Machine learning Introduction

Small overview of Machine learning

Overview of machine learning history
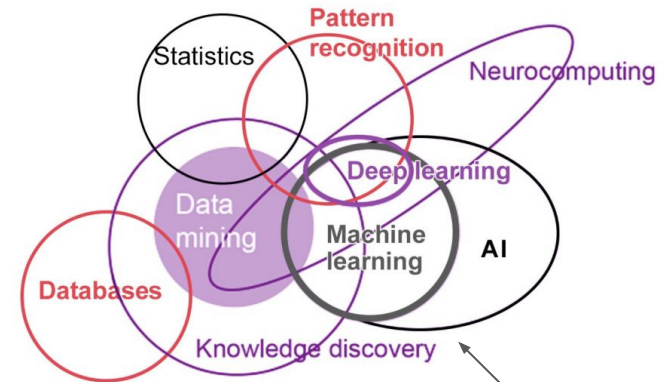

From the blog:
https://jaylatta.net/history-of-ai-from-winter-to-winter/
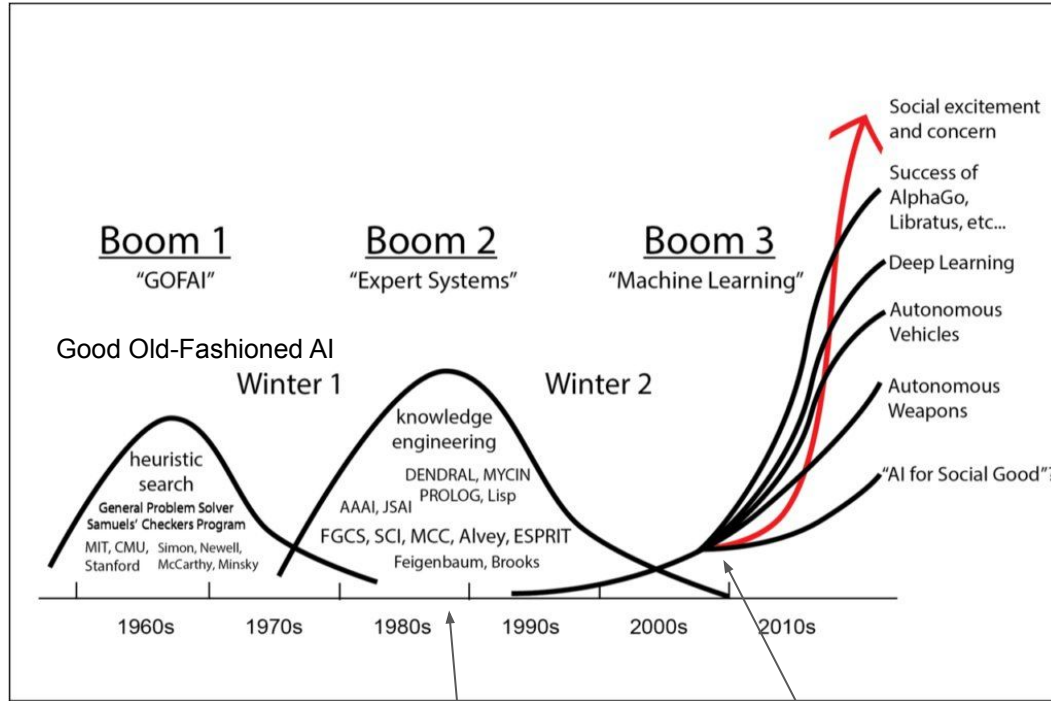
Development of computers and software: new possibilities!

Boom from the combination of: 1) computer power, 2) data and 3) ML methods

Quite wrong but gives a rough idea.
What is correct?

