Computer Vision

Lecture 1: Introduction to Artificial Intelligence (AI)

Yuriy Kochura
iuriy.kochura@gmail.com
@y_kochura

Today

- Intelligence vs Artificial Intelligence
- A brief history of AI
- Machine learning definitions and paradigm
- Types of Machine Learning
- Deep Learning concept
- Applications of Deep Learning
- Perceptron: Forward propagation
- Common activation functions

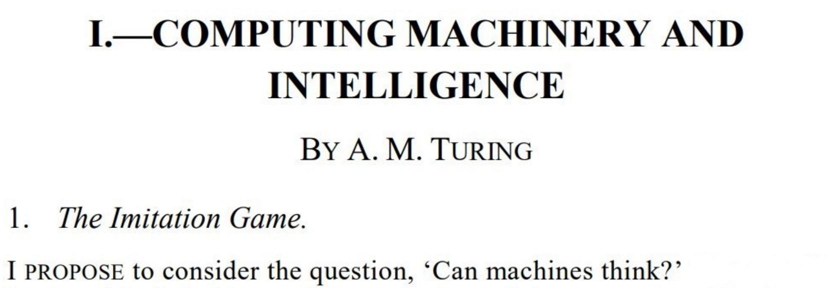
What is Computer Vision?

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do. [Wikipedia]

Computer Vision is used in image understanding, search, apps, mapping, medicine, drones, and self-driving cars. Core visual recognition tasks to many of these applications are image classification, localization and object detection. Recent success of neural networks (aka “deep learning”) have greatly advanced the performance of these state-of-the-art visual recognition systems.

Artificial Intelligence

Can machines think?



I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?'

— Alan Turing, 1950

Image source: biography

—

Credits: Alan Turing, 1950.

In the process of trying to imitate an adult human mind we are bound to think a good deal about the process which has brought it to the state that it is in. We may notice three components,

a. The initial state of the mind, say at birth,

b. The education to which it has been subjected,

c. Other experience, not to be described as education, to which it has been subjected.

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed.

— Alan Turing, 1950

—

Credits: Alan Turing, 1950.

What is intelligence?

- Intelligence is about

To be able to learn to make decisions to achieve goals

- Learning, decisions, and goals are all central

What is artificial intelligence?

- In a general sense

Any technique that enable computers to mimic human behavior

What is artificial intelligence?

- In a narrow sense

Capability of an engineered system to acquire, process and apply knowledge and skills

- Knowledge are facts, information and skills acquired through experience or education.

—
Credits: ISO/IEC TR 24028:2020(en), 2020.

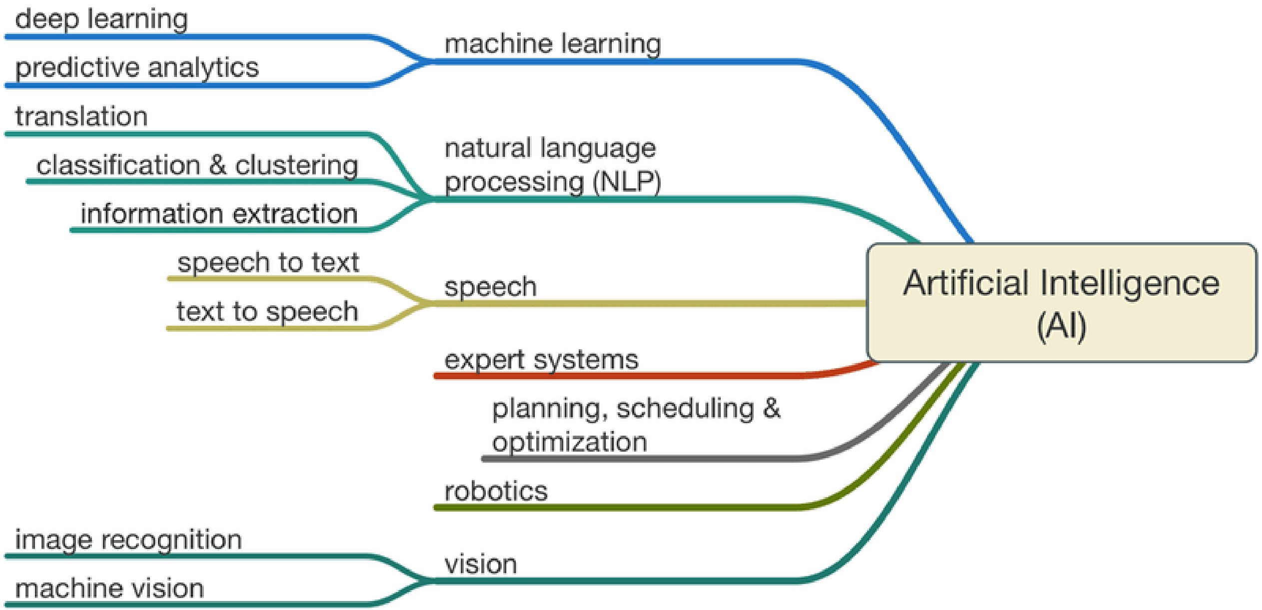
A brief history of AI

- 1940—1952: Early days
    - 1943: McCulloch & Pitts: Boolean circuit model of brain
    - 1950: Turing's ''Computing Machinery and Intelligence''
- 1952–1956: The birth of AI
    - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
    - 1956: Dartmouth meeting: ''Artificial Intelligence'' adopted
- 1956–1974: The golden years
    - 1958: Frank Rosenblatt invented perceptron (simple neural network)
    - 1964: Bobrow's program that solves algebra word problems
    - 1965: Robinson's complete algorithm for logical reasoning
- 1974–1980: The first AI winter
- 1980–1987: Expert systems industry boom
- 1987—1993: Expert systems industry busts: the second AI winter
- 1993–2011: Statistical approaches
    - Resurgence of probability, focus on uncertainty
    - General increase in technical depth
    - Intelligent agents
- 2011–present: Deep Learning, Big Data and AI
    - Big data, big compute, neural networks
    - AI used in many industries

—

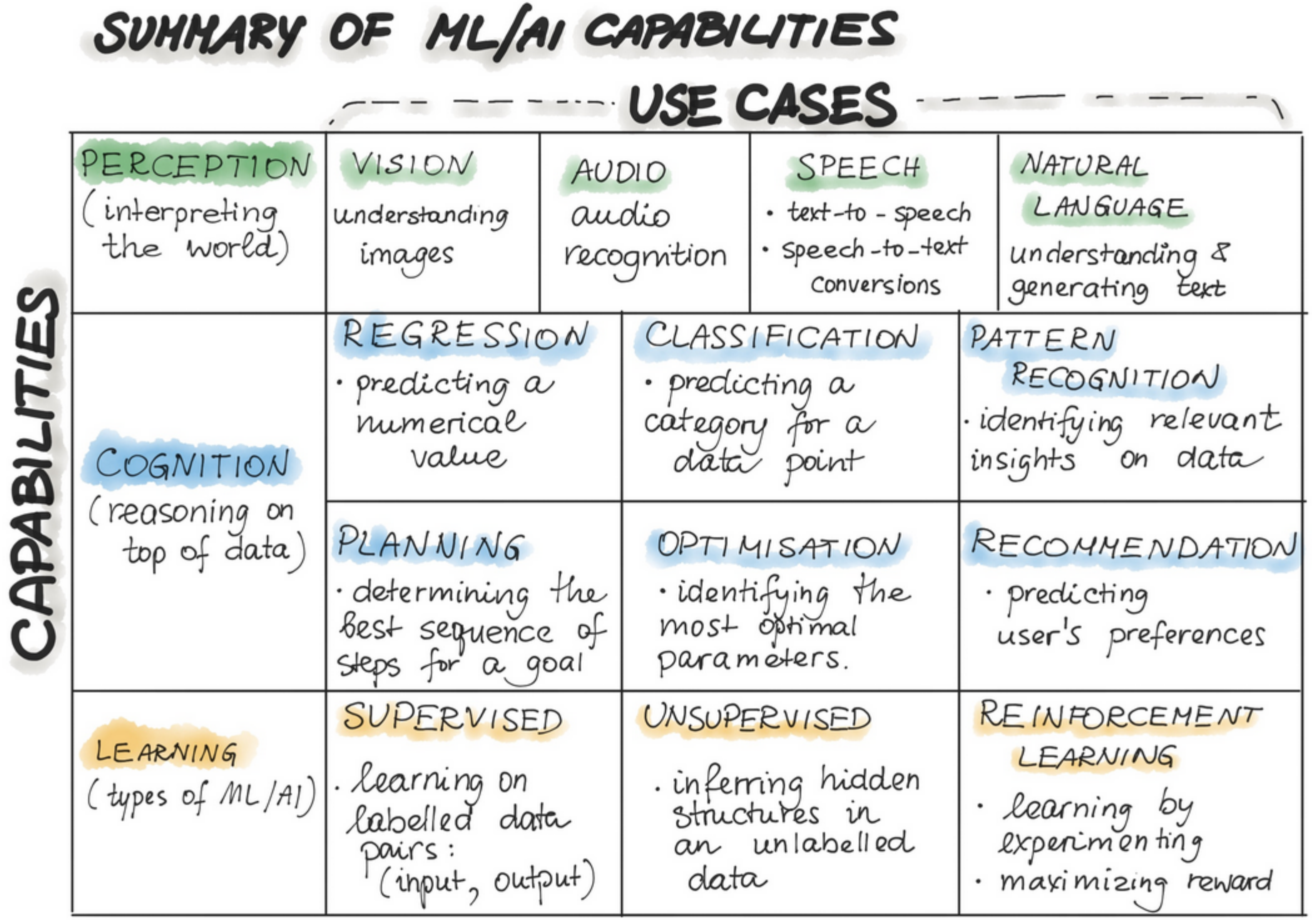
Credits: Wikipedia - History of artificial intelligence

AI is a rich field



—

Image Source: Marizel B. and Ma. Louella Salenga, 2018.


