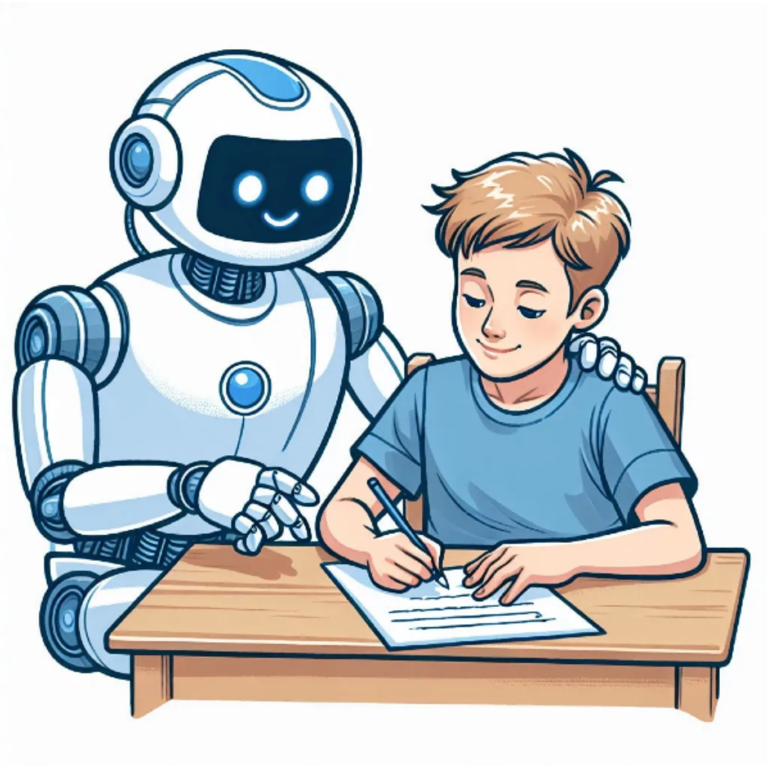
AI & Machine Learning in Physics



PHYS3151 (6 credits)

Time & Place : Tue 16:30-17:20; 17:30-18:20 KKL 201
Fri 17:30-18:20 KKL 201

Teachers: Zi Yang Meng (zymeng@hku.hk), HOC 231

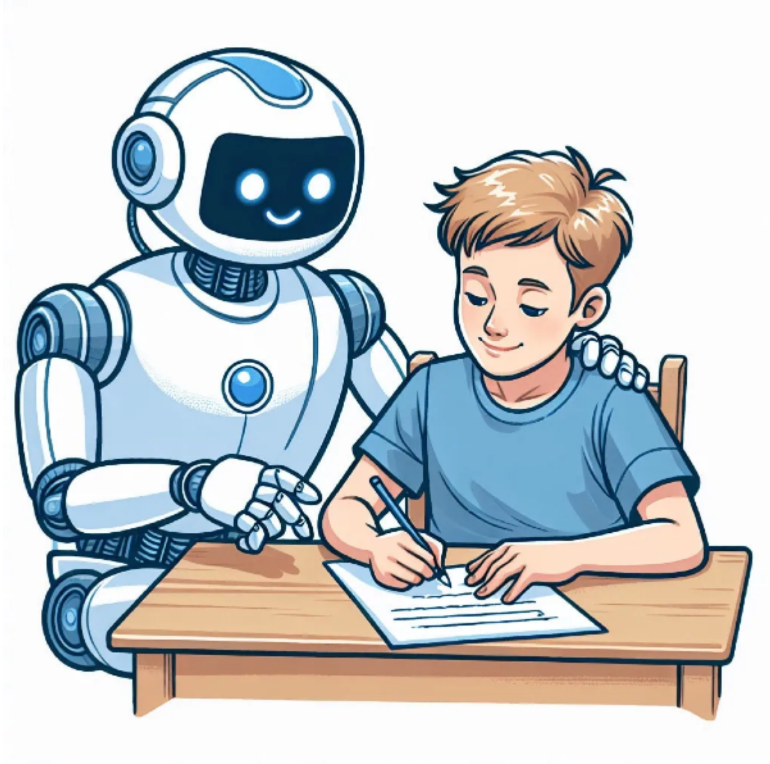
Tutor: Tim-Lok Chau (Justin) (justintlchau@connect.hku.hk)

AI & Machine Learning in Physics

Teaching Materials:

https://quantummc.xyz/teaching/hku-phys3151-machine-learning-in-physics-2025/

Slides / Reading materials
Python notebooks
Assignments



Assessment Methods and Weighting

- Assignments 30%
- Presentation 20%
- Project report 20%
- Exam. 30%

AI & Machine Learning in Physics

Literature: Books

there are many, actually too many

- Ethem Alpaydin, Introduction to Machine Learning, Third Edition, MIT Press 2014
- Simon Haykin, Neural Networks and Learning Machines, Third Edition, Pearson 2009
- Stuart Russell, Peter Norvig, Artificial Intelligence, Third Edition, Pearson 2010

Literature: Online material

Andrew Ng, Stanford University http://www.holehouse.org/mlclass/

https://www.youtube.com/playlist?list=PLLssT5z_DsK-h9vYZkQkYNWcItqhIRJLN

Neuroscience For Kids http://faculty.washington.edu/chudler/neurok.html

Michael Nielsen, scientist at home, the best reading material for NN

http://neuralnetworksanddeeplearning.com

AI & Machine Learning in Physics

In the era of AI & Big data

ChatGPT

In April 2017, AlphaGo vs. Jie Ke

- The machine played perfect···
- I am so behind, unbelievable···
- AlphaGo is not the God, but it is a superior species than human being···

AlphaGo

Smart Robots https://www.bostondynamics.com/ **Self-driving Car**

AI & Machine Learning in Physics

AlphaGo-1

In March 2016, AlphaGo played with Lee Sedol in Seoul.

- ➢ its psychological aspects (its relentless concentration) no human can match it···
- ➢ The Go skill has improved surprisingly ...

AI & Machine Learning in Physics

AlphaGo-2

In April 2017, AlphaGo vs. Jie Ke

- The machine played perfect…
- I am so behind, unbelievable…
- AlphaGo is not the God, but it is a superior species than human being…

AI & Machine Learning in Physics

What happened to AlphaGo?

1. In January 2016, researchers revealed that AlphaGo could play 1 million games in 4 weeks. This means that AlphaGo could play 30,000 games per day. How many games could Lee and Ke play? 10 games/day, 82 years (30,000 days) , only 300,000 games only. AlphaGo has played 300,000,000 games after march 2016!!!→ **Big Data!!!**

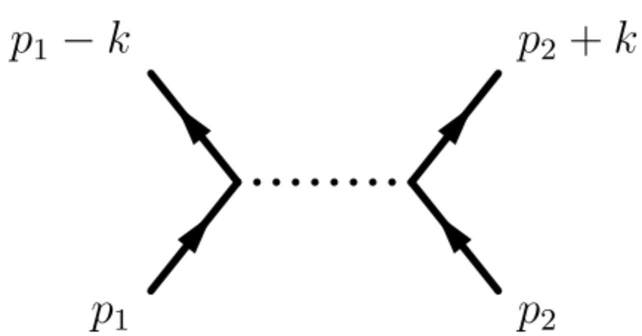
2. Till now, only ~2 million games in total. AlphaGo actually learned from data generated by itself→ **Self learning!!!**

3. Software/Algorithm: 12 layers → 40 layers. **Make it more complicated!!!**

4. Hardware: TPU1 → TPU2. **New Hardware is necessary!!!**

Tensor processing unit (TPU) by google

Hideki Yukawa 汤川秀树

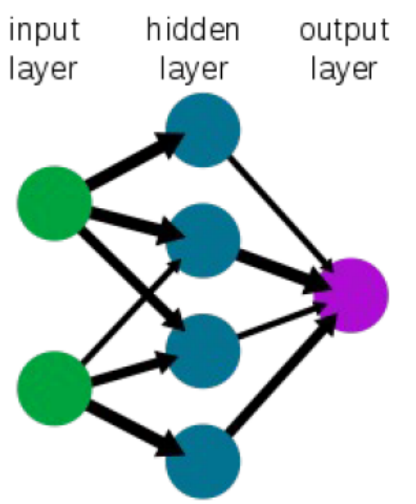


AI & Machine Learning in Physics

Challenges 1: models are more complicated

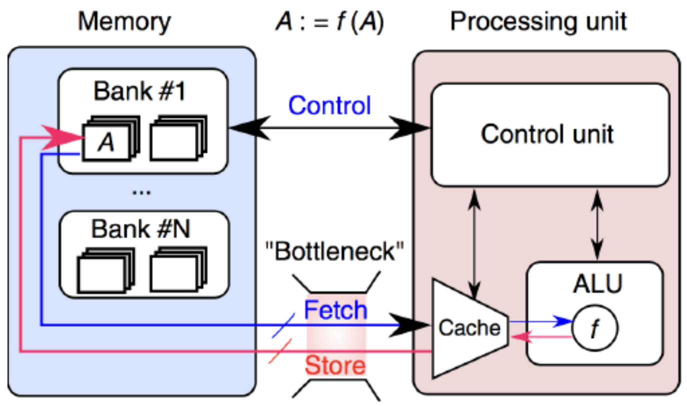
~ 100 layers, ~ 10^6 weights/parameters

A simple neural network



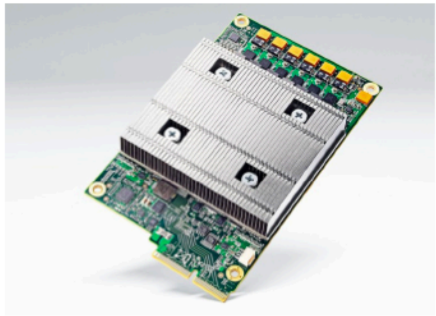
Challenges 2: memory bottleneck

Data fetch is much expensive than data process

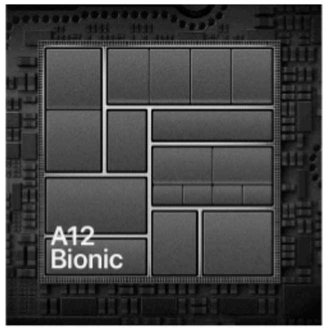


Large on-chip memory, bring computing and memory closer, using low precision computing.

