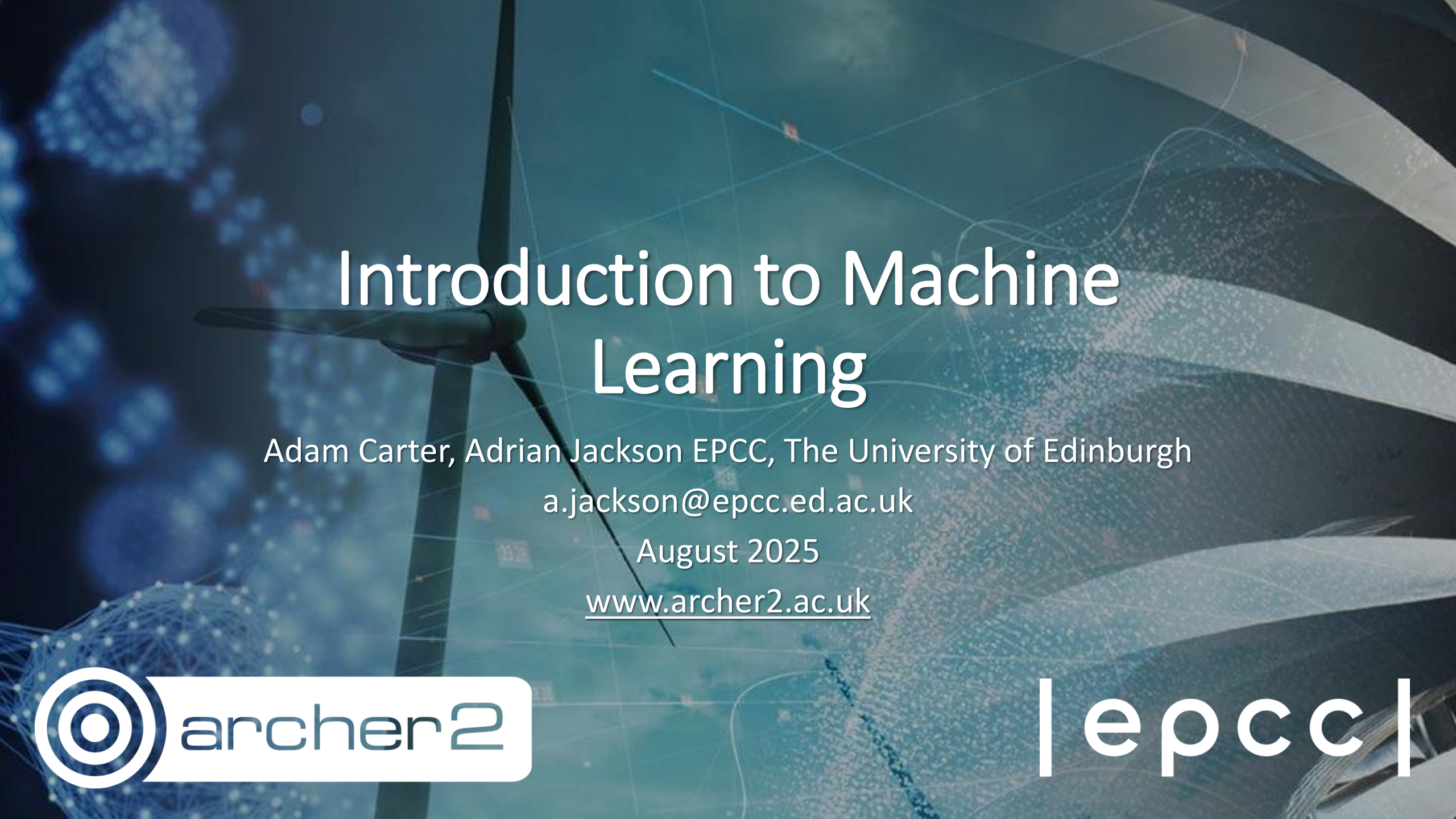


Introduction to Machine Learning

Adam Carter, Adrian Jackson EPCC, The University of Edinburgh
a.jackson@epcc.ed.ac.uk

