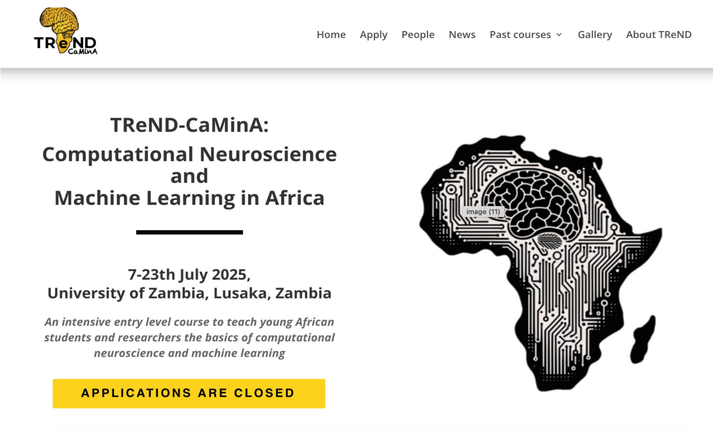
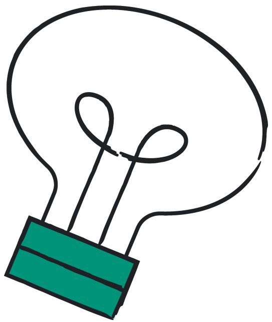


Machine Learning
& Neuroscience

Presentation inspired from Martino Sorbaro's slides at TReND CaMinA



Imane Hamzaoui

جامعة نيويورك ابوظبي
NYU ABU DHABI

esi

Research assistant at NYUAD

I've obtained a Masters and Engineering degrees in Computer Science specialized in Intelligent Systems and Data at Ecole Nationale superieure d'informatique.

I did my final year internship at New York University Abu Dhabi where I'm working on exploring the potential of fruit fly brain inspired neural networks.

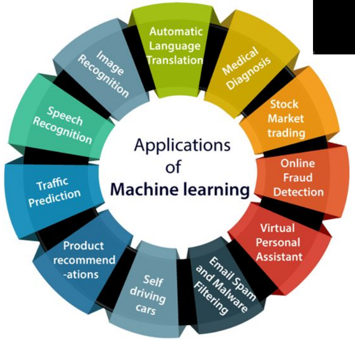
My research interests revolve around Neuromorphic Computing, In-Memory Computing, and Deep Learning in general.

Women Techmakers

1 Machine Learning

Why do we need Machine Learning

image classification



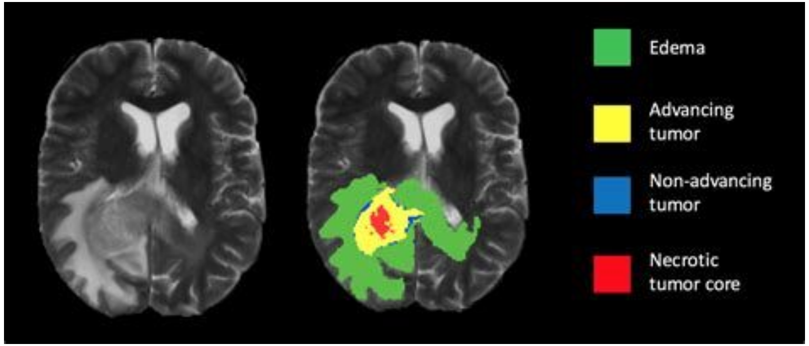
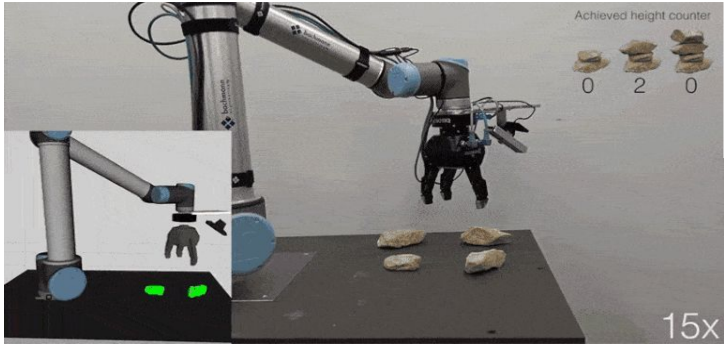
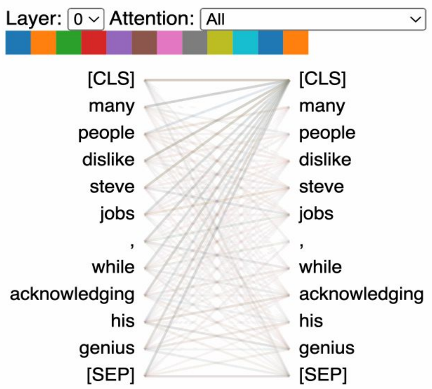


image segmentation



reinforcement learning



natural language processing

Deep Learning has revolutionized various domains, from computer vision, natural language processing, to reinforcement learning

When not to use ML !