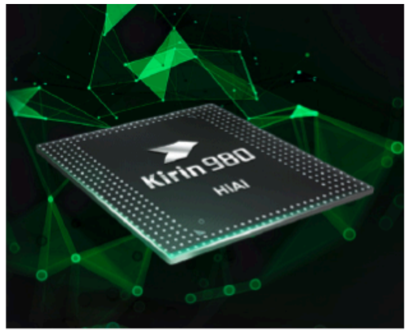
Neural Processing unit (NPU) for AI computing



TPU by Google



A12 Bionic by Apple



Kirin 980 by Huawei

AI & Machine Learning in Physics

Challenges 3: energy consumption

AlphaGo:
- ❑ 176 GPUs, 1202 CPUs
- ❑ 150, 000 Watts

Jie Ke:
- ❑ 1.2L Human Brain
- ❑ ~20 Watts

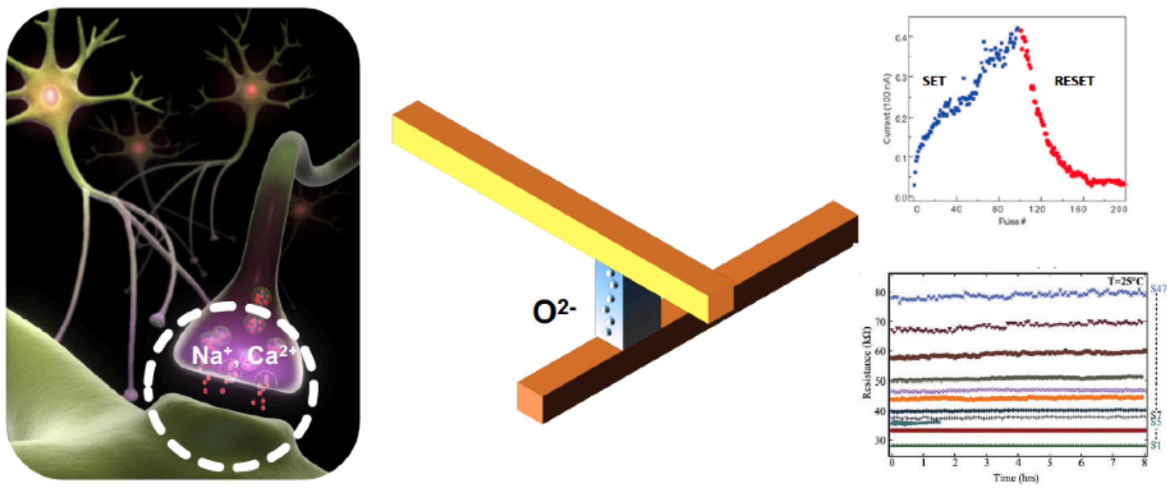
Huge power gap between human brain and CMOS-based AI system

➢ It is much needed to develop **new hardware** with **new device** and **new architecture and new algorithm.**

AI & Machine Learning in Physics

Bio-synapse vs. Electrical synapse

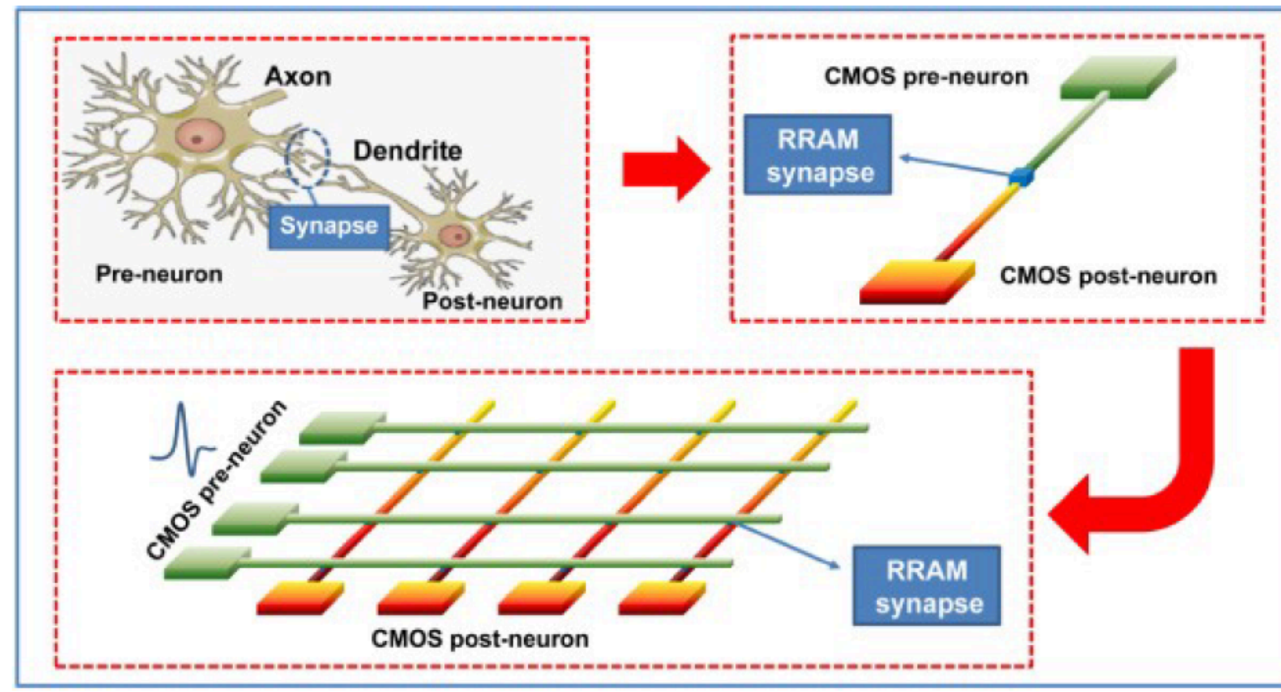
Bio-synapse **Memristor-Electrical synapse**



Bio-synapse conductance change through the Na^{+}, CA^{2+} ions movement

Memristor device conductance change through O^{2-} ions movement

Neuromorphic computing



$$\mathbb{R}^{N\times N}$$

$$Ax = b$$

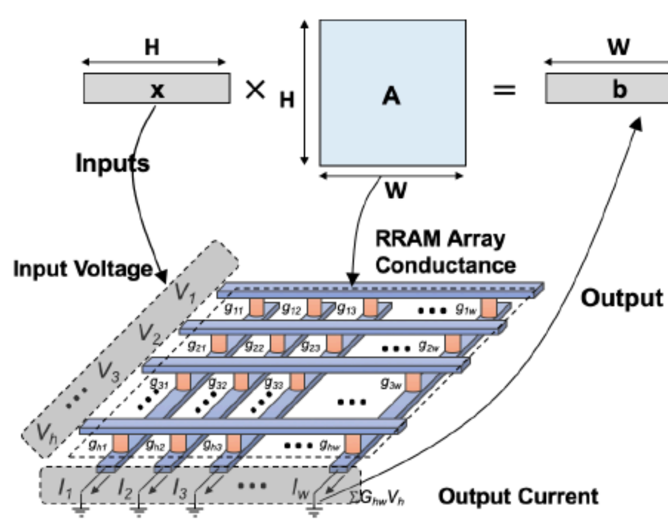
$$O(N^2)$$

$$x = A^{-1}b$$

$$O(N^3)$$

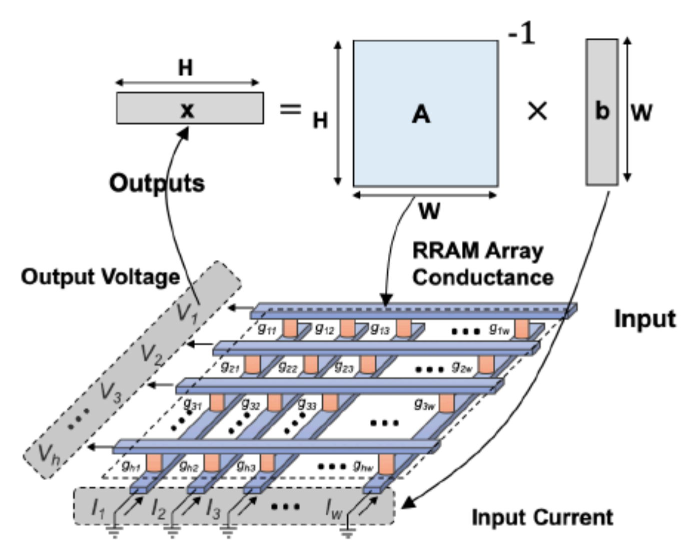
- Vector-Matrix-Multiplication (VMM)

$\boldsymbol{A}x = b$



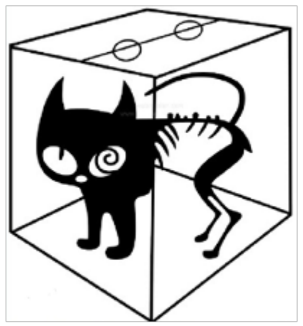
❑ Linear equation solver

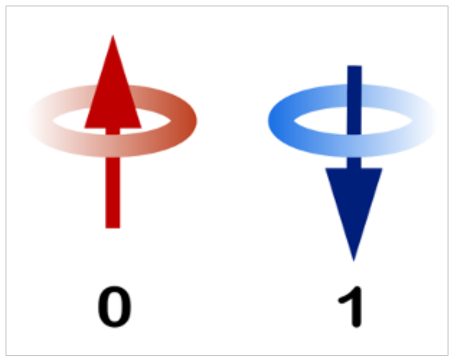
$\boldsymbol{A}x = b \,{-}{>}\, x = \boldsymbol{A}^{-1}b$

