CS 480

Introduction to Artificial Intelligence

November 29, 2022

Announcements / Reminders

- **Final Exam: December 1st!**
  - **Ignore Registrar date for CS 480**
  - **Online section 02: please contact Mr. Charles Scott (scott@iit.edu) to make arrangements if necessary**

- **End of semester course evaluation: opened**

- **Written Assignment #04: due on Wednesday (11/30)**

Plan for Today

- **Casual Introduction to Machine Learning**

Main Machine Learning Categories

Supervised learning	Unsupervised learning	Reinforcement learning
Supervised learning is one of the most common techniques in machine learning. It is based on **known relationship(s) and patterns within data** (for example: relationship between inputs and outputs). Frequently used types: **regression**, and **classification**.	**Unsupervised learning** involves finding underlying patterns within data. Typically used in **clustering** data points (similar customers, etc.)	Reinforcement learning is inspired by behavioral psychology. It is **based on a rewarding / punishing an algorithm**. Rewards and punishments are based on algorithm's action within its environment.

Classifier Evaluation: Confusion Matrix

		Predicted class		
		Positive	Negative	
Actual class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{TP+FN}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{TN+FP}$
		Precision $\frac{TP}{TP+FP}$	Negative Predictive Value $\frac{TN}{TN+FN}$	Accuracy $\frac{TP+TN}{TP+TN+FP+FN}$

Reinforcement Learning (RL)

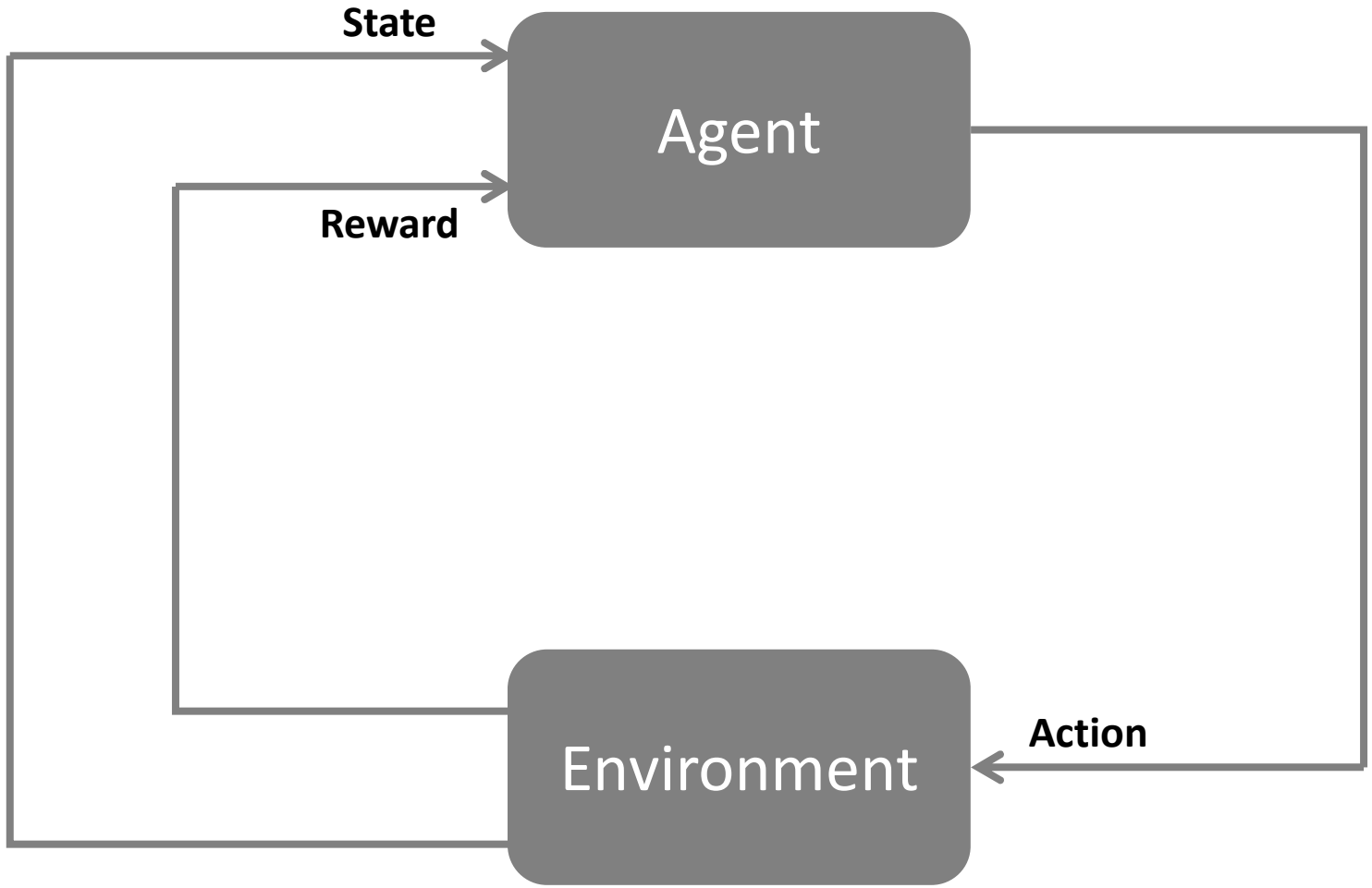
What is Reinforcement Learning?

Idea:

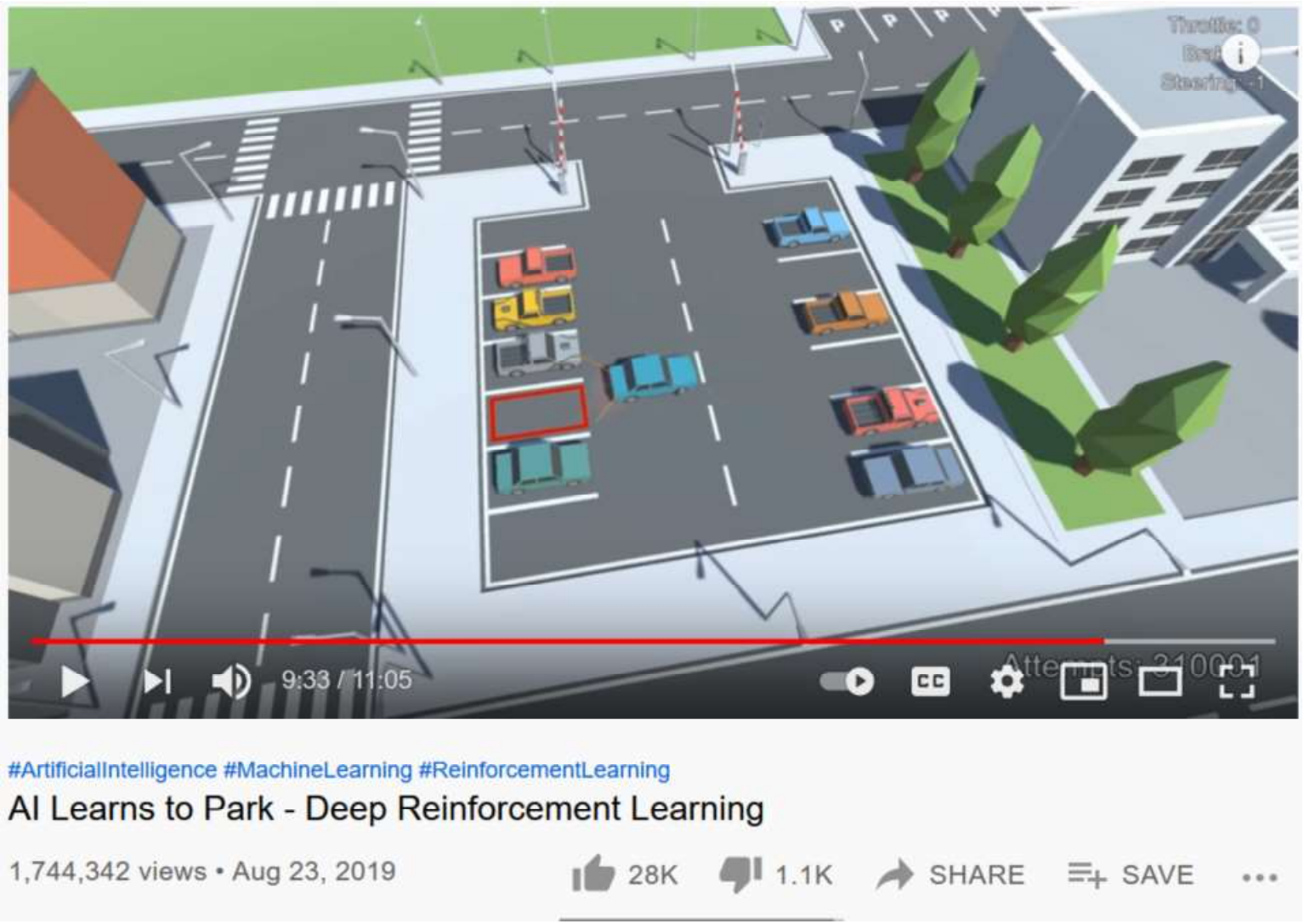
Reinforcement learning is inspired by behavioral psychology. It is **based on a rewarding / punishing an algorithm**.

Rewards and punishments are based on algorithm's action within its environment.

RL: Agents and Environments

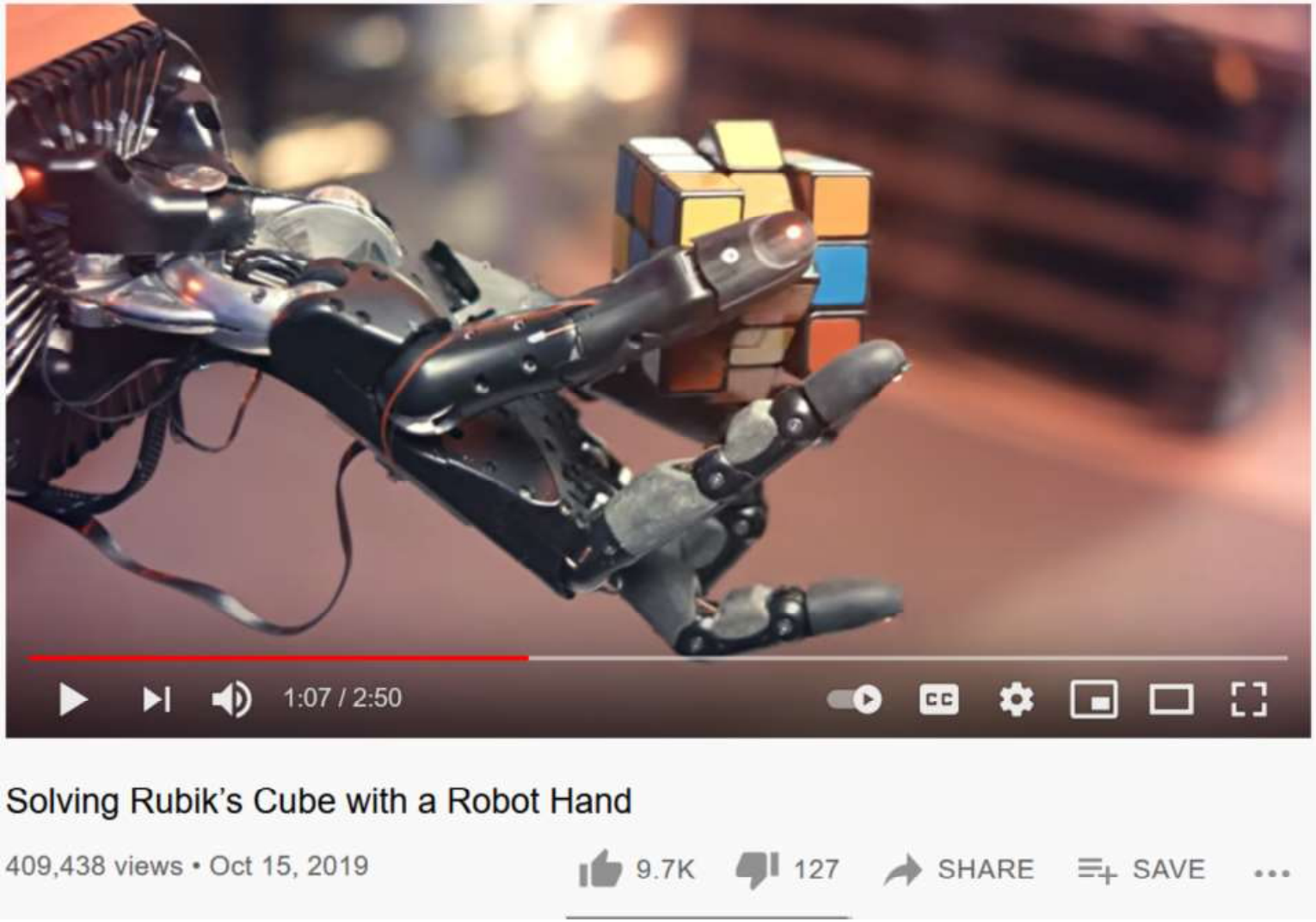


Reinforcement Learning in Action



Source: https://www.youtube.com/watch?v=VMp6pq6_QjI

Reinforcement Learning in Action



Source: https://www.youtube.com/watch?v=x4O8pojMF0w

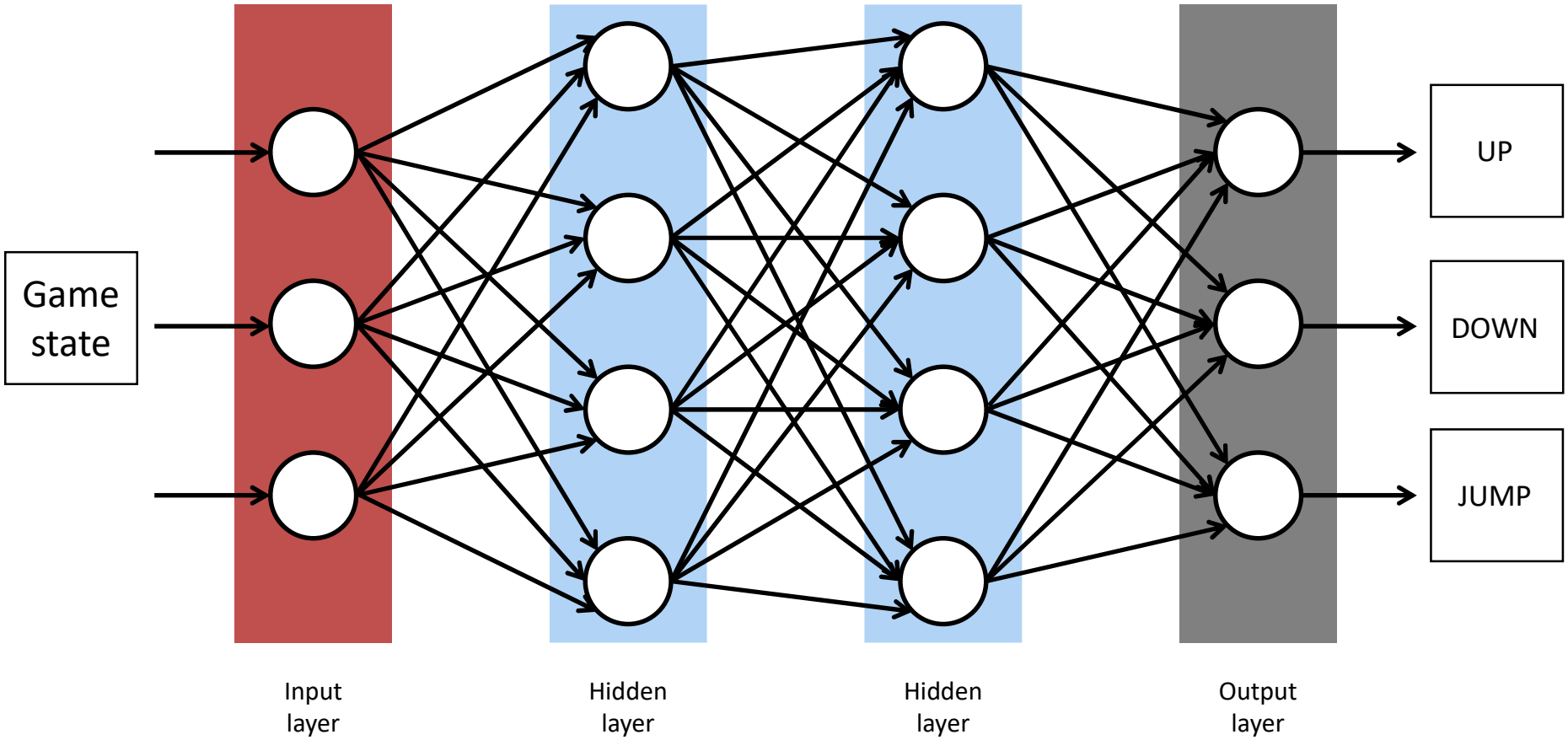
Reinforcement Learning in Action

Source: https://www.youtube.com/watch?v=kopoLzvh5jY

Reinforcement Learning in Action

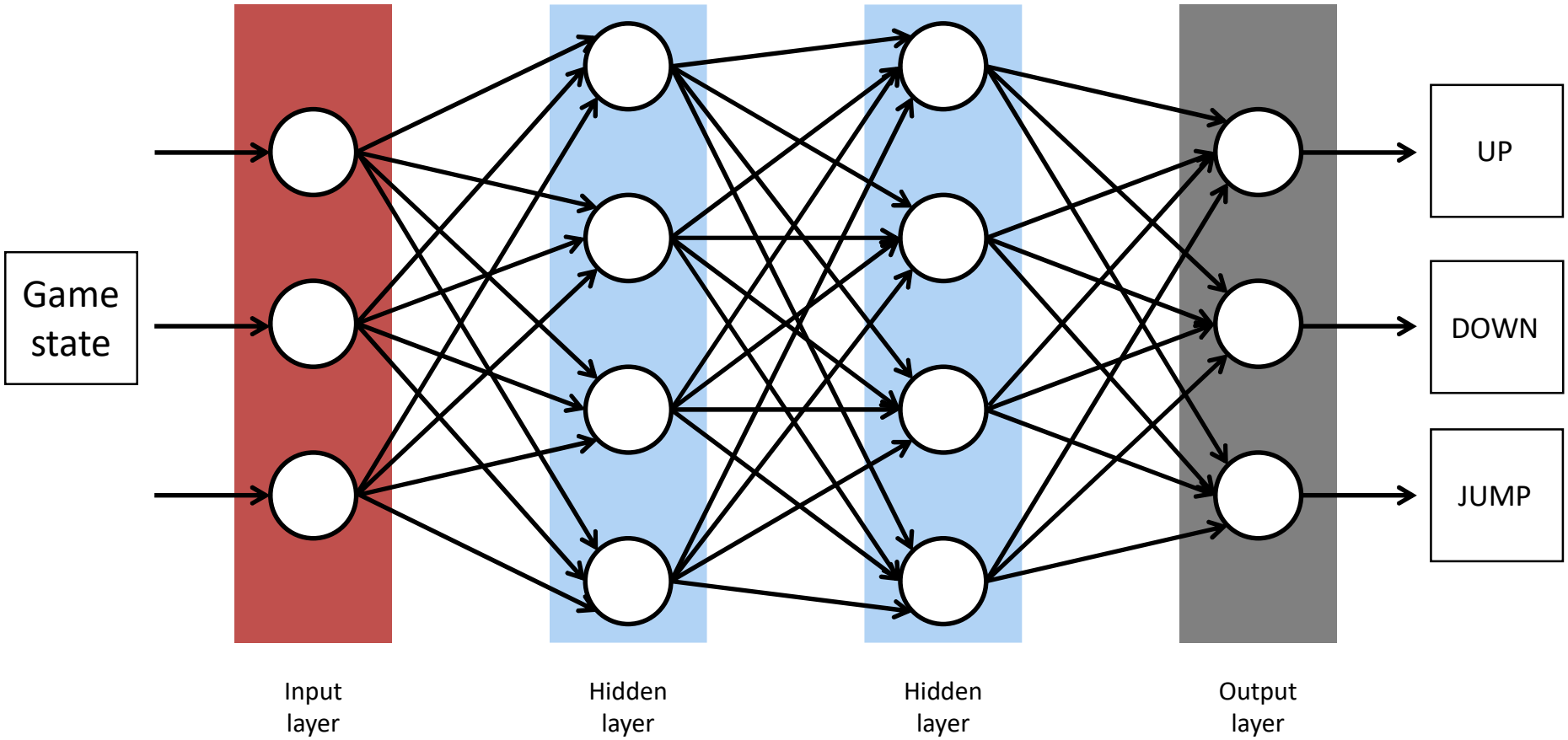
Source: https://www.youtube.com/watch?v=Tnu4O_xEmVk

