INM431: Machine Learning

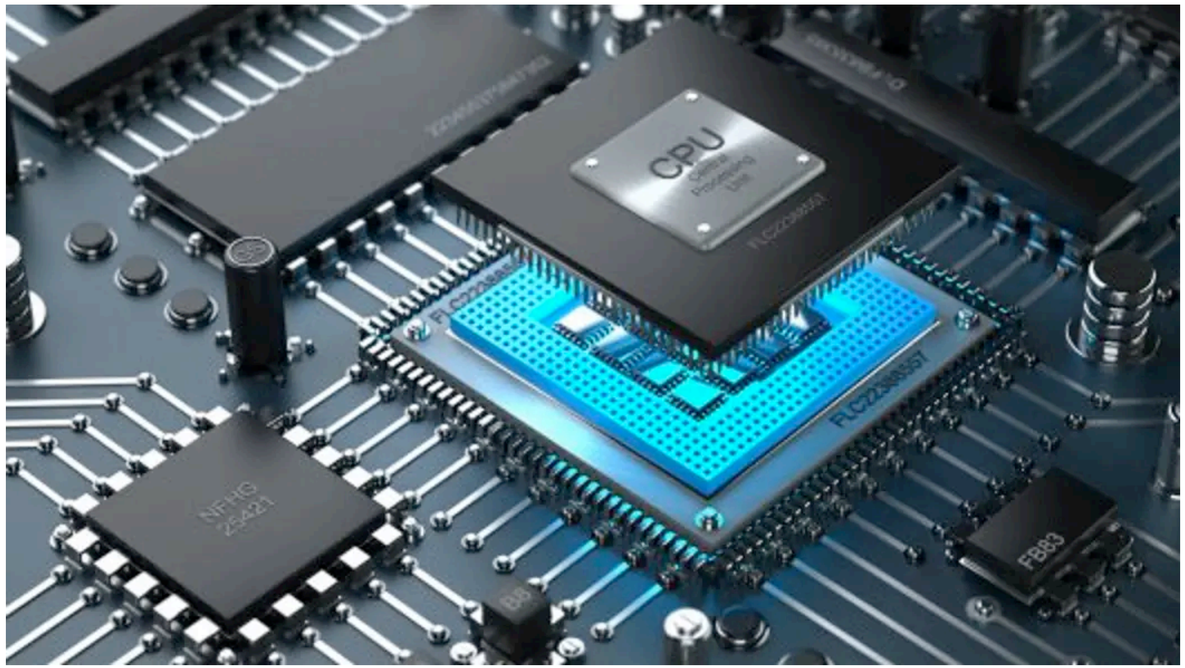
Introduction

Pranava Madhyastha (pranava.madhyastha@city.ac.uk;https://pmadhyastha.github.io)