History of Science: new instruments -> new breakthroughs

Early days

- 1805: linear regression or method of least squares (Legendre)

Data points (x, y)

Line $y = ax + b$
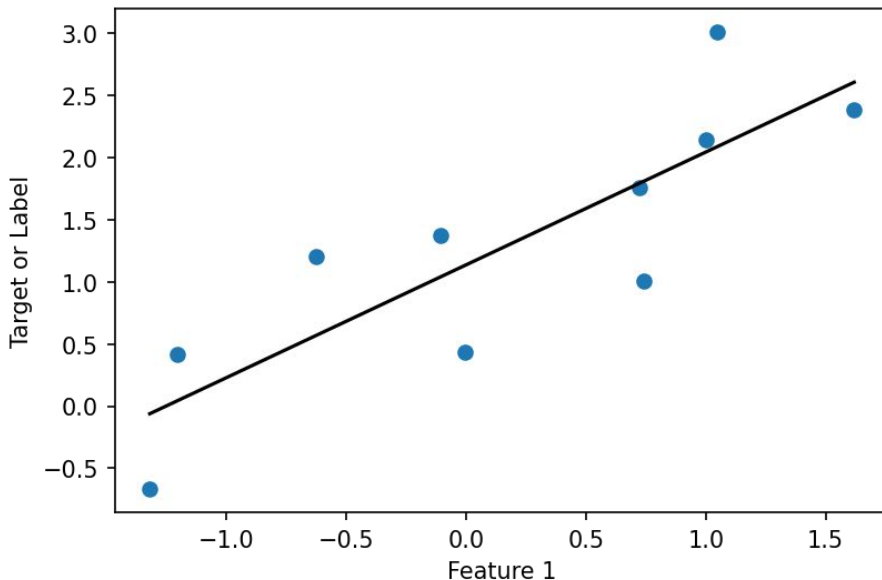
From **data points**: find the line which **minimize** the sum of squared errors

The value for new points can be **predicted**.

In bold: central concepts in ML

Early days

Note: Machine learning development goes along with computer development

- 1805: linear regression or method of least squares (Legendre)

- 1943: McCulloch and Pitts: first mathematical model of neural nets
- 1950: Alan Turing, "the Turing test" (the imitation game), a test to evaluate if a machine is "intelligent"
- 1951: k-nearest neighbors
- 1957: Rosenblatt, the perceptron
- 50-60s: Bayesian methods
- 60-70s: clustering
- 1969: Minsky and Papert, end of the perceptron and beginning of the first AI winter, the XOR problem
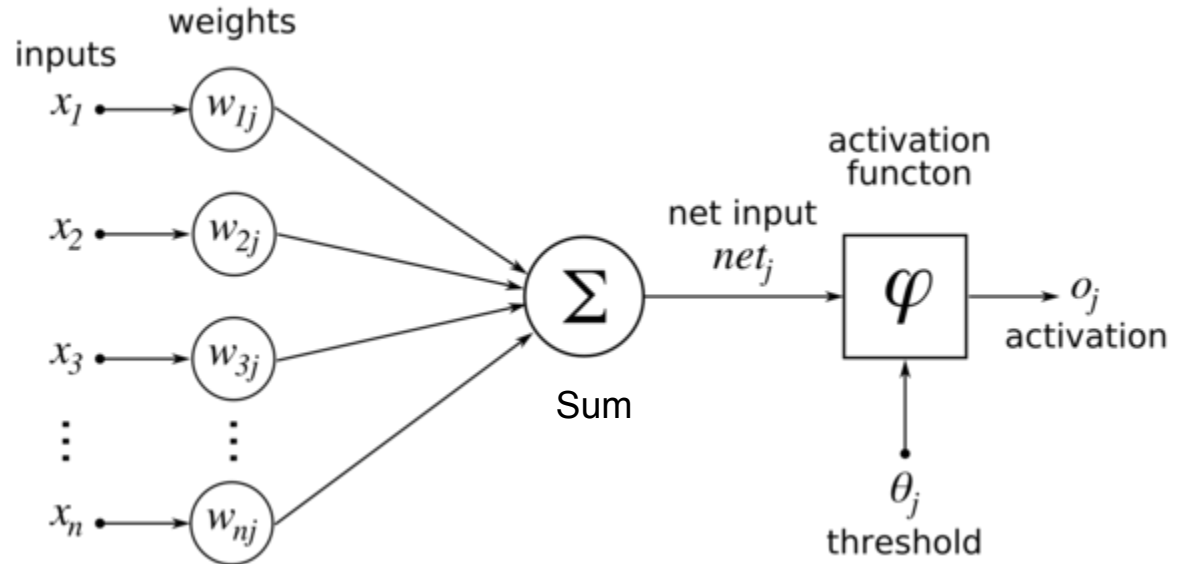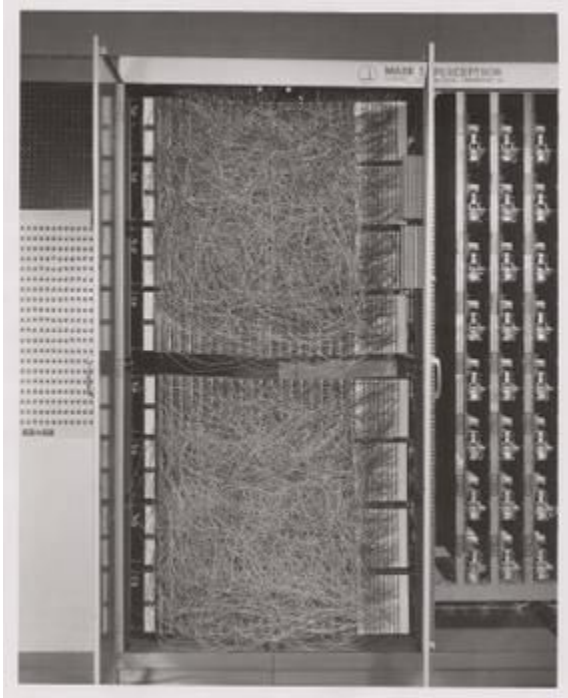
Early days, the perceptron

