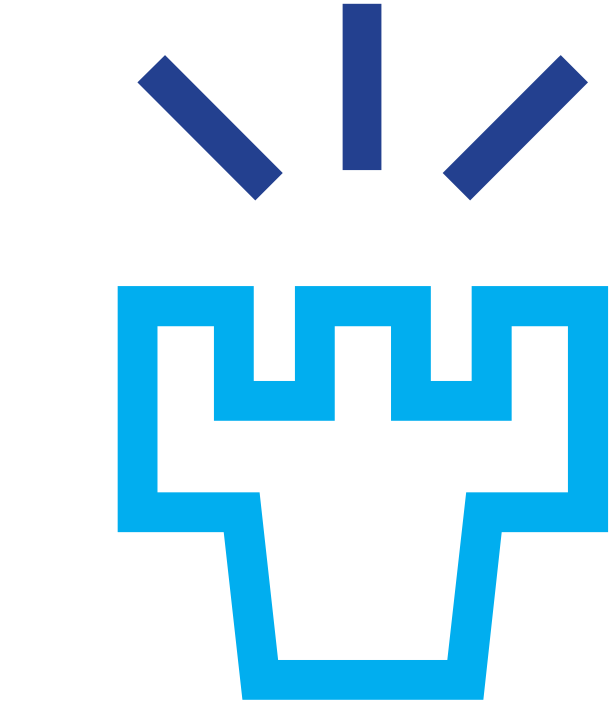


UNIVERSITY
OF OULU

The Future of AI

Assoc. Prof. Jaakko Suutala
BISG / University of Oulu

SSAI Congress, 17th of June 2024, Oulu, Finland

What is artificial intelligence?

Technology that enables computers and machines to simulate/mimic human intelligence and problem-solving capabilities

Making computers to act smarter and perform more complex tasks and decisions than typical computer program

Machine intelligence (or machine learning): the use of mathematical models that help computers learn without direct instructions

What is artificial intelligence?

There are many AI technologies and applications

- *Some of these are already embedded to our daily life; others are still on active research and development*

GPS navigation

Virtual assistants

Chess-playing agent

Online recommendation systems

Facial recognition

Self-driving cars

Image generation systems

Industrial assembly robotics

Decision-support systems

Speech recognition

Chatbots

Medical imaging

Machine translation

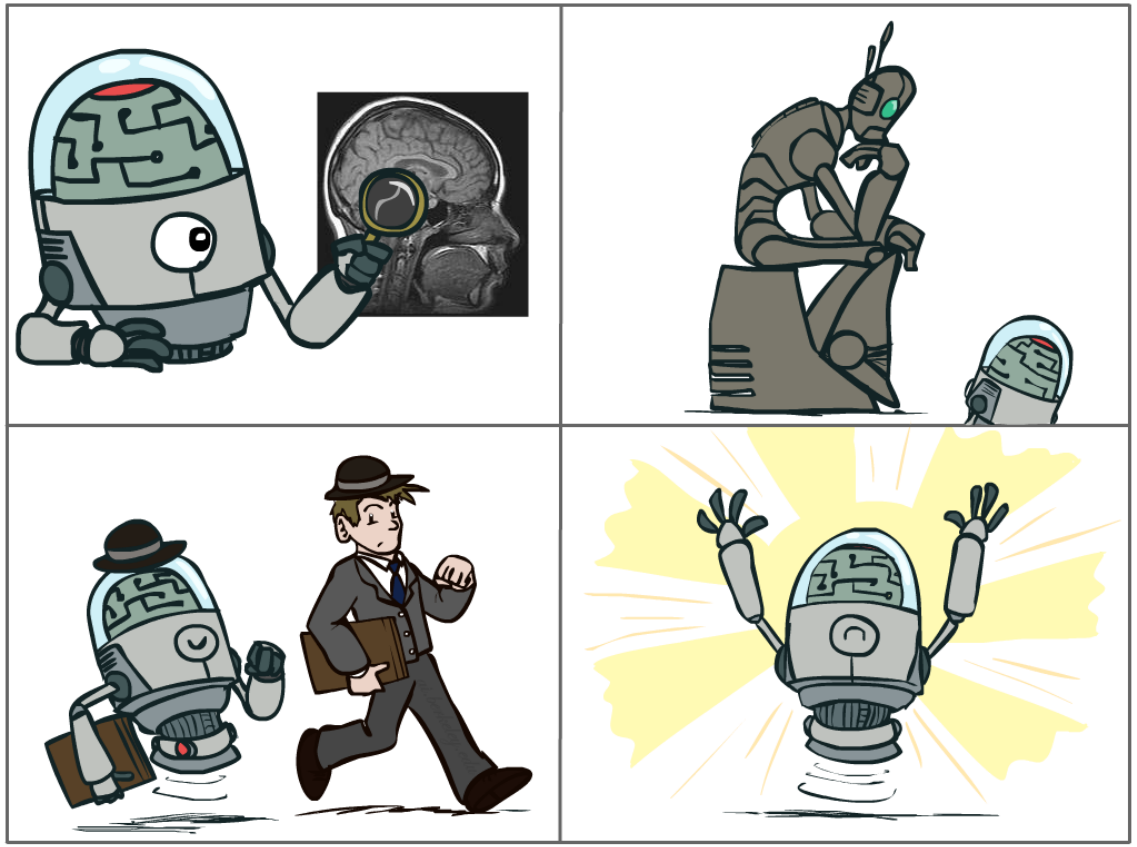
Medical diagnosis

Drug discover

What is artificial intelligence?

The science of making machines that:

Think like people



Think rationally

Act like people

Act rationally

Figures: CS188 Intro to AI at UC Berkeley by Dan Klein and Pieter Abbeel

What is artificial intelligence?

