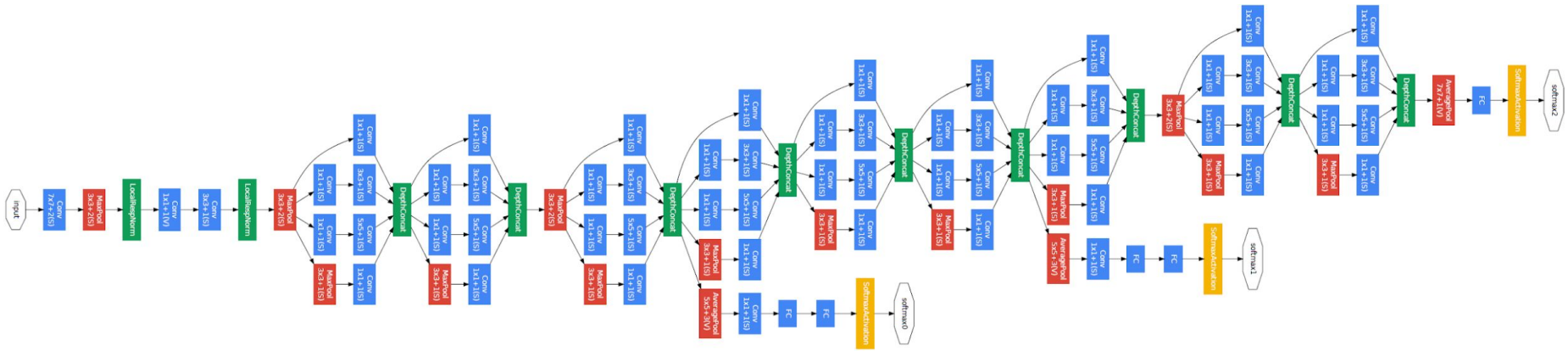
Inception V1 (2014)



https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvlc-2014-image-classification-c2b3565a64e7

CNN architectures

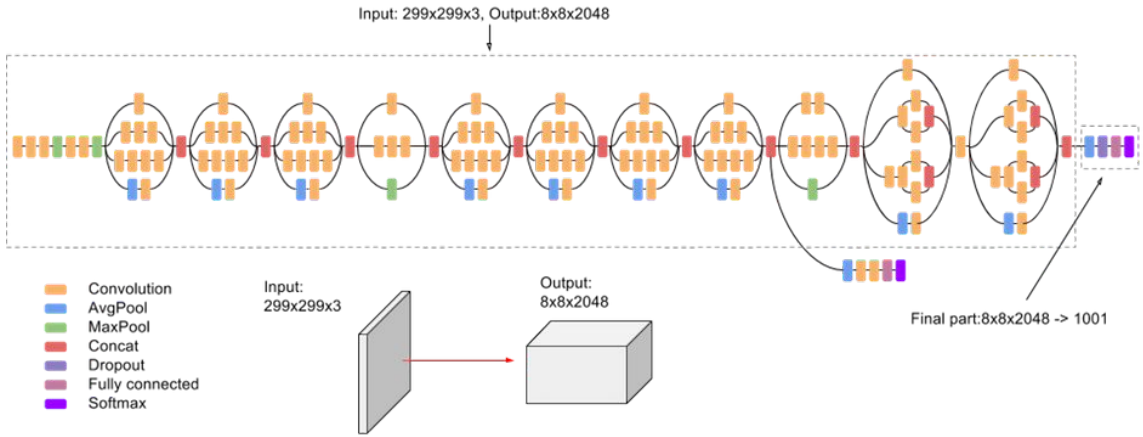
Inception V1 (2014) - GoogLeNet



https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvlc-2014-image-classification-c2b3565a64e7

CNN architectures

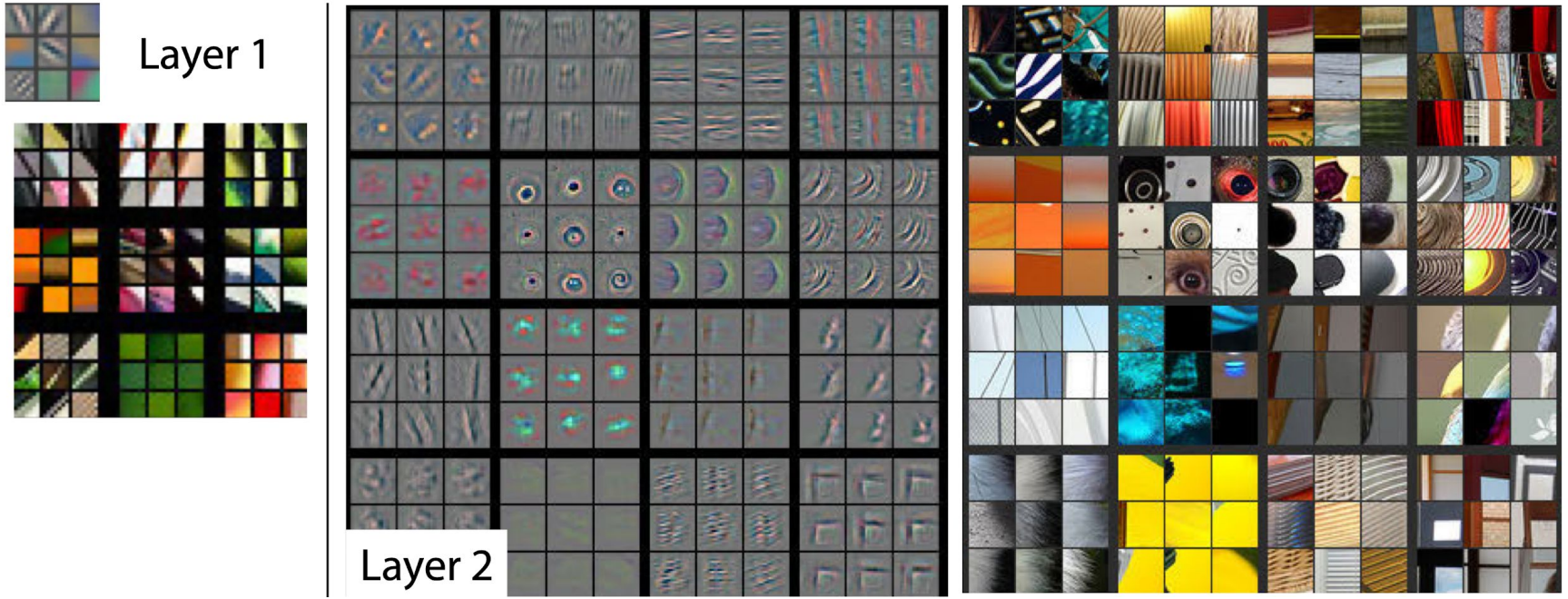
Inception V3



https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvlc-2014-image-classification-c2b3565a64e7

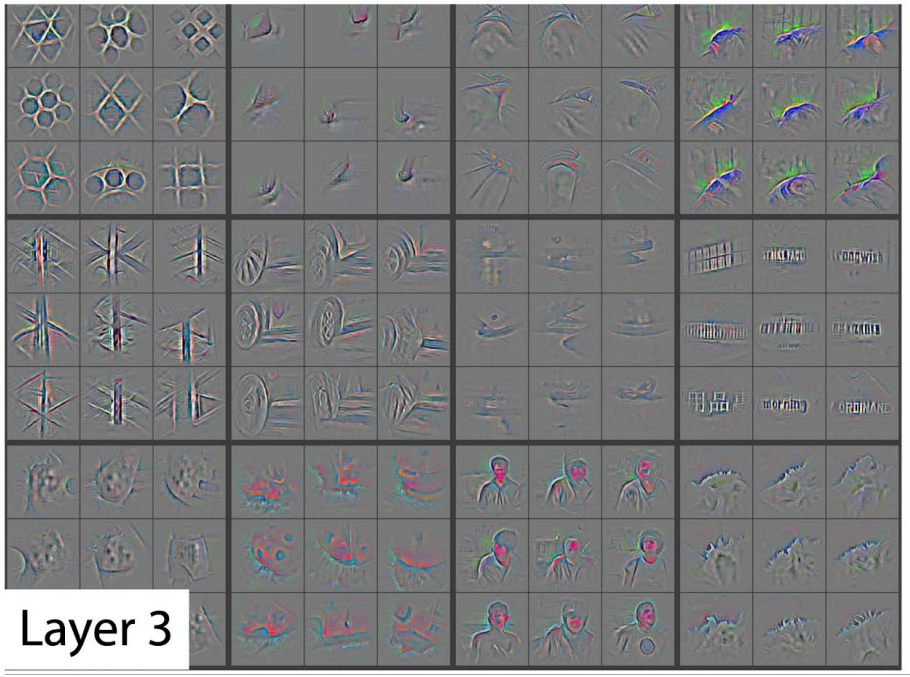
How does a filter look like?