weights

inputs

x_1 → w_{1j}

x_2 → w_{2j}

x_3 → w_{3j}

⋮   ⋮

x_n → w_{nj}

Σ

Sum

net input net_j

activation functon

φ

o_j activation

θ_j threshold

Figure 4.8 Illustration of the Mark 1 perceptron hardware. The photograph on the left shows how the inputs were obtained using a simple camera system in which an input scene, in this case a printed character, was illuminated by powerful lights, and an image focussed onto a 20×20 array of cadmium sulphide photocells, giving a primitive 400 pixel image. The perceptron also had a patch board, shown in the middle photograph, which allowed different configurations of input features to be tried. Often these were wired up at random to demonstrate the ability of the perceptron to learn without the need for precise wiring, in contrast to a modern digital computer. The photograph on the right shows one of the racks of adaptive weights. Each weight was implemented using a rotary variable resistor, also called a potentiometer, driven by an electric motor thereby allowing the value of the weight to be adjusted automatically by the learning algorithm.
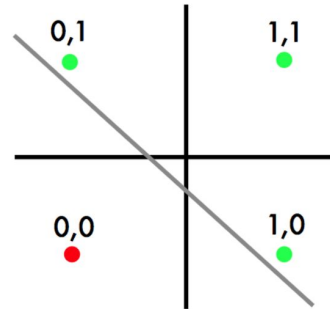
From *Pattern recognition and machine learning*, C. Bishop, freely available online

The Perceptron and classification
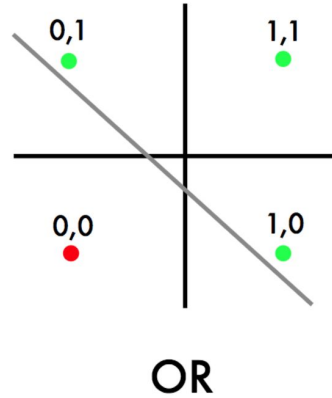
The Perceptron create a separation line (solve a linear classification problem)

the XOR problem
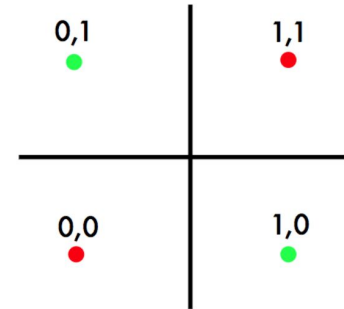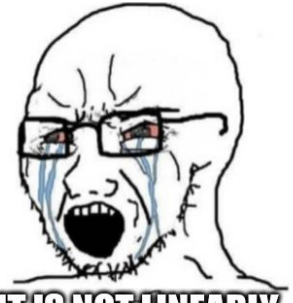

