

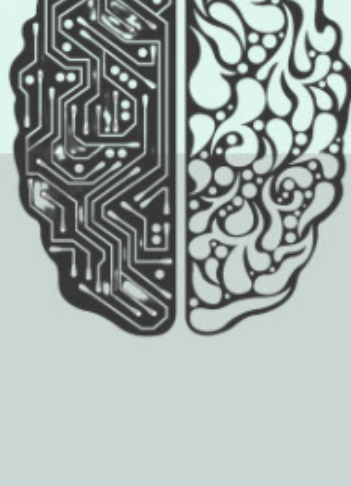
WHAT IS ARTIFICIAL INTELLIGENCE?

Machine Learning?

Deep Learning?

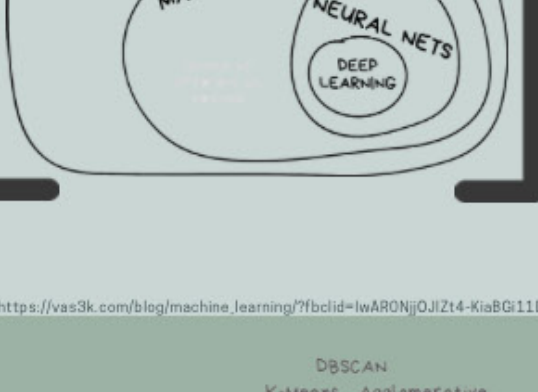
10010010000101010111011001

01011101000101010011000101

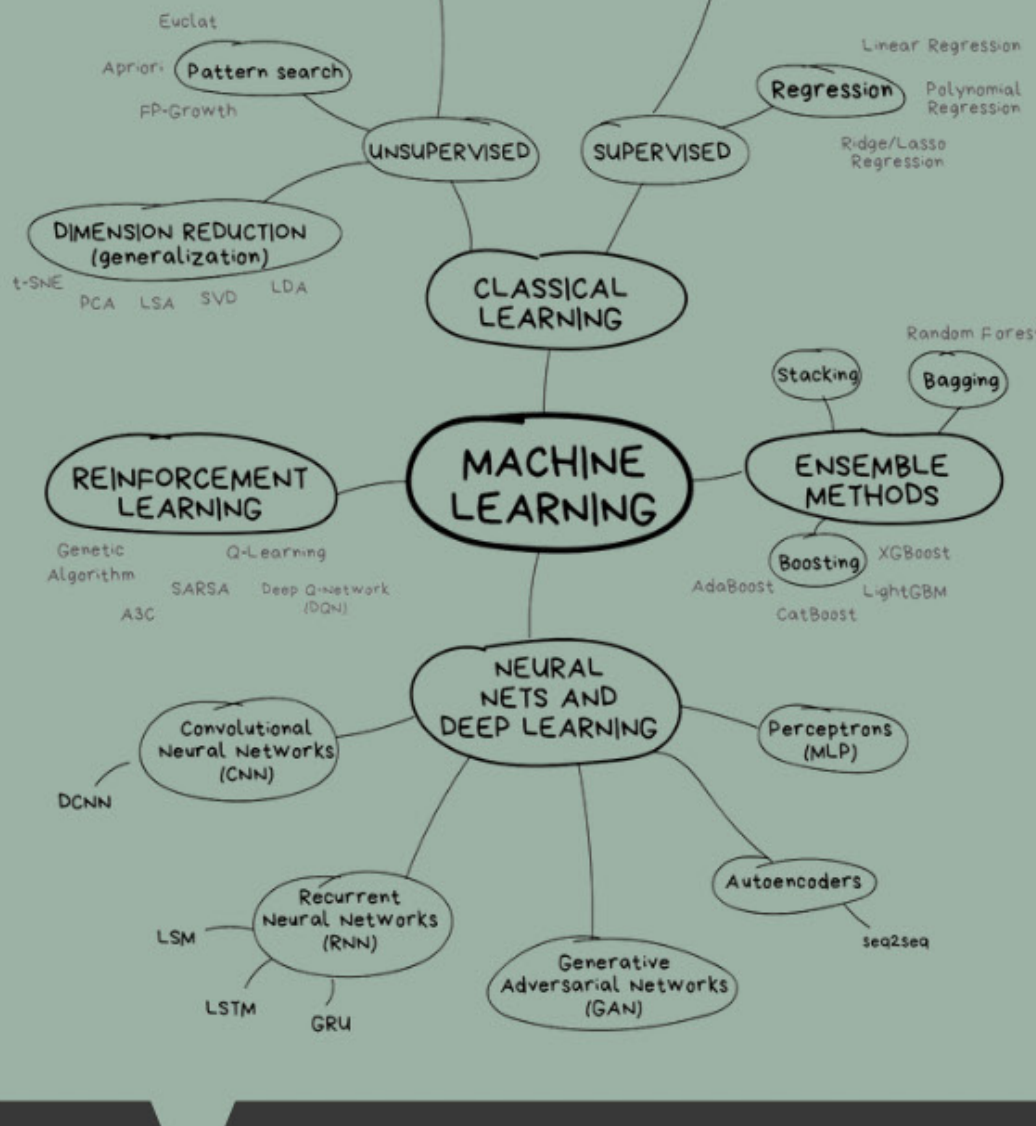


Artificial intelligence (AI) is the catch-all term for the creation and study of "intelligent machines". AI consists of Machine Learning, Neural Networks & Deep Learning.

LET'S UNPACK THIS



Machine Learning is made up of three parts:
 1. The Data used by the system which can be made up of relevant text, audio, images, etc.
 2. The Features or Parameters which allow the machine to capture information from the data
 3. The Training Methods or Algorithms which "teach" the machine the correct outputs and processes.



As the mind map on the left demonstrates, there are many types of Machine Learning systems. Artificial Intelligence creators may choose a specific type of Machine Learning depending on the type of task they plan for the machine to perform.