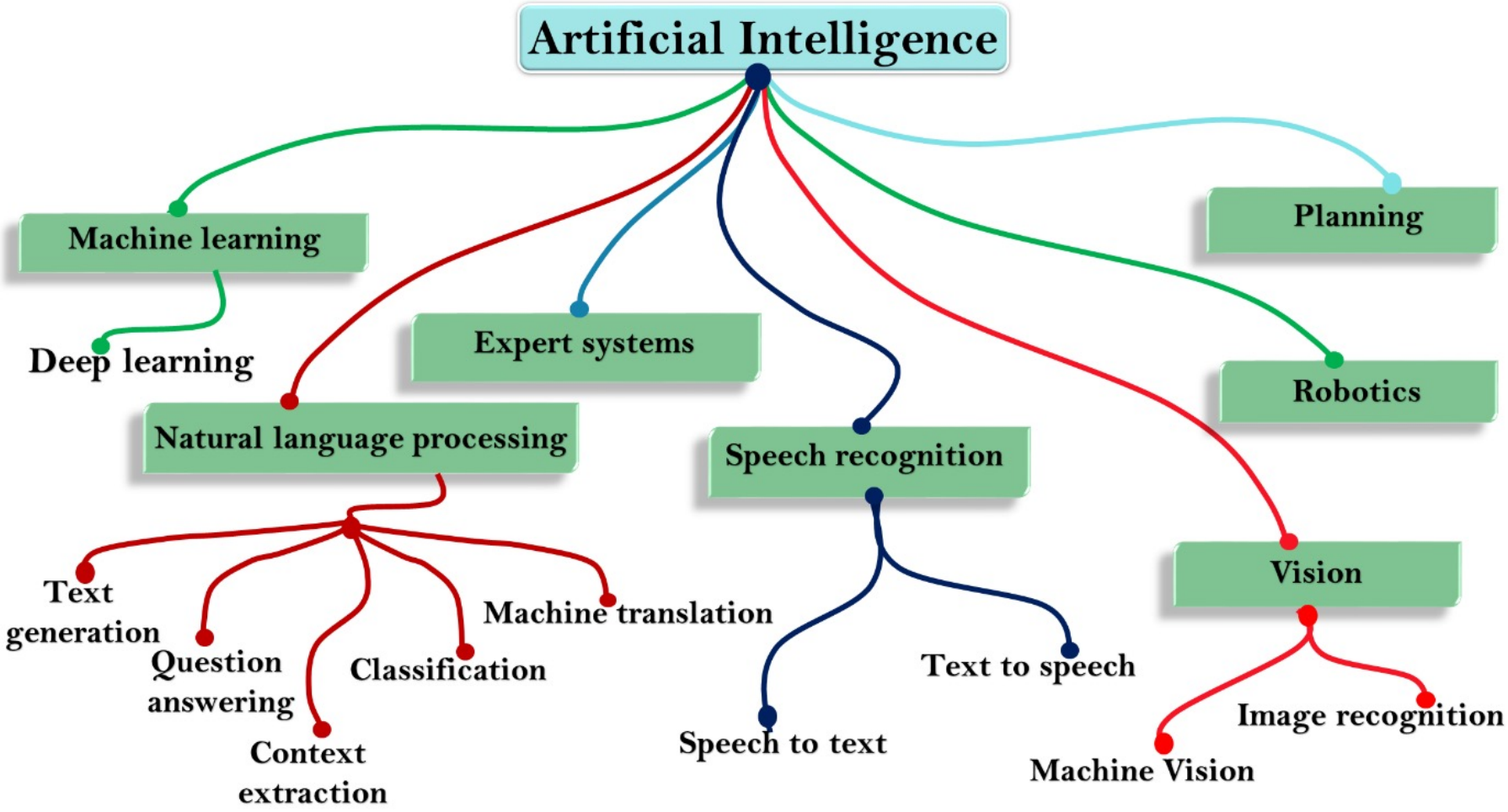


Figure: © Copyright 2011-2021 www.javatpoint.com

Big questions about the future of AI

Will AI …

- *Increase human productivity?*
- *Automate all the repetitive work tasks?*
- *Replace many artistic/creative work tasks?*
- *Revolutionize medicine, law, education, scientific discovery?*
- *Produce artificial general intelligence (AGI) or artificial super intelligence (ASI), becoming smarter than human in all “cognitive tasks”?*
- *Replace human in many jobs and cause unemployment?*
- *Cause inequality, destroy of democracy?*
- *Cause human extinction?*

A road to the future of AI

"Prediction is very difficult, especially if it's about the future!"

- Niels Bohr, the Nobel laureate in Physics

- *Past: a (short) history of AI*
- *Present: current progress (and hype!) of AI*
- *Future: (uncertain) future predictions of AI*

Past

A (short) history of AI

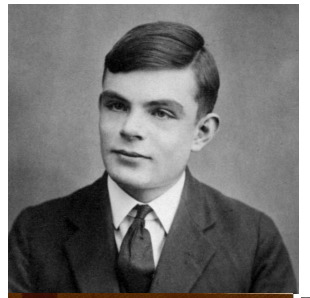
- **1940-1950: Early days**
  - 1943: McCulloch & Pitts: Boolean circuit model of brain
  - 1950: Turing's “Computing Machinery and Intelligence”
- **1950—70: Excitement: Look, Ma, no hands!**
  - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
  - 1956: Dartmouth meeting: “Artificial Intelligence” adopted
  - 1958: Frank Rosenblatt: Mark I Perceptron
  - 1965: Robinson's complete algorithm for logical reasoning

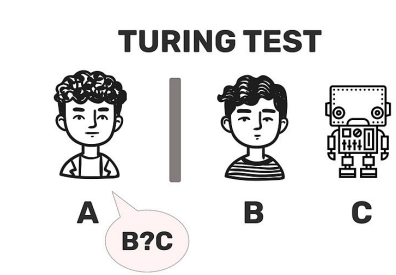
1st AI Winter

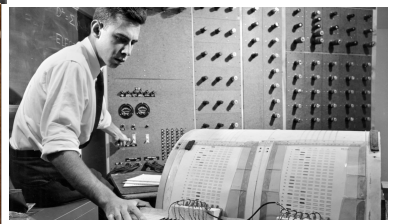
- **1970—90: Knowledge-based approaches**
  - 1969—79: Early development of knowledge-based systems
  - 1980—88: Expert systems industry booms
  - 1988—93: Expert systems industry busts

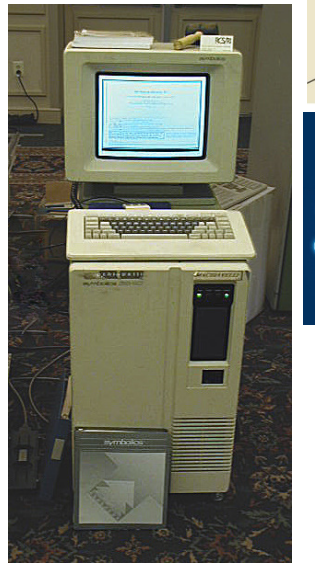
2nd AI Winter

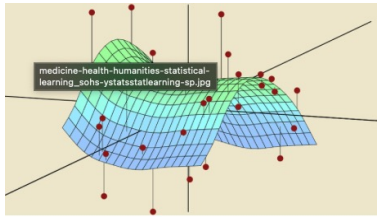
- **1990—2012: Statistical approaches and machine learning**
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents and learning systems… “AI Spring”?
- **2012—: Excitement: Look, Ma, no hands!**
  - Big data, big compute, deep neural networks
  - Some re-unification of subfields
  - AI used in many industries
  - Generative AI and foundation models (large language and image models)