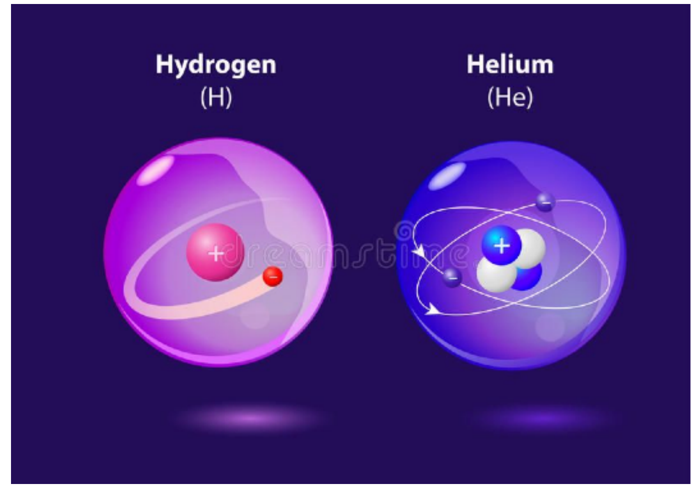


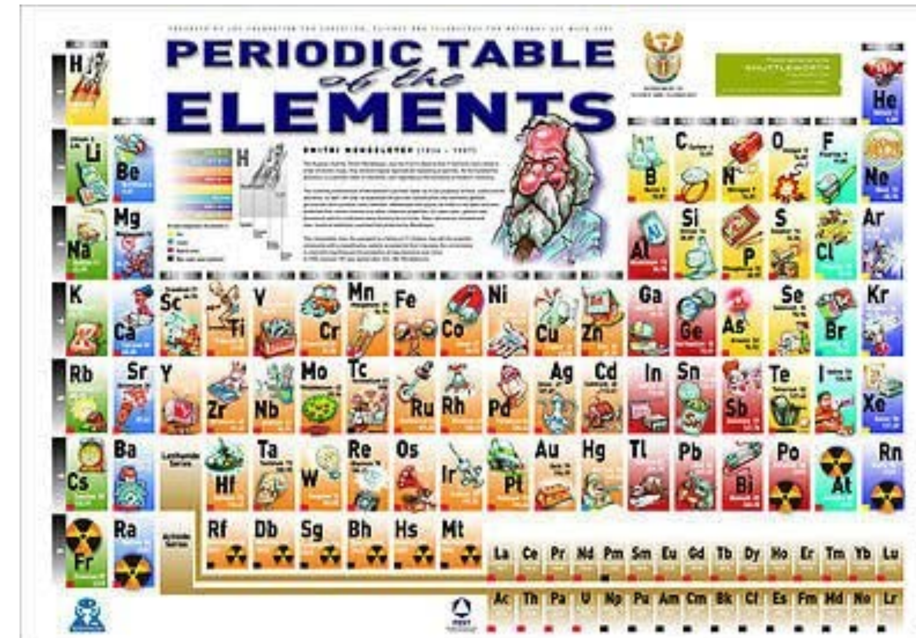
Computing with physical law in memory: Ohm' Law and Kirchhoff' Law

Computation on quantum bit and quantum entanglement









Lead to the famous "exponential wall"

2^N Multi-electron atoms, cannot be solved exactly

Krishna and Radha playing chaturanga

$N = 10 \quad 2^{10} = 1{,}024 \sim 10^3$

$N = 20 \quad 2^{20} = 1{,}048{,}576 \sim 10^6$

$N = 30 \quad 2^{30} = 1{,}073{,}741{,}824 \sim 10^9$

$N = 40 \quad 2^{40} = 1{,}099{,}511{,}627{,}776 \sim 10^{12}$

$N = 50 \quad 2^{50} = 1{,}125{,}899{,}906{,}842{,}624 \sim 10^{15}$

Wheat grains on chessboard — Sissa ibn Dahir, inventor of Chaturanga

$2^{64} - 1$ = 18,446,744,073,709,551,615 grains of wheat, weighing about 1,199,000,000,000 tons.
About 1,645 times the global production of wheat.

Solving exponentially complex problem in polynomial time

Computation and AI could solve the energy crisis

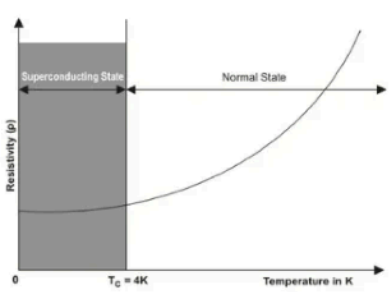
CLP 中電 electric power transmission at high voltage

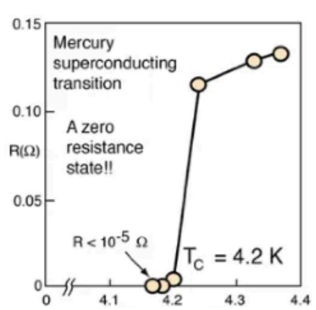
High speed rail

Maglev (magnetic levitation) bullet train with 600 k/h

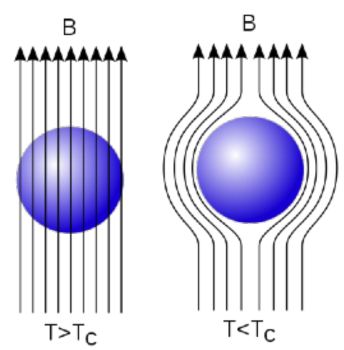
- **Understanding Quantum Metals**

What is Superconductivity?



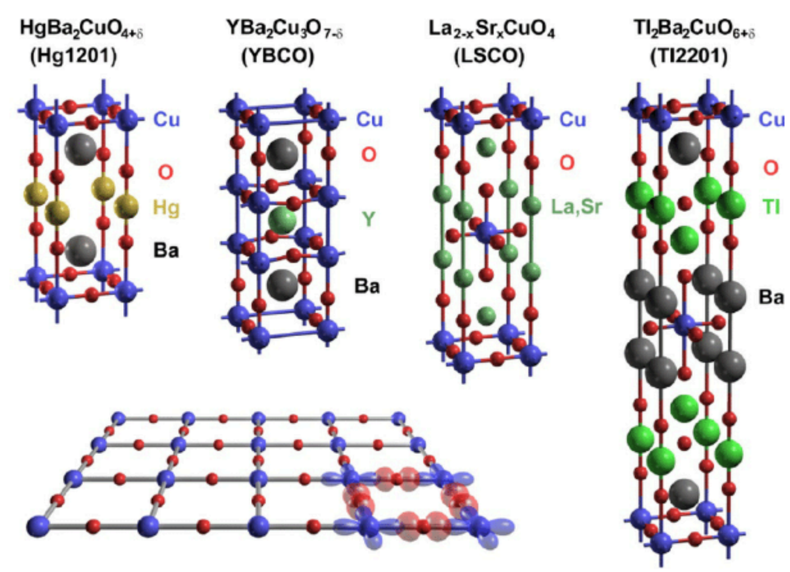


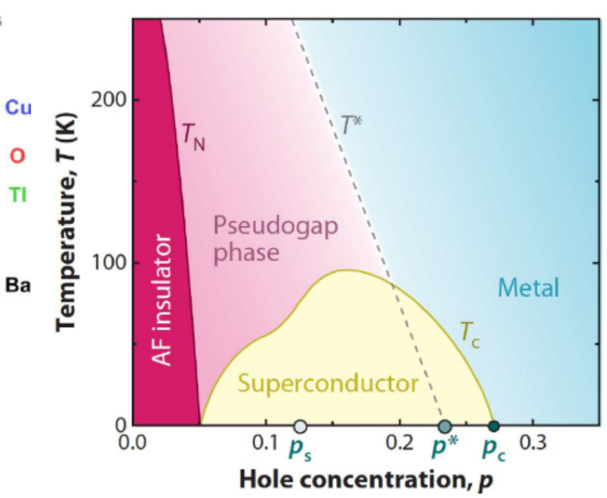
Dutch Physicist
Heike Kamerlingh Onnes
in 1911



