HeiCAD
Heine Center for Artificial Intelligence and Data Science

HeiCADLearn

Practical Introduction to AI and Data Science for Doctoral Researchers in Medicine and Biology

Dr Joana Sarah Grah

Joana.Grah@hhu.de

heicad.hhu.de

Accompanying GitHub repository for this course:
https://github.com/JoanaGrah/HeiCADLearn/Intro_MedBio_July2021

@heicad_hhu

@HeiCAD.hhu

@HeiCAD_HHU

Day 1

1. Introduction
2. What is AI?
3. What can we already do using AI?
   What are the current limitations of AI?
   What are ethical/legal concerns?

BREAK 

4. A short introduction
   to data science in Python

Day 2

1. From linear regression to neural networks

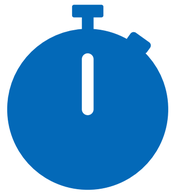
BREAK 

2. AI in practice:
   Deep learning for biology and medicine

Day 3

Deep learning example

Introduction



30 sec

- Name
- Department / Research Area
- Why did you sign up for this workshop and what do you expect from it?

Learning Objectives

At the end of the workshop you should

- be able to **classify the importance** of AI technologies
- be able to explain **basic terms and methods**
- be able to assess whether the use of AI is **necessary / useful**
- be able to **analyse your data** as research questions and problems arise **and apply basic data science and AI tools**
- have the ability to **differentiate** where “traditional” mathematical / statistical methods should be preferred for analysis and in which cases (and if there is sufficient qualitative / quantitative data) AI methods can be used

What is AI?

Discussion

miro

- How would you define AI?
- Where do you already encounter AI in your everyday life?
- What preconceptions and myths do you associate with AI?
- How is AI depicted in the literature, on TV, in the media?

What is AI?

https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

"Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.

AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."

What is AI?

https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

"A **general AI** system is intended to be a system that can perform most activities that humans can do. **Narrow AI** systems are instead systems that can perform one or few specific tasks. Currently deployed AI systems are examples of narrow AI.
In the early days of AI, researchers used a different terminology (weak and strong AI).
There are still many open ethical, scientific and technological challenges to build the capabilities that would be needed to achieve general AI, such as common sense reasoning, self-awareness, and the ability of the machine to define its own purpose."

What is AI?

Yann LeCun (VP and Chief AI Scientist, Facebook):
It's clear to him that AI still has a long, long way to go before it approaches anything near the intelligence of a baby, or even an animal. Oh, and if you don't mind, he'd really like it if we all stopped using Terminator pictures on AI articles.

https://www.theverge.com/2017/10/26/16552056/a-intelligence-terminator-facebook-yann-lecun-interview

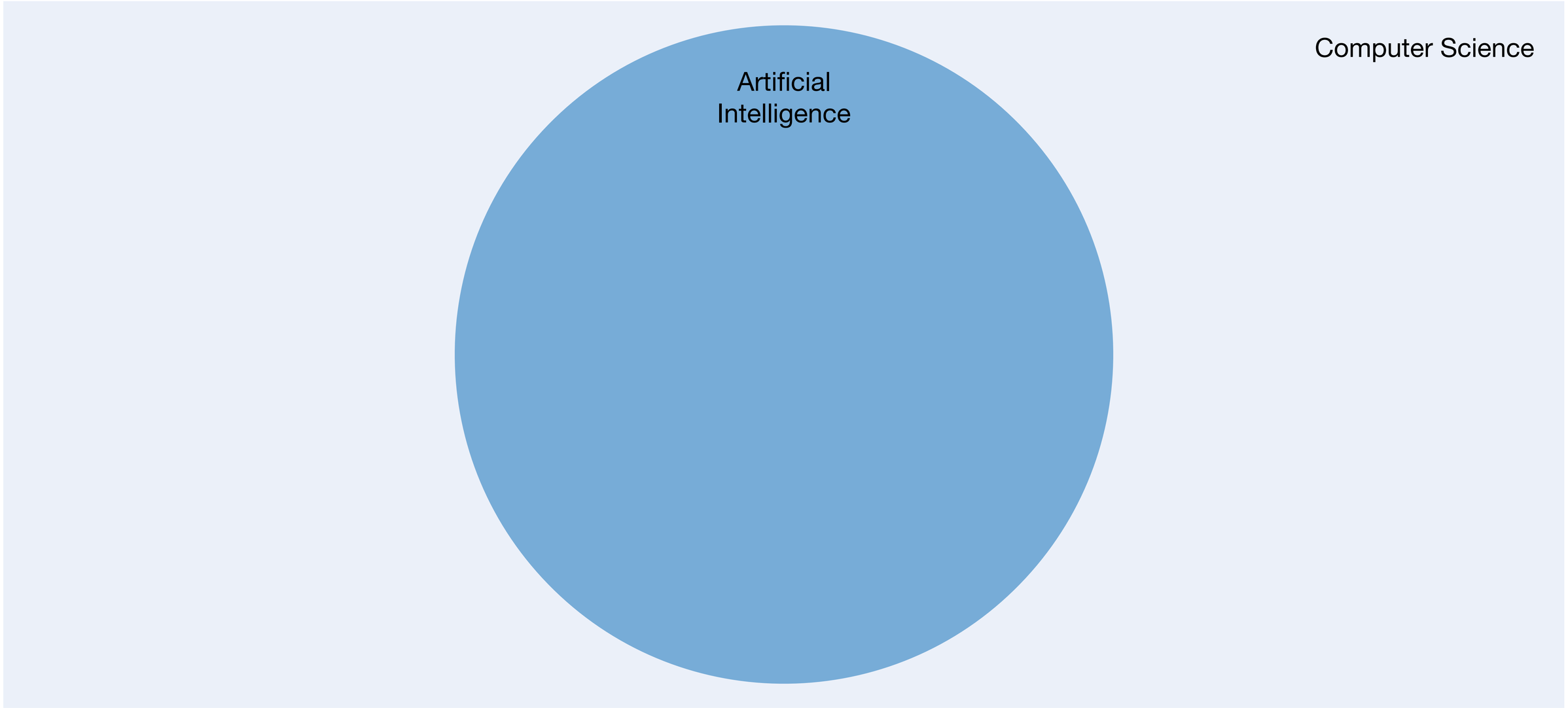
Rodney Brooks (Professor of Robotics (emeritus) at MIT):
"None of our AI systems have any intentionality. They're all very narrow tools: solve this sub-problem, solve this sub-problem. They don't have the intentionality, really, even of an insect, let a lone a small mammal. People hear AI and they think they're these smart machines with wants and desires and capabilities. They don't have any of that. I wish they did, but that's not what we've been able to do in the last, really, 55 years of AI research."

https://www.roboticsbusinessreview.com/rbr/rodney_brooks_ai_systems_very_narrow_tools/