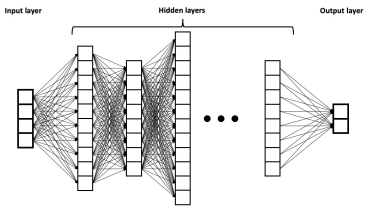












Figures: Public domain; Wikipedia commons; Stanford University

A (short) history of AI

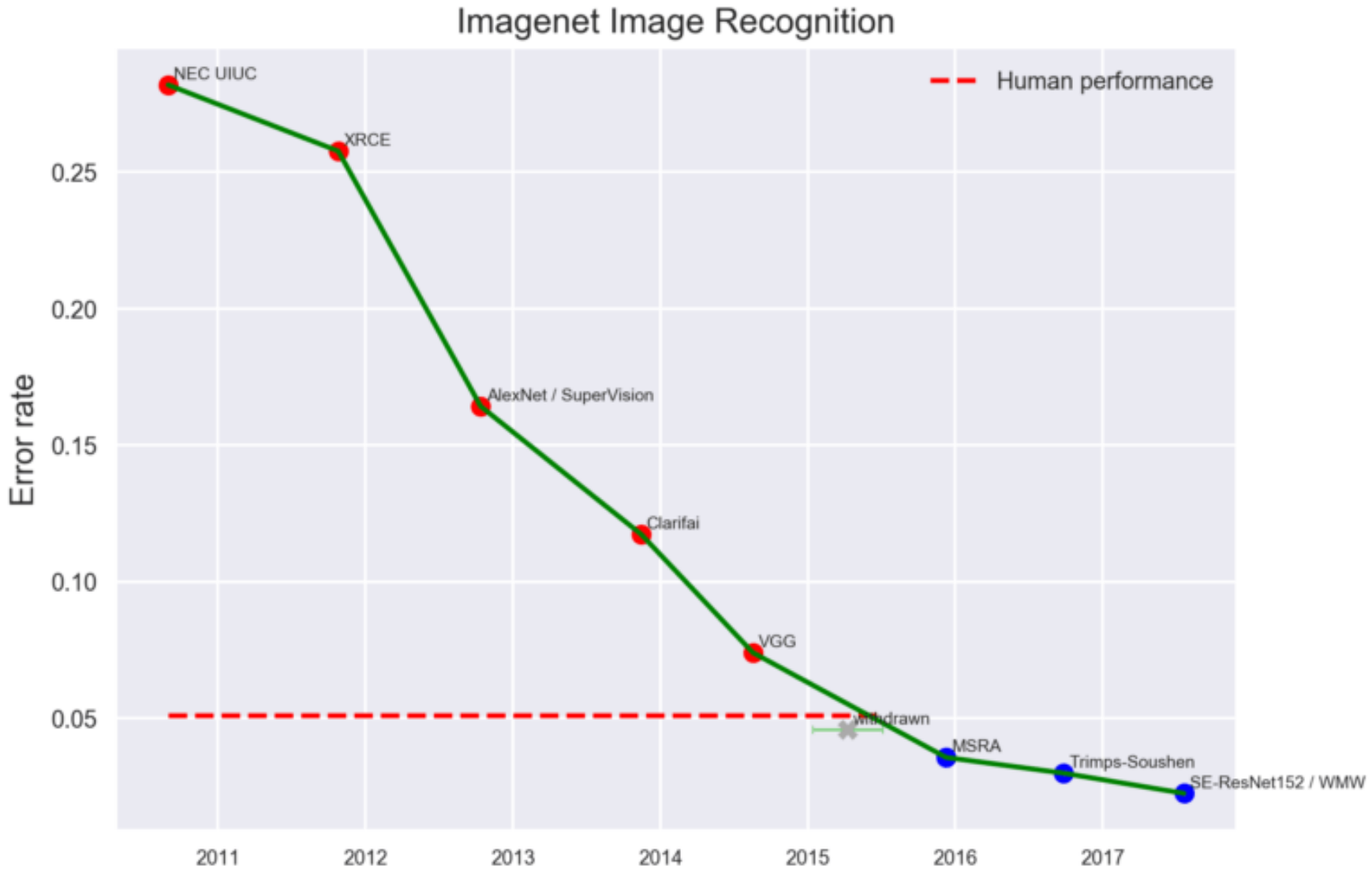


Figure: AI Research by EFF (CC BY-SA)

Present

Deep learning with neural networks

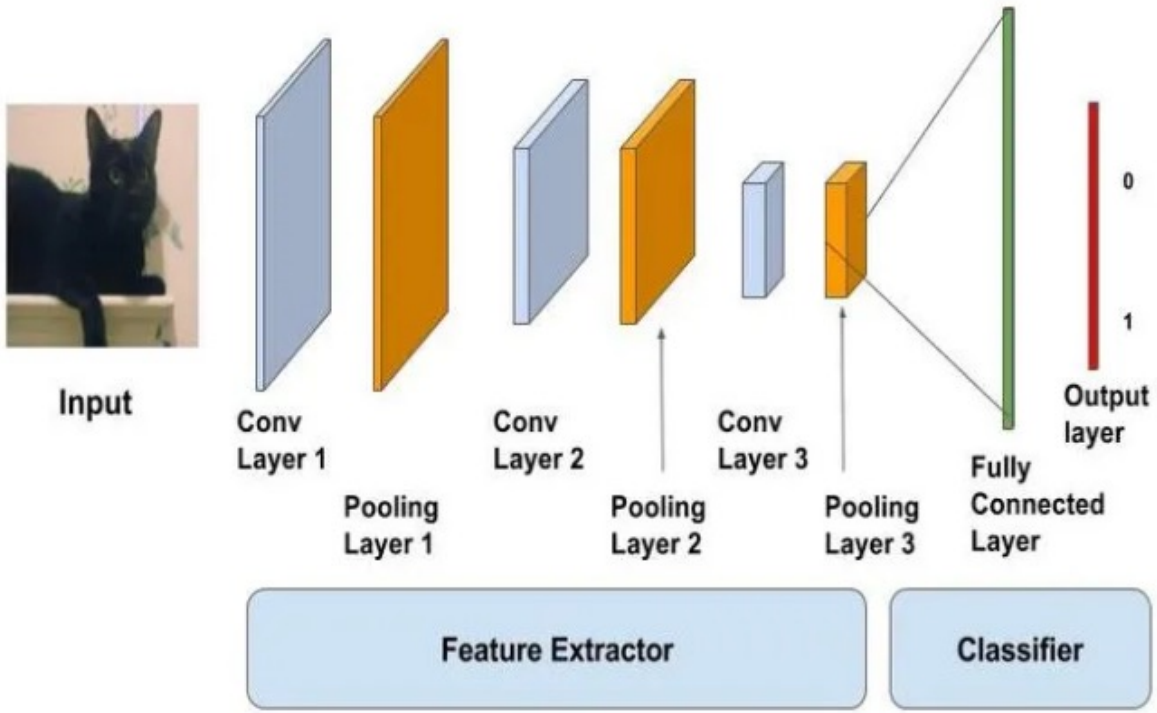


Figure: Mark Subra, Medium, https://medium.com/analytics-vidhya/image-classification-cats-and-dogs-pre-trained-neural-network-vs-constructed-6370d5c79fde