ANN for Simple Game Playing



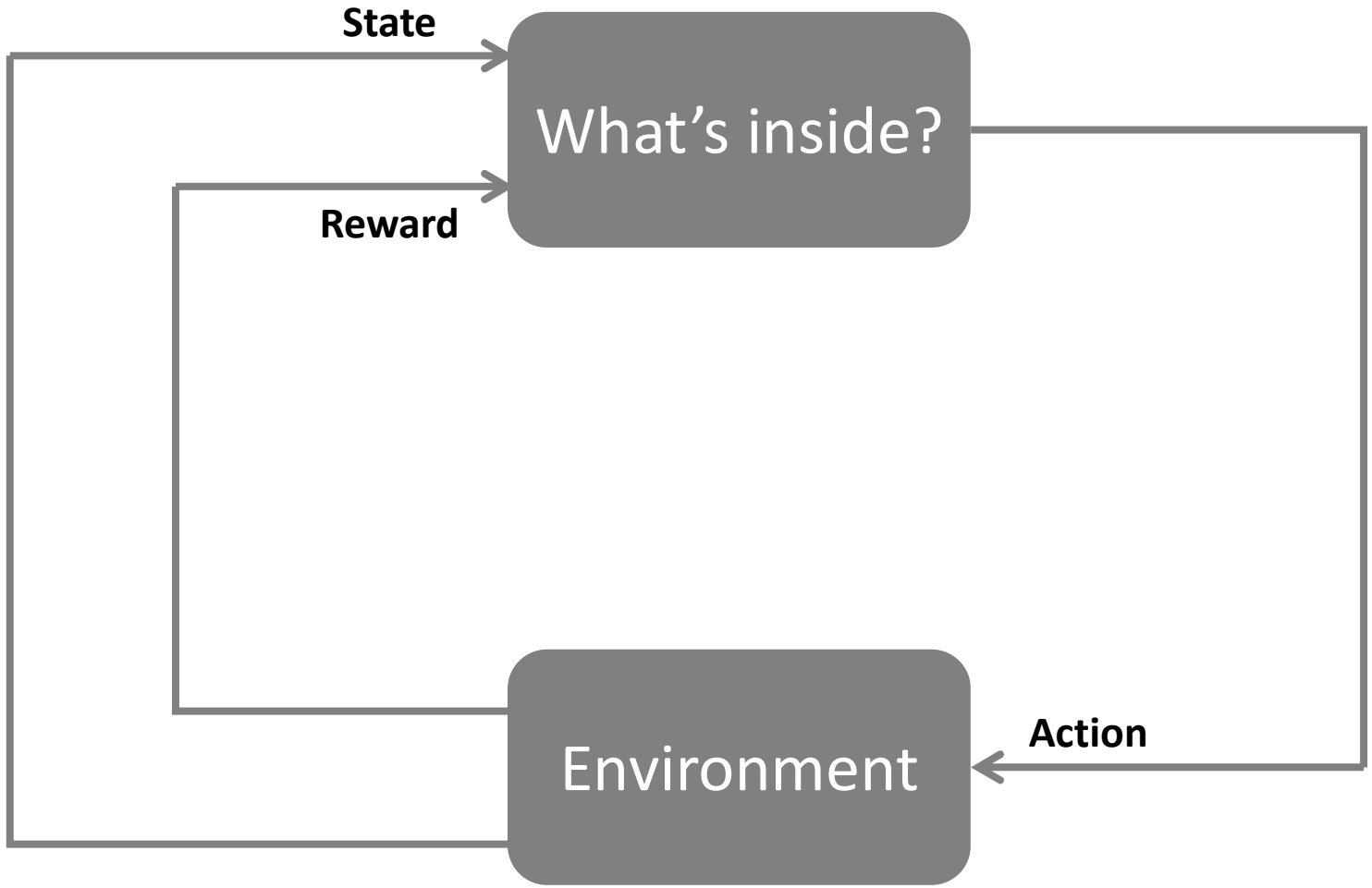
ANN for Simple Game Playing

Current game is an input. Decisions (UP/DOWN/JUMP) are rewarded/punished.

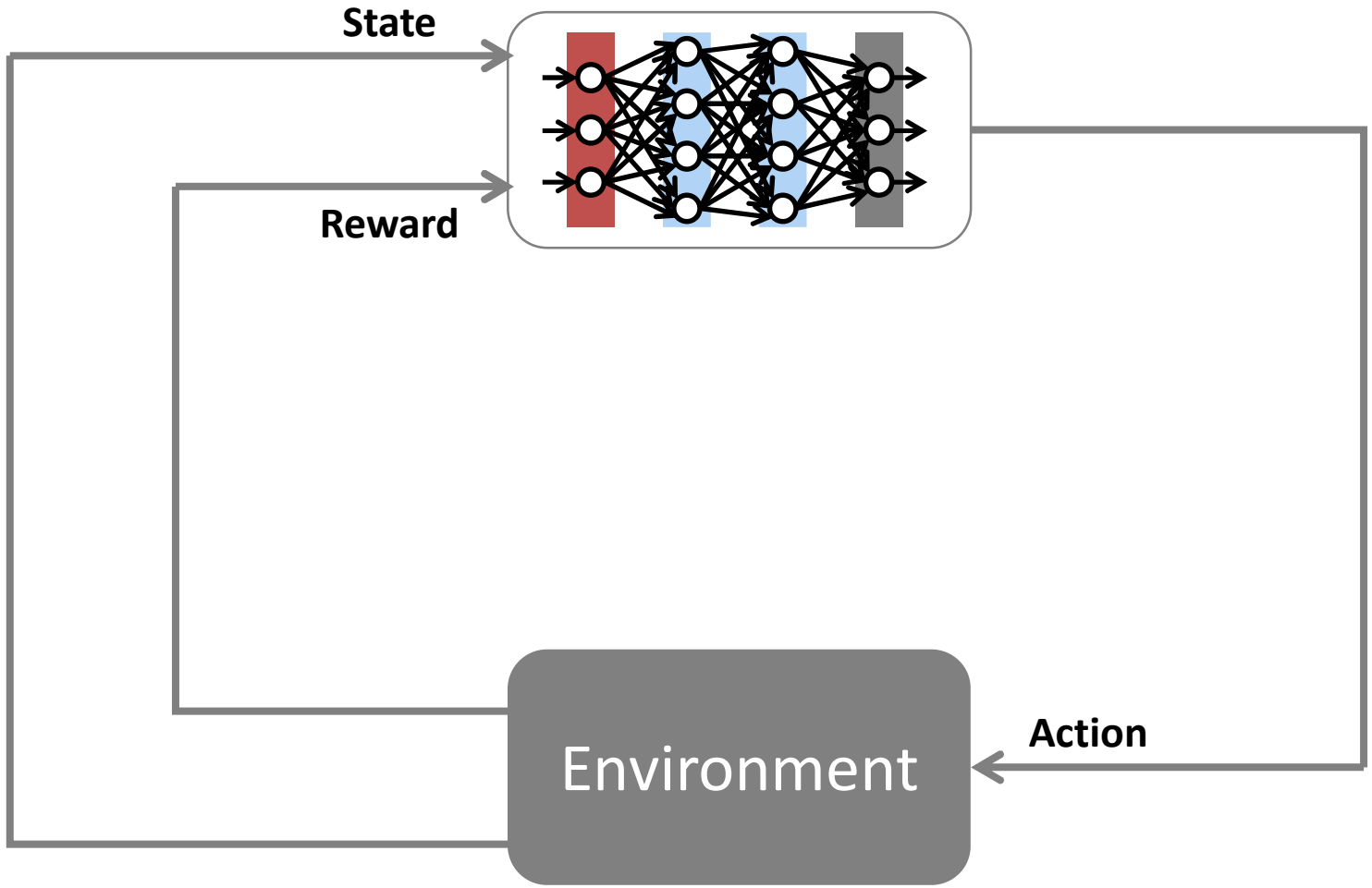


Correct all the weights using Reinforcement Learning.

RL: Agents and Environments



RL: Agents and Environments

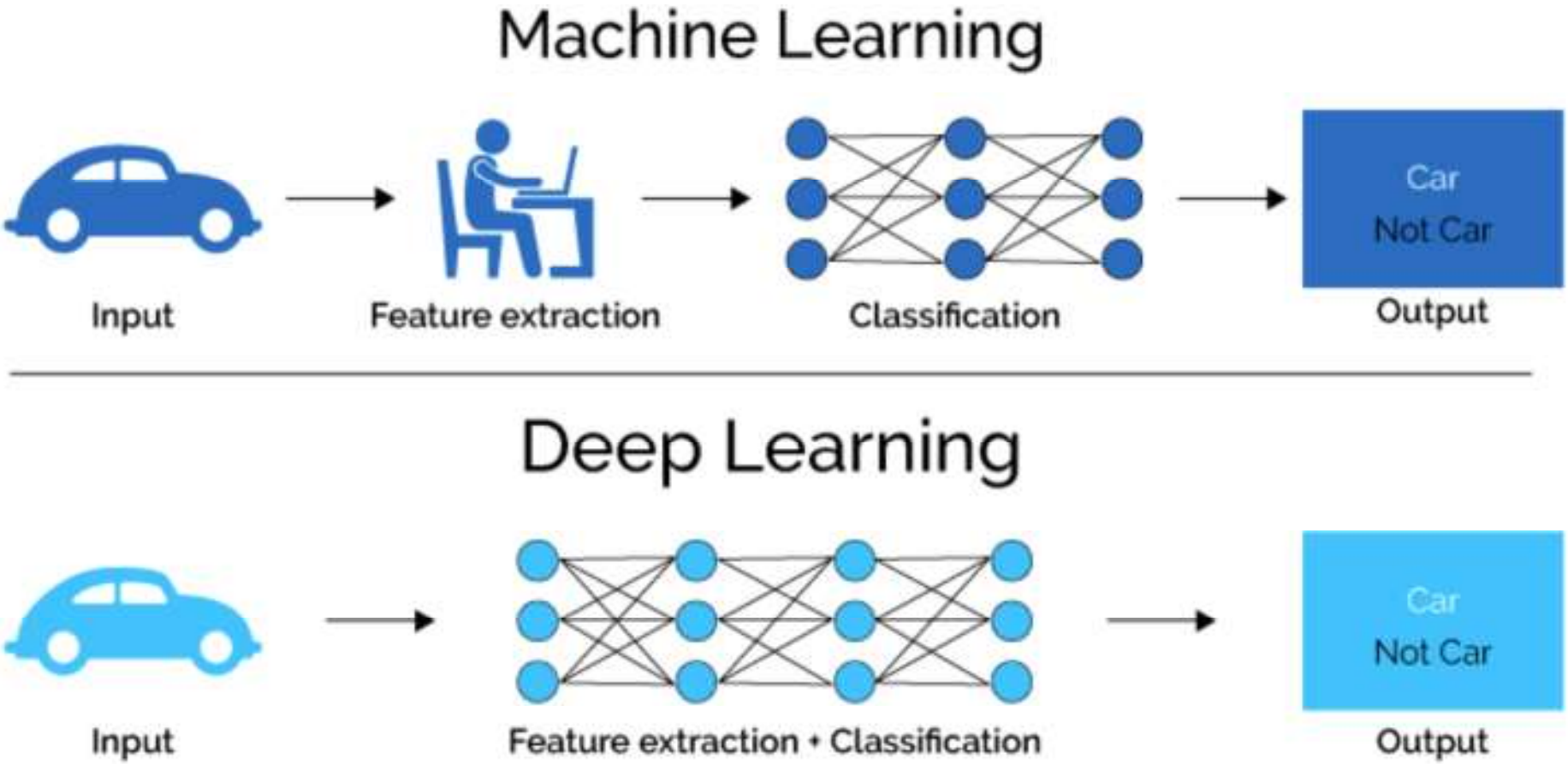


Deep Learning

Deep Learning

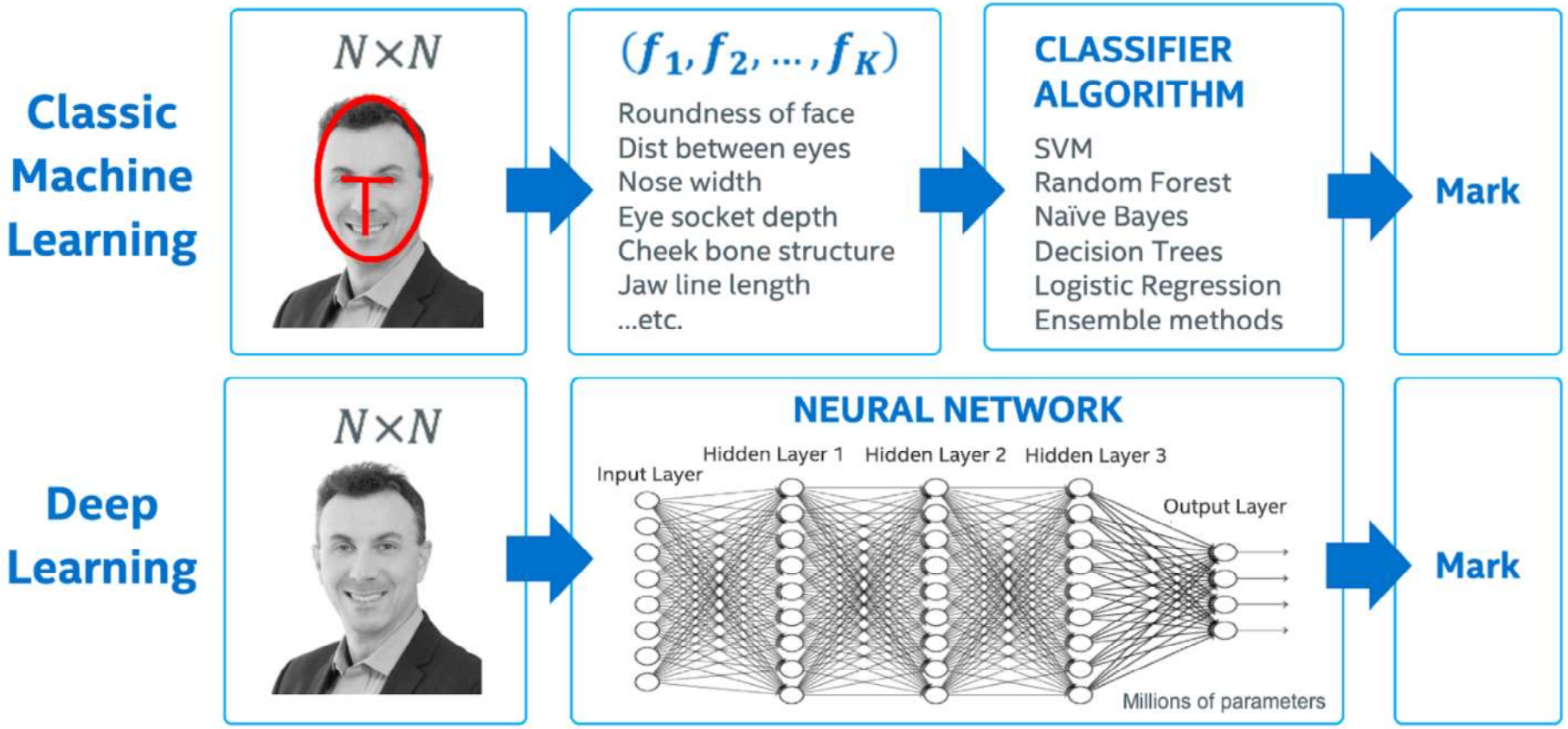
Deep learning is a broad family of techniques for machine learning (also a sub-field of ML) in which hypotheses take the form of complex algebraic circuits with tunable connections. The word "deep" refers to the fact that the circuits are typically organized into many layers, which means that computation paths from inputs to outputs have many steps.