SUMMARY OF ML/AI CAPABILITIES
USE CASES
CAPABILITIES
PERCEPTION
(interpreting the world)
VISION
understanding images
AUDIO
audio recognition
SPEECH
• text-to-speech
• speech-to-text conversions
NATURAL LANGUAGE
understanding & generating text
COGNITION
(reasoning on top of data)
REGRESSION
• predicting a numerical value
CLASSIFICATION
• predicting a category for a data point
PATTERN RECOGNITION
• identifying relevant insights on data
PLANNING
• determining the best sequence of steps for a goal
OPTIMISATION
• identifying the most optimal parameters.
RECOMMENDATION
• predicting user's preferences
LEARNING
(types of ML/AI)
SUPERVISED
• learning on labelled data pairs: (input, output)
UNSUPERVISED
• inferring hidden structures in an unlabelled data
REINFORCEMENT LEARNING
• learning by experimenting
• maximizing reward

"Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don't think AI will transform in the next several years."

— Andrew Ng

—
Credits: Andrew Ng: Artificial Intelligence is the New Electricity, 2017.

Machine Learning

What is machine learning?

A. L. Samuel*

Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress

Image source: wikipedia

Field of study that gives computers the ability to learn without being explicitly programmed.
— Arthur Samuel, 1959

—
Credits: Arthur Samuel, 1959.

What is machine learning?

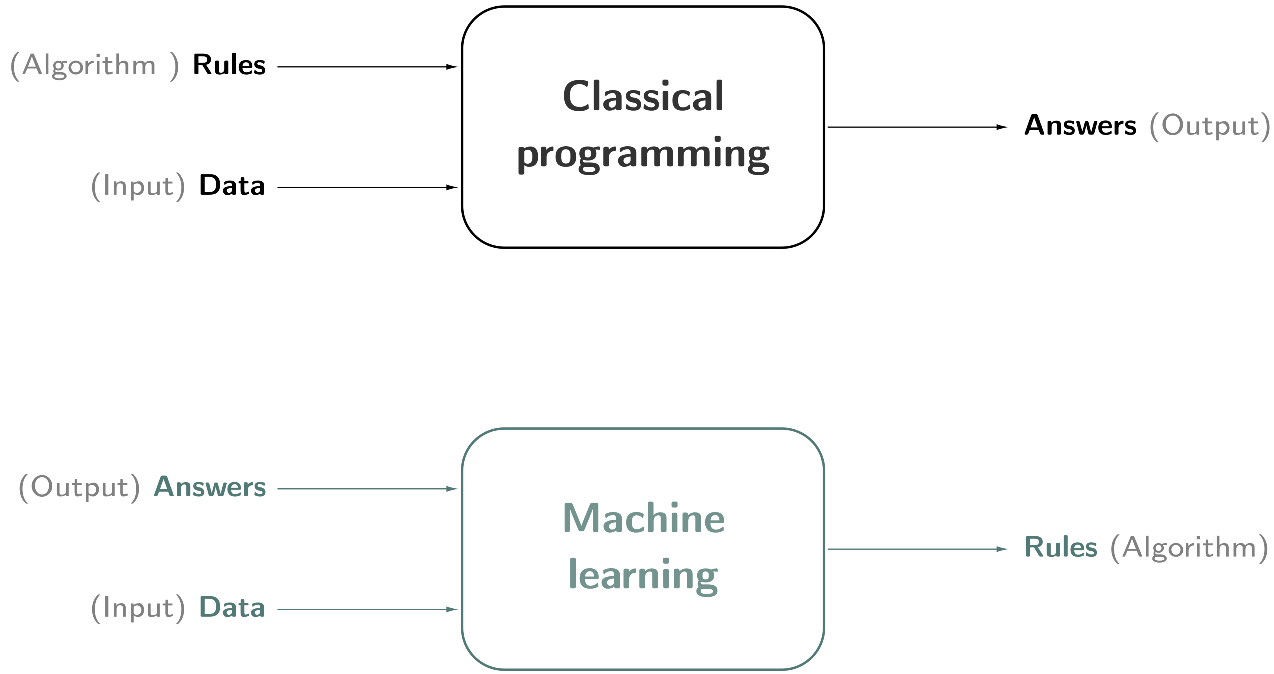
Machine Learning is the study of computer algorithms that improve automatically through experience.

— Tom Mitchell, 1997

Image source: Tom Mitchell's Home Page

—
Credits: Tom Mitchell, 1997.

Machine learning: a new programming paradigm



Types of Machine Learning

Контрольоване навчання Supervised learning	Напівконтрольоване навчання Semi-supervised learning	Неконтрольоване навчання Unsupervised learning	Навчання з підкріплення Reinforcement learning
Дані: (x, y) x – приклад, y – мітка	**Дані:** (x, y) та x, $\lvert(x, y)\rvert < \lvert x\rvert$ x – приклад, y – мітка	**Дані:** x x – приклад, немає міток!	**Дані:** пари **стан-дія**
Мета – знайти функцію відображення $x \longrightarrow y$	**Мета** – знайти функцію відображення або категорію $x \longrightarrow y$	**Мета** – знайти правильну категорію.	**Мета** – максимізація загальної винагороди, отриманої агентом при взаємодії з навколишнім середовищем.
Приклад  Це є яблуко.	**Приклад**  Це є яблуко.	**Приклад**  Цей об'єкт схожий на інший.	**Приклад**  Їжте це, бо це зробить вас сильнішим.

What is Deep Learning?

Artificial Intelligence

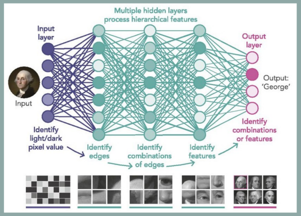
Any technique that enable computers to mimic human behavior

Machine Learning

Ability to learn without explicitly being programmed

Deep Learning

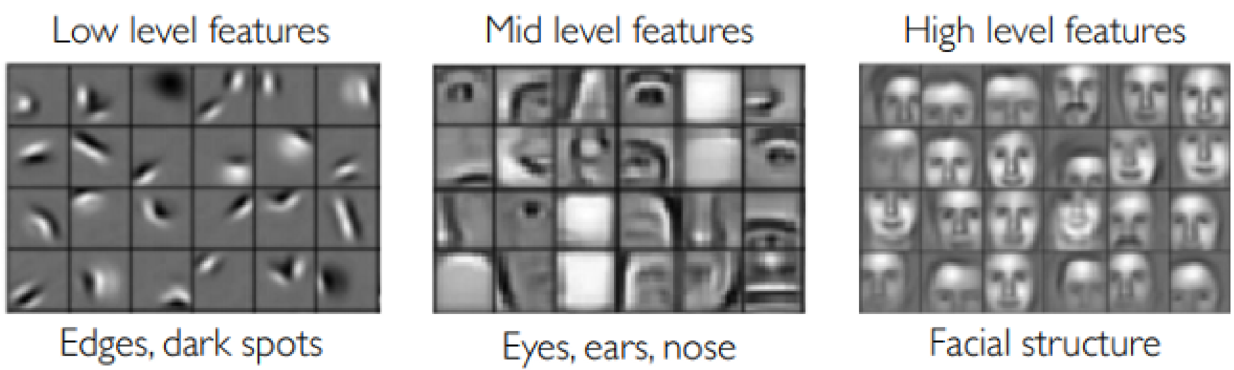
Extract patterns from data using neural networks



Why Deep Learning?

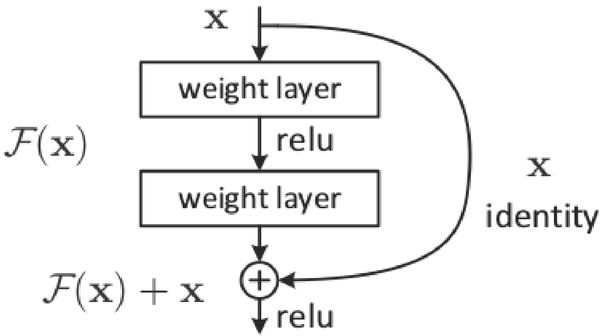
Hand engineered features are time consuming, brittle, and not scalable in practice

Can we learn the underlying features directly from data?