August 2025

www.archer2.ac.uk

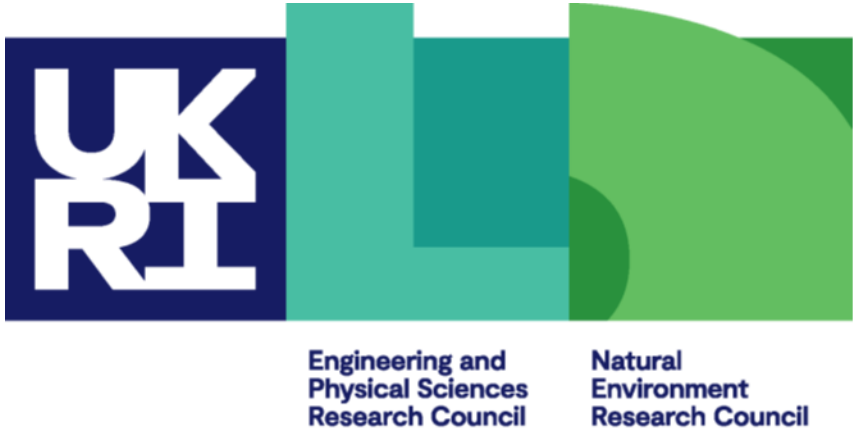
archer2

epcc

Reusing this material

Partners

epcc



epcc

THE UNIVERSITY of EDINBURGH

Hewlett Packard Enterprise

What is Machine Learning?

- Getting a computer to do something, without explaining to the computer exactly how to do it
  - "Learn from these examples" – **Supervised Learning**
    - e.g., given many examples of pictures of cats and dogs with the label "cat" or "dog", learn how to identify whether a new, unlabeled image is a cat or a dog
  - "Find patterns in this data" – **Unsupervised Learning**
    - e.g., given information about all your customers, can you group your customers into sets that are similar in some respect, perhaps so you could treat them as a group in a marketing campaign?
  - "Try to achieve this goal by learning from your experience" – **Reinforcement Learning**
    - e.g. get a robot to learn how to climb over a box by trying different movements of its legs to see which ones are most effective

Why is it so powerful?

- You can use (basically) the same algorithms for tasks in multiple domains, e.g.,
  - identifying genes that are similar to each other and identifying customers that are similar to each other
  - predicting which of your customers are likely to unsubscribe from your service and predicting the upcoming weather from current rain radar images
  - identifying tumours in medical images and identifying words spoken to a voice assistant
  - Classifying images of cats and dogs, and classifying email as spam or not spam
  - Teaching a robot to walk, and predicting the next word you will type in a text message
  - Responding to a prompt in ChatGPT and creating an image from a text prompt

Some other definitions of machine learning

- “the field of study that gives computers the ability to learn without explicitly being programmed.” – Arthur Samuel, IBM (~1959)
  - This is possibly a paraphrase, but it is widely accepted that the term was coined by Arthur Samuel
- “Machine learning (ML) is a field of study in artificial intelligence concerned with the development and study of statistical algorithms that can effectively generalize and thus perform tasks without explicit instructions” – Wikipedia (accessed November, 2023)
- “Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.” – IBM
- “it's the science of getting computers to learn without being explicitly programmed” – Andrew Ng, Stanford/Coursera/…

NLP/LLMs is not (just) ML and vice versa

epcc

- NLP/LLMs requires
  - Semantic analysis
  - Syntactic analysis
  - Combining both systems to process and respond to language questions
- Leverages multiple ML approaches and other systems to deliver NLP functionality
  - NLP systems are complex data processing, conversion, and optimisation processes
  - Couple together the language models with image generators, code generators, etc…, which are ML systems in their own right