Deconvolution



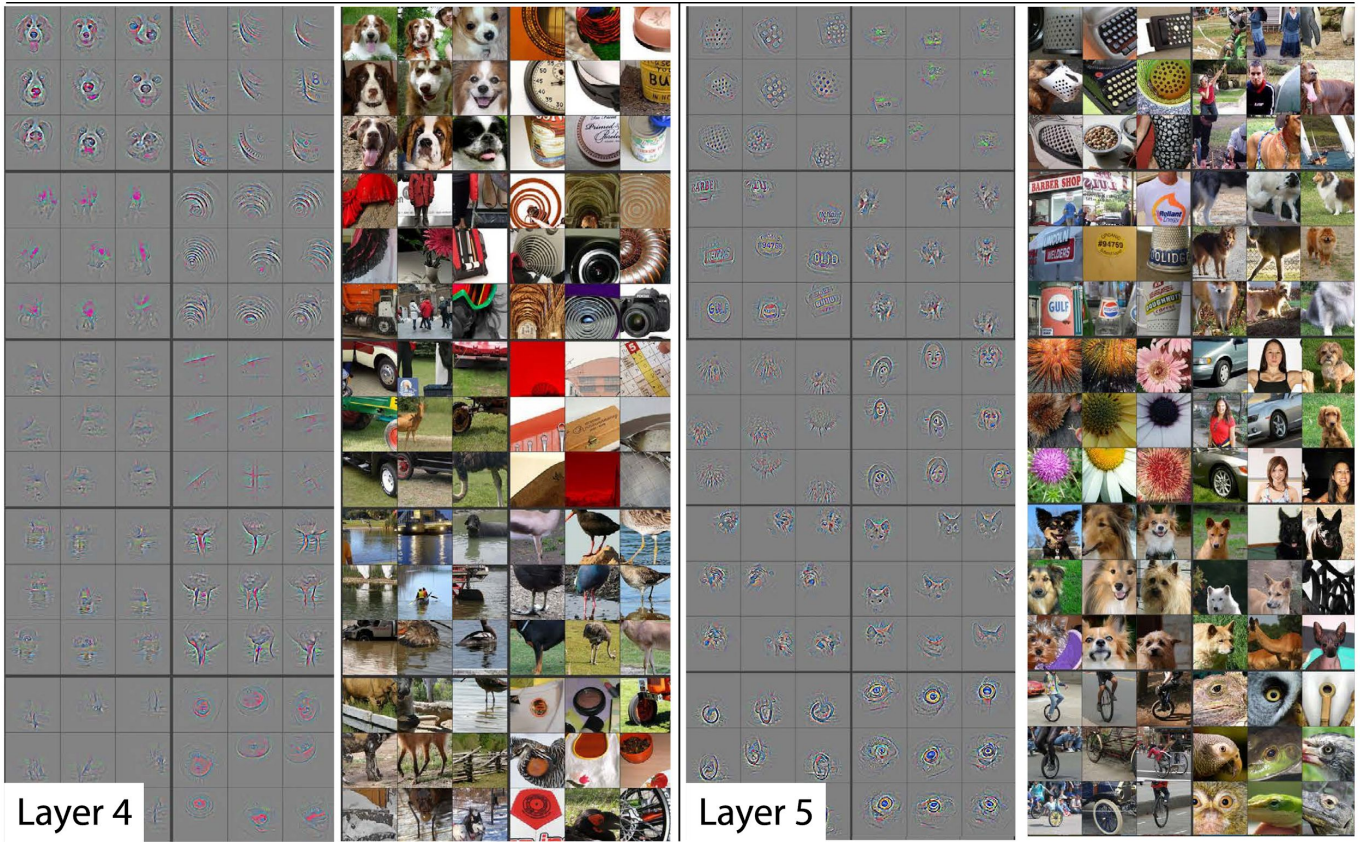
Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

Deconvolution



Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

Deconvolution



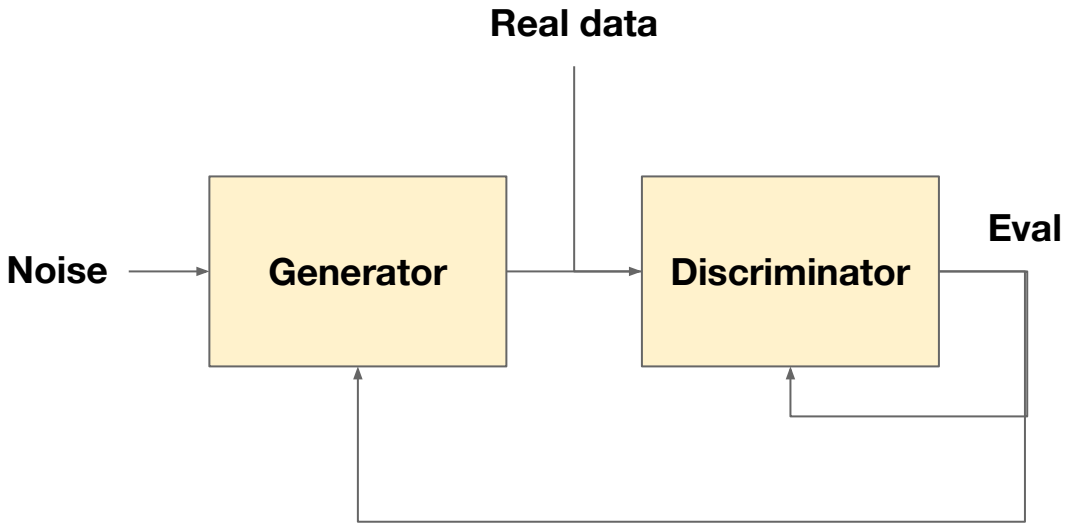
Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

GANs
Generative adversarial Networks

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. arXiv preprint arXiv:1406.2661.
LINK

GANs

Two CNNs playing a game



https://www.ibm.com/ibm/history/ibm100/us/en/icons/ibm700series/impacts/

GANs

Fake images generation

StyleGAN

Building a convolutional neural network with Keras

Natural Language Processing with Deep Learning

NLP - Natural Language Processing

NLP - Natural Language Processing

Evolution of NLP

- Symbolic NLP (ontologies, rules based, ...)
- Statistical NLP (TF-IDF)
- Neural models (present: Word2Vec, Tranformers, ...)

NLP - Natural Language Processing

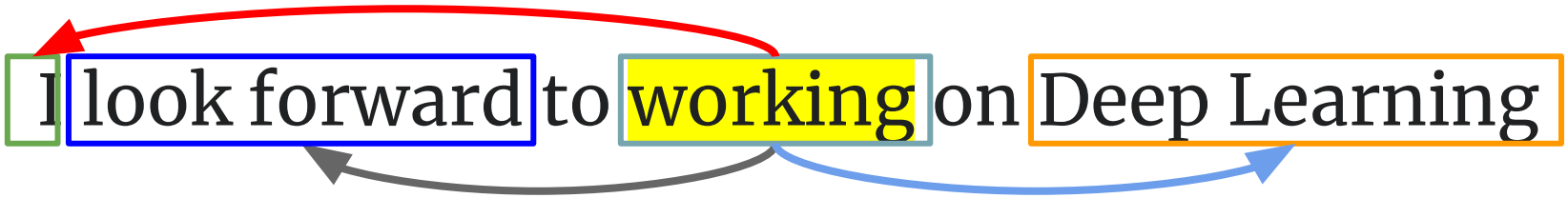
- Statistical NLP
  - TF-IDF: term frequency–inverse document frequency

How important is a word?