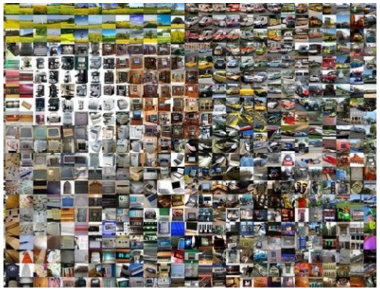
Why does DL work now?

Algorithms (old and new)



More data

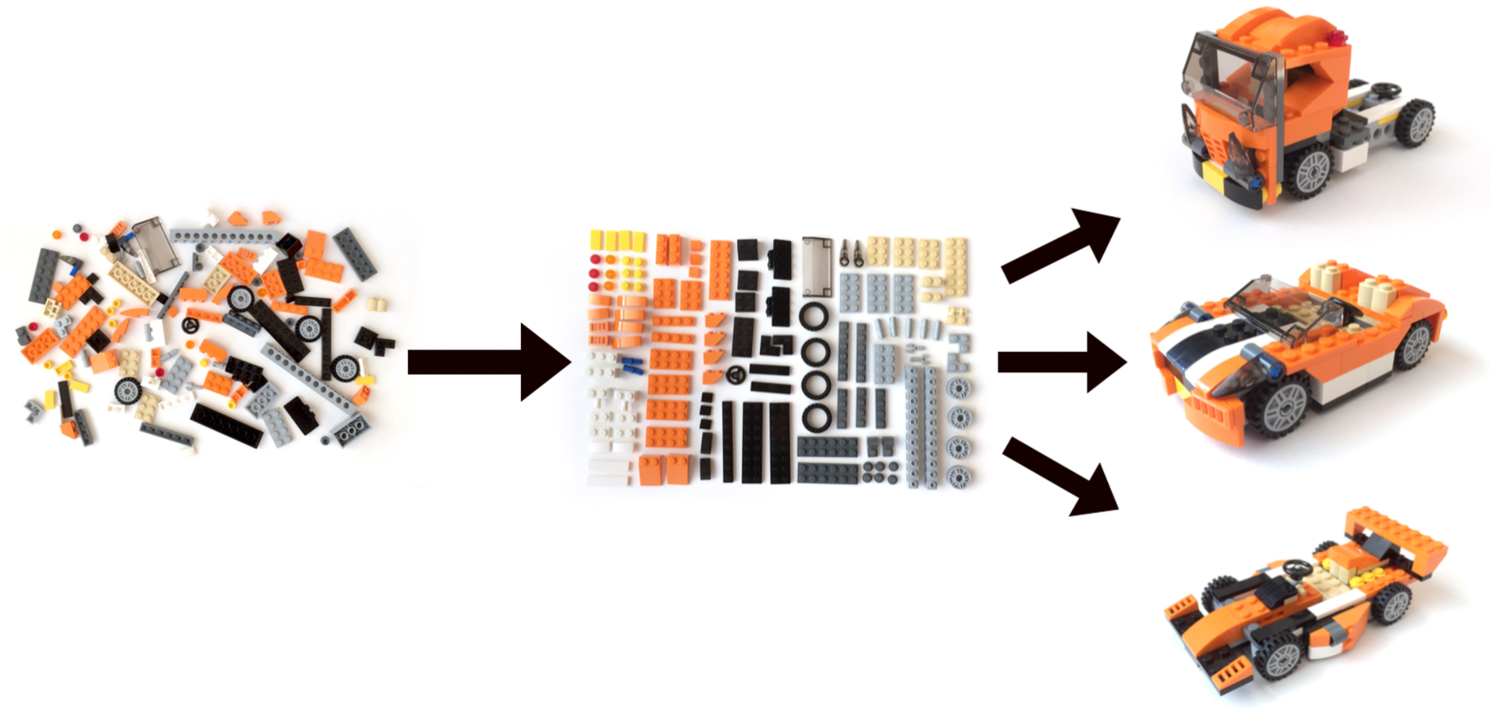
- Larger Datasets
- Easier Collection & Storage



Software (New models and improved techniques)

Faster compute engines

DL as an architectural language



—

Image source: http://chelseamarzean.com/post-the-atomic-workflow/, 2016.

Image source: Christopher Bishop

For the last forty years we have programmed computers; for the next forty years we will train them.

— Chris Bishop, 2020.

—
Credits: Chris Bishop on "The Real AI Revolution" NeurIPS 2020

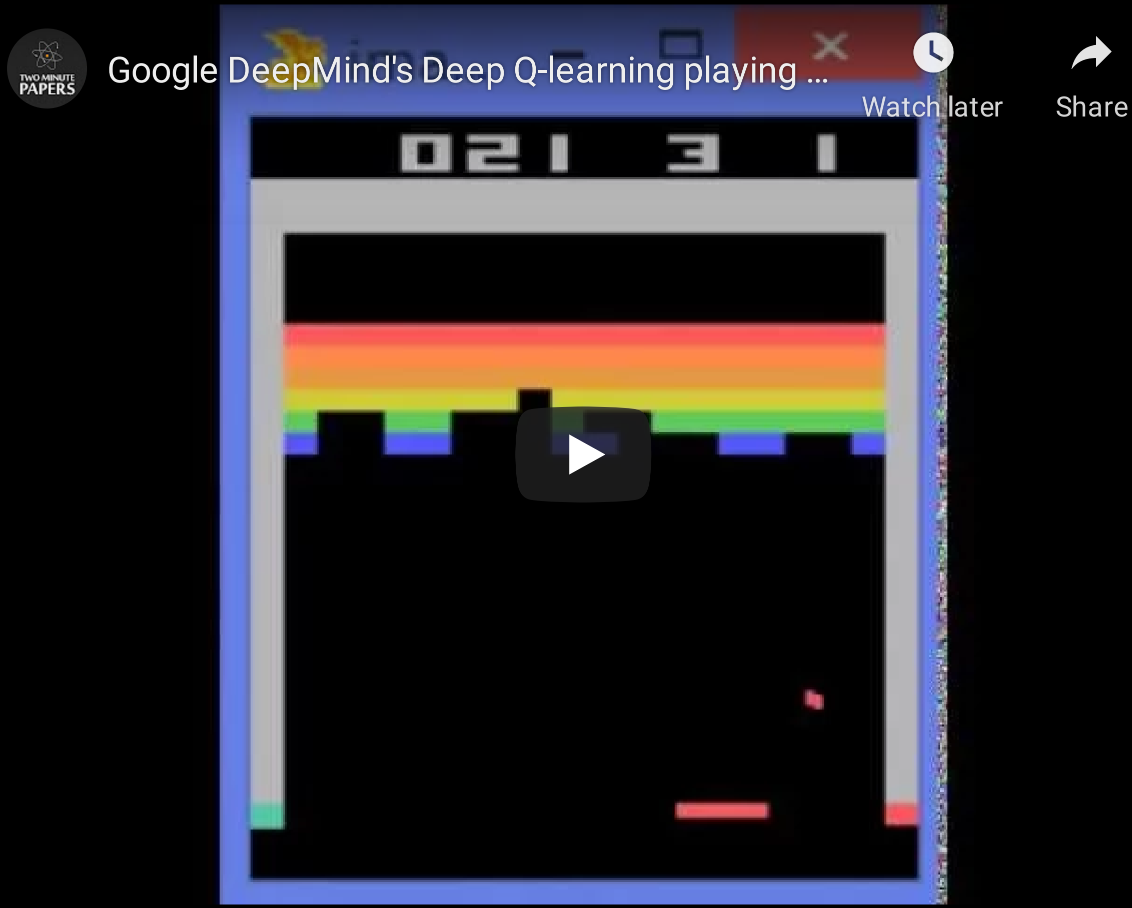
*"ACM named **Yann LeCun**, **Geoffrey Hinton**, and **Yoshua Bengio** recipients of the **2018 ACM A.M. Turing Award** for conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing."*