Computer science

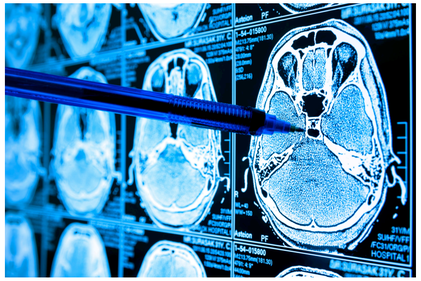
Data

Program



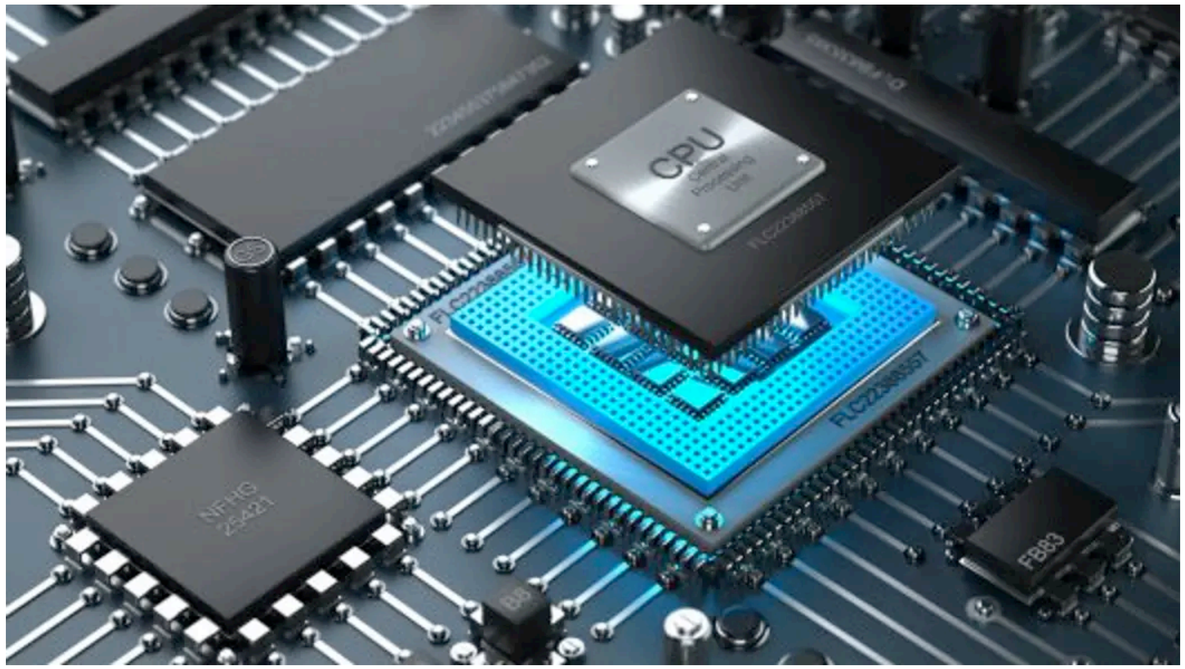
Output

Unfortunately



Data

Program

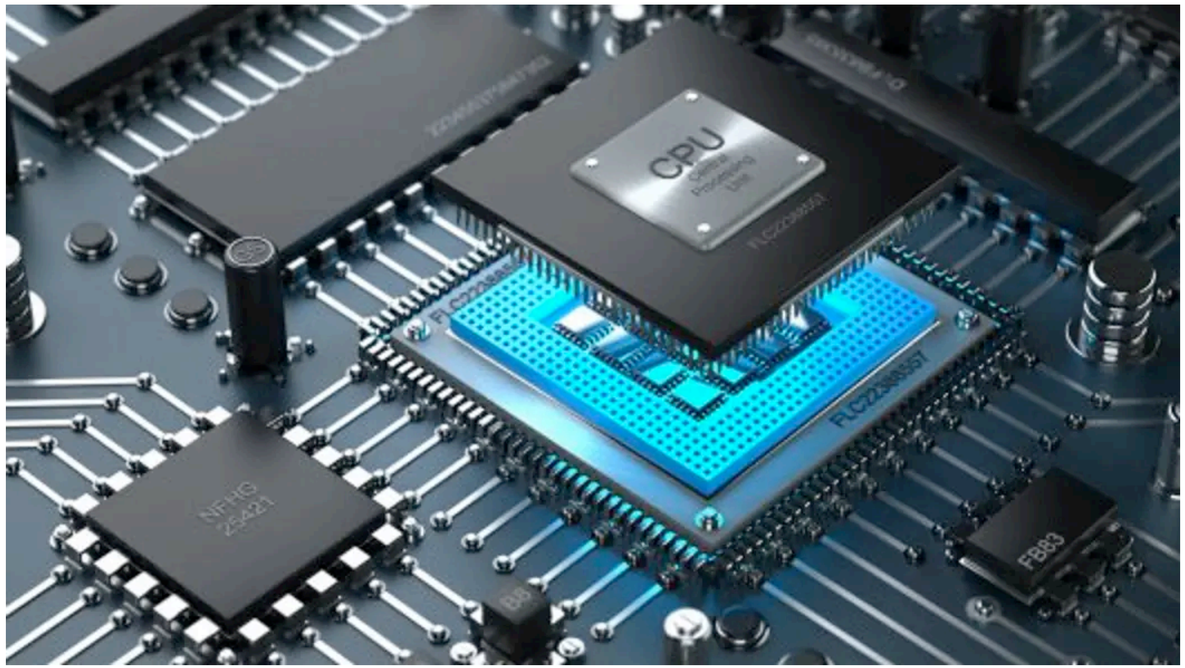


Output

Machine learning

Data

Output



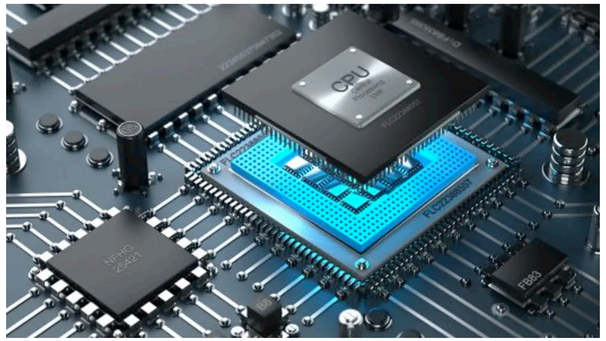
Program

Machine learning

Training

Data

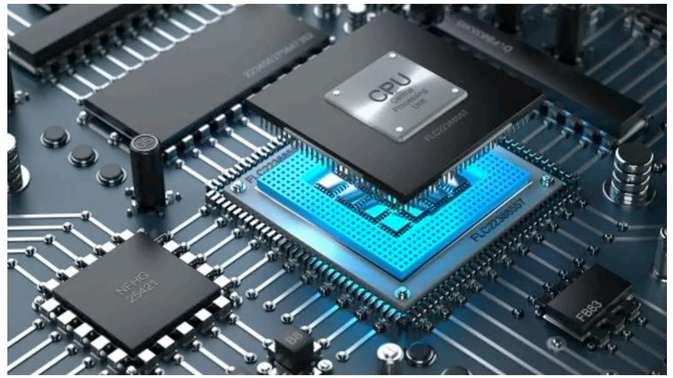
Output



Testing

Data

Program



Output

Formally

Kevin Murphy: “ …, we define machine learning as a set of methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data or to perform other types of decision making under uncertainty … ”

Tom Mitchell: “ … a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measure by P, improves with experience E, …”

A large umbrella

Intersection of computer science, statistics, neuroscience/biology, engineering, optimisation, etc.,

