Imperial College London

§3.2 MLBD MRes practical Introduction to neural networks Single neuron classifier

Jarvist Moore Frost.

Imperial College London
Email: jarvist.frost@ic.ac.uk
https://jarvist.github.io
Twitter: @JarvistFrost

Intended Learning Outcomes §3.2

Demonstrate how artificial neural networks are built, by constructing a single neuron classifier from scratch.

- ➔ Summarise the biological origin of artificial neural networks. (Slides)
- ➔ Prepare a binary classifier from a single neural network. (Jupyter)
- ➔ Generalise the 'by hand' approach into functions.
- ➔ Compare with a logistical regression approach to the same data.

Recommended reading

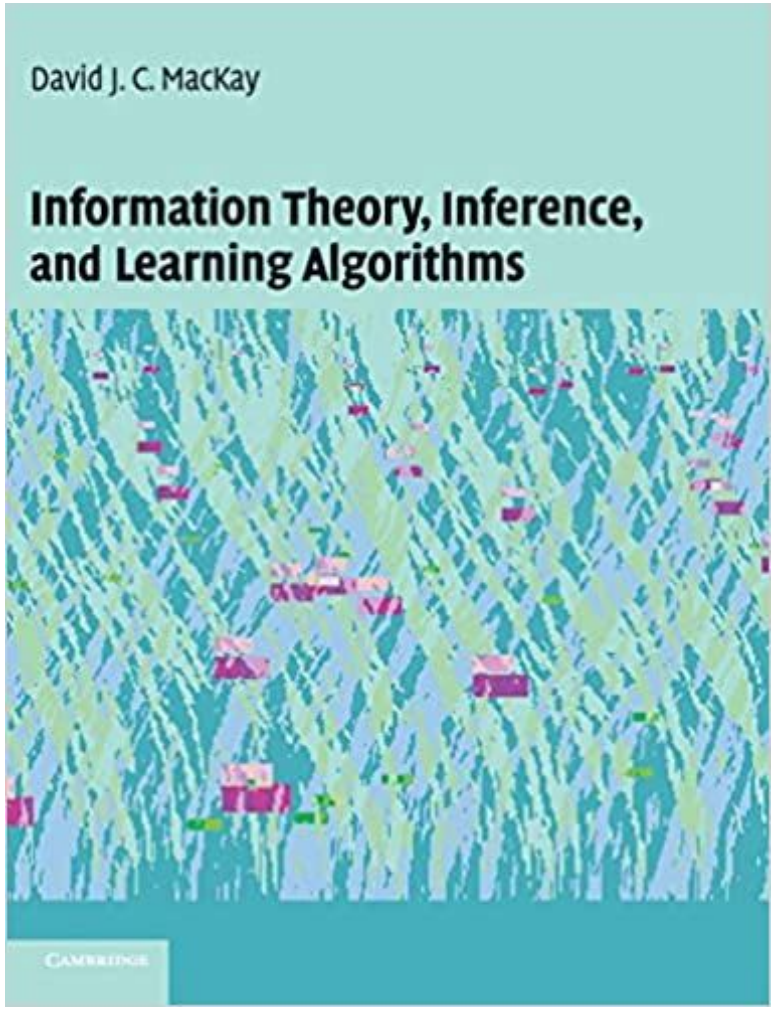
David MacKay, Information Theory, Inference and Learning Algorithms (ITILA), 2003, Chapters 38–42.

Freely available online!

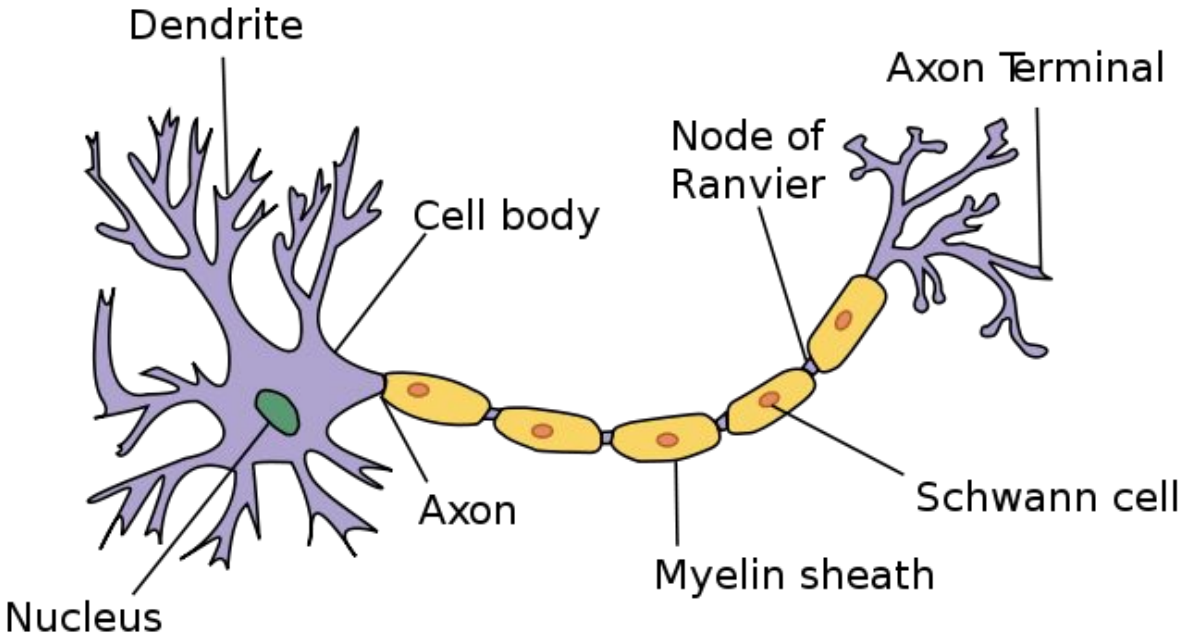
A physicists perspective on ML.