Do we need a ML model? Can we not use a model based on real theoretical understanding of the subjects

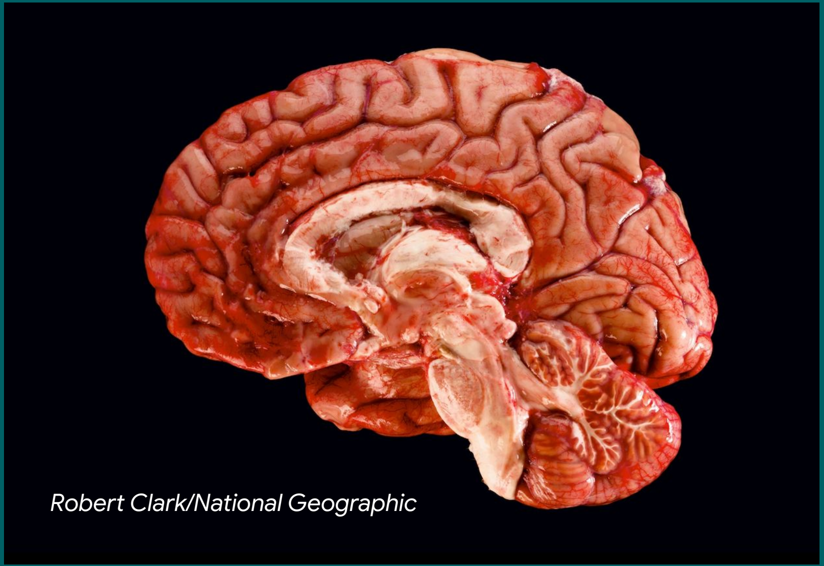
Are enough data available? ML is data-hungry.

Are the data stationary and complete? Generalizing is hard and risky.

What performance guarantees do we need? Is it ok to only have “statistical” guarantees?

Do we need interpretability? ML methods are often “black-box” and don’t provide explanations for their decisions – there are ethical concerns related to this.

What’s the largest neural network you know?



ResNet152? 60 million parameters

GPT3? 175 billion parameters

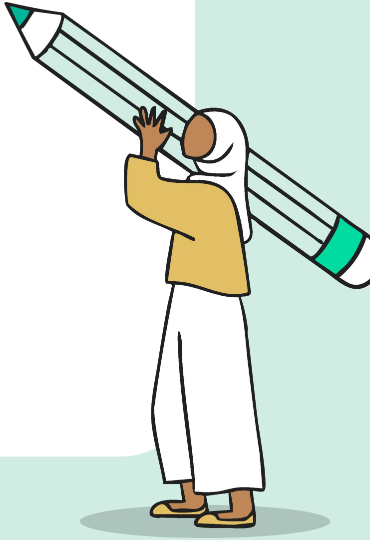
The brain contains about 80-100 billion neurons connected to each other by 100+ trillion synapses (that’s 10^{14}!).

Each of these neurons is a lot more complicated than a ReLU. The tiny worm *C. Elegans* exhibits complex behaviour with just 302 neurons and 7000 synapses.