superconductors at ~ -270°C

High-temperature superconductors at ~ -100°C

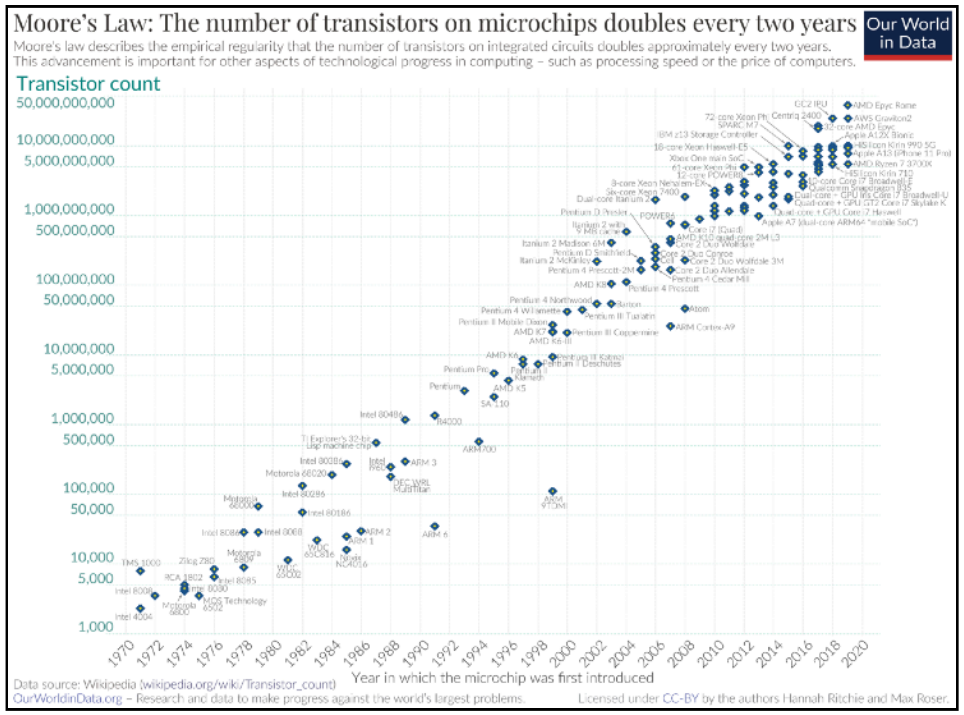


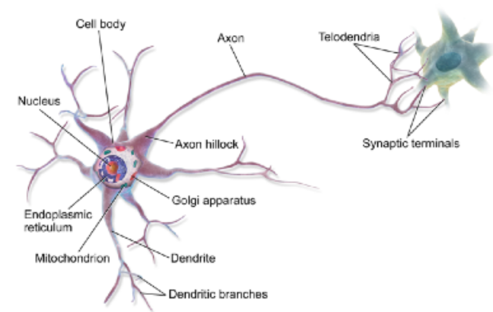


Computation becomes easy

50 years of supercomputer tracks Moore's law

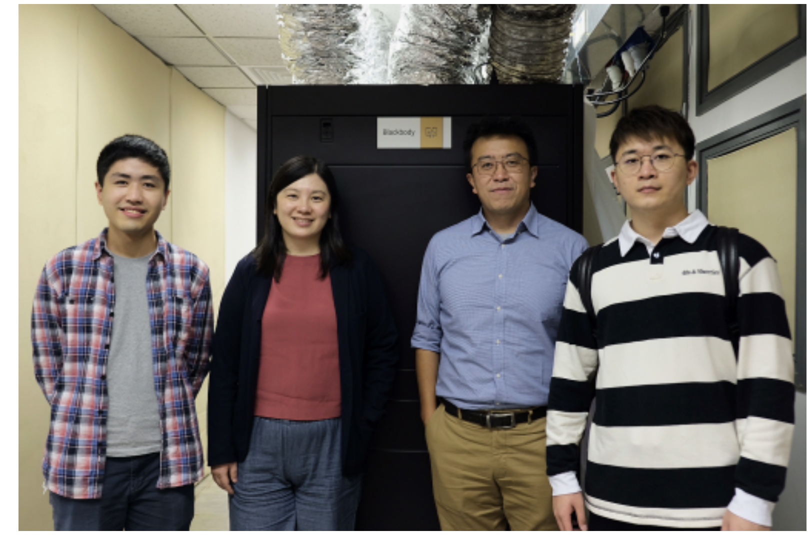
 # transistors doubles every 2 years





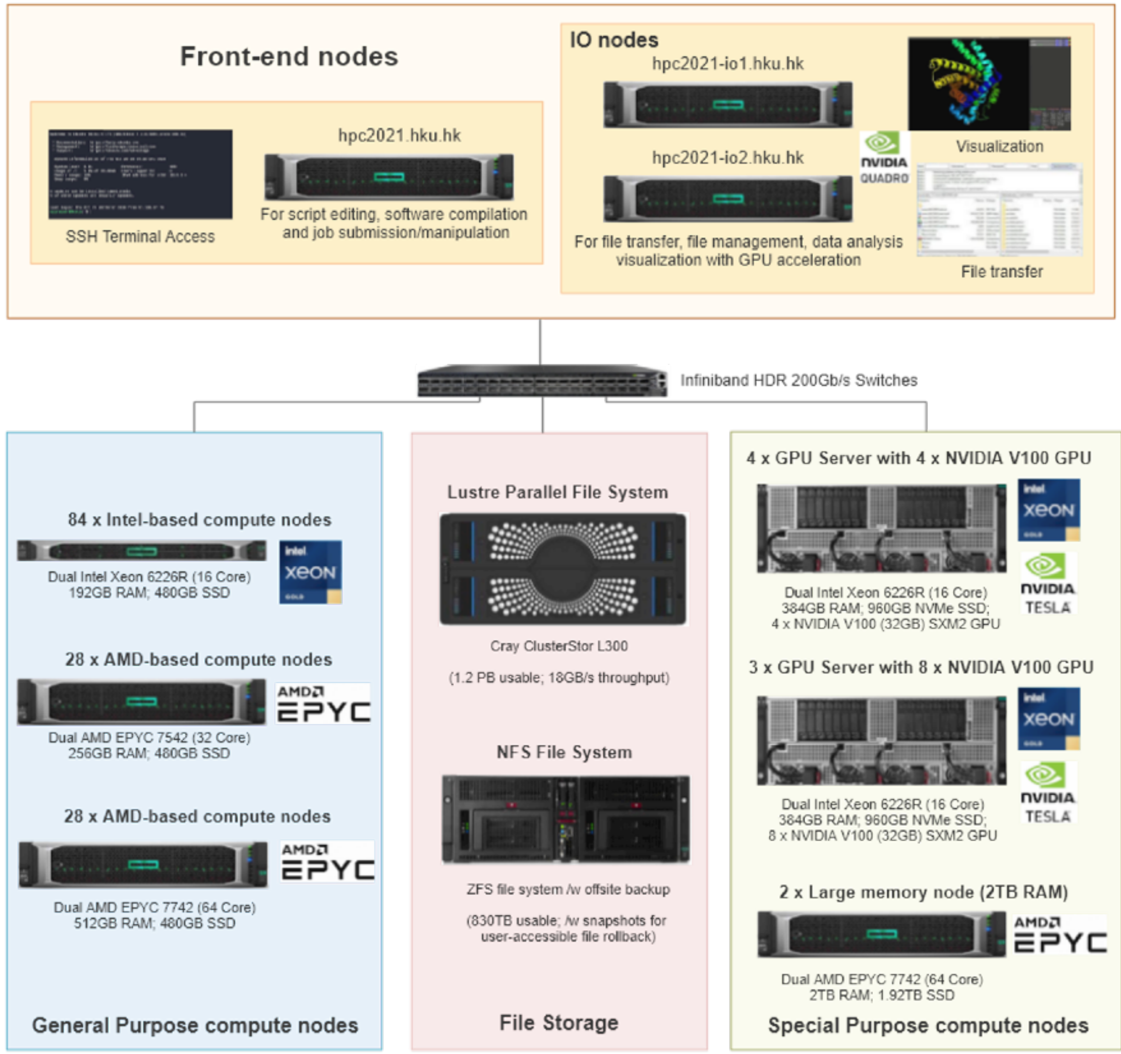
	Supercomputer	Personal Computer	Human Brain
Computational Units	32,000 Xeon CPUs 10^12 transistors	4 CPUs, 10^9 transistors	10^11 neurons
Cycle time	10^-9 sec	10^-9 sec	10^-3 sec
Operations/sec	10^15	10^10	10^17
Memory updates/sec	10^14	10^10	10^14
Weight / Space	150 tons / Basketball court	1 Kg / A4 Paper	1.5 Kg / 1/6 basketball
Power consumption	500 megawatt	100 watt	20 watt

Our own Blackbody



HPC2021

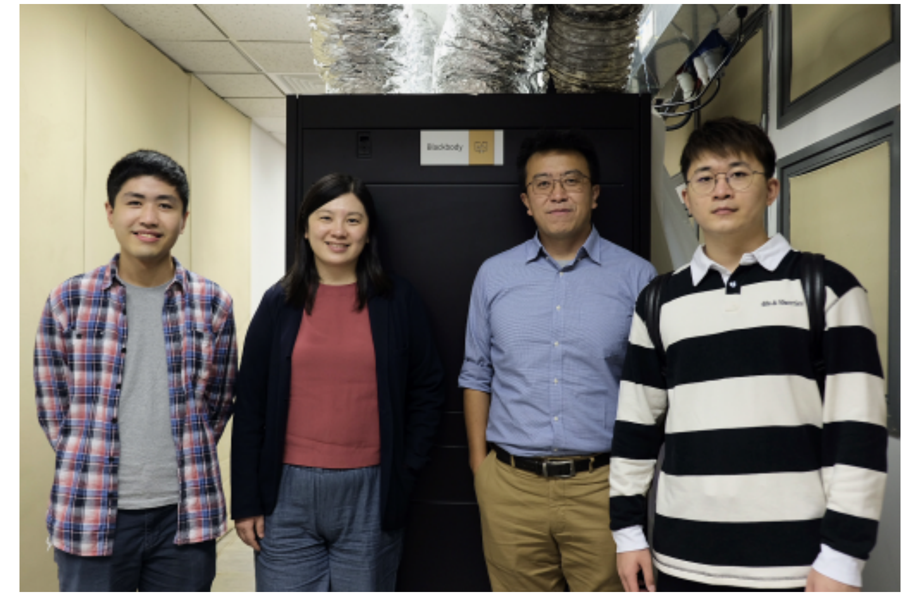
8064 cpu cores



Our own Blackbody

1024 cpu cores

2022/09



Tianhe-II: 16,000 node, 24 Intel Xeon E5 core CPU, 384,000 in total

2023/11

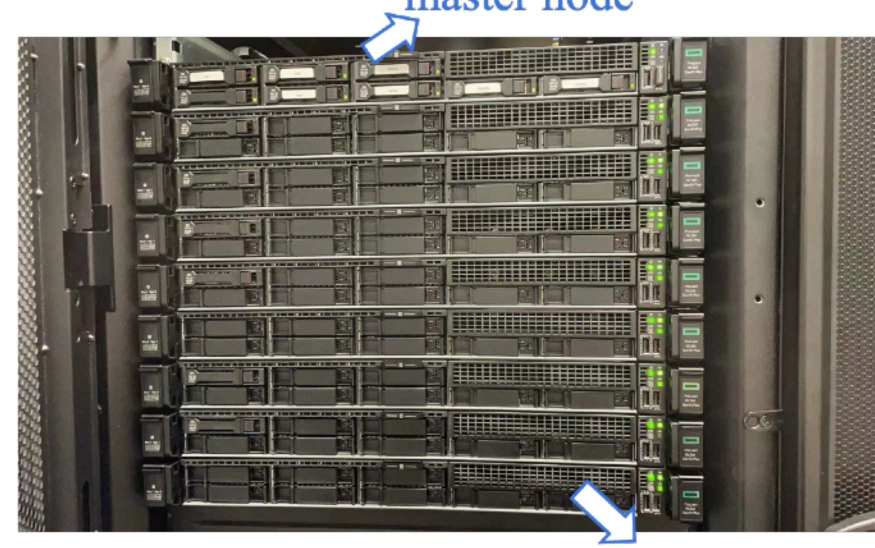
AMD 7702P (64 core) x 2 x 10 = 1280 cores
AMD 7573X (32 core) x 2 x 1 = 64 cores
AMD 7763 (64 core) x 2 x 7 = 896 cores
AMD 9654 (96 core) x 2 x 2 = 384 cores