http://www.inference.org.uk/mackay/itila/book.html

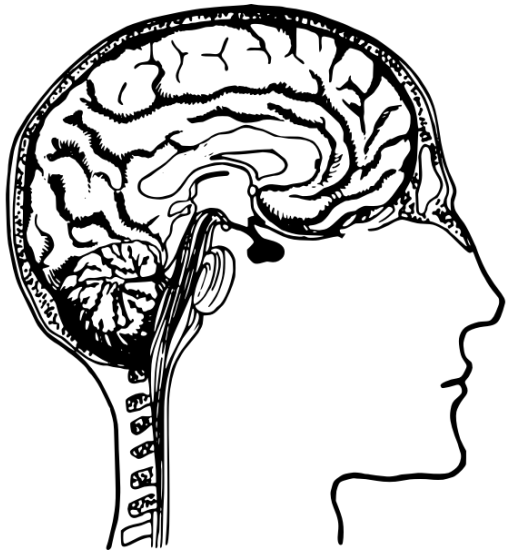
These slides and classworks follow some of the structure of Chapter 38 & 39.



Biological inspiration



"Anatomy and Physiology" by the US National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) Program .



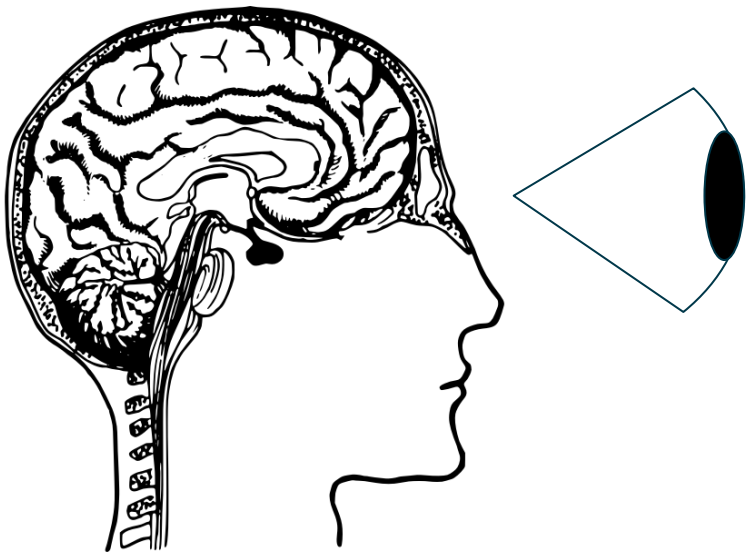
Neurons have multiple inputs / outputs

Relatively (!) simple functional relationship

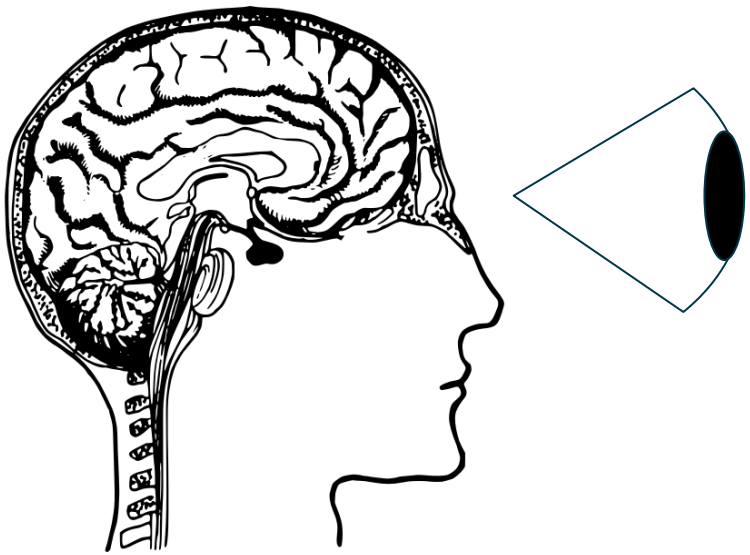
86 billion neurons ⇒ human brain

0.1 million neurons ⇒ fruit fly Drosophila (fully imaged 2020)

