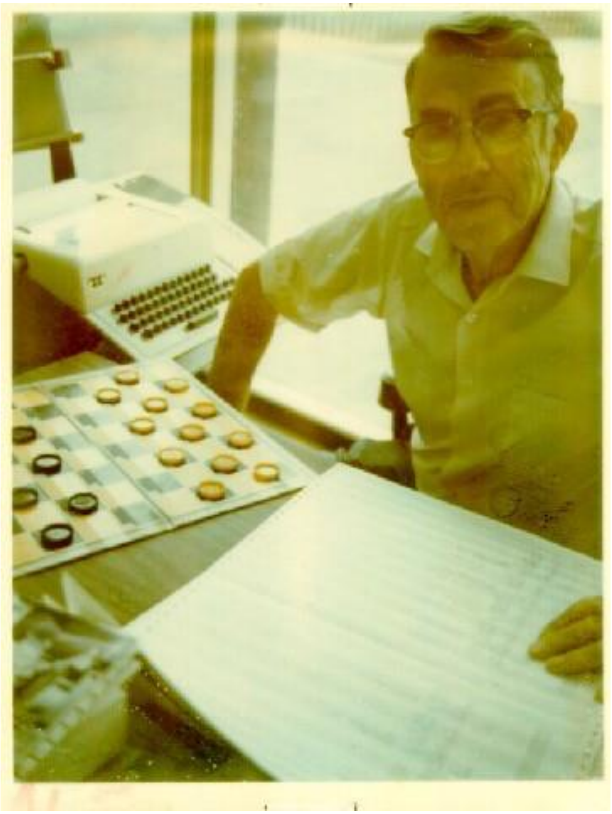
MACHINE LEARNING INTRO

Dr. Brian Mc Ginley

WHAT IS MACHINE LEARNING (ML)?

- Samuel, 1959:
  - "Field of study that gives computers the ability to learn without being explicitly programmed"
- Witten & Frank, 1999:
  - Learning is changing behaviour in a way that makes performance better in the future

Arthur Samuel, 1901-1990



WHAT IS MACHINE LEARNING (ML)?

- Algorithms that automatically improve performance through experience
  - Often this means defining a model by hand, and use data to fit its parameters
- There are problems that are difficult for humans but easy for computers
  - E.g. calculating large arithmetic problems
- And there are problems easy for humans but difficult for computers
  - E.g. recognising a picture of a person from the side
- Machine Learning often tries to leverage the things computers are good at (large arithmetic and repetitive problems) to solve things that we are usually able to do more easily.

WHY ML?

- The real world is complex – difficult to hand-craft solutions.
  - Think about how many ”if” statements would be needed!
- ML is the preferred framework for applications in many fields:
  - Computer Vision
  - Natural Language Processing
  - Speech Recognition
  - Robotics
  - . . .
- Humans can typically parse sentences, but computers do not get the context.
  - Humans are not consistent with grammar etc. but computers expect consistency.

