

# Deep Learning (CS324)

## Introduction

Prof. Jianguo Zhang

Department of Computer Science and  
Engineering

SUSTech

# Deep Learning

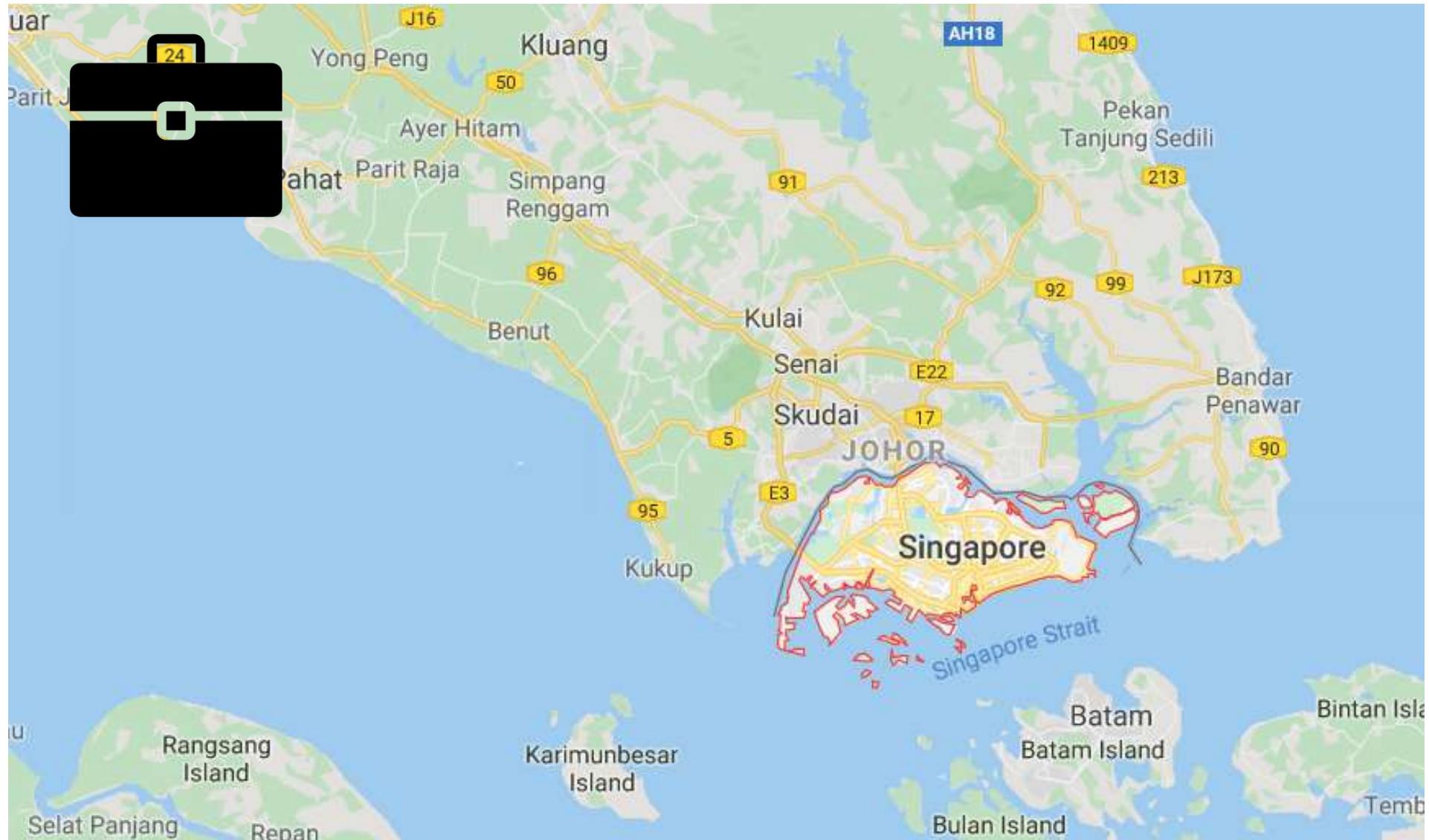
- This course (CS324) will provide an introduction to the field of deep learning, from both a theoretical and practical point of view
- Instructor: Prof. **Jianguo Zhang**, [zhangjg@sustech.edu.cn](mailto:zhangjg@sustech.edu.cn)
- Lab instructor: Dr. Weiyu Wang
- Lab TA/SA:
  - Jincan Lou
  - Zide Yang,
  - Sunbin Chen
  - Biaoxiang Chen
- **On line teaching:** Blackboard on the university website.

# Who am I?

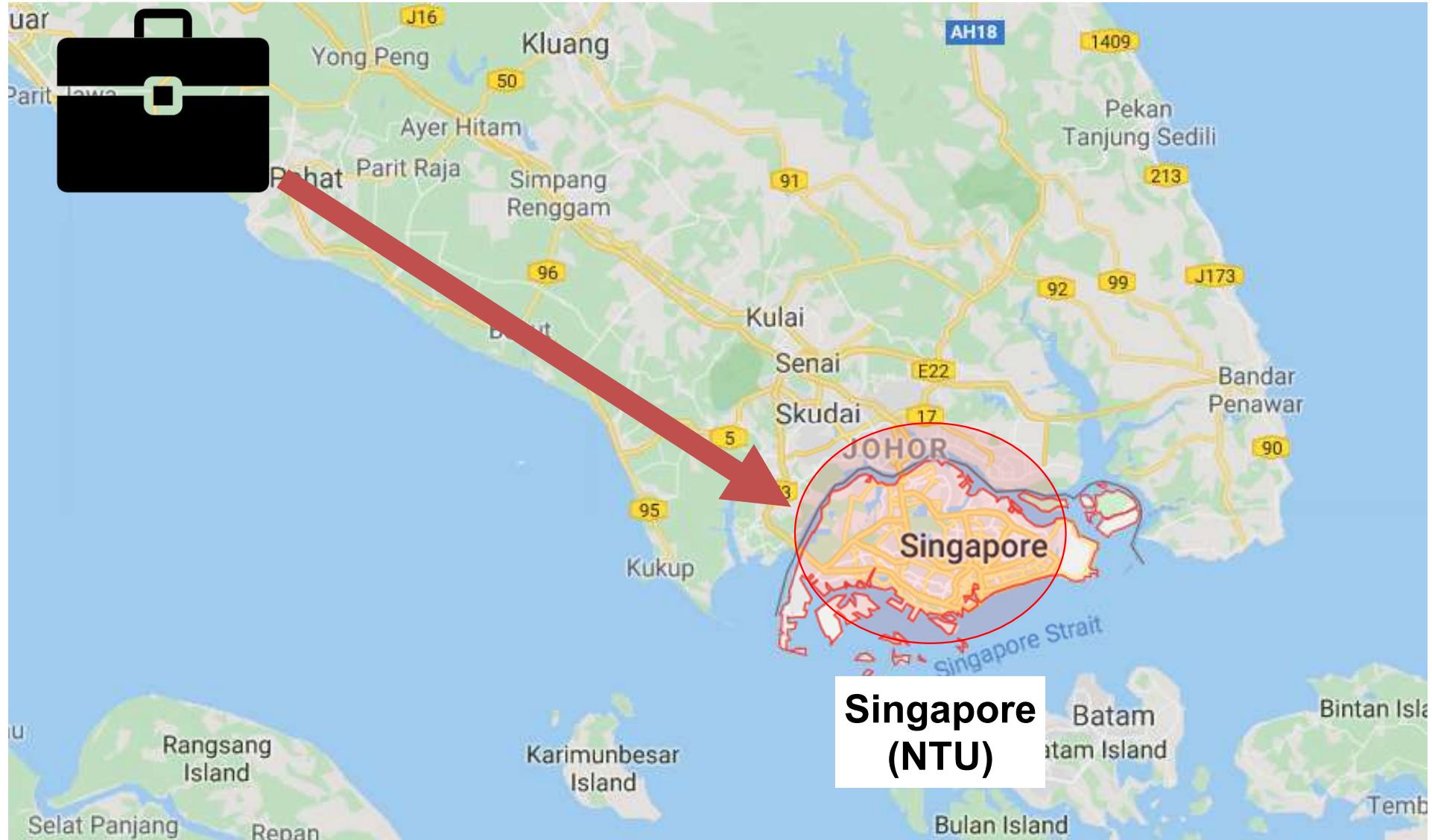
# Who am I?



# Who am I?



# Who am I?

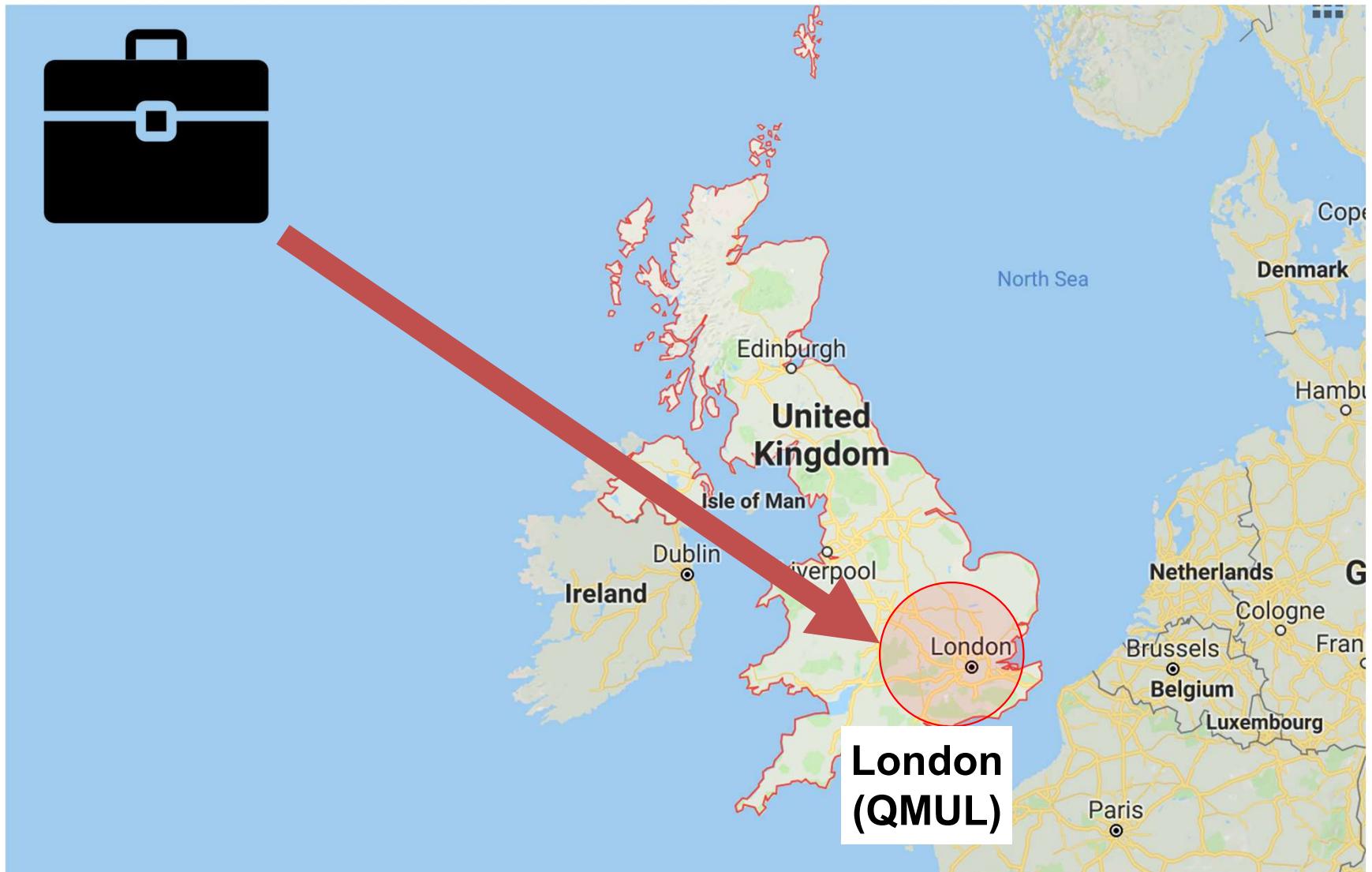


**Singapore  
(NTU)**

# Who am I?



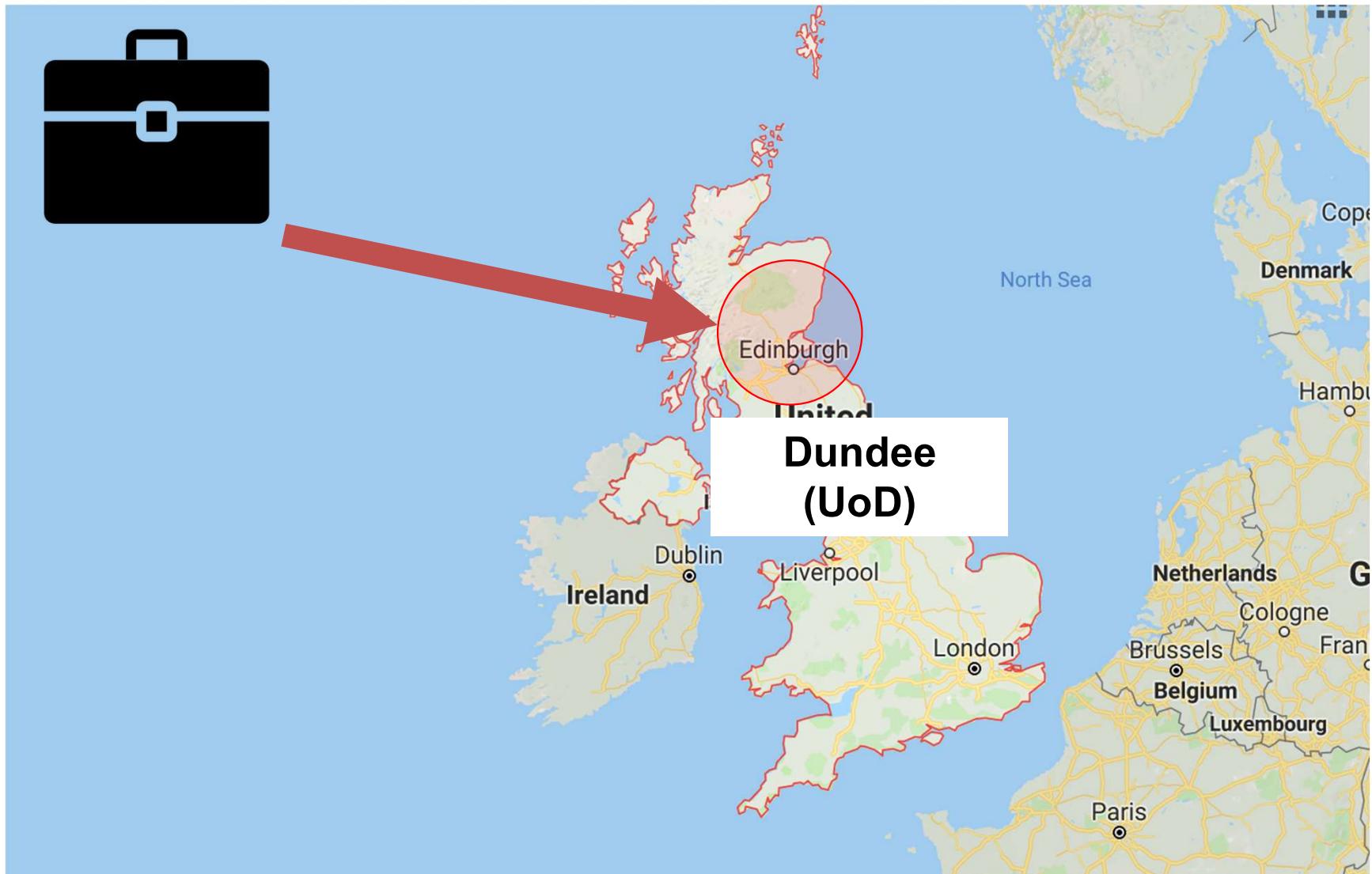
# Who am I?



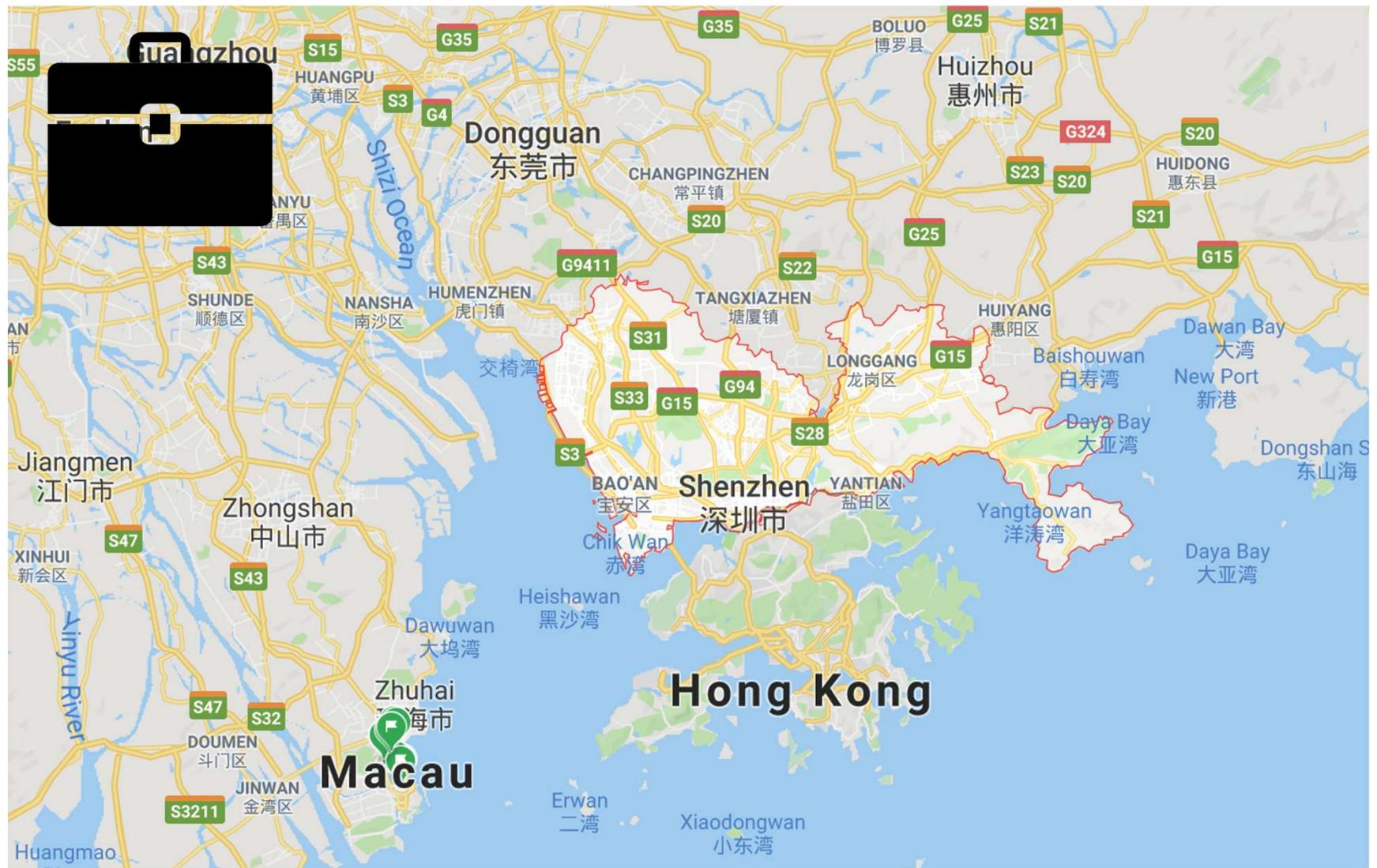
# Who am I?



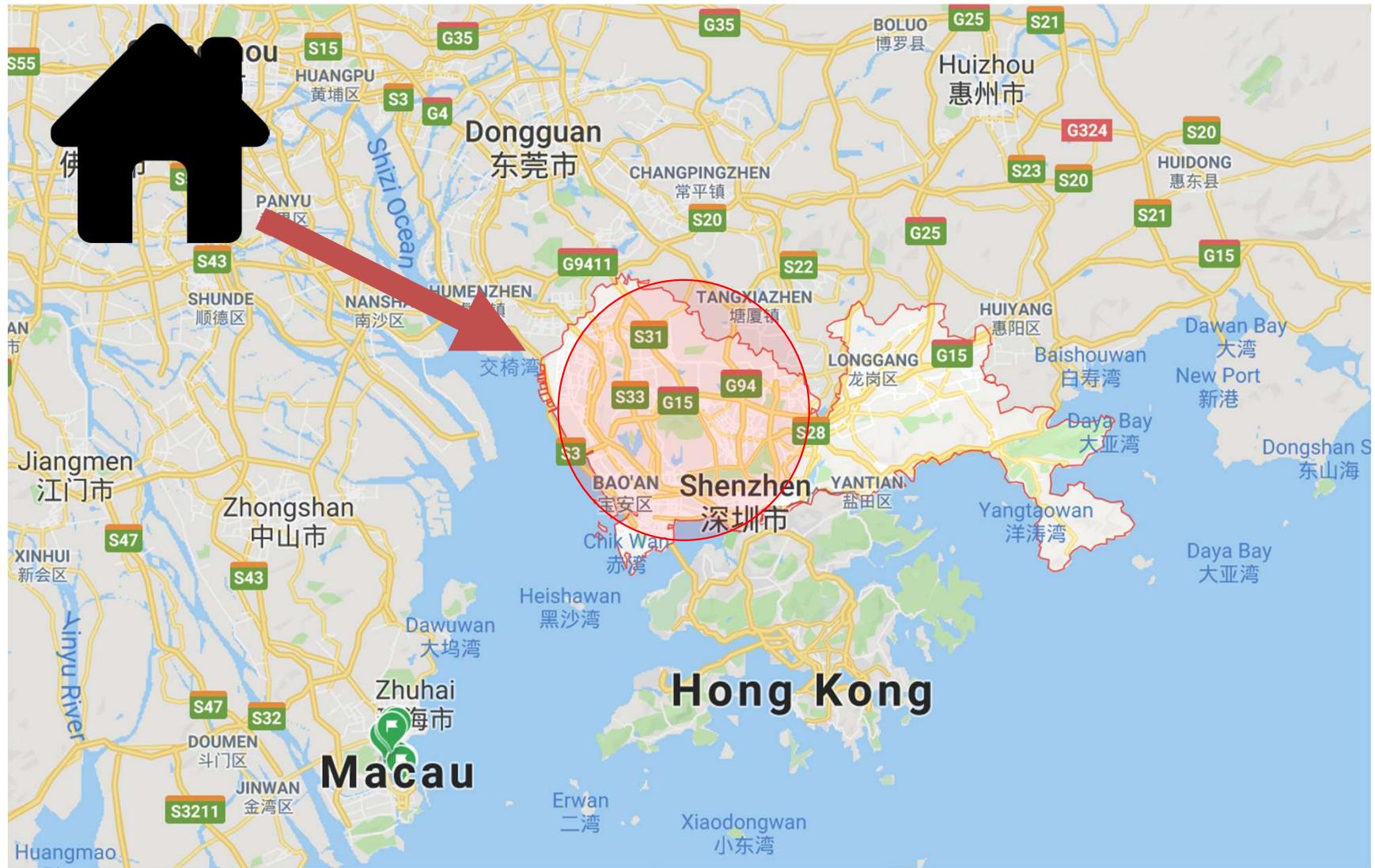
# Who am I?