Machine Learning vs. Deep Learning



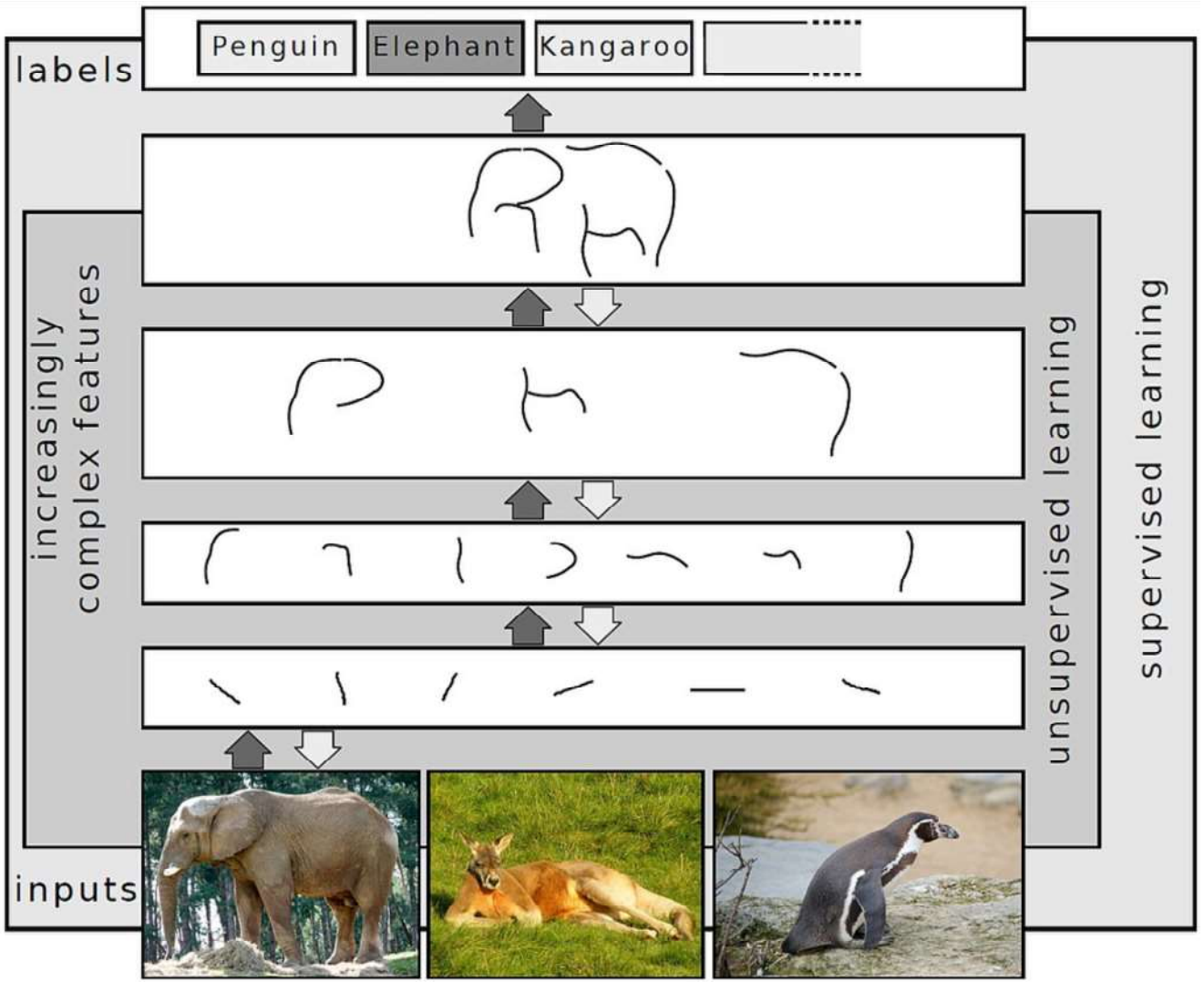
Source: https://www.quora.com/What-is-the-difference-between-deep-learning-and-usual-machine-learning

Machine Learning vs. Deep Learning



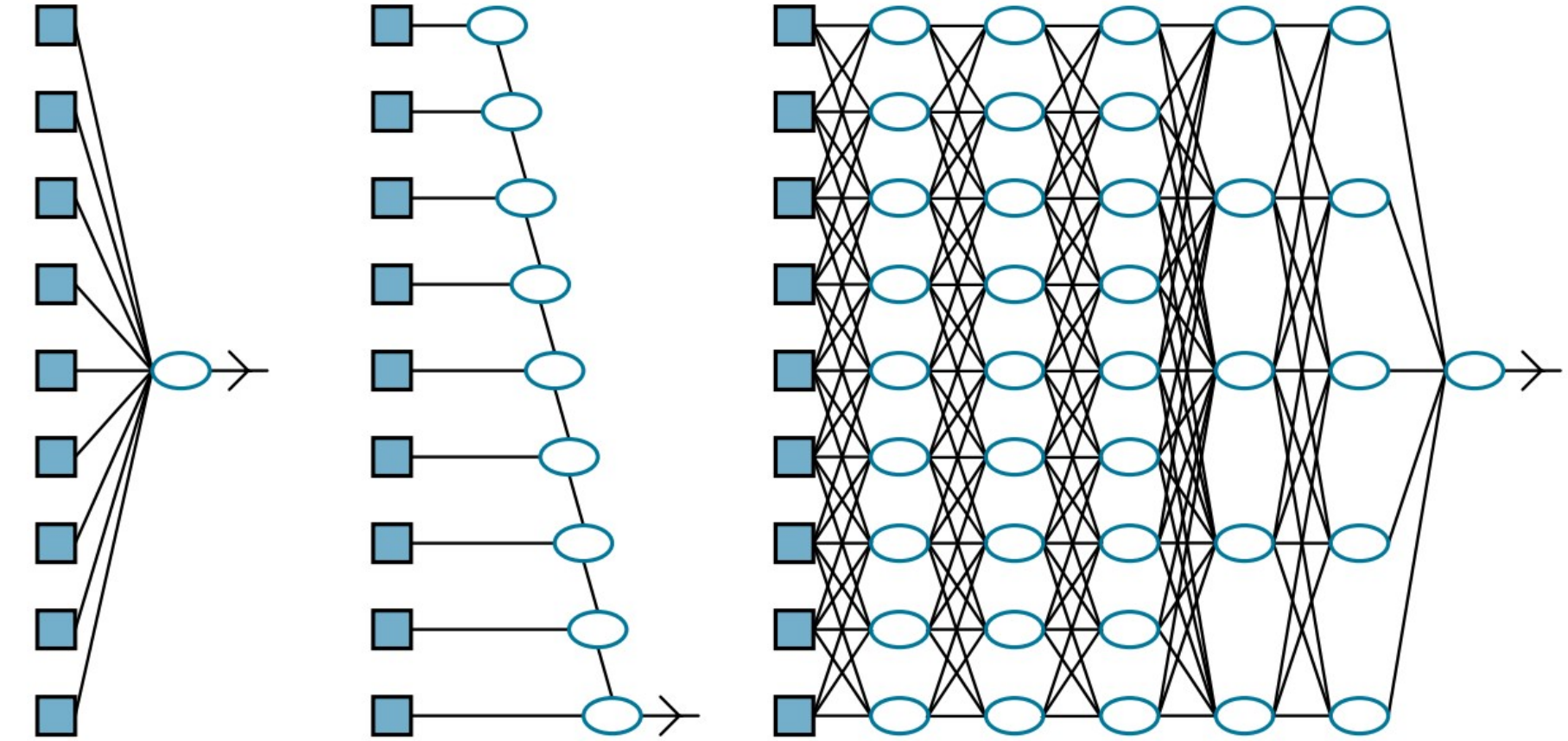
Source: https://www.intel.com/content/www/us/en/artificial-intelligence/posts/difference-between-ai-machine-learning-deep-learning.html

Deep Learning: Feature Extraction



Source: https://en.wikipedia.org/wiki/Deep_learning

Shallow vs. Deep Models



Shallow Model

Shallow Model

Deep Model

Longer computation path

Convolutional Neural Networks

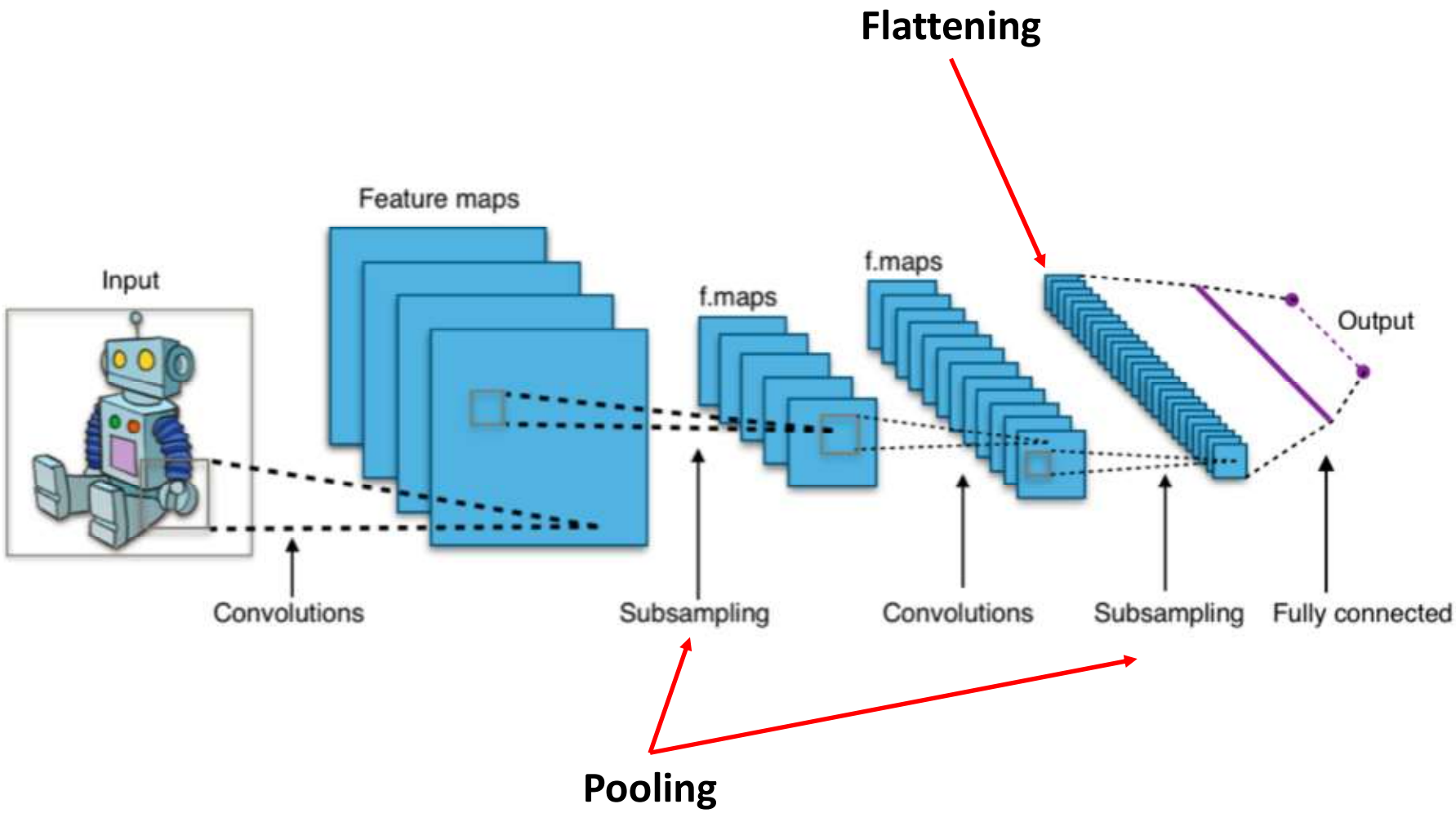
The name Convolutional Neural Network (CNN) indicates that the network employs a mathematical operation called convolution.

Convolutional networks are a specialized type of neural networks that use convolution in place of general matrix multiplication in at least one of their layers.

CNN is able to successfully capture the spatial dependencies in an image (data grid) through the application of relevant filters.

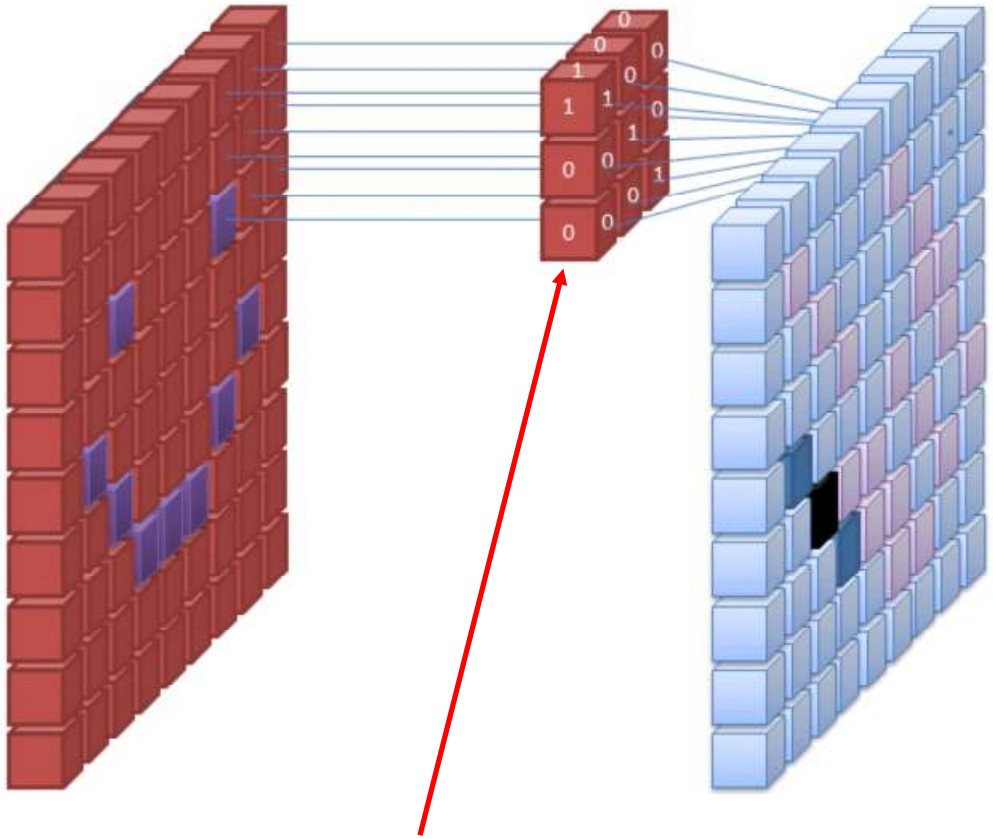
CNNs can reduce images (data grids) into a form which is easier to process without losing features that are critical for getting a good prediction.

Convolutional Neural Networks



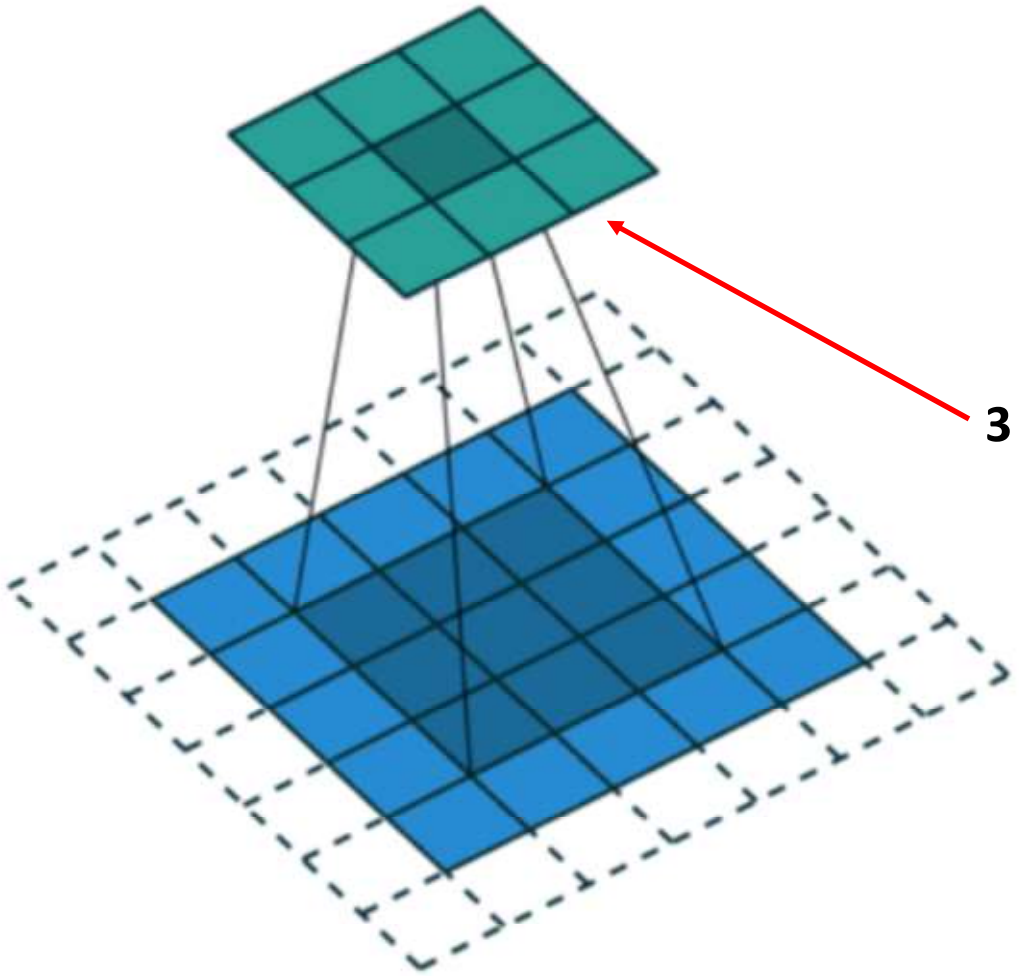
By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45679374

Convolution: The Idea



Source: https://commons.wikimedia.org/wiki/File:Convolutional_Neural_Network_NeuralNetworkFilter.gif

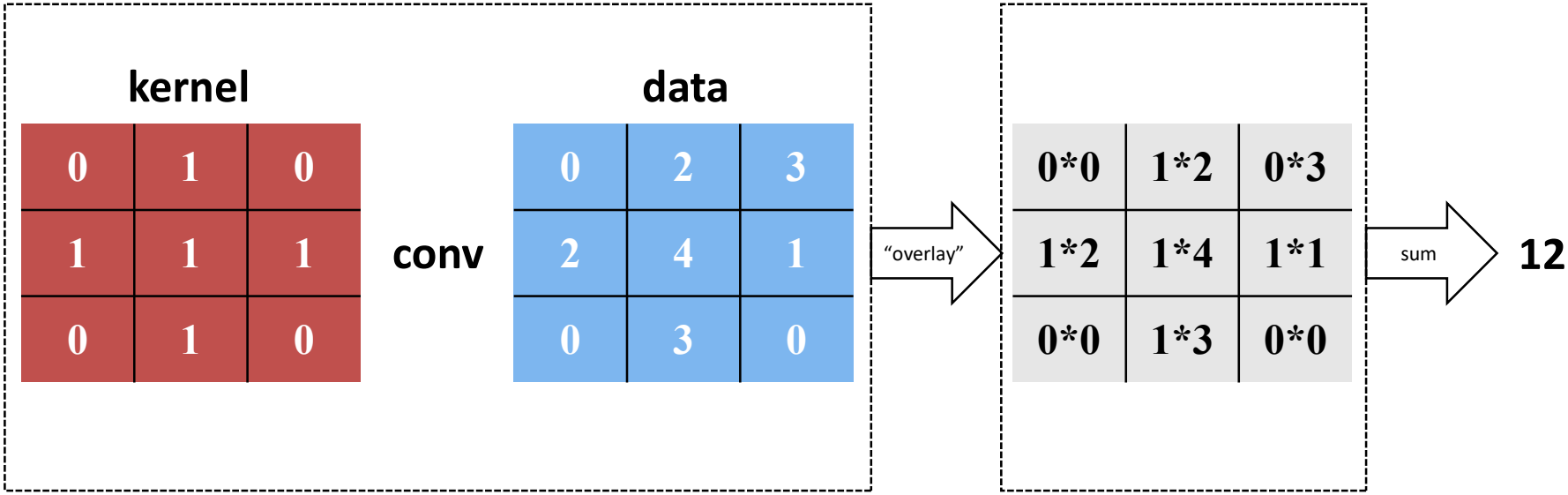
Kernel / Filter: The Idea



Source: https://commons.wikimedia.org/wiki/File:Convolution_arithmetic_-_Padding_strides.gif

Convoluting Matrices

Convolution (and Convolutional Neural Networks) can be applied to any grid-like data (tensors: matrices, vectors, etc.).