Women Techmakers

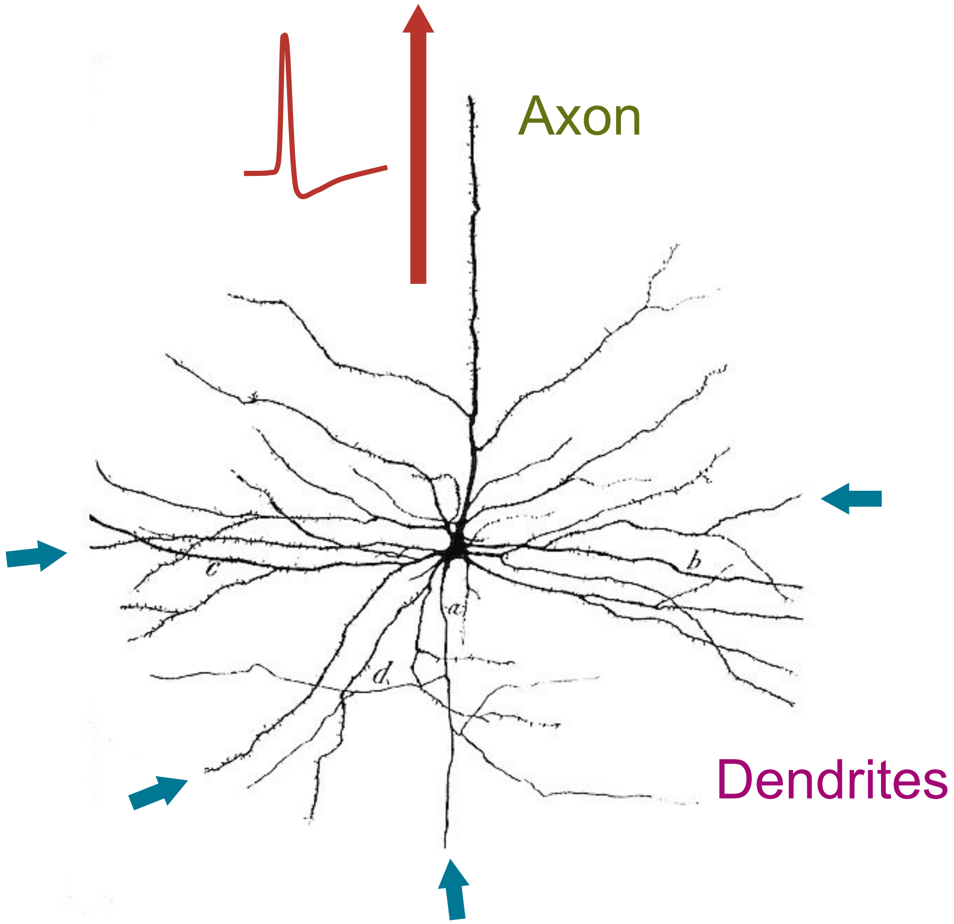
02 Neural Networks

Biology & ML

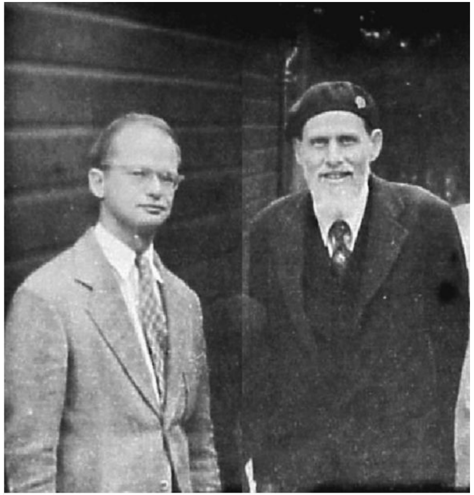


‘Real Neuron’

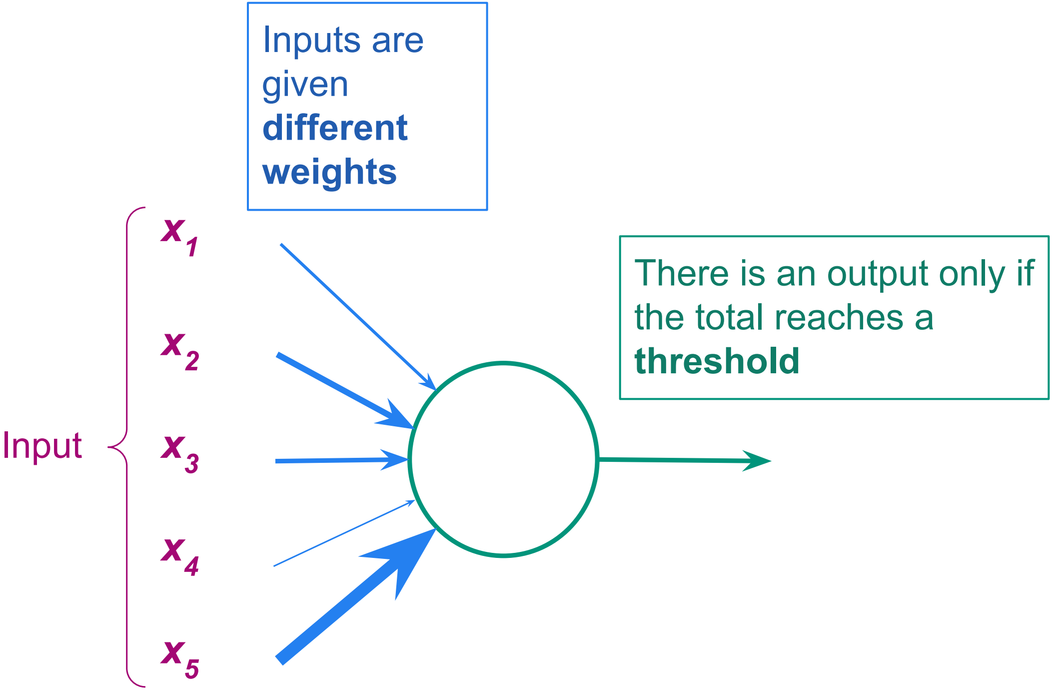
A real neuron is characterized by a tree structure where the branches are dendrites and the center is an axon connecting it to other neurons, the axons and dendrites connect in a place called the soma.