Machine Learning is analogous to the process of training your dog a new trick: You decide on the trick (ex. rollover), the method that you will use to train it (ex. positive reinforcement with treats) & the outcome should be a dog who can rollover (if it fails then you may try a new method).

WHERE IS AI IN MY LIFE?



SPAM MAIL FILTERING, FACE ID, ANTI-VIRUS SOFTWARE, SMART HOME TECH, GRAMMARLY, SHOPPING RECOMENDATINS, ASSISTED DRIVING EATURES, SURVIELLENCE, ANTI-VIRUS SOFTWARE, FRAUD DETECTION & BANKING

BUT, SHOULD WE FEAR AI?

SUPERINTELLIGENT AI

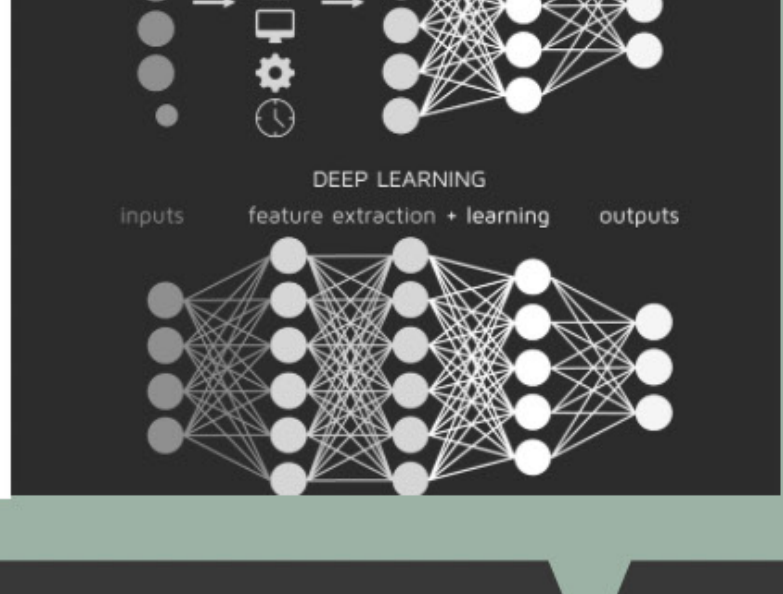
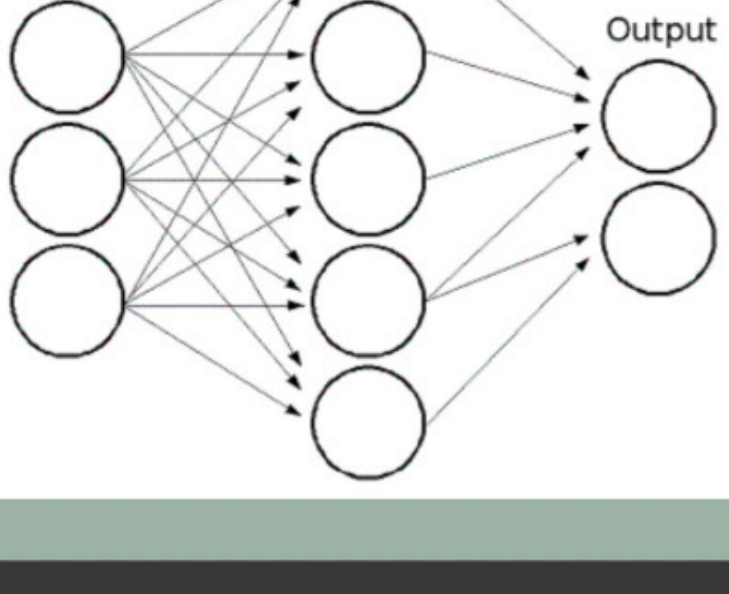
Basic Machine Learning AI systems such as our email's spam filter function are not the AI machines that bring about the existential dread we feel when watching Black Mirror's Metalhead.