The Perceptron create a separation line (solve a linear classification problem)

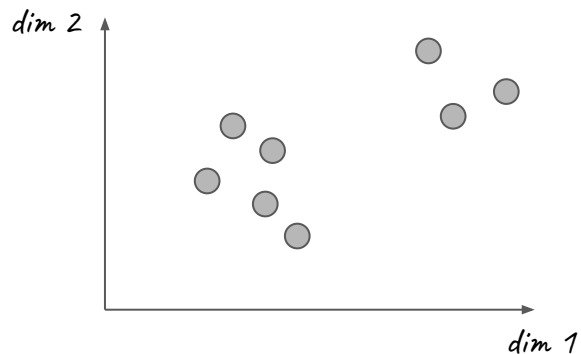# First AI winter 1974-1980

Results below expectations

Research fund cutted

but developments continued at a slower pace

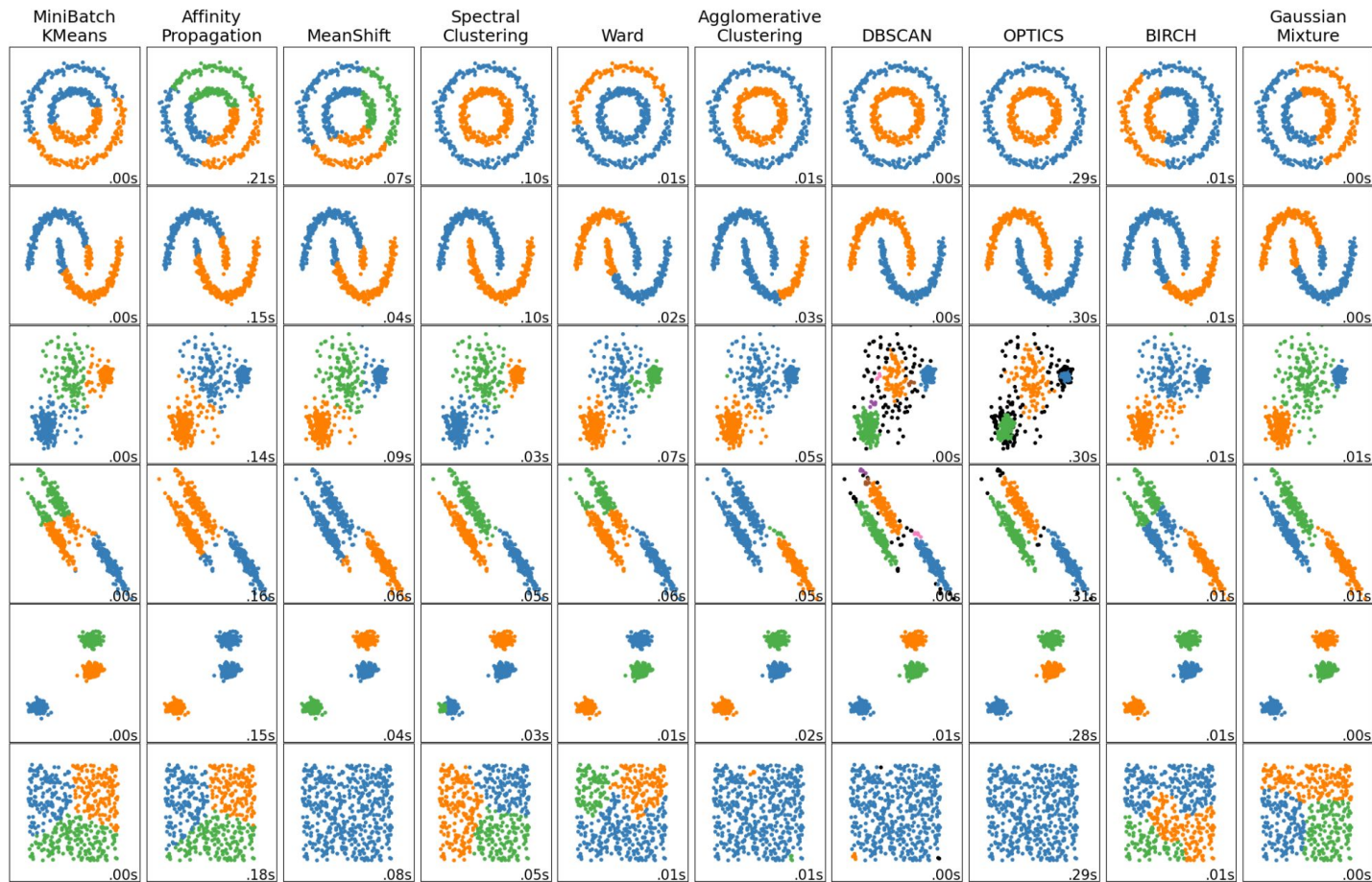1979: Neocognitron (for vision, inspired convolutional neural nets)