# Who am I?



# Who am I?

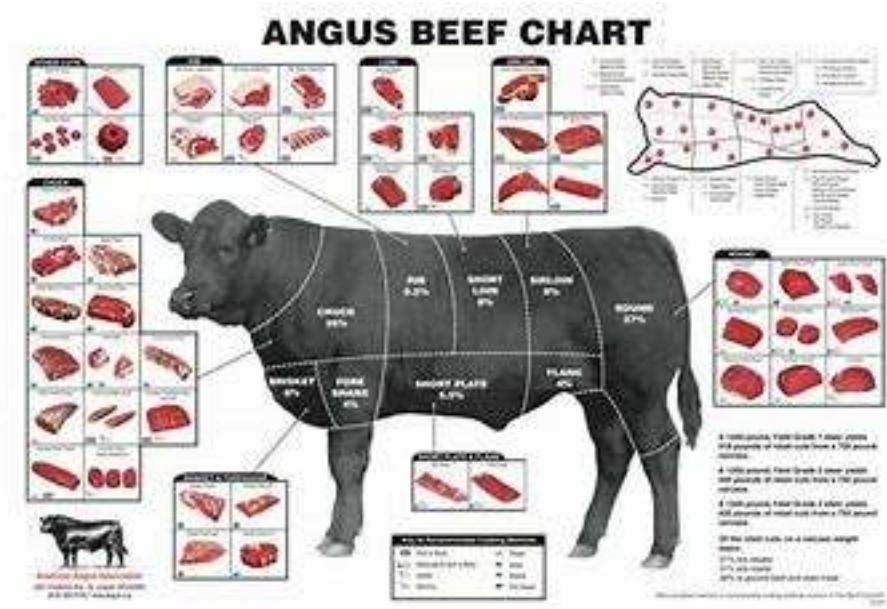


# A bit introduction about myself



**Jianguo Zhang**  
**Former Affiliation**  
**Computing,**  
**School of Science and Engineering**  
**University of Dundee**





# My working experience



**2002, PhD - CAS**

**2002-03, NTU, Singapore**

**2003-05: Lear group, France**

**2005-2007 - Queen Mary, University of London, UK**

**2007- 2010 – Queen’s University, Belfast, UK**

**2010 – 2019 – University of Dundee, UK**

**2019 – present, SUSTech, China**



**OLD: Research directions:** Medical image processing (retina, OPT, tracking 3D objects, etc) and surveillance, Biometrics, action recognition etc

- Brain Tumour Segmentation (MR/Microscopy)**
- OPT Colorectal Polyp Analysis**
- Endoscopic Polyps in Colons**
- Human Action/Activity Recognition (RGB-D)**
- Robust Person Re-identification**



## SUSTech: Research directions: Computer Vision and Medical Image Analysis

- Multimodal Brain Lesion Segmentation (MR/CT)**
- Semi- and Self-supervised Learning**
- Attack and Defence in Deep Learning Models**
- Human Action/Activity Recognition/Localisation**
- Lifelong Learning for Person Re-identification**
- Attention and Uncertainties**



# My Publications



## Journals:

IJCV

IEEE TPAMI

MIA

IEEE TMI

IEEE TIP

NeuroImage

Scientific report

Nature

## Conferences:

ICCV

CVPR

ECCV

MICCAI

AAAI

NeuRIPS

ICML

IJCAI

ICLR

If you are interested in doing a PhD,  
please come to see me at some  
stage.

# Communications and Feedback

1. Via Email, [zhangjg@sustech.edu.cn](mailto:zhangjg@sustech.edu.cn)
2. Pop into my office: RM313, Southern Block,  
College of Engineering, based on appointment
3. Time: In class and labs

# Who am I?

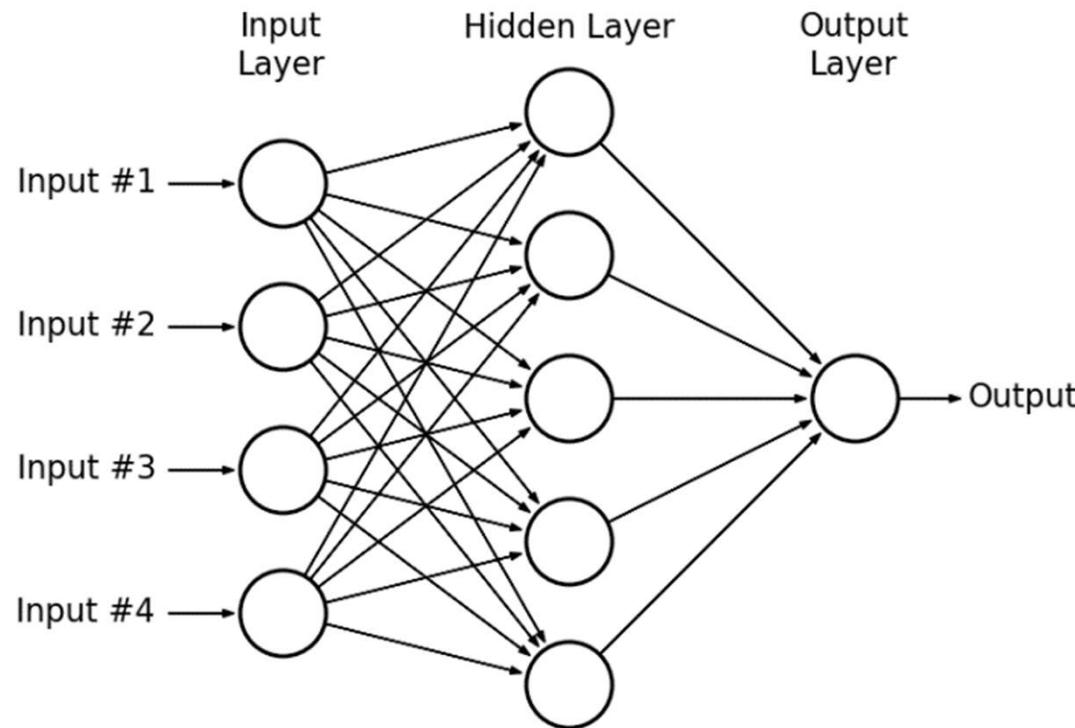
- My research interests in AI:
  - Computer Vision
  - Pattern Recognition and multimodal analysis
    - Object recognition
    - Image segmentation
  - Medical Image Analysis
    - Multimodal Brain lesion (tumour) segmentation
  - Machine Learning
    - OOD, Calibration, Uncertainty

# Module content

- Topics we will cover in this module:

# Module content

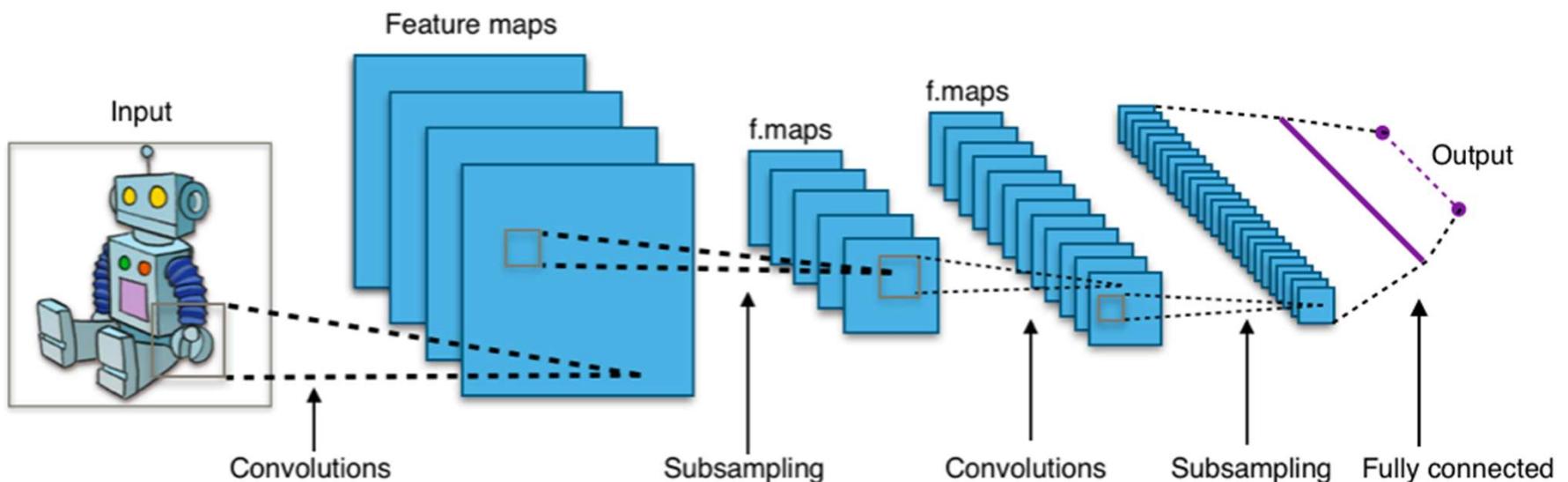
- Topics we will cover in this module:



**Multi-layer perceptrons (MLPs)**

# Module content

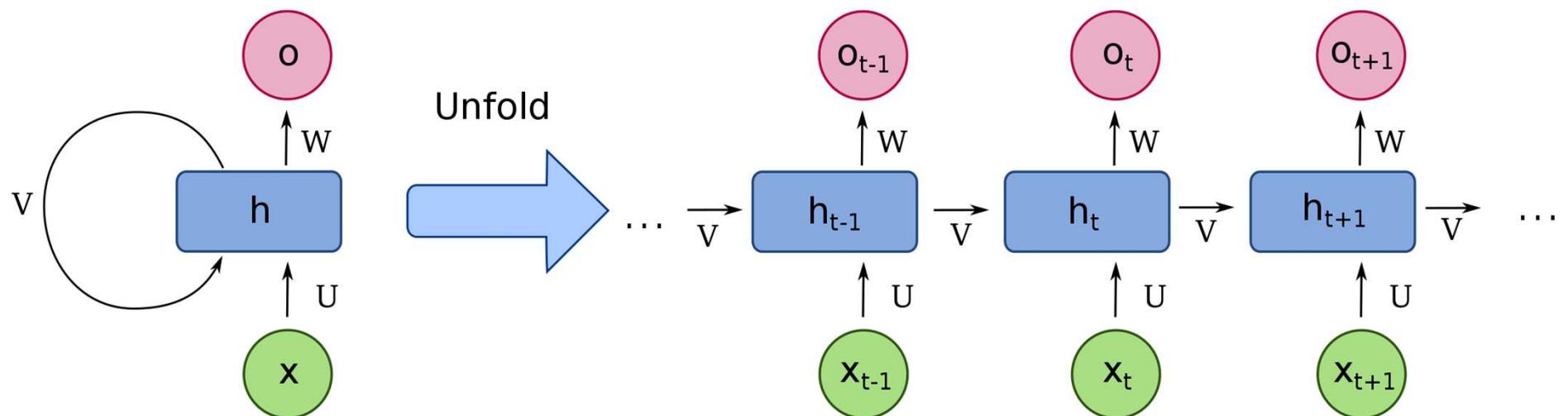
- Topics we will cover in this module:



## Convolutional neural networks (CNNs)

# Module content

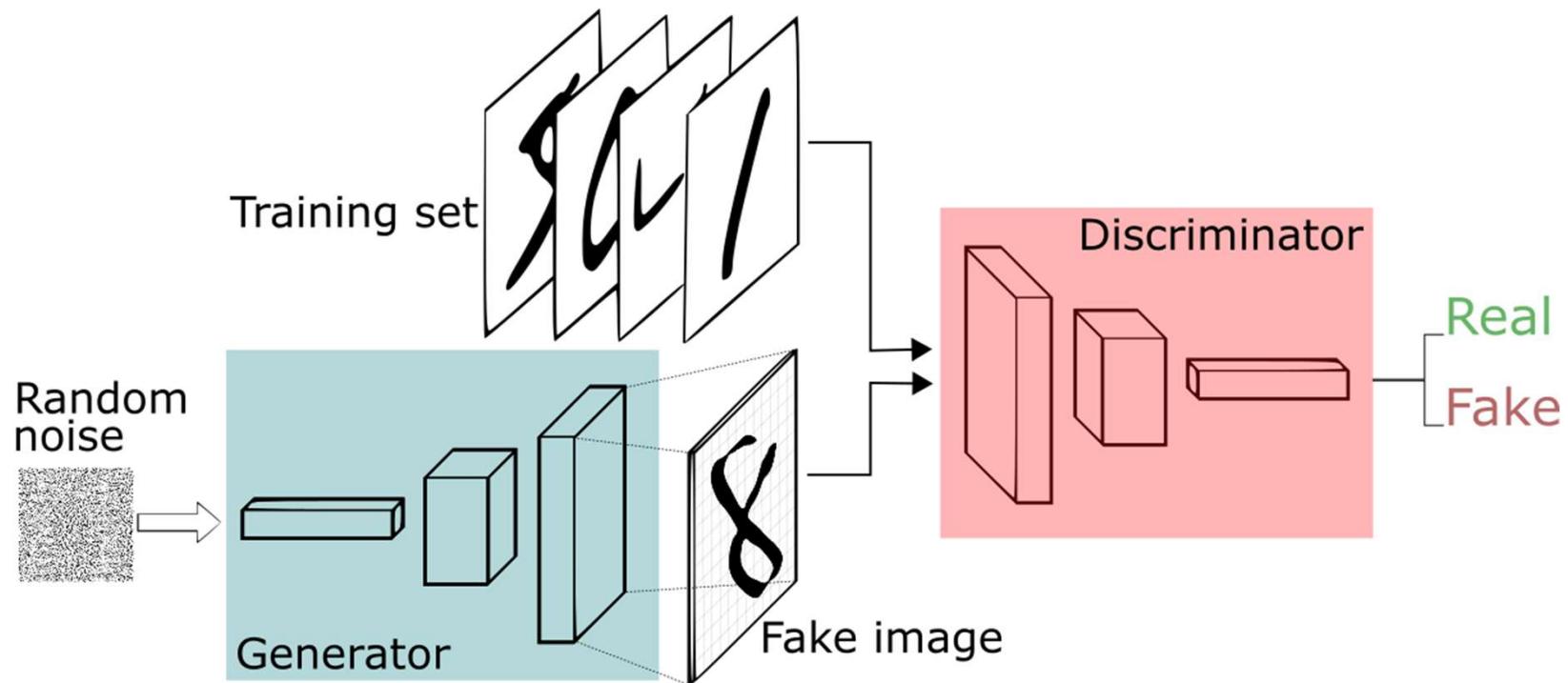
- Topics we will cover in this module:



## Recurrent neural networks (RNNs)

# Module content

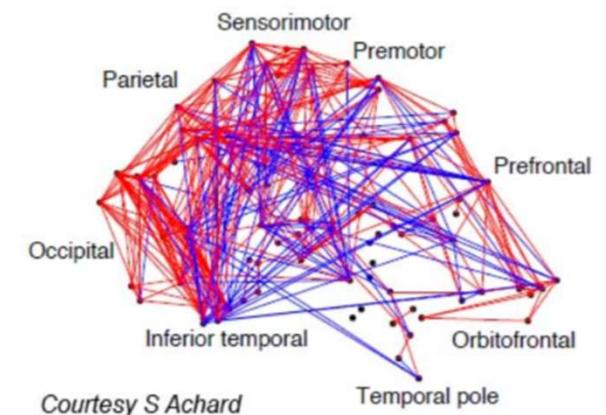
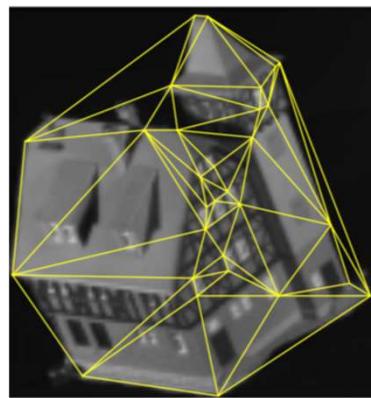
- Topics we will cover in this module:



**Generative adversarial networks (GANs)**

# Module content

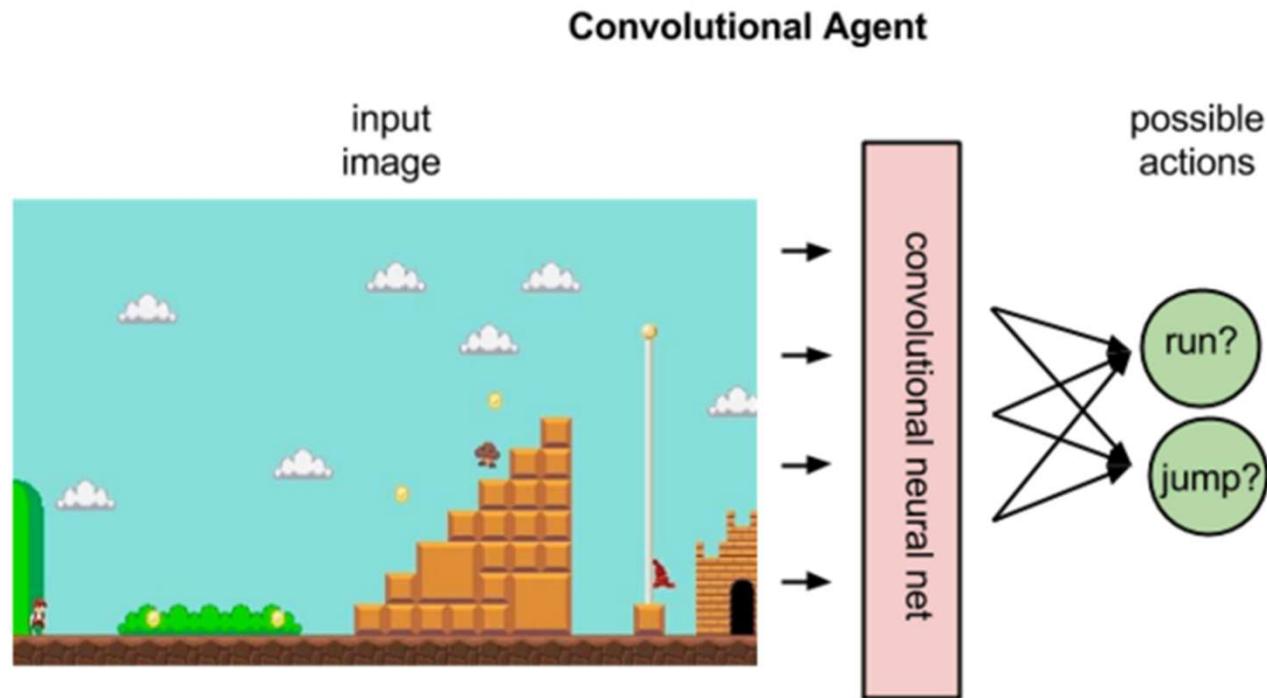
- Topics we will cover in this module:



**Deep learning on graphs**

# Module content

- Other Topics we will not cover in this module:

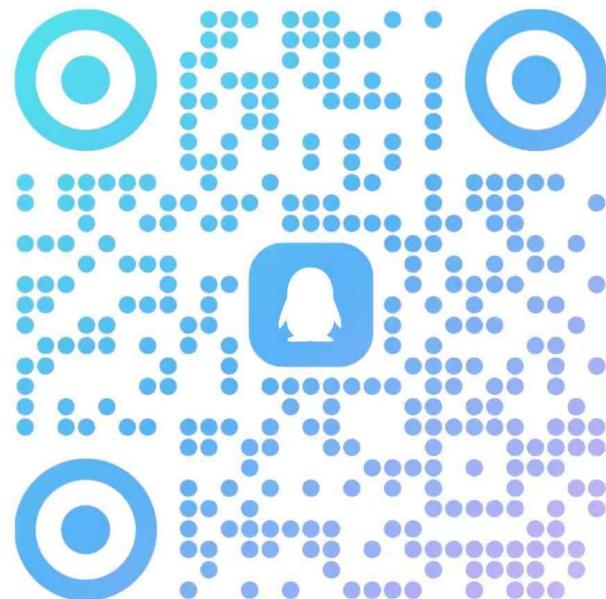


**Deep reinforcement learning**



Deep Learning Fall...