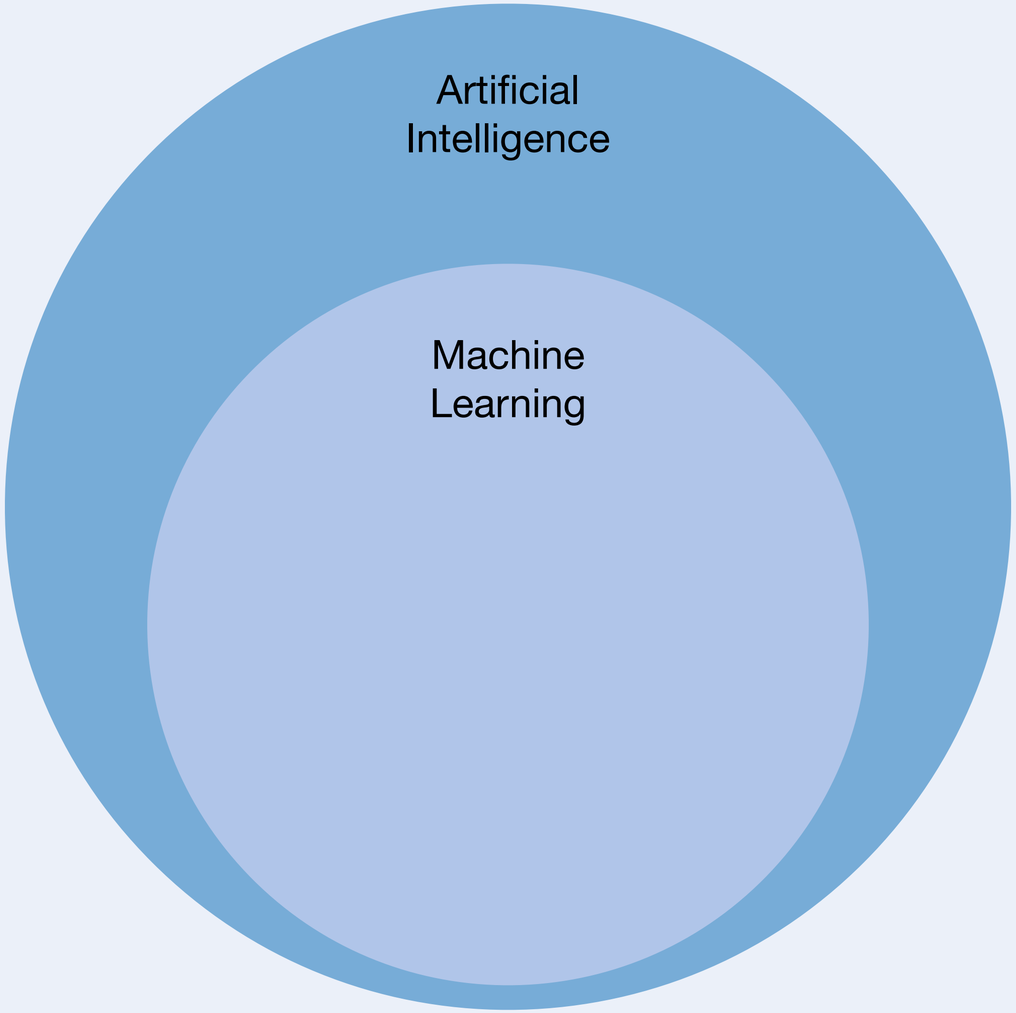
Stephen Hawking (Theoretical Physics, Cambridge)
"Success in creating AI could be the biggest event in the history of our civilisation. But it could also be the last – unless we learn how to avoid the risks. Alongside the benefits, AI will also bring dangers like powerful autonomous weapons or new ways for the few to oppress the many."
"We cannot predict what we might achieve when our own minds are amplified by AI. Perhaps with the tools of this new technological revolution, we will be able to undo some of the damage done to the natural world by the last one – industrialisation."

https://www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of

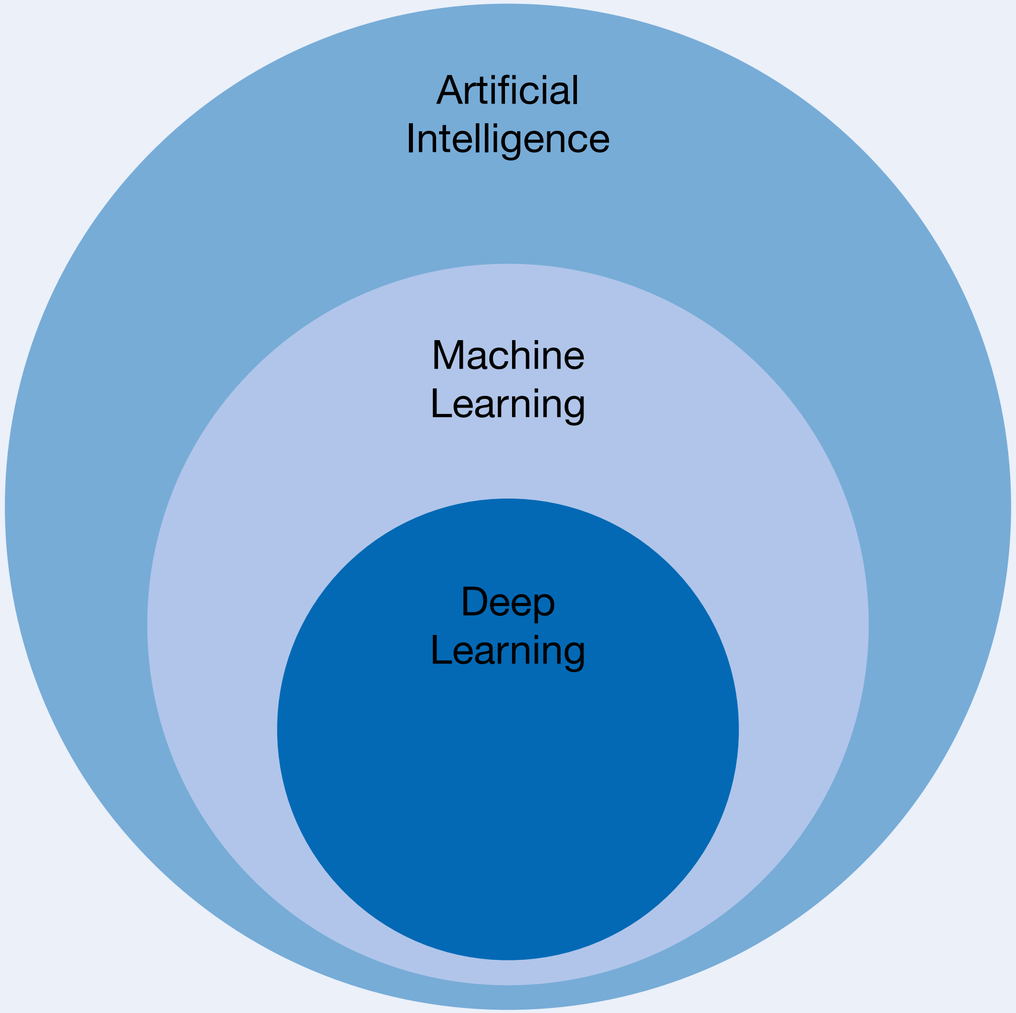
What is AI?