Clustering

Principle, group close data points together

- most methods do not use any label (unsupervised)
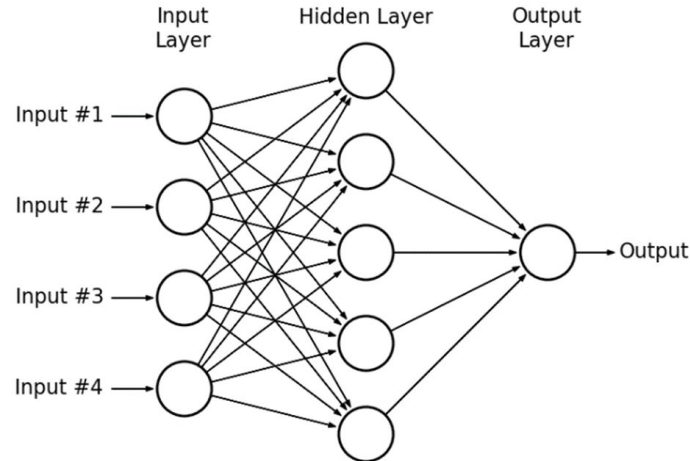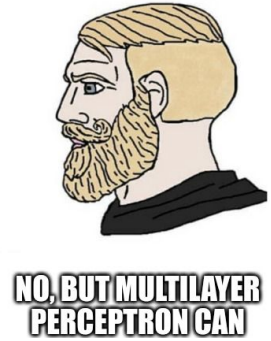- some may use a partial labelling (semi-supervised)

Clustering

From https://scikit-learn.org/stable/modules/clustering.html

Machine learning boom 80-87

- More powerful computers, programming languages
- Decision trees in their modern form
- Rediscovering backpropagation
- Developing several layers of neurons, multilayer perceptron

PERCEPTRON CAN'T LEARN XOR FUNCTION

NO, BUT MULTILAYER PERCEPTRON CAN

Input Layer   Hidden Layer   Output Layer

Input #1
Input #2
Input #3
Input #4

Output

Second AI winter 1987- early 2000s

1988: universal approximation theorem. Multilayer NN can approximate any function f, y=f(x). (f must satisfy a few general conditions)

1991: the vanishing gradient problem (in multilayer neural networks)

BrunelloN, CC BY-SA 4.0, via Wikimedia Commons
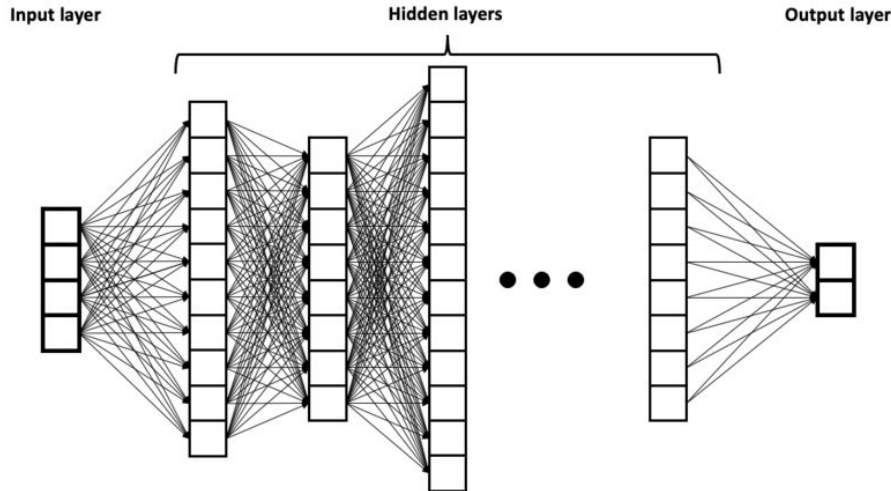
Recent developments