Intel(R) Xeon(R) Gold 6226 (12 cores) 2 x 2 =48 cores (head node)
Intel(R) Xeon(R) Platinum 9242 (48 cores) 2 x 4 = 384 cores (computation node)

3056 CPU cores

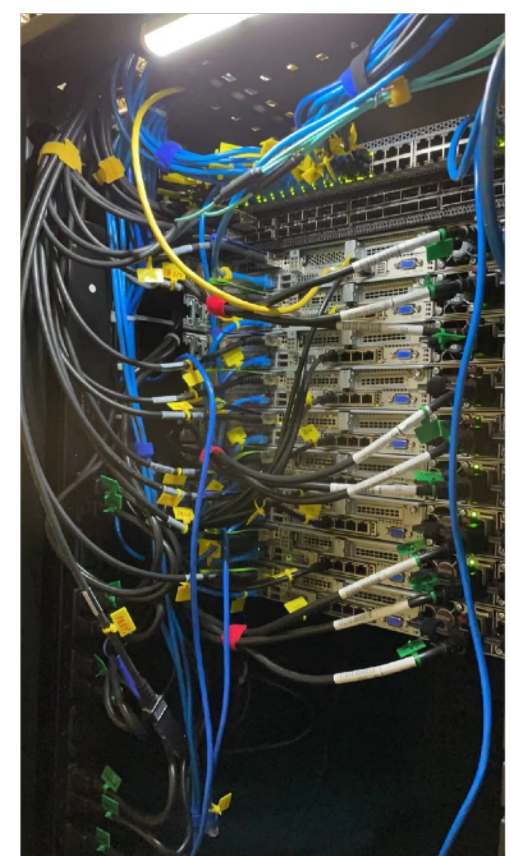
Our Blackbody Cluster in Room 311 of CYM Building

master node



compute node

Cable Management & Labeling



2 x AMD 7702 64C 2.0 GHz
512 GB RAM DDR4-3200
2 x 480 GB RI SSD RAID 1

1024 cpu cores

Exhaust Pipes and Inrow Cooling

Electricity

Total Electricity: 24 kW
Current Usage:
cluster ~ 5.4 kW (computing nodes 0.55 kW × 8, head node, storage)
cooling ~ 8 kW (each ~ 2.4 kW)

Around 10 kW electricity for future use.

AI & Machine Learning Basics

Tradition: Task —> algorithm (algorithm for loop, sorting)

Big data era: Don't have algorithm —> lack in knowledge, make up for in data

Approximation detect certain patterns or regularities, Data Mining

- Model with some parameters, model can be predictive or descriptive.
- Learning is the execution of a computer program to optimise the parameters of the model using the training data or past experience.
- Using theory of statistics, math and physics: building mathematical models, making inference from a sample
- Using computer science: efficient algorithm to solve the optimisation problem, store and process big data; representation and algorithmic solution for inference needs to be efficient
- The computational efficiency may be as important as predictive accuracy

AI & Machine Learning Basics

Infer hidden association rule from observed data
In the era of “big data”

Basket analysis

- In retail, associations between products bought by the customers
- People do not buy at random
- There are certain patterns (association rule) in the data, machine extract them

Conditional probability P(Y|X), Y is the product one would like to condition on X,

P(chips | beer) = 0.8   P(bread | milk) = 0.71

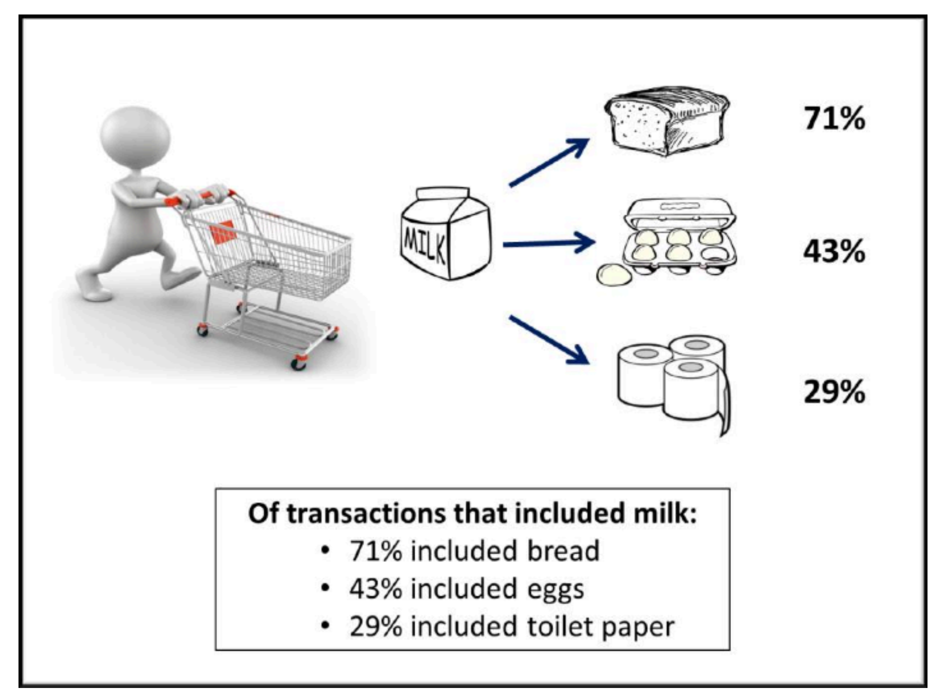
P(Y|X, D) where D is the set of customs attributes, gender, age, martial …

Books / Music / Shows :
P (Game of Thrones | Fantasy, male) =
P (The daily show | Comedy, layman) =
P (Last Week Tonight with John Oliver | Comedy, sophisticated) =
P (Late-night with Seth Meyers | Comedy, politics) = …

Webpages:

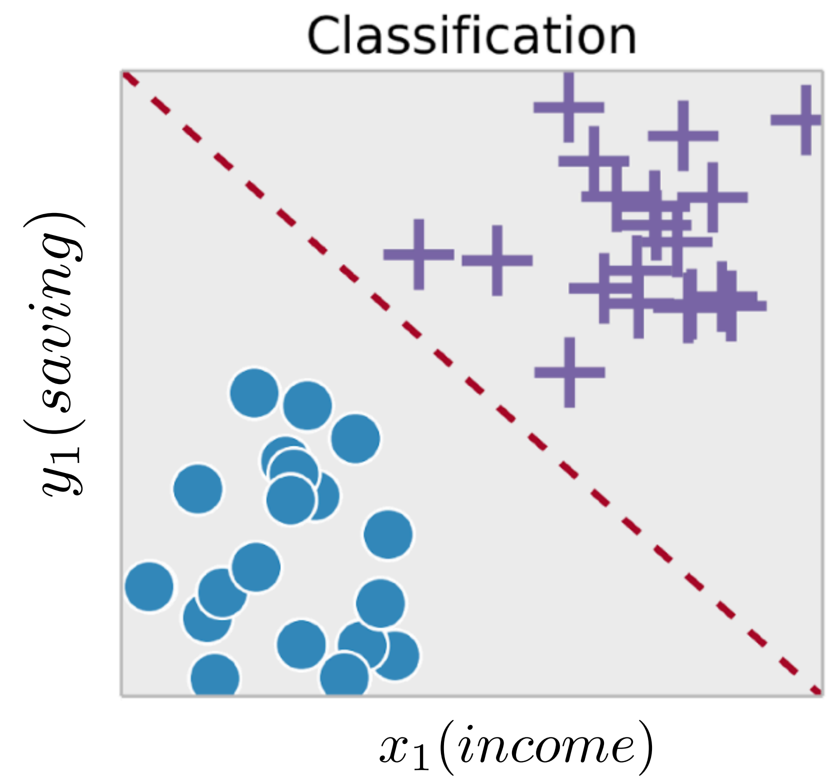
Social medias:

In spam email detection
In Fintech, credit application, stock market
In medical diagnosis, COVID-19
In Science, physics, astronomy, biology, …



AI & Machine Learning Basics

Classification: input —> classifier —> output



Discriminant (two classes):

Banks classify credit for low-risk and high-risk customers

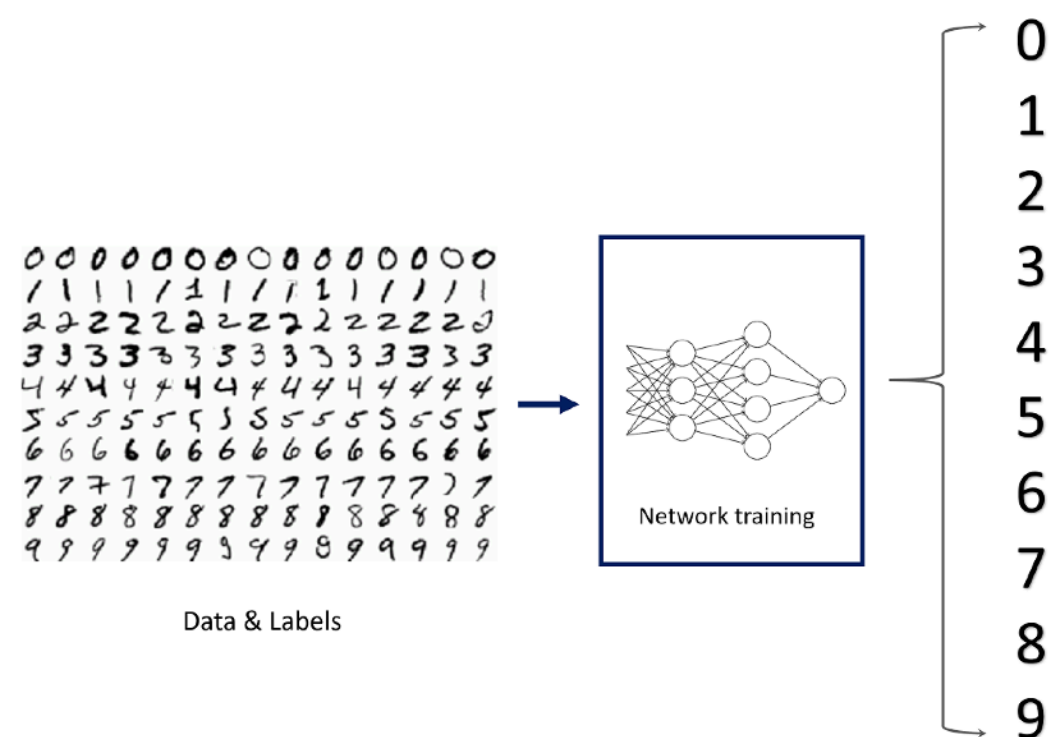
income, saving, profession, age, past financial history, …