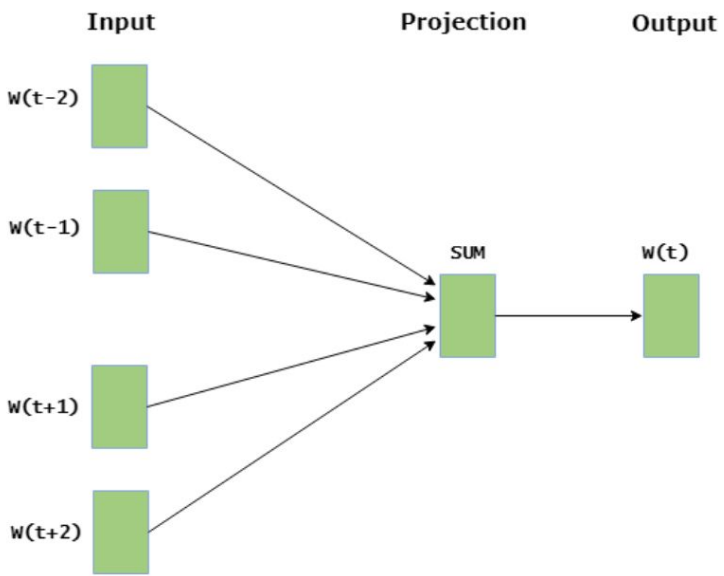
I look forward to working on Deep Learning

NLP - Natural Language Processing

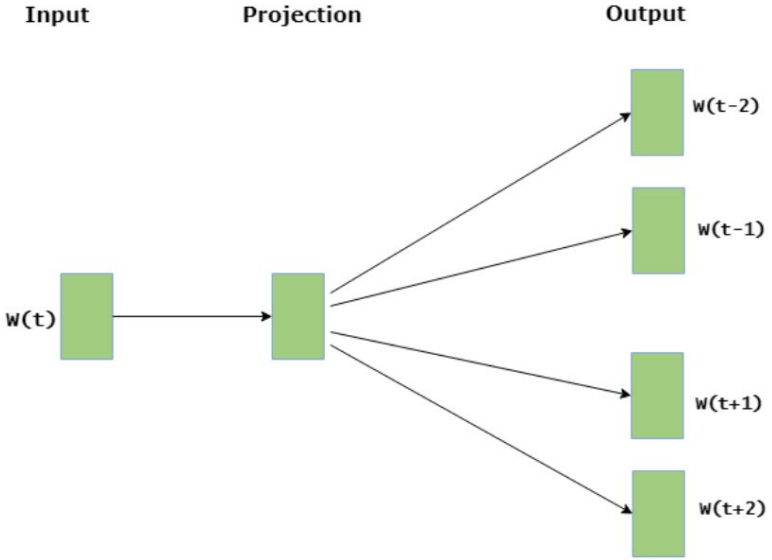
- Neural NLP: the importance of the context



- Word2Vec: neural networks to build word embedding models



Continuous Bag of Words (CBOW)



Skip Gram

https://www.geeksforgeeks.org/python-word-embedding-using-word2vec/

NLP – Natural Language Processing

- The sky is
  - **blue**
  - is covered with dark clouds.
  - is red
  - is falling

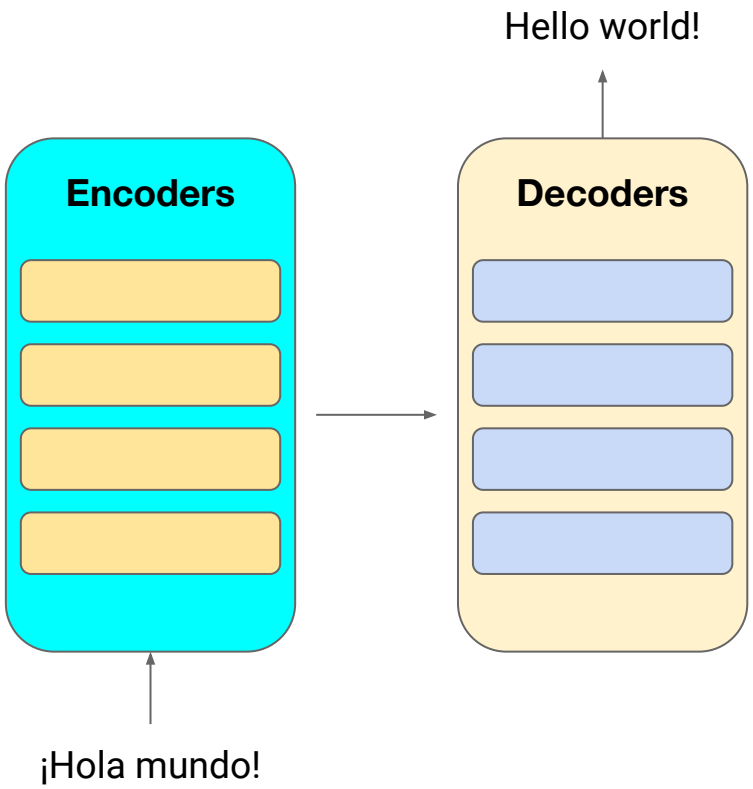
Transformers

Attention Is All You Need

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *arXiv preprint arXiv:1706.03762*.