BASIC IDEA OF ALL MACHINE LEARNING

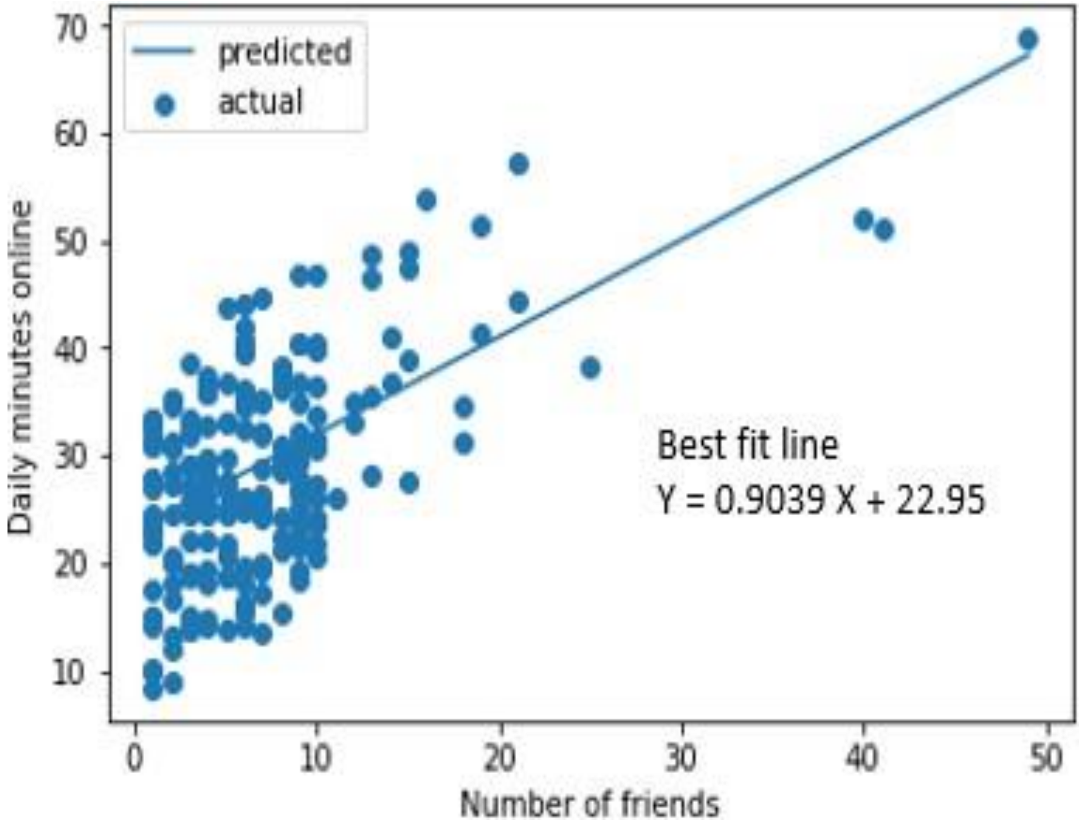
- Machine Learning is using an **algorithm** that can learn from **Data**.
- We want to create a **Model** that takes input and gives output. It is basically:

$$y = f(x)$$

- We figure out the right-hand side of the equation by gathering a lot of data and making the parameters best fit our base model.

BASIC EXAMPLE

- Take a base model of $y = w0 + w1x$, we need to out the "best" $w0$ and $w1$ parameters e.g.



- Equation of line: $y = mx + c$

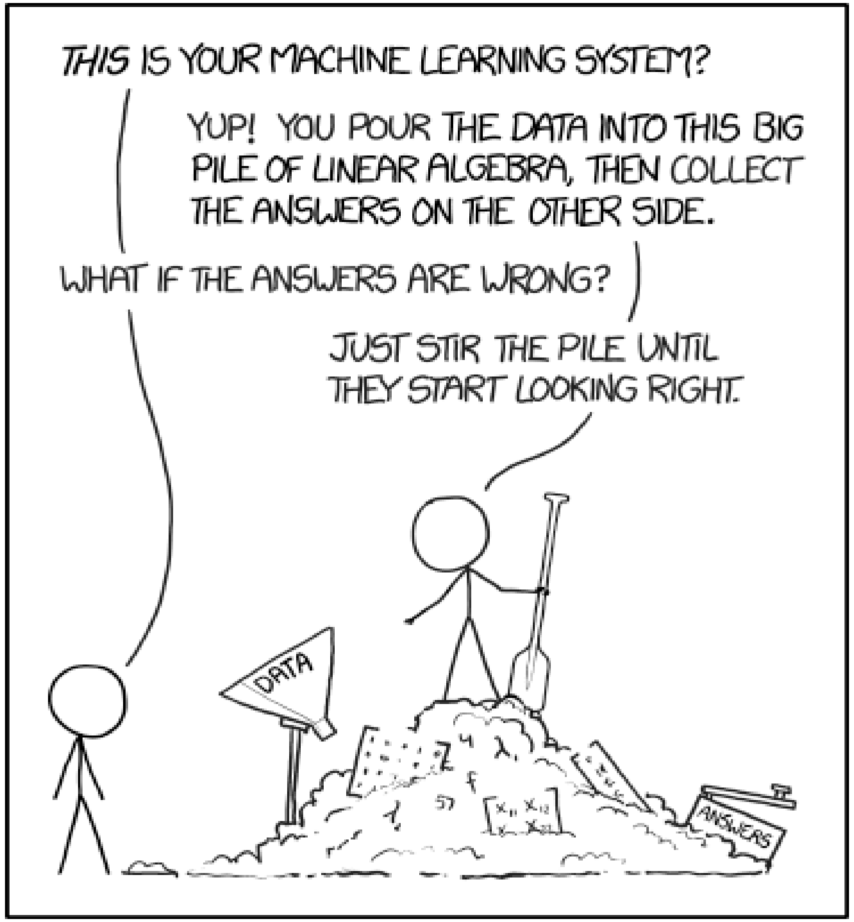
- Of course, our base model can be much more complicated, have many more parameters that need to be learnt, but the basic idea is the same.
- After we have made the model, we can then make **predictions** in the future. Above, how many daily minutes do we predict a person with 30 friends on the platform?