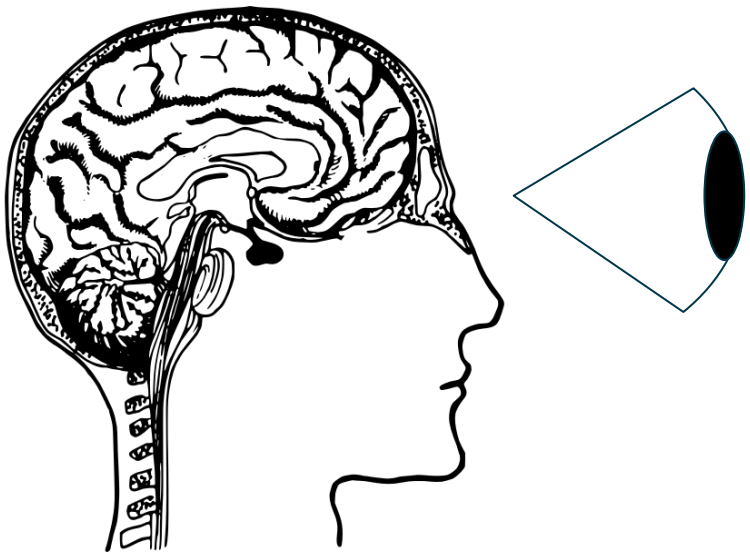
Biological inspiration: memory



Biological inspiration: memory

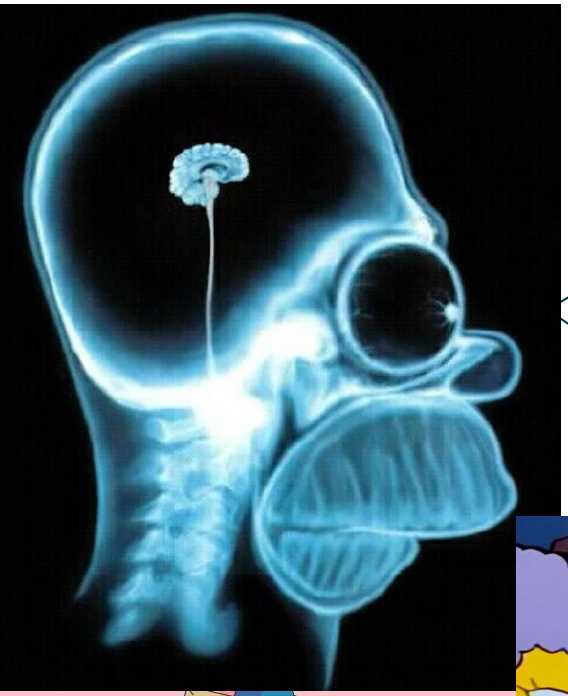


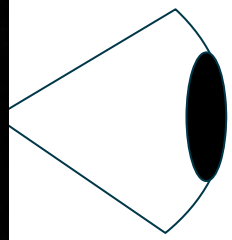
Associative, extrapolative, memory

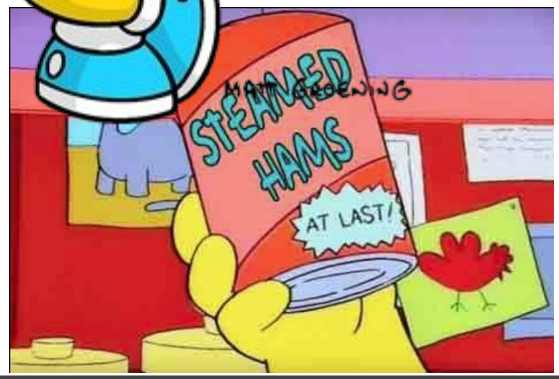


Vs. address based memory in a library / computer.