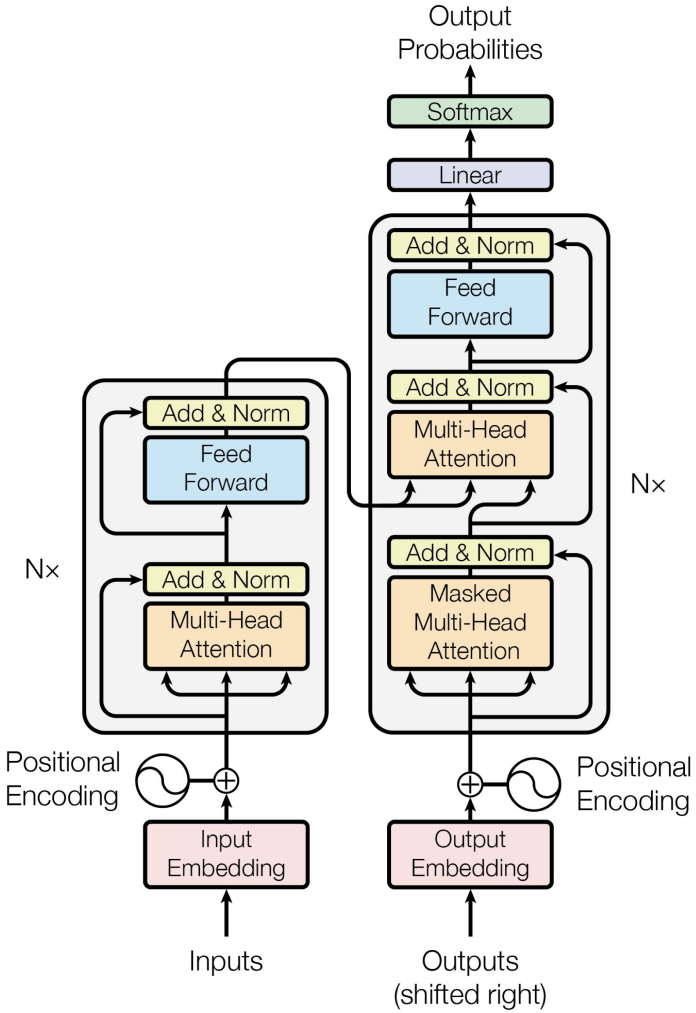
The Transformer architecture

Architectures for Sequence to Sequence transforming



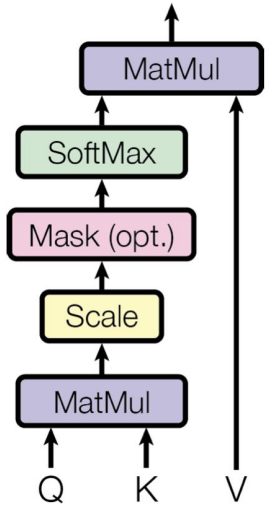
The Transformer architecture

Architectures for Sequence to Sequence transforming



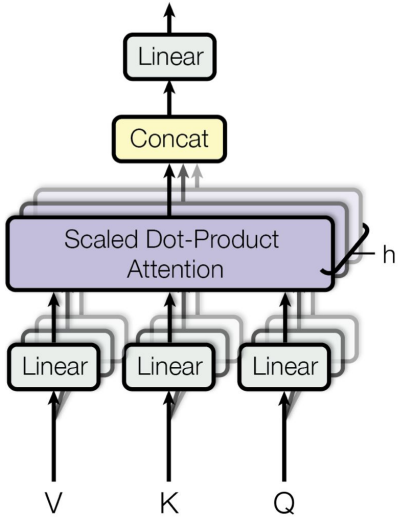
The Transformer architecture

Scaled Dot-Product Attention



Mapping between a query and key-value pairs

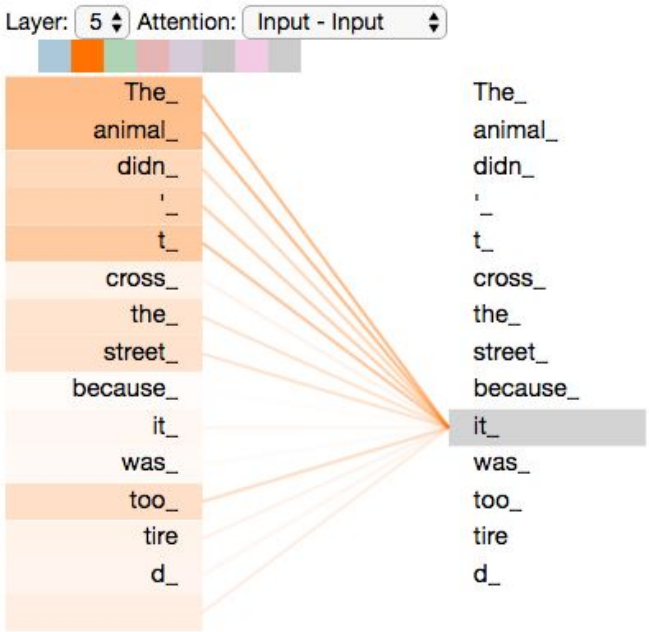
Multi-head attention



Several projections of the queries, keys and values

The Transformer architecture

Attention



The Transformer architecture

Architectures for Sequence to Sequence transforming

Tensor2Tensor notebook

Transformers

Transformers

Tensorflow & PyTorch

https://github.com/huggingface/transformers

Sentence Tranformers

PyTorch

https://github.com/UKPLab/sentence-transformers

Arquitectures

Some models available at HuggingFace library

BERT (Google)

DistilBERT (HuggingFace)

GPT, GPT2 (OpenAI)

XLM (Facebook)

XLM-RoBERTa (Microsoft)

...

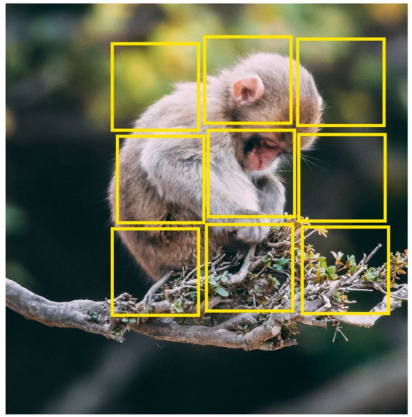
Text generation

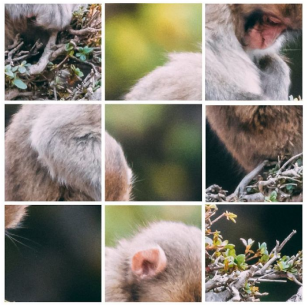
https://transformer.huggingface.co/doc/gpt2-large

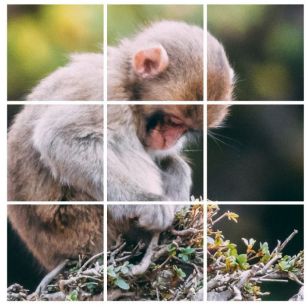
Trending topics in deep learning

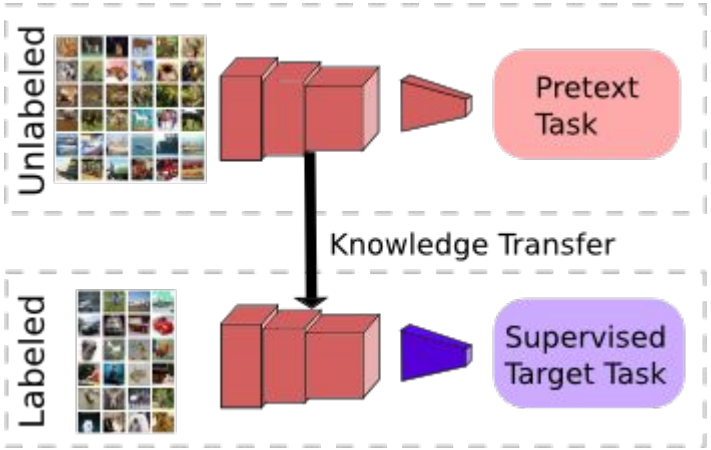
Self-supervised learning

The main idea of Self-Supervised Learning is to generate the labels from unlabeled data, according to the structure or characteristics of the data itself, and then train on this unsupervised data in a supervised manner.