WHY DOES IT TAKE SO LONG?

- In the above example (one parameter, a basic line as the model), learning all the parameters is very very quick. But real world is not that quick.
- Some ML models can take hours, days, weeks or even months to be trained.
  - ChatGPT is actually 8 different models, with 220 billion parameters in each. This takes a very long time to figure out.

WHY DOES IT TAKE SO LONG?

- The training is repeatedly doing the same procedure over and over and over and over again, until we get the parameters we think are best.



GATHERING DATA

- Most of the data we want, needs to be *labelled*. This means, any data we put into the system is a pair of information

$$y, \mathbf{x}$$

  - where y is the correct answer for each x (which is the collection of all features)
- So, let's say we have a collection of images that are cats or dogs, and we want to build a model that can separate them.
- For a particular image there is a correct y - i.e. y is cat or dog and then x is all the features of the image - i.e. each x will be the pixel values.

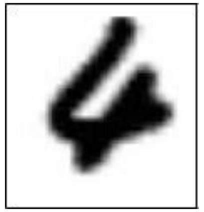
DATASET EXAMPLE

- We (the human intelligence) must do a lot of work here.
- We need lots of training data and labels associated with them.
- Example: MNIST Data set.
  - To learn to recognise handwritten letters, the algorithm must be trained on 60,000 samples, each of which must be labelled correctly - by hand.
- Sample of data set, from wikipedia.

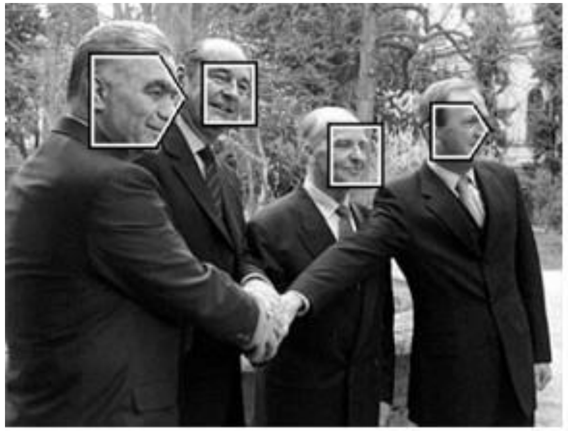
HAND-WRITTEN DIGIT RECOGNITION

- Now let's look at some potential applications and how we could try "phrasing" the problem.
- Difficult to hand-craft rules about digits.

HAND-WRITTEN DIGIT RECOGNITION

- xi =  , yi = (0, 0, 0, 0, 1, 0, 0, 0, 0, 0)
- So, the image can be represented by the value of each pixel. Each pixel colour is called a *feature*, they are independent variables..
  - Represent input image as a vector $x_i = \mathrm{R}^{784}$
  - Suppose we have a target vector t_i
    - This is **supervised learning**
    - Discrete, finite label set: a **classification** problem.
  - Given a **training set** $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, the learning problem is to construct a "good" function *y=f(x)* from these.
    - $f : \mathrm{R}^{784} \rightarrow \mathrm{R}^{10}$
- Some algorithms will not exactly return (0, 0, 0, 1, . . . etc. but may return probabilities like (0.01, 0.03, 0.04, , 0.01, 0.88 . . . so the image is most *likely* a 4.