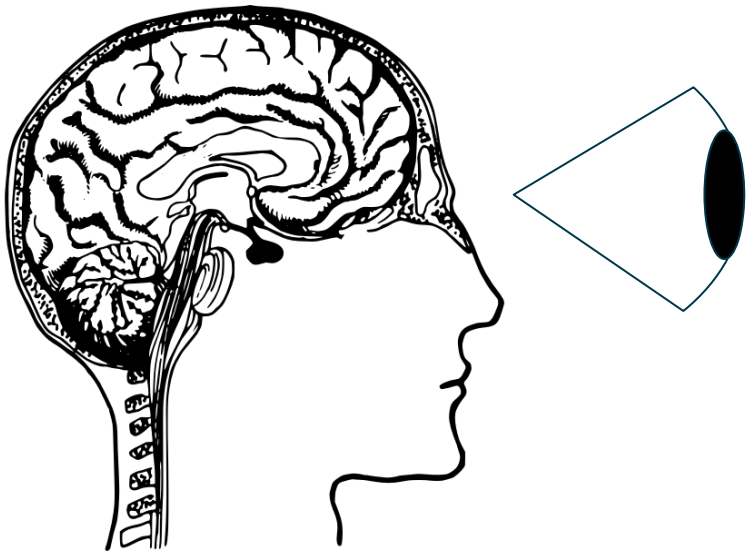
Spontaneous recall





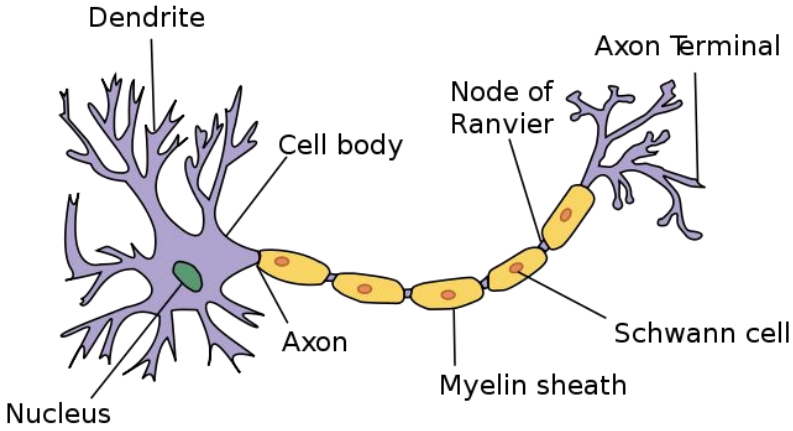


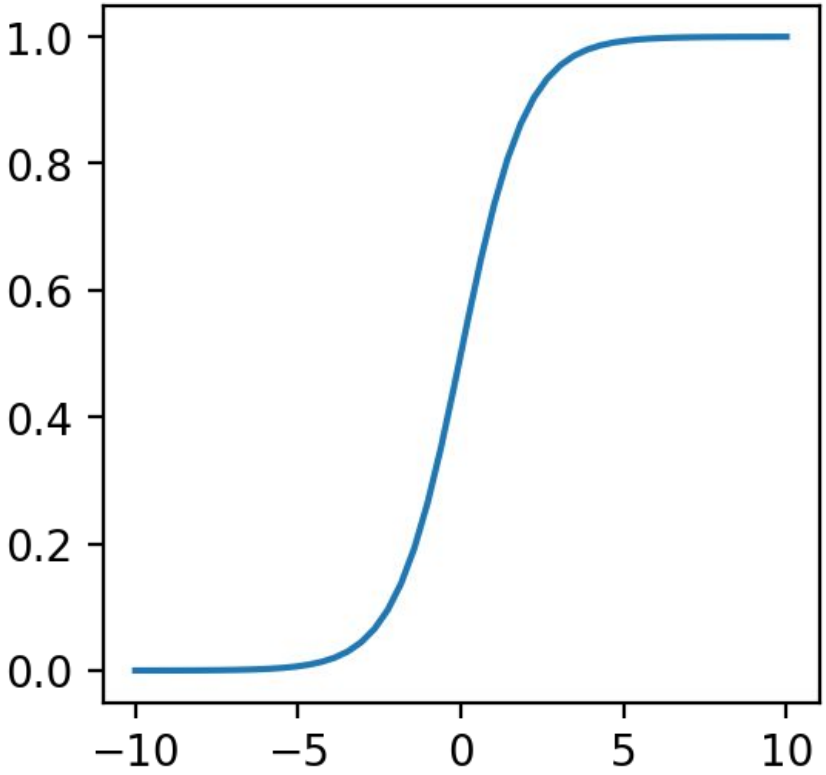
Human perception: fault tolerant



Yuo cna porbalby raed tihs esaliy desptie teh msispeillgns.

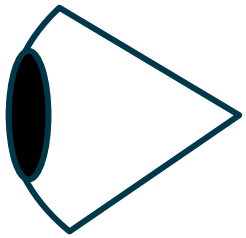
Biology → code





Abstract (and vastly simplify!) neuron behaviour to a sigmoid / tanh

The Mark 1 Perceptron (1957>)