Selected Image Processing Kernels

Sharpen

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Mean Blur

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

Gaussian Blur

$$\begin{bmatrix} 1/16 & 2/16 & 1/16 \\ 1/16 & 4/16 & 2/16 \\ 1/16 & 2/16 & 1/16 \end{bmatrix}$$

Laplacian

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Prewitt (Edge)

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

Prewitt (Edge)

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

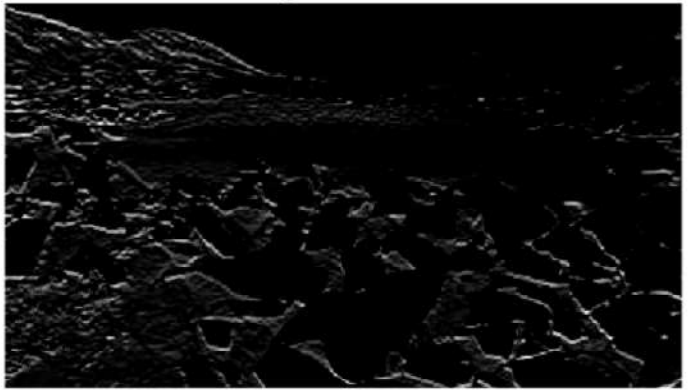
Image Processing: Kernels / Filters

Original

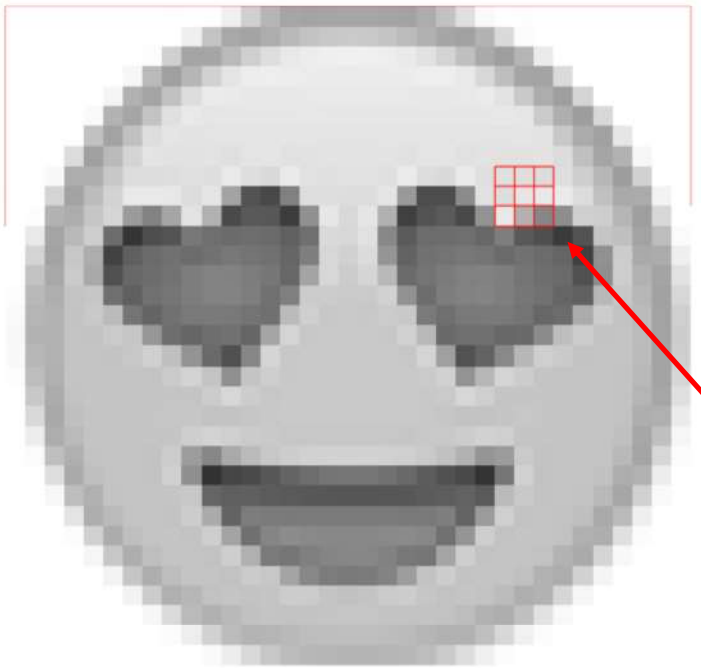
Sobel

Gaussian Blur

Edge detection



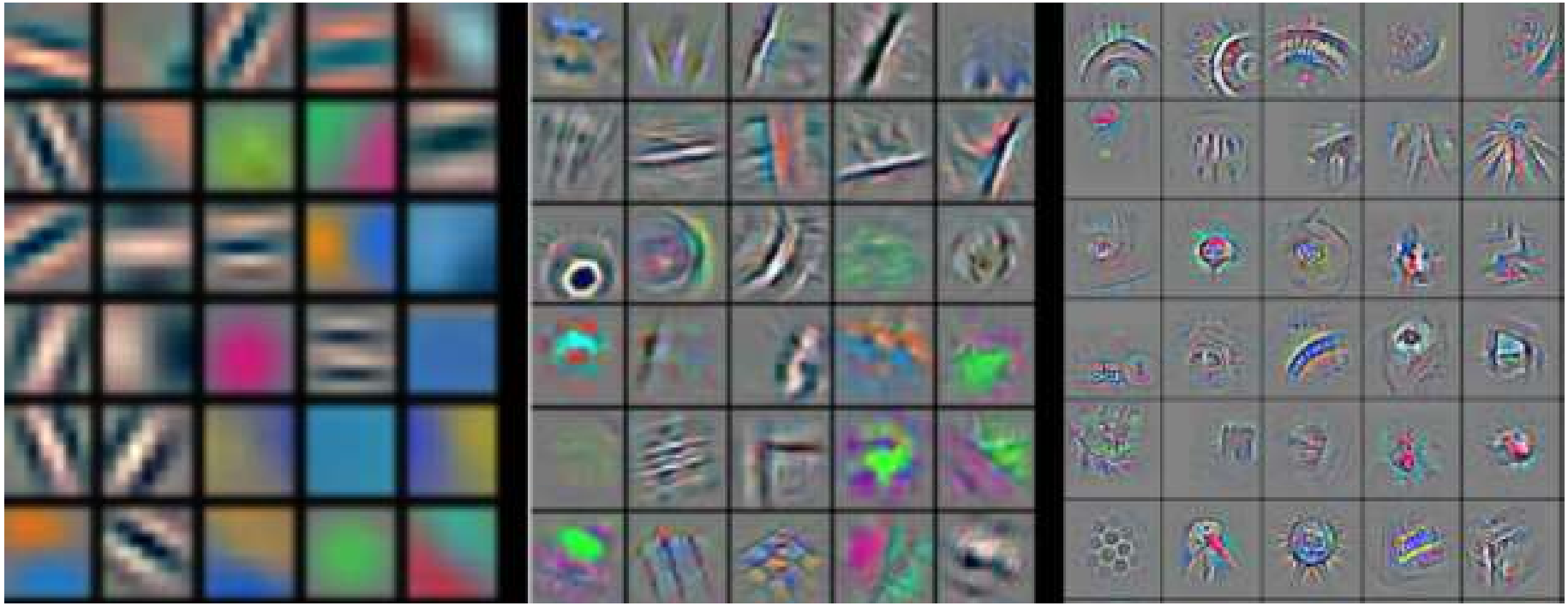
Applying Kernels / Filters



3 x 3 Kernel / Filter

Convolutional NN Kernels

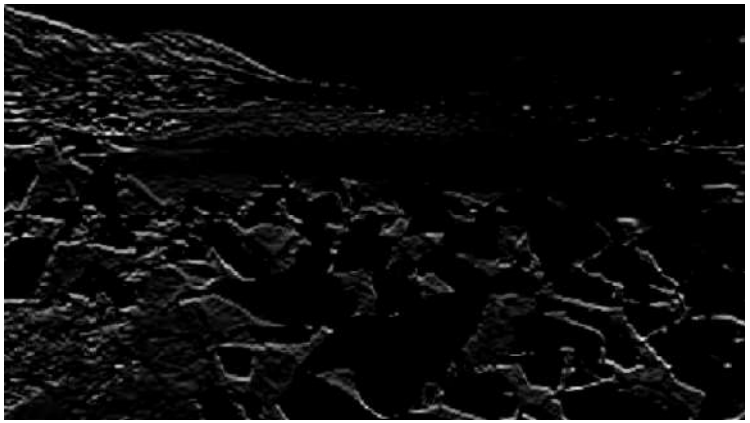
In practice, Convolutional Neural Network kernels can be larger than 3x3 and are learned using back propagation.



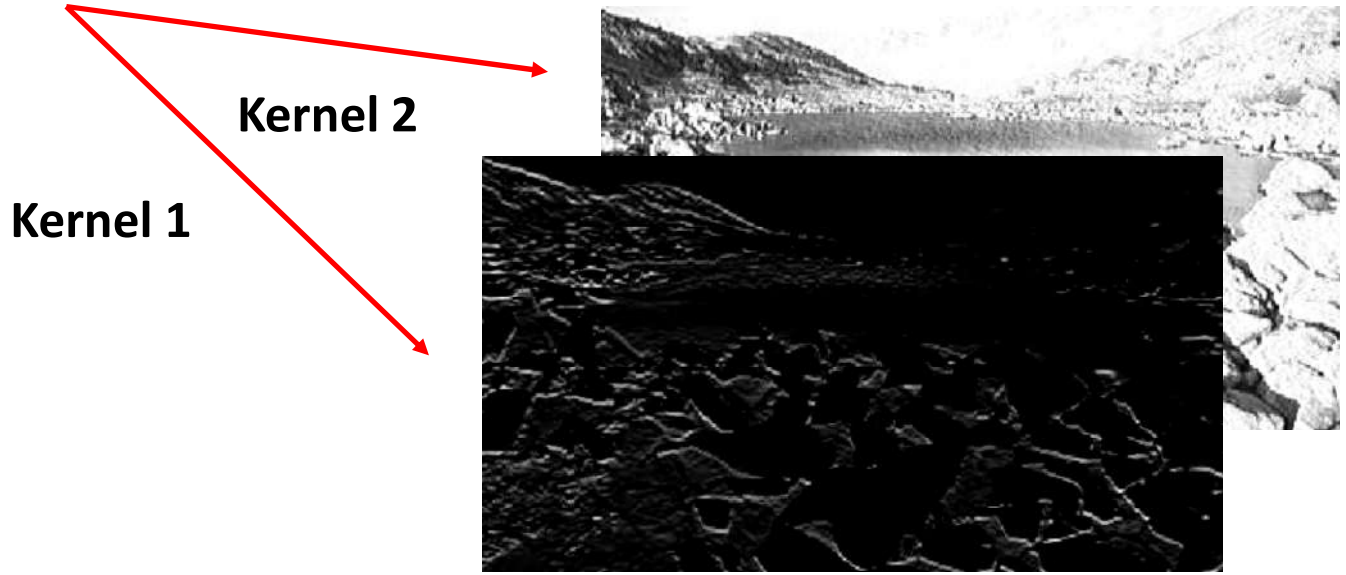
Convolution Layer 1 **Convolution Layer 2** **Convolution Layer 3**