群号: 790947325



扫一扫二维码，加入群聊



QQ

# Tentative Schedule: Week 1

- **Introduction:** course overview; information on lecture & labs schedule; assessments structure and rules; what is deep learning and why are you here?
- **[Lab 1]** Introduction to Pytorch and assignments overview

# Tentative Schedule: Week 2

- **Where it all started:** brief recap of linear algebra; tensors; brief history of machine learning and recap of fundamental concepts; biological neurons; the perceptron
- **[Lab 2]** Assignment #1 is made available (MLP and back-propagation)

# Tentative Schedule: Week 3

- **Going deep:** shallow networks and the hidden layer; multi-layer perceptron; gradient descent; back-propagation
- **[Lab 3]** Working on assignment #1

# Tentative Schedule: Week 4

- **Optimisation:** batch gradient descent; stochastic gradient descent; challenges in optimisation; advanced techniques
- **[Lab 4]** Working on assignment #1

# Tentative Schedule: Week 5

- **Regularisation and good practices:** input normalisation; l1 and l2 regularisation; dropout; learning rate; weight initialisation
- **[Lab 5]** Working on assignment #1

# Tentative Schedule: Week 6

- **Convolutional neural networks, part one:** what are CNNs and what makes them special; their importance in computer vision; CNNs modules; how to train a CNN
- **[Lab 6]:** Assignment #2 is made available (CNNs and RNNs)

# Tentative Schedule: Week 7

- **Convolutional neural networks, part two:** popular modern CNNs architectures; vanishing gradients; inception model
- **[Lab 7]:** Working on assignment #2

# Tentative Schedule: Week 8

- **Recurrent neural networks:** sequential data; RNNs; back-propagation through time; exploding and vanishing gradients; LSTM architectures
- **[Lab 8]:** Working on assignment #2

# Tentative Schedule: Week 9

- **Auto-encoders:** supervised versus unsupervised learning; manifold hypothesis; PCA; kernel PCA; auto-encoders
- **[Lab 9]:** Working on assignment #2

# Tentative Schedule: Week 10

- **Generative adversarial networks:** generative vs discriminative models; generative adversarial networks; variants of GANs
- **[Lab 10]:** Assignment #3 is made available (GANs and VAEs)

# Tentative Schedule: Week 11

- **Explicit generative models:** variational inference; variational auto-encoders; restricted Boltzmann machines (optional); deep Boltzmann machines (optional);
- **[Lab 11]:** Working on assignment #3

# Tentative Schedule: Week 12

- **Adversarial learning – Attack and defense**  
: the safety issues of deep net; why deep net can be easily fooled; generating adversarial examples; different attack methods; defense;
- **[Lab 12]:** Working on assignment #3

# Tentative Schedule: Week 13

- **Beyond vectors:** graph-based learning and neural networks, part one: graph-based vs vector-based learning, advantages and problems; deep learning on graph data; graph CNNs
- **[Lab 13]:** Working on assignment #3

# Topics might not be covered in detail

- **Deep reinforcement learning:** what is reinforcement learning; Bellman equation; deep RL; Q-learning; stability problems; policy-based deep RL; model-based deep RL

# Tentative Schedule: Week 14-15

- **Students presentations** on selected papers about state-of-the-art architecture. The format of the presentations will vary depending on the size of the cohort
- Revision, Q&A

# Tentative Schedule: Week 14-15

- **Students presentations** on selected papers about state-of-the-art architecture. The format of the presentations will vary depending on the size of the cohort
- Revision, Q&A
- Flipped classroom activity: critical appraisal of deep learning, reflections and class discussion

# Tentative Schedule: Week 14-15

- **Students presentations** on selected papers about state-of-the-art architecture. The format of the presentations will vary depending on the size of the cohort
- Revision, Q&A
- ~~Flipped classroom activity: critical appraisal of deep learning, reflections and class discussion (?)~~

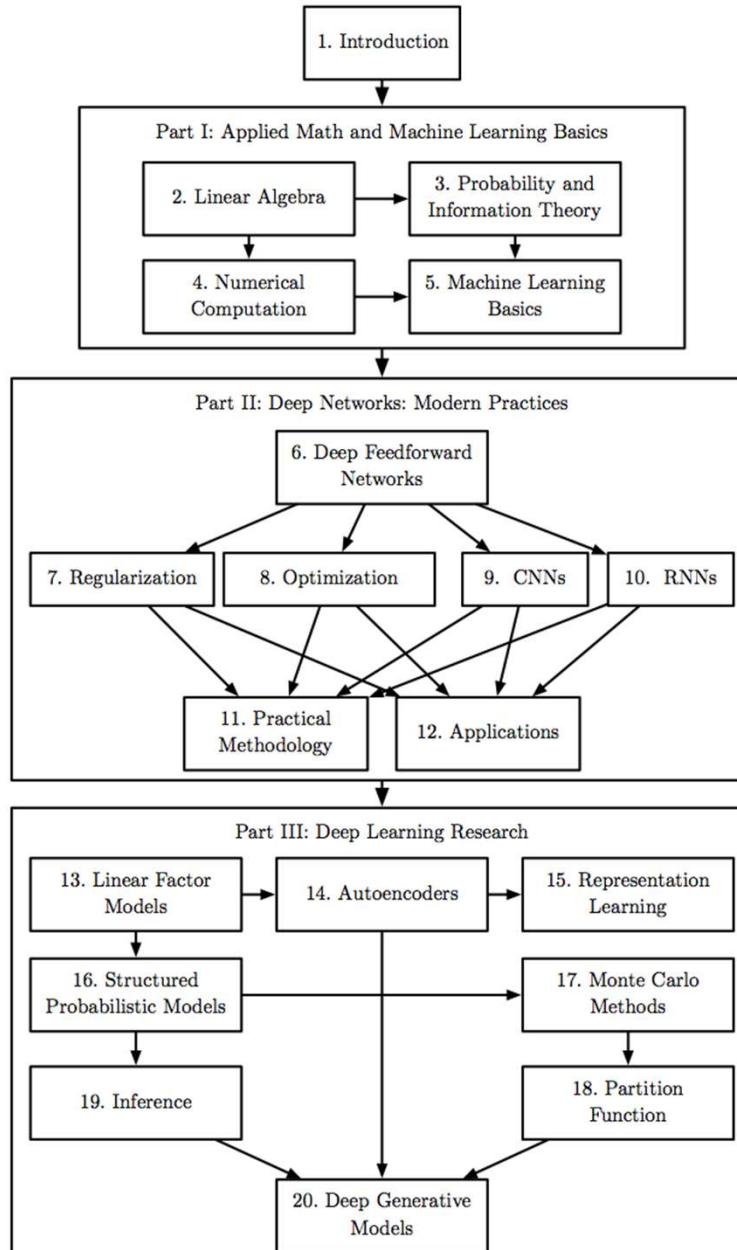
# Assessment

- **Final exam (40%)**
- **practical assignments** worth 15%+15% +15% (**45% total**)
  - Plagiarism will not be tolerated (remember to always cite your sources)
- Student **presentation (5%)**
  - On an advanced topic (to be decided 2/3 into the course) or on (Demo on Chips)
- Demo AI on Chips (**10%**) --- Online real-time GAN-based face interpolation (demo) or other demo

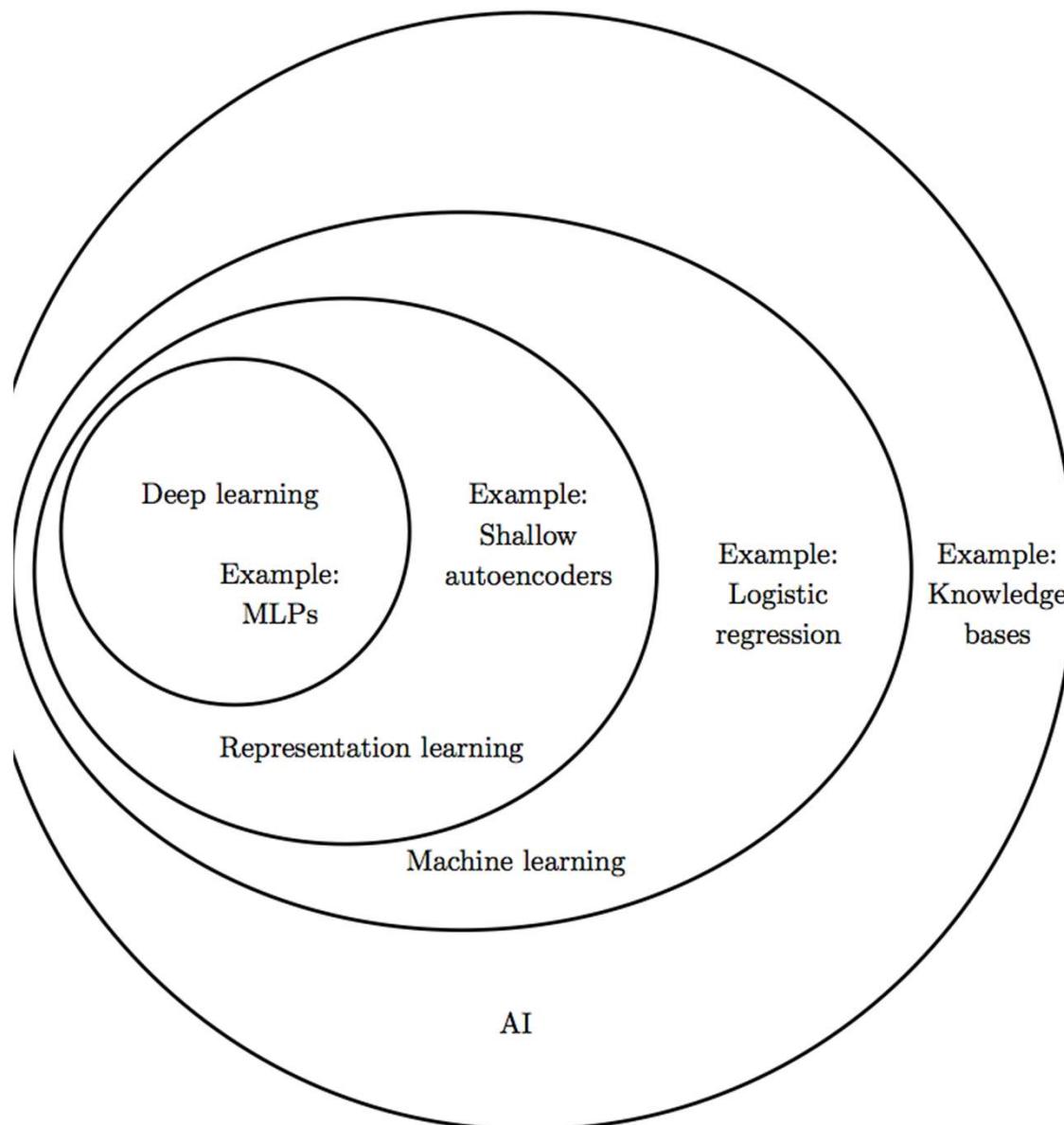
# Reading List

- **Main textbook:**
  - Goodfellow, Ian, Yoshua Bengio, Aaron Courville. Deep learning. Vol. 1. Cambridge: MIT press, 2016.  
<https://www.deeplearningbook.org/>
- *Other recommended books:*
  - Chollet, Francois. Deep learning with python. Manning Publications Co., 2017.
  - Nielsen, Michael A. Neural networks and deep learning. Vol. 25. USA: Determination press, 2015.
- But, most importantly, come to the lectures and study what's in the slides!

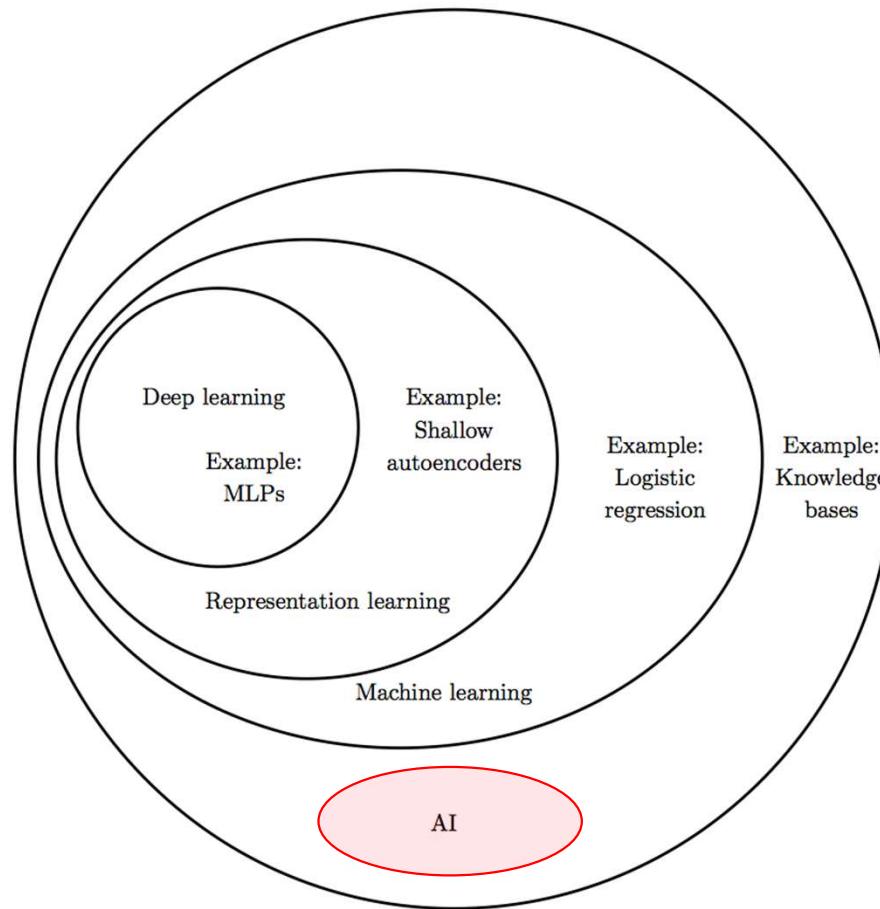
# Textbook organisation



# What is Deep Learning?

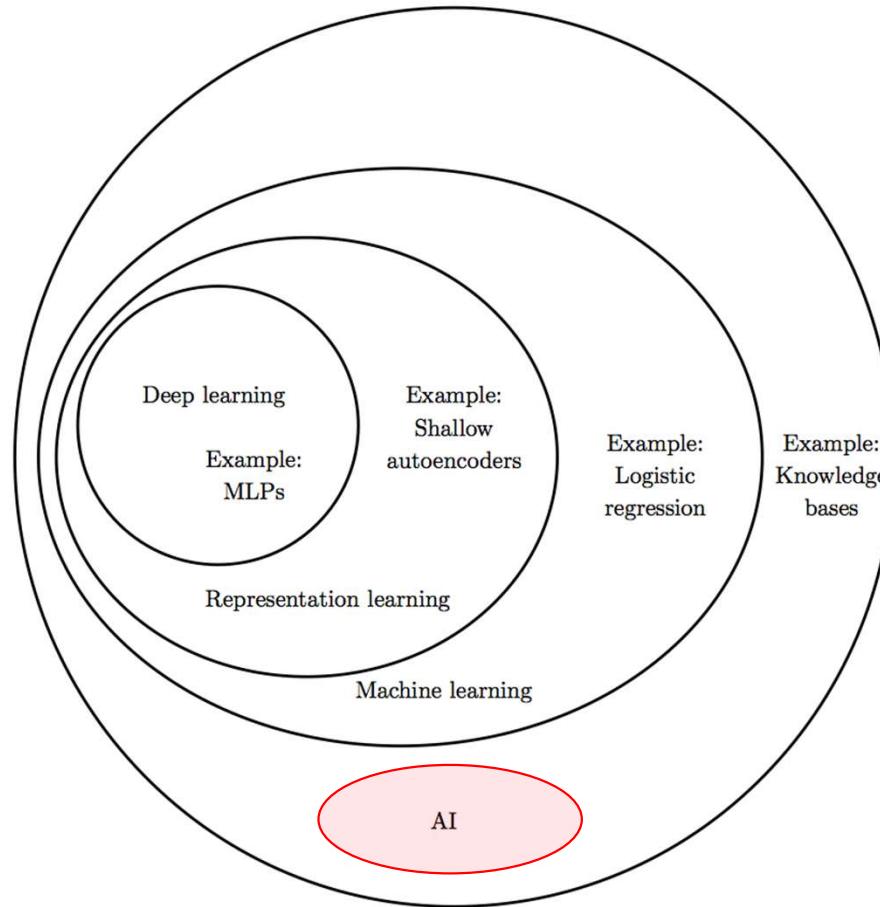


# What is Deep Learning?



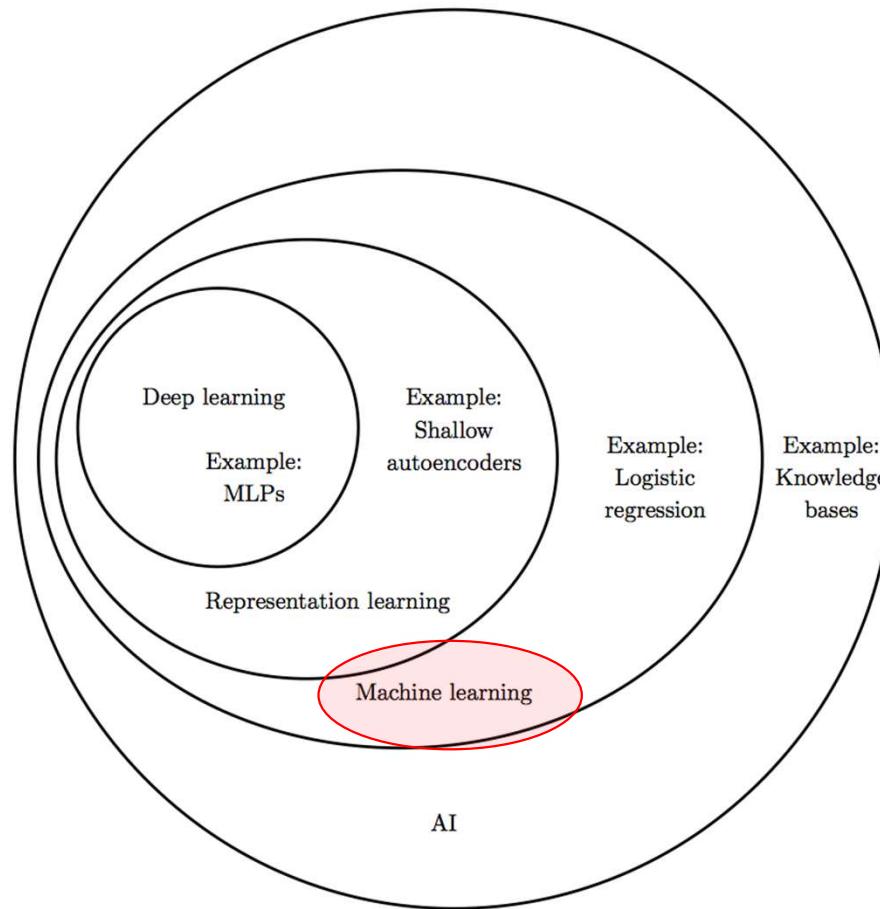
Can we make “intelligent” machines?  
Automate labour, assist in diagnosis, understand speech...

# What is Deep Learning?



Knowledge bases use hardcoded inference rules  
to reason automatically about statements...

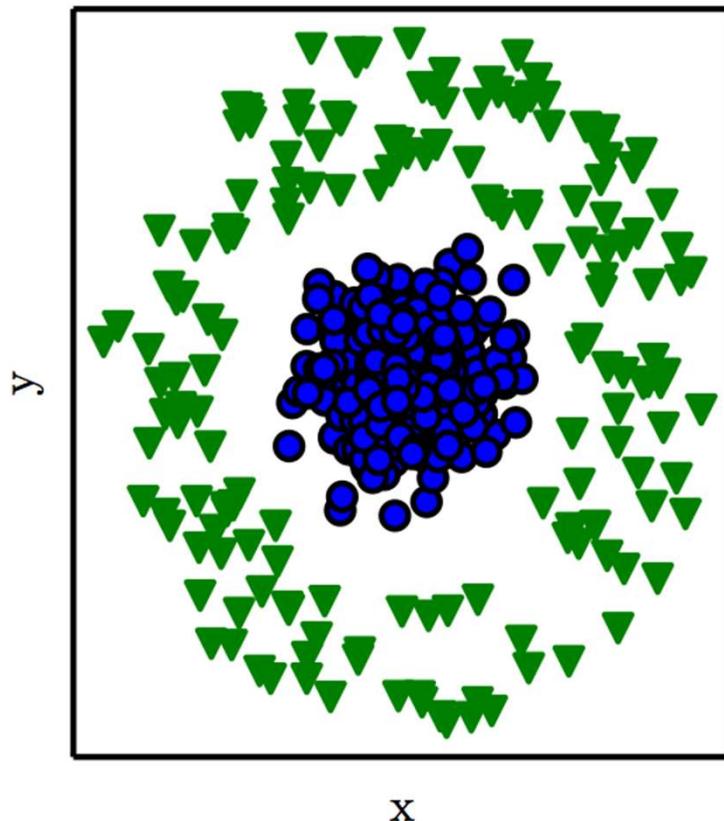
# What is Deep Learning?



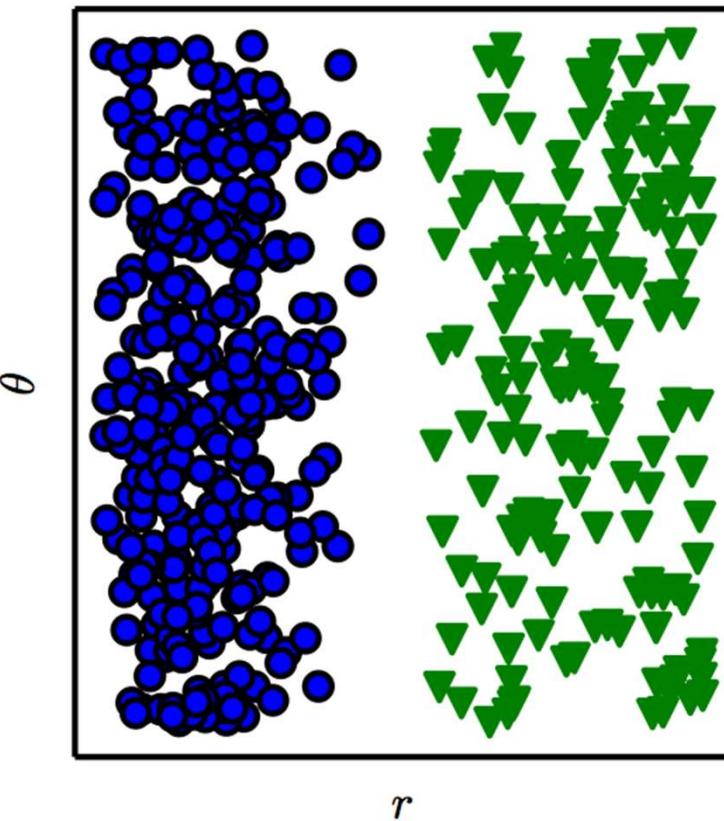
...but knowledge is difficult to hard-code, expensive and not flexible.  
Machine learning comes to the rescue: machines acquire  
their own knowledge by finding patterns in raw data!