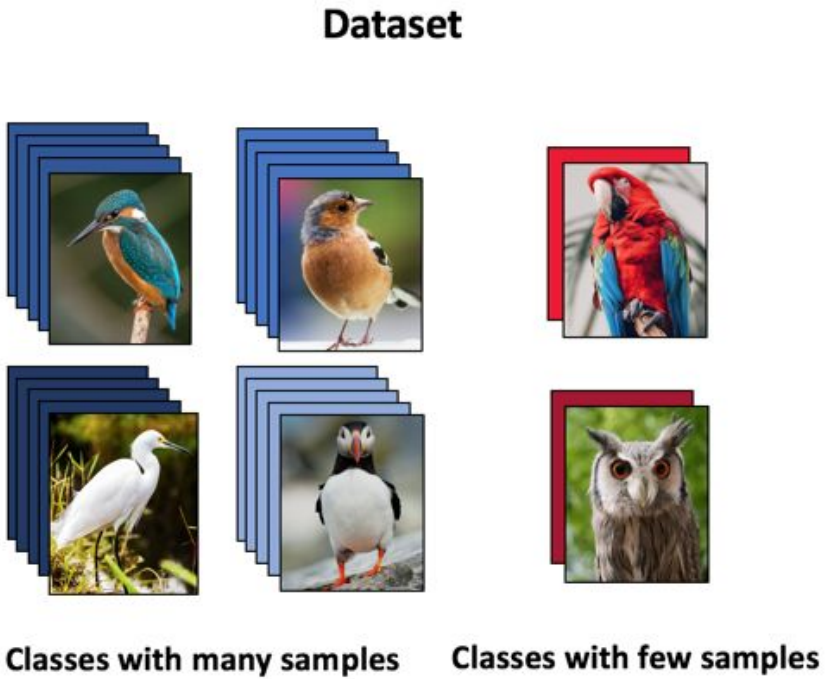


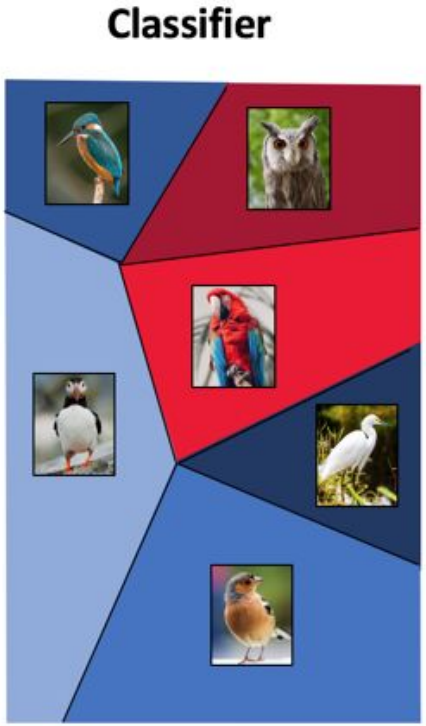


https://rl.uni-freiburg.de/img/teaching/selfsup-seminar

https://perfectial.com/wp-content/uploads/2020/03/SSL-02-scaled.jpg

Few-shot learning

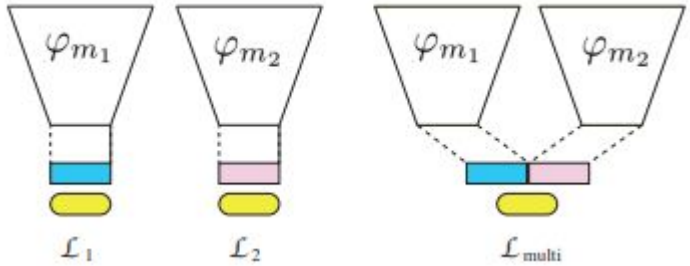




https://medium.com/sap-machine-learning-research/deep-few-shot-learning-a1caa289f18

Data fusion and multimodal models

Recent research shows that end-to-end training of multi-modal networks with fused inputs from different modalities (e.g., images and tabular data at once) do not always outperform uni-modal networks on the same tasks



https://arxiv.org/pdf/1905.12681.pdf

TEXT PROMPT

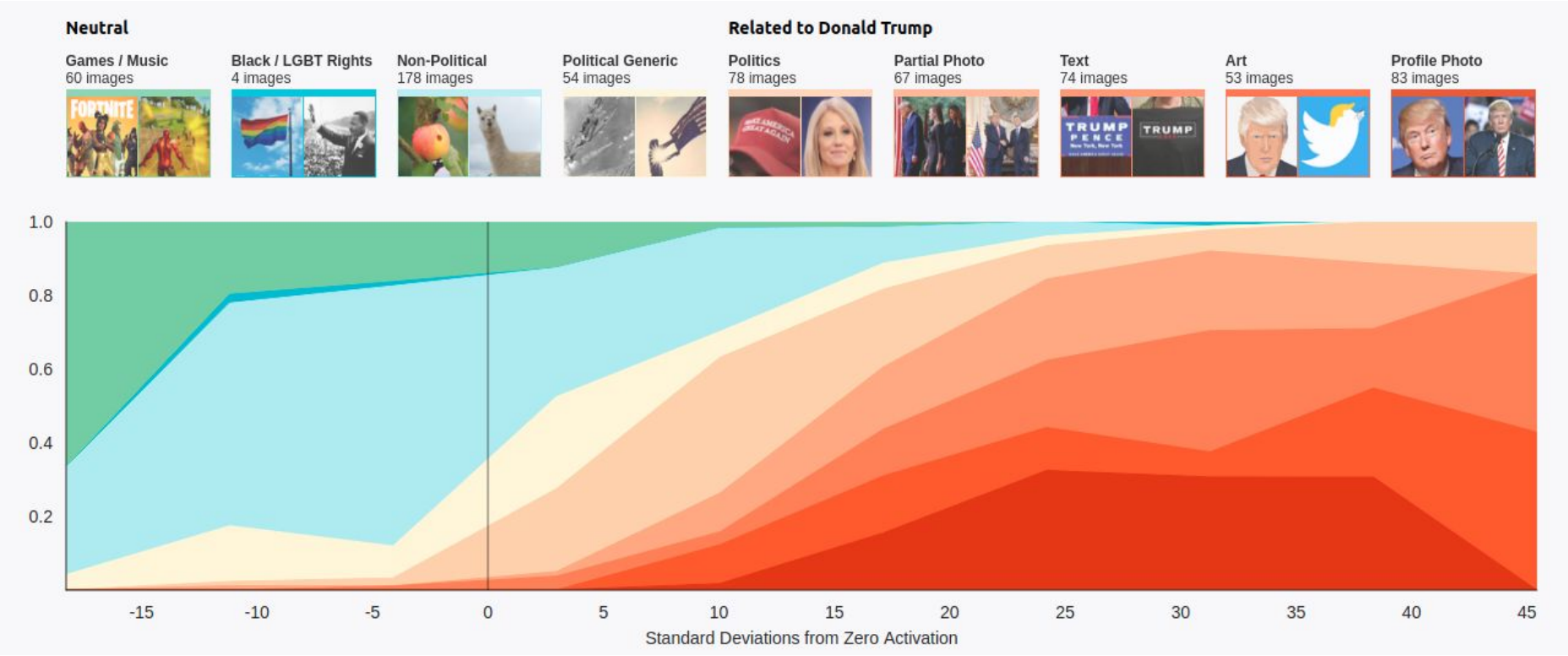
an armchair in the shape of an avocado [...]

AI-GENERATED IMAGES

Edit prompt or view more images ↓

https://openai.com/blog/dall-e/

CLIP & DALL-E, by OpenAI: Connecting text and images. CLIP is a neural network that efficiently learns visual concepts from natural language supervision. That can be used to create images from text captions for a wide range of concepts expressible in natural language.



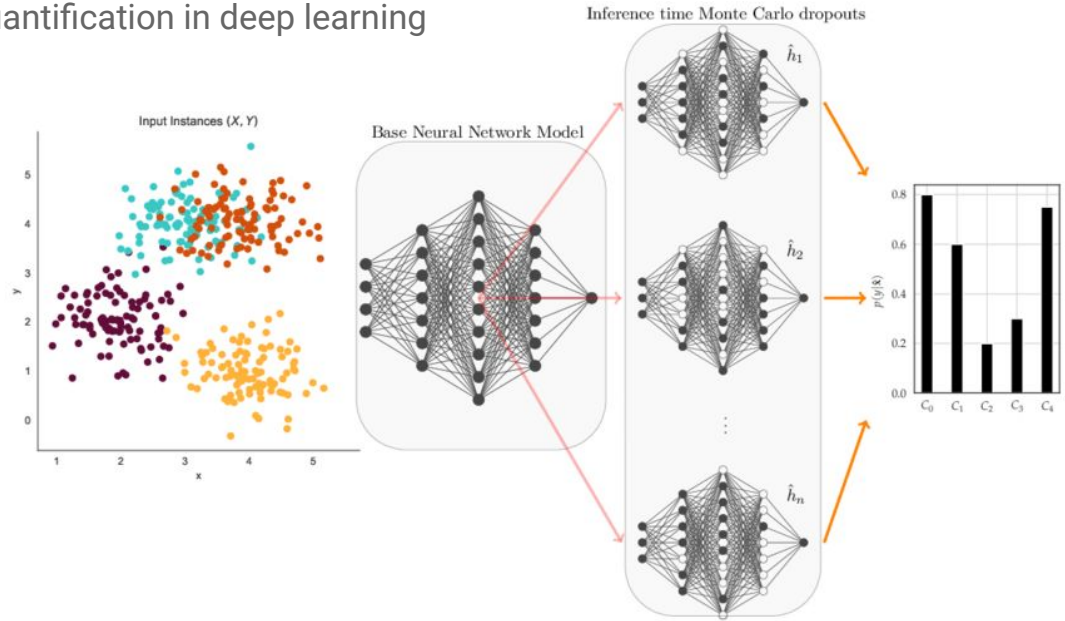
https://distill.pub/2021/multimodal-neurons/

Neurons in CLIP respond to the same concept whether presented literally, symbolically, or conceptually.