What is AI?


Computer Science
Artificial
Intelligence
Machine
Learning

What is AI?


Computer Science
Artificial Intelligence
Machine Learning
Deep Learning

What is AI?

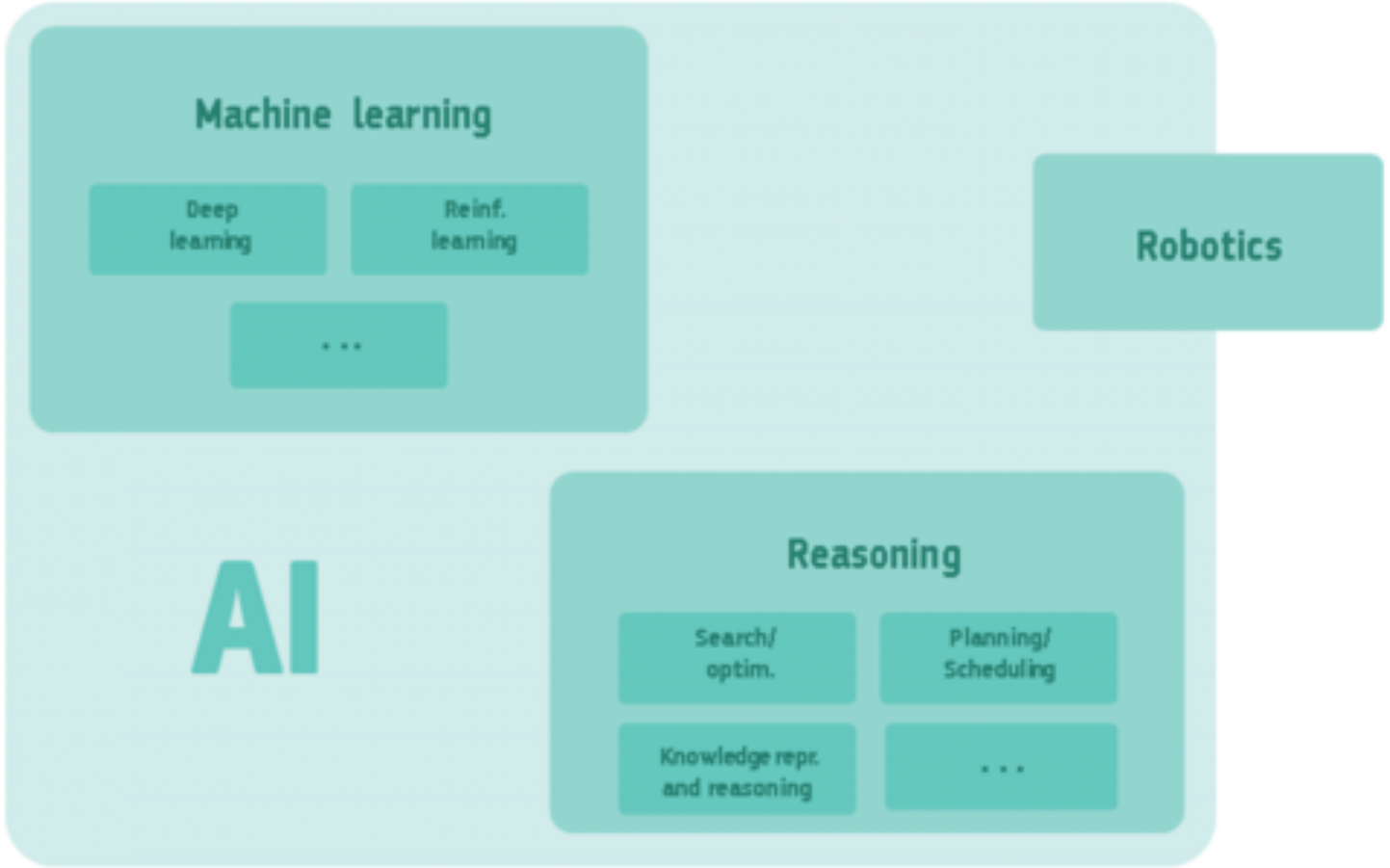


Figure 2: A simplified overview of AI's sub-disciplines and their relationship.
Both machine learning and reasoning include many other techniques, and robotics includes techniques that are outside AI. The whole of AI falls within the computer science discipline.

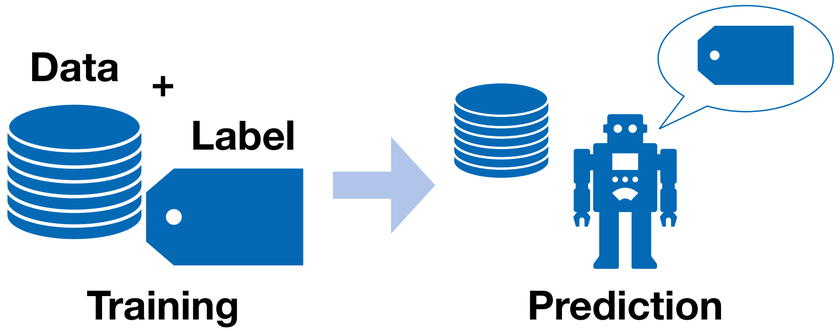
https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

What is Machine Learning?

The process of computers changing the way they carry out tasks by **learning from new data**, without a human being needing to give instructions in the form of a program
(Cambridge Dictionary)

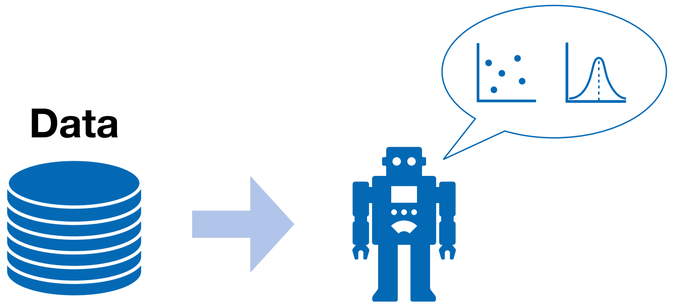
Machine learning can be broadly defined as **computational methods using experience to improve performance or to make accurate predictions**. Here, *experience* refers to the past information available to the learner, which typically takes the form of electronic data collected and made available for analysis. This data could be in the form of digitised human-labeled training sets, or other types of information obtained via interaction with the environment. In all cases, its quality and size are crucial to the success of the predictions made by the learner.
(Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.)