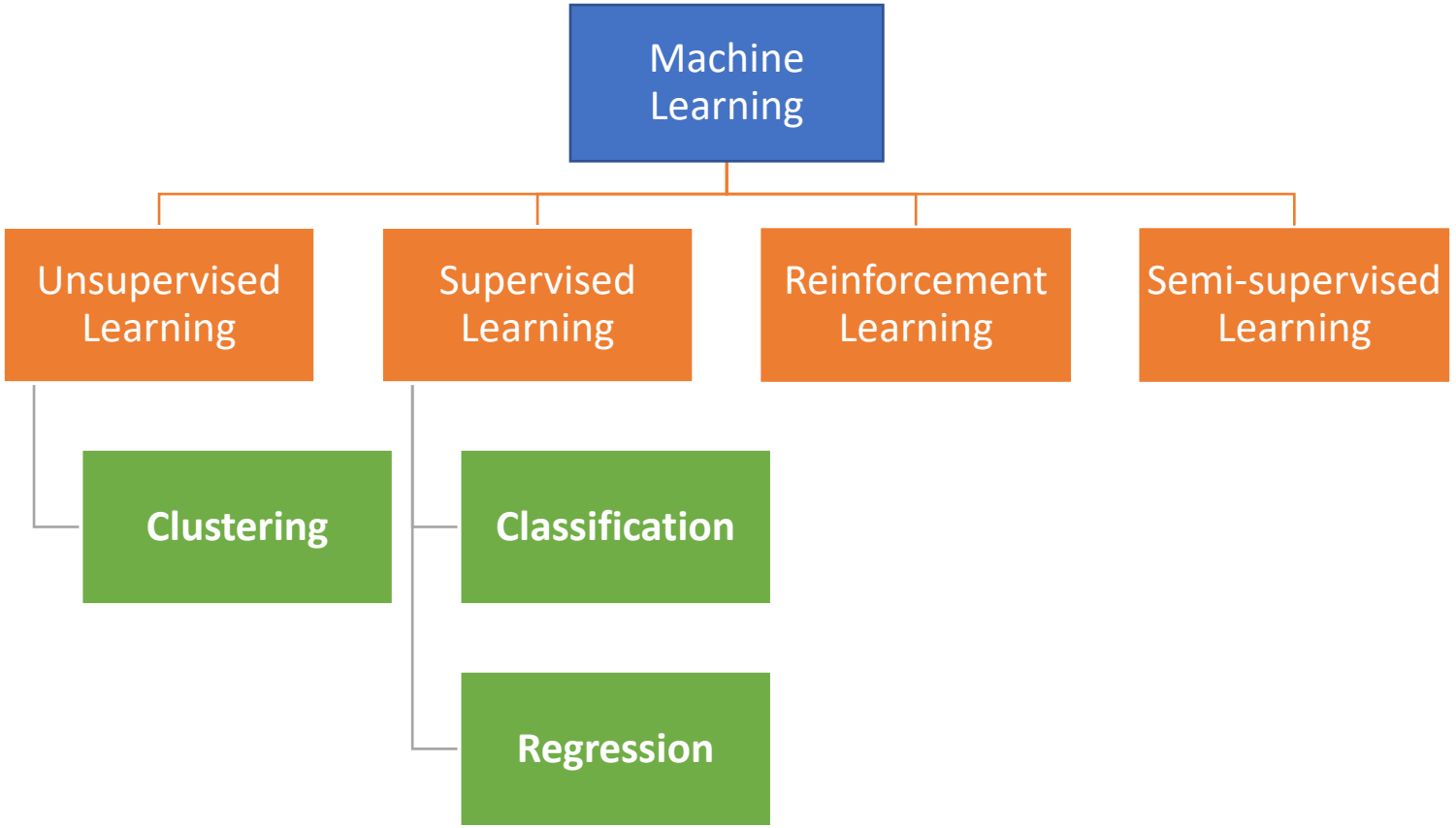
What do the machines learn?

epcc

- Central to Machine Learning is the idea of a **model**
- A model is learned from data using a **machine learning algorithm** and the model can then be used to make predictions/classifications from new, previously unseen data (in the case of supervised learning) or to cluster/group the data (in the case of unsupervised learning) or to decide on a course of action in an observed environment (reinforcement learning)
- Generally speaking, with ML, machines do *not* learn facts and they do *not* learn to reason, they learn models/patterns
  - No reason why this isn't possible in future, but that's not how ML currently works, even in ChatGPT!