Applications and successes

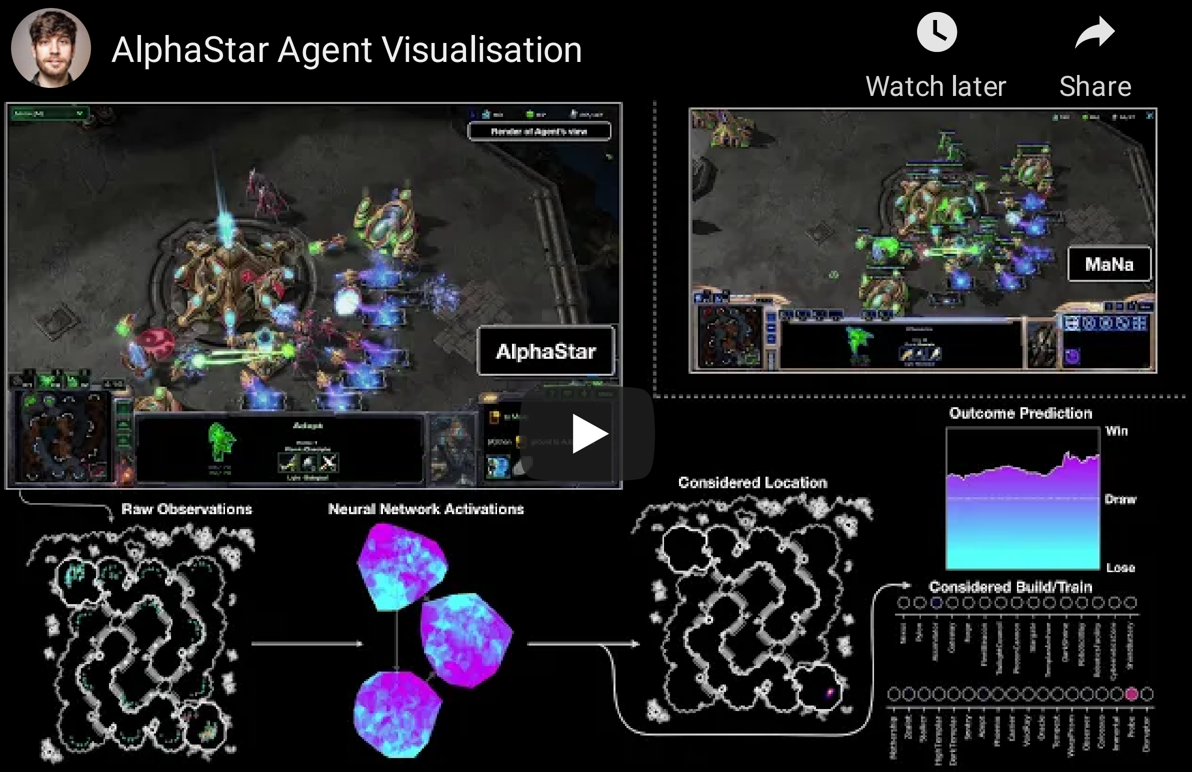
Detectron2: A PyTorch-based modular object detection ...

Copy link

Object detection, pose estimation, segmentation (2019)



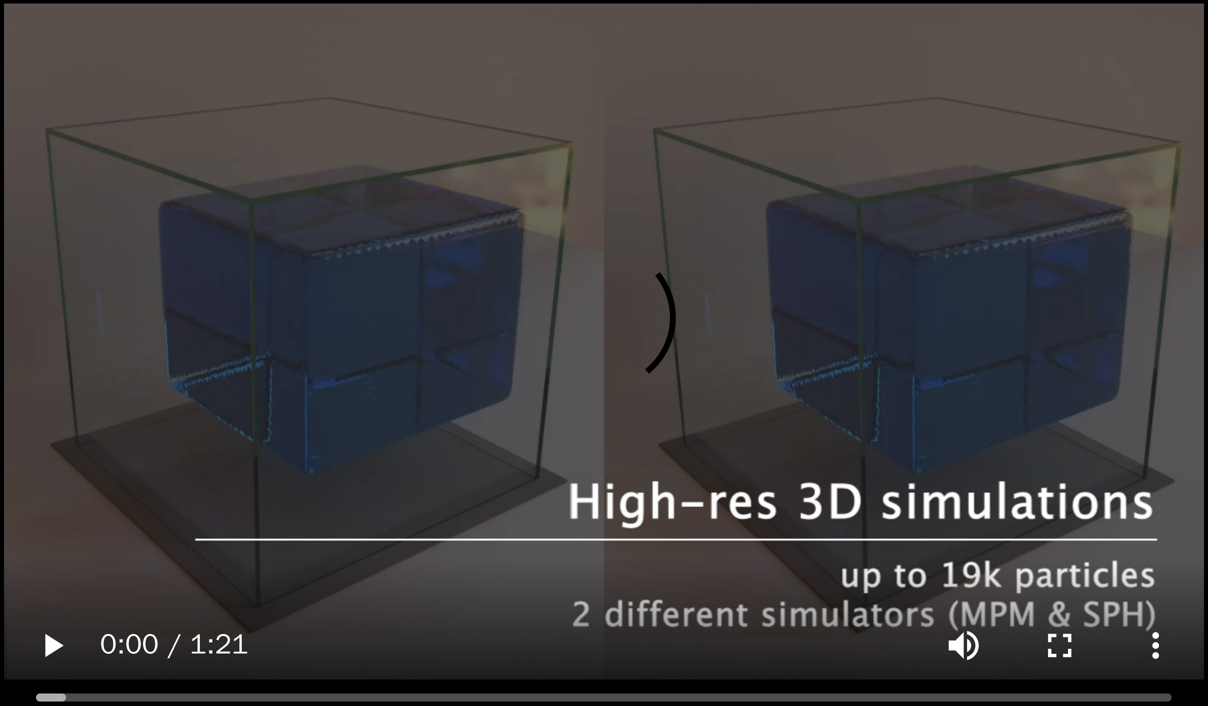
Reinforcement learning (Mnih et al, 2014)



Strategy games (Deepmind, 2016-2018)

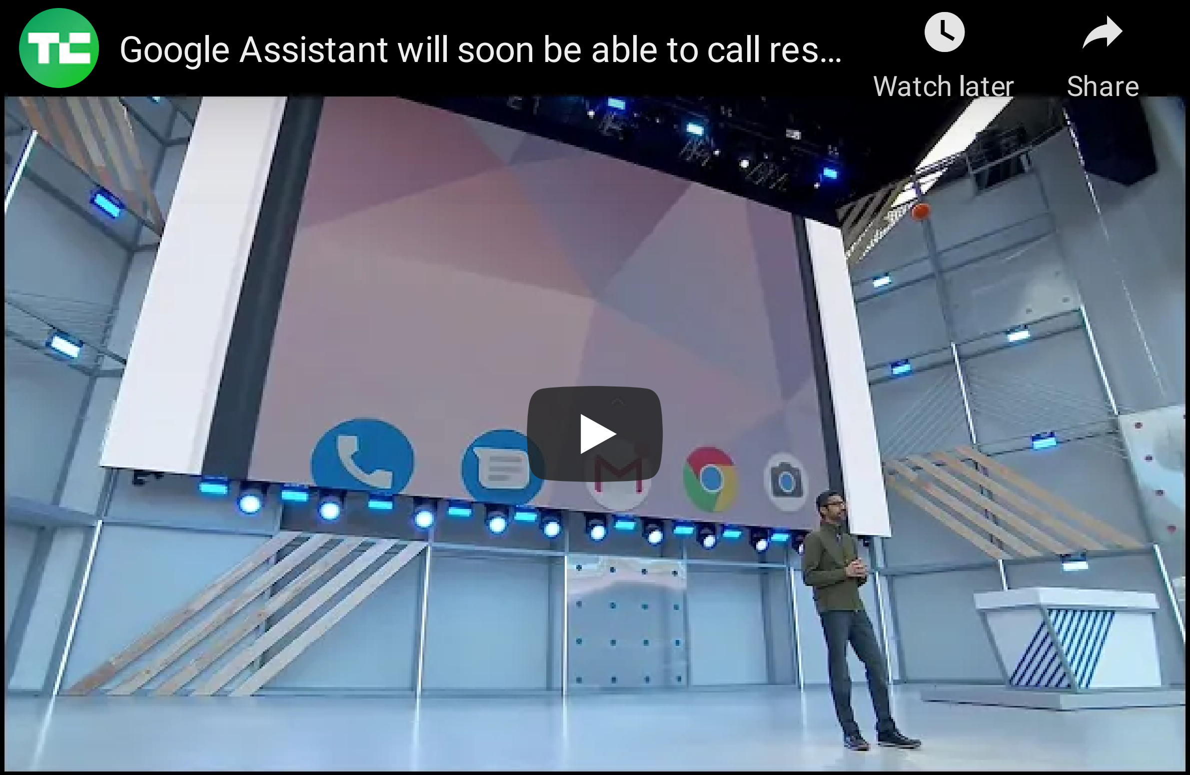
Autonomous cars (NVIDIA, 2016)

Autopilot (Tesla, 2019)