The McCulloch-Pitts neuron

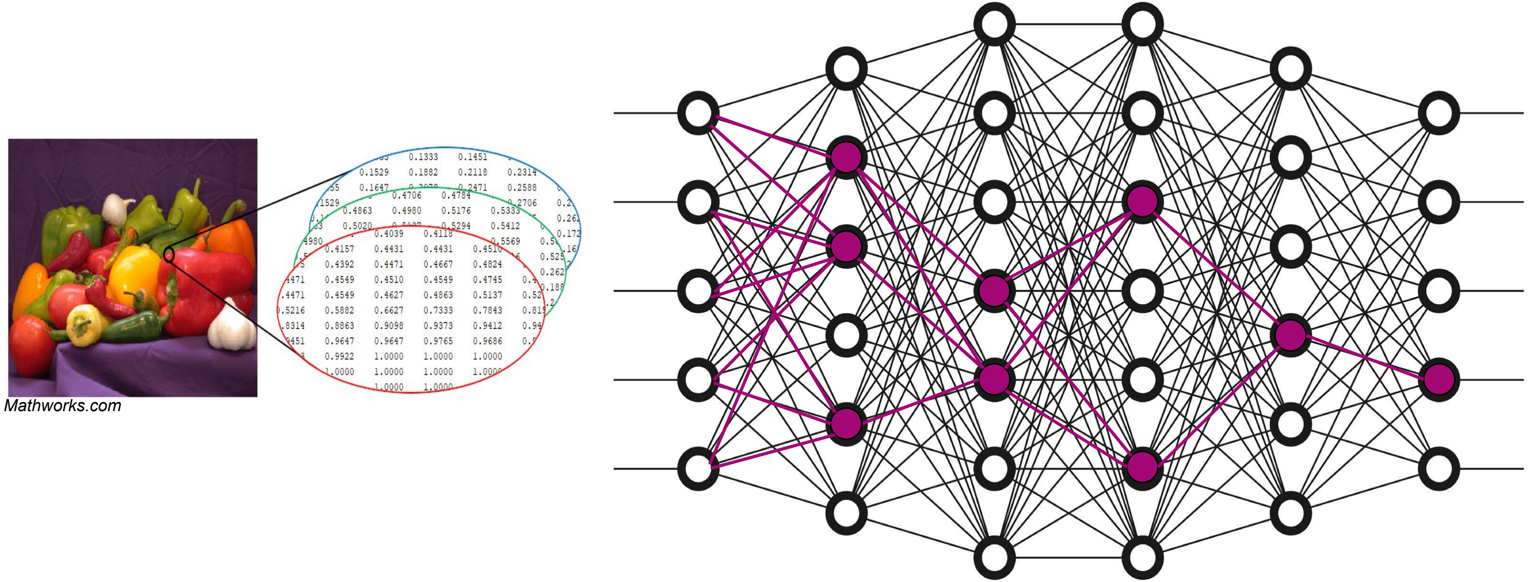


A logical calculus of the ideas immanent in nervous activity (1943)



Any kind of mathematical or logical operation can be expressed as a combination of these "neural" operations