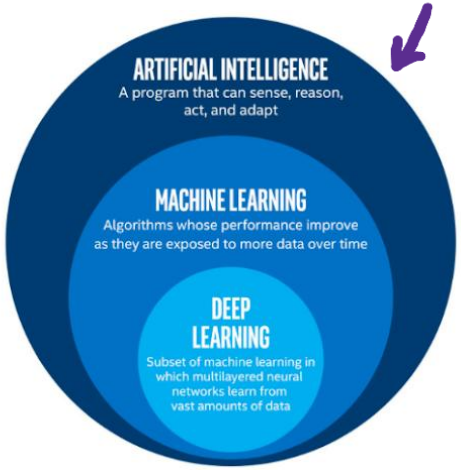
Types of Machine Learning

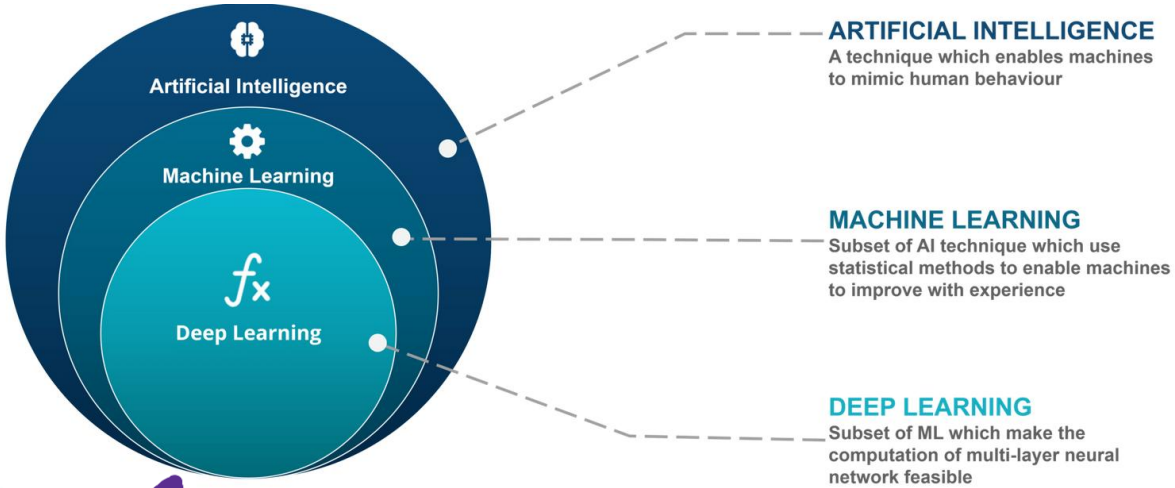
- Unsupervised Learning
  - Clustering
    - kMeans
    - DBSCAN
    - Hierarchical
- Supervised Learning
  - Regression
    - Linear Regression
  - Classification
    - kNN
    - Naïve Bayes
    - Logistic Regression
    - Neural Networks