Machine learning fits a model to the past data,
calculate the risk for a new application,
decide to accept or refuse
We have a rule that fits the past data, if the future is similar to the past

Predictions: decide new customer is low-risk and high-risk

Pattern recognition (multiple classes):

Handwritten character recognition, MNIST database
Face recognition, medical diagnosis,
Speech recognition, time series, machine translation,
natural language processing



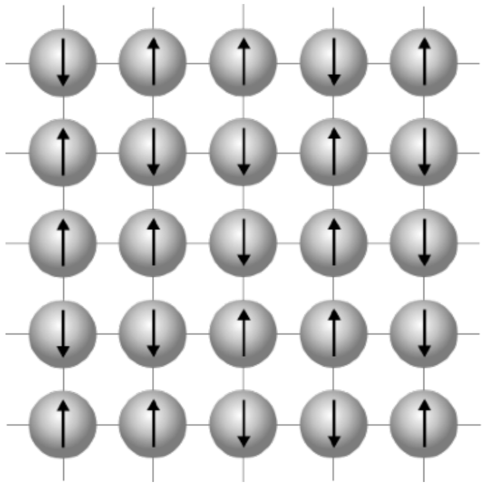
Knowledge extraction: rule is simpler than data

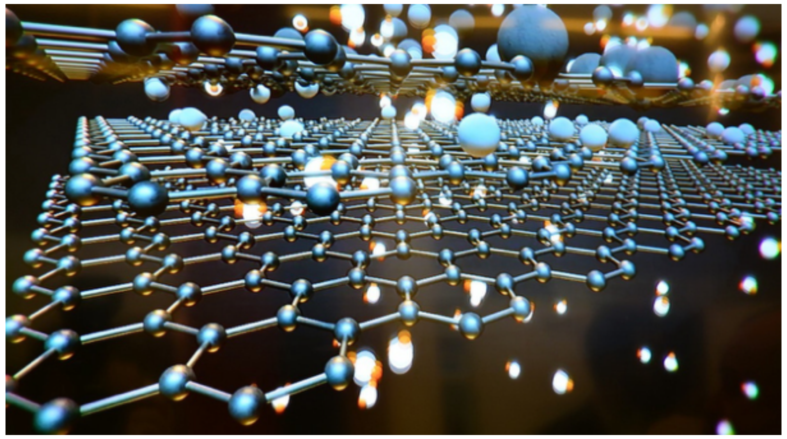
Phase transition, Landau-Ginzburg paradigm, order parameters

AI & Machine Learning Basics

Ising model, continuous phase transition
workhorse for statistical physics

https://mattbierbaum.github.io/ising.js/

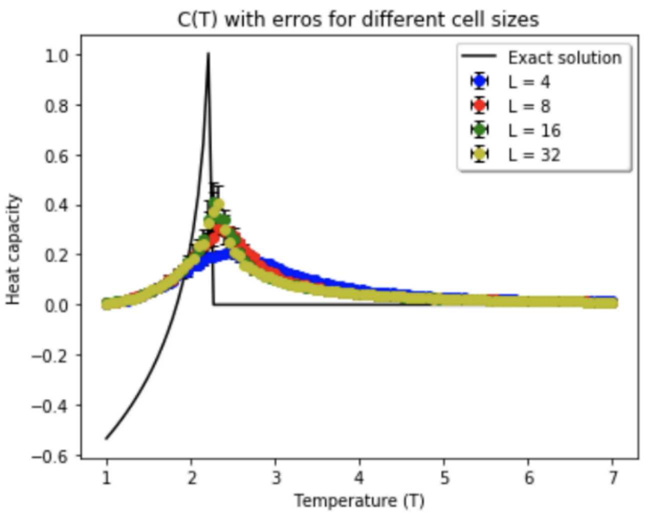


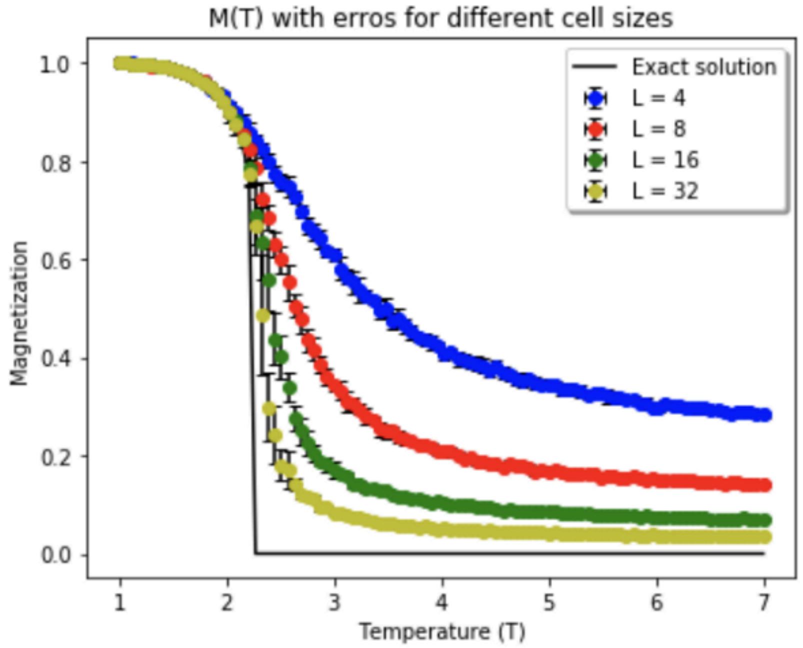


$$H = -J \sum_{\langle i,j \rangle} S_i^z S_j^z \qquad S_i^z = \pm 1$$

Configuration space: 2^N

https://en.wikipedia.org/wiki/Ising_model#/media/File:Ising_quench_b10.gif





$$m = \frac{1}{N} |\sum_{i=1}^{N} S_i^z|$$

$$m(T) = |T - T_c|^{\beta} \text{ with } \beta = 1/8 \text{ in 2D}$$

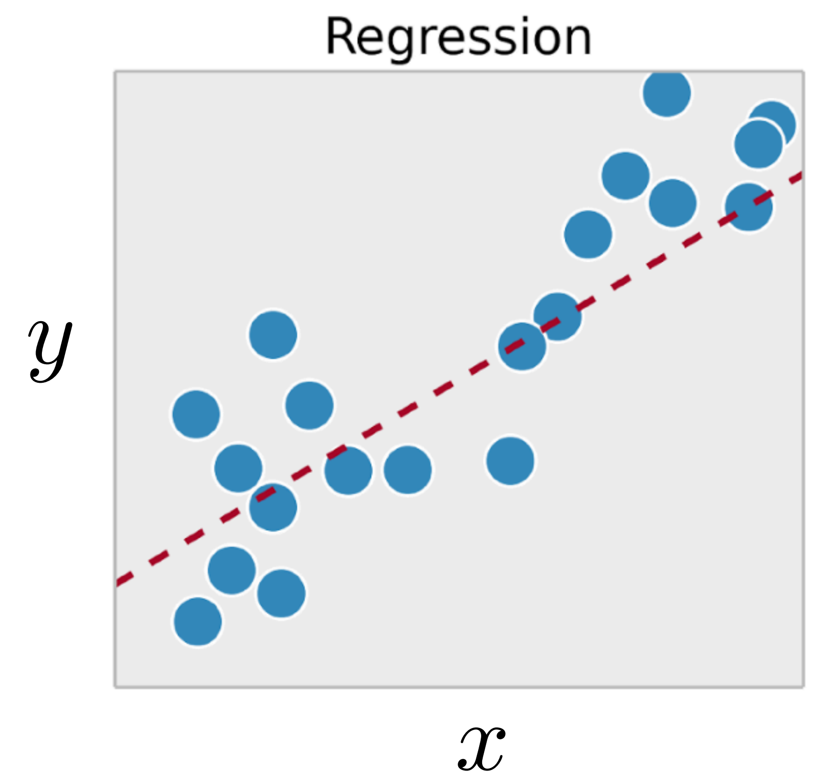
AI & Machine Learning Basics

Regression: $$y = h_\Theta(x) = \Theta \cdot x$$

$$\{(x_j^{(i)}, y^{(i)}), \theta_j\};\ j = 1, 2, \cdots, N;\ i = 1, 2, \cdots, M; N < M$$

$$y^{(i)} = \theta_0 + \theta_1 x_1^{(i)} + \theta_2 x_2^{(i)} + \cdots + \theta_N x_N^{(i)}$$

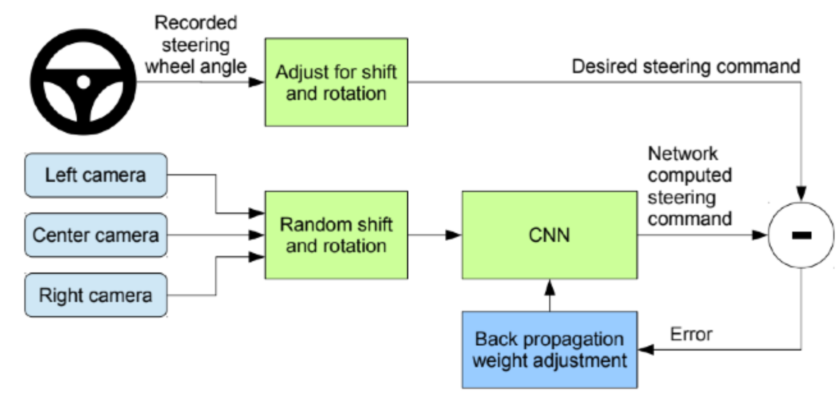
$$\begin{bmatrix} 1 & x_1^{(1)} & x_2^{(1)} & \cdots & x_N^{(1)} \\ 1 & x_1^{(2)} & x_2^{(2)} & \cdots & x_N^{(2)} \\ & \vdots & & \vdots & \\ 1 & x_1^{(M)} & x_2^{(M)} & \cdots & x_N^{(M)} \end{bmatrix} \cdot \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(M)} \end{bmatrix}$$