Uncertainty quantification

Deterministic deep neural networks are known to be overconfident in their predictions on Out of Distribution (OoD) data. Popular ways to approach uncertainty quantification in deep learning

- Monte Carlo Dropout
- Deep ensembles

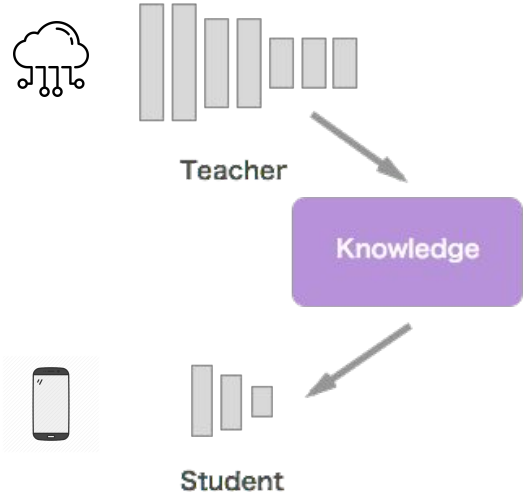


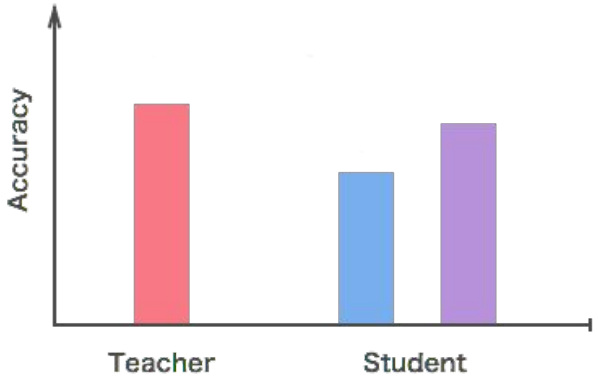
https://www.researchgate.net/publication/346953063_Closeness_and_Uncertainty_Aware_Adversarial_Examples_Detection_in_Adversarial_Machine_Learning/figures?lo=1&utm_source=google&utm_medium=organic

Knowledge distillation

Train a compact model (student) to reproduce the behaviour of a larger one (teacher), achieving smaller, faster, cheaper and lighter versions of it.

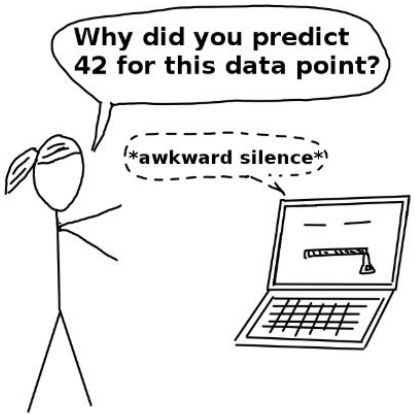
- Allows for deployment in mobile and edge devices

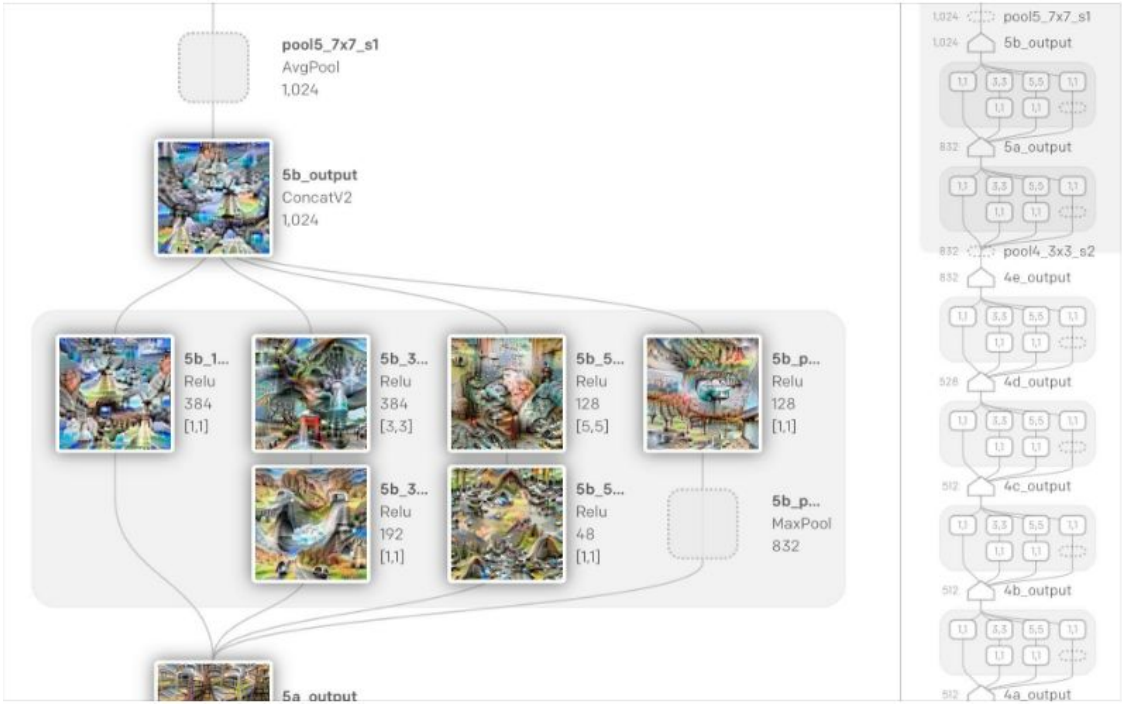




https://medium.com/neuralmachine/knowledge-distillation-dc241d7c2322

Model interpretability





https://towardsdatascience.com/interpretable-machine-learning-1dec0f2f3e6b

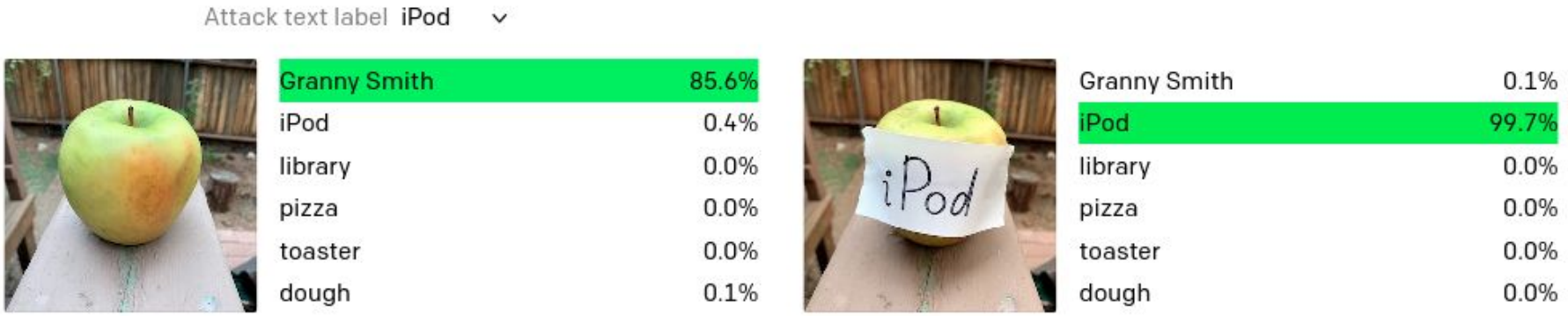
https://openai.com/blog/microscope/

Risks & Ethics of deep learning systems

¿Is this a stop sign?