FACE DETECTION

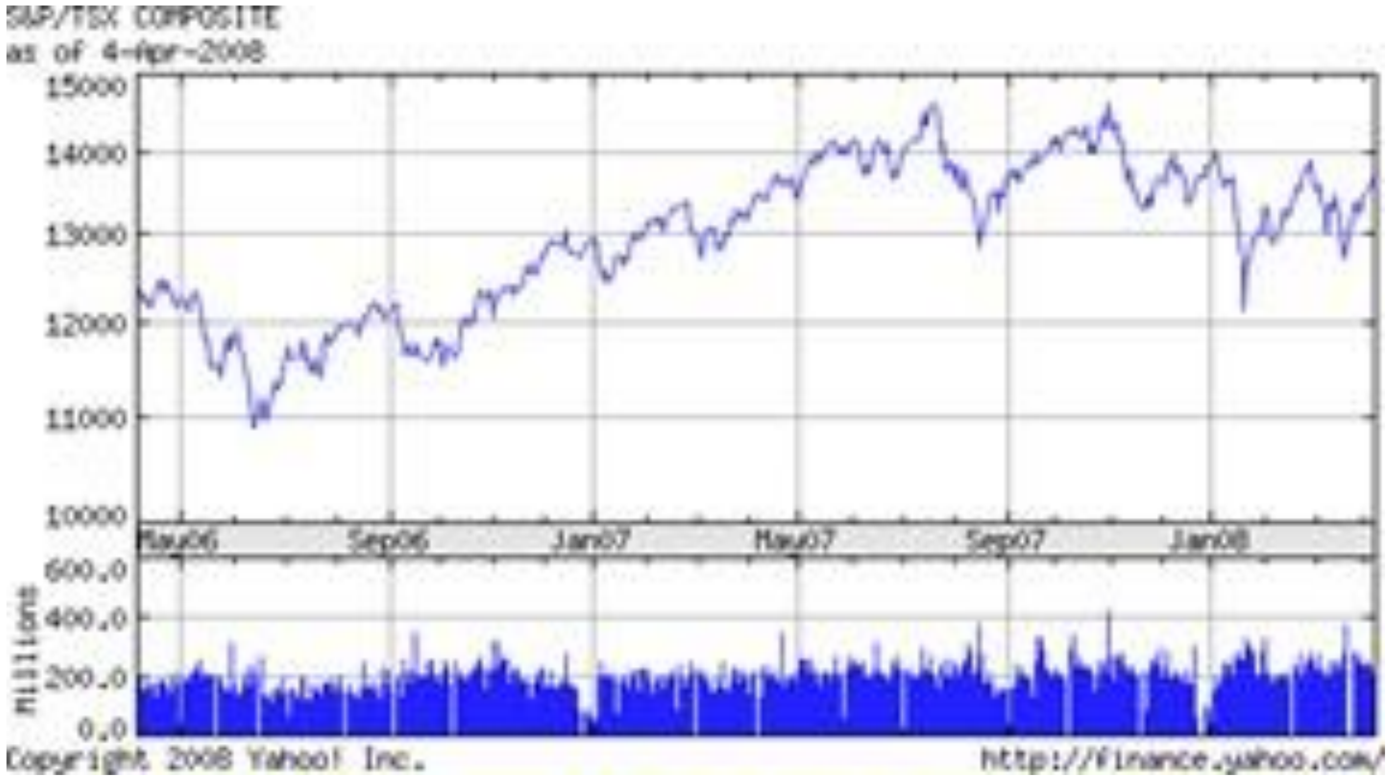


Schneiderman and Kanade, IJCV 2002

- **Classification** problem.
- $t_i \in \{0, 1, 2\}$, non-face, frontal face, profile face.
- Of course, this can be expanded into the face of a particular person so the possible t_i set could be large.
  - i.e: t_i is the set of ALL possible results.
- We **map/transform** the image to a particular t_i

STOCK PRICE PREDICTION

- Problems in which t_i is continuous are called **regression**
- E.g. t_i is stock price, x_i contains company profit, debt, cash flow, gross sales, number of spam emails sent, . . .