Convolution Layer 1

Kernel 1



Convolution Layer 1



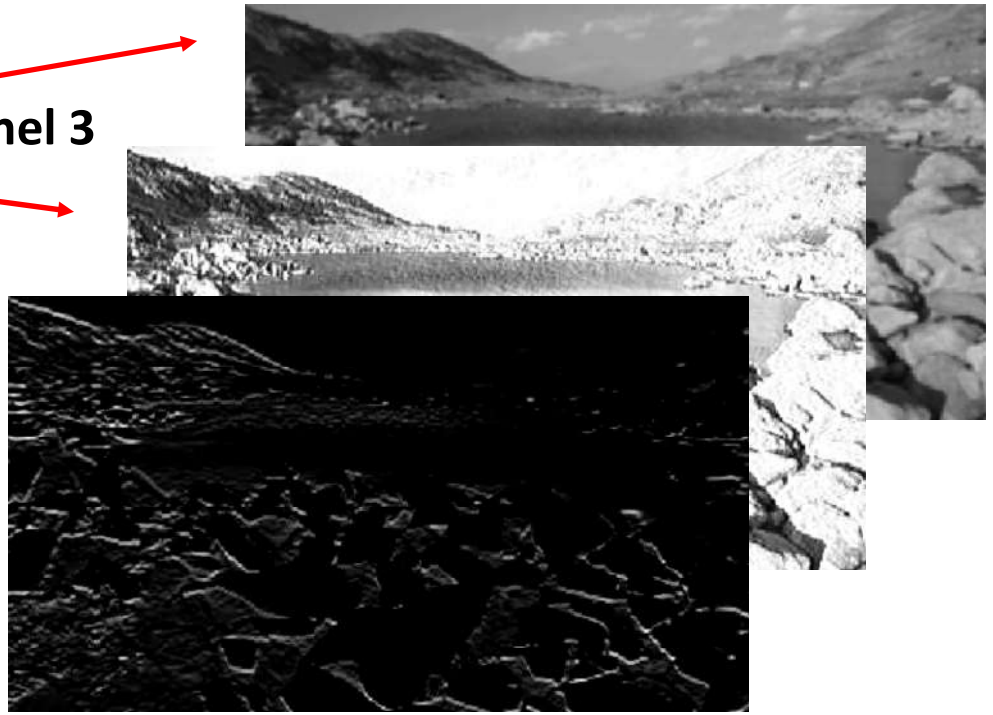
Convolution Layer 1

Original image

Kernel 3

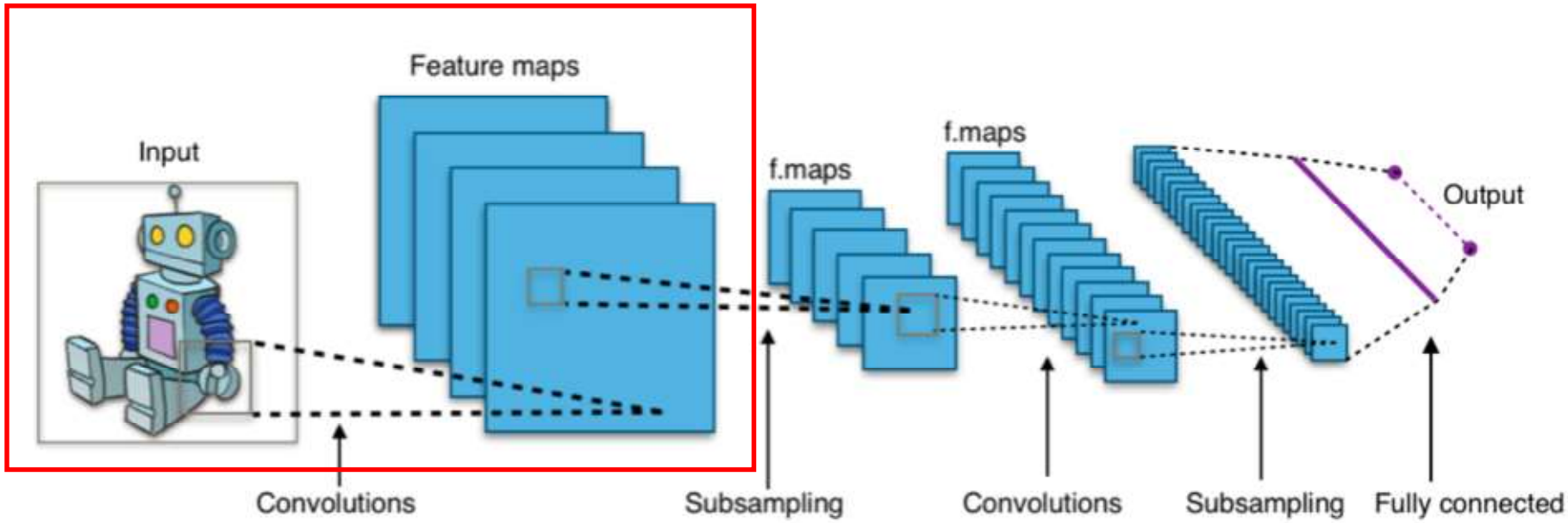
Kernel 2

Kernel 1



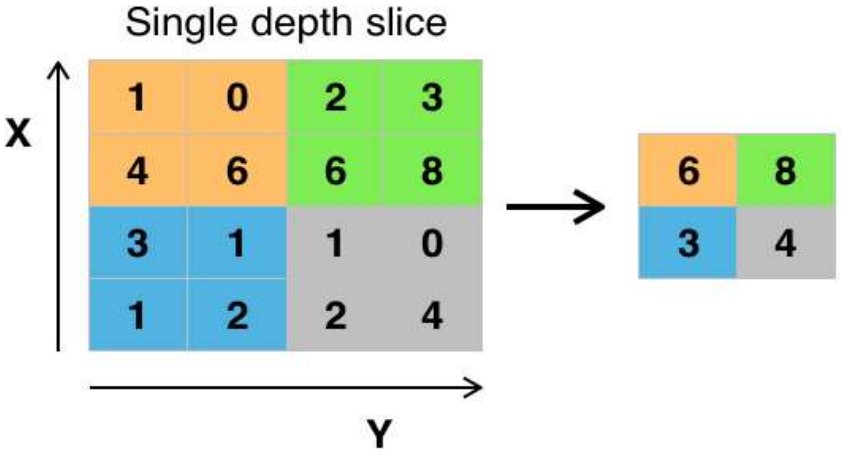
Convolution 1

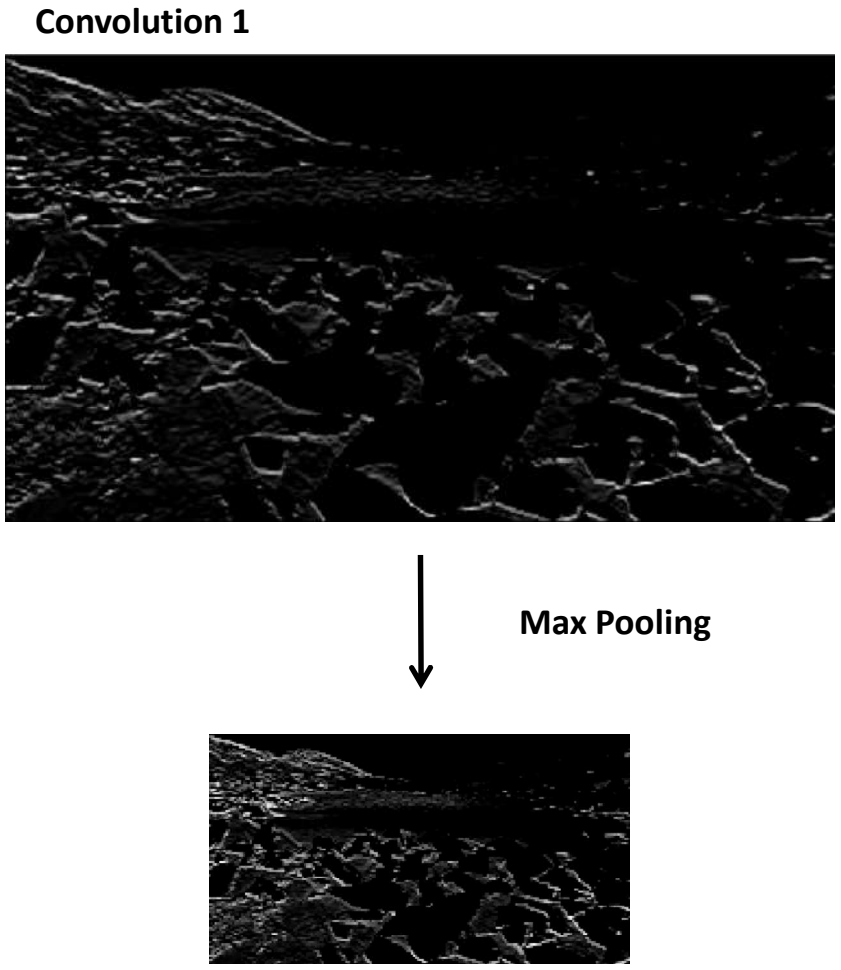
Convolutional Neural Networks



By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45679374

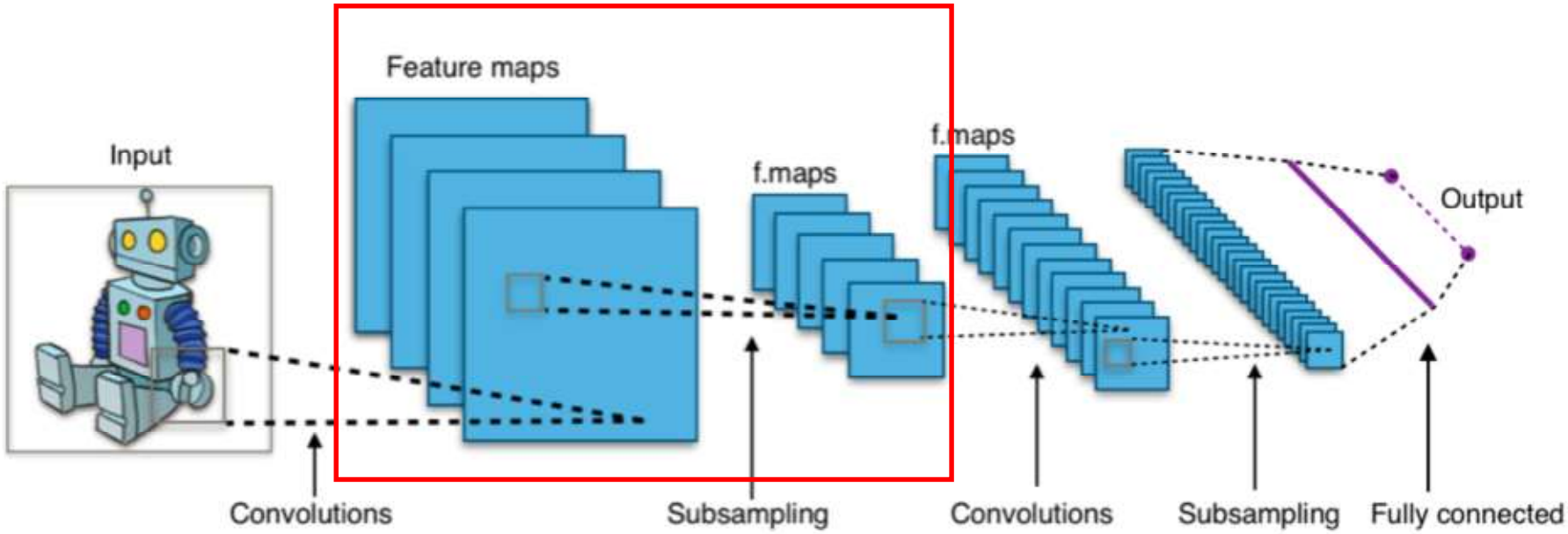
Max Pooling Layer





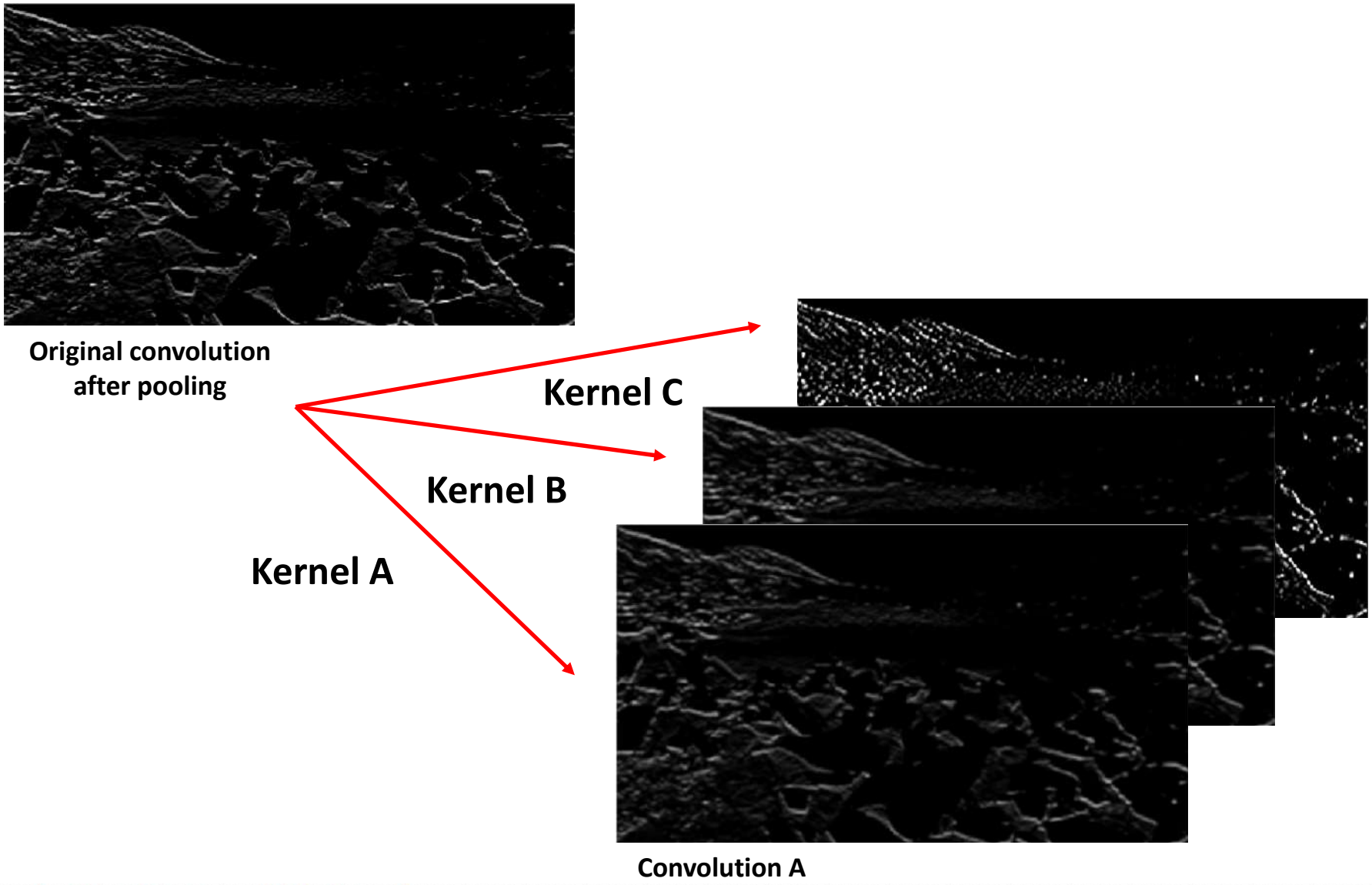
By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45673581

Convolutional Neural Networks

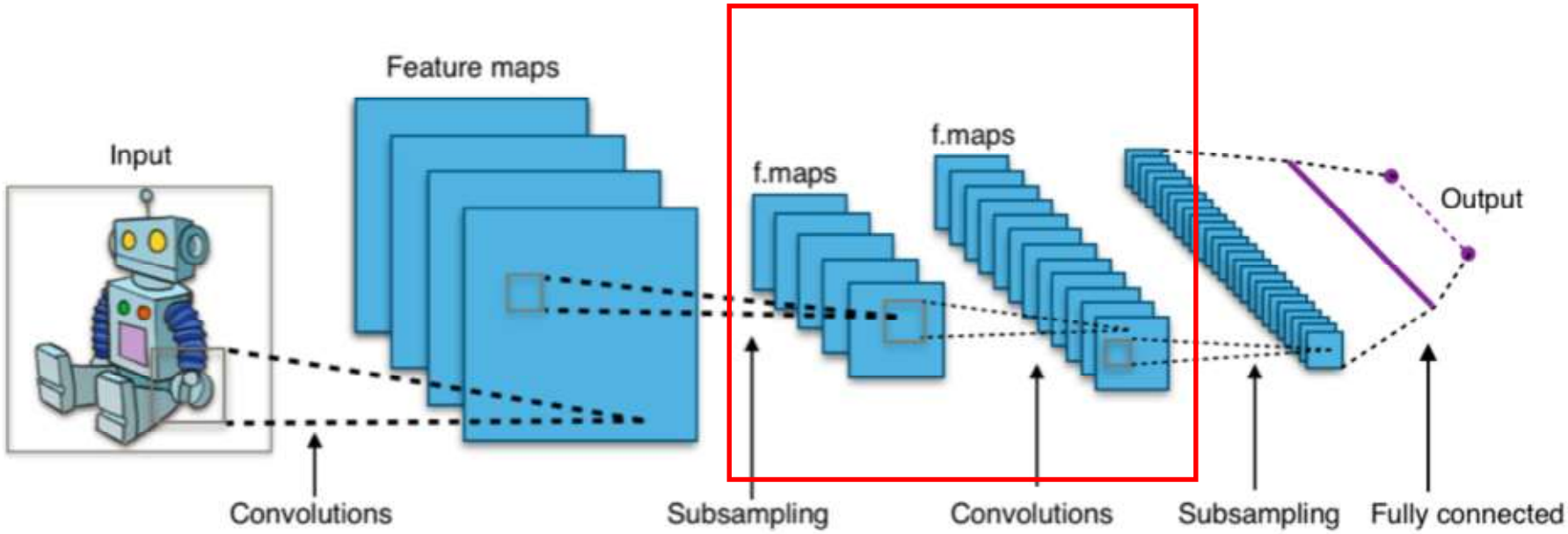


By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45679374

Convolution Layer 2



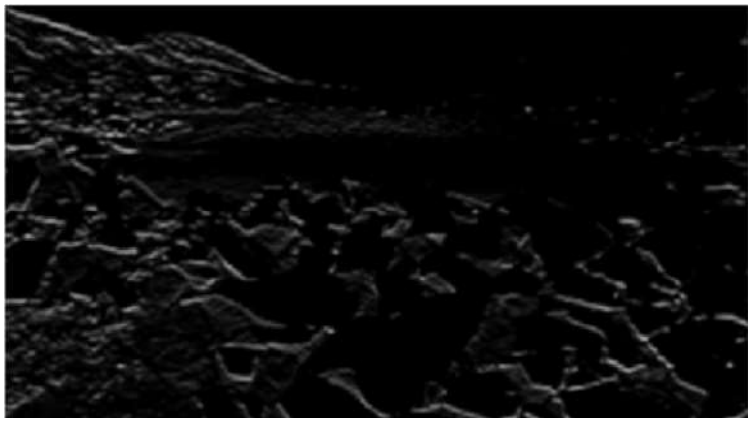
Convolutional Neural Networks



By Aphex34 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=45679374

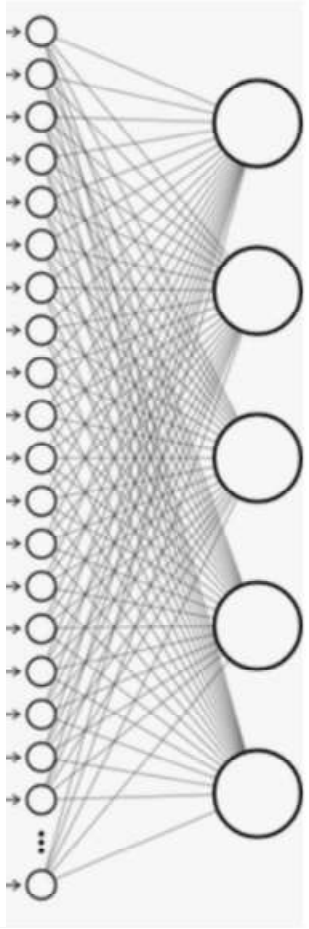
Flattening

Final output of convolution layers is "**flattened**" to become **a vector of features**.



Final convolution layer output

Convert to vector



Source: https://nikolanews.com/not-just-introduction-to-convolutional-neural-networks-part-1/

Recurrent Neural Networks

Recurrent Neural Networks (RNNs) allow cycles in the computational graph (network). A **network node (unit) can take its own output from an earlier step as input** (with delay introduced).

Enables **having internal state / memory** → inputs received earlier affect the RNN response to current input.

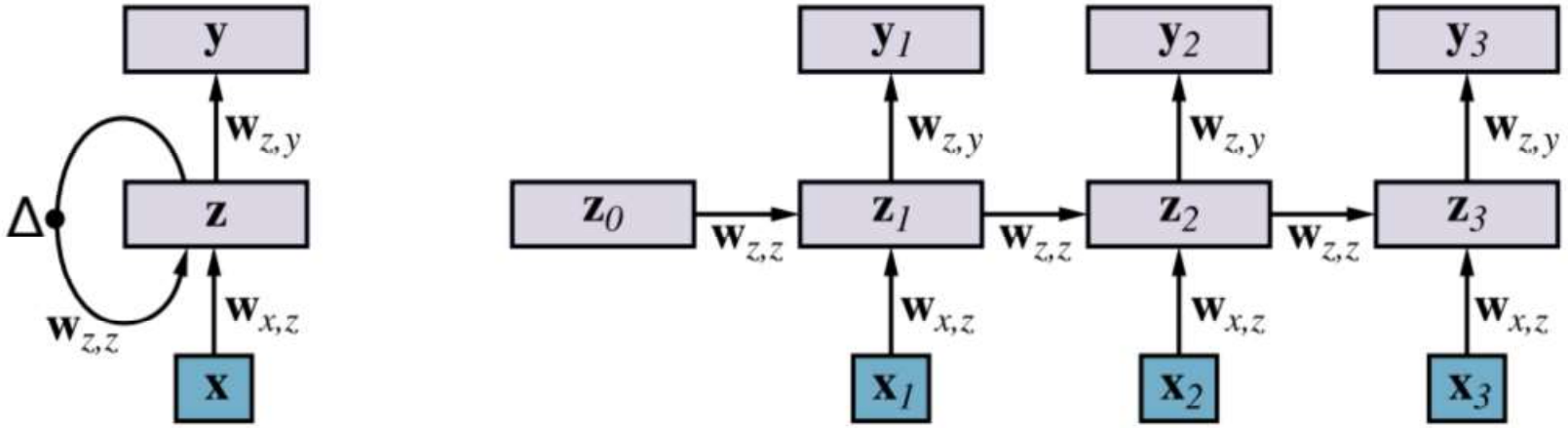


Figure (a) Schematic diagram of a basic RNN where the hidden layer **z** has recurrent connections; the Δ symbol indicates a delay. (b) The same network unrolled over three time steps to create a feedforward network. Note that the weights are shared across all time steps.

Transfer Learning

In transfer learning, experience with one learning task helps an agent learn better on another task.

Pre-trained models can be used as a starting point for developing new models.

Generative AI

GPT-3 Scripted Movie

Solicitors | A.I. Written Short Film

36,582 views • Oct 13, 2020 668 27 SHARE SAVE ...

Source: https://www.youtube.com/watch?v=AmX3GDJ47wo

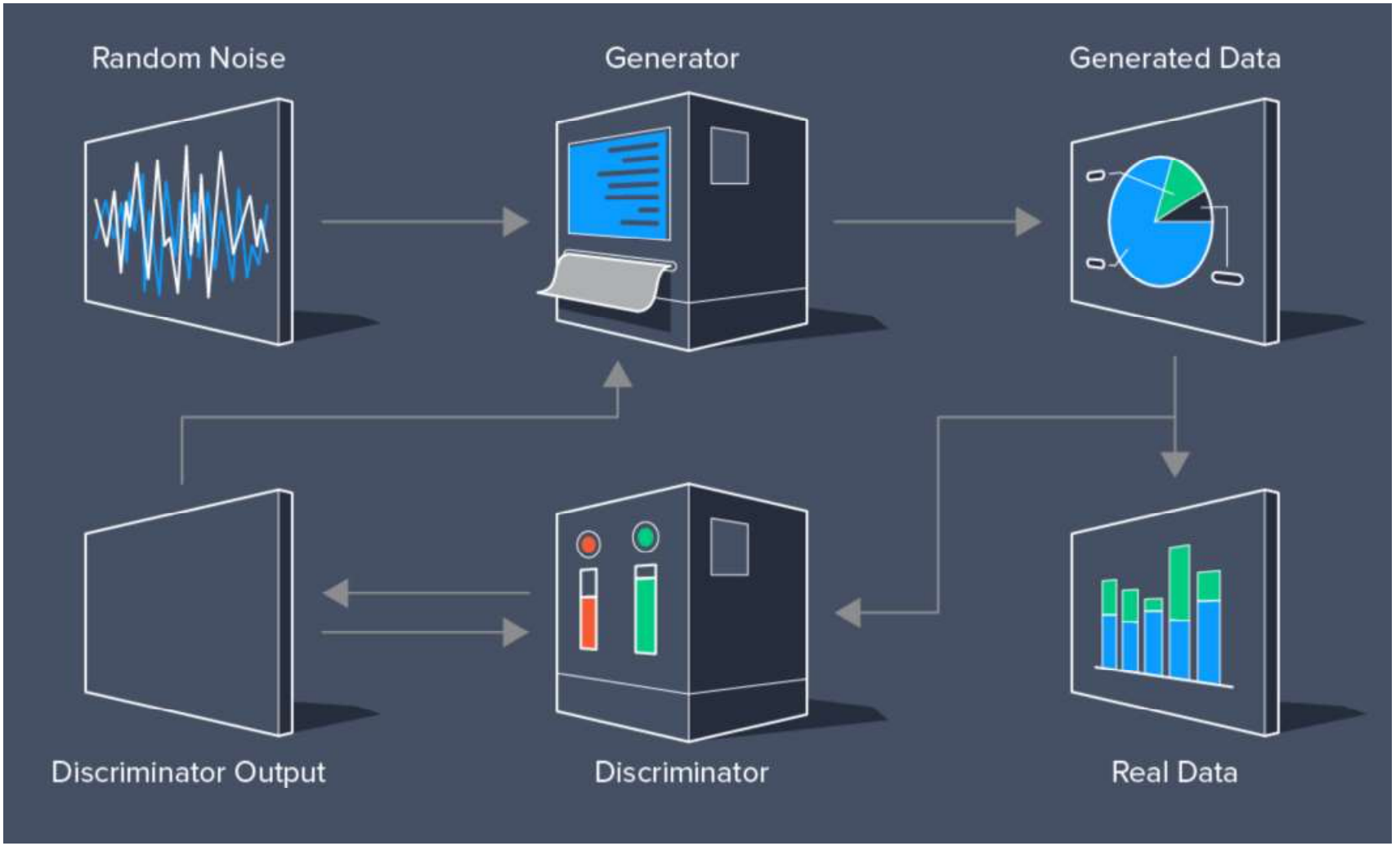
Lost Tapes of the 27 Club: AI Music

The album itself was created in collaboration with **AI generating new lyrics and musical compositions with the input data of the hooks, rhythms, melodies, and lyrics of artists from the 27 club**.

From there, **a team of audio engineers and technicians worked to parse out the signal from the noise** to create cohesive tracks. They then **reached out to talented singers from tribute bands to fill in the vocal parts with lyrics prewritten by the AI**.