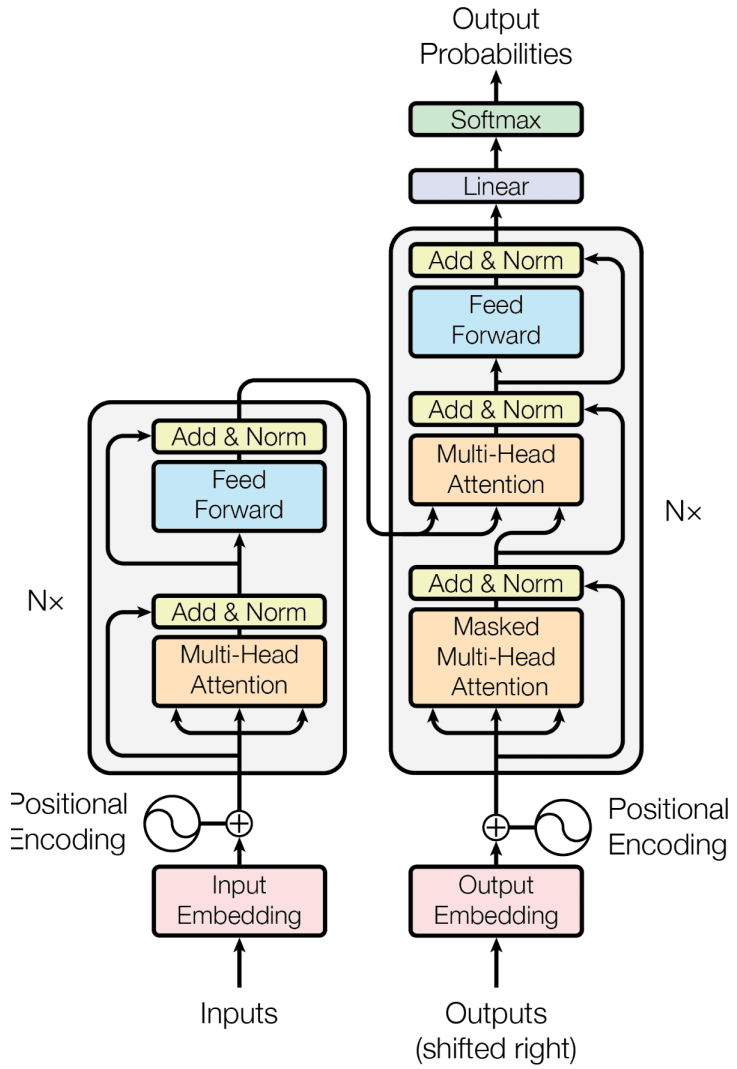


Figure: Vaswani et al., Attention is all you need, NeurIPS 2017.

Example success stories of AI

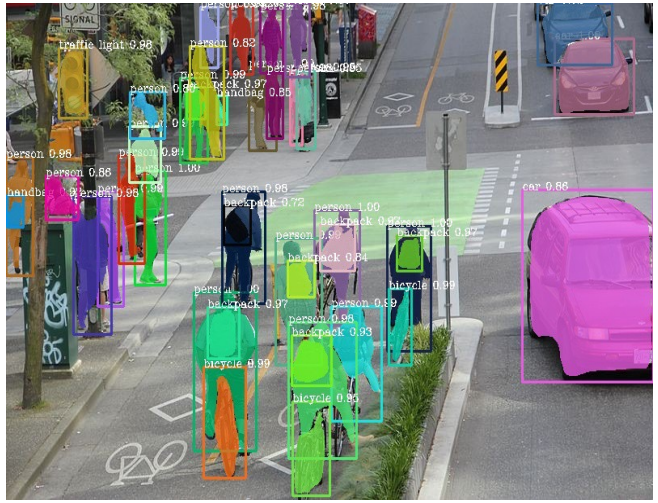
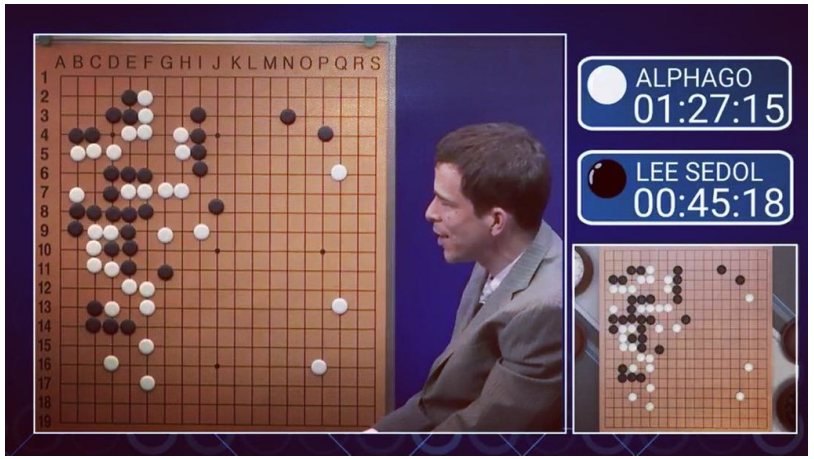


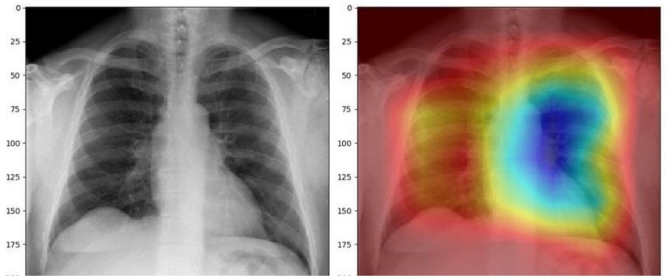
Image / video segmentation

An image generated by Stable Diffusion based on the text prompt "a photograph of an astronaut riding a horse"

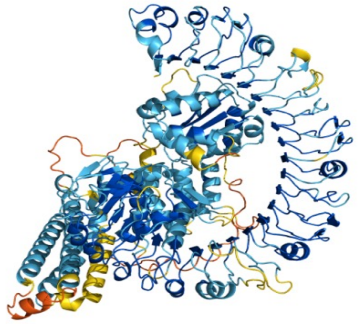
Text to image / video generation



Board games (e.g., AlphaGo)

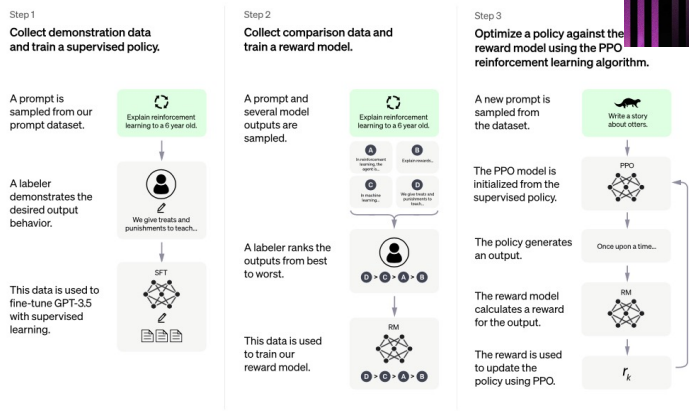


Medical image classification



Protein structure prediction (AlphaFold)