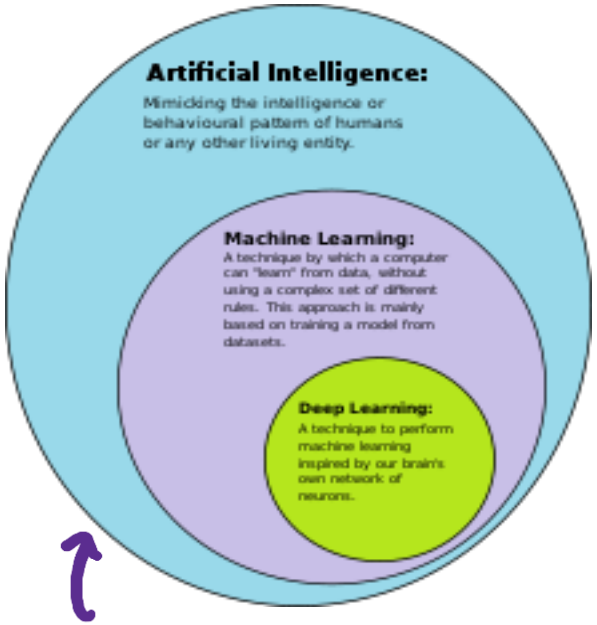
ML, Deep Learning & AI

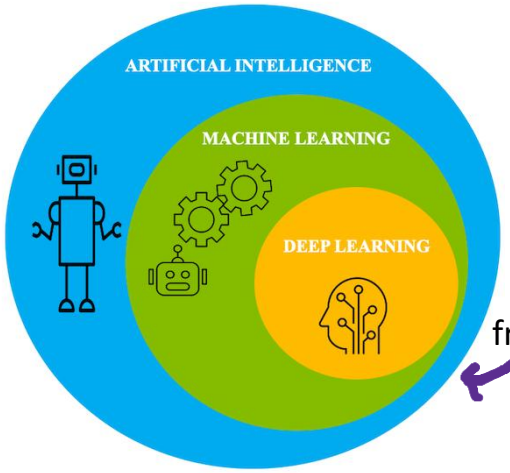
from https://builtin.com/machine-learning/what-is-deep-learning