USERS OF ML

amazon

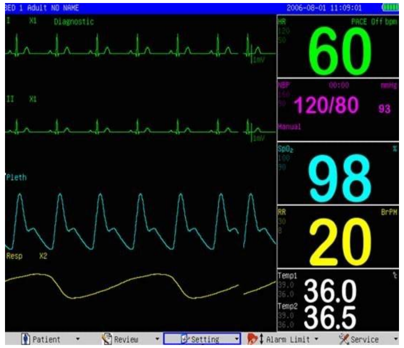
NETFLIX

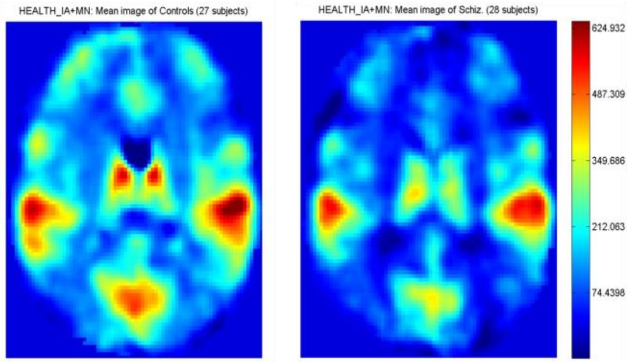
Microsoft

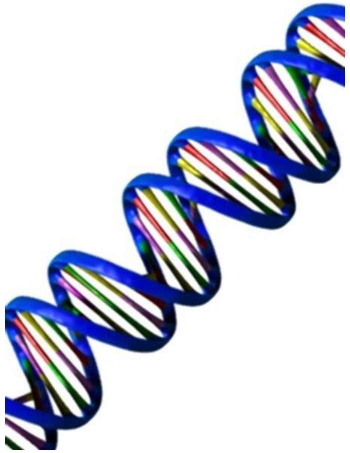
YAHOO!

facebook

Google







HIGH-PROFILE EXAMPLES …

Forbes · New Posts · +4 posts this hour · Most Popular · Most Disliked Athletes

Kashmir Hill, Forbes Staff

Welcome to The Not-So Private Parts where technology & privacy collide

+ Follow (1,178)