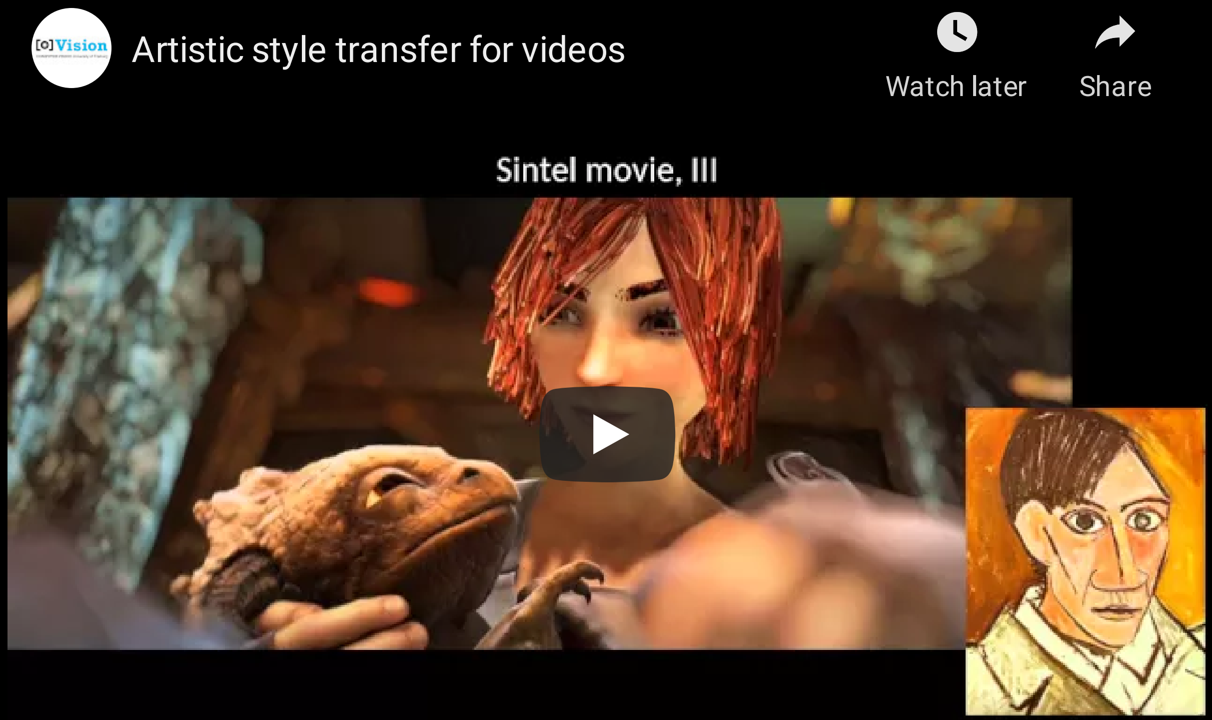


Physics simulation (Sanchez-Gonzalez et al, 2020)

AI for Science (Deepmind, AlphaFold, 2020)



Speech synthesis and question answering (Google, 2018)



Artistic style transfer (Ruder et al, 2016)

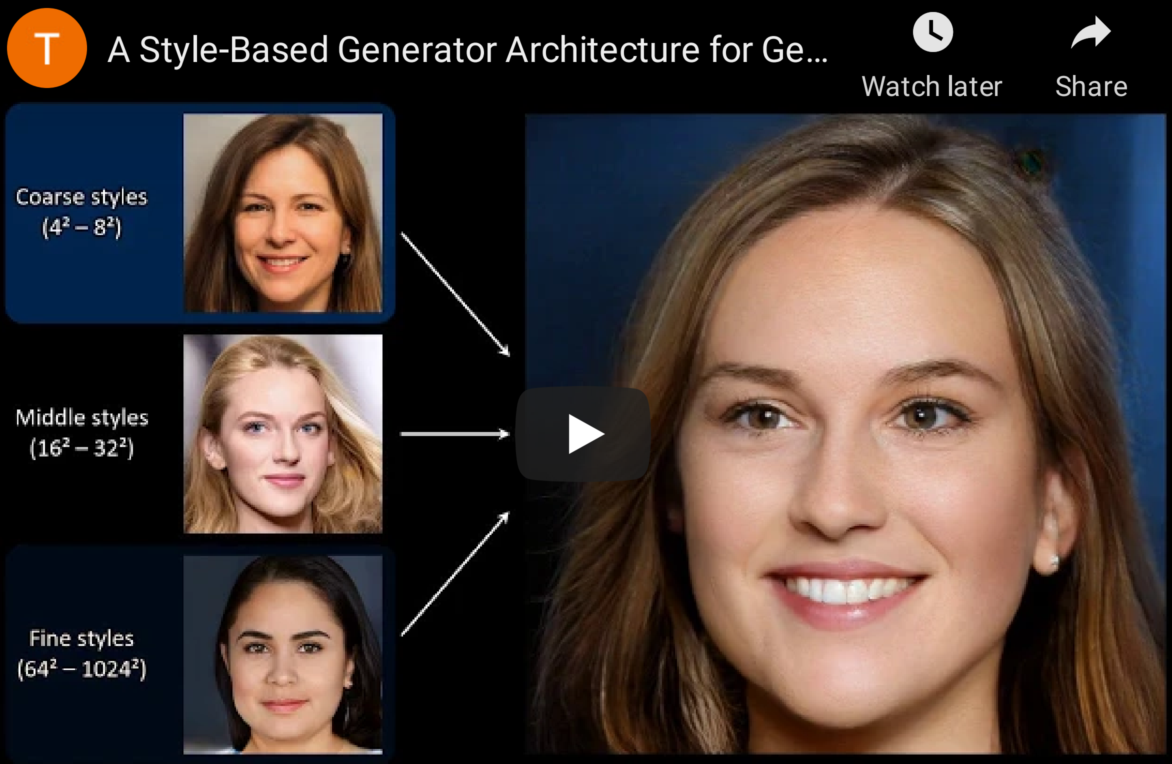
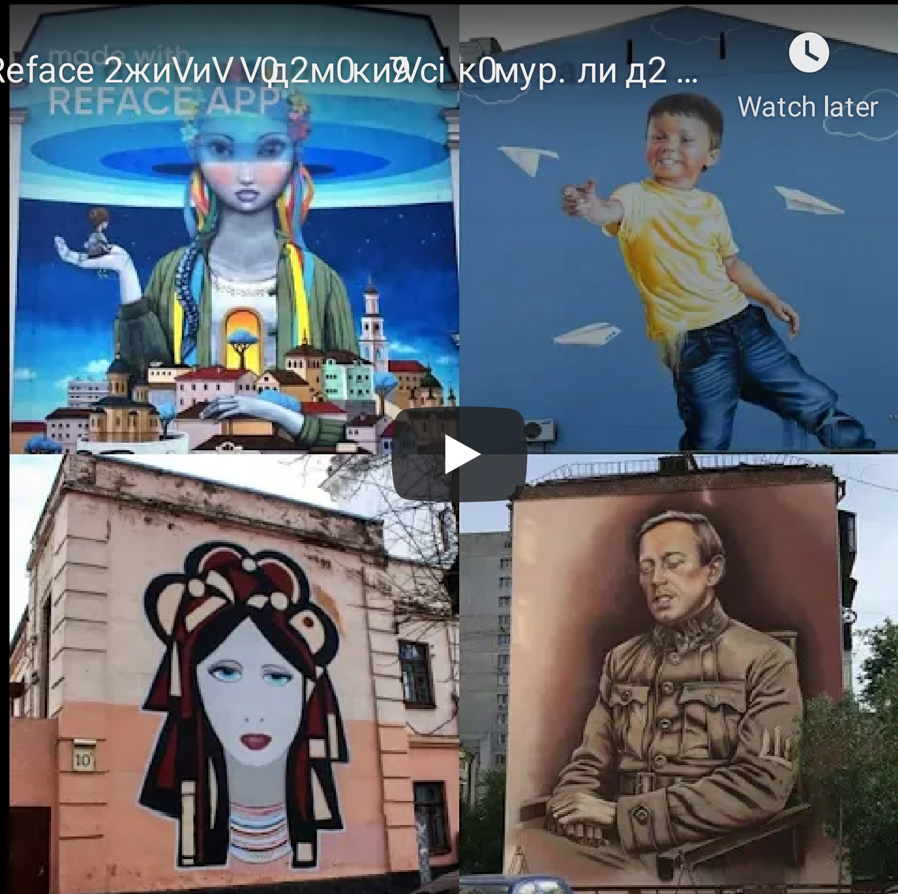


Image generation (Karras et al, 2018)

Music composition (NVIDIA, 2017)

Dali Lives (2019)

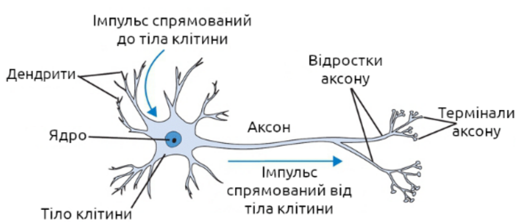
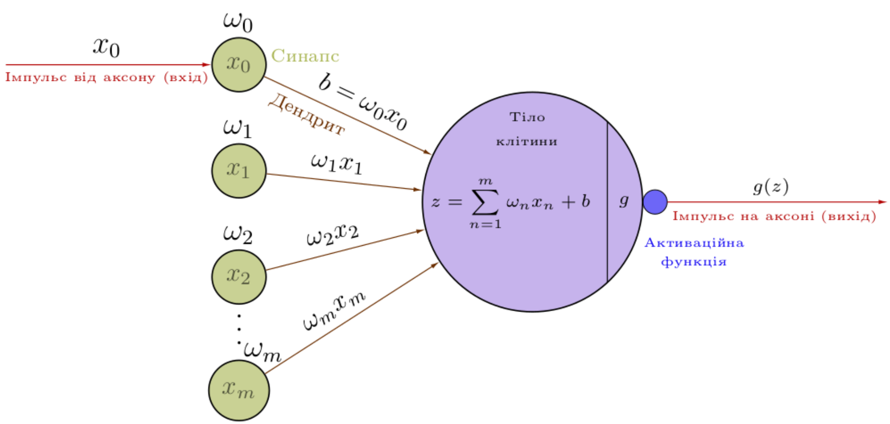