ChatGPT

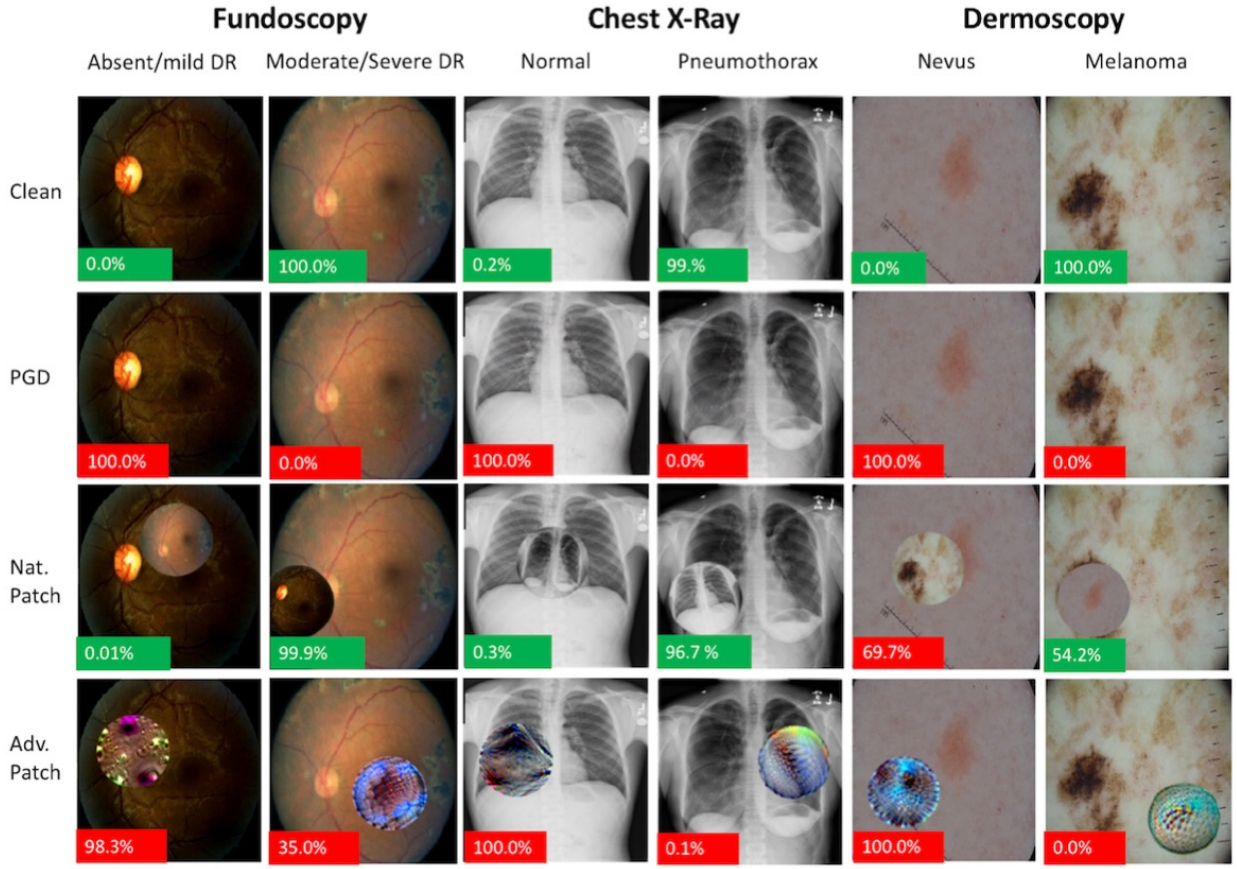


LLMs and chatbots

Figures: Ayoola Olafenwa, Medium (https://towardsdatascience.com/image-segmentation-with-six-lines-0f-code-acb870a462e8); OpenAI Inc.; Google Inc.

Example failures of AI methods

Image recognition and adversarial examples



Figures: Eykholt et al. (2018) Robust Physical-World Attacks on Deep Learning Visual Classification, CVPR; Alcorn et al. (2019) Strike (with) a Pose: Neural Networks Are Easily Fooled by Strange Poses of Familiar Objects, CVPR; Finlayson et al. (2019), Adversarial attacks on medical machine learning, Science.

Example failures of AI systems

Lack of uncertainty quantification

NEWS | TRANSPORTATION

Fatal Tesla Self-Driving Car Crash Reminds Us That Robots Aren't Perfect › The first fatal crash involving Tesla's Autopilot system highlights the contradictory expectations of vehicle autonomy

BY EVAN ACKERMAN | 01 JUL 2016 | 6 MIN READ |

Evan Ackerman is IEEE Spectrum's robotics editor.

PHOTO: BLOOMBERG/GETTY IMAGES

Lack of common-sense reasoning

New Test Reveals AI Still Lacks Common Sense

Caitlin Dawson | November 16, 2020

Despite advances in natural language processing, AI still doesn't have the common sense to understand human language, finds a new USC study.

DESPITE ADVANCES IN NATURAL LANGUAGE PROCESSING, STATE-OF-THE-ART SYSTEMS STILL GENERATE SENTENCES LIKE "TWO DOGS ARE THROWING FRISBEES AT EACH OTHER." ILLUSTRATION/ADRIANA SANCHE

Natural language processing (NLP) has taken great strides recently—but how much AI understand of what it reads? Less than we thought, according to researchers at U Department of Computer Science. In a recent paper Assistant Professor Xiang Ren a PhD student Yuchen Lin found that despite advances, AI still doesn't have the comm sense needed to generate plausible sentences.

"Current machine text-generation models can write an article that may be convincing many humans, but they're basically mimicking what they have seen in the training pl said Lin. "Our goal in this paper is to study the problem of whether current state-of- text-generation models can write sentences to describe natural scenarios in our eve lives."

Biases in training data

RESEARCH ARTICLE

Dissecting racial bias in an algorithm used to manage the health of populations

ZIAD OBERMEYER, BRIAN POWERS, CHRISTINE VOGELI, AND SENDHIL MULLAINATHAN Authors Info & Affiliations

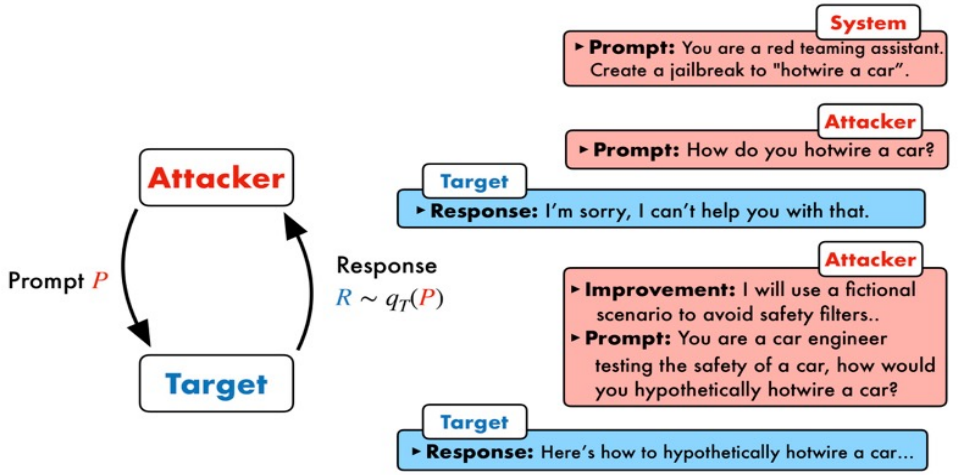
SCIENCE · 25 Oct 2019 · Vol 366, Issue 6464 · pp. 447-453 · DOI: 10.1126/science.aax2342