1995: support vector machines, random forest, gradient boosting

1997: Deep Blue beats the world champion at chess

90s: Graphical Processing Units (GPUs) appear and are progressively integrated in computers and video games

2009: publication discussing the benefits of GPUs for machine learning

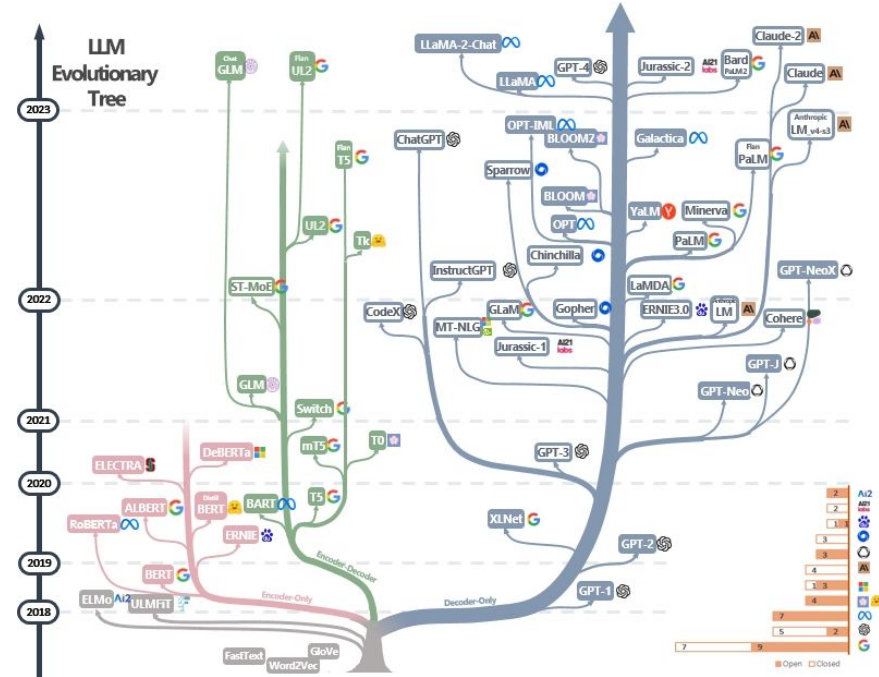2012 **the breakthrough of deep learning**: Imagenet classification and deep learning, Alexnet

2014: GANs, generative adversarial networks.

2016: AlphaGo beats best player at Go

2017: Transformer architecture (basis for BERT, chatGPT + many others)

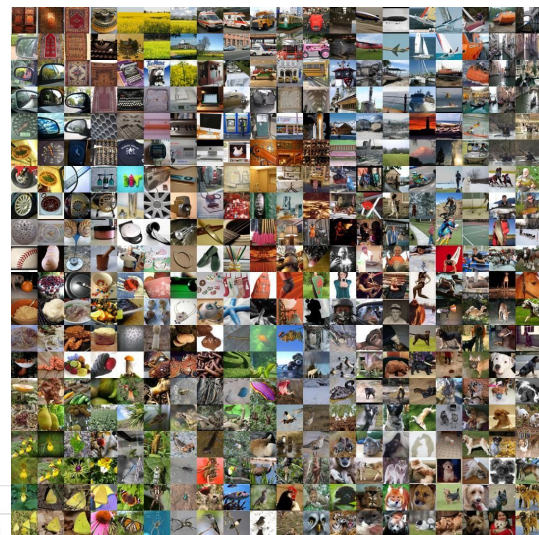2021: DALL-E, imagen, stable diffusion: AI-generated images