Supervised Learning



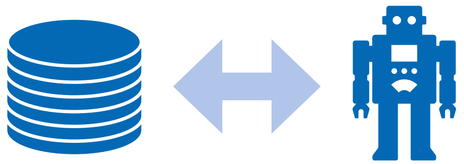
Maximise prediction accuracy

Unsupervised Learning



Find structure in (unlabelled) data

Reinforcement Learning

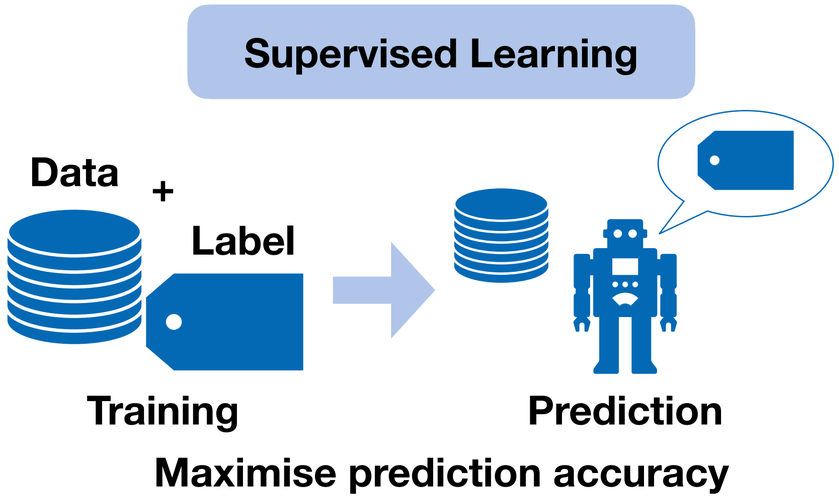


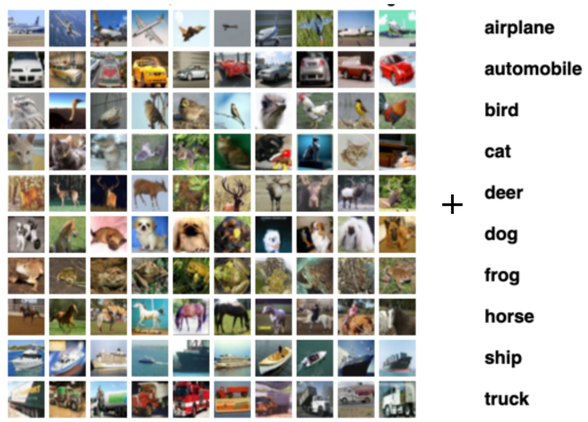
Act to maximise rewards

What is Machine Learning?

The process of computers changing the way they carry out tasks by **learning from new data**, without a human being needing to give instructions in the form of a program
(Cambridge Dictionary)

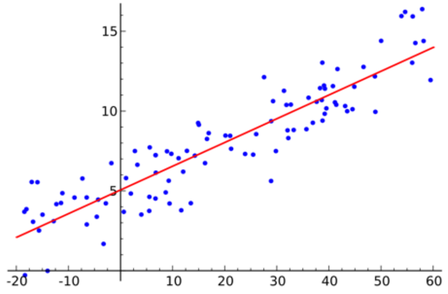
Machine learning can be broadly defined as **computational methods using experience to improve performance or to make accurate predictions**. Here, *experience* refers to the past information available to the learner, which typically takes the form of electronic data collected and made available for analysis. This data could be in the form of digitised human-labeled training sets, or other types of information obtained via interaction with the environment. In all cases, its quality and size are crucial to the success of the predictions made by the learner.
(Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.)





https://www.cs.toronto.edu/~kriz/cifar.html

Josef Steppan, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons

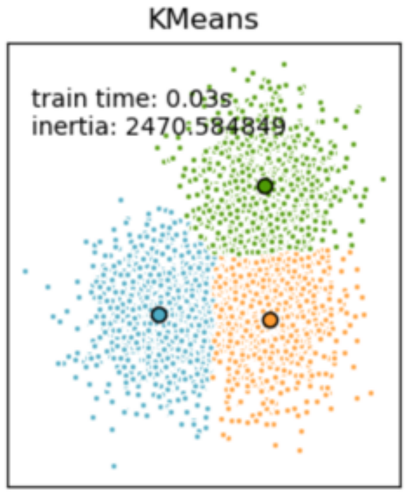


Sewaqu, Public domain, via Wikimedia Commons

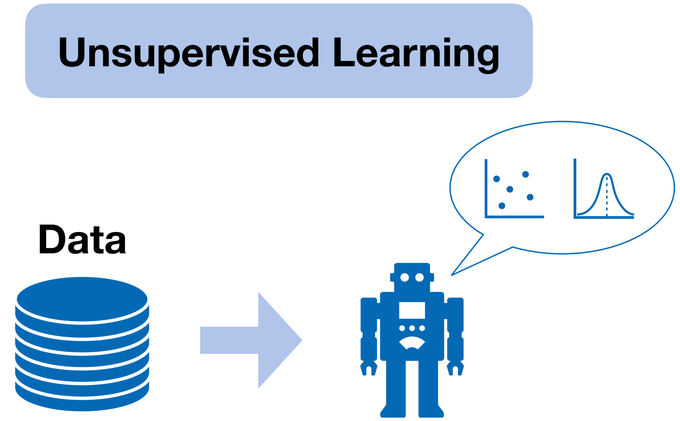
What is Machine Learning?

The process of computers changing the way they carry out tasks by **learning from new data**, without a human being needing to give instructions in the form of a program
(Cambridge Dictionary)

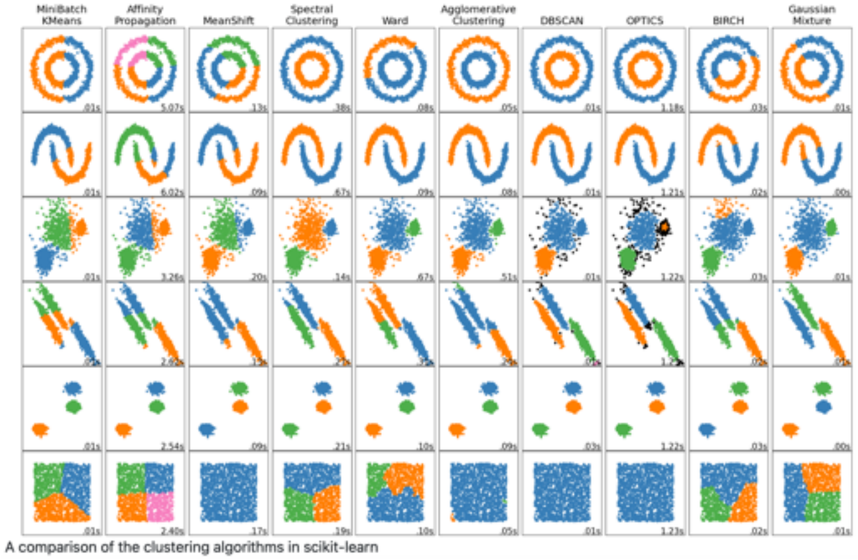
Machine learning can be broadly defined as **computational methods using experience to improve performance or to make accurate predictions**. Here, *experience* refers to the past information available to the learner, which typically takes the form of electronic data collected and made available for analysis. This data could be in the form of digitised human-labeled training sets, or other types of information obtained via interaction with the environment. In all cases, its quality and size are crucial to the success of the predictions made by the learner.
(Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.)