Reface revived the famous Kyiv murals for Kyiv Day (2021)

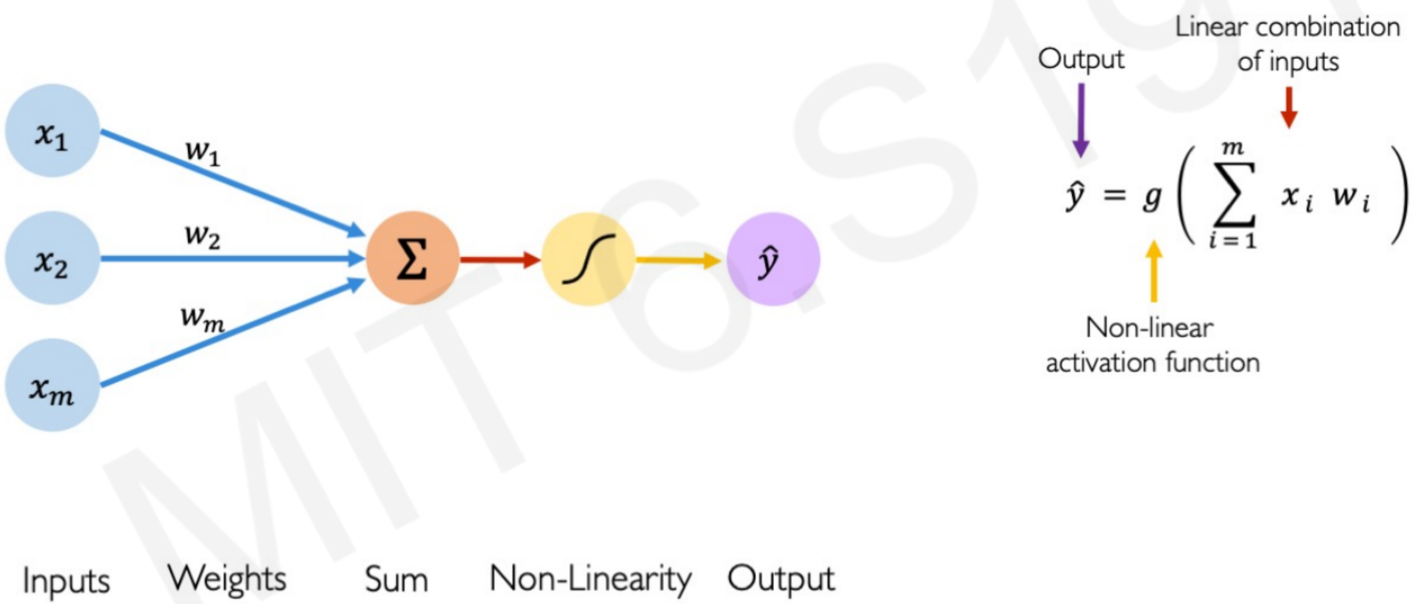
The Preceptron

The structural building block of deep learning

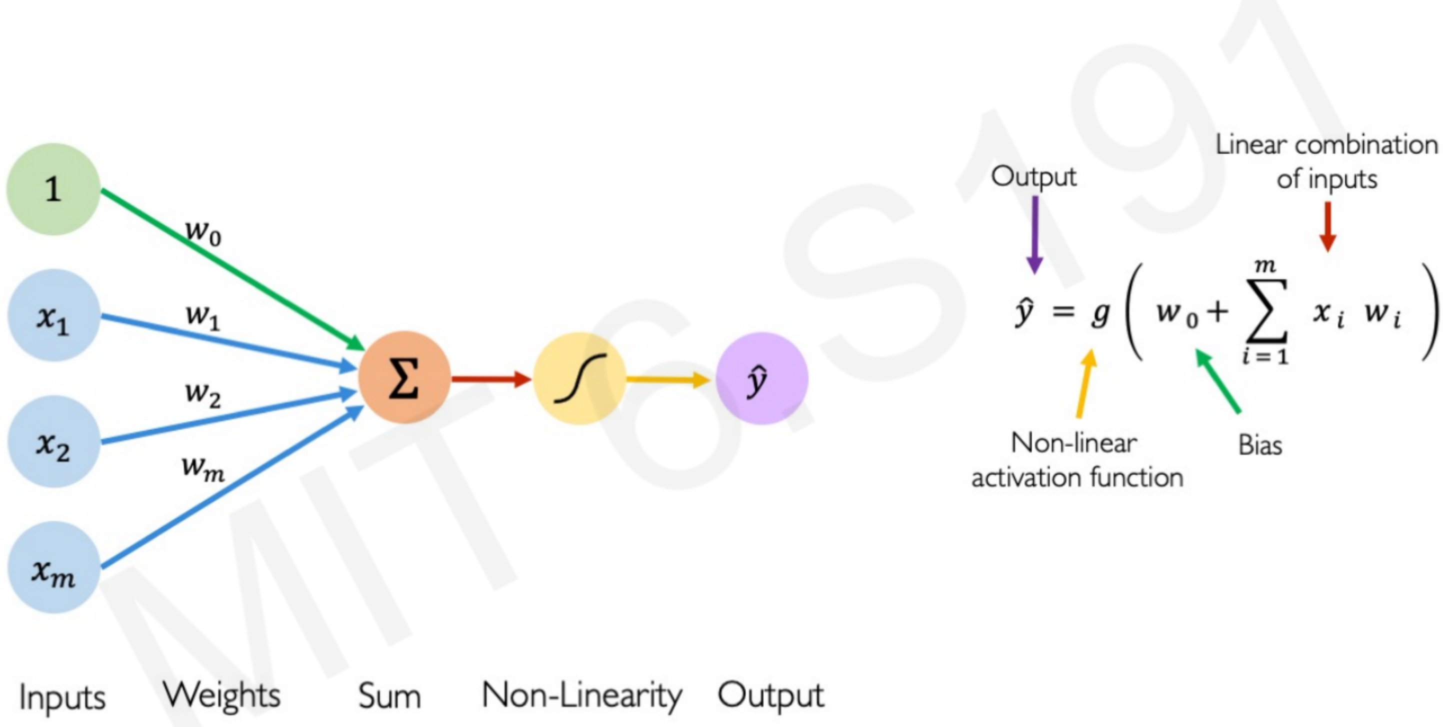
Biological vs Artificial neuron

<table>
<tr><th>Біологічний нейрон</th><th>Штучний нейрон</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

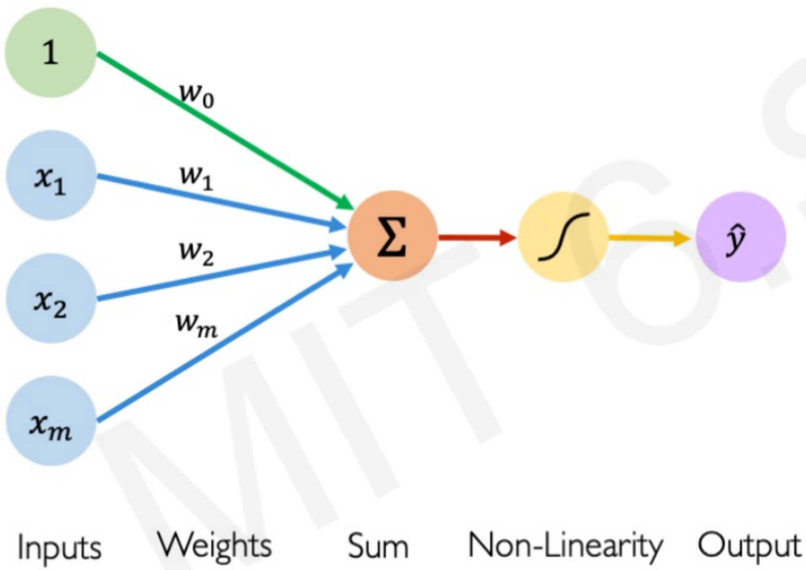
The Perceptron: Forward Propagation



The Perceptron: Forward Propagation



The Perceptron: Forward Propagation

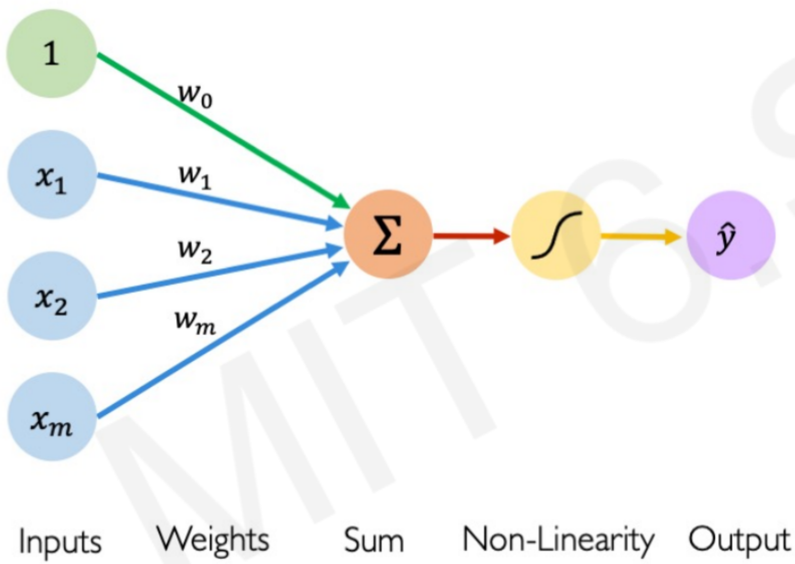


$$\hat{y} = g\left(w_0 + \sum_{i=1}^{m} x_i \, w_i \right)$$

$$\hat{y} = g\left(w_0 + \boldsymbol{X}^T \boldsymbol{W} \right)$$

where: $\boldsymbol{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ and $\boldsymbol{W} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$