# Representation matters

Cartesian coordinates

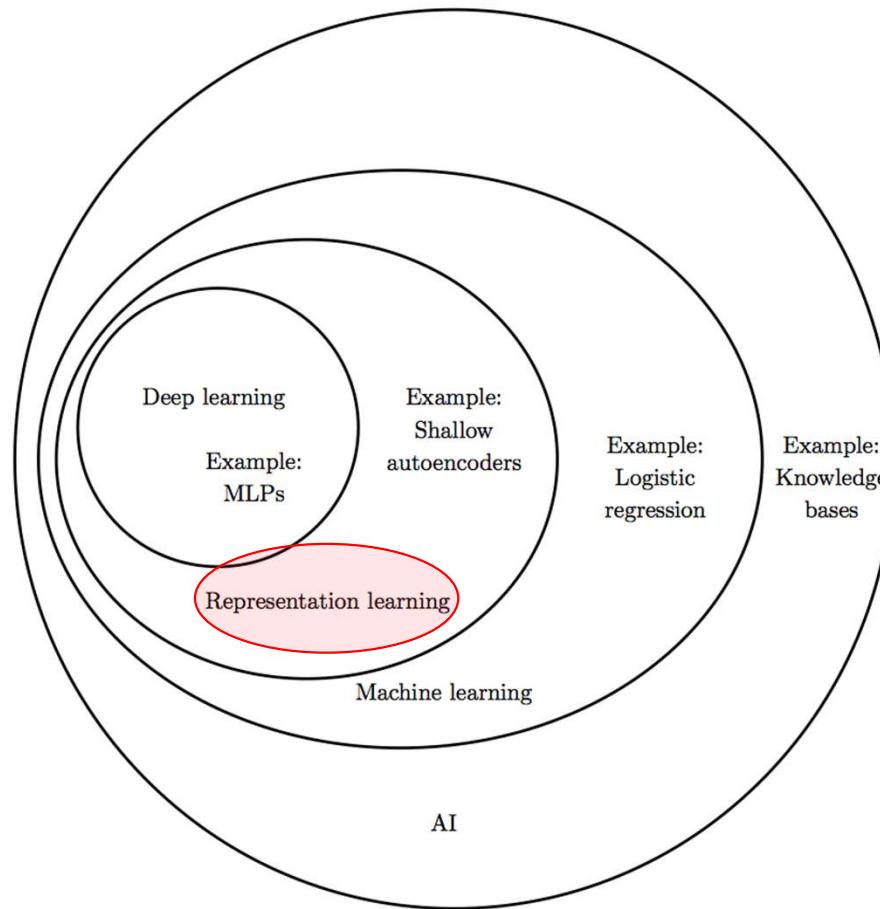


Polar coordinates



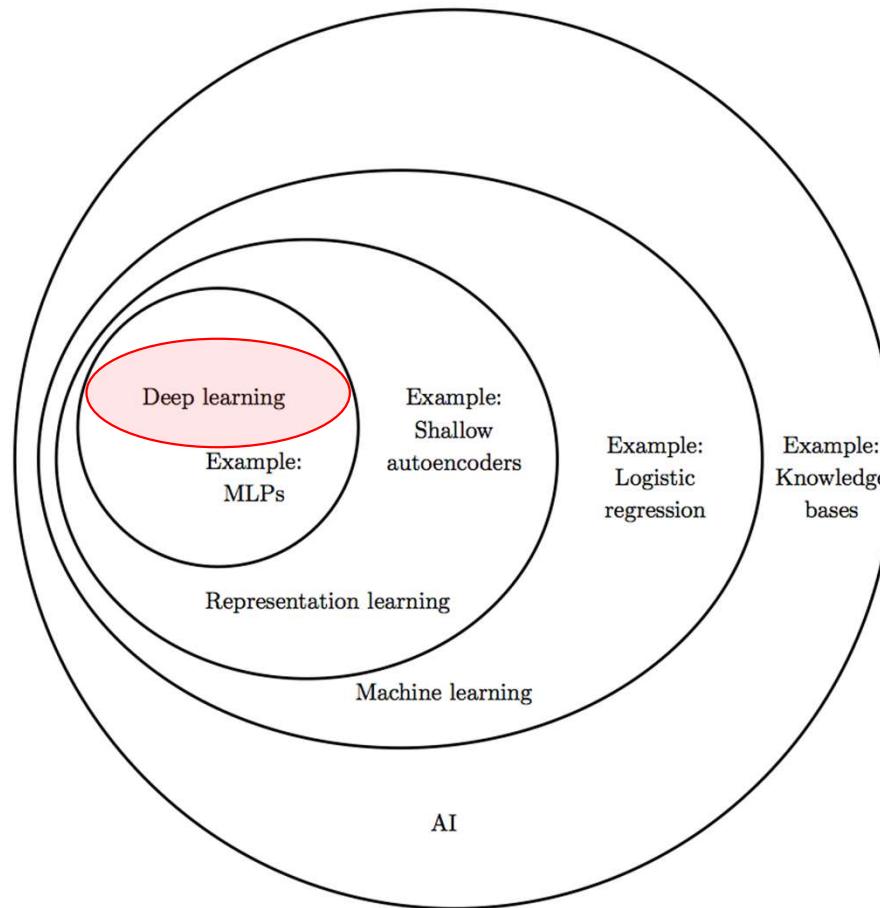
But we still need to feed the machines with a representation of the data...

# What is Deep Learning?



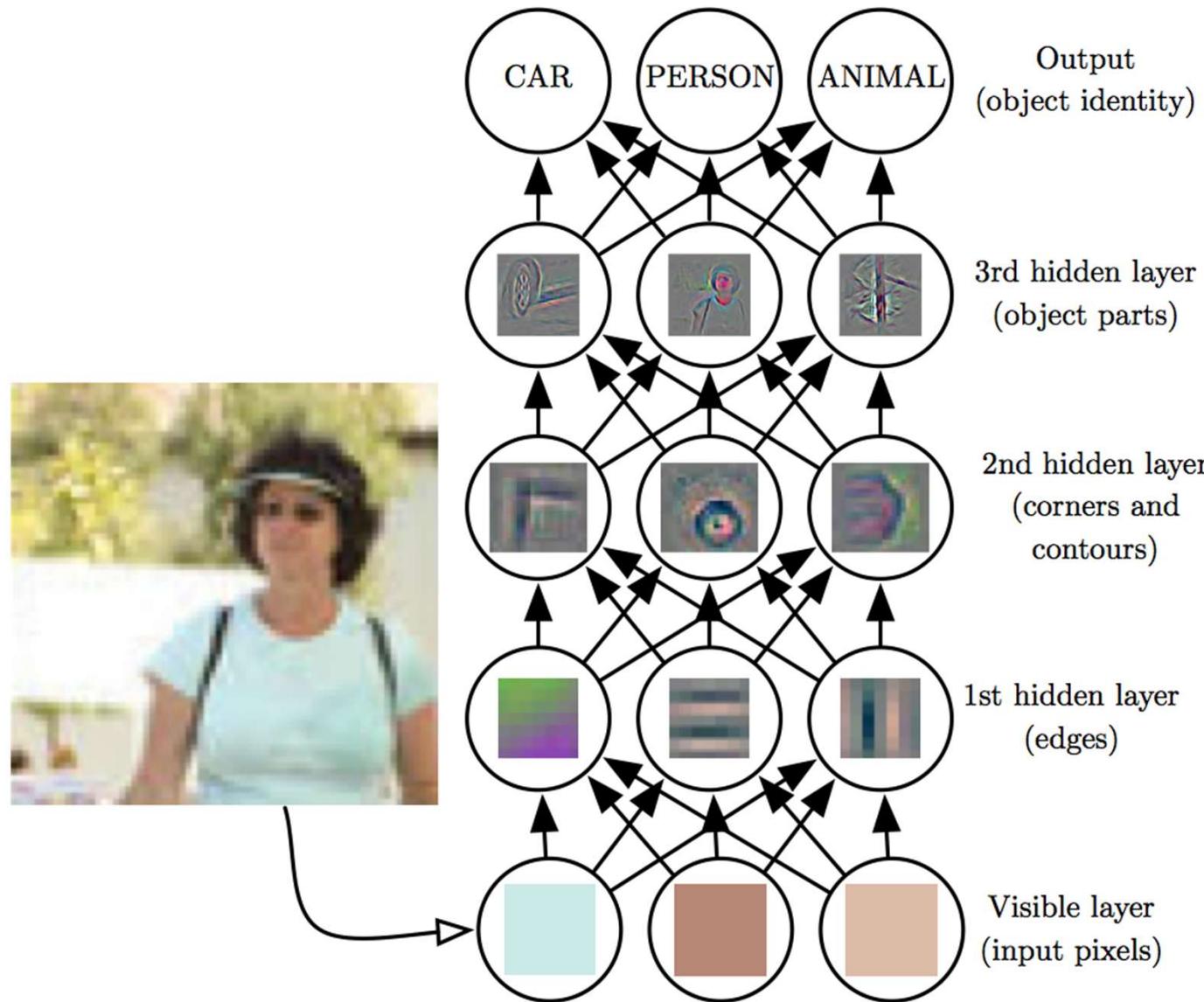
What about learning the representation instead? Representation learning comes to the rescue! Autoencoders learn to encode and decode data by preserving information but adding nice properties

# What is Deep Learning?

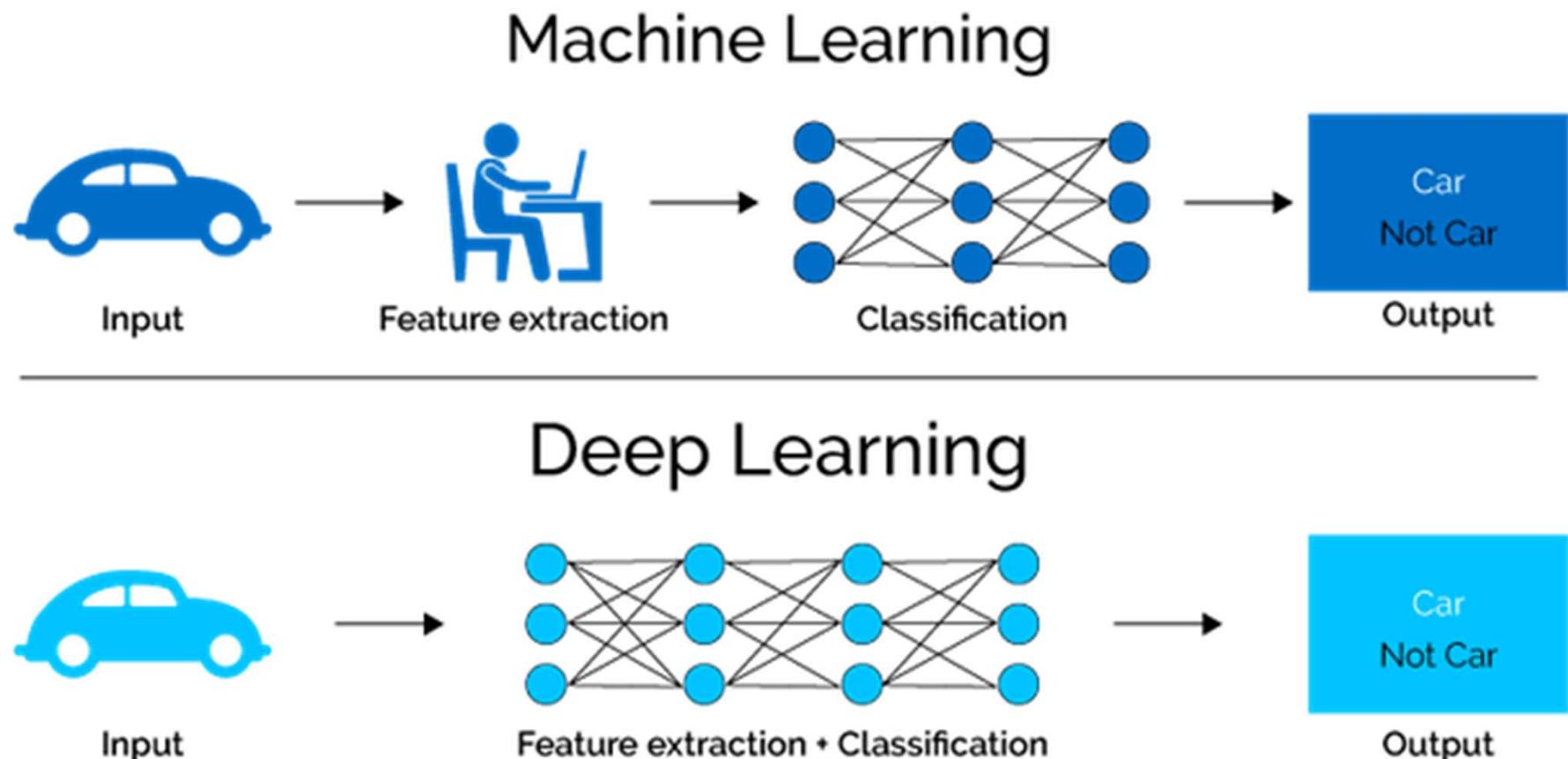


But sometimes the features are too complex to be automatically extracted from raw data in this way. They need to be represented in terms of hierarchy of more and more complex concepts: from shallow to deep learners!

# Complexity through composition



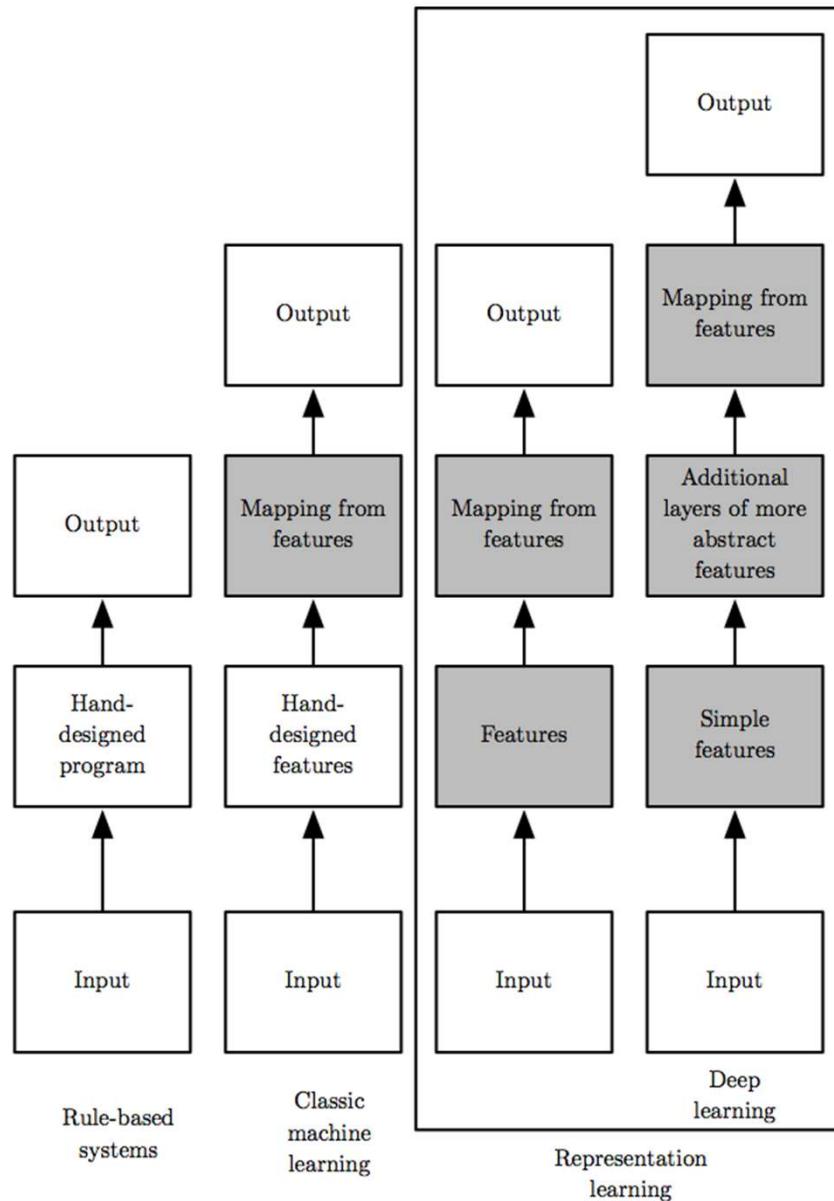
# Machine Learning vs Deep Learning



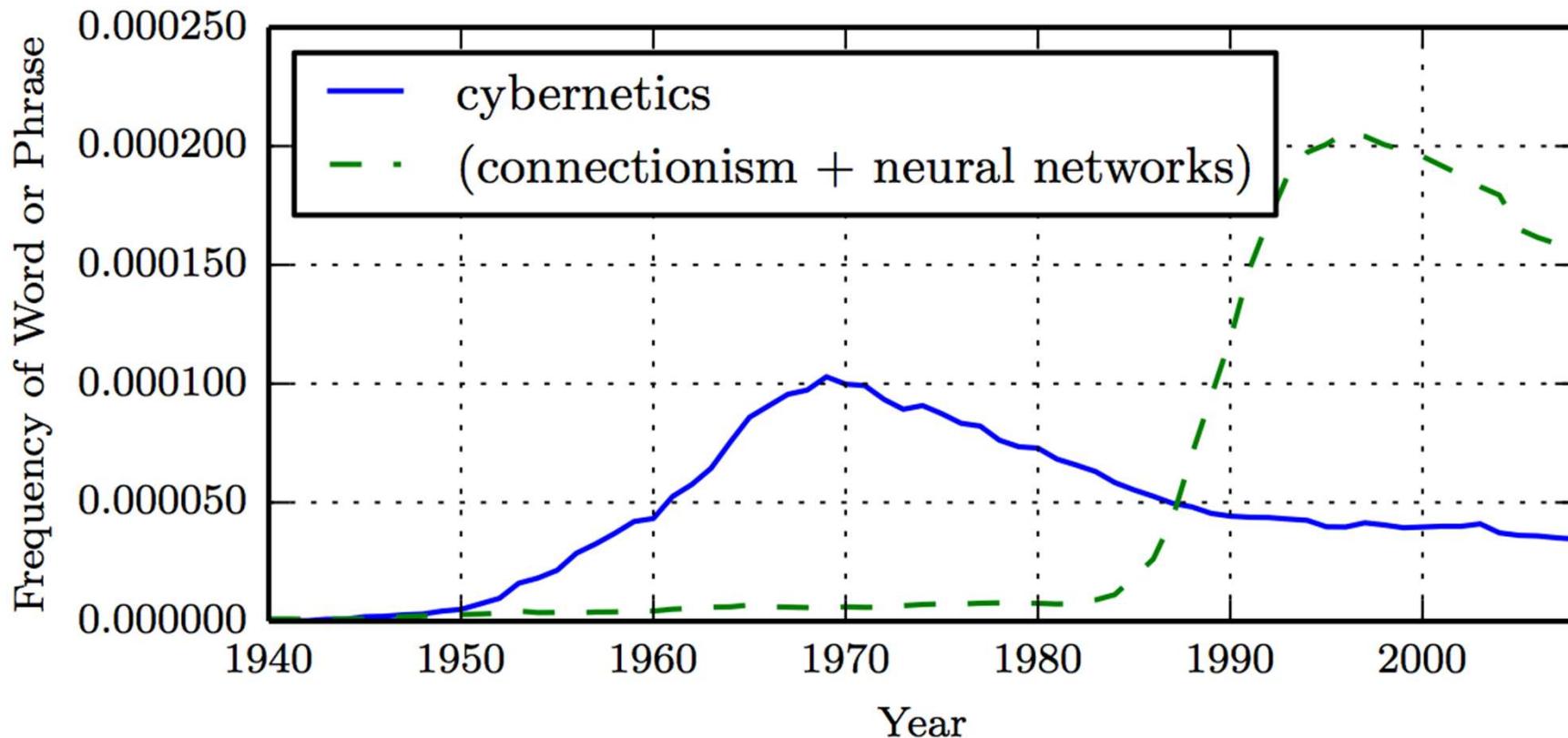
Source <https://codeutsava.in/blog/40>

Another way of seeing this: deep learning learns **not only** the classification but also the representation

# Yet another perspective

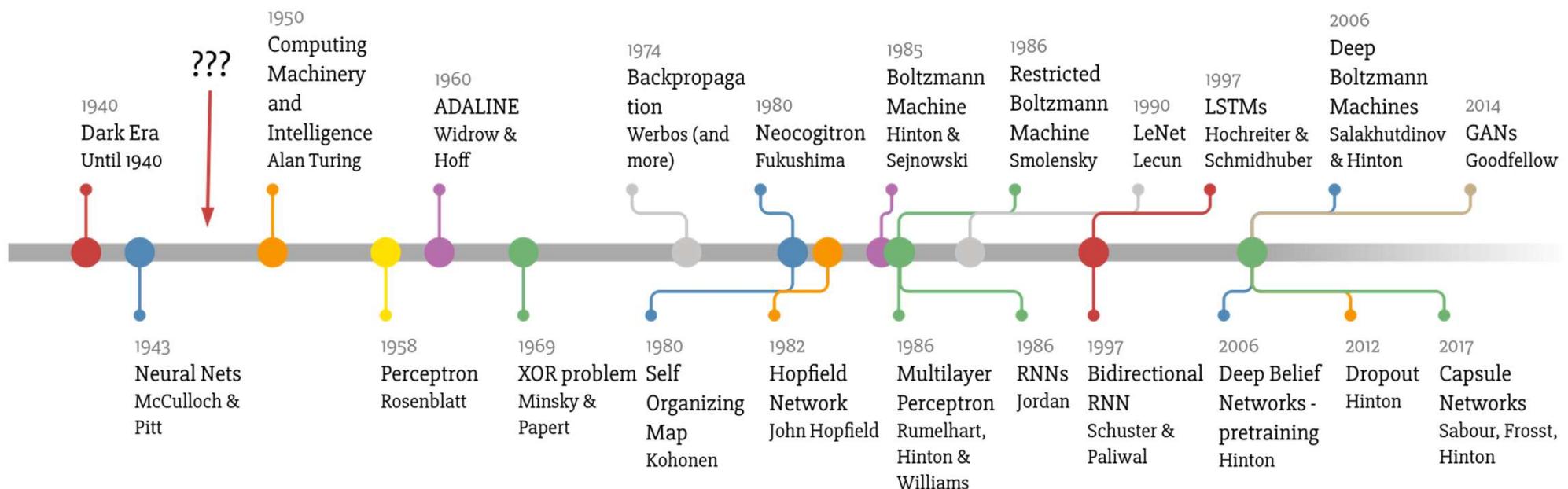


# The many names of deep learning



Deep learning has been around for a loooong time...