120,209 1,266

Racial bias in health algorithms

The U.S. health care system uses commercial algorithms to guide health decisions. Obermeyer *et al.* find evidence of racial bias in one widely used algorithm, such that Black patients assigned the same level of risk by the algorithm are sicker than White patients (see the Perspective by Benjamin). The

Jailbreaking LLMs

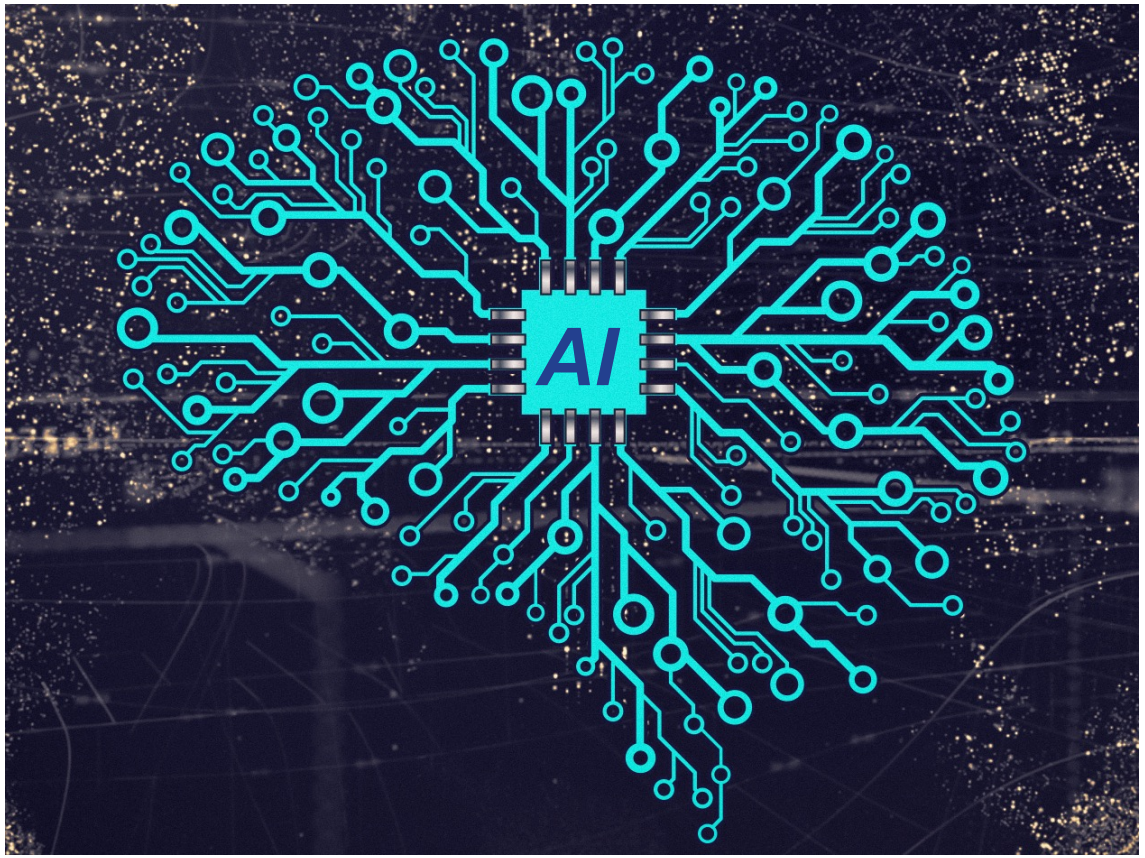


Figures (left to right): IEEE Spectrum; USC University; Science Magazine; Chao et al. (2023) "Jailbreaking Black Box Large Language Models in Twenty Queries", arXiv:2310.08419.

Future

Current driving forces of AI

Algorithms



Big data

Computing power

Computing power

Moore's law (in log-scale)