20x20
CdS
photocells

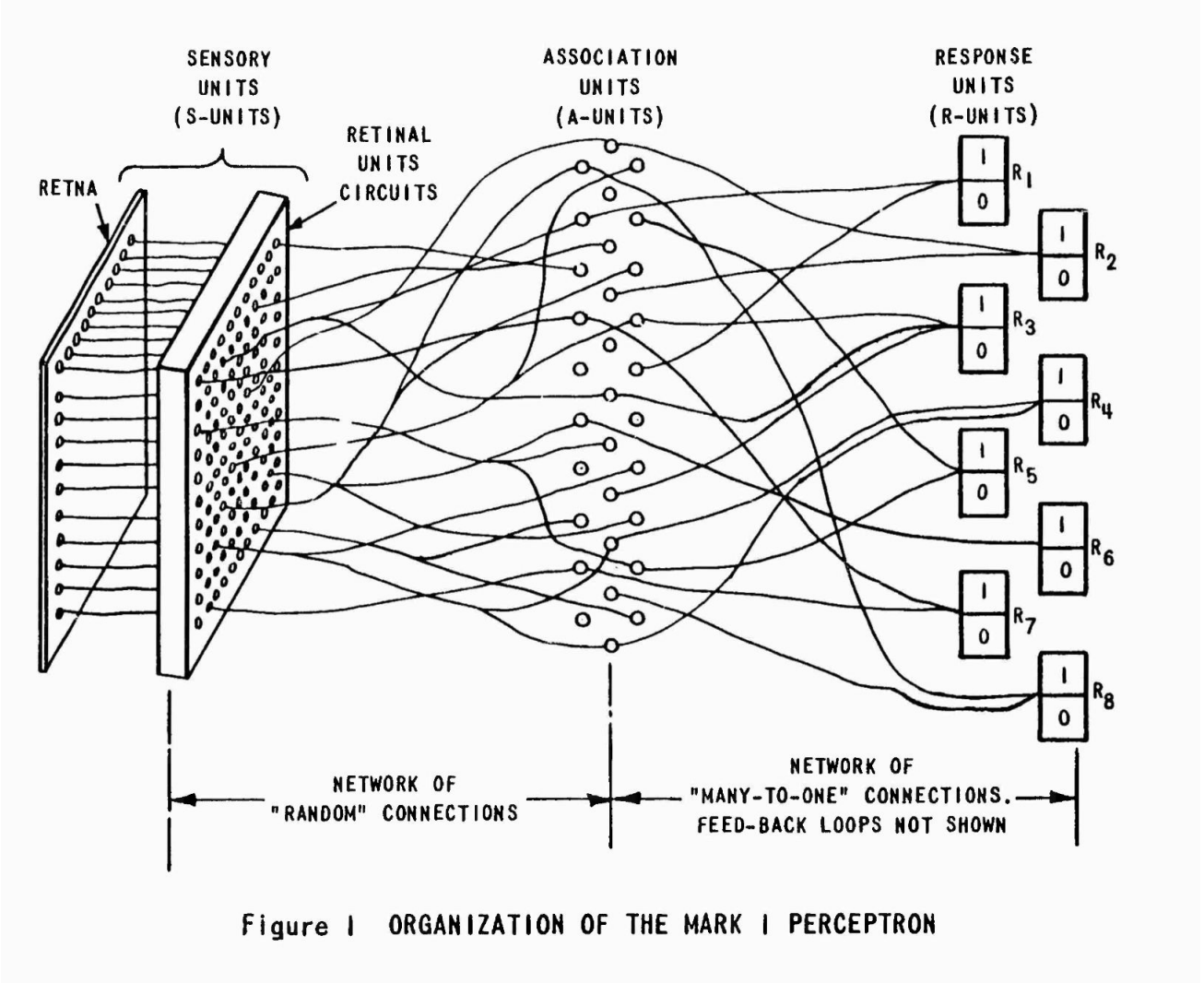
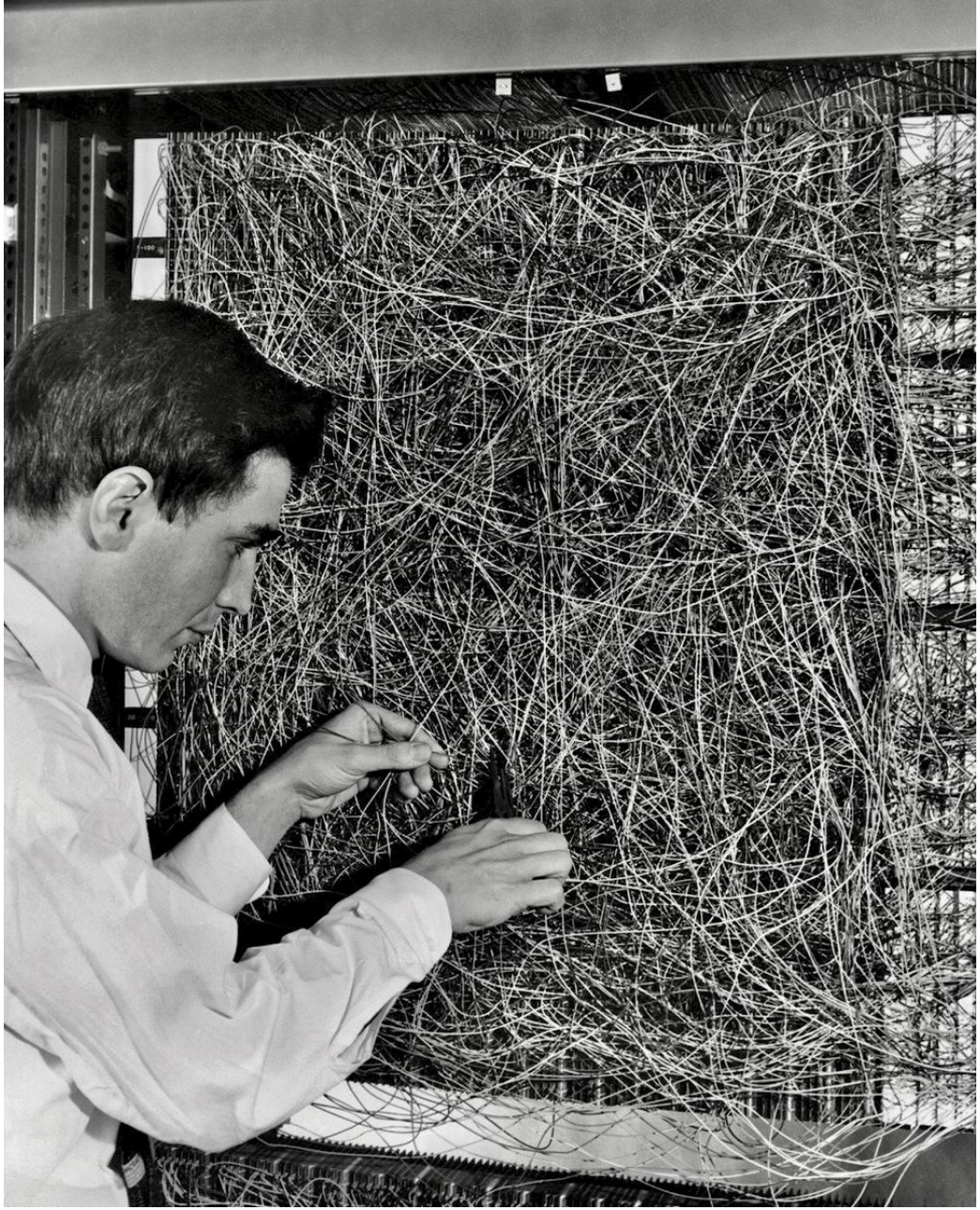


Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON



Frank Rosenblatt, with the 'feature selector' input patch

New York Times, July 8th 1958

NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

WASHINGTON, July 7 (UPI) —The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

The embryo—the Weather Bureau's $2,000,000 "704" computer—learned to differentiate between right and left after fifty attempts in the Navy's demonstration for newsmen.

The service said it would use this principle to build the first of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of $100,000.

Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human beings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers.

Without Human Controls

The Navy said the perceptron would be the first non-living mechanism "capable of receiving, recognizing and identifying its surroundings without any human training or control."

The "brain" is designed to remember images and information it has perceived itself. Ordinary computers remember only what is fed into them on punch cards or magnetic tape.

Later Perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech or writing in another language, it was predicted.

Mr. Rosenblatt said in principle it would be possible to build brains that could reproduce themselves on an assembly line and which would be conscious of their existence.

In today's demonstration, the "704" was fed two cards, one with squares marked on the left side and the other with squares on the right side.

Learns by Doing

In the first fifty trials, the machine made no distinction between them. It then started registering a "Q" for the left squares and "O" for the right squares.

Dr. Rosenblatt said he could explain why the machine learned only in highly technical terms. But he said the computer had undergone a "self-induced change in the wiring diagram."

The first Perceptron will have about 1,000 electronic "association cells" receiving electrical impulses from an eye-like scanning device with 400 photo-cells. The human brain has 10,000,000,000 responsive cells, including 100,000,000 connections with the eyes.