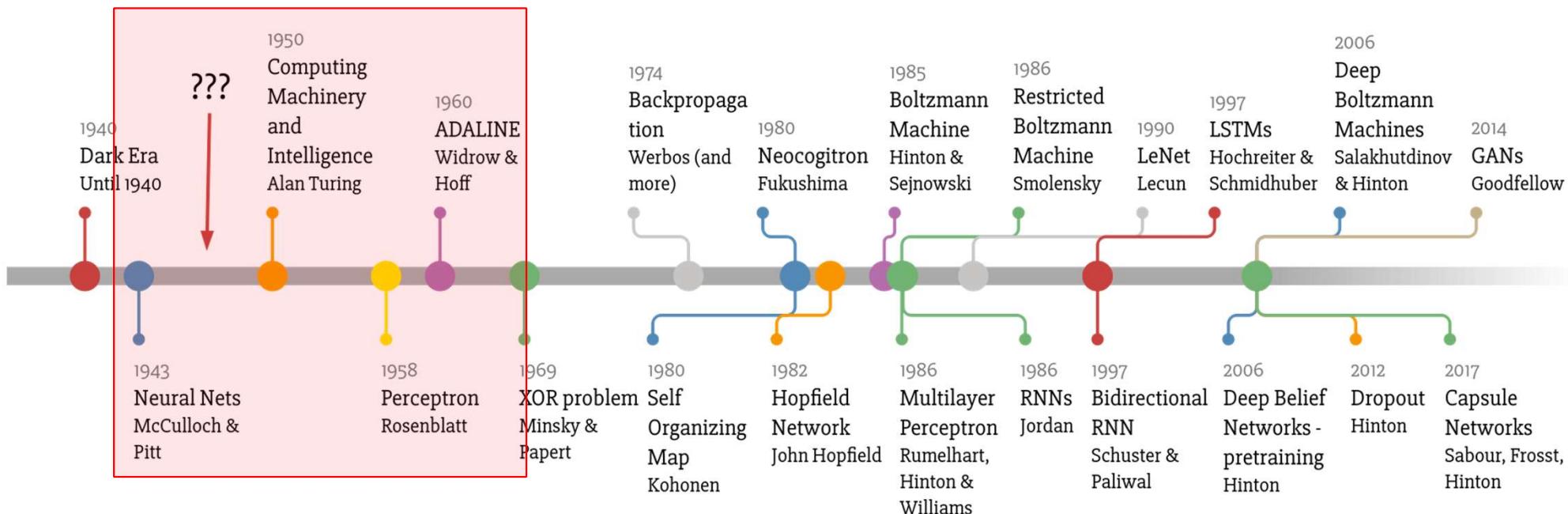
# The many names of deep learning



source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

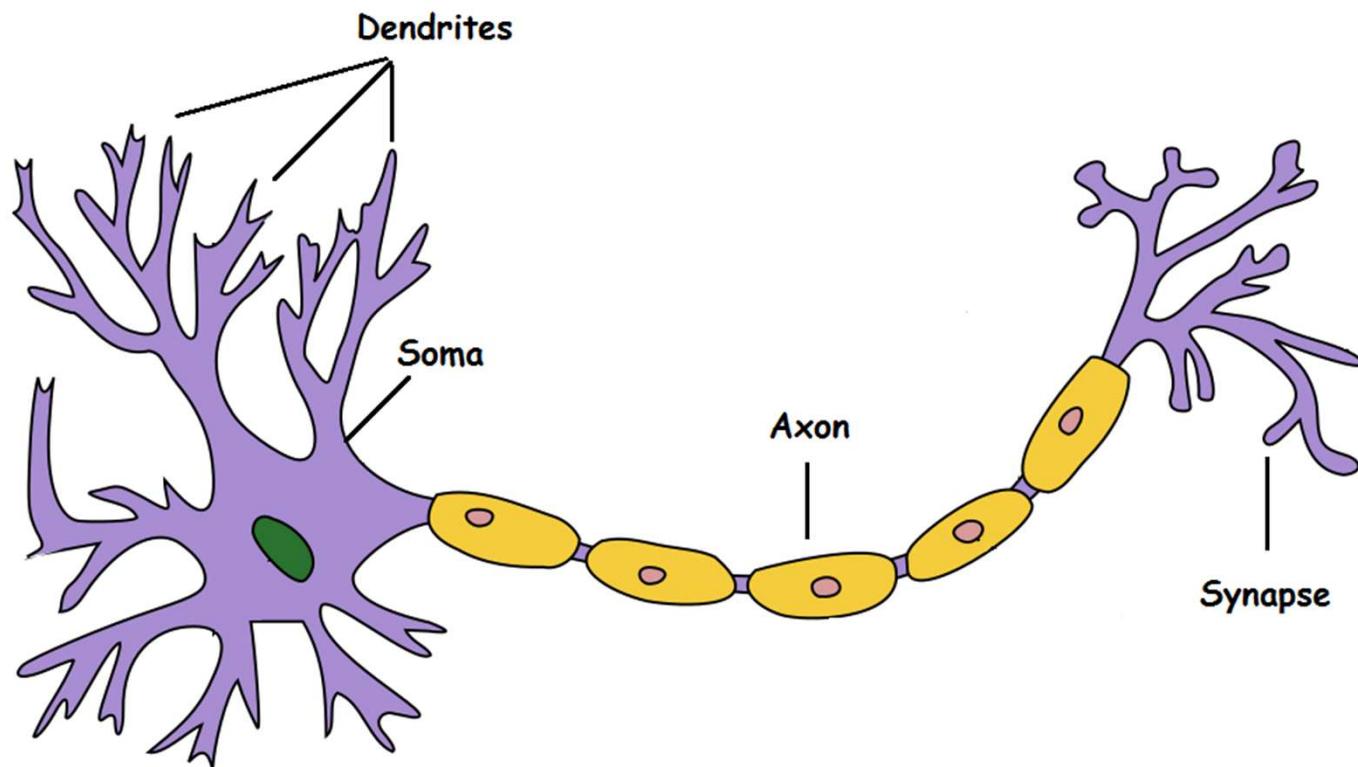
# The many names of deep learning

## First wave: cybernetics



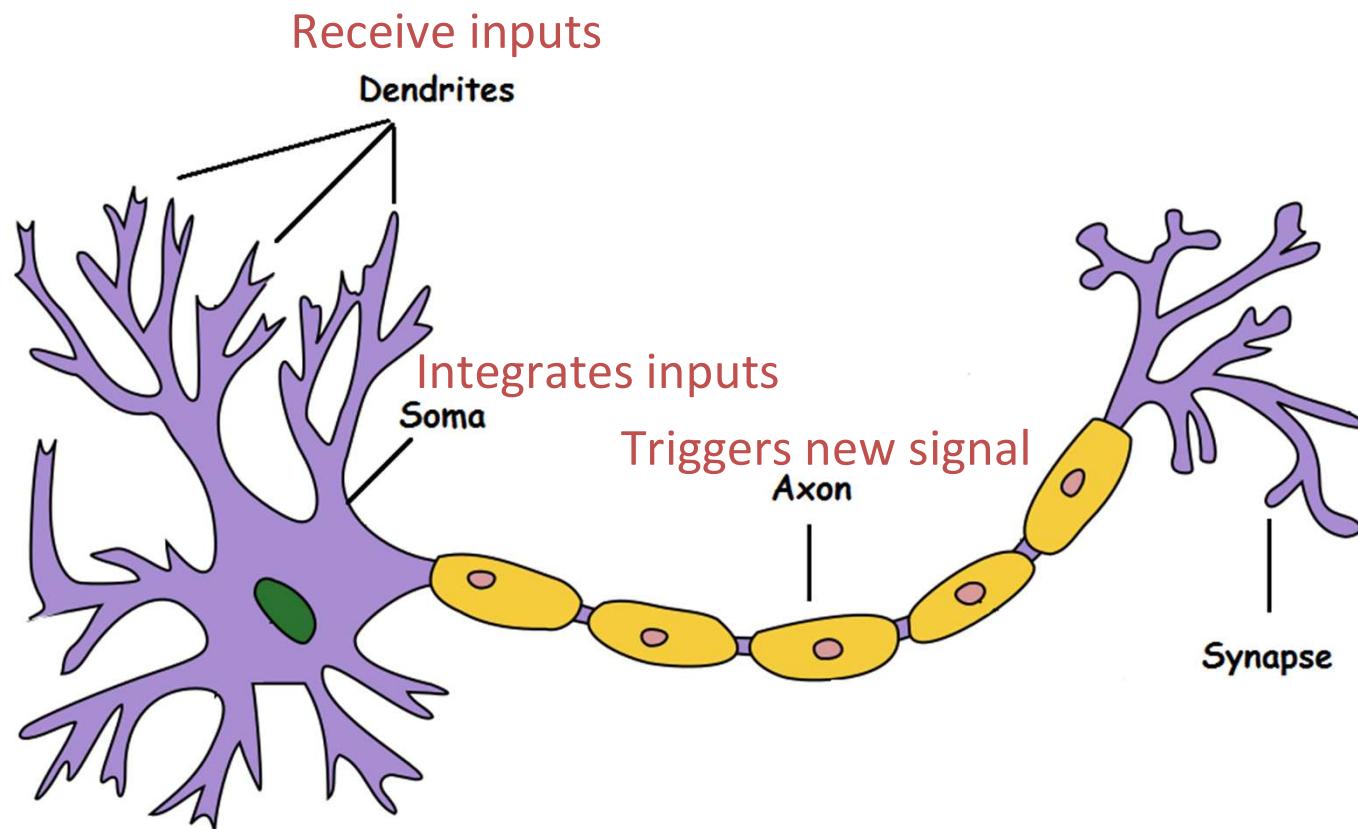
source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

# Biological neurons



By Quasar Jarosz at English Wikipedia, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=7616130>

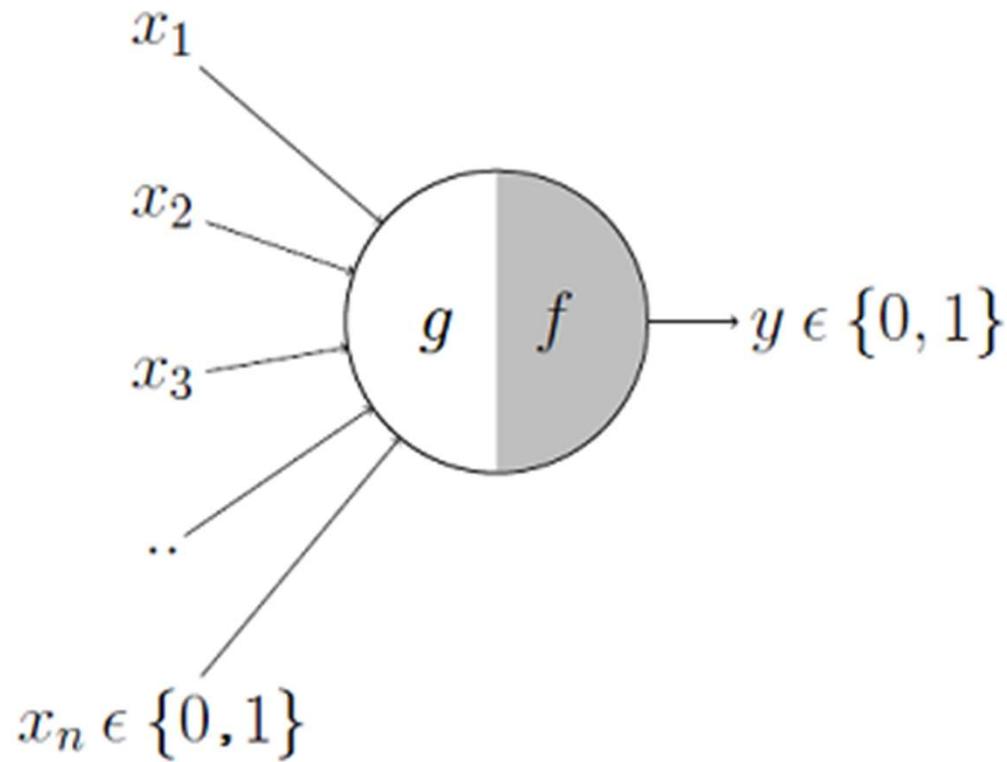
# Biological neurons



By Quasar Jarosz at English Wikipedia, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=7616130>

# Artificial neurons

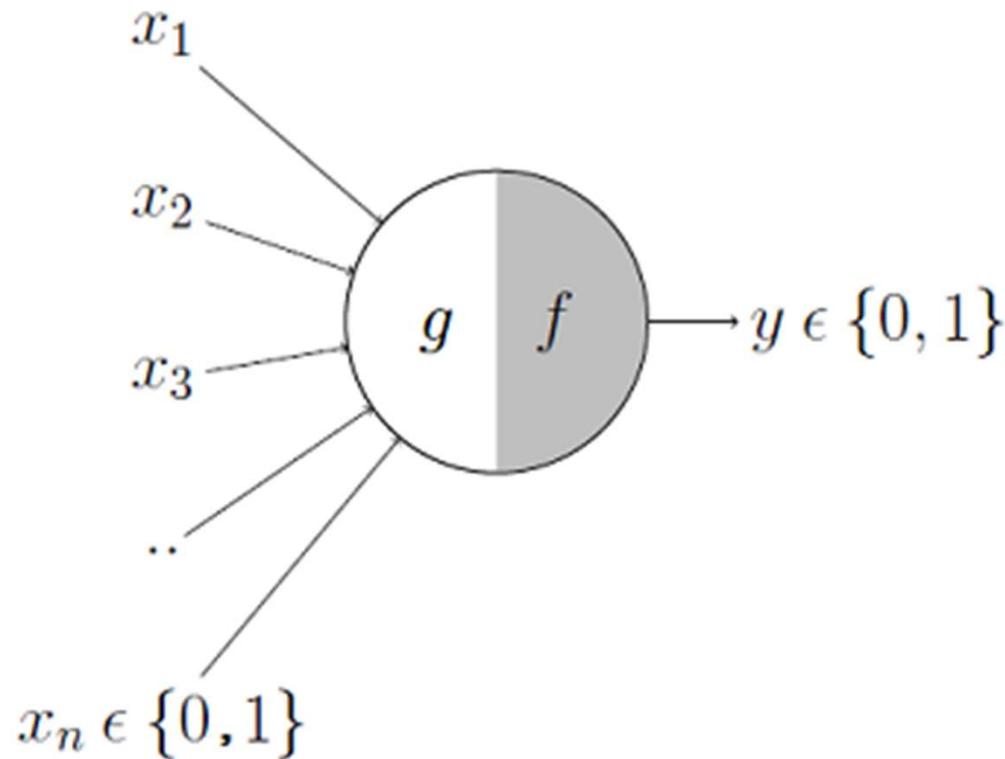
McCulloch-Pitts neuron



source: <https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1>

# Artificial neurons

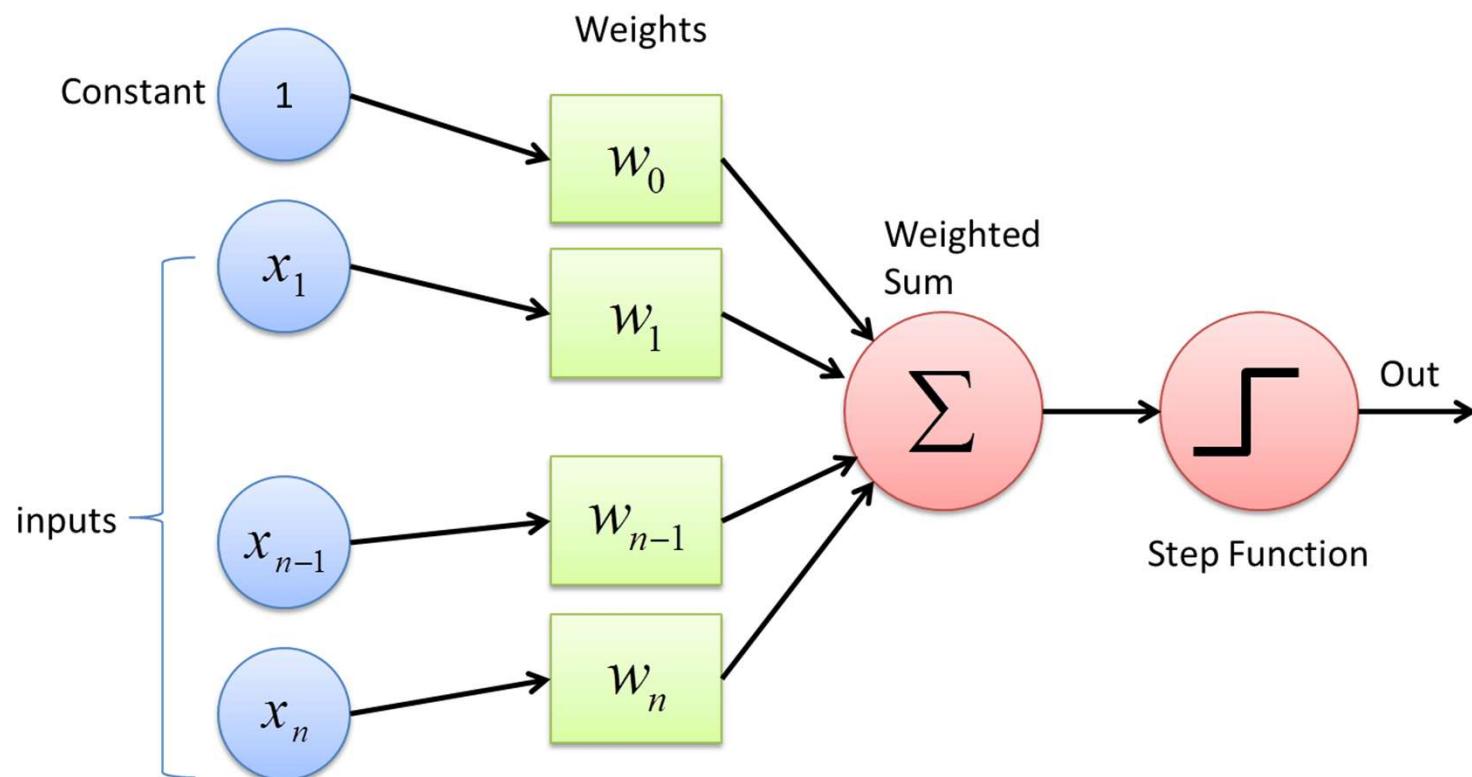
Can implement AND, OR, NOT, etc., but



source: <https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1>

# Artificial neurons

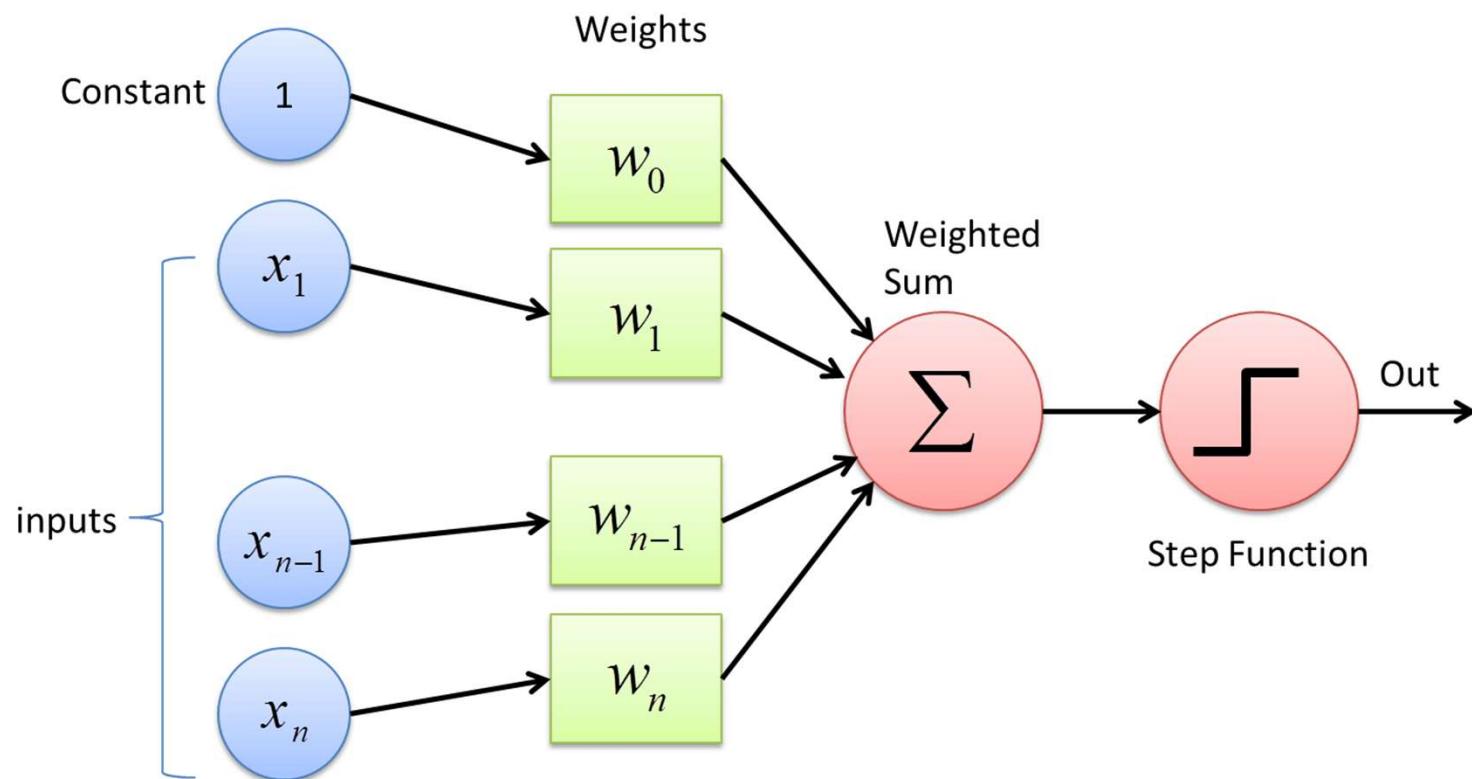
Perceptron by Rosenblatt - same as MP neuron but weighted sum