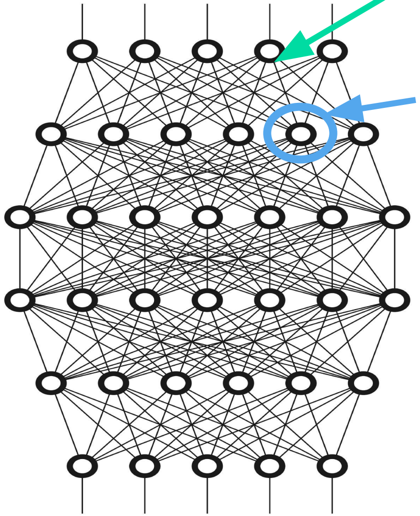
Neuron to network: deep learning



Mathworks.com

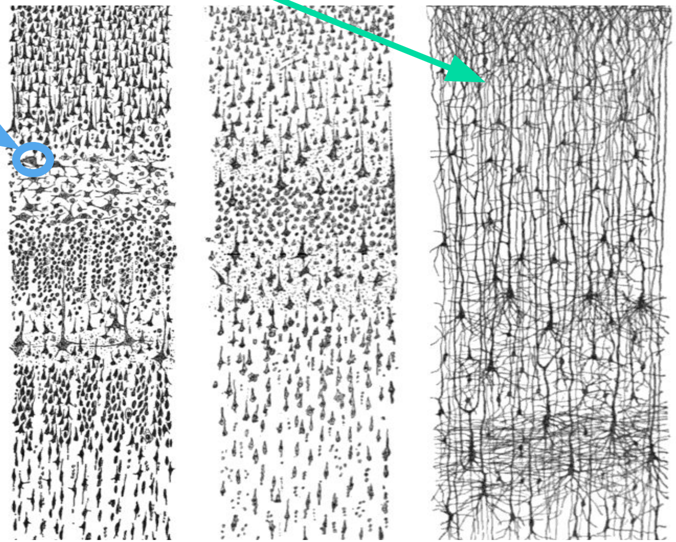
High-level structural similarities

Connections are weighted



Composed of many individual units ("neurons")

Network structure, somewhat layered

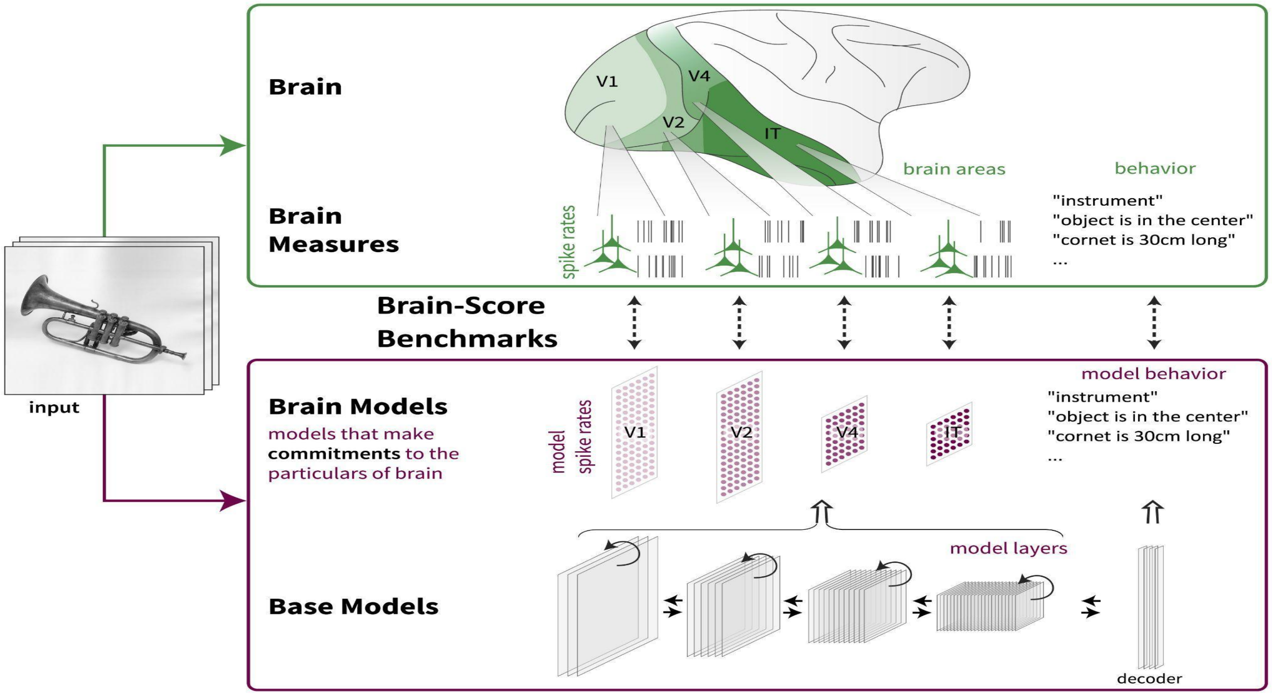


Schematic of an artificial neural network (multilayer perceptron)