Statistics:

- How much data is needed?
- When can we be confident in our prediction?

Computer science:

- How fast do they run?
- What is the computational power requirement?

Formally

Intersection of computer science, statistics, neuroscience/biology, engineering, optimisation, etc.,

Statistics:

- How much data is needed?
- When can we be confident in our prediction?

Computer science:

- How fast do they run?
- What is the computational power requirement?

A very brief history

A statistical beginning

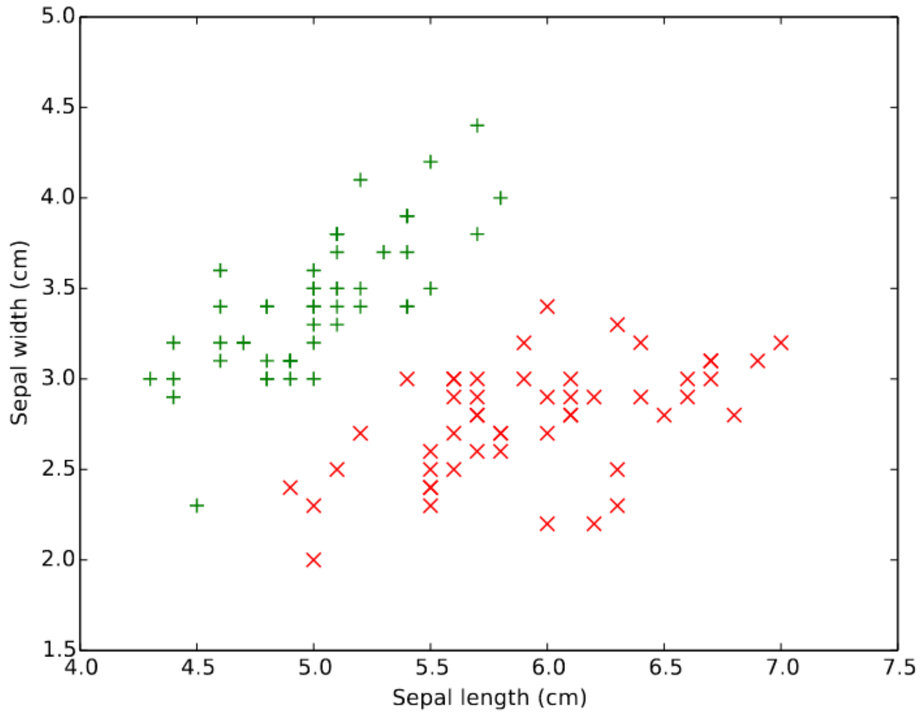
Three types of iris: setosa, versicolour, virginica

Task: classify flowers into these three

Ronald Fisher

For each flower: sepal width (feat1), sepal length (feat2), petal width (feat3), petal length (feat4)