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

# Artificial neurons

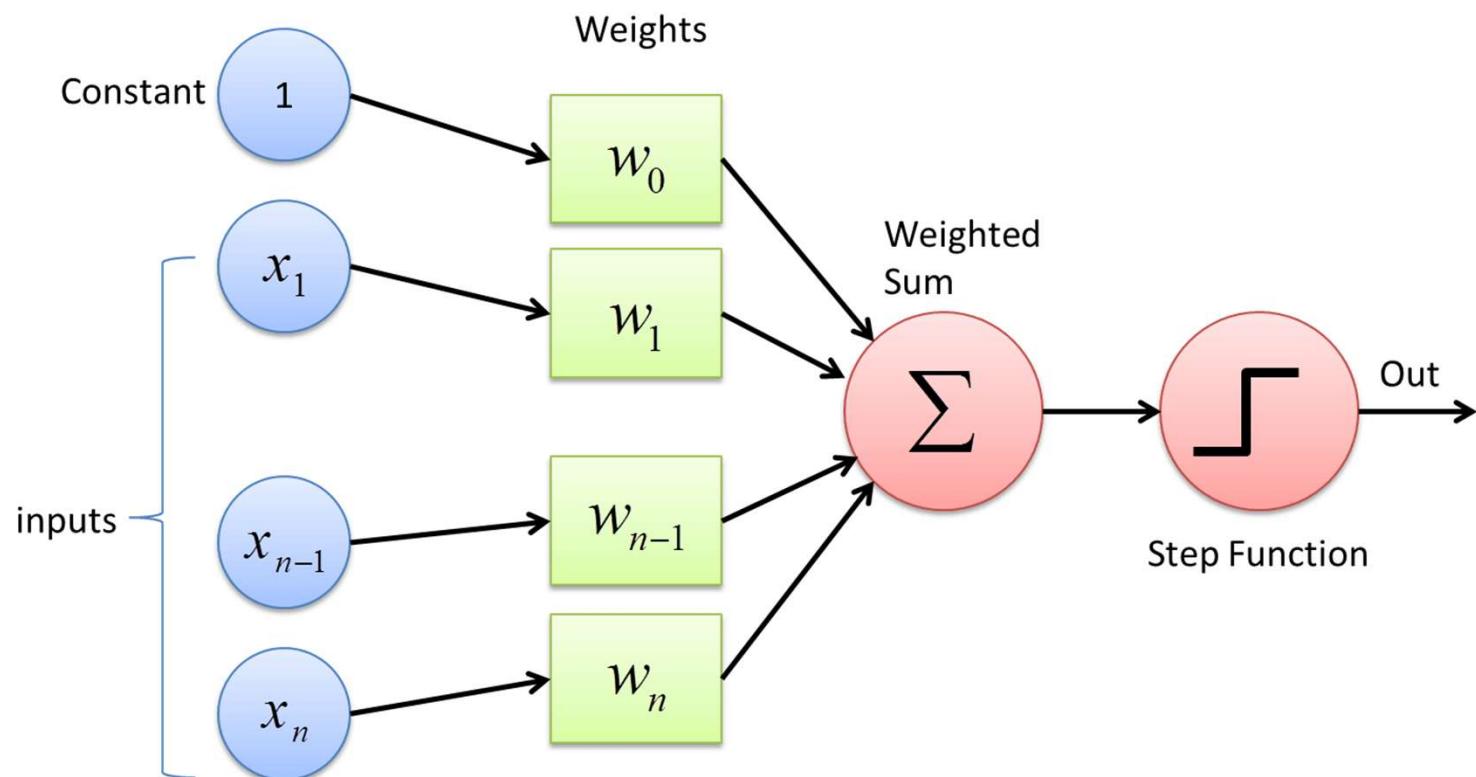
Main innovation is the learning algorithm



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

# Artificial neurons

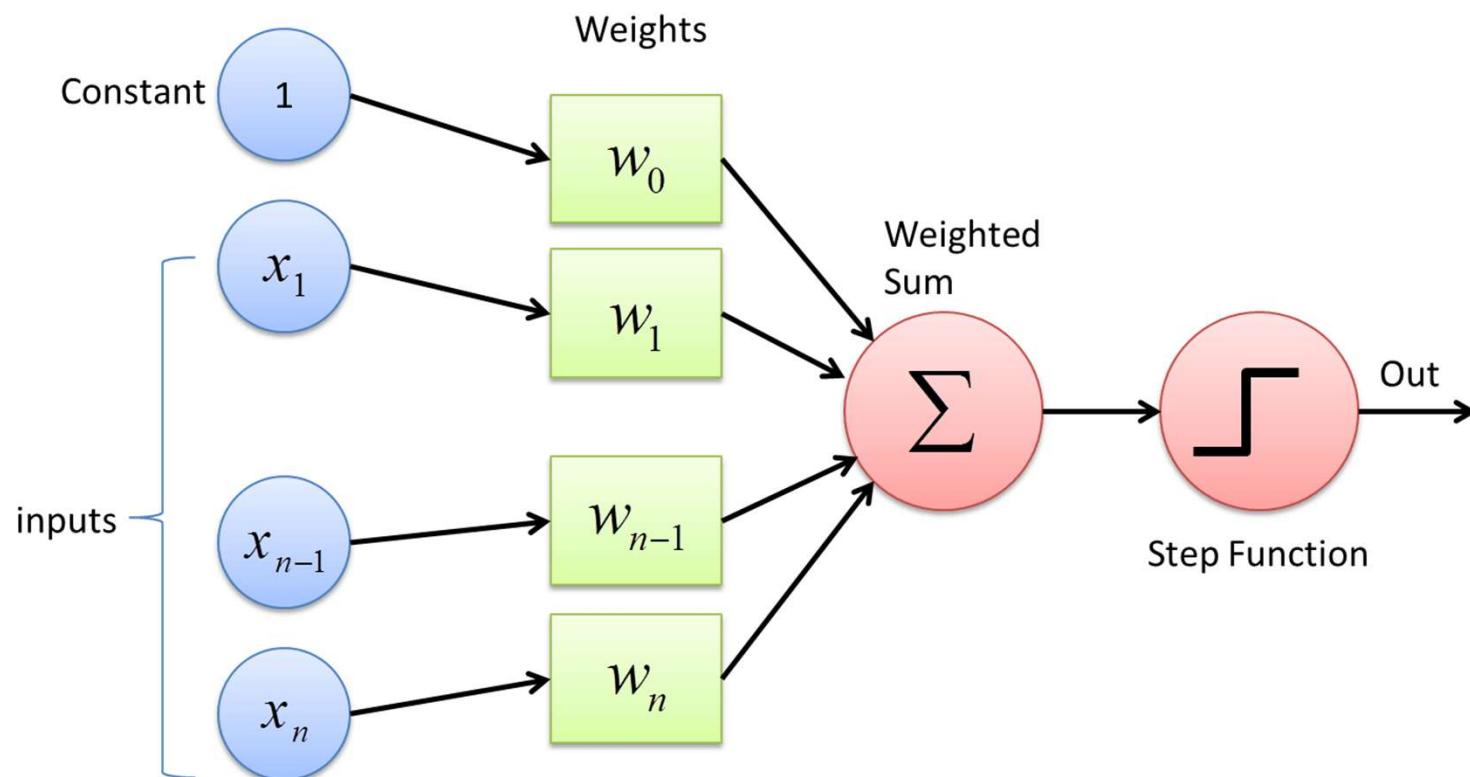
1. Start with random weights



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

# Artificial neurons

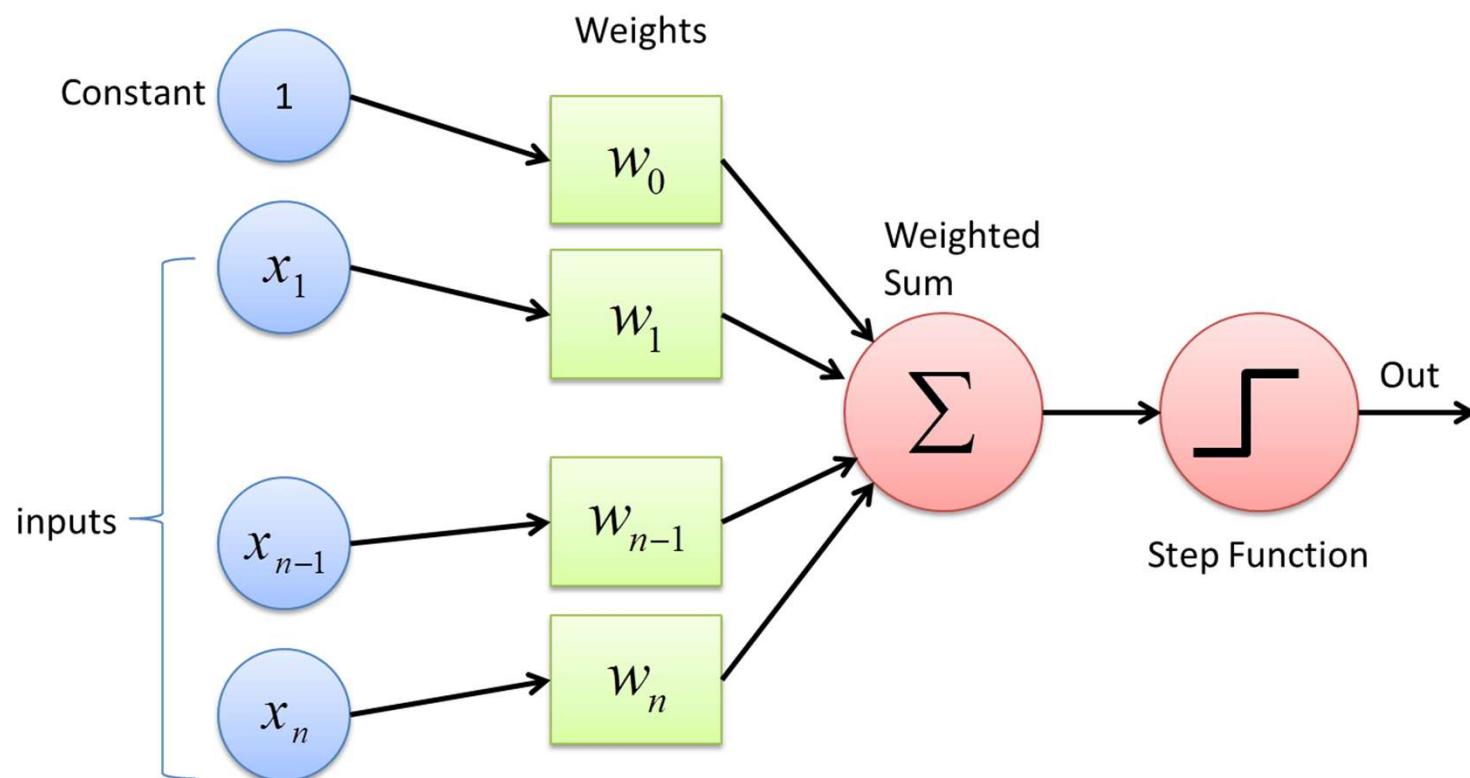
2. Take one sample  $\mathbf{x}_j = [x_1, \dots, x_n]$  and predict  $y_j$



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

# Artificial neurons

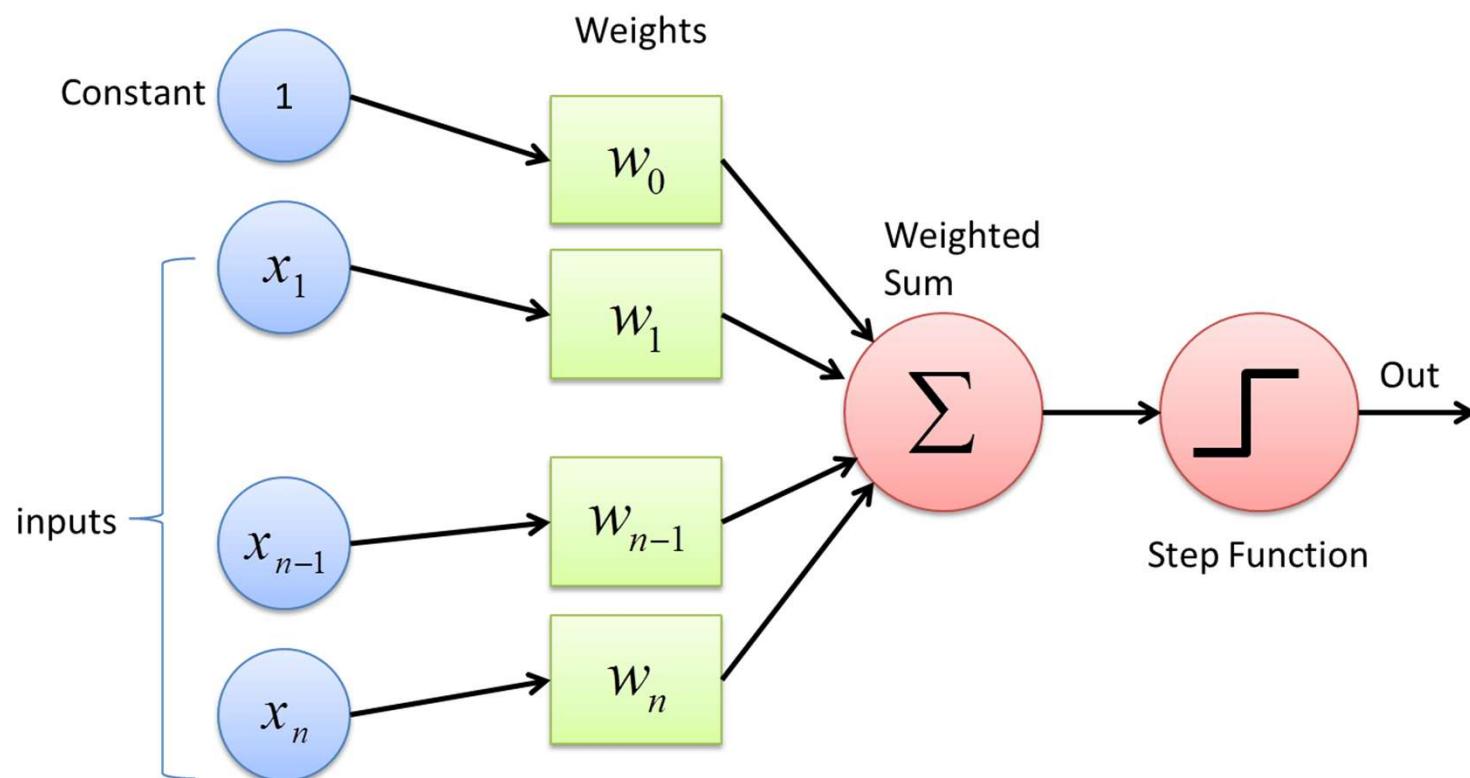
3. If  $y_j = 0$  but  $\hat{y}_j = 1$  then decrease weights



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

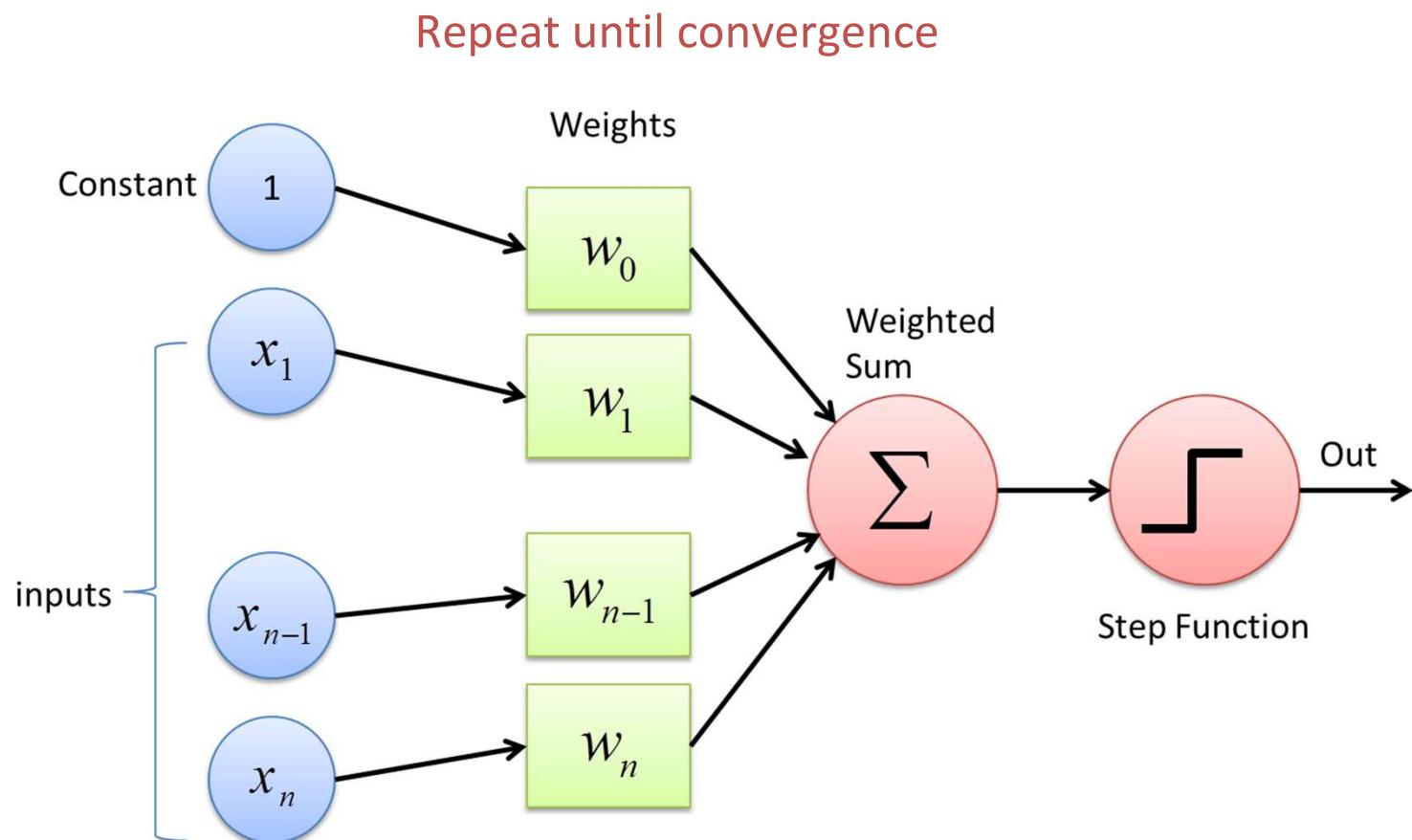
# Artificial neurons

3. If  $y_j = 1$  but  $\hat{y}_j = 0$  then increase weights



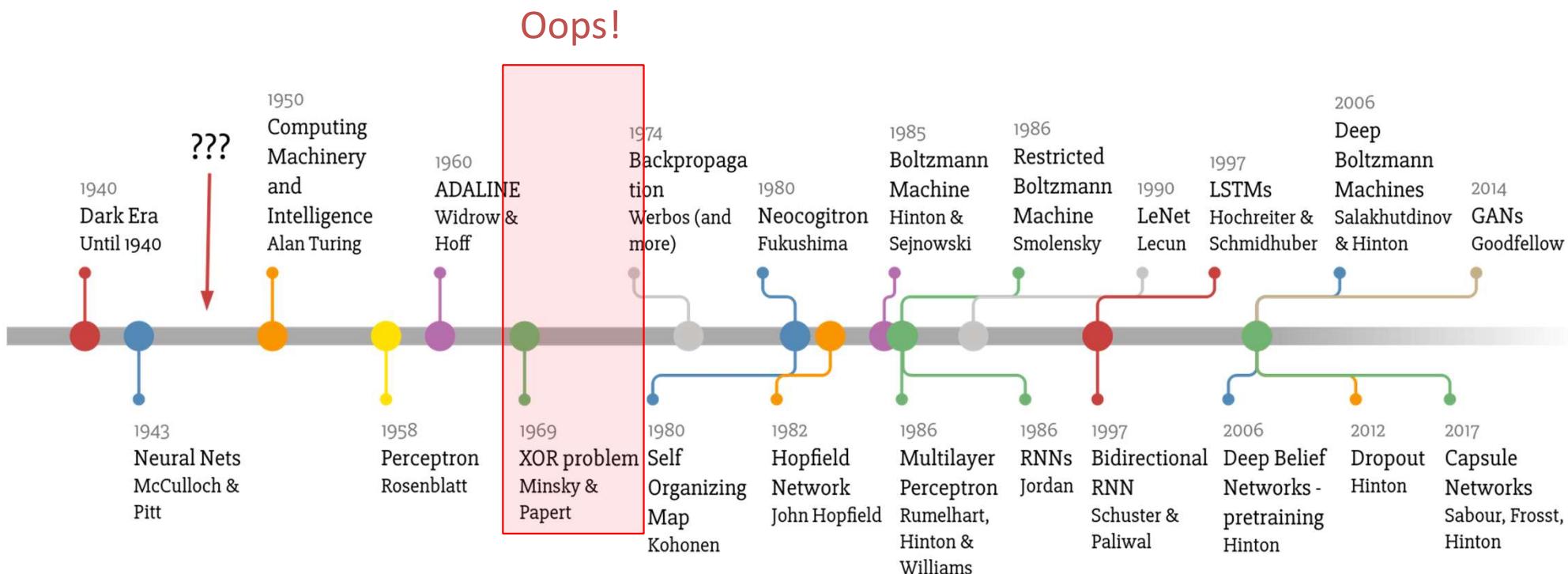
source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

# Artificial neurons



source: <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>

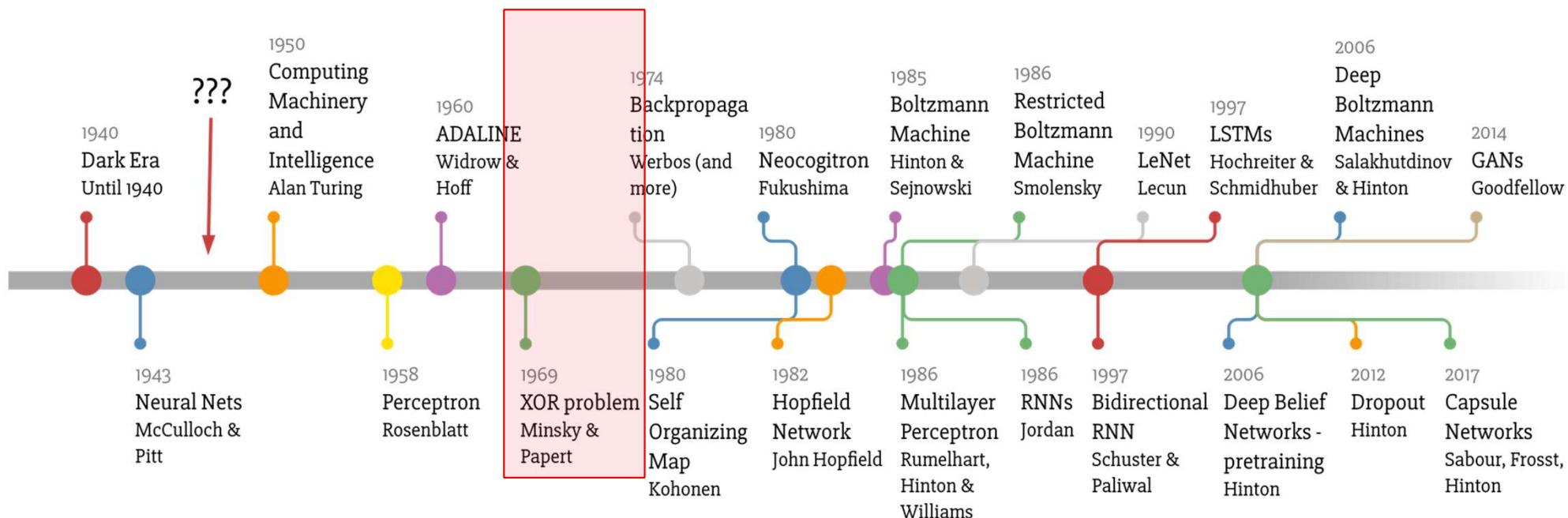
# The many names of deep learning



source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

# The many names of deep learning

MP neurons can't solve simple XOR problems! (non linearly separable)



source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

# The many names of deep learning

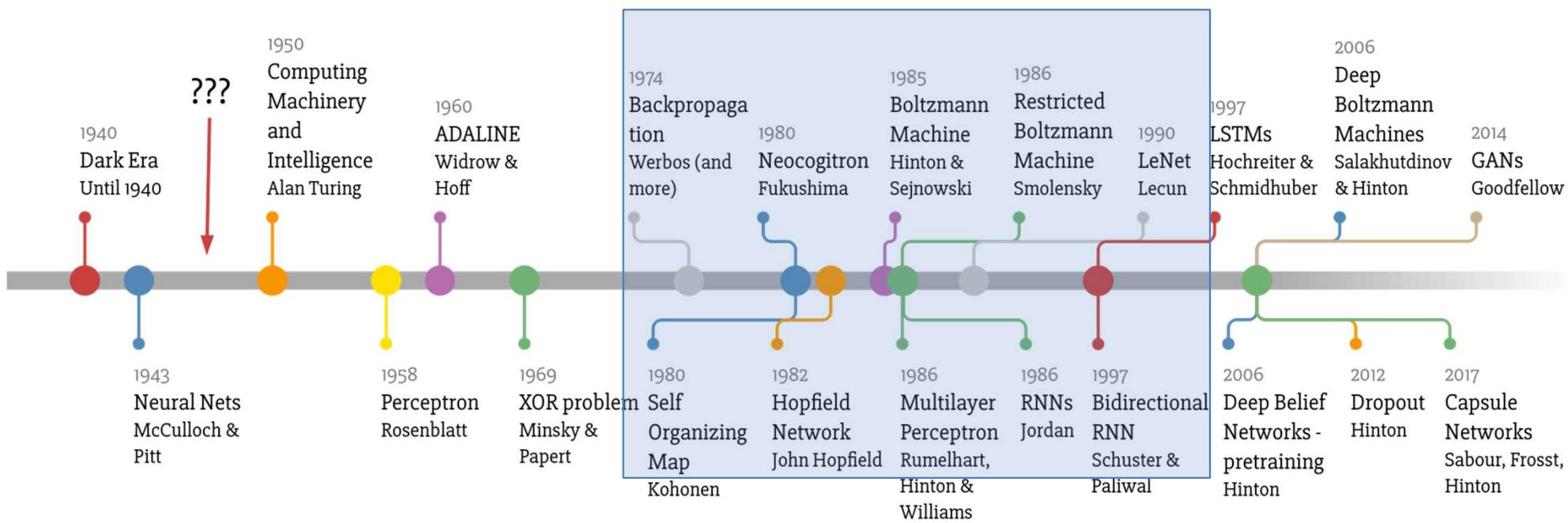




Brace yourself....