https://scikit-learn.org/stable/modules/clustering.html#k-means



Find structure in (unlabelled) data



A comparison of the clustering algorithms in scikit-learn

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

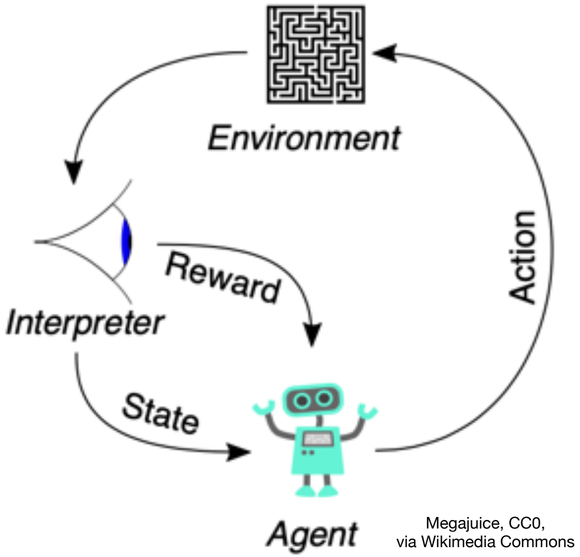
What is Machine Learning?

The process of computers changing the way they carry out tasks by **learning from new data**, without a human being needing to give instructions in the form of a program
(Cambridge Dictionary)

Machine learning can be broadly defined as **computational methods using experience to improve performance or to make accurate predictions**. Here, *experience* refers to the past information available to the learner, which typically takes the form of electronic data collected and made available for analysis. This data could be in the form of digitised human-labeled training sets, or other types of information obtained via interaction with the environment. In all cases, its quality and size are crucial to the success of the predictions made by the learner.
(Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.)

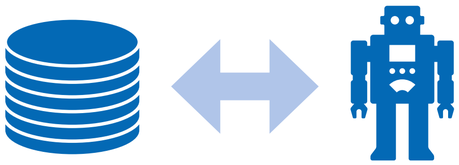
Chad Miller from Orlando, Florida, US of A,
CC BY-SA 2.0 <https://creativecommons.org/licenses/by-sa/2.0>,
via Wikimedia Commons

Ian Maddox, CC BY-SA 4.0
<https://creativecommons.org/licenses/by-sa/4.0>,
via Wikimedia Commons



Megajuice, CC0,
via Wikimedia Commons

Reinforcement Learning



Act to maximise rewards

What is Deep Learning?

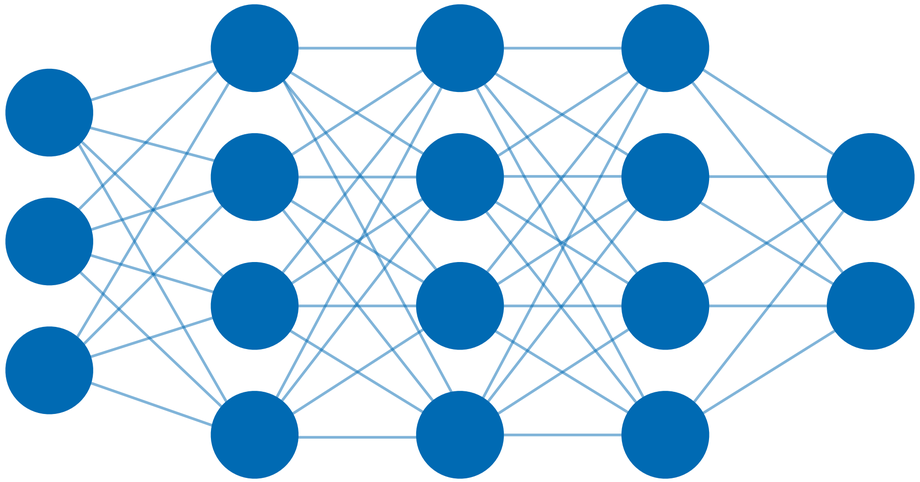
A type of artificial intelligence that uses algorithms (= sets of mathematical instructions or rules) based on the way the human brain operates
(Cambridge Dictionary)

A type of machine learning based on artificial neural networks in which multiple layers of processing are used to extract progressively higher level features from data
(Oxford Languages)

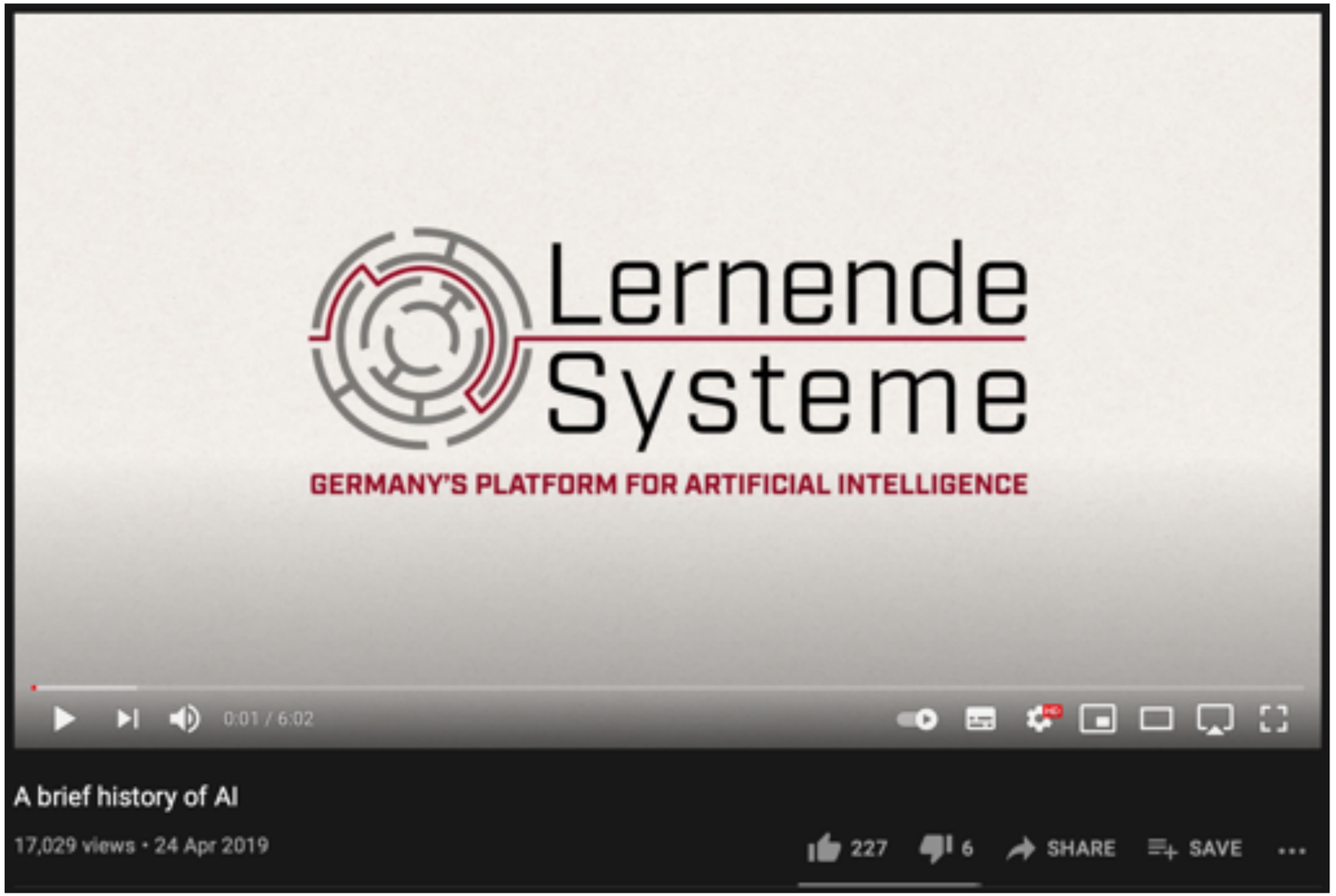
What is Deep Learning?