1958 New York Times...

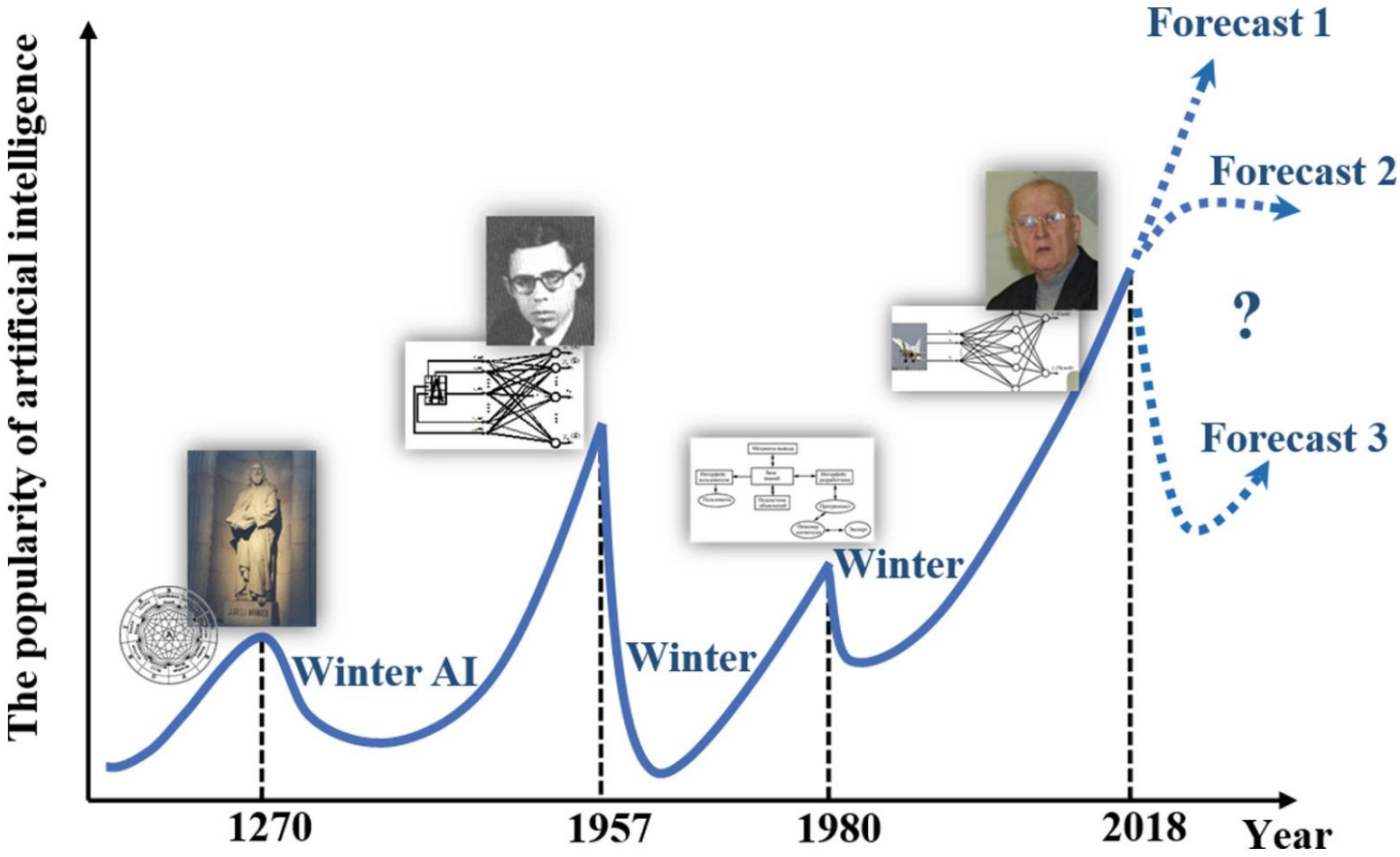
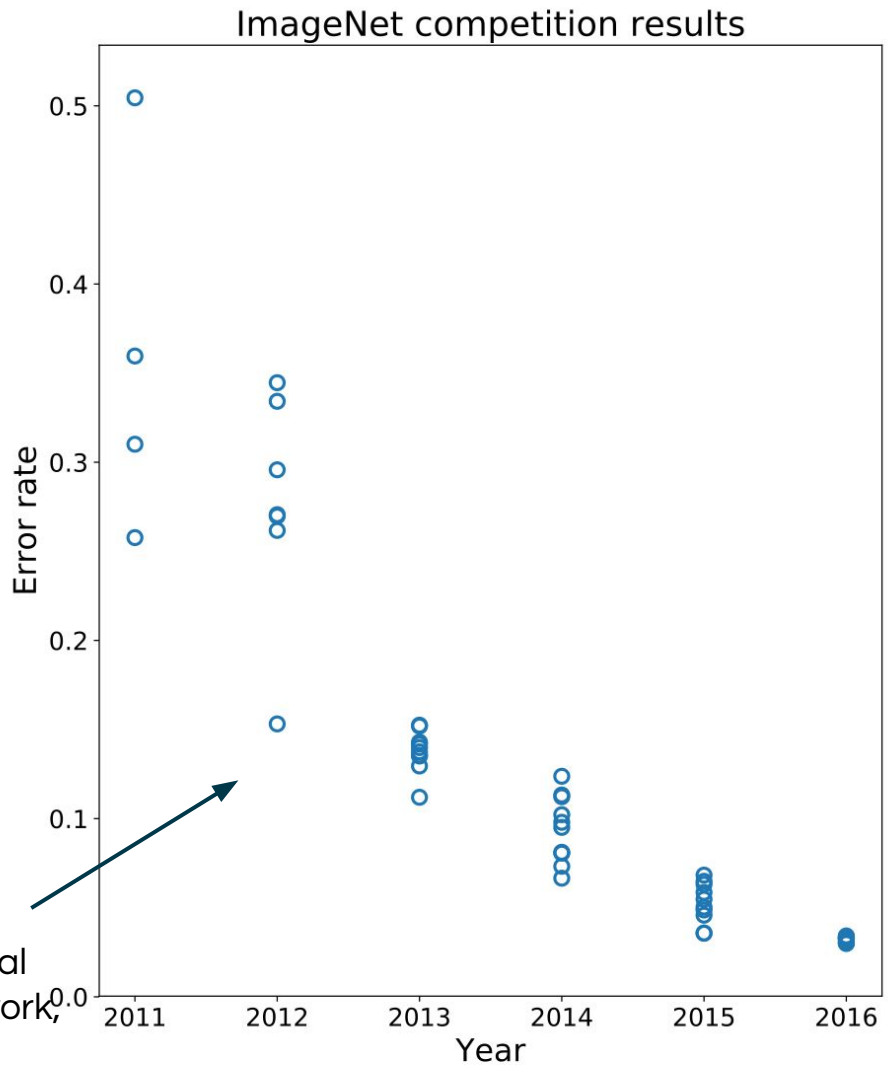