# The many names of deep learning

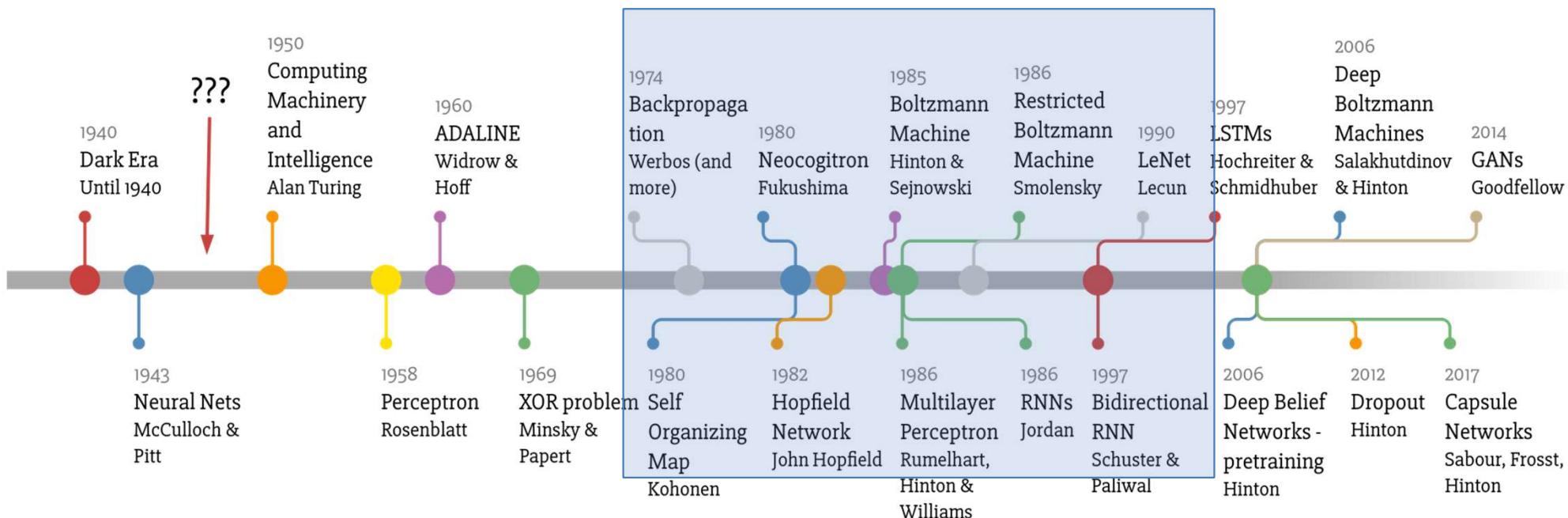
Second wave: back propagation, MLP, lots of cool ideas...



source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

# The many names of deep learning

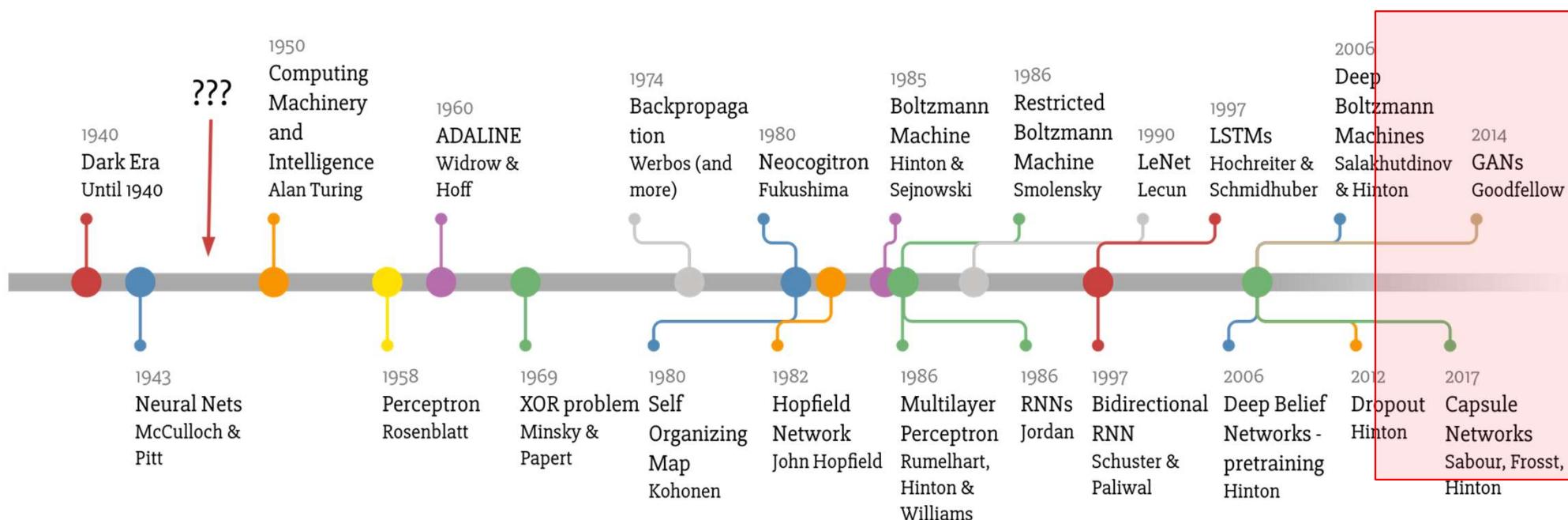
...but too much hype. Until...



source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

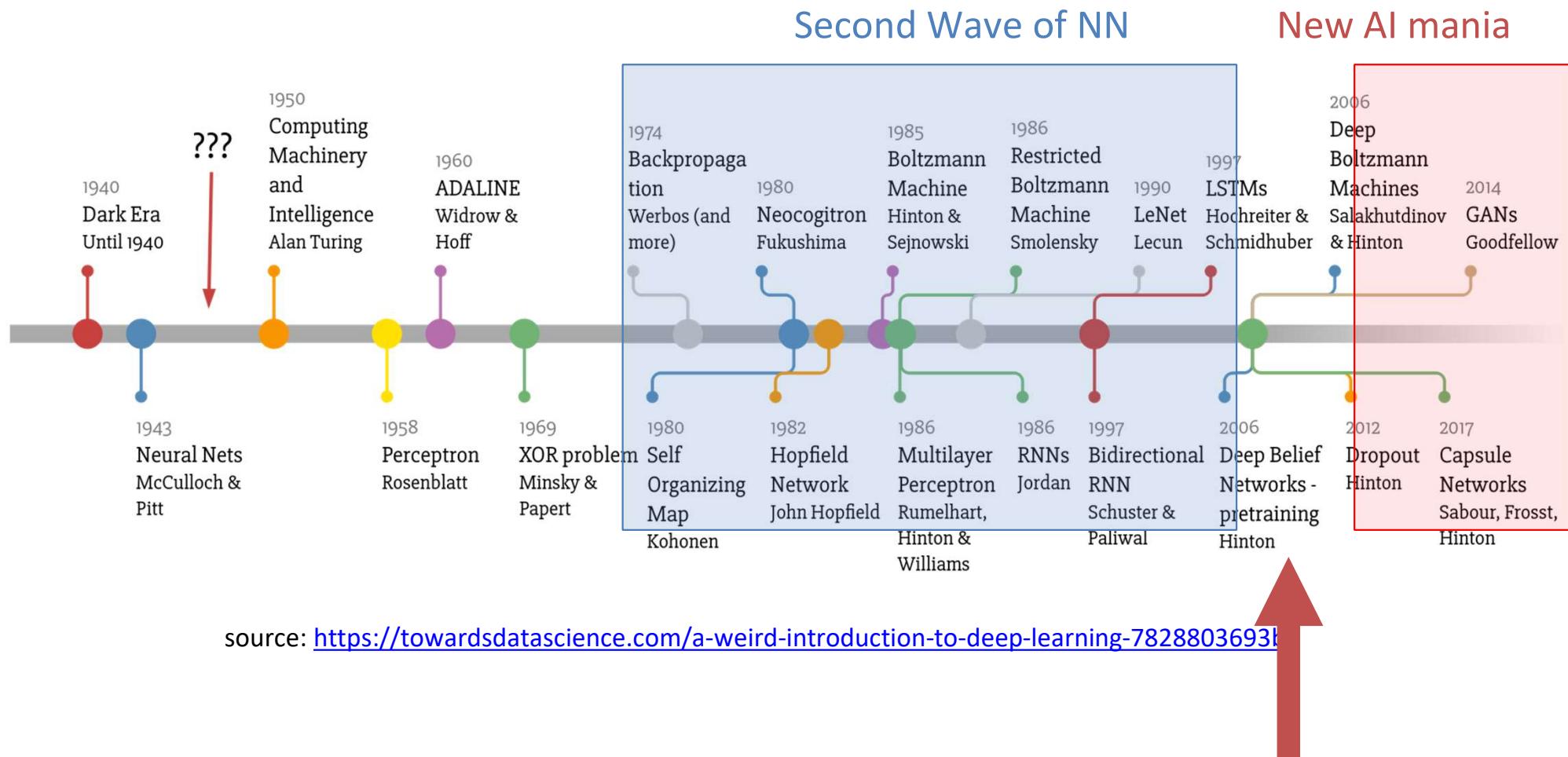
# The many names of deep learning

...NN blooming again



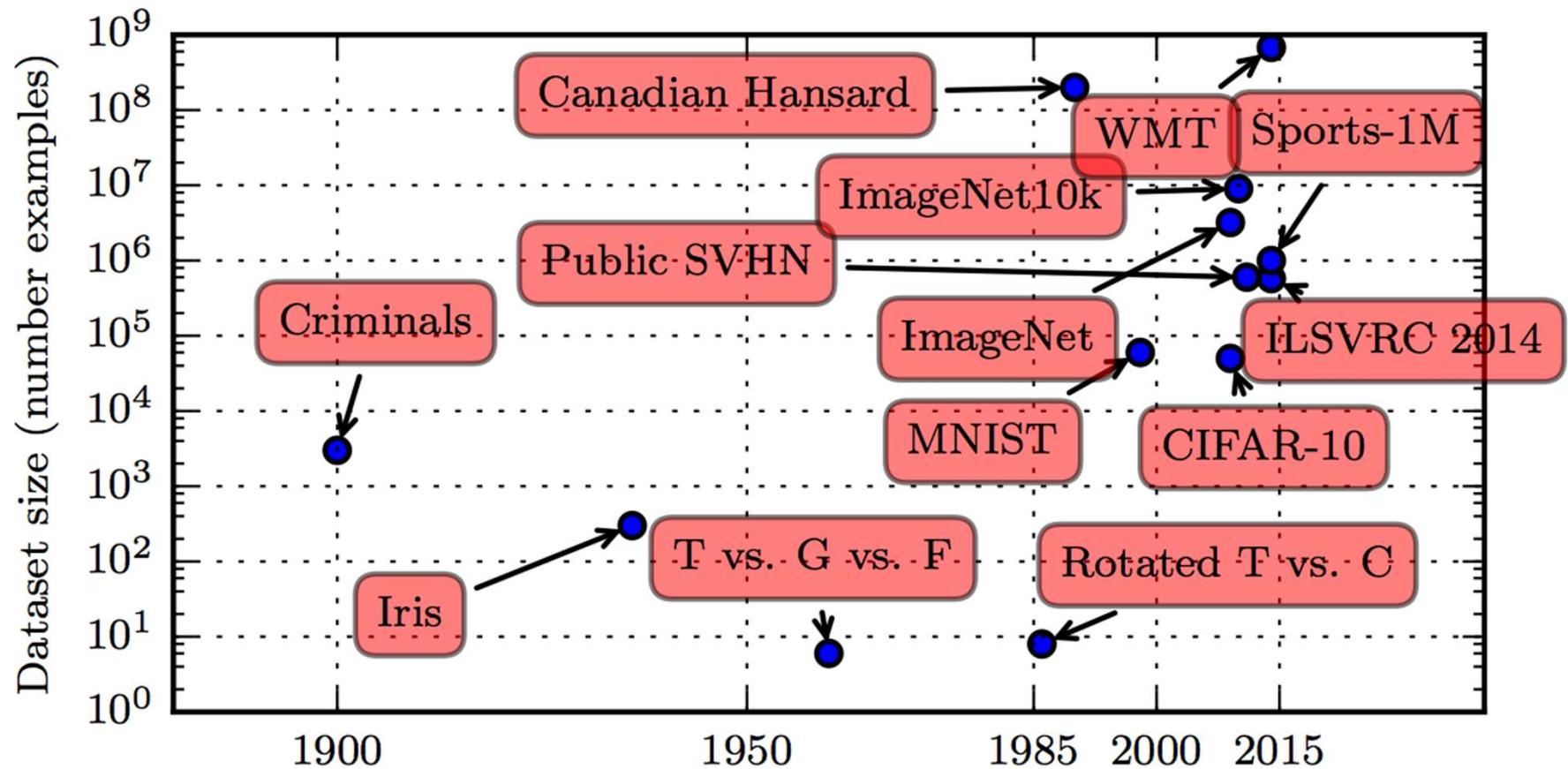
source: <https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0>

# The many names of deep learning



What happened here??

# Hungry learners



Deep learning is Big Data Hungry! More data, better performing models

# Hungry learners

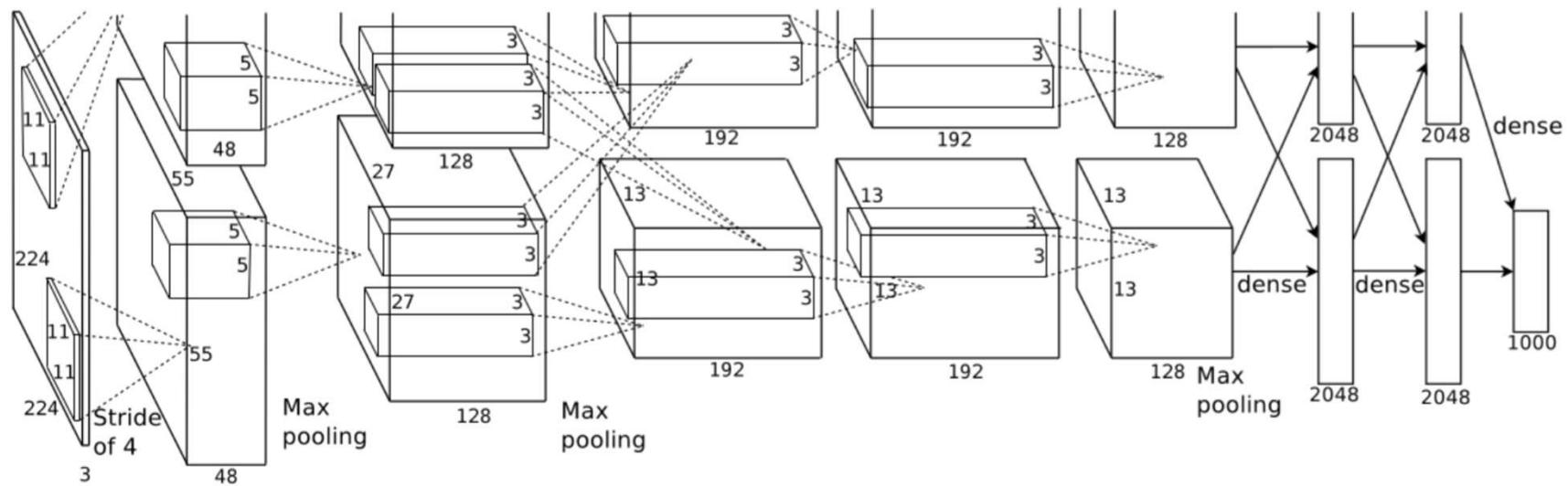
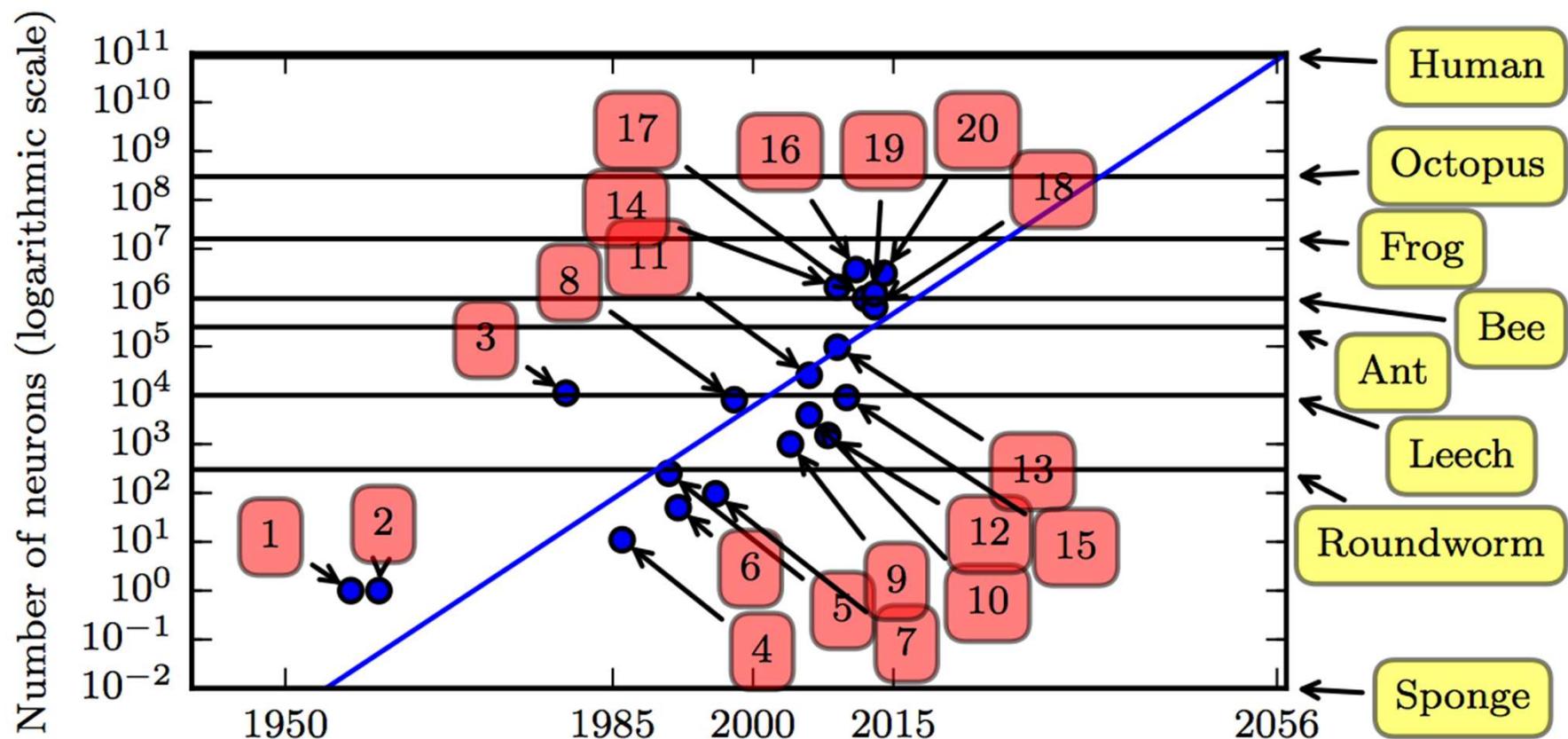