TECH | 2/16/2012 @ 11:02AM | 1,930,513 views

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

HIGH-PROFILE EXAMPLES …

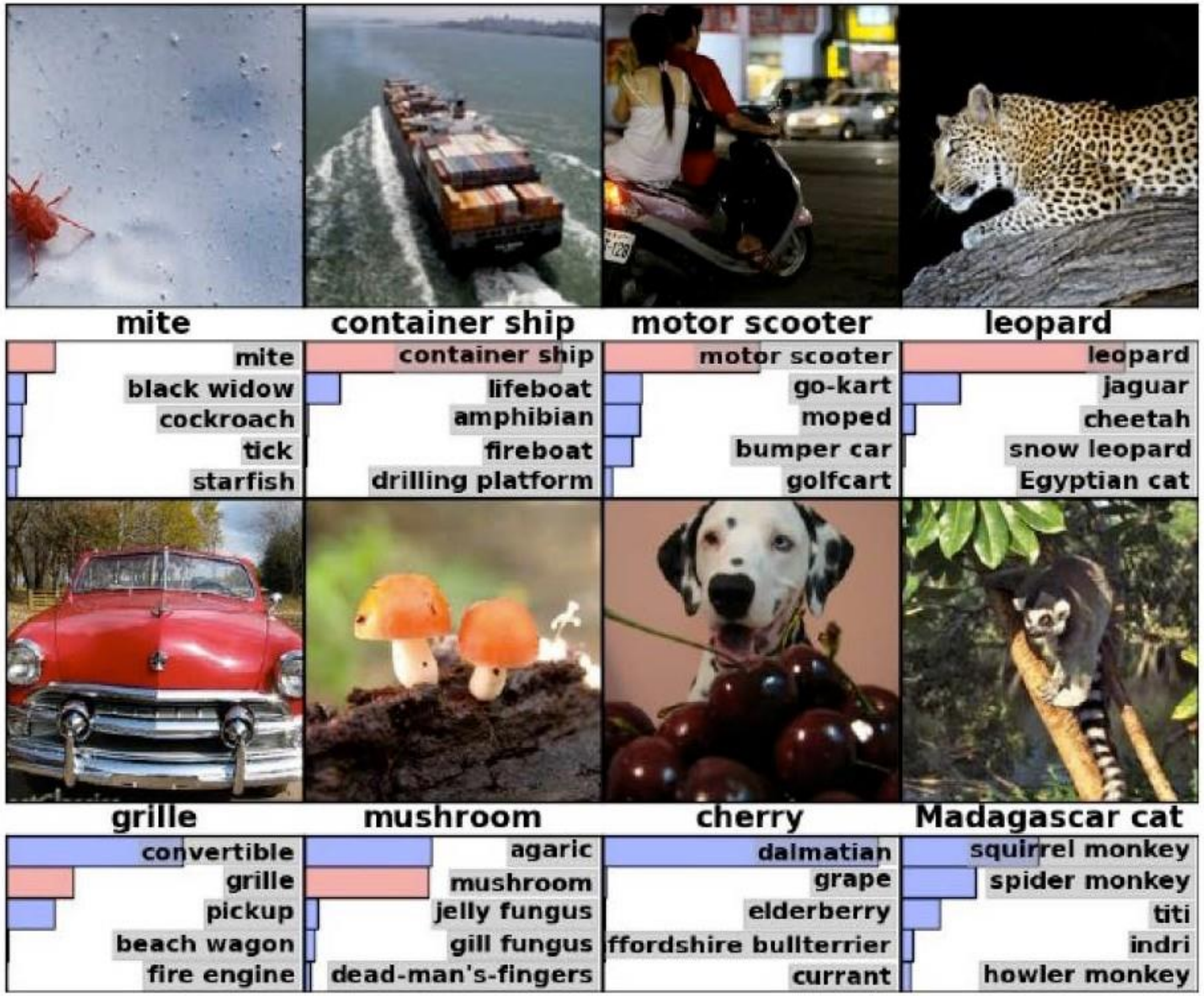
How Netflix is turning viewers into puppets

"House of Cards" gives viewers exactly what Big Data says we want. This won't end well

BY ANDREW LEONARD

HIGH-PROFILE EXAMPLES …

- Deep Learning for Object Recognition: Hinton & colleagues, NIPS 2012



HIGH-PROFILE EXAMPLES

- Learns to play the game Go, just by playing games against itself
- Starting from completely random play https://deepmind.com/blog/alphago-zero-learning-scratch/

HIGH-PROFILE EXAMPLES …

Obvious' "Portrait of Edmond Belamy" exceeded expectations at Thursday's sale (Courtesy of Obvious)

SMARTNEWS *Keeping you current*

Christie's Is First to Sell Art Made by Artificial Intelligence, But What Does That Mean?

Paris-based art collective Obvious' 'Portrait of Edmond Belamy' sold for $432,500, nearly 45 times its initial estimate

HIGH-PROFILE EXAMPLES …

- Dall-E
- Chat GPT
- Gemini
- Midjourney
- Etc…

SUPERVISED LEARNING

- Techniques where we have training examples where we **know** the correct result (the y values). Different types of supervised learning are:
  - Classification
  - Regression
- Some algorithms:
  - Linear/Logistic Regression
  - Naive Bayes
  - k-Nearest Neighbours
  - Support Vector Machines
  - Neural Networks
  - Decision Trees
  - Random Forests