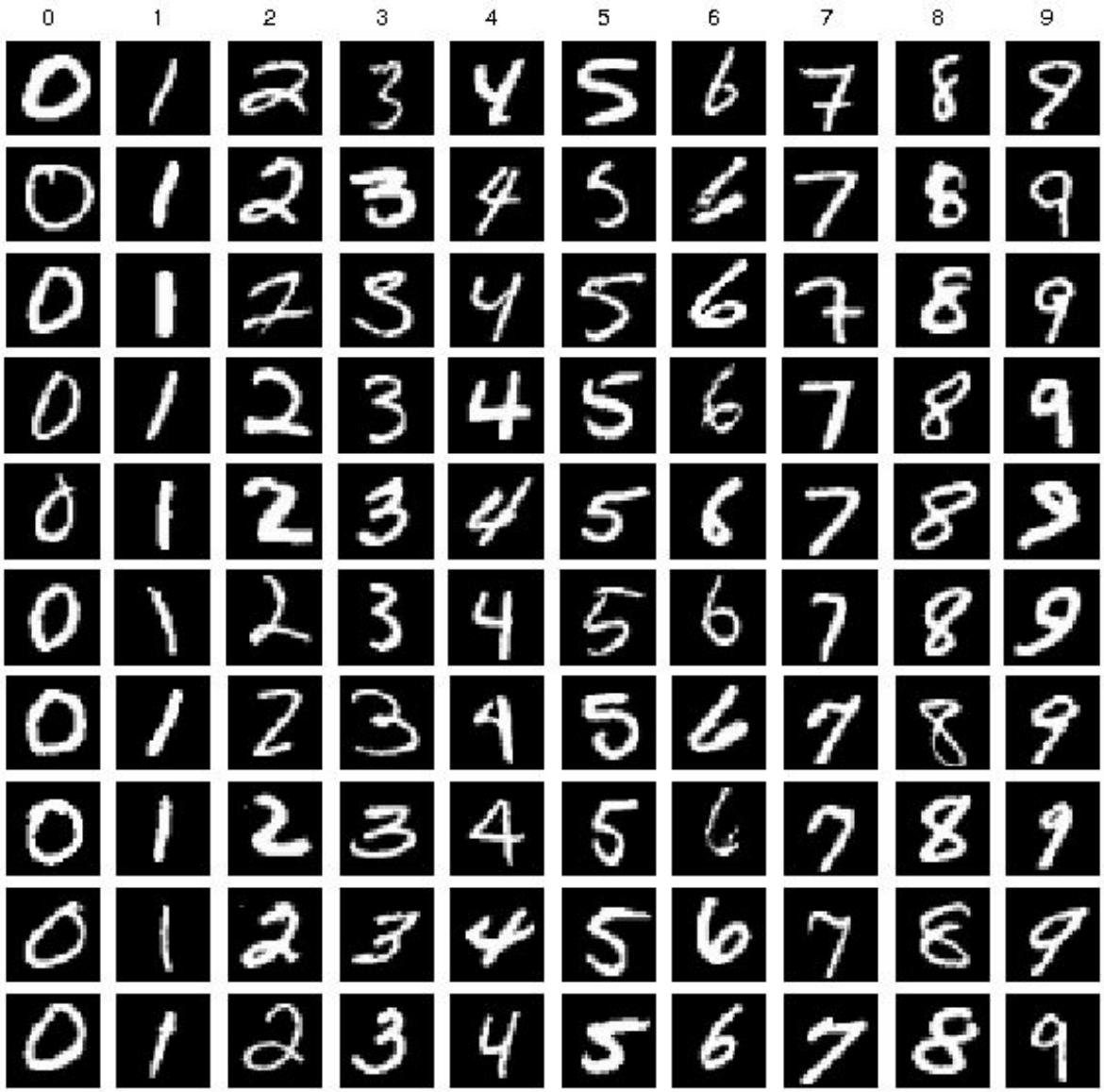
Fig. 1. Periods of surges and falls artificial intelligence. The variants of predictions of the future.

Yasnitsky, L. N. (2019). *Whether Be New "Winter" of Artificial Intelligence? Lecture Notes in Networks and Systems, 13–17.* doi:10.1007/978-3-030-22493-6_2
Paper argues that 'Forecast 3' & a new winter is inevitable!



AlexNet,
A deep
convolutional
neural network,
16% error

In 2012 neural networks mostly solved computer vision (Imagenet = 14 million annotated images.)



MNIST - a relatively simple, 'toy', dataset.

§3.2 Practical

- Add missing code:
  - Sigmoid activation function
  - Neuron function
  - Training loop
- You can now train a model & visualise the decision boundary!
  - Document how the training performs, with different training rates ('eta') and weight decay ('alpha)
  - Compare training with a simple linear classifier
  - How does the neuron fair on the more difficult Gaussian dataset?
- Advanced concepts
  - Change the activation function. (Think about the gradients.)
  - Often neural networks have an additional 'bias' input. Add this to your code.
  - Batch training - currently all data is used to build the single gradient.
  - What happens if you try and use the method for regression (against a function)?
  - Can you improve the regression performance by adding extra neurons side-by-side, each fitting a different part of the function?
  - Compare regression to Gaussian processes: https://jarvist.github.io/2021-PhysicsMachineLearningPracticum/02_GaussianProcessPotentialEnergySurface.html
- Suggested homework / self-study
  - David MacKay, Information Theory, Inference and Learning Algorithms (ITILA), 2003, Chapters 38–42.
  - PyTorch '60 minute Blitz' https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html
  - https://fluxml.ai/tutorials/2020/09/15/deep-learning-flux.html - Julia ML library, the tutorial is based on the Pytorch 60 minute Blitz