The Perceptron: Forward Propagation

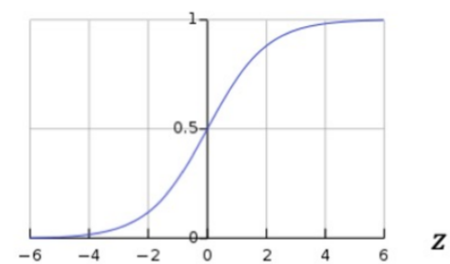


Activation Functions

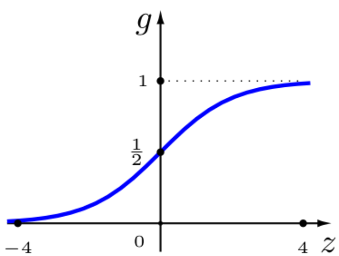
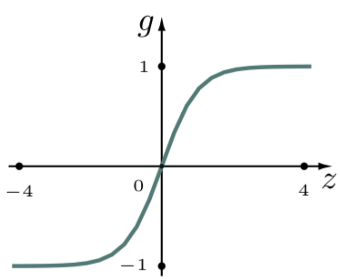
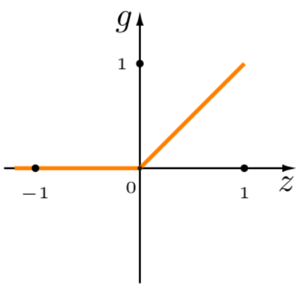
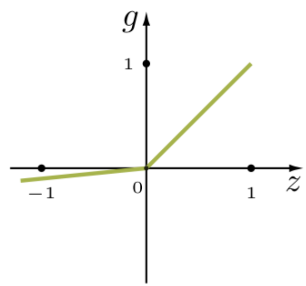
$$\hat{y} = g\left(w_0 + \boldsymbol{X}^T \boldsymbol{W} \right)$$

- Example: sigmoid function

$$g(z) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

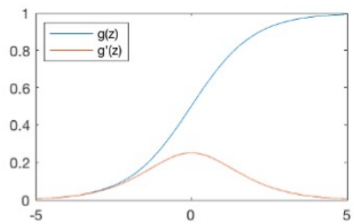


Common Activation Functions

Sigmoid	**Tanh**	**ReLU**	**Leaky ReLU**
$g(z) = \frac{1}{1+e^{-z}}$	$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	$g(z) = max(0, z)$	$g(z) = max(\epsilon z, z)$ $\epsilon \ll 1$
			

Common Activation Functions

Sigmoid Function

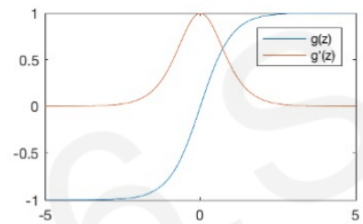


$$g(z) = \frac{1}{1 + e^{-z}}$$

$$g'(z) = g(z)(1 - g(z))$$

```
tf.math.sigmoid(z)
```

Hyperbolic Tangent

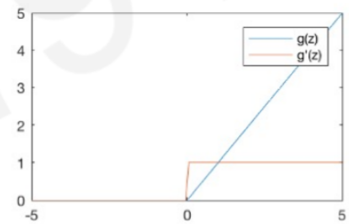


$$g(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$$

$$g'(z) = 1 - g(z)^2$$

```
tf.math.tanh(z)
```

Rectified Linear Unit (ReLU)



$$g(z) = \max(0, z)$$

$$g'(z) = \begin{cases} 1, & z > 0 \\ 0, & \text{otherwise} \end{cases}$$

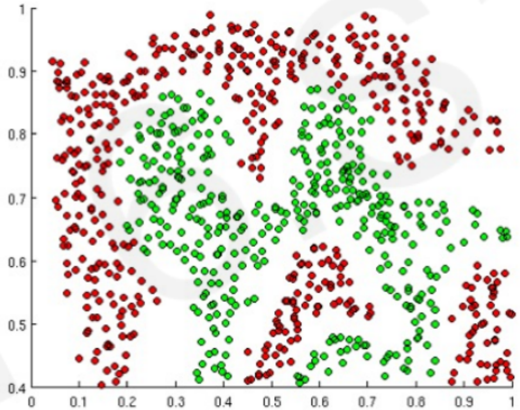
```
tf.nn.relu(z)
```

`TensorFlow code blocks`

NOTE: All activation functions are non-linear

Importance of Activation Functions

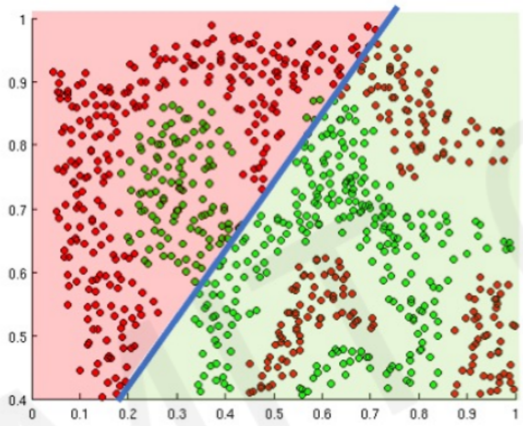
The purpose of activation functions is to ***introduce non-linearities*** *into the network*



What if we wanted to build a neural network to distinguish green vs red points?

Importance of Activation Functions

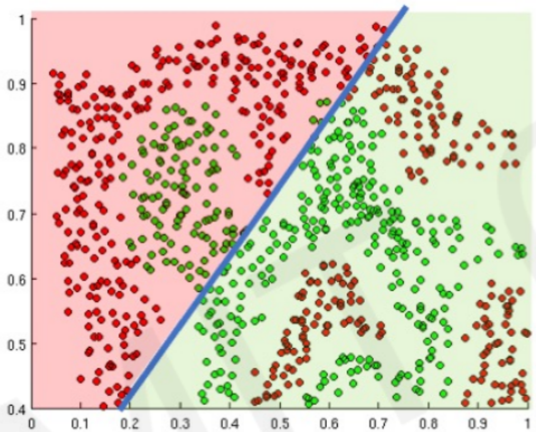
The purpose of activation functions is to ***introduce non-linearities*** *into the network*



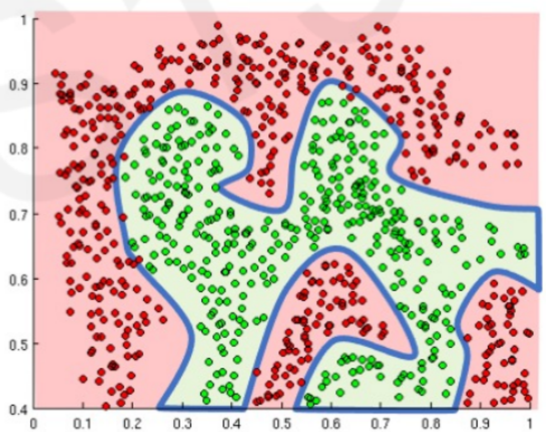
Linear activation functions produce linear decisions no matter the network size

Importance of Activation Functions

The purpose of activation functions is to ***introduce non-linearities*** *into the network*