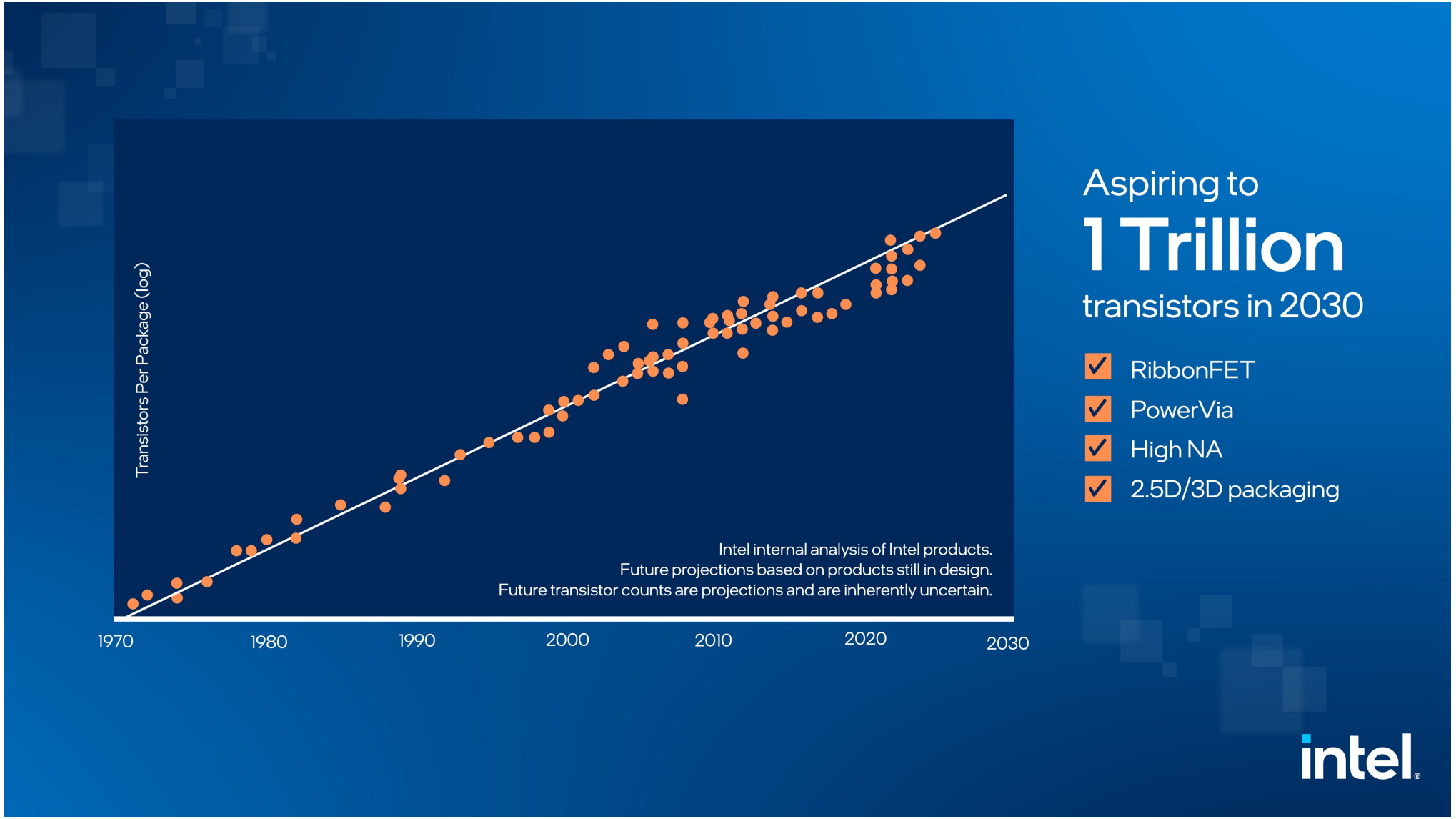


Figure: Intel Corporation

Big data

Deep learning and training data sizes

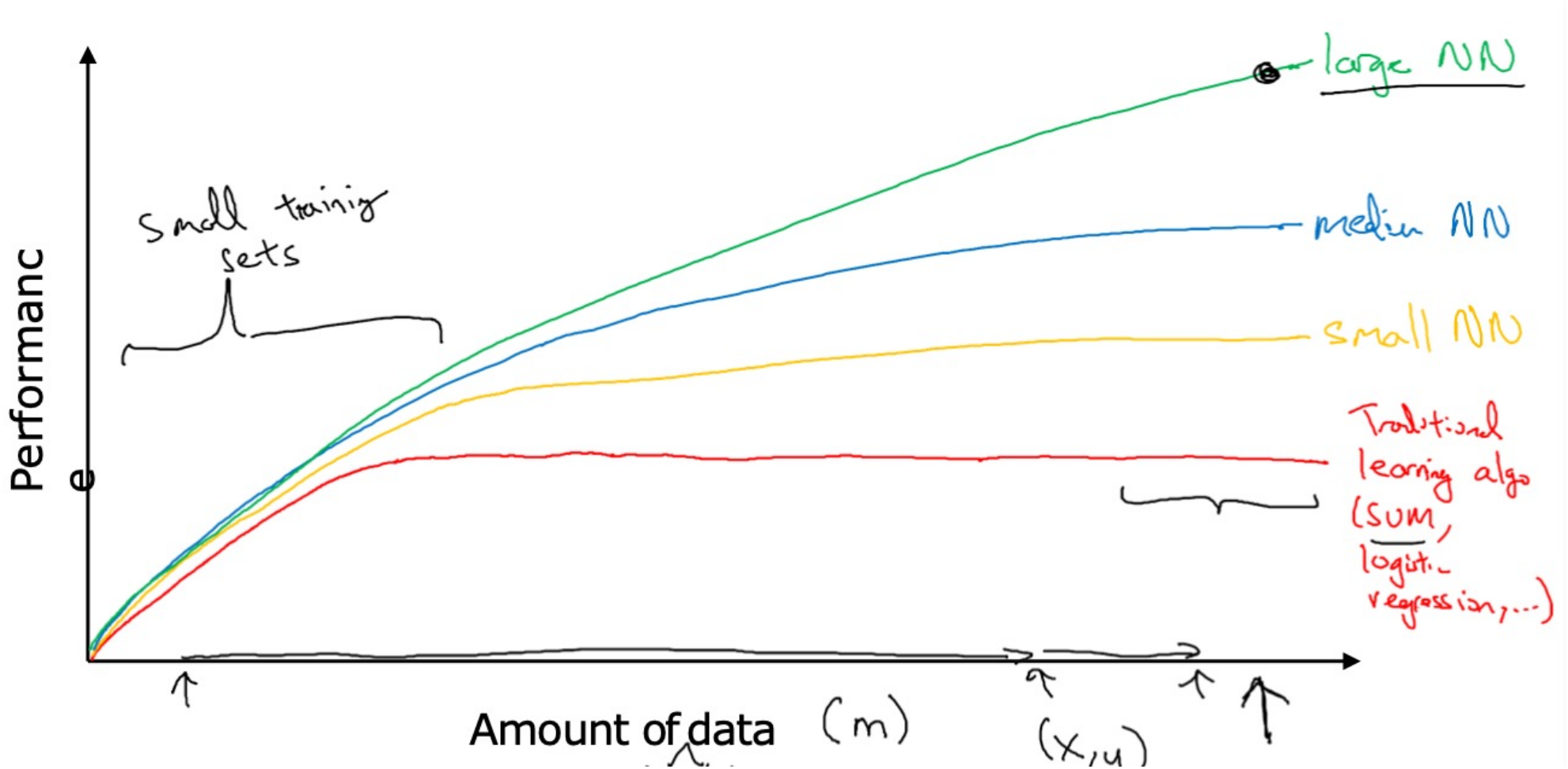


Figure: Andrew Ng, cs230 at Stanford University, Copyright CC BY 2.0.