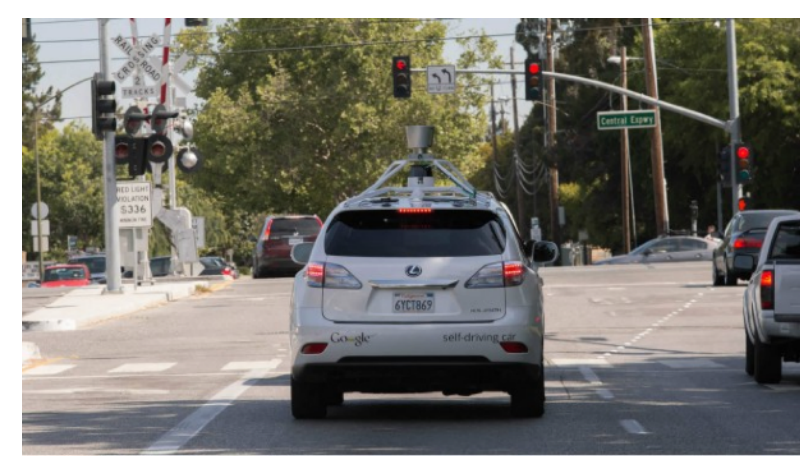
Prediction & forecasting: $$\underline{\underline{X}} \cdot \underline{\Theta} = \underline{Y}$$

Least squares by Legendre 1805 and Gauss 1809
normal equation, Gradient descent and Conjugate Gradients, Lagrange multiplier

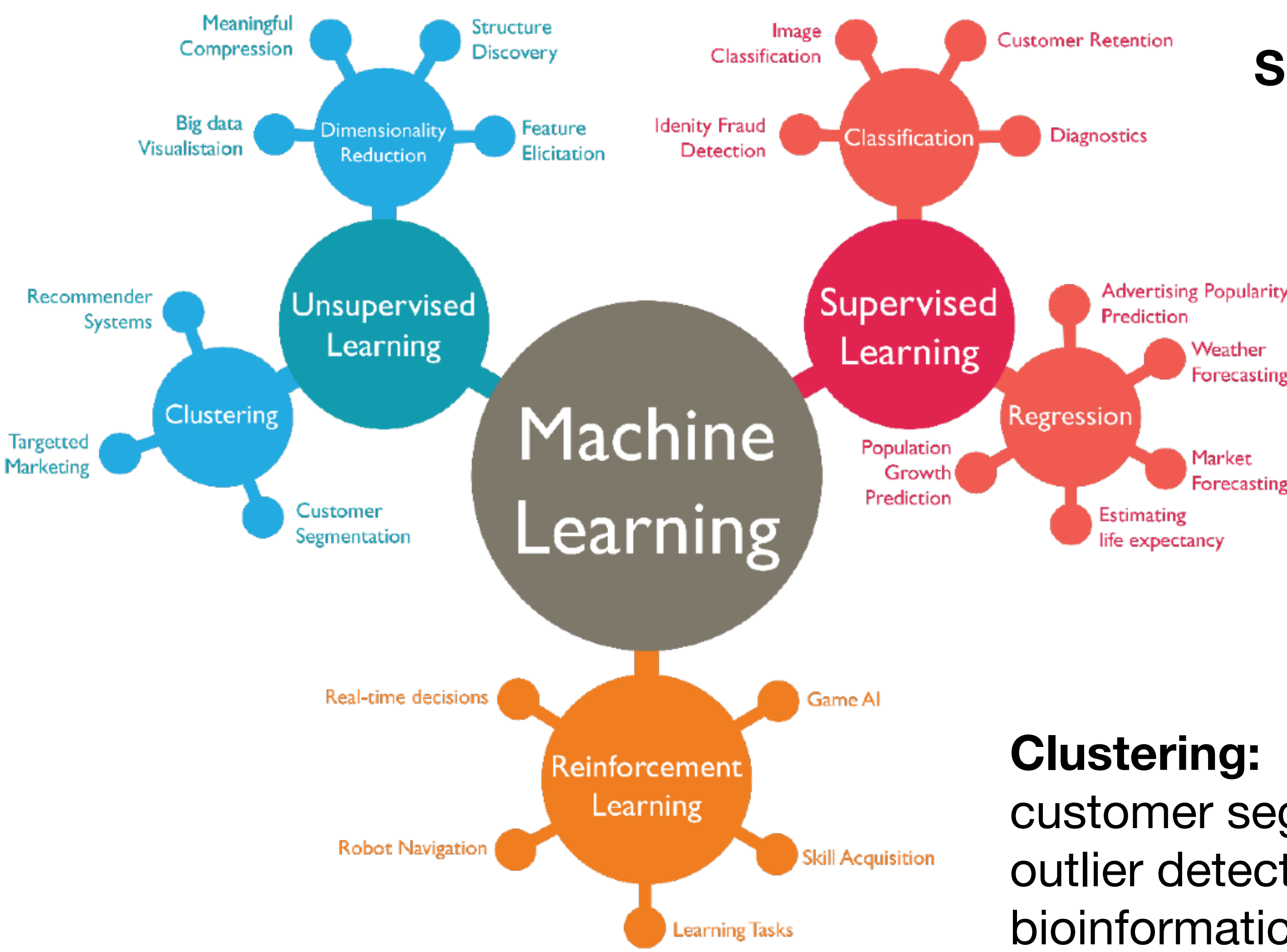
- Model / hypothesis $$h_\Theta(x)$$
- Loss / cost function $$\mathcal{L}(\Theta|X) = \sum_{i=1}^{M} L(y^{(i)}, h_\Theta(x^{(i)}))$$
- Optimisation procedure $$\Theta^* = \arg\min_\Theta \mathcal{L}(\Theta|X)$$



Example: Self-Driving car
Input: sensors on the car, video caramel, GPS, …;
Output: steering wheel;
Training data: monitoring and recording the action of human driver



AI & Machine Learning Basics



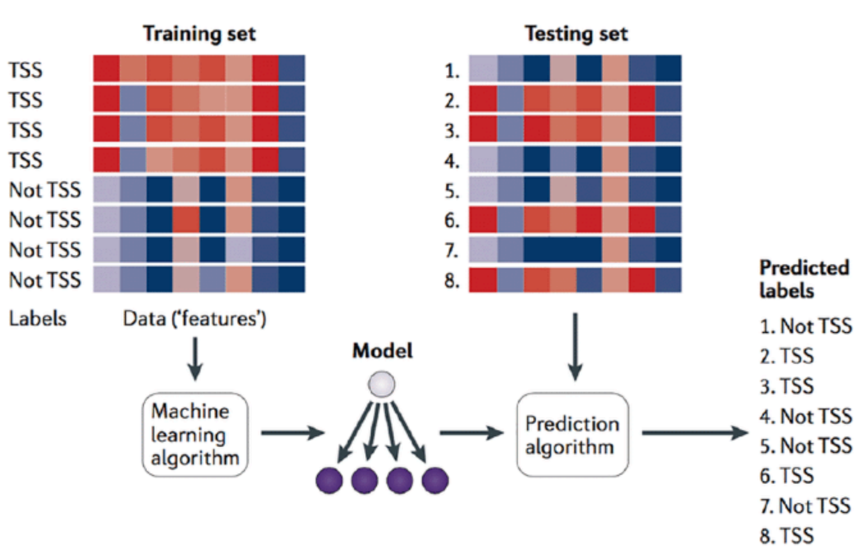
Supervised Learning: Classification & Regression

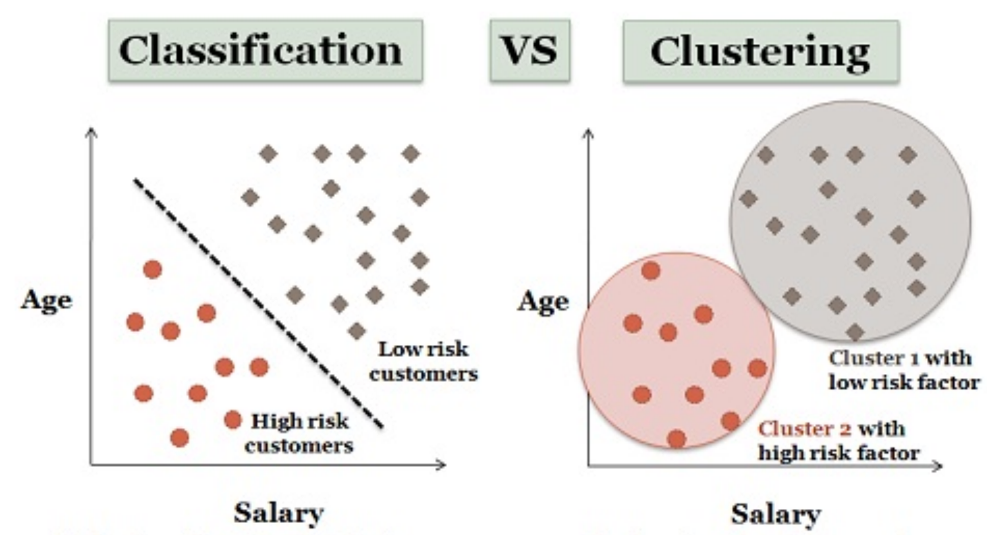
Input —> machine/model —> Output
Correct outputs are provided by the supervisor

Unsupervised Learning: only have input data

Find regularities from the input

Clustering:
customer segmentation, customer relationship management, outlier detection; Image compression
bioinformatics: DNA, RNA, amino acids, Motif, Proteins, sequence alignments





Risk classification for the loan payees on the basis of customer salary

AI & Machine Learning Basics

Reinforcement learning (policy generation):
Single action is not important, good policy is the sequence of correct actions.