The autonomous robotic quadrupeds in Metalhead are similar to Boston Dynamics 'Spot' seen here. Currently Spot is used for inspection, exploration and documentation.

Superintelligent AI is referred to when the AI machine is created with intelligence that exceeds human intelligence. Superintelligent AI requires Neural Networks and Deep Learning to problem solve & 'act' intelligently. Often this is accompanied by autonomy by the agent. Thus, fear of something autonomous & smarter than we could ever be follows naturally.

NEURAL NETWORK DEEP LEARNING



BEYOND THE PSYCHOLOGICAL FEAR

Given the complexity of Neural Networks and Deep Learning systems, similar to the human brain, we are unable to pinpoint what 'hidden layers' direct the output. Thus, we have to think of ways to prevent potential harms in creating these agents.

For instance, we can assume that in Metalhead the quadrupeds were created with the a different intention in mind, say helping police similar to police dogs, but they then begin to see all humans the same way that they were programmed to see criminals.

PREVENTATIVE OPTIONS



Attempt to mimic the human evolution of cooperation into the architecture of the machine

(Hubinger, Evan, & Haggström, Ollie)

BUT, THESE TOOLS ARE NOT GUARANTEED TO WORK

REGARDLESS OF WHETHER THE SAFETY PRECAUTIONS ARE IN PLACE WE CONTINUE TO BUILD SUPERINTELLIGENT AI

1001001000010101011101100101011101000101010011000101000110110111001010

SO, WHAT SHOULD WE FEAR?

Hint: Do not fear your spam mail filter

- Algorithms?**
 - Though helpful for most functions, the fear would be trying to fight with a machine that can predict your future actions with near accuracy.
- Boston Dynamics' Spot?**
 - Currently, Spot is being used to do routine inspections autonomously, but it doesn't have enough freedom to call for worry, but maybe soon?
- Superintelligent AI?**
 - It seems that the existential fears that we have are driven out of us by machines that could be autonomous, problem solving and have access to large amounts of data.

DO WE CONTINUE?

LONGTERMISM AND AI

If there is a possibility that we are unable to stop a machine from eradicating human existence (Metalhead) where should we stand? Contra Peter Singer's belief that AI Safety is not the most pressing concern for altruists, individuals such as Iaroslav Elistratov argue that the existential risk is high and very serious (HELIX).



Arguably, superintelligent AI may be the most concerning issue due to predictions that machine intelligence equating humans will be developed by 2045 with a 31% probability that this development will be 'bad' or 'extremely bad' for humanity (HELIX). Thus, by the Maxipok Rule, we should intervene.

MAXIPOK RULE:

"MAXIMIZE THE PROBABILITY OF AN 'OK OUTCOME', WHERE AN OK OUTCOME IS ANY OUTCOME THAT AVOIDS EXISTENTIAL CATASTROPHE" (BOSTROM)

WHAT DO YOU THINK?

1001001000010101011101100101011101000101010011000101000110110111001010

RESOURCES IN THIS INFOGRAPHIC:

Bostrom, Nick. 2013. "Existential Risk Prevention as Global Priority". University of Oxford.
 Elistratov, Iaroslav. 2020. "Is AI Safety 'Rather Speculative Long-Termism?'. Helix: A Journal of Interdisciplinary Research. Vol. 17.
 Haggström, Ollie. "Strategies for an unfriendly oracle AI with reset button". Chalmers University of Technology and the Institute for Future Studies.
 Hubinger, Evan. May 29, 2020. "An overview of AI proposals for building safe advanced AI". MIRA.
 JAY SHAH: <https://blog.statsbot.co/neural-networks-for-beginners-d99f2235efca>
 vas3k: https://vas3k.com/blog/machine_learning/?fbclid=IwAR0Nj0JZt4-KiaBG11DskBHAa2d6xaUchKpZdDch7pxS5sbcrcZkUBJA