

Machine Learning for Physicists

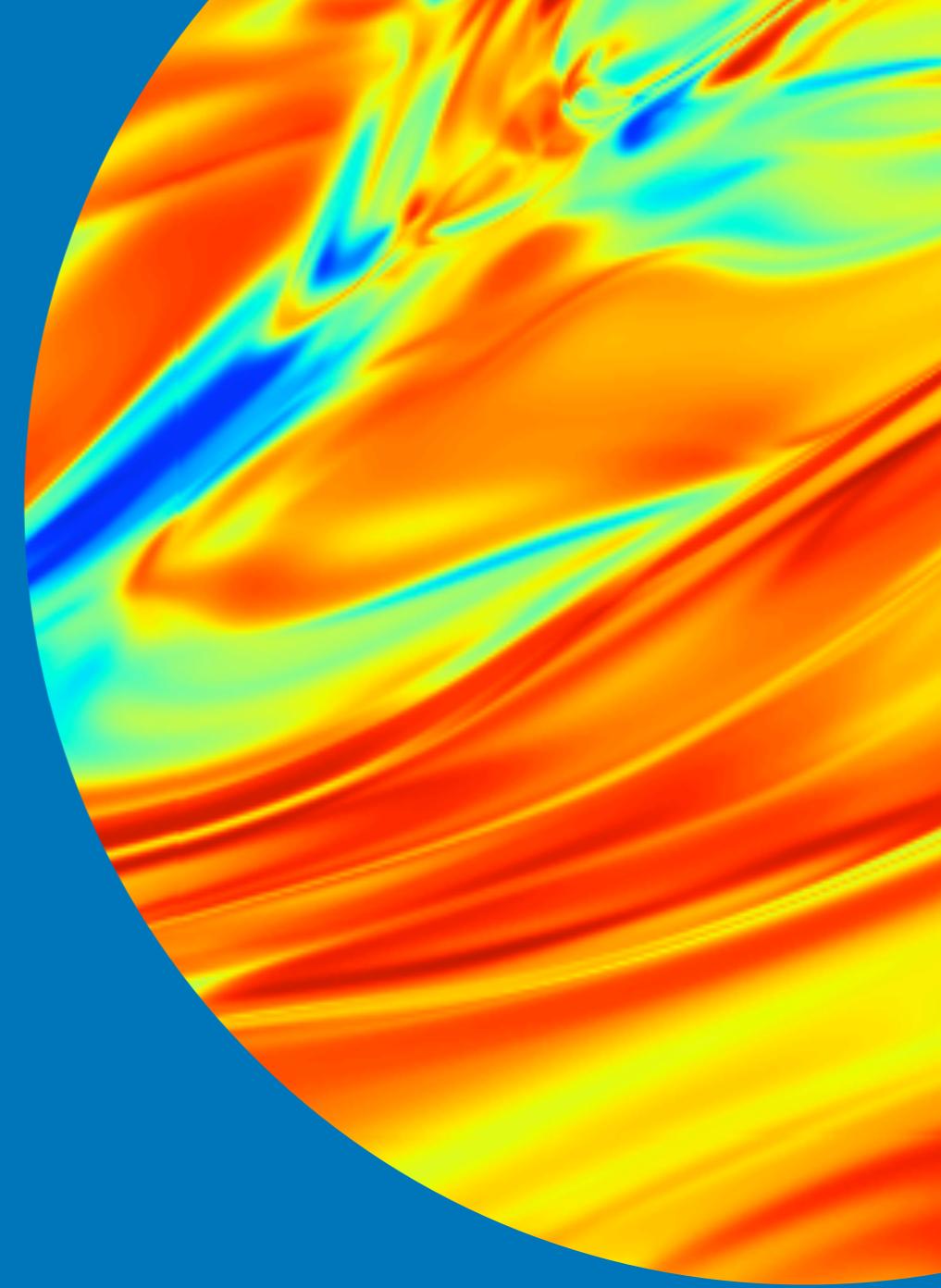
Summer 2025

Florian Marquardt

University of Erlangen-Nuremberg

& Max Planck Institute for the Science of Light

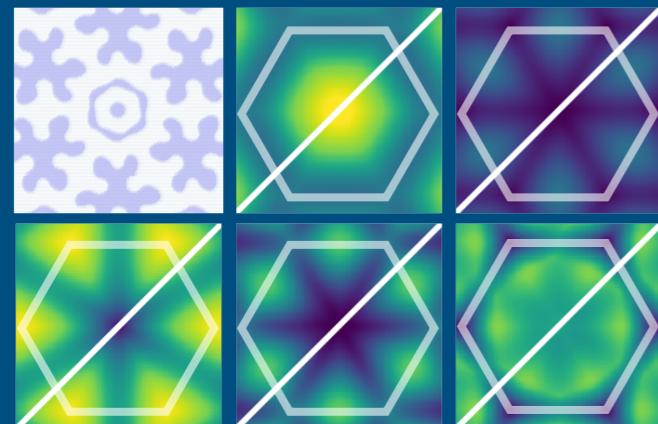
Florian.Marquardt@fau.de



Machine Learning & Physics @ Marquardt group