Game playing:

AlphaGo is CNN with 12 convolution layers

Robot navigation:

Correct sequence of action to reach the goal state from an initial state

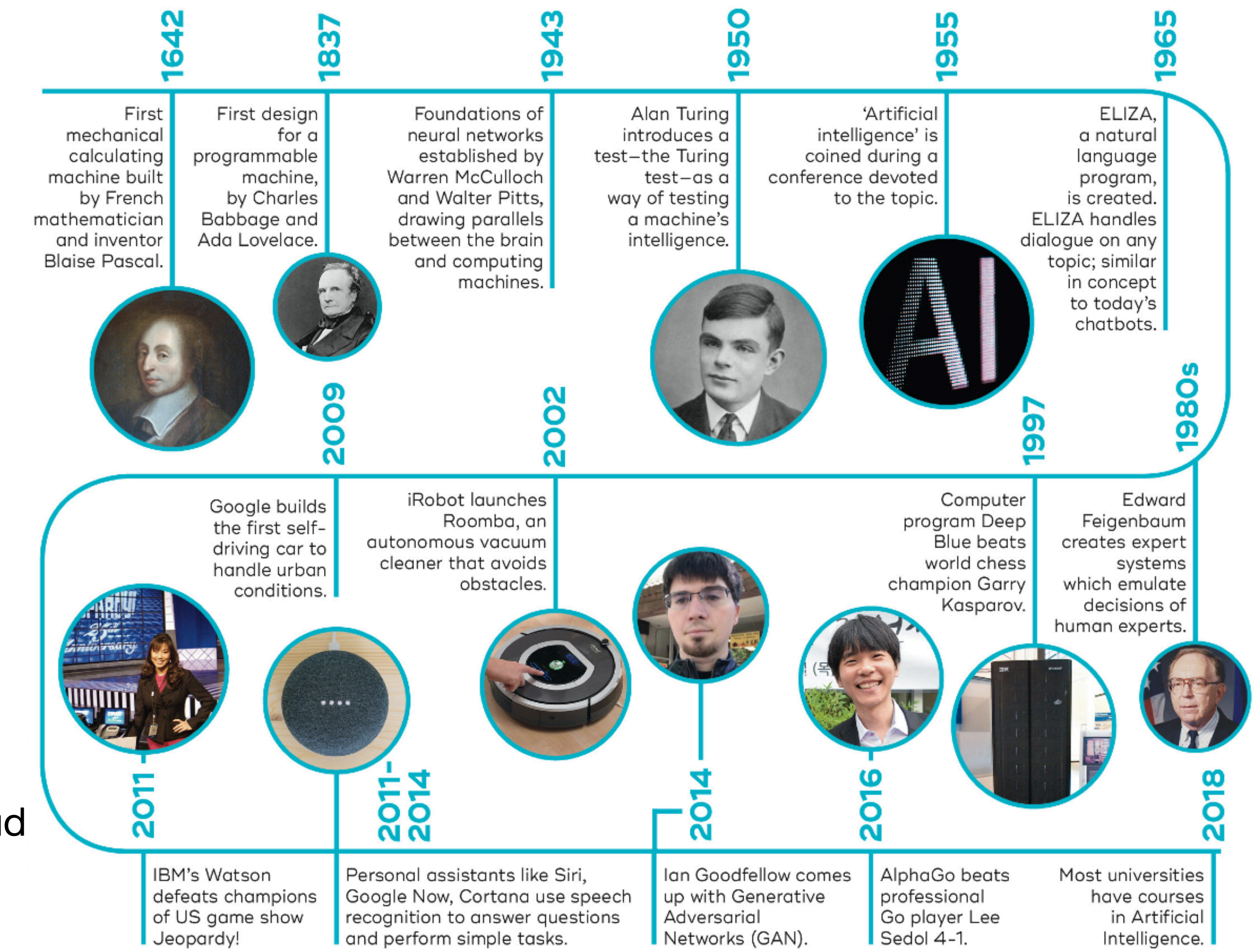
Watch this!

https://www.bostondynamics.com/spot

AI & Machine Learning Basics

A bit of history

Church-Turing thesis:
If a human could not distinguish between responses from a machine and a human, the machine could be considered “intelligent”.



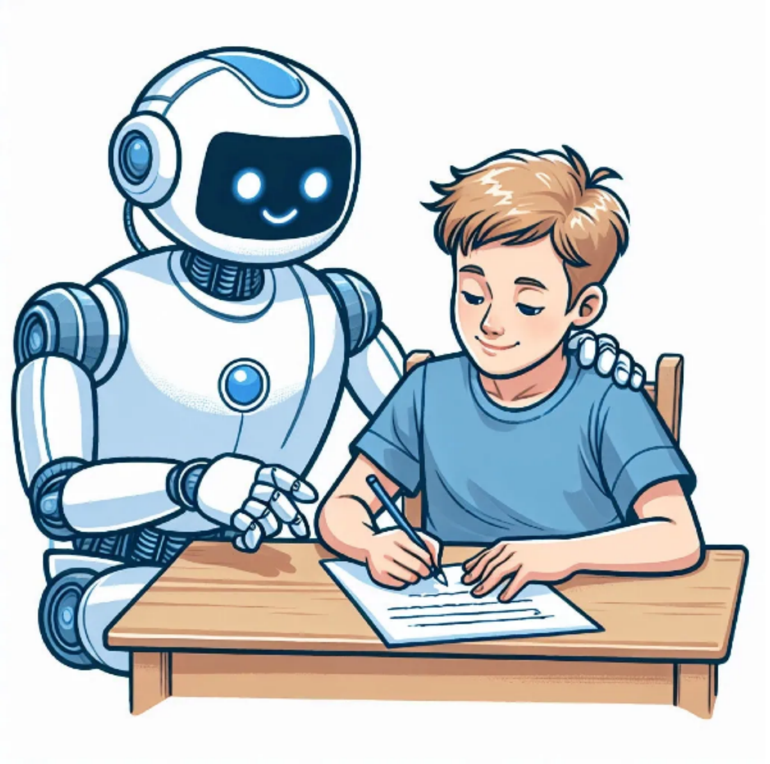
1943, McCullouch and Pitts, artificial neutrons
1955, workshop at Dartmouth College, Allen Newell (CMU), Herbert Simon (CMU), John McCarthy (MIT), Marvin Minsky (MIT), …
1958, Rosenblatt, perceptron
1974, first AI winter
1987, second AI winter
1997, IBM Deepblue vs Kasparov, logistics, data mining, medical diagnosis, …
2016, Alpha Go vs Lee Sedol
2017, Alpha Go vs Ke Jie
Xbox, Smartphone, affordable neural networks, cloud computing, internet of things …
2020, COVID-19

AI & Machine Learning Basics

Software: Python https://www.python.org/about/gettingstarted/

- Numpy and Scipy: Core toolboxes for scientic computing (mathematical functions, linear algebra routines, random number generators, optimisation, etc.)
- Matplotlib: Visualisation
- Pandas: Data manipulation
- Seaborn: Statistical visualisation
- Scikitlearn: Machine Learning
- Pytorch: Deep Learning

```
import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
import sklearn as skl
import torch
```



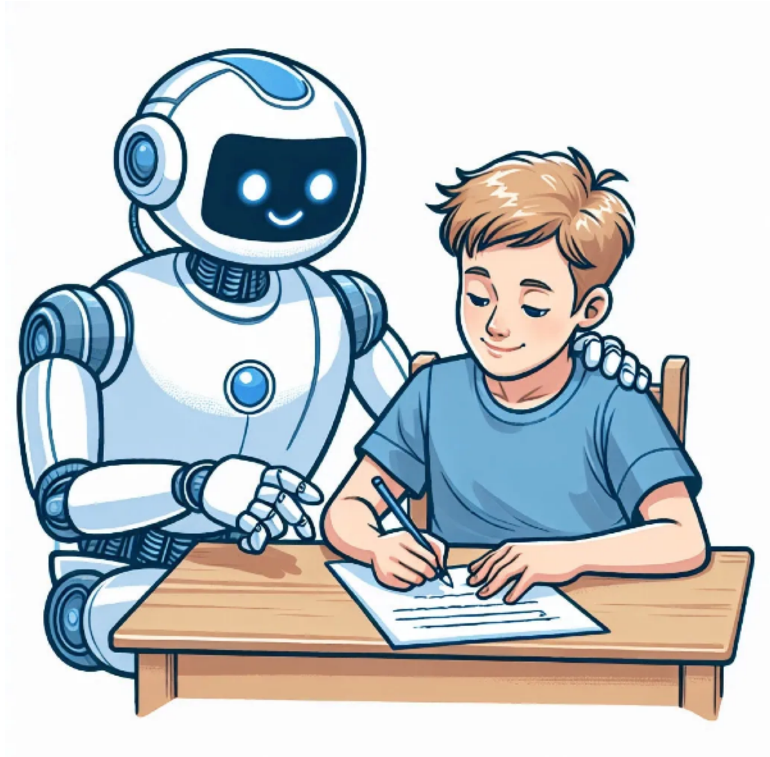
Content

0. Introduction

1. Regression
  1.1 Multivariate Linear Regression (curve fitting)
  1.2 Regularization (Lagrange multiplier)
  1.3 Logistic Regression (Fermi-Dirac distribution)
  1.4 Support Vector Machine (high-school geometry)

2. Dimensionality Reduction/feature extraction
  2.1 Principal Component Analysis (order parameters)
  2.2 Recommender Systems
  2.3 Clustering (phase transition)

Content



3. Neural Networks

3.1 Biological neural networks

3.2 Mathematical representation

3.3 Factoring biological ingredient

3.4 Feed-forward neural networks

3.5 Learning algorithm

3.6 Universal Approximation Theorem