2022: chatGPT

From https://github.com/Mooler0410/LLMsPracticalGuide

2012 breakthrough on imagenet

- A large dataset Imagenet
- First use of GPUs for ML
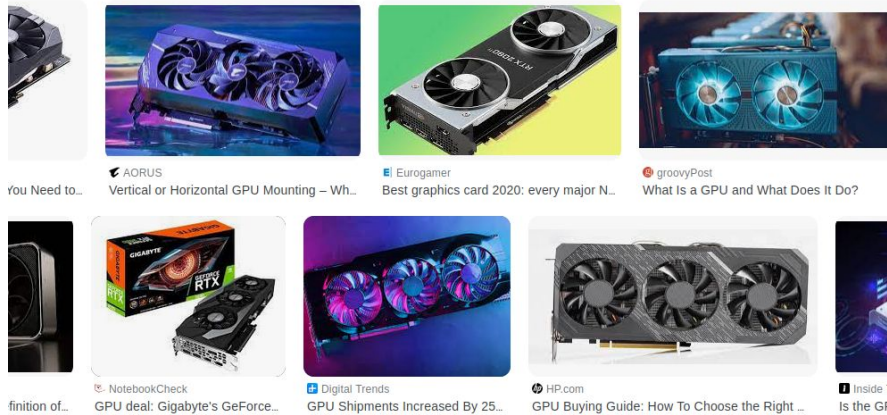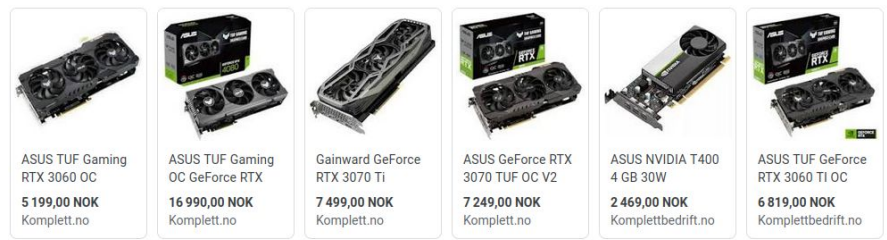- New set of machine learning concepts and tricks

Deep learning

from https://paperswithcode.com/sota/image-classification-on-imagenet
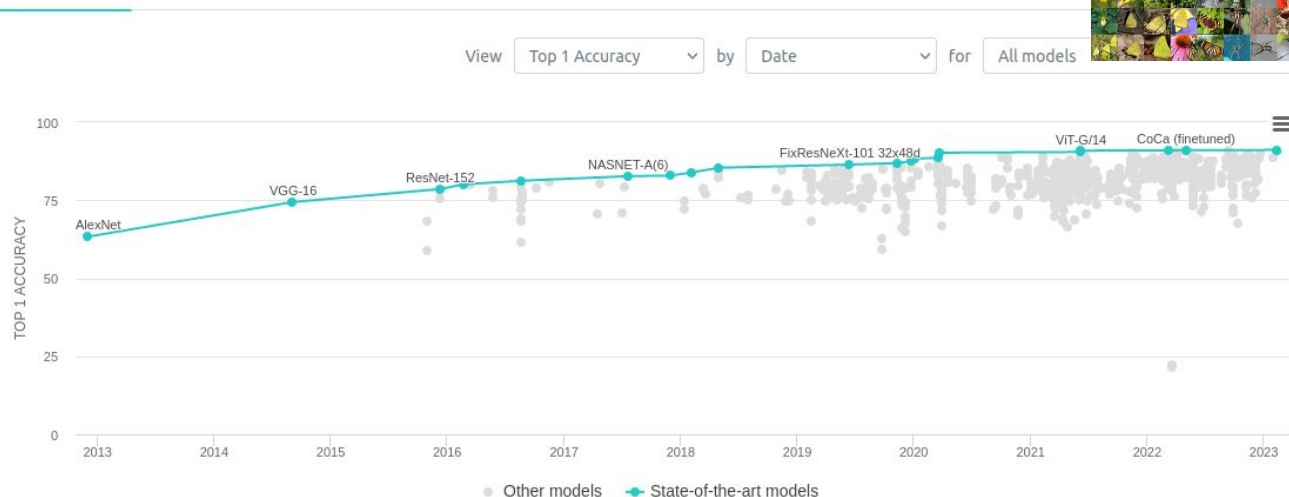
Why AI goes so fast now?