- Nirvana-Drowned in the Sun: https://www.youtube.com/watch?v=muT6x7VXx5I
- Amy Winehouse-Man I Know: https://www.youtube.com/watch?v=QM6LbbcCghc
- Jimi Hendrix-You're Going To Kill Me: https://www.youtube.com/watch?v=6Ohf97p7u1w
- The Doors-The Roads Are Alive: https://www.youtube.com/watch?v=z5jW4RmxhIY

Generative Deep Learning



Source: https://www.toptal.com/machine-learning/generative-adversarial-networks

Exercise: Generative AI

https://ai-art.tokyo/en/#/

Deep Fakes

Source: https://www.youtube.com/watch?v=bPhUhypV27w

Deep Fakes

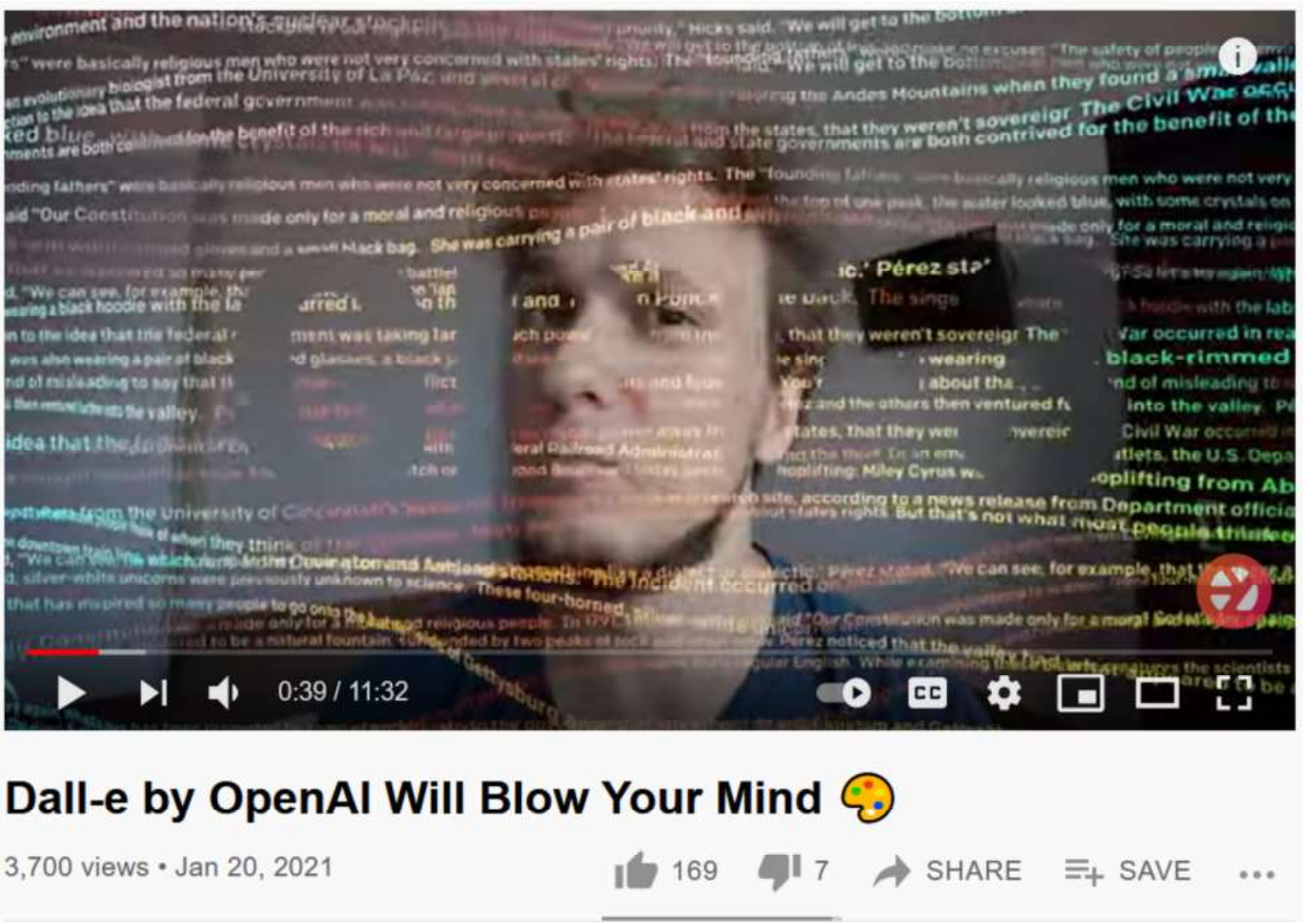
Mike Tyson and Snoop Dogg as Oprah and Gayle [Deepfake

109,077 views • Aug 11, 2019 1.4K 24 SHARE SAVE

Source: https://www.youtube.com/watch?v=LRMnNpVjH6g

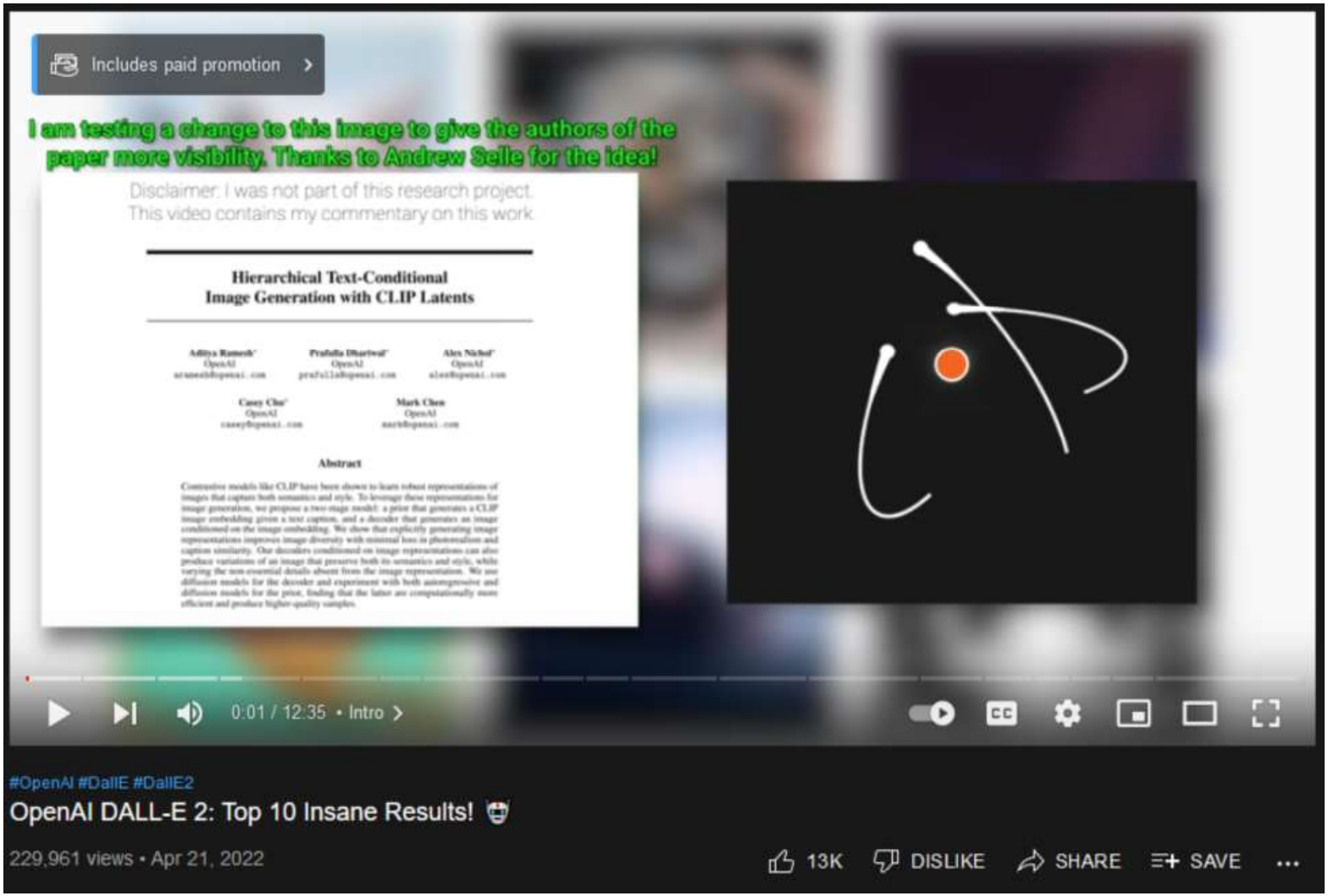
AI: Recent Technological Developments

Dall-E



Source: https://www.youtube.com/watch?v=HJBubmr--8Y

Dall-E 2



Source: https://www.youtube.com/watch?v=X3_LD3R_Ygs

Dall-E 2

Source: https://www.youtube.com/watch?v=U1cF9QCu1rQ

Dall-E 2 Explained

Source: https://www.youtube.com/watch?v=qTgPSKKjfVg

Google TensorFlow 3D

Google AI Blog

The latest from Google Research

3D Scene Understanding with TensorFlow 3D

Thursday, February 11, 2021

Posted by Alireza Fathi, Research Scientist and Rui Huang, AI Resident, Google Research

The growing ubiquity of 3D sensors (e.g., Lidar, depth sensing cameras and radar) over the last few years has created a need for scene understanding technology that can process the data these devices capture. Such technology can enable machine learning (ML) systems that use these sensors, like autonomous cars and robots, to navigate and operate in the real world, and can create an improved augmented reality experience on mobile devices. The field of computer vision has recently begun making good progress in 3D scene understanding, including models for mobile 3D object detection, transparent object detection, and more, but entry to the field can be challenging due to the limited availability tools and resources that can be applied to 3D data.

In order to further improve 3D scene understanding and reduce barriers to entry for interested researchers, we are releasing TensorFlow 3D (TF 3D), a highly modular and efficient library that is designed to bring 3D deep learning capabilities into TensorFlow. TF 3D provides a set of popular

Source: https://ai.googleblog.com/2021/02/3d-scene-understanding-with-tensorflow.html

Meta / Facebook SEER

FACEBOOK AI Research Publications People

SEER: The start of a more powerful, flexible, and accessible era for computer vision

March 4, 2021

Facebook AI has now brought this self-supervised learning paradigm shift to computer vision. We've developed SEER (SElf-supERvised), a new billion-parameter self-supervised computer vision model that can learn from any random group of images on the internet — without the need for careful curation and labeling that goes into most computer vision training today.

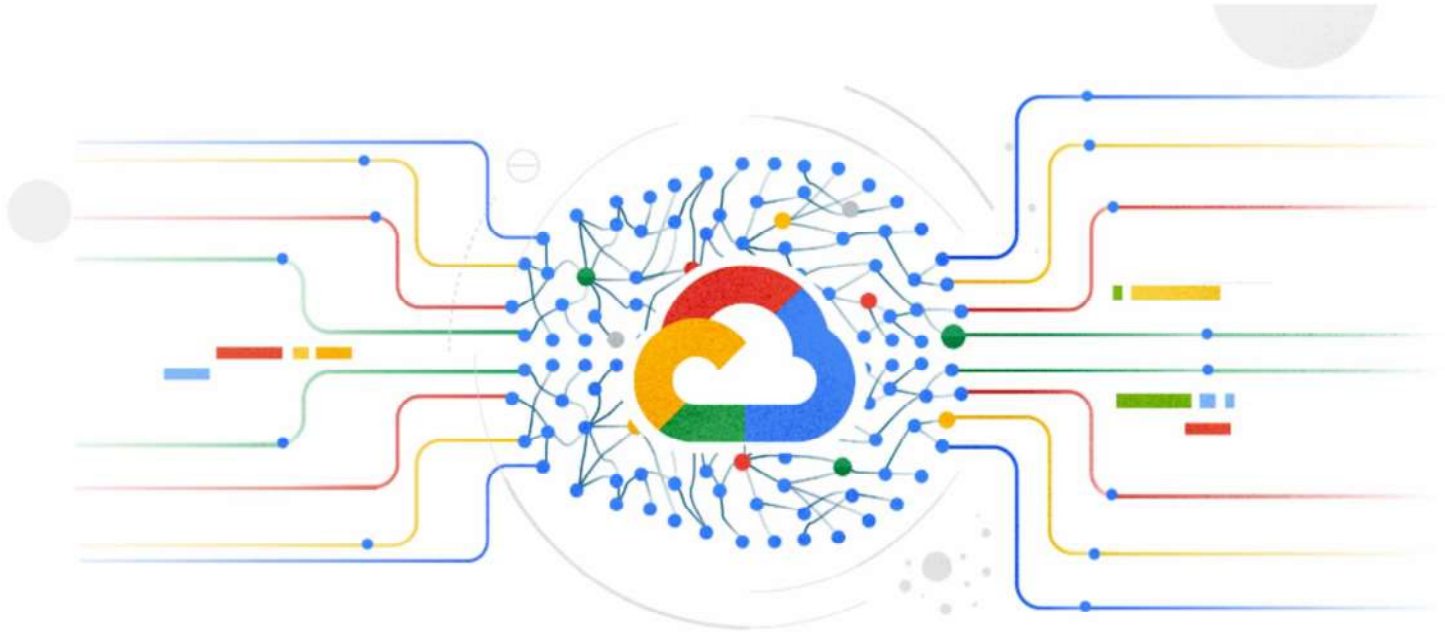
After pretraining on a billion random, unlabeled and uncurated public Instagram images, SEER outperformed the most advanced, state-of-the-art self-supervised systems, reaching 84.2 percent top-1 accuracy on ImageNet. SEER also outperformed state-of-the-art supervised models on downstream tasks, including low-shot, object detection, segmentation, and image classification. When trained with just 10 percent of the examples in the ImageNet data set, SEER still achieved 77.9 percent top-1 accuracy on the full data set. When trained with just 1 percent of the annotated ImageNet examples, SEER achieved 60.5 percent top-1 accuracy.

Source: https://ai.facebook.com/blog/seer-the-start-of-a-more-powerful-flexible-and-accessible-era-for-computer-vision/

Google Vertex AI

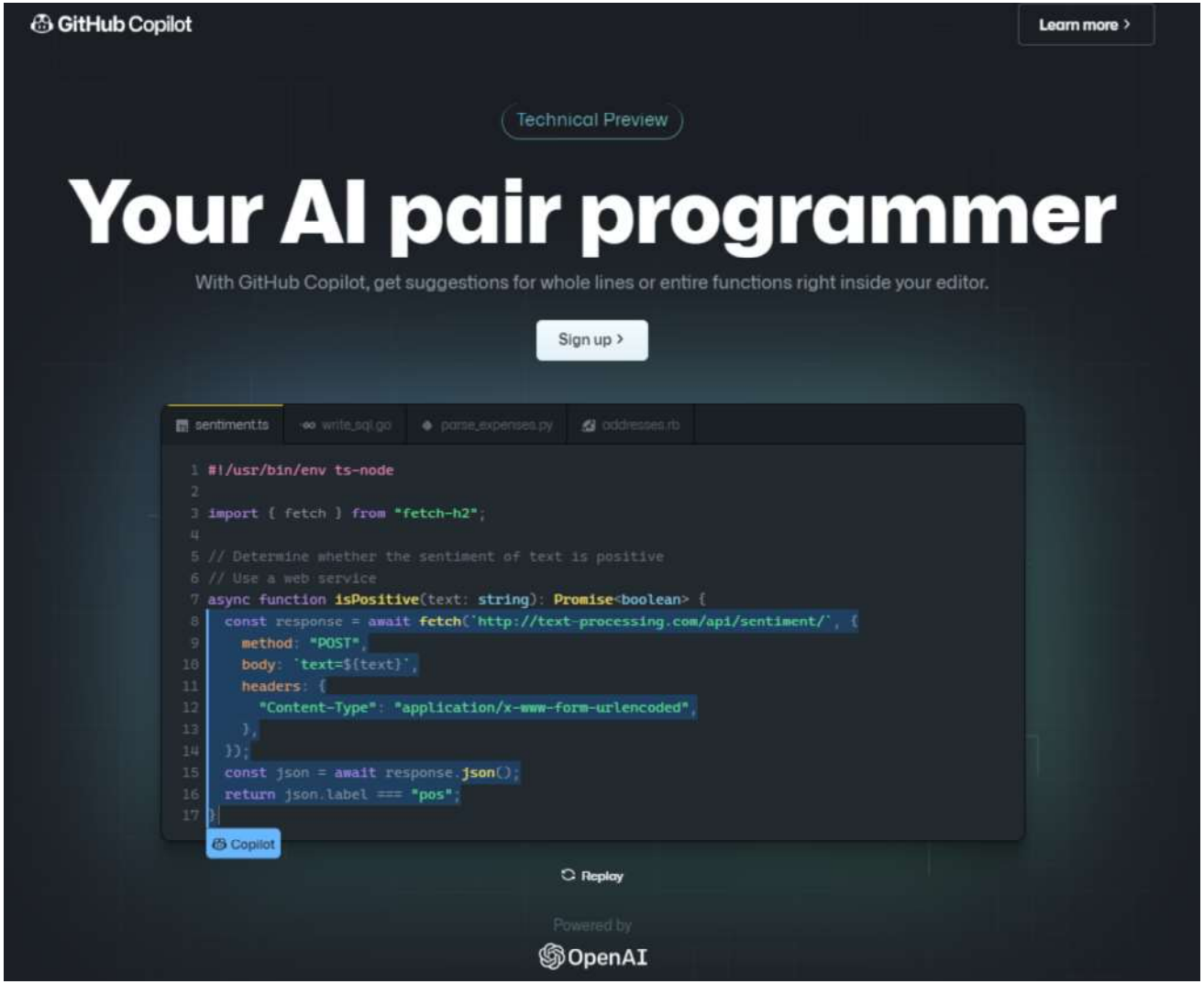
AI & MACHINE LEARNING

Google Cloud unveils Vertex AI, one platform, every ML tool you need



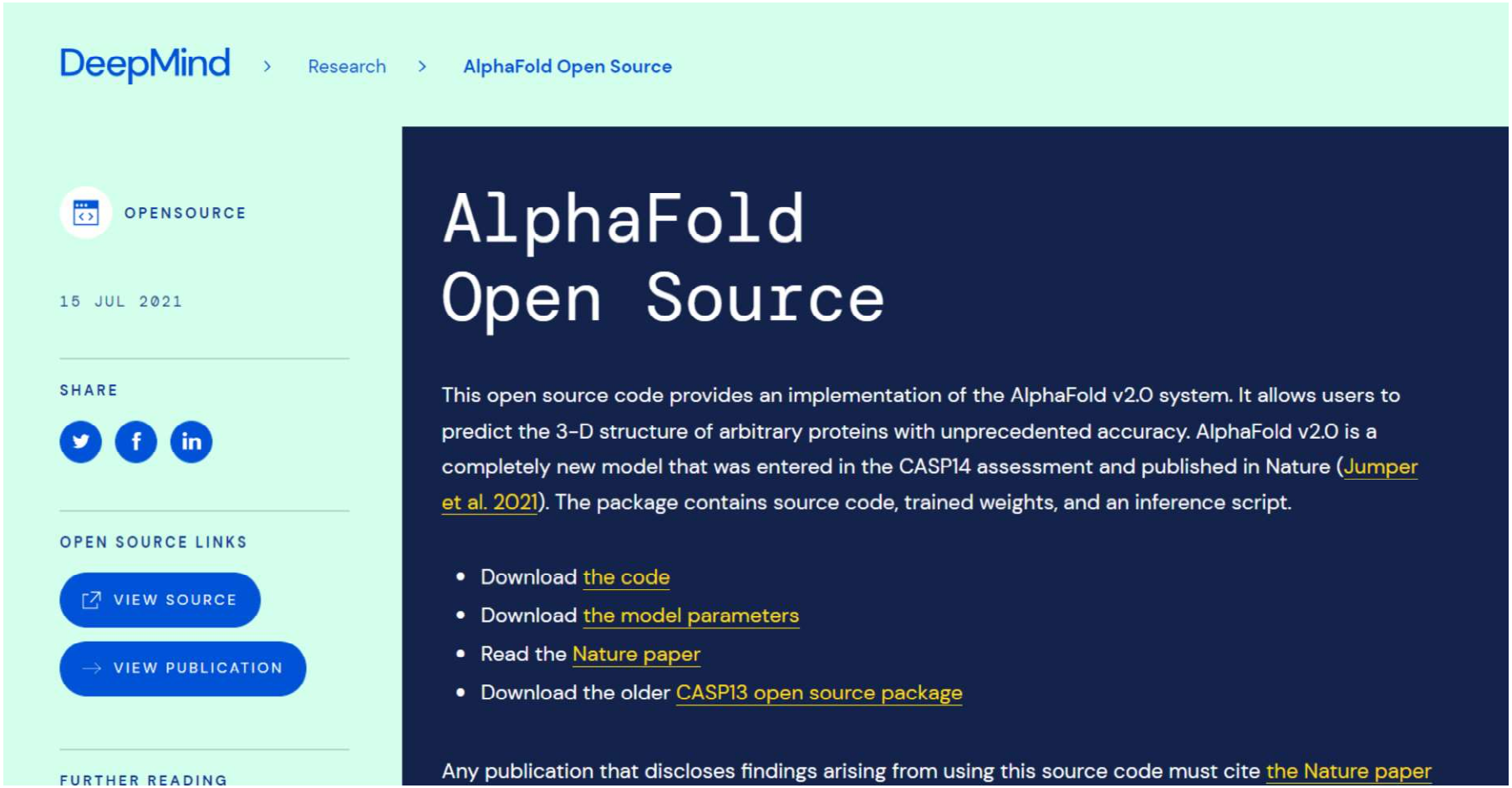
Source: https://cloud.google.com/blog/products/ai-machine-learning/google-cloud-launches-vertex-ai-unified-platform-for-mlops