Drawing of the human visual cortex by Santiago Ramón y Cajal (Nobel laureate 1906)

Slide by Sorbaro, 2024, TreND-CAMinA

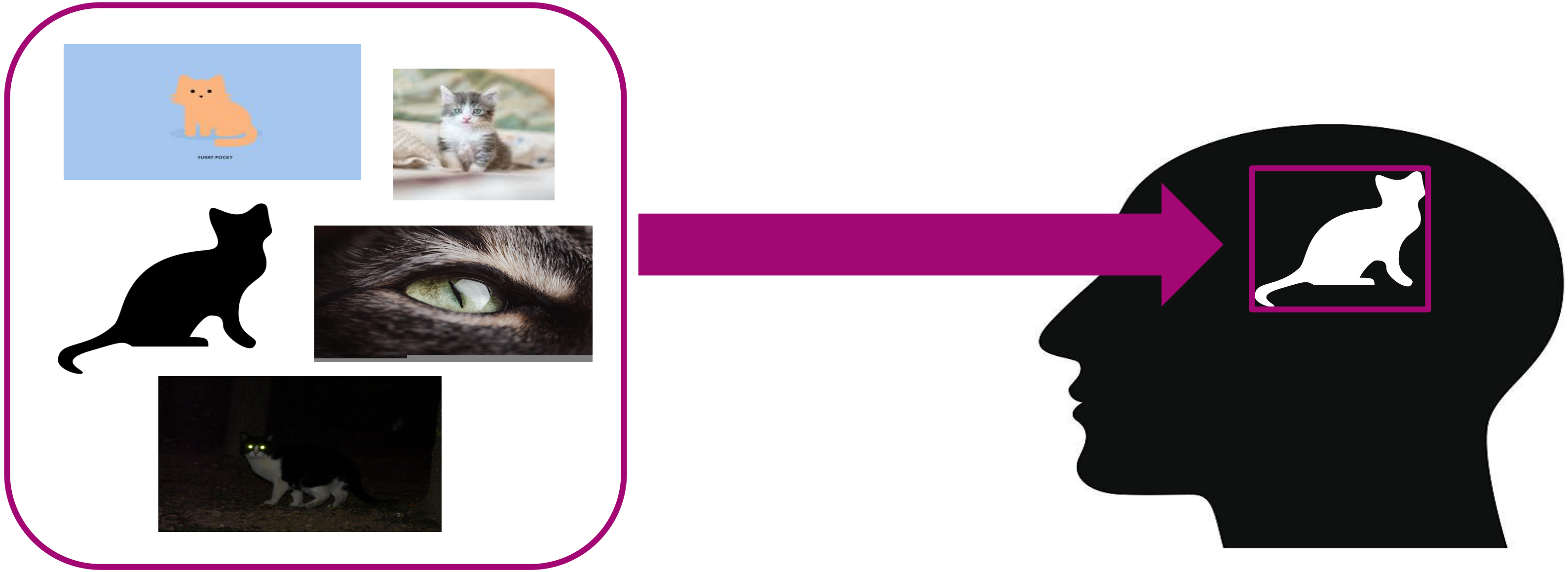
Similarity of representations – CNNs and visual cortex



- CNNs trained to do object recognition predict the activity of brain cortex on same image
- It works better than actually fitting cortex data!
- Layer-by-layer correspondence:
  - early layers == retina, V1
  - intermediate layers == V2/V4
  - last layers == IT cortex

Figure from Schrimpf et al., *Neuron* (2020)

Representations



Summary

Common aspects

Structural

- Nonlinear units
- Connected in a network
- Plastic connections

Empirically observed

- Encode information in analogous ways (similar **representations**)

Differences

Structural

- Spikes
- Continuous time
- A lot more recurrence
- Complexity and diversity
- Dale's law

Learning

- Backpropagation is not biologically plausible
- Loss function? No clue

Limitations of Backpropagation

Credit assignment

Updating the weights based on the weight transport across the network.