A type of artificial intelligence that uses algorithms (= sets of mathematical instructions or rules) based on the way the human brain operates
(Cambridge Dictionary)

A type of machine learning based on artificial neural networks in which multiple layers of processing are used to extract progressively higher level features from data
(Oxford Languages)



A Brief History of AI

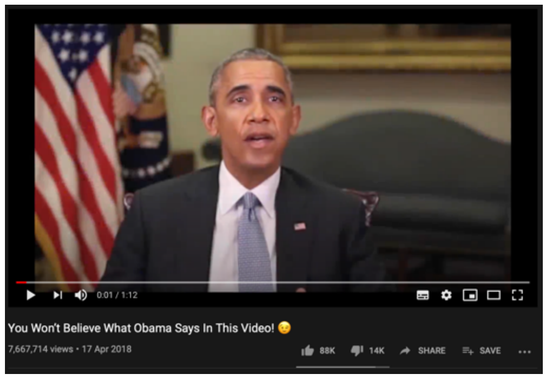


What AI can already do

Recognise and analyse speech/language

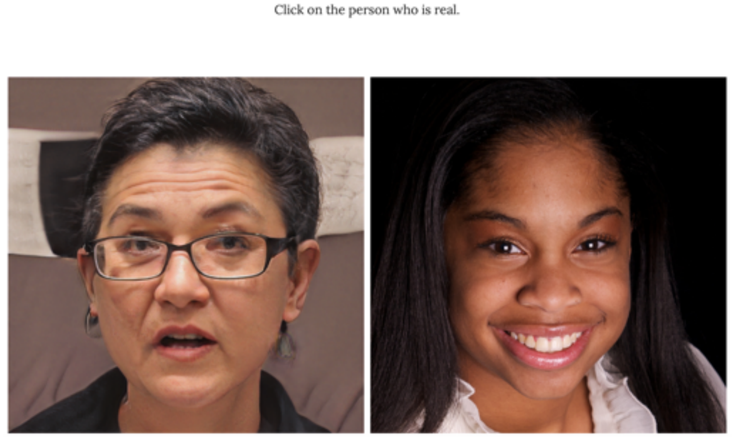
https://commons.wikimedia.org/w/index.php?curid=47040540

Produce deep fakes



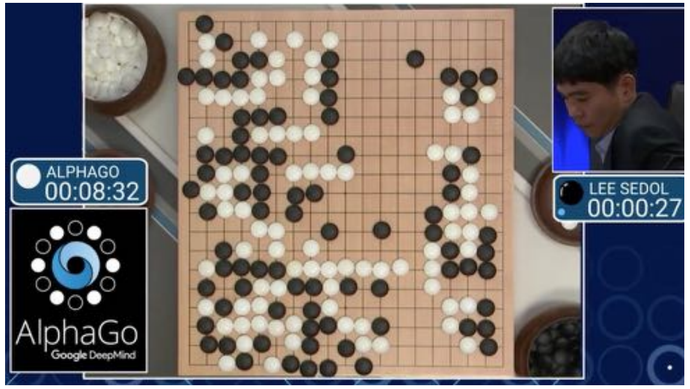
https://www.youtube.com/watch?v=cQ54GDm1eL0

Generate human faces



https://www.whichfaceisreal.com/index.php

Beat human opponents



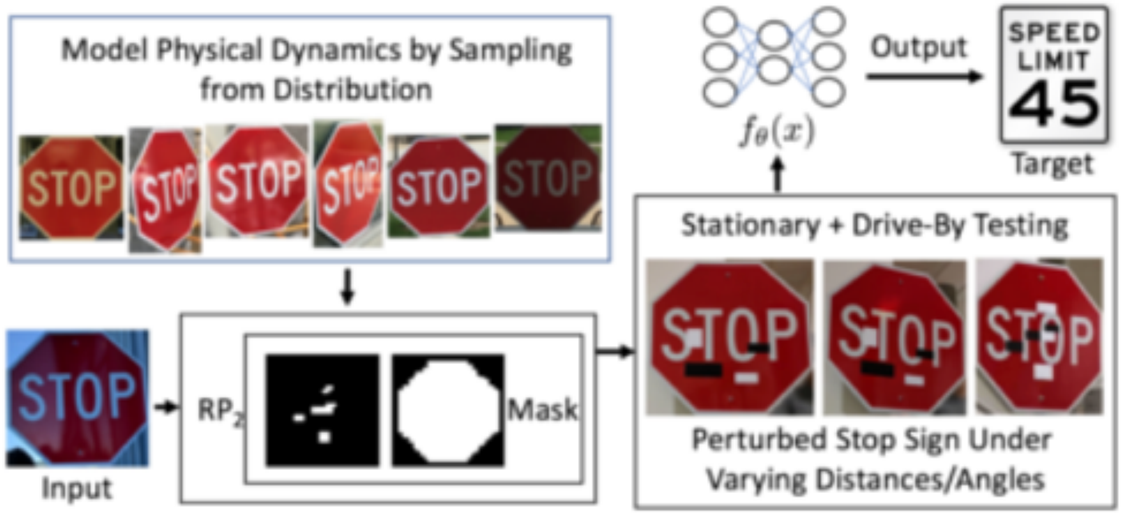
https://www.bbc.com/news/technology-35785875