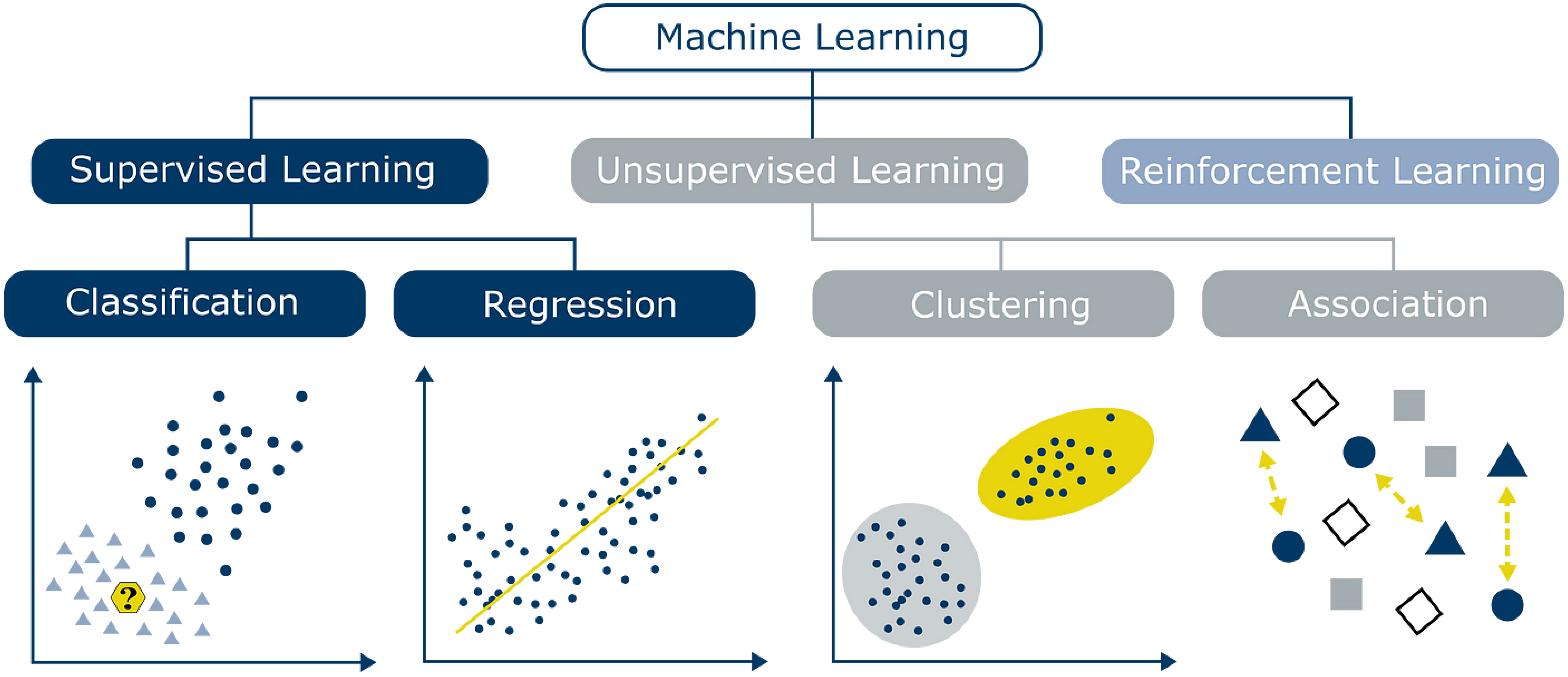
UNSUPERVISED LEARNING

- Techniques where there is no "right" answer known, but the algorithm tries to find some **structure/patterns** in the data.
  - Principal Component Analysis (Dimensionality Reduction)
  - k-mean Clustering
  - PageRank (Google)
- Unsupervised learning techniques will group things together but we do not necessarily know what the groups are.
- Will do less unsupervised learning in this module

OTHER TYPES

- There are other types of Machine Learning, e.g. Reinforcement Learning.
  - (outside the scope of this module).

TYPES OF MACHINE LEARNING



MACHINE LEARNING

- Often referred to as narrow AI.
- We are only concerned with a specific task.
- So the goal of machine learning is not to develop general intelligence or a universal learning algorithm.
- Machine learning seeks an algorithm that learns a particular task well
- And should result in a system that probably carries out the task correctly on most occasions.

WHAT DO WE NEED, TO DO MACHINE LEARNING?

- The Task - define or limit the task
- The Experience - This is the data, more data = more experience
- The Performance measure - We need a good way to measure this.
- The Learning Algorithm - The recipe by which we will improve our performance.
- The Intelligence - the Network - the brain.