Update locking

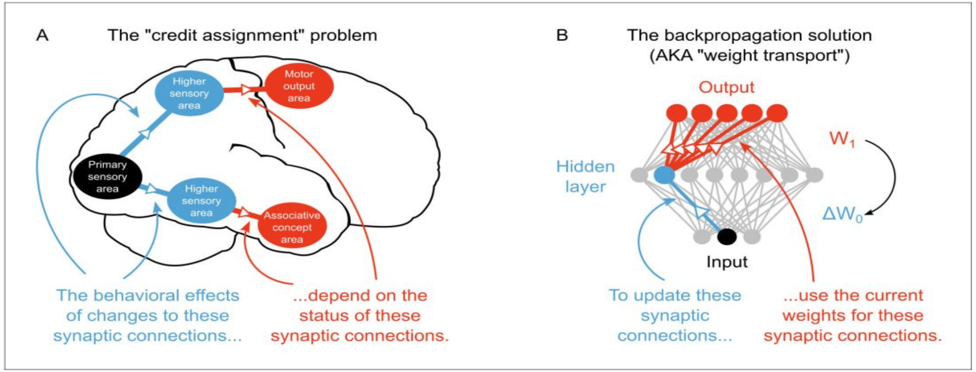
Needs to complete forward pass to calculate updates, which requires more memory and time.

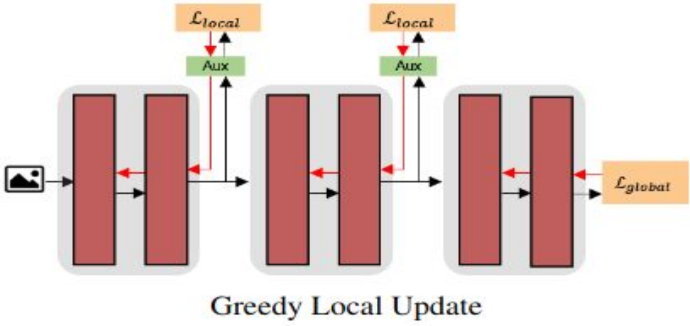
Global loss function

Storing the network weights and activations incurs intensive consumption of memory while waiting for the loss calculation.

Non local synaptic plasticity

Neurons get information and updates from neurons in the other side of the network, while neurons in the brain observe neighboring neurons only.





Figures from Guerguiev et al. *eLife* 2017 | Bhatti, H. I., & al HAET *ICML* (2022) |

The quest for bio-plausible alternative to backpropagation

Neuron Selectivity as a Biologically Plausible Alternative to Backpropagation

C.J. Norsigian
Department of Bioengineering
University of California, San Diego
cnorsigi@eng.ucsd.edu

Jesper Pedersen
Department of Physics
University of California, San Diego
jrpeders@ucsd.edu

Vishwajith Ramesh
Department of Bioengineering
University of California, San Diego
vramesh@eng.ucsd.edu

Hesham Mostafa
Institute for Neural Compu
University of California, Sa
hmmostafa@ucsd.ed

Towards Biologically Plausible Deep Learning

Yoshua Bengio[1], Dong-Hyun Lee, Jorg Bornschein, Thomas Mesnard and Zhouhan Lin
Montreal Institute for Learning Algorithms, University of Montreal, Montreal, QC, H3C 3J7
[1]CIFAR Senior Fellow

Biol(
time for learning in recurrent neural nets

Guillaume Bellec*, Franz Scherr*, Elias Hajek , Darjan Salaj , Robert Legenstein ,

Institute for Tl

Towards deep learning with segregated dendrites

Jordan Guerguiev, Timothy P Lillicrap, Blake A Richards
University of Toronto Scarborough, Canada; University of Toronto, Canada; DeepMind, United Kingdom; Canadian Institute for Advanced Research, Canada

Forward Learning with Top-Down Feedback: Empirical and Analytical Characterization

Ravi Francesco Srinivasan [1 2] Francesca Mignacco [3 4] Martino Sorbaro [5 6] Maria Refinetti [7 8] Avi Cooper [9] Gabriel Kreiman [10 11] Giorgia Dellaferrera [10 11 1 5]
[1] ETH Zurich, Switzerland

tion in the Brain

10.1016/j.tics.2018.12.005

Top-down vs. bottom-up modelling

Show it can perform the desired function

Adjust until

Biological model of neurons, synapses, small networks

Bottom-up modelling

Neural network that is functional on task

Adjust until

More biologically plausible and interesting

Top-down modelling

The advent of deep learning **shifted** the computational neuroscience community more towards this way of thinking.