from https://www.edureka.co/blog/ai-vs-machine-learning-vs-deep-learning/

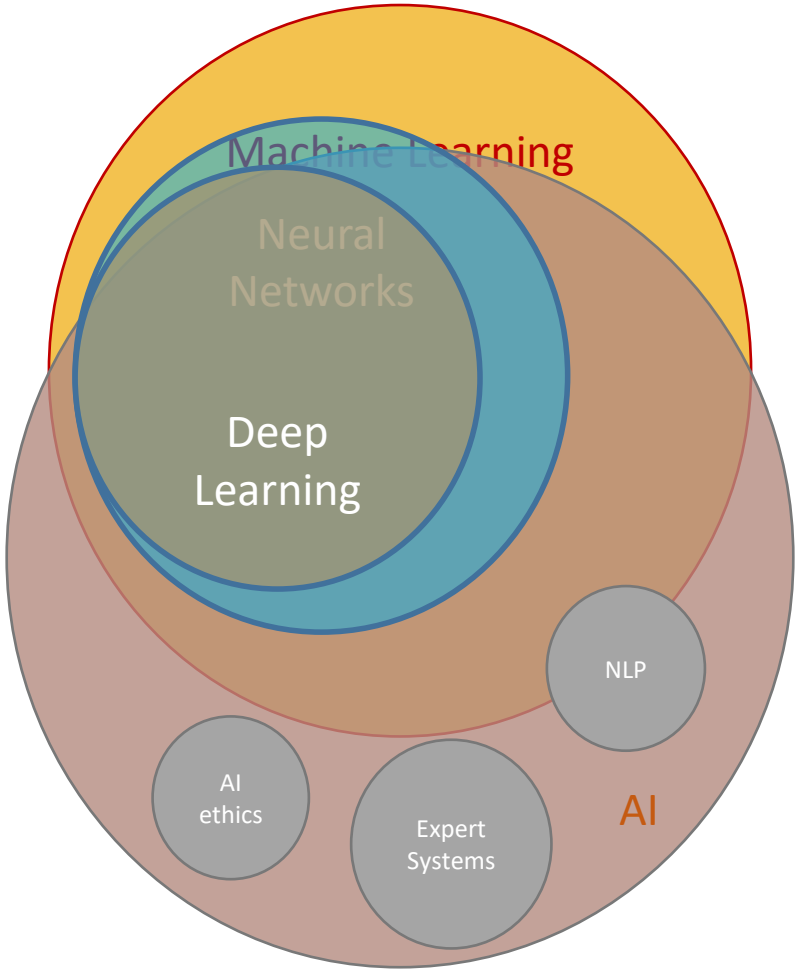




from https://developer.ibm.com/articles/an-introduction-to-deep-learning/

from https://commons.wikimedia.org/wiki/File:AI-ML-DL.svg

…or maybe…


Machine Learning
Neural Networks
Deep Learning
NLP
AI ethics
Expert Systems
AI

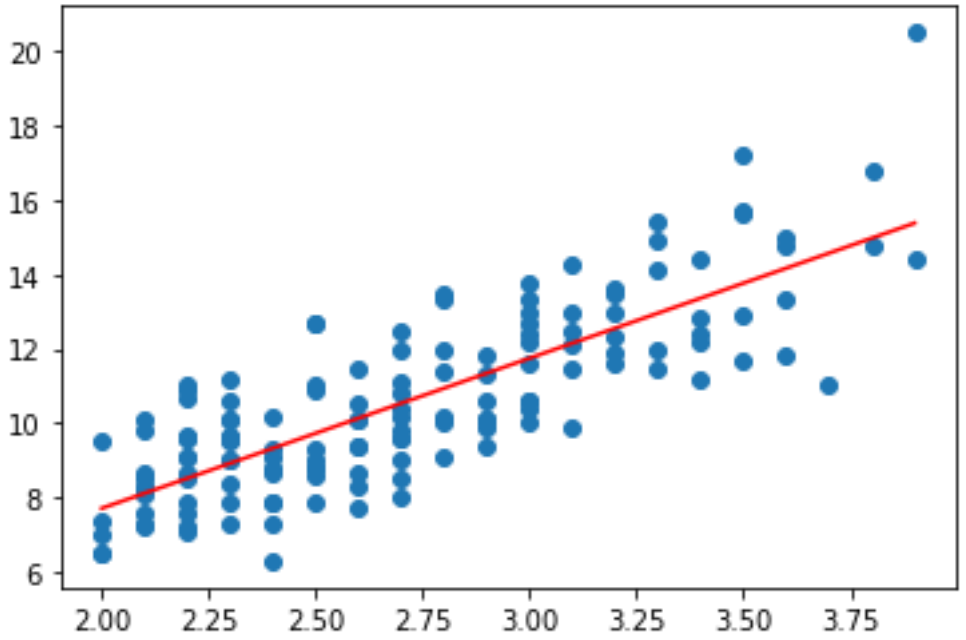
Machine Learning

epcc

- Implementing these approaches requires some software and frameworks
  - How do you “learn” data patterns or formats or features
- Different approaches
  - Currently dominate ones are deep neural networks (DNNs)
    - Large number of parameters that can be updated automatically
    - Loss function that represents the aggregate of these and the outputs of the network
    - Update the parameters based on the optimisation of the loss function
    - Converge to minimal loss and small(ish) weights
  - Other approaches also available
    - Graph Neural Networks: The network that is built encodes some spatial locations (i.e. points on earth, place in a social network graph, etc…

Linear Regression as an example of ML

- Given a new value of data point with known *x* but unknown *y* can you predict *y*?
- The **model** here is the red line
- The model is **learned** using a process which finds the best fit for the existing datapoints: the "**training data**"
- We sometimes say we **fit** the model to the training data, or that we **train** the model using the training data
- Key point: In ML, like here, we often decide on the shape of the model (here a straight line) but we let the machine learning algorithm find what the *best* straight line is...



Some important ideas

- Models are only as good as the data they're trained on
  - Possibility for bias, changing underlying conditions
  - Rare events can be harder to model (less training data)
- Machine learning models are often opaque:
  - It can be difficult to "understand" why they make a certain prediction
- There are different ways of deciding if a model is "good"
  - Would you prefer your AI poisonous mushroom predictor to be 98% accurate, or would you prefer that it was only accurate 95% of the time but erred on the "safe" side when unsure?
- Training models is often computationally expensive
  - Although modern computing resources make this more accessible than ever
  - Training is usually much more resource intensive than prediction/inference

Training models

- The process of training models is often iterative
  - That is, machine learning algorithms often repeatedly use the training data to get better and better models (for some definition of better)
  - Increased training time can lead to better models
- Sometimes in the process of training models, you can build a model which suffers from **overfitting**
  - The model represents the training data very well but does not generalise well to new, unseen data
- Once models are trained, then they can be used to react to or generate new data
  - Inference phase has different computational requirements to the training phase
  - Generally much faster than network training
    - No iterative processes, no differentiation of functions, etc…
    - Simply pass the data through the network with fixed weights, requiring a (relatively) small number of matrix-vector operations