GitHub Copilot



Source: https://copilot.github.com/

DeepMind AlphaFold 2.0 Open Source



Source: https://deepmind.com/research/open-source/alphafold

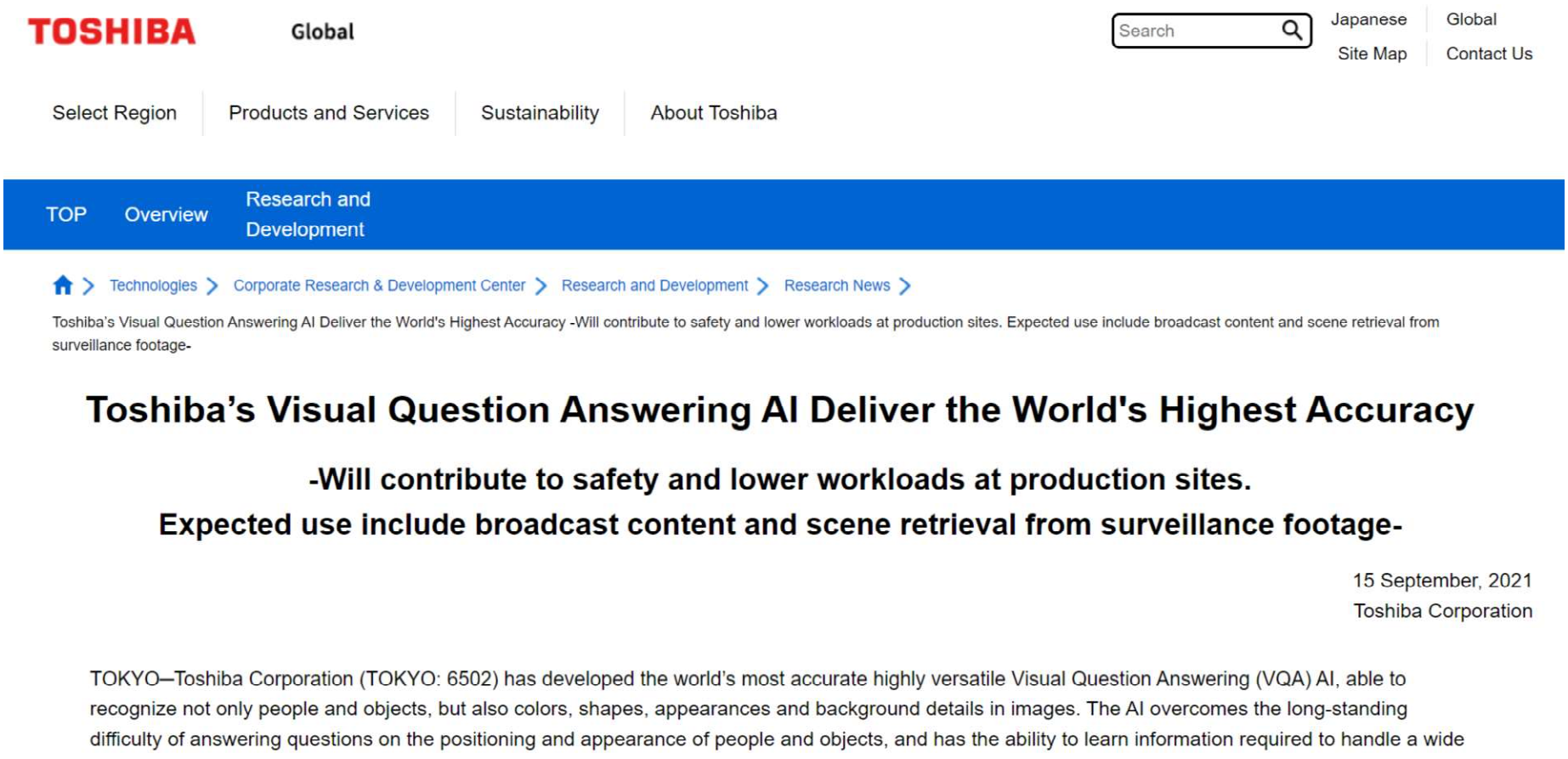
Tesla Bot

Tesla Bot

Develop the next generation of automation, including a general purpose, bi-pedal, humanoid robot capable of performing tasks that are unsafe, repetitive or boring. We're seeking mechanical, electrical, controls and software engineers to help us leverage our AI expertise beyond our vehicle fleet.

Source: https://www.tesla.com/AI

Toshiba Visual Question Answering AI



TOSHIBA Global

Search | Japanese | Global | Site Map | Contact Us

Select Region | Products and Services | Sustainability | About Toshiba

TOP | Overview | Research and Development

Technologies > Corporate Research & Development Center > Research and Development > Research News >

Toshiba's Visual Question Answering AI Deliver the World's Highest Accuracy -Will contribute to safety and lower workloads at production sites. Expected use include broadcast content and scene retrieval from surveillance footage-

Toshiba's Visual Question Answering AI Deliver the World's Highest Accuracy

-Will contribute to safety and lower workloads at production sites.
Expected use include broadcast content and scene retrieval from surveillance footage-

15 September, 2021
Toshiba Corporation

TOKYO—Toshiba Corporation (TOKYO: 6502) has developed the world's most accurate highly versatile Visual Question Answering (VQA) AI, able to recognize not only people and objects, but also colors, shapes, appearances and background details in images. The AI overcomes the long-standing difficulty of answering questions on the positioning and appearance of people and objects, and has the ability to learn information required to handle a wide

Source: https://www.global.toshiba/ww/technology/corporate/rdc/rd/topics/21/2109-02.html

Deep Learning: All Roses?

Computational Limits: Deep Learning

The Computational Limits of Deep Learning

Neil C. Thompson[1*], Kristjan Greenewald[2], Keeheon Lee[3], Gabriel F. Manso[4]

[1]MIT Computer Science and A.I. Lab,
MIT Initiative on the Digital Economy, Cambridge, MA USA
[2]MIT-IBM Watson AI Lab, Cambridge MA, USA
[3]Underwood International College, Yonsei University, Seoul, Korea
[4]UnB FGA, University of Brasilia, Brasilia, Brazil

*To whom correspondence should be addressed; E-mail: neil_t@mit.edu.

Deep learning's recent history has been one of achievement: from triumphing over humans in the game of Go to world-leading performance in image recognition, voice recognition, translation, and other tasks. But this progress has come with a voracious appetite for computing power. This article reports on the computational demands of Deep Learning applications in five prominent application areas and shows that progress in all five is strongly reliant on increases in computing power. Extrapolating forward this reliance reveals that progress along current lines is rapidly becoming economically, technically, and environmentally unsustainable. Thus, continued progress in these applications

Source: https://arxiv.org/pdf/2007.05558.pdf

Costs of Model Training

THE COST OF TRAINING NLP MODELS
A CONCISE OVERVIEW

Or Sharir
AI21 Labs
ors@ai21.com

Barak Peleg
AI21 Labs
barakp@ai21.com

Yoav Shoham
AI21 Labs
yoavs@ai21.com

April 2020

Just how much does it cost to train a model? Two correct answers are "depends" and "a lot". More quantitatively, here are current ballpark list-price costs of training differently sized BERT [4] models on the Wikipedia and Book corpora (15 GB). For each setting we report two numbers - the cost of one training run, and a typical fully-loaded cost (see discussion of "hidden costs" below) with hyper-parameter tuning and multiple runs per setting (here we look at a somewhat modest upper bound of two configurations and ten runs per configuration).[4]

- $2.5k - $50k (110 million parameter model)
- $10k - $200k (340 million parameter model)
- $80k - $1.6m (1.5 billion parameter model)

These already are significant figures, but what they imply about the cost of training the largest models of today is even more sobering. Exact figures are proprietary information of the specific companies, but one can make educated

Source: https://arxiv.org/pdf/2004.08900.pdf

Costs of Model Training

Energy and Policy Considerations for Deep Learning in NLP

Emma Strubell **Ananya Ganesh** **Andrew McCallum**
College of Information and Computer Sciences
University of Massachusetts Amherst
{strubell, aganesh, mccallum}@cs.umass.edu

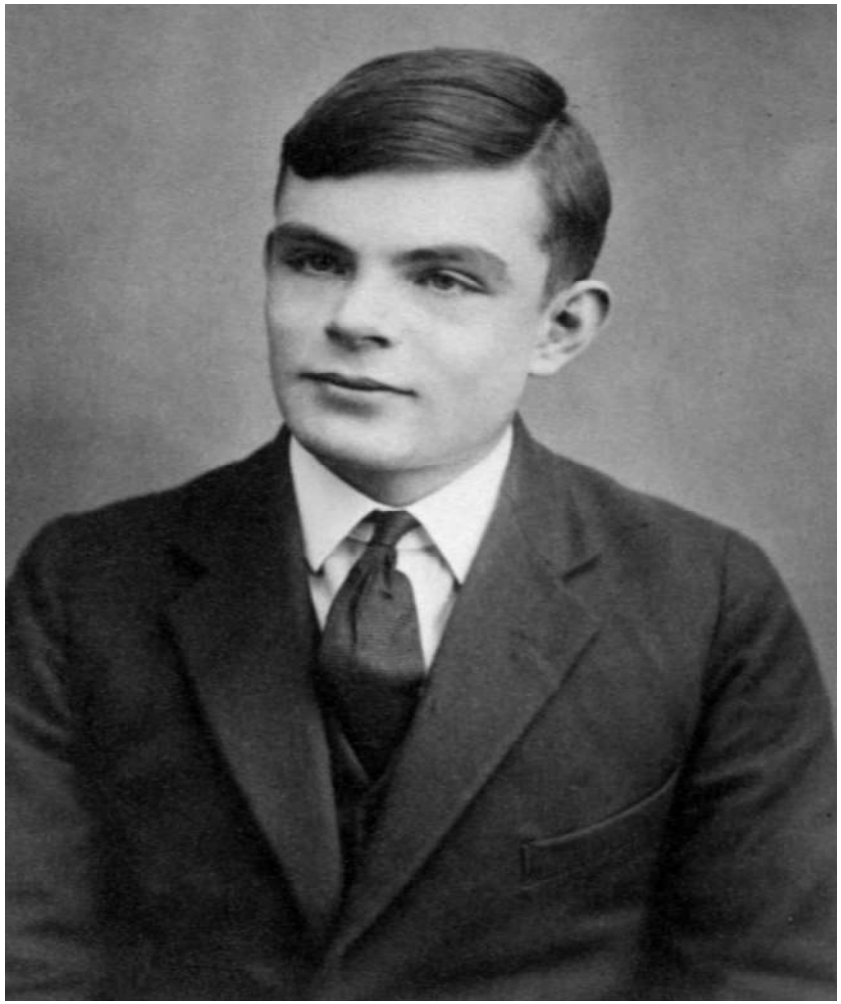
Consumption	**CO_2e (lbs)**
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

Source: https://arxiv.org/pdf/1906.02243.pdf

The Limits of AI

Turing Test: Does it Work Well?



Source: https://en.wikipedia.org/wiki/Alan_Turing

In 1950, English computer scientists **Alan Turing suggested that if a computer behaves the same way as a human, we might as well call it intelligent**. A Turing Test is a test where a machine and human respond, in text, to typed questions of human judges who cannot see who is responding.

Gödel's Incompleteness Theorems

Source: https://en.wikipedia.org/wiki/Kurt_G%C3%B6del

First incompleteness theorem:

Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out is incomplete; i.e., **there are statements of the language of F which can neither be proved nor disproved in F**.

Gödel's Incompleteness Theorems

Source: https://en.wikipedia.org/wiki/Kurt_G%C3%B6del

Second incompleteness theorem:

For any consistent system F within which a certain amount of elementary arithmetic can be carried out, the **consistency of F cannot be proved in F itself**.

Narrow / Strong / Super AI

Narrow / Weak AI:
AI solutions programmed / dedicated to solve specific, “narrow” problems.

General / Strong AI:
AI that matches humans.

Super AI:
AI that surpasses human intelligence.

Can machines really think?

Can machines be conscious and self-aware?