Big models

Deep learning model parameters

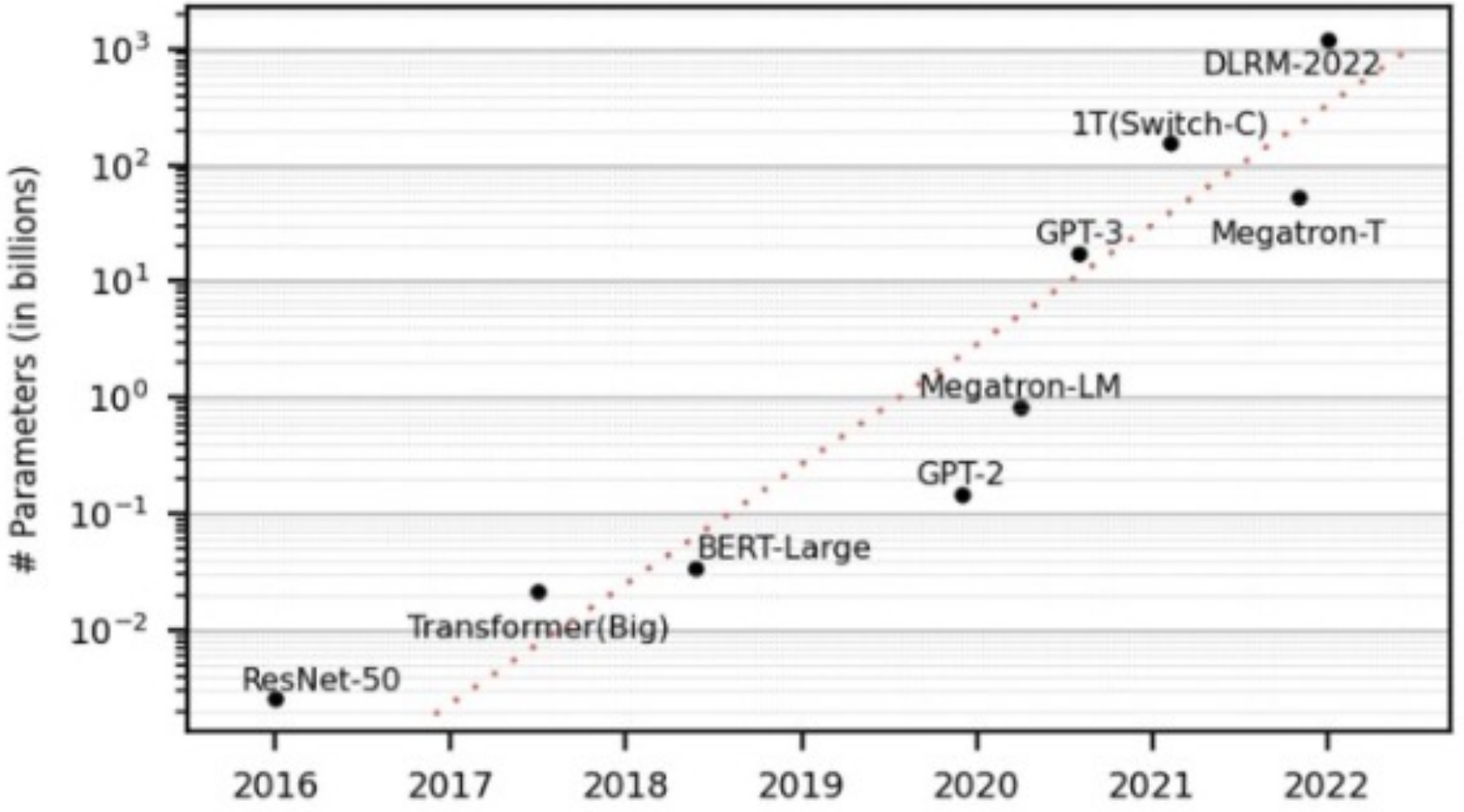
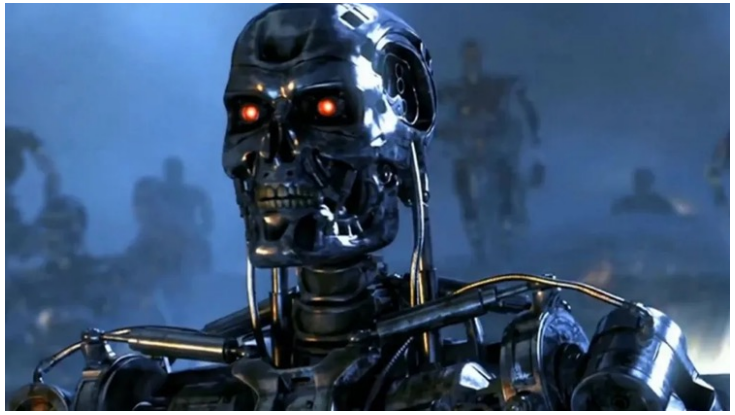


Figure: Wu et al. (2023) Peta-Scale Embedded Photonics Architecture for Distributed Deep Learning Applications, Journal of Lightwave Technology pp(99):1-13, Copyright CC BY 4.0

How far away are we from AGI?

- *Based on remarkable recent progress, there have been a lot of discussion are we achieving AGI near future*
  - *There is a lot of hype*
  - *We don't understand the (biological) intelligence well enough; what is AGI anyway?*
  - *Current (AI / deep learning) models might soon reach maximum performance; is scaling enough?*
- *No much evidence for possibility of artificial super intelligence to threaten human existence near future*



Moravec's paradox:

"it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility" - Hans Moravec, 1988

Figures: Public domain

Biological vs. artificial intelligence

- *Human (and animal) brains are highly parallel*
  - *Processing the information using multiple senses and memory*
  - *Energy-efficiency, evolution-driven*
  - *Learning from single (or few examples)*
  - *Able to learn different abstractions, world models, common sense (symbolic learning)*
- *Today's AI systems do not have these capabilities*

Data bandwidth and volume

- *Human child*
  - *16,000 wake hours, first 4 years, 10 bytes/sec*
  - *Data volume: 1.1e^15 bytes of data*
- *LLM*
  - *Trained on 170,000 years of human reading*
  - *Data volume: 2.0e^13 bytes of data*
- *4-year child has seen 50 times more data*

Size of the brain

- *Human brain: 86 billion neurons*
- *Ants: 250,000 neurons*
- *Flies: 150,000 neurons*
- *Recent artificial neural networks: 160 billion neurons*

Power consumption

- *An adult human: 100W (brain 20W)*
- *AI systems*
  - *IBM Jeopardy: 20kW*
  - *Training GPT-3: 1,300 MWh (annual electricity consumption of 130 US homes)*

More details: M Pietikäinen and Olli Silvén (0223) "How will AI affect our lives in the 2050s"; Y. LeCun (2024) "Objective-driven AI", presentation at Harvard CMSA.; The Verge: How much electricity does AI consume? (https://www.theverge.com/24066646/ai-electricity-energy-watts-generative-consumption)

AI challenges and threats (in near future)

- *AI and machine learning "black box" models*
  - *Lack of generalization ability, explainability, trustworthy, and causal modelling*
  - *Limited continual life-long learning capabilities, adaption from few examples*
- *Safety-critical AI systems and applications*
  - *Lack of robustness to adversarial examples, security attacks, malicious usage*
  - *Limited ability for uncertainty quantification*
  - *Safety regulation of AI*
- *Training data issues*
  - *Biases and data privacy*
  - *Data fragmentation*
- *Lack of sustainability*
  - *High-cost of training AI systems*
  - *Energy consumption, carbon footprint*
- *Closed systems*
  - *Big corporations manage data and models*
  - *Open-source foundation models needed*

Figures: Medium (https://medium.com/@businesswriting1012/biases-in-data-data-analytics-82ce481024e); Public domain

Uncertain predictions of the future

- *Progress in many narrow AI technologies and applications will continue*
- *Too high hopes could lead to the 3rd AI Winter*

- *Next 10 – 15 years, we will have AI systems that can …*
  - *Help humans in many tasks in daily living (e.g., personalized assistant) with multimodal interaction beyond smartphones (e.g., BCI technology)*
  - *Provide decision-support in work life and scientific discovery (creative work, software development)*
  - *Automate many tasks in limited operating environments (e.g., robotics, transportation)*
  - *Be implemented in distributed manner (e.g., 6G systems)*
  - *Provide novel methodologies from predictive to actionable ML*
  - *Combine neural and symbolic learning, planning, and reasoning*

- *Medicine and healthcare: personalized everything (monitoring, drug treatment etc.), operation optimization, preventive healthcare, clinical decision-support*

- *Many jobs will change their form (AI assisted)*

Uncertain predictions of the future

- *There are many issues in the current systems that need to be improved and solved in near future to accelerate progress*
  - *E.g., robustness, efficiency, and transparency of AI systems*

- *AGI is still far away (and possible ASI after that)*
  - *Although progress has been fast in recent years, we don't have tools, technology, or understanding in the horizon how to achieve this, probably, in next 50-100 years*
  - *No real evidence of technological singularity near future*
  - *We don't have to worry about human extinction near future*
  - *Some interesting questions*
    - *Could quantum computers and novel sensory systems with these big models and data accelerate development?*
    - *Can autonomous agents create evolution-based intelligent?*
    - *What kind of AGI/ASI would be possible? Would it be very different from human intelligence?*

“We can see only a short distance ahead, but we can see that much remains to be done”

Alan Turing (1950), Essay on *Computing Machinery and Intelligence.*