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

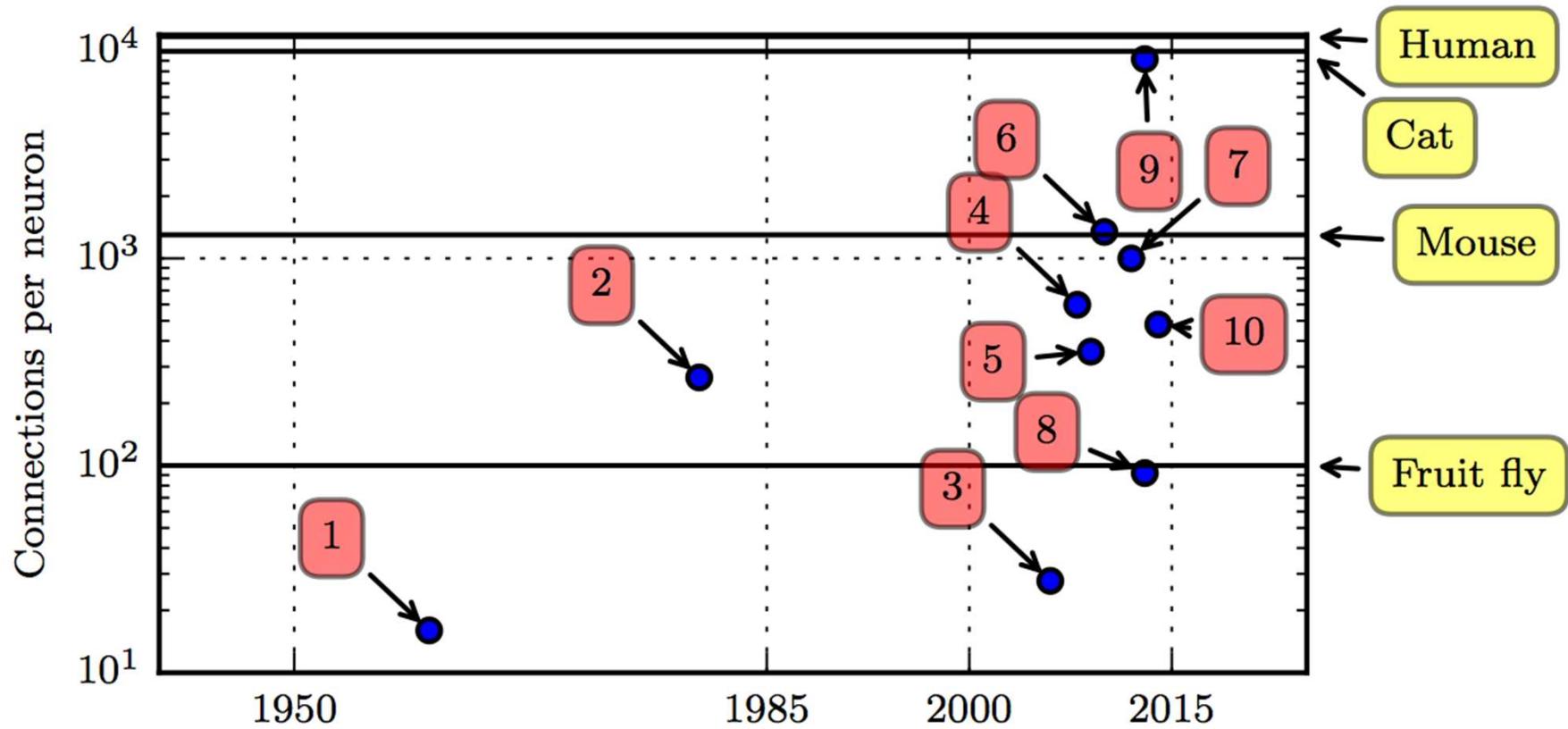
Deep learning is Resources Hungry!  
Advent of general purpose GPUs made the difference

# Hungry learners



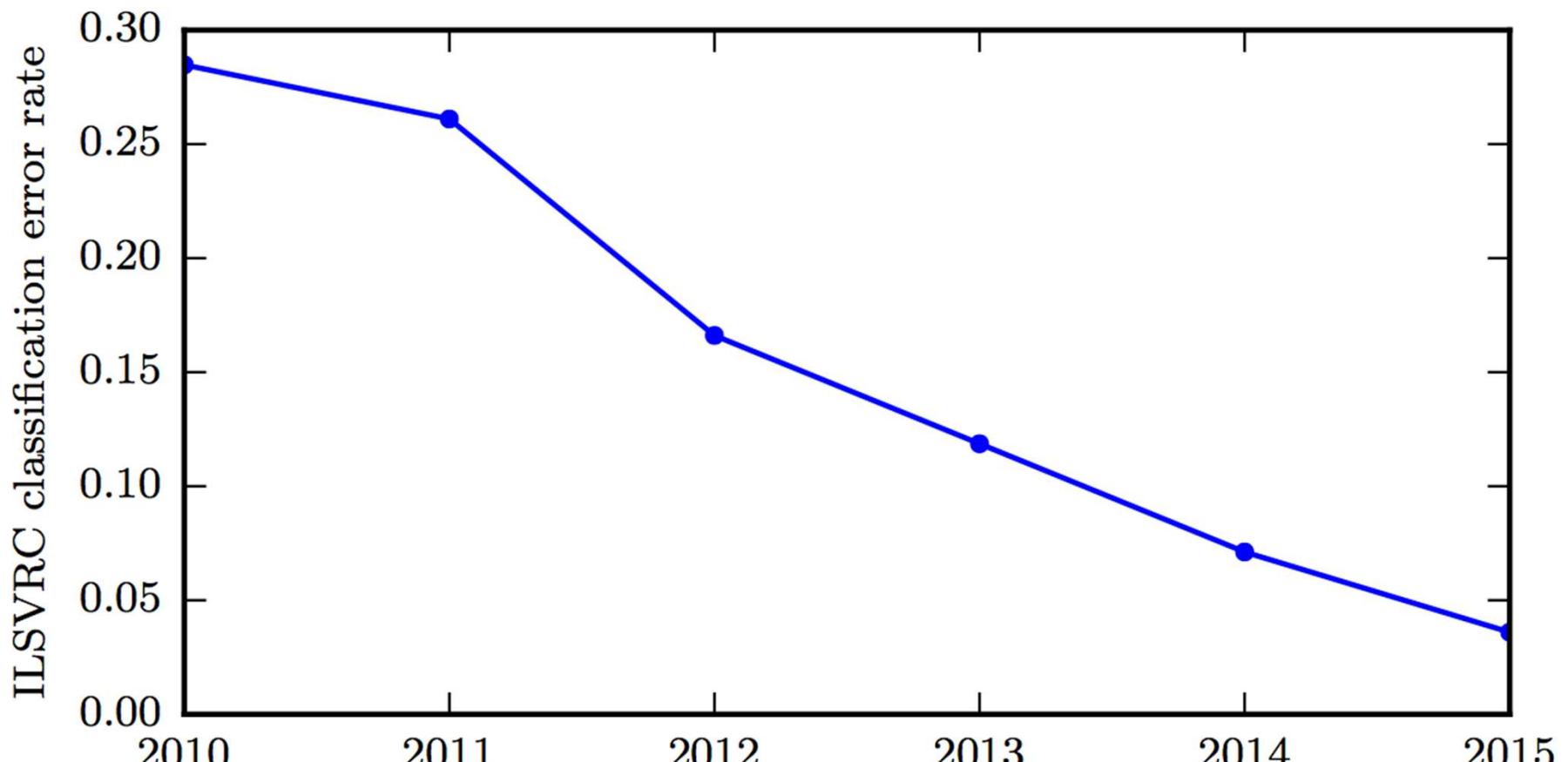
More computational resources means more neurons and thus better models

# Hungry learners



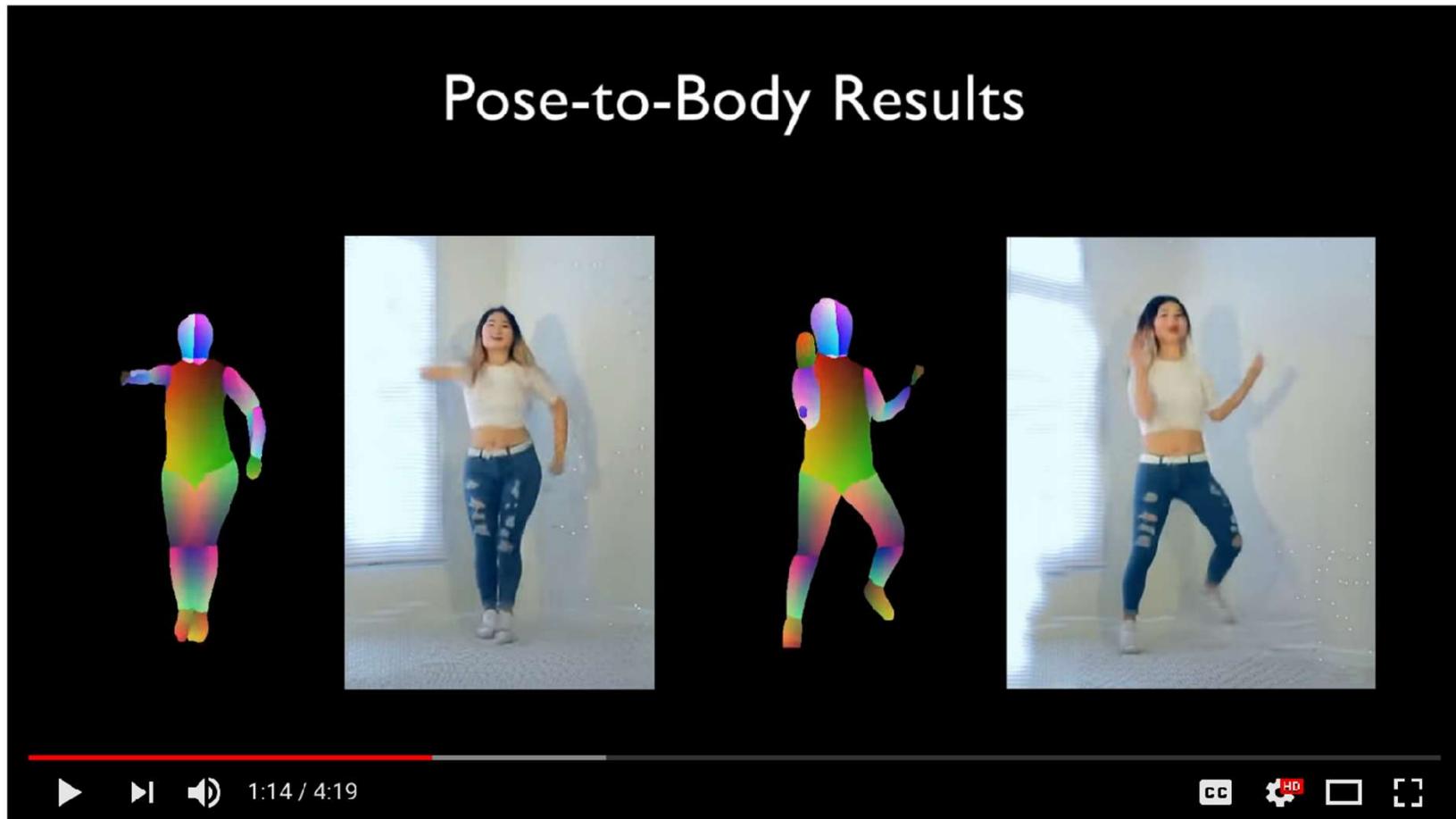
But note: biological neural networks are not very dense, compared to deep networks

# The party crasher



Deep learning achieved astounding performance on many tasks

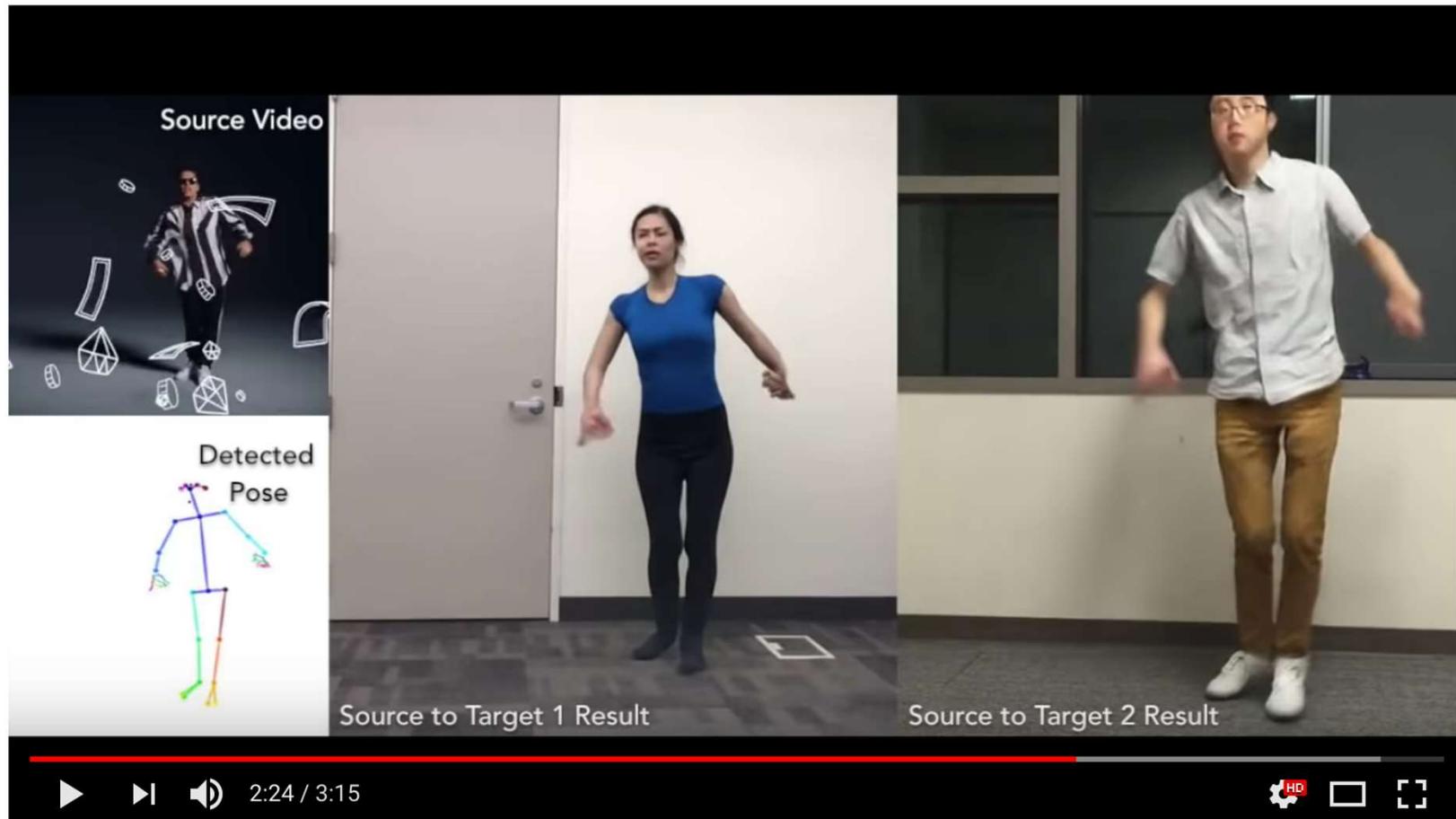
# Some cool applications of DL



(Wang et al, 2018)

<https://www.bilibili.com/video/BV1vE411B7LS?from=search&seid=14546211848937884840>

# Some cool applications of DL



(Chan et al 2018)

# Some cool applications of DL

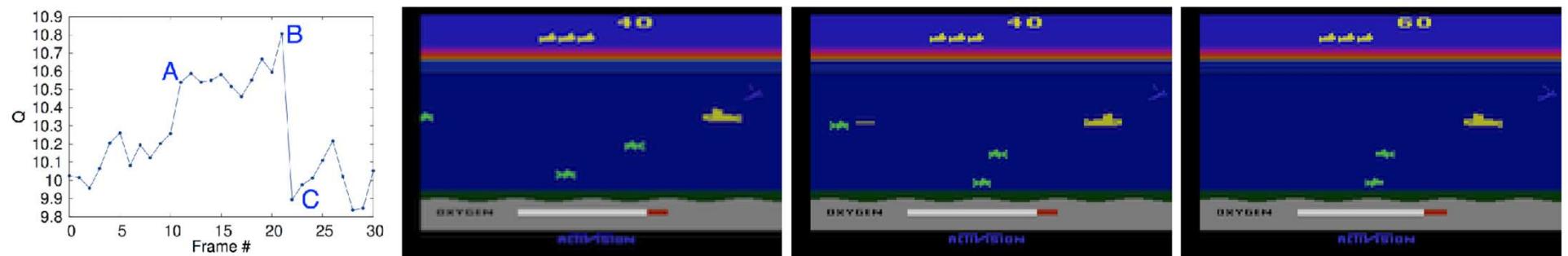


Figure 3: The leftmost plot shows the predicted value function for a 30 frame segment of the game Seaquest. The three screenshots correspond to the frames labeled by A, B, and C respectively.

([Mnih et al 2013](#))

[Demo 1](#) [Demo 2](#)

# Some cool applications of DL



<https://www.youtube.com/watch?v=S1OwOd-war8>

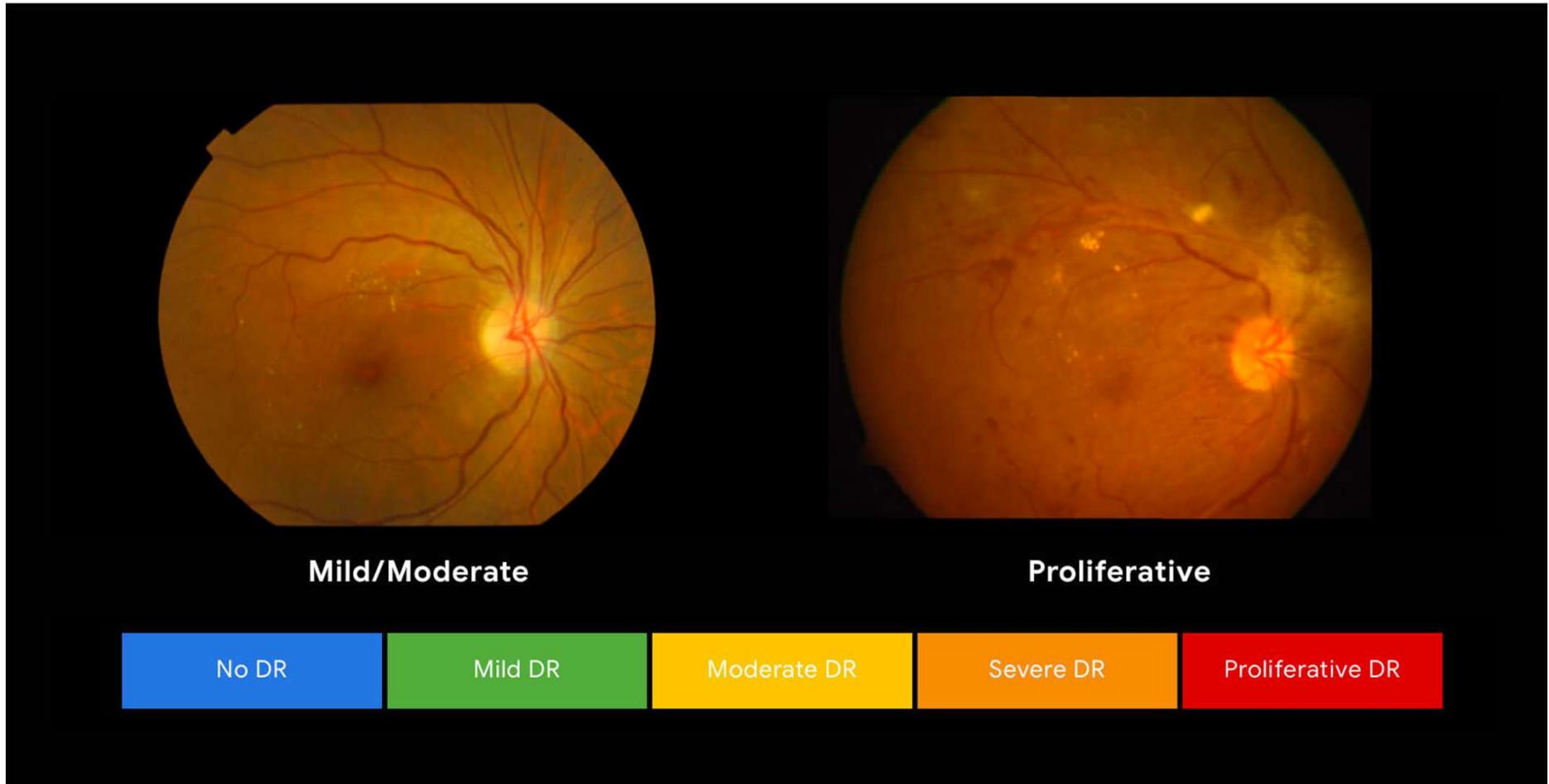
<https://www.youtube.com/watch?v=S1OwOd-war8>

MORE VIDEOS



([WayMo](#))

# Some cool applications of DL



(Google Brain)

# Some cool applications of DL



(Video classification)

# Some cool applications of DL



("deep" art)

# Are you trying to impress me?

- Well, these are all **very complex problems**:
  - Lots of different game states in Game AI, e.g.,  $((10)^{10})^{48}$  possible GO games!



# Are you trying to impress me?

- Well, these are all **very complex problems**:
  - Natural language processing is very high dimensional and vague



# Are you trying to impress me?

- Well, these are all **very complex problems**:
  - Computer vision is really challenging:
    - For 256x256 resolution,  $2^{524288}$  possible images!! That's a lot, btw
    - Massive visual and semantic variation



# OpenAI -- SORA

- Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

# SORA

- We have watched the video at the beginning of this class.

# How Can I learn deep learning well? By ChatGPT

ATTENDING  
MY  
LECTURES

YJ How can I learn deep learning well

 Learning deep learning can be challenging, but with the right approach, you can develop a solid understanding of this important field. Here are a few steps to help you learn deep learning well:

1. Start with the basics: It's important to have a solid foundation in linear algebra, calculus, and statistics, as these concepts form the backbone of deep learning. If you need to refresh your knowledge, consider taking an online course or reading a book on these topics.
2. Get hands-on experience: Practical experience is essential when learning deep learning. Start by working on small projects and experimenting with different deep learning algorithms and models. As you gain more experience, you can work on more complex projects.
3. Learn from experts: There are many experts in the field of deep learning, and learning from their experiences can be extremely valuable. Read their blog posts, watch their videos and tutorials, and consider enrolling in a course taught by an expert.
4. Join a community: Joining an online community of deep learning enthusiasts can provide you with valuable resources and support as you learn. You can ask questions, get feedback on your work, and connect with others who are also learning.
5. Stay up-to-date: Deep learning is a rapidly evolving field, and it's important to stay up-to-date with the latest developments. Follow key researchers and practitioners on social media, read academic papers, and attend conferences and workshops to stay informed.
6. Practice, practice, practice: Deep learning is a complex field, and the best way to master it is through consistent practice. Work on as many projects as you can, and don't be afraid to make mistakes. The more you practice, the better you will become.

By following these steps, you can develop a strong foundation in deep learning and become proficient in using this technology [to solve real-world problems](#).

 Regenerate response