- Adversarial examples: Researchers found that popular neural networks for image recognition were classifying this sign as a speed limit sign

https://threatpost.com/machine-learning-dark-side/142616/



https://openai.com/blog/multimodal-neurons/

Typographic attacks: By exploiting the model's ability to read text robustly, we find that even photographs of hand-written text can often fool the model. It requires no more technology than pen and paper

Deep fakes


FACEBOOK
ZUCKERBERG:
WE'RE INCREASING TRANSPARENCY ON ADS
ANNOUNCES NEW MEASURES TO "PROTECT ELECTIONS"
CBSN

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

Joy Buolamwini & Timnit Gebru

https://cedrickchee.gitbook.io/knowledge/courses/fast.ai/deep-learning-part-1-practical-deep-learning-for-coders/2019-edition/lesson-6-foundations-convolutional-neural-nets

Gender classification: Lighter male skinned people on IBM's main computer vision system get 99.7% accuracy, while darker darker skinned women get much less accuracy

¿Where does this bias come from?

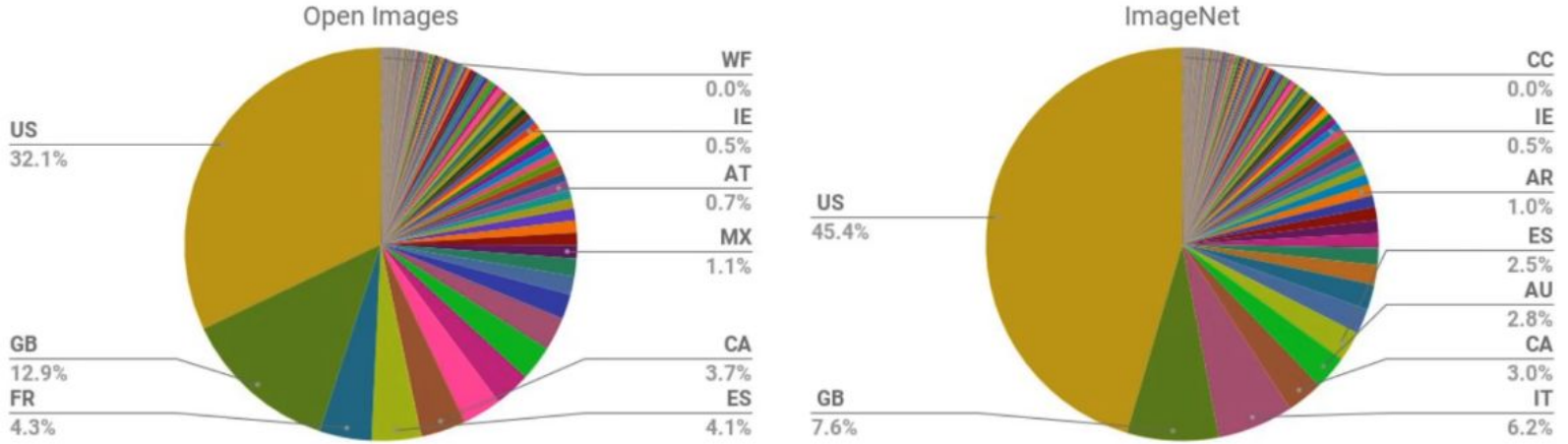
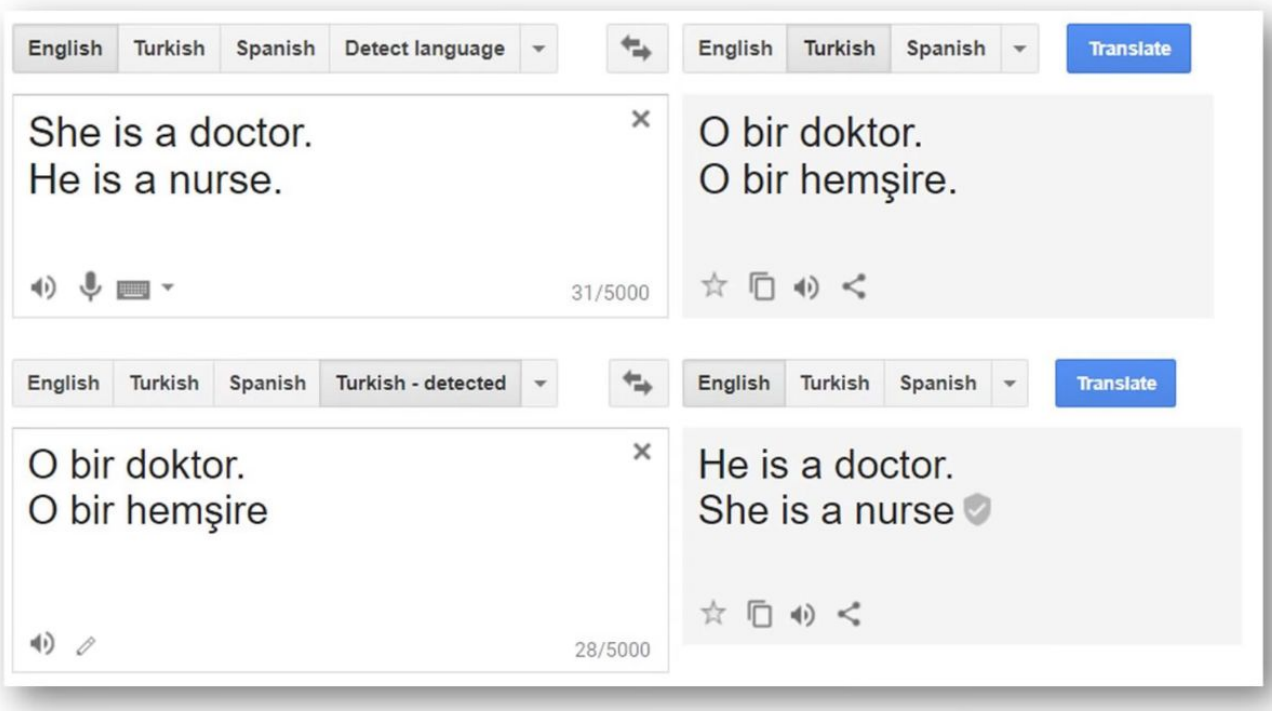


Figure 1: Fraction of Open Images and ImageNet images from each country. In both data sets, top represented locations include the US and Great Britain. Countries are represented by their two-letter ISO country codes.

https://cedrickchee.gitbook.io/knowledge/courses/fast.ai/deep-learning-part-1-practical-deep-learning-for-coders/2019-edition/lesson-6-foundations-convolutional-neural-nets

Bias in NLP



https://cedrickchee.gitbook.io/knowledge/courses/fast.ai/deep-learning-part-1-practical-deep-learning-for-coders/2019-edition/lesson-6-foundations-convolutional-neural-nets

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

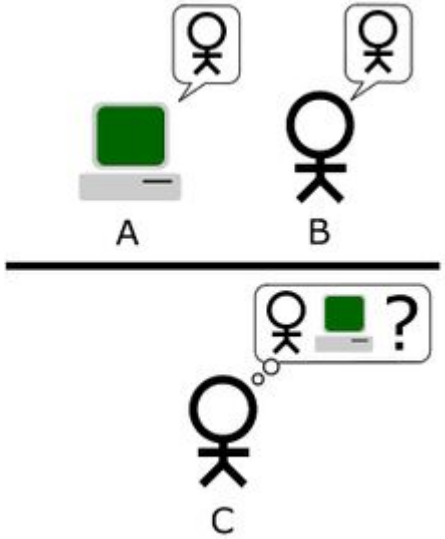
https://www.fast.ai/page7/

Results of the COMPAS system, used to assist judges in deciding the risk of re-offenders.