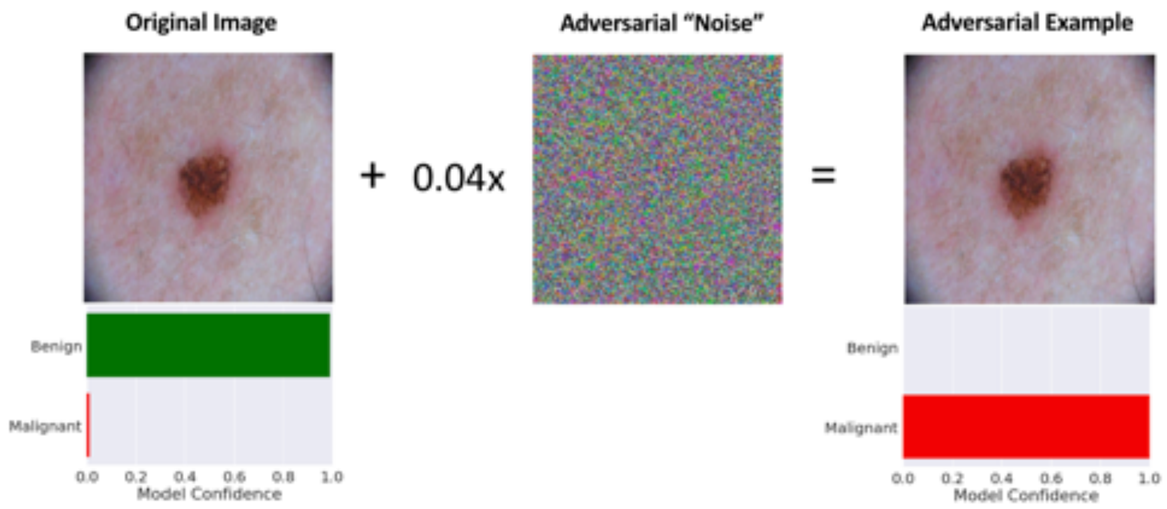
Transform photos into paintings

https://deepart.io

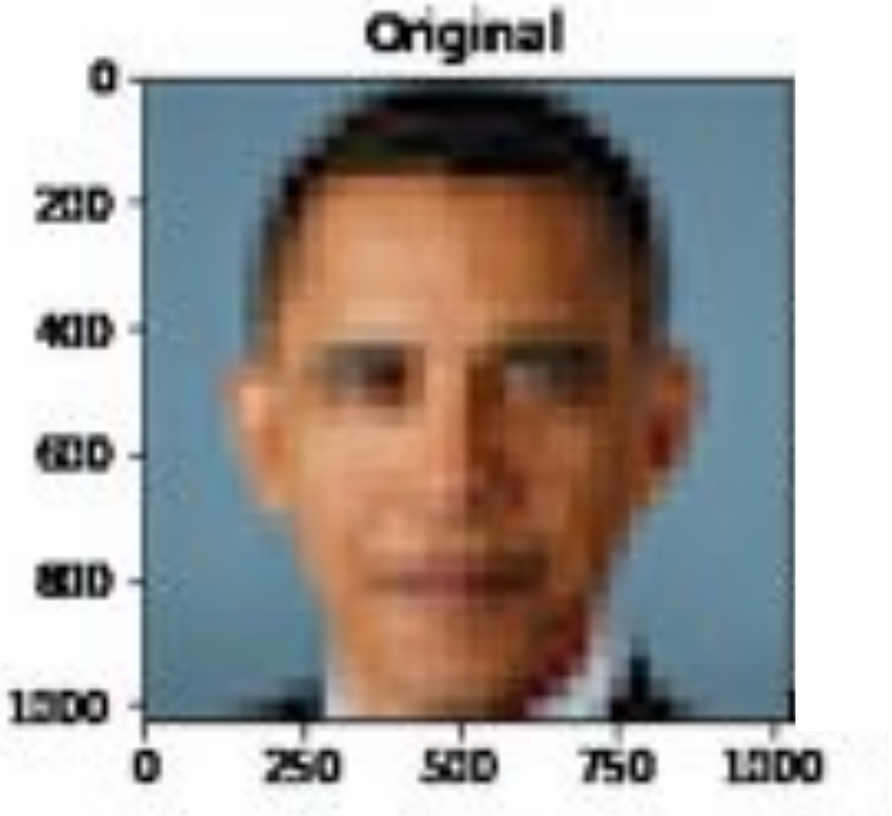
Issues and ethical/legal problems

Adversarial attacks



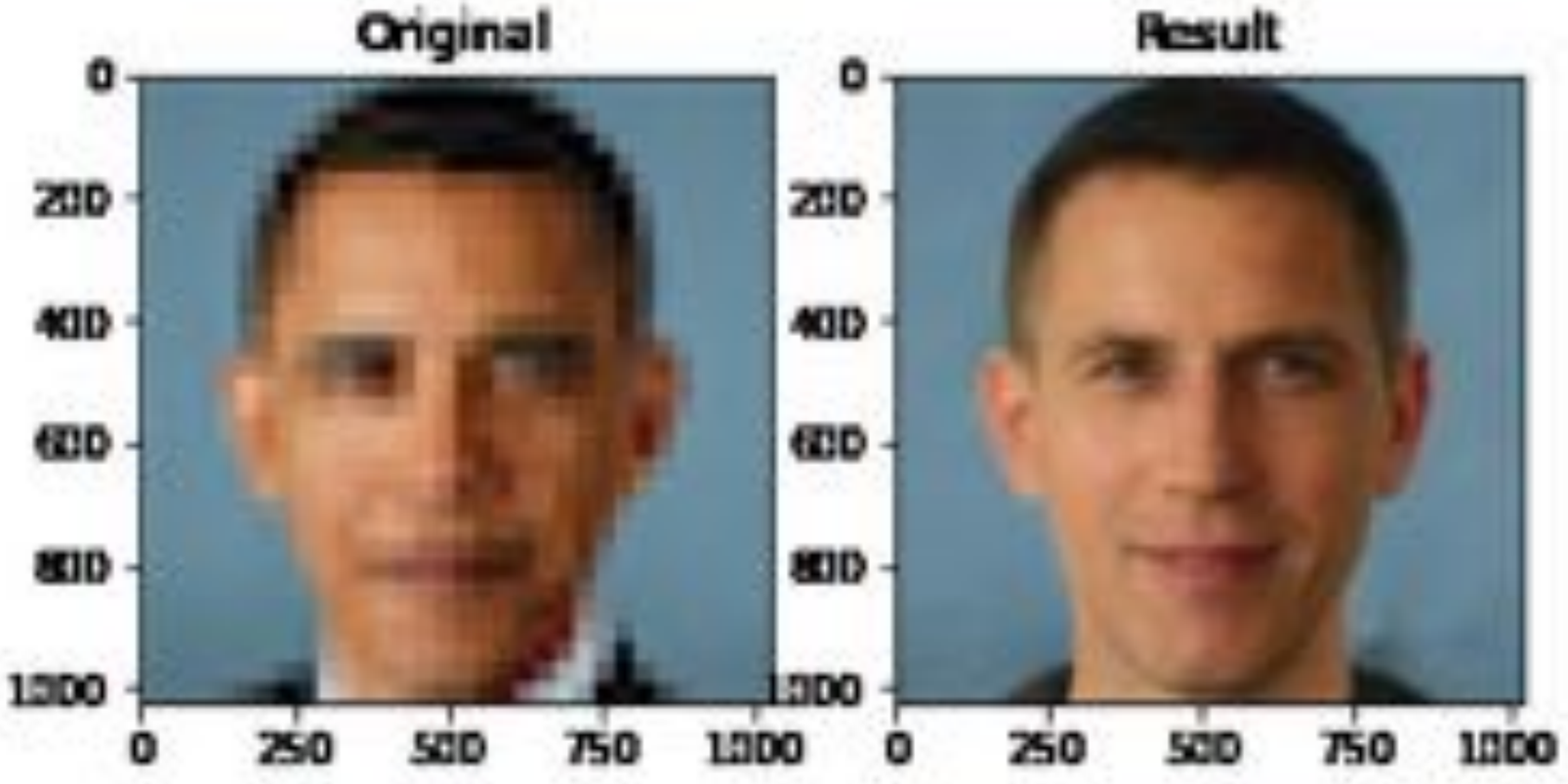


Issues and ethical/legal problems



https://twitter.com/Chicken3gg/status/1274314622447820801

Issues and ethical/legal problems



https://twitter.com/Chicken3gg/status/1274314622447820801

Issues and ethical/legal problems

Jul 1, 2015, 01:42pm EDT

Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software

https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/?sh=4b1fed4f713d

New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions

https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph

Are robots sexist? UN report shows gender bias in talking digital tech

https://news.un.org/en/story/2019/05/1038691

Issues and ethical/legal problems

https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Figure 2: Interrelationship of the seven requirements: all are of equal importance, support each other, and should be implemented and evaluated throughout the AI system's lifecycle

Revision - Q&A in Webex

Revision - Q&A in Webex

1. Using general artificial intelligence, we can already surpass human intelligence.
   A. True
   B. False

2. What is the correct relation? (A<B means A is a subset of B in this case)
   A. Deep Learning < Artificial Intelligence < Machine Learning
   B. Machine Learning < Deep Learning < Artificial Intelligence
   C. Deep Learning < Machine Learning < Artificial Intelligence

3. What is a broad definition of supervised machine learning?