Learning rules are now a lot more involved in neuroscience research.

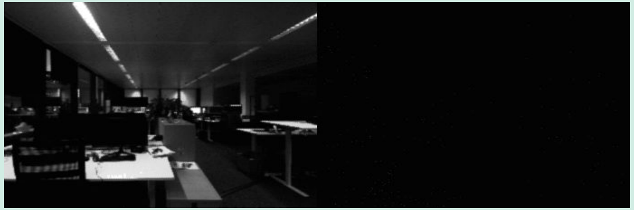
The brain is low-power: neuromorphic engineering

The brain uses about **15-25 W** of power to run its billions of neurons and perform advanced cognition.

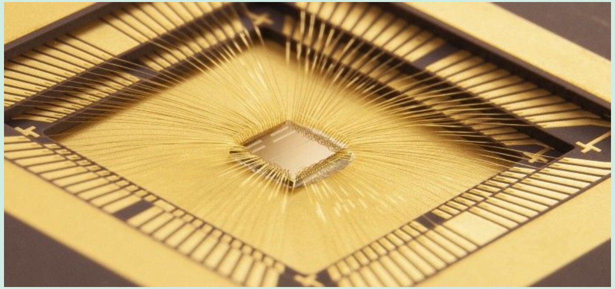
Clearly, deep learning technology has a lot to learn.

Neuromorphic engineering designs **specialised hardware** that simulates neurons on-chip.

These can be used for always-on battery-powered devices.



Dynamic Vision Sensors are "spiking cameras" used for low-power, low-latency, high dynamic range uses.
They behave similarly to artificial retinas.



A convolutional neuromorphic chip (SynSense AG, Zurich)

The brain can do amazing things

- Out of distribution generalisation (novel environments, …)
- Excellent transfer learning, few-shot learning, continual learning
- Multi-modal integration (motor, auditory, visual, language…)
- Abstract planning, multi-step reasoning

Studying neuroscience and cognitive science can help us understand what is missing in AI

DEEP LEARNING NEEDS A PREFRONTAL CORTEX

Jacob Russin
Psychology Department
University of California Davis
Davis, California
jlrussin@ucdavis.edu

Randall C. O'Reilly
Psychology Department
Center for Neuroscience
Computer Science Department
University of California Davis
oreilly@ucdavis.edu

Yoshua Bengio
MILA
Université de Montréal
CIFAR Senior Fellow

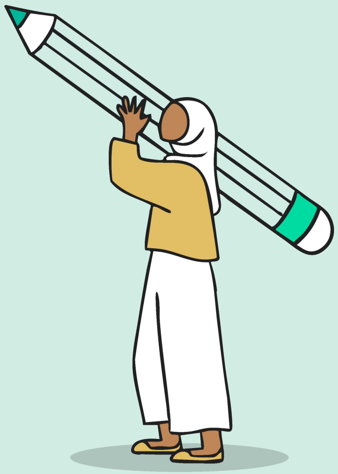
ABSTRACT

Research seeking to build artificial systems capable of reproducing elements of human intelligence may benefit from a deeper consideration of the architecture and learning mechanisms of the human brain. In this brief review, we note a connection between many current challenges facing artificial intelligence and the functions of a particular brain area — the prefrontal cortex (PFC). This brain area is known to be involved in executive functions such as reasoning, rule-learning, deliberate or controlled processing, and abstract planning. Motivated by the hypothesis that these functions provide a form of out-of-distribution robustness currently not available in state-of-the-art AI systems, we elaborate on this connection and highlight some computational principles thought to be at work in PFC, with the goal of enhancing the synergy between neuroscience and machine learning.

What’s next?

- **Feedback alignment**: use a random matrix in the backward pass (Lillicrap et al. Nat. Comm. 2014)
- Dendritic and **three-factor** learning: a reward signal modulates Hebbian learning; exploits complex structure of biological neurons
- **Target propagation** instead of gradient propagation (Lee et al. ICLR 2015)
- **Eligibility traces** (Bellec et al. Nature Comm. 2020)
- …

Active field of research!



Thank you!