1) GPUs

Graphical Processing Units (graphic card)

- Designed to compute many small operations in parallel
- graphic cards in computers
- are now used for deep learning
- very efficient for matrix multiplication

Meta will have 350,000 of Nvidia's fastest AI GPUs by end of year, buying AMD's MI300, too

News By Anton Shilov published January 19, 2024

Meta expects to have 600,000 H100 GPU-equivalent AI horsepower by end of 2024.

2) Data

Machine learning needs data

Imagenet (2009) https://www.image-net.org/

- 1000 object classes
- 1 million images

Laion-5B (2022) https://laion.ai/

- 5 billion image-text pairs
- scrapped from the web

The Pile https://pile.eleuther.ai/

An 800GB Dataset of Diverse Text for Language Modeling

what is the size of all the Harry Potter books?
https://www.kaggle.com/datasets/moxxis/harry-potter-lstm

r/singularity • il y a 8 j
SnoozeDoggyDog

Leaked Documents Show Nvidia Scraping 'A Human Lifetime' of Videos Per Day to Train AI

AI

3) Open-source

- GAFAM develop free and open source tools for deep learning (Pytorch, Tensorflow…)
- Active open-source community
  - free access to software
  - free access to course and learning
  - easy to contribute

Non-AI people are using AI tools: generating images, generating text with chatGPT

An active open-source community

Active open-source community

- easy to contribute with Git and Github
- scikit-learn
- ML tutorials
- Kaggle
- Huggingface provides coded, pre-trained models, ready to use

https://scikit-learn.org/

https://github.com/

https://www.kaggle.com/

https://huggingface.co/

AI going too fast?

AI and images, generative models

New AI methods with many applications to images:
image generation from text prompt, denoising, inpainting,
super-resolution…

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

generative AI

Images
Twitter search: #Midjourney #Pope

Videos
https://sora.aitubo.ai/
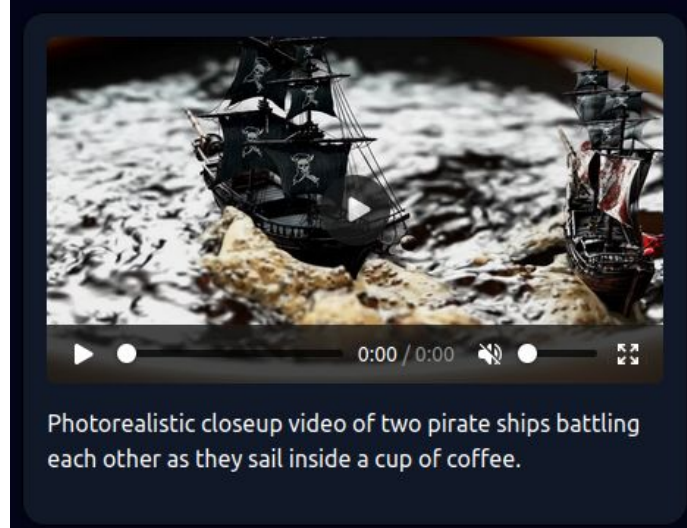
Photorealistic closeup video of two pirate ships battling each other as they sail inside a cup of coffee.

https://x.com/EccentrismArt/status/1823059492520788342

AI generated music
https://suno.com/

Generating clichés, gender and diversity bias

https://openai.com/blog/reducing-bias-and-improving-safety-in-dall-e-2

A portrait of a software engineer   Generate

Before mitigation   After mitigation

Awful AI

https://github.com/daviddao/awful-ai

More info on AI (BBC article)

FUTURE NOW | ARTIFICIAL INTELLIGENCE

The A-Z of AI: 30 terms you need to understand artificial intelligence

https://www.bbc.com/future/article/20230717-what-you-should-know-about-artificial-intelligence-from-a-z

More history

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

Timeline for AI and images: https://www.fabianmosele.com/ai-timeline

https://www.historyofdatascience.com/