Seperation


5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
Sepal width (cm)
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
Sepal length (cm)

A very brief history (early 1930s)

Three types of iris: setosa, versicolour, virginica

Task: classify flowers into these three

For each flower: sepal width (feat1), sepal length (feat2), petal width (feat3), petal length (feat4)

$$f(w) = w_1\text{feat1} + w_2\text{feat2} + w_3\text{feat3} + w_4\text{feat4}$$

Linear regression

Linear separation

Three types of iris: setosa, versicolour, virginica

Task: classify flowers into these three

For each flower: sepal width (feat1), sepal length (feat2), petal width (feat3), petal length (feat4)

$$f(w) = w_1\text{feat1} + w_2\text{feat2} + w_3\text{feat3} + w_4\text{feat4}$$

Linear regression

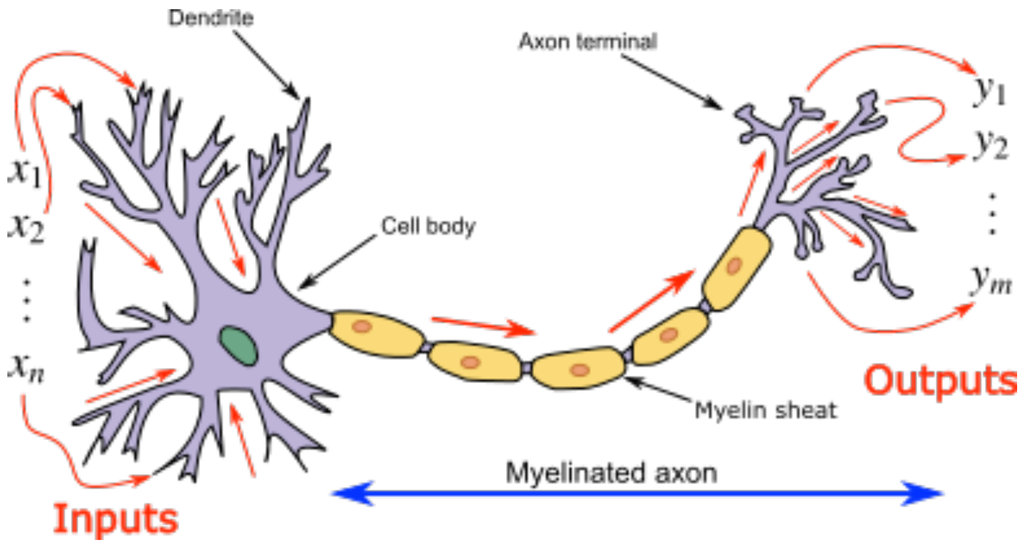
Brief history - Turing

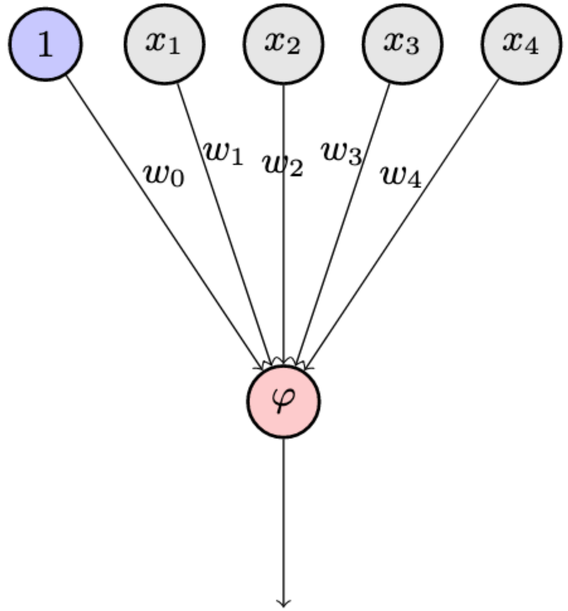
Turing Test: The Imitation Game

Learning Machines (Computing Machinery and Intelligence. Mind (1950))

“Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child’s? If this were then subjected to an appropriate course of education one would obtain the adult brain.”