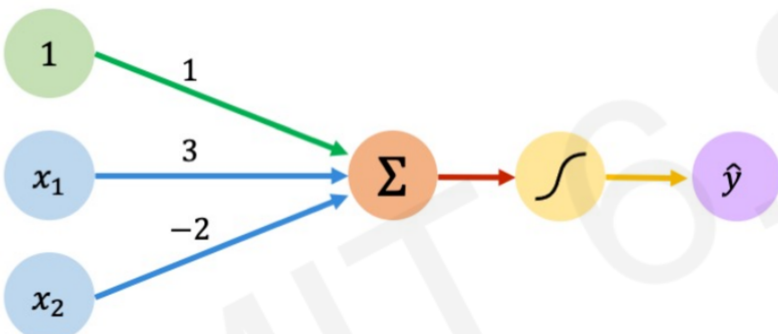


Linear activation functions produce linear decisions no matter the network size



Non-linearities allow us to approximate arbitrarily complex functions

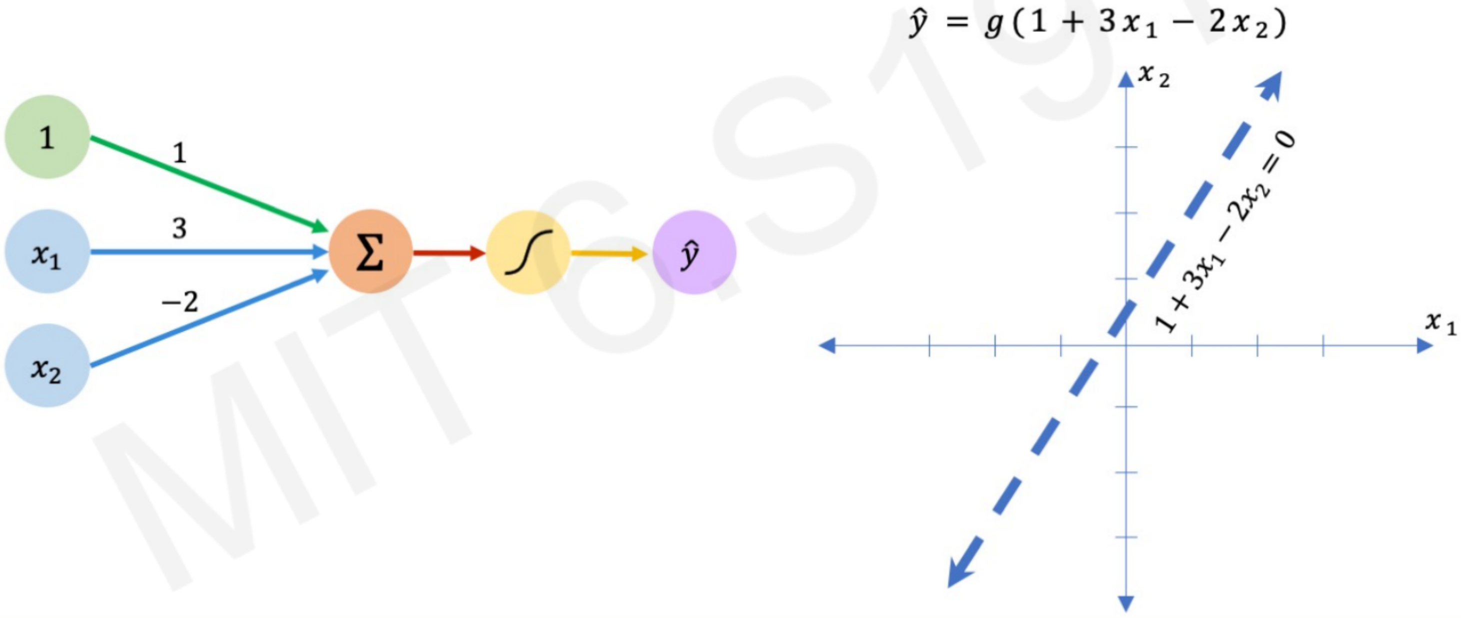
The Perceptron: Example



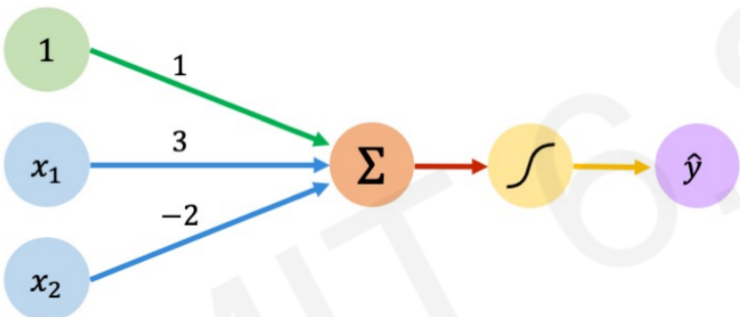
We have: $w_0 = 1$ and $\boldsymbol{W} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

$$\hat{y} = g\left(w_0 + \boldsymbol{X}^T \boldsymbol{W} \right)$$
$$= g\left(1 + \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 3 \\ -2 \end{bmatrix}\right)$$
$$\hat{y} = g\big(\underbrace{1 + 3x_1 - 2x_2}_{\text{This is just a line in 2D!}}\big)$$

The Perceptron: Example

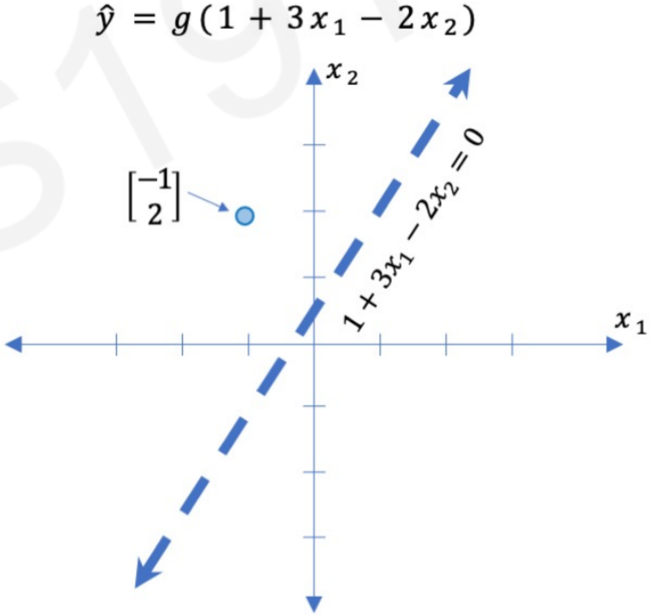


The Perceptron: Example

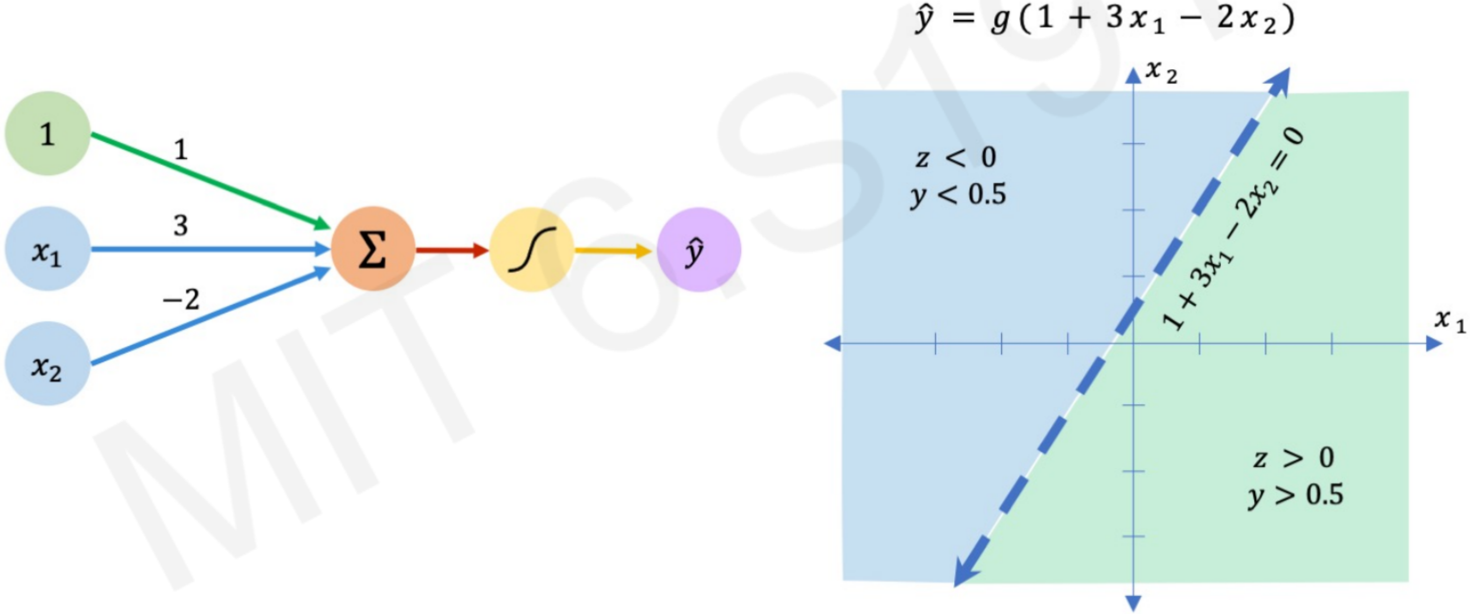


Assume we have input: $\boldsymbol{X} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

$$\hat{y} = g\big(1 + (3*-1) - (2*2)\big)$$
$$= g(-6) \approx 0.002$$



The Perceptron: Example



The end