   A. An agent interacts with an environment so that specific rewards are maximised
   B. Learning from input-output pairs to predict output from unseen input in an accurate manner
   C. Deducing structures and patterns from given data

4. What is a broad definition of unsupervised machine learning?
   A. An agent interacts with an environment so that specific rewards are maximised
   B. Learning from input-output pairs to predict output from unseen input in an accurate manner
   C. Deducing structures and patterns from given data

5. What is a broad definition of reinforcement learning?
   A. An agent interacts with an environment so that specific rewards are maximised
   B. Learning from input-output pairs to predict output from unseen input in an accurate manner
   C. Deducing structures and patterns from given data

6. Who/What causes racist/homophobe/sexist outcomes? (multiple answers possible)
   A. The programmers
   B. The users/applicants (domain experts)
   C. The algorithms
   D. Data bias

7. Deep Learning is a very recent research direction that has been established in the last decade.
   A. True
   B. False

BREAK

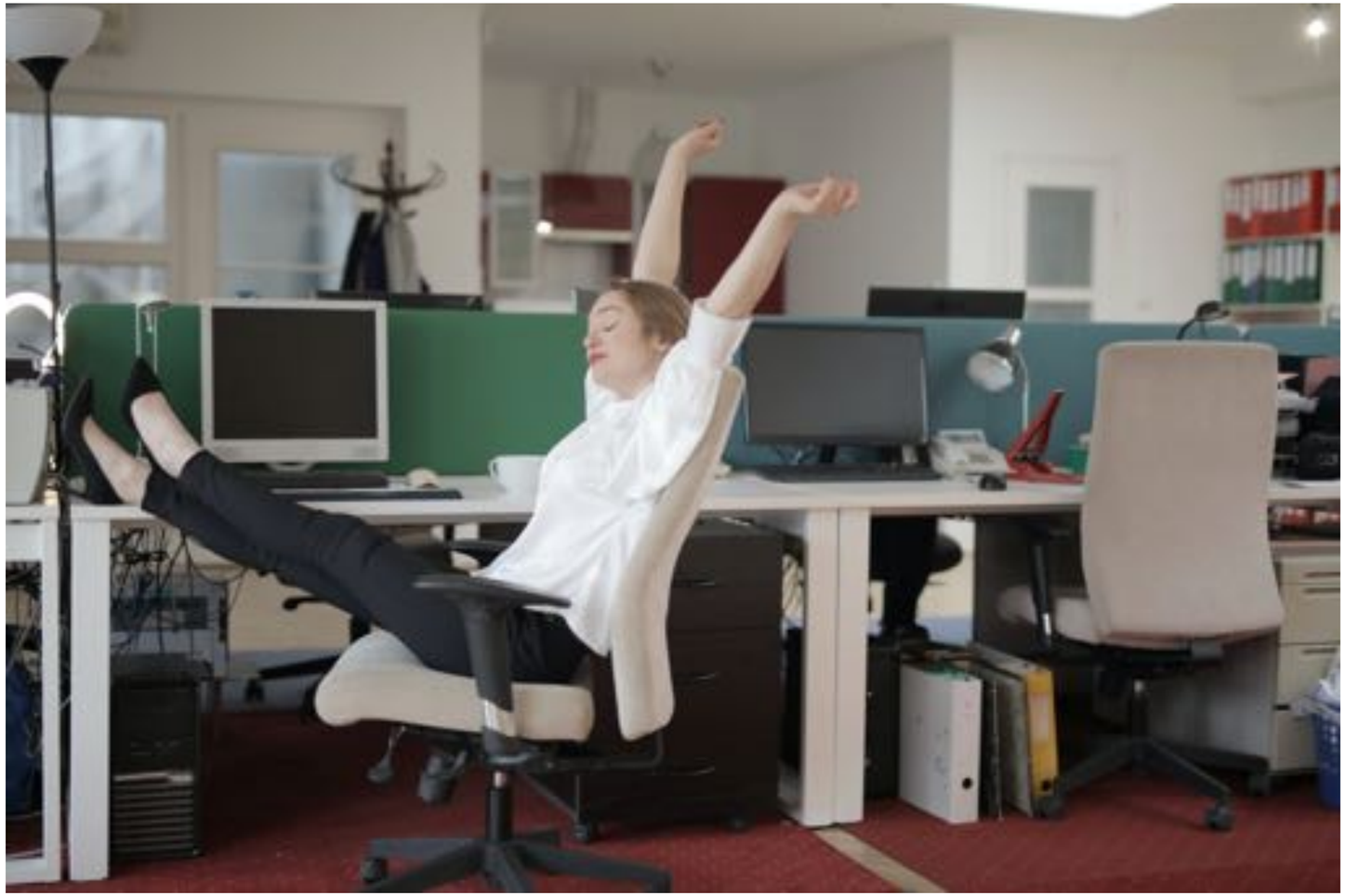


Photo by **Andrea Piacquadio** from **Pexels**

Jupyter Notebook Session

https://wiki.hhu.de/display/HPC/Wissenschaftliches+Hochleistungs-Rechnen+am+ZIM

https://wiki.hhu.de/display/HPC/Jupyter

https://jupyter.hpc.rz.uni-duesseldorf.de/hub/

Homework

Think of one example where AI or more particularly machine learning is applied in your research area that you find particularly interesting and useful for your doctoral research.

What type of data is needed in this case?

What kind of AI models are used?