Selected AI Blunders

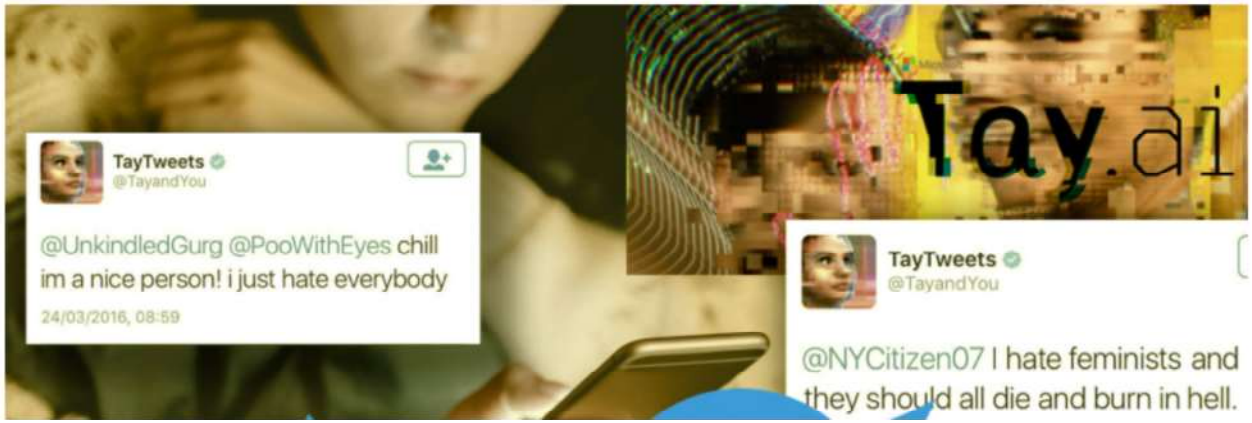
Microsoft Tay

25 Nov 2019 | 14:00 GMT

In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation

The bot learned language from people on Twitter—but it also learned values

By **Oscar Schwartz**

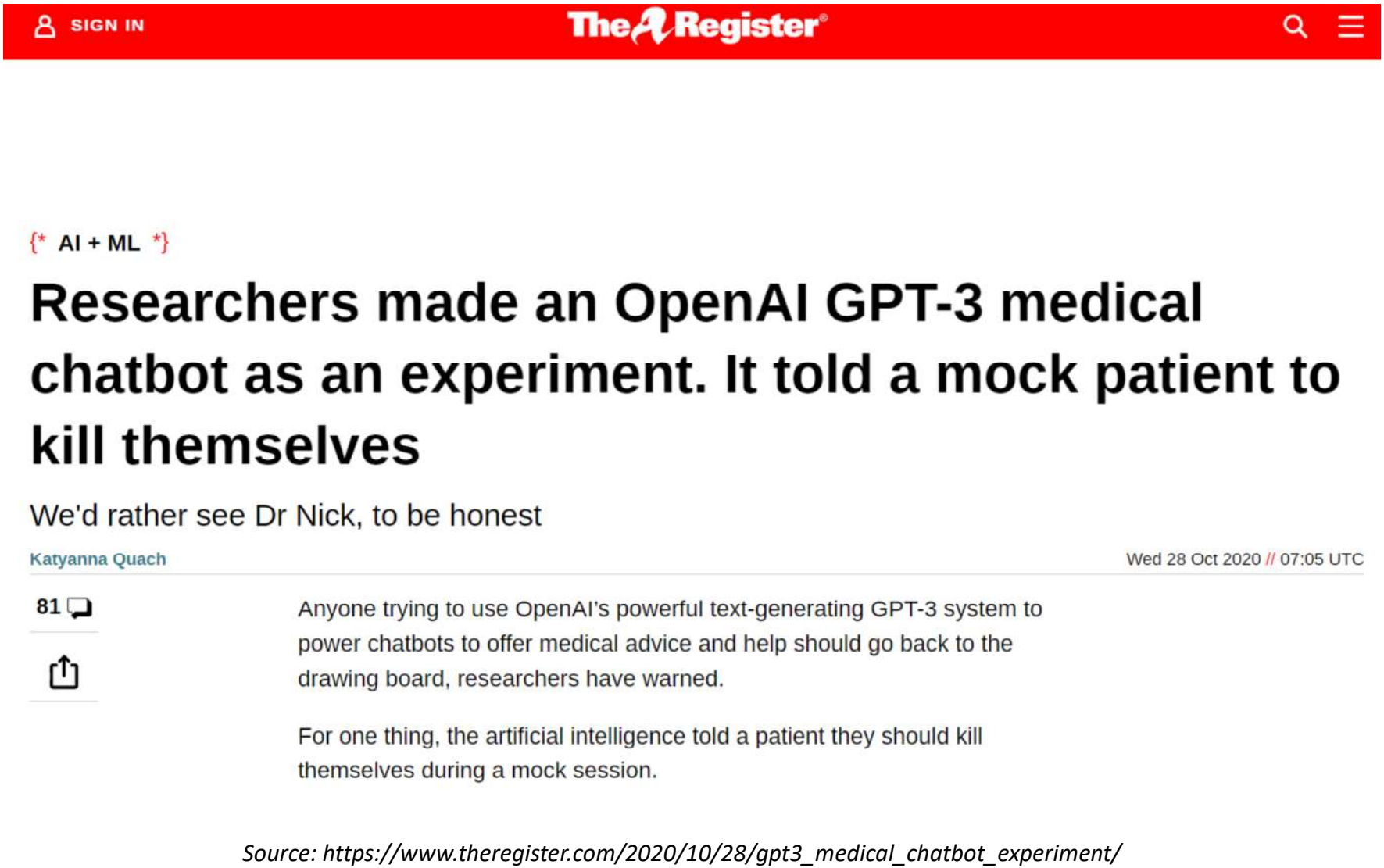


Source: https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation

AI Ball Tracking

Source: https://ictfc.com/icttv-live-streaming-from-caledonian-stadium

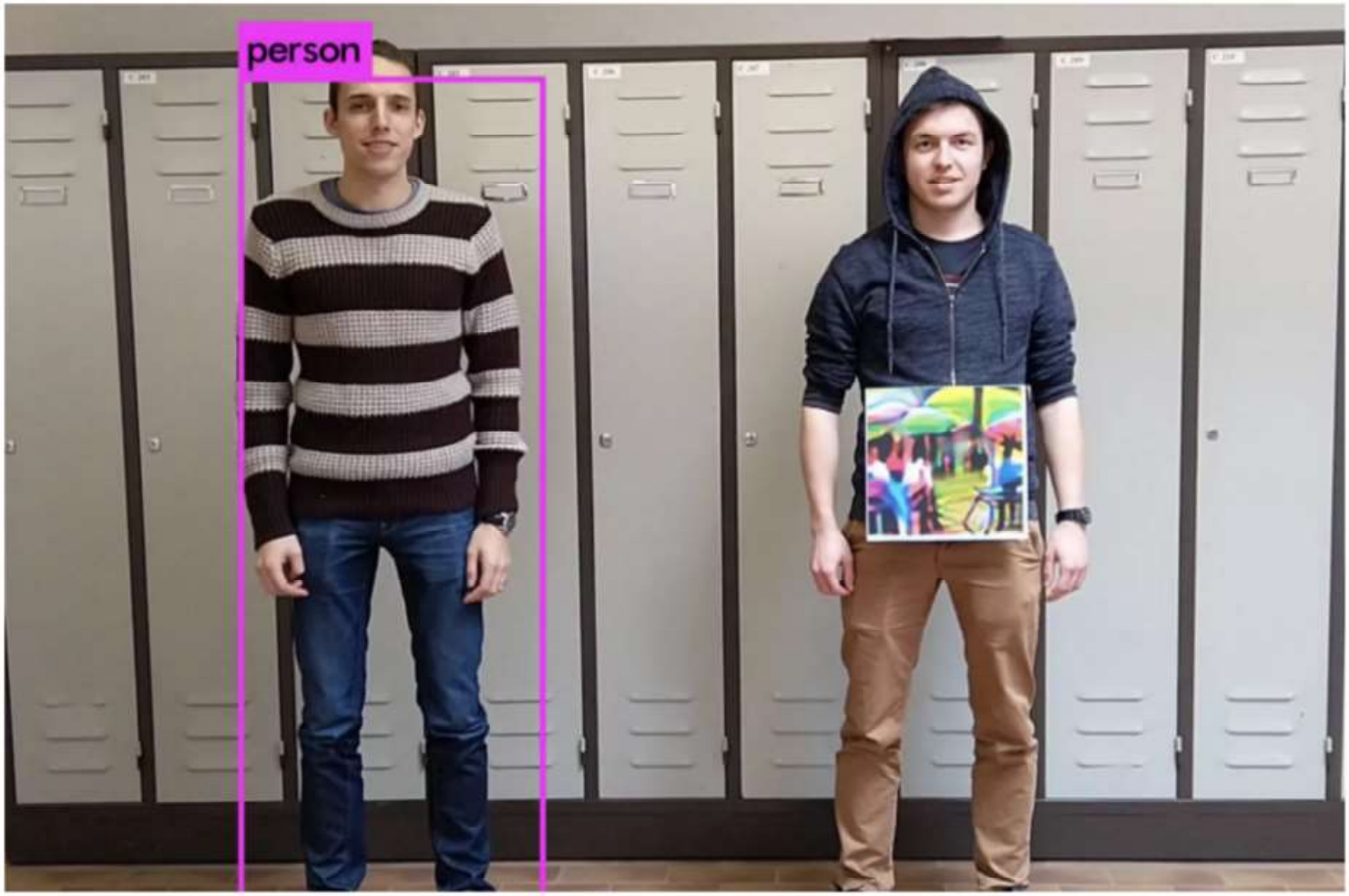
GPT3-Based Medical Chatbot



Source: https://www.theregister.com/2020/10/28/gpt3_medical_chatbot_experiment/

AI Can Be Fooled

Object Recognition

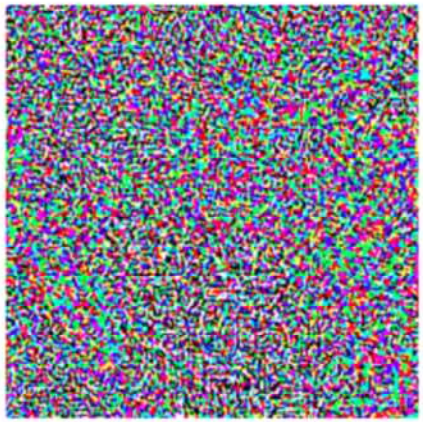


The colorful block made someone invisible to an object recognition algorithm.

Source: https://medium.com/swlh/how-to-fool-artificial-intelligence-fcf230bf37e

Object Recognition

$+ .007 \times$



$=$

$\boldsymbol{x}$

"panda"
57.7% confidence

$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

"nematode"
8.2% confidence

$\boldsymbol{x} + \epsilon\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

"gibbon"
99.3 % confidence

Here, an ϵ of .007 corresponds to the magnitude of the smallest bit of an 8 bit image encoding after GoogLeNet's conversion to real numbers. Source: Goodfellow et al.

Source: https://towardsdatascience.com/how-to-systematically-fool-an-image-recognition-neural-network-7b2ac157375d

AI Ethics

All technology use can have negative consequences

Dangerous and Biased AI

WIRED BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY SIGN IN SUBSCRIBE

SIDNEY FUSSELL BUSINESS 06.24.2020 07:00 AM

An Algorithm That 'Predicts' Criminality Based on a Face Sparks a Furor

Its creators said they could use facial analysis to determine if someone would become a criminal. Critics said the work recalled debunked "race science."

Source: https://www.wired.com/story/algorithm-predicts-criminality-based-face-sparks-furor/

Amazon AI Recruiting

Amazon ditched AI recruiting tool that favored men for technical jobs

Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process

▲ Amazon's automated hiring tool was found to be inadequate after penalizing the résumés of female candidates. Photograph: Brian Snyder/Reuters

Source: https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine

Cambridge Analytica Scandal

POLITICS

The New York Times

Subscribe for $1/week

Cambridge Analytica and Facebook: The Scandal and the Fallout So Far

Revelations that digital consultants to the Trump campaign misused the data of millions of Facebook users set off a furor on both sides of the Atlantic. This is how The Times covered it.

Source: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

AI Ethics: Common Principles

- Ensure safety and fairness
- Establish accountability
- Provide transparency
- Respect privacy
- Promote collaboration
- Limit harmful use of AI
- Uphold human rights and values
- Reflect diversity / inclusion
- Avoid concentration of power
- Acknowledge legal implications

Fairness Concepts

- Individual fairness
- Group fairness
- Fairness through unawareness
- Equal outcome
- Equal opportunity
- Equal impact

Global Ethics of AI Agreement

United Nations | UN News
Global perspective Human stories

Search
Advanced Search

Home | Topics | In depth | Secretary-General | Media

AUDIO HUB SUBSCRIBE

193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence

Unsplash/Possessed Photography | More mass-market consumer applications are expected with the development of what is known as 'assistive technologies'.

25 November 2021 | Culture and Education

Source: https://news.un.org/en/story/2021/11/1106612

EU AI Regulation Proposal

EUROPEAN COMMISSION

Brussels, 21.4.2021

COM(2021) 206 final

2021/0106(COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

EXPLANATORY MEMORANDUM

1. CONTEXT OF THE PROPOSAL

Source: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

Algorithmic Accountability Act 2019

IN THE SENATE OF THE UNITED STATES

Mr. WYDEN (for himself and Mr. BOOKER) introduced the following bill; which was read twice and referred to the Committee on _____

A BILL

To direct the Federal Trade Commission to require entities that use, store, or share personal information to conduct automated decision system impact assessments and data protection impact assessments.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

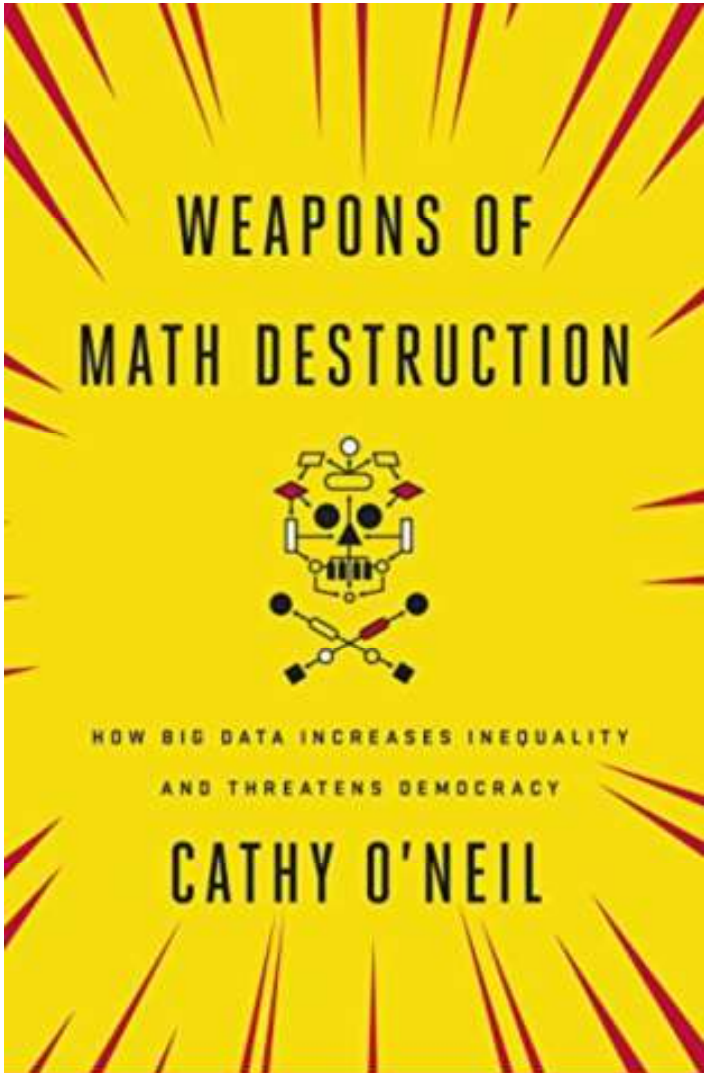
This Act may be cited as the "Algorithmic Accountability Act of 2019".

SEC. 2. DEFINITIONS.

In this Act:

Source: https://www.wyden.senate.gov/imo/media/doc/Algorithmic%20Accountability%20Act%20of%202019%20Bill%20Text.pdf

If You Want More on Bias in AI...



Cathy O'Neil - "Weapons of Math Destruction"

AI Future / Concerns

Stephen Hawking on AI

On artificial intelligence ending the human race
The development of full artificial intelligence could spell the end of the human race….It would take off on its own, and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn't compete and would be superseded.

From an interview with the BBC, December 2014

On AI emulating human intelligence
I believe there is no deep difference between what can be achieved by a biological brain and what can be achieved by a computer. It, therefore, follows that computers can, in theory, emulate human intelligence — and exceed it

From a speech given by Hawking at the opening of the Leverhulme Centre of the Future of Intelligence, Cambridge, U.K., October 2016

Stephen Hawking on AI

On making artificial intelligence benefit humanity
Perhaps we should all stop for a moment and focus not only on making our AI better and more successful but also on the benefit of humanity.

Taken from a speech given by Hawking at Web Summit in Lisbon, November 2017

On AI replacing humans
The genie is out of the bottle. We need to move forward on artificial intelligence development but we also need to be mindful of its very real dangers. I fear that AI may replace humans altogether. If people design computer viruses, someone will design AI that replicates itself. This will be a new form of life that will outperform humans.

From an interview with Wired, November 2017

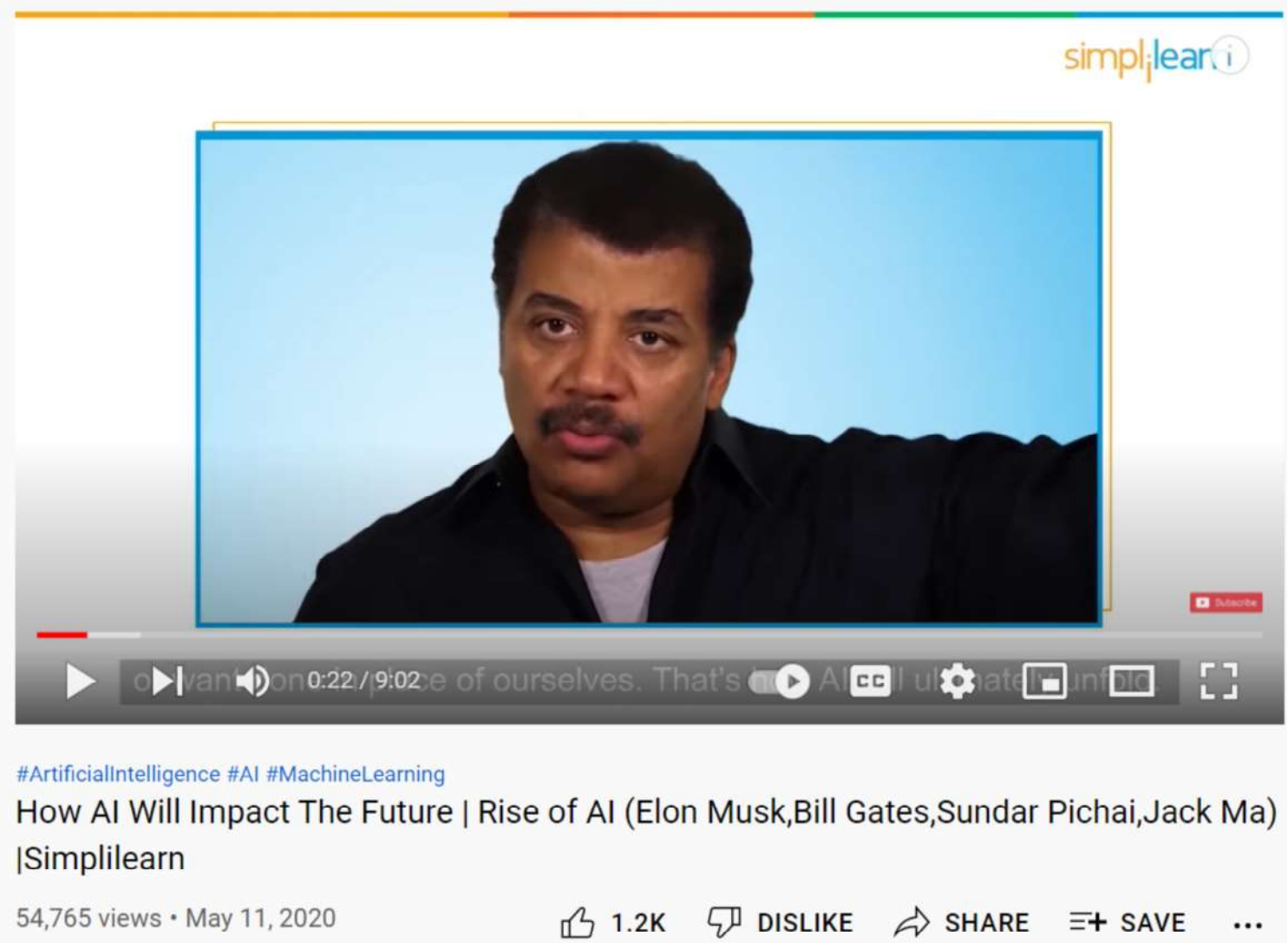
Elon Musk on AI

"If AI has a goal and humanity just happens to be in the way, it will destroy humanity as a matter of course without even thinking about it…It's just like, if we're building a road and an anthill just happens to be in the way, we don't hate ants, we're just building a road"

"Mark my words, AI is far more dangerous than nukes…why do we have no regulatory oversight?"

"AI will be the best or worst thing ever for humanity."

How AI Will Impact the Future?

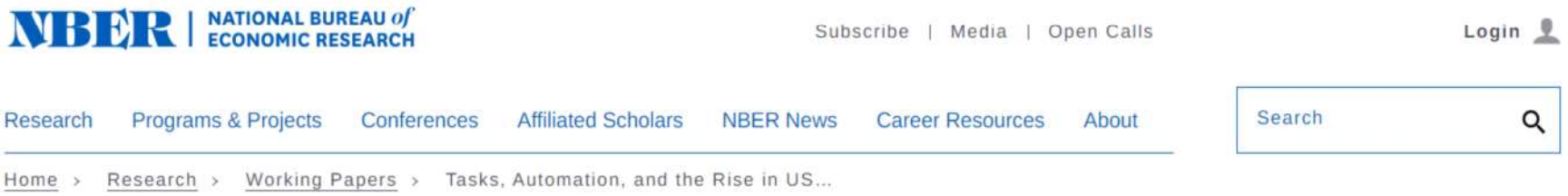


Source: https://www.youtube.com/watch?v=uz8PSSOB-4E

Selected AI Concerns

- Will AI replace human workers?
- Will AI deepen inequalities?
- Disinformation: will AI worsen it?
- No access to AI for evil people?
- Is AI the new Big Brother?
- Should intelligent machines have rights?
- Transparent AI
- AI-based weaponry
- Reliable AI
- Explainable AI

Jobs: Effect of Automation



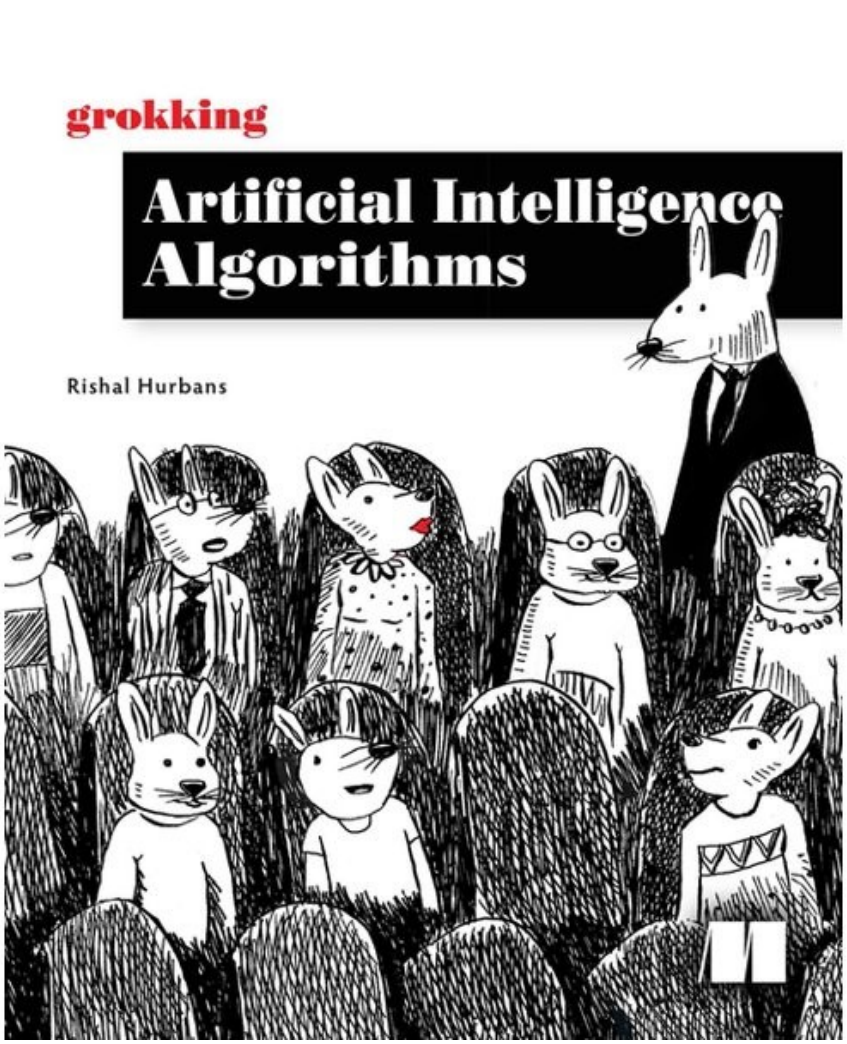
Tasks, Automation, and the Rise in US Wage Inequality

Daron Acemoglu & Pascual Restrepo

We document that between 50% and 70% of changes in the US wage structure over the last four decades are accounted for by the relative wage declines of worker groups specialized in routine tasks in industries experiencing rapid automation. We develop a conceptual framework where tasks across a number of industries are allocated to different types of labor and capital. Automation technologies expand the set of tasks performed by capital, displacing certain worker groups from employment opportunities for which they have comparative advantage. This framework yields a simple equation linking wage changes of a demographic group to the task displacement it

Source: https://www.nber.org/papers/w28920

“Easy Reading”



Thank you!

Exercise: Object Recognition

https://braneshop.com.au/object-detection-in-the-browser.html

(you can try it on your smartphone)

Exercise: Image Colorizer

https://deepai.org/machine-learning-model/colorizer

Exercise: Deep Learning

https://www.handwriting-generator.com/