A very brief history - perceptron





$$\text{sign}(w_0 + w_1 x_1 + w_2 x_2 + \cdots)$$

Brief history - Machine learning

The name: a trick to get funding

1960-1980: an era filled with expert systems (rule based algorithms); top-down systems

1980-1990: statistics & optimisation

Brief history - Deep Blue (1997)

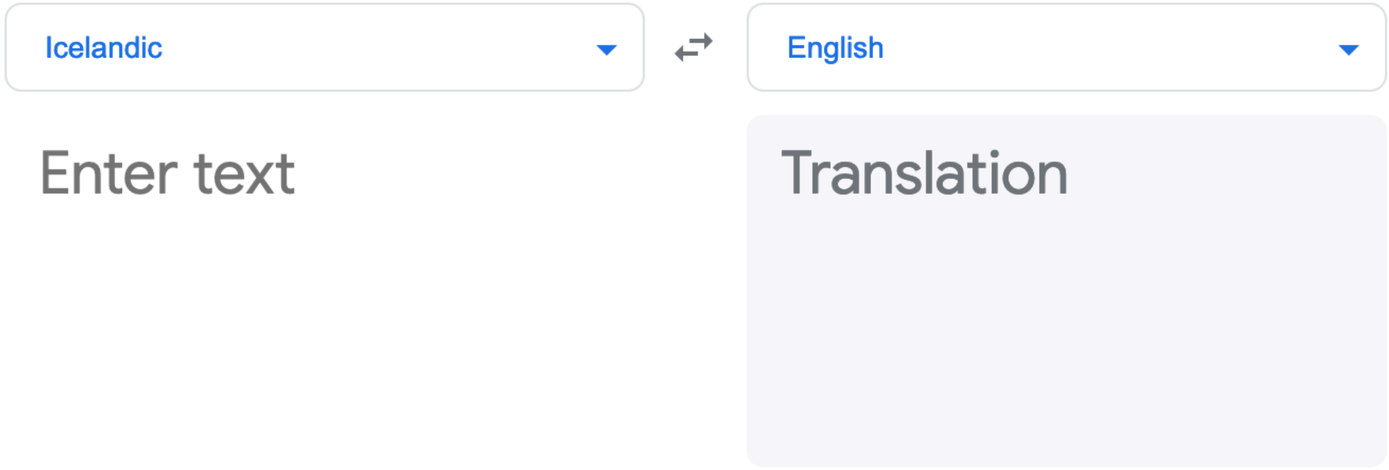
Deep Blue (IBM) wins against Gary Kasparov

The winning move is made due to a machine learning system that is trained over significant amount of data!

A breakthrough moment in popular science history

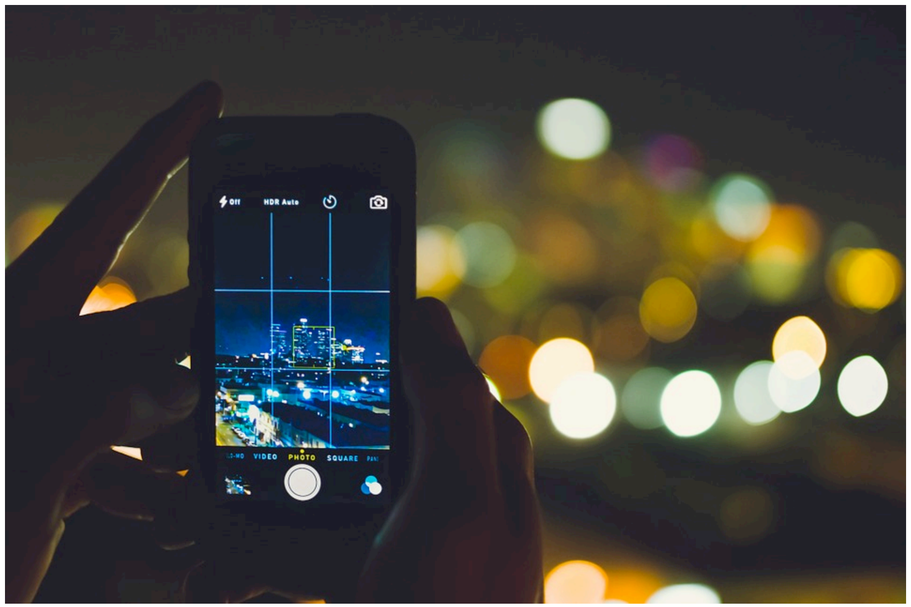
Omnipresence

Machine translation



Active question answering & reasoning

Omnipresence



Limitations

Data limits discovery

Computational efficiency and machines are becoming ever so powerful

But, it is hard to design new ML algorithms

Effective data usually saturates and large data can limit deeper analysis

This course

This course provides an introduction to machine learning, focusing on the theoretical foundations.