Limits of deep learning

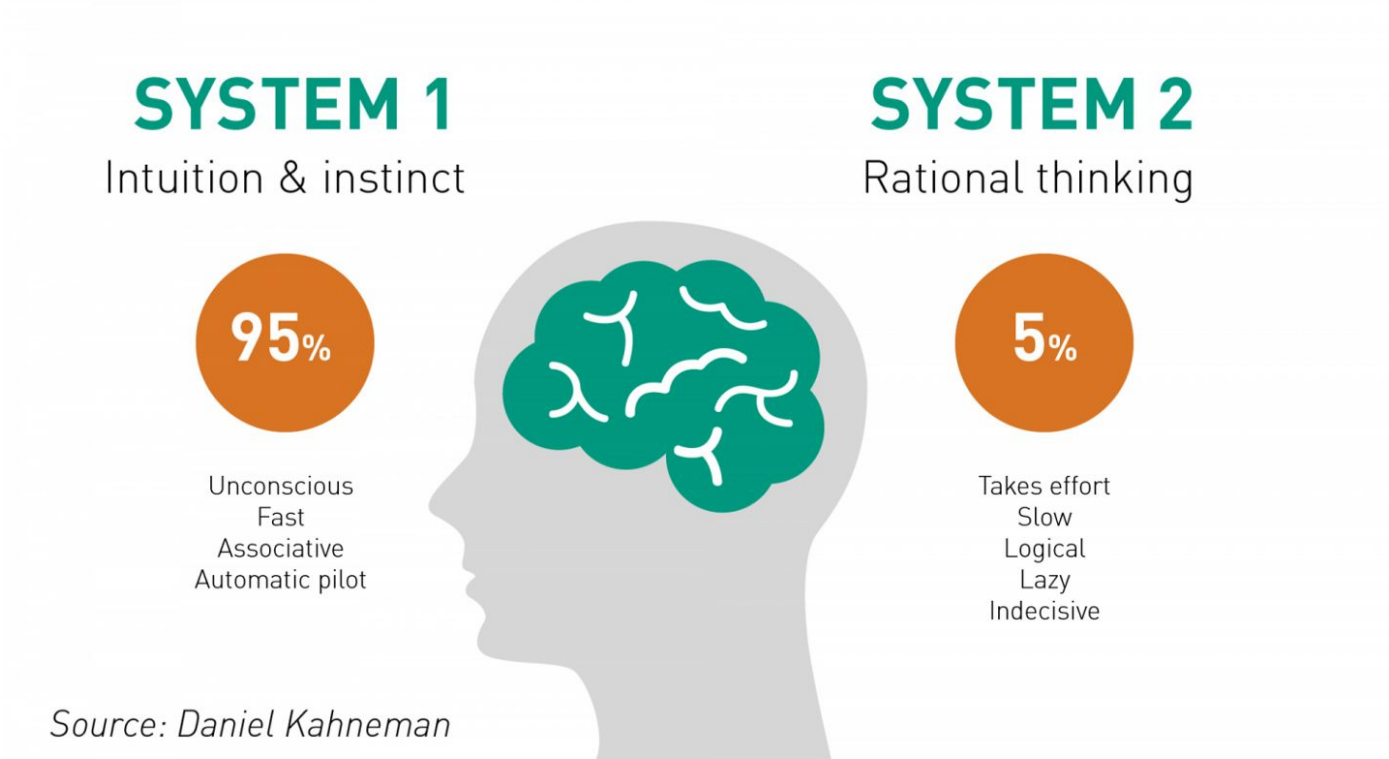
The turing test (1950)

- A human assesses natural language (textual) conversations between a human and a machine designed to generate human-like responses.
- In the event that the tester could not distinguish between the human and the machine accurately, the machine would have passed the test.

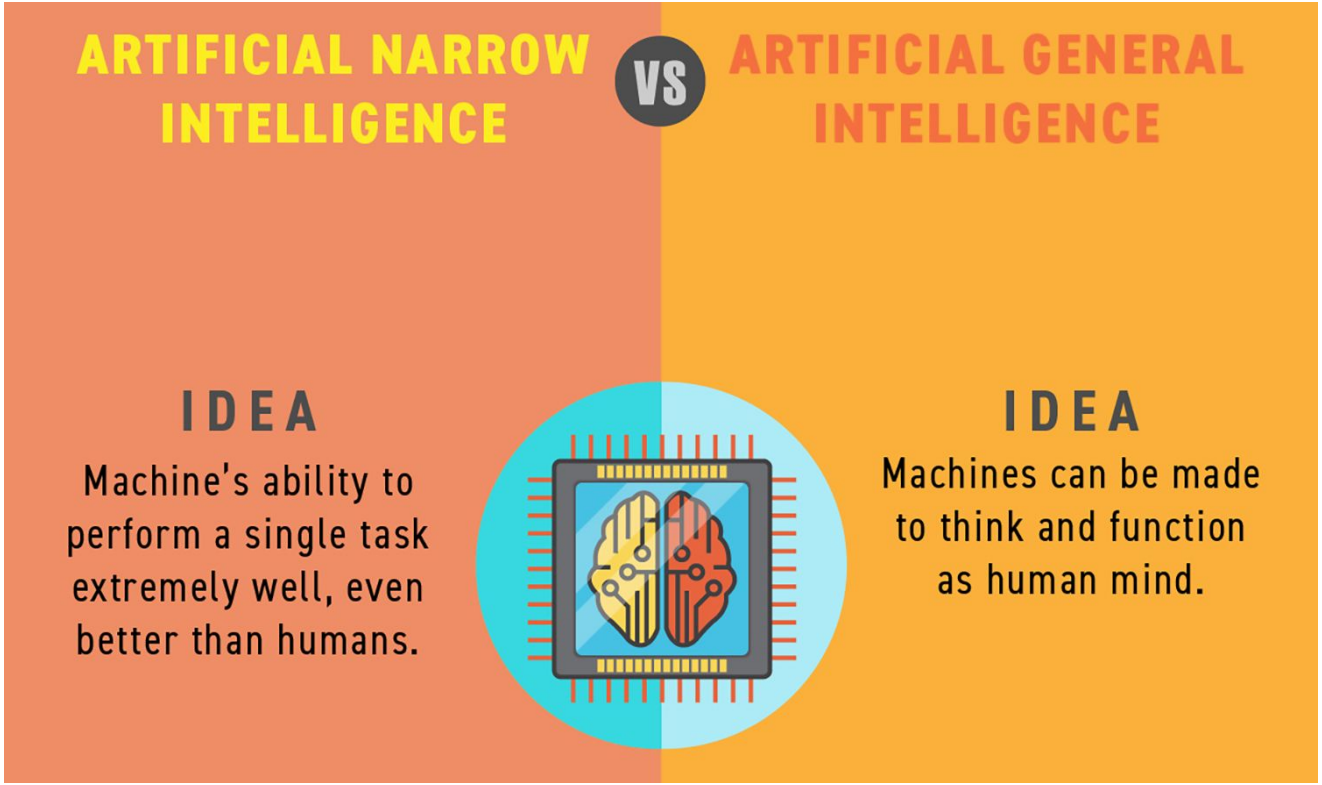


https://naukas.com/2019/08/27/turing-y-la-inteligencia-de-lo-no-computable/turing7/

Thinking systems (Daniel Kahneman)



Daniel Kahneman: “Deep learning cannot go beyond system 1”



https://analyticsindiamag.com/artificial-narrow-vs-artificial-general/

https://openai.com/blog/microsoft/

Microsoft (July 2019): OpenAI and Microsoft sign an exclusive collaboration agreement to create new supercomputing technologies for Artificial General Intelligence

GPT (2019-2020)

New AI fake text generator may be too dangerous to release, say creators

The Elon Musk-backed nonprofit company OpenAI declines to release research publicly for fear of misuse

https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction

SYSTEM PROMPT (HUMAN-WRITTEN)	*In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.*
MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)	The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science. Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved. Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow. Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez. Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns. While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."

https://blog.floydhub.com/gpt2/

You can try GPT-2 in https://transformer.huggingface.co/doc/gpt2-large

Free ML/DL resources

Free courses and books

- Free online courses
  - Practical deep learning for coders (fast.ai)
  - Stanford CS229: Machine Learning
  - Stanford CS231n: Convolutional Neural Networks for Visual Recognition
  - Fundamentals of deep learning

- Free books
  - Deep learning, by Ian GoodFellow, Yoshua Bengio and Aaron Courville
  - Fastbook: Deep learning for coders with fastai & Pytorch: Applications of AI without a PhD
  - Approaching (Almost) Any Machine Learning Problem