Theory Division / Max Planck Institute for the Science of Light

ML for photonic systems



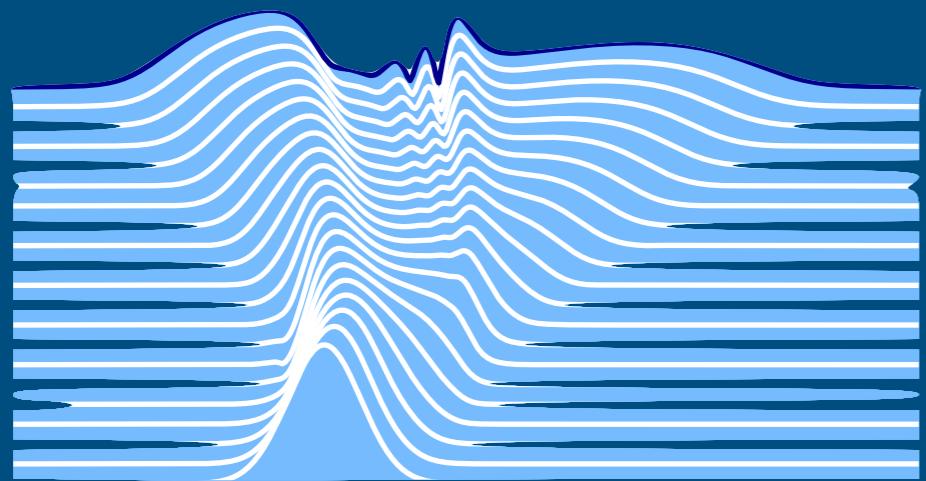
e.g. predicting and optimizing
topological band structures

Artificial scientific discovery



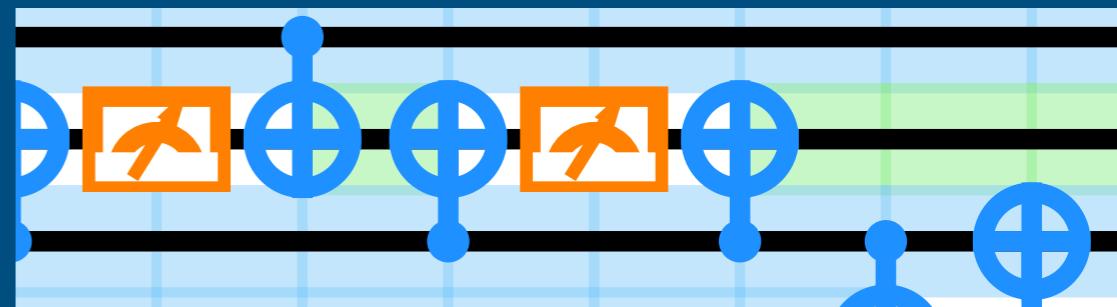
e.g. discovery of collective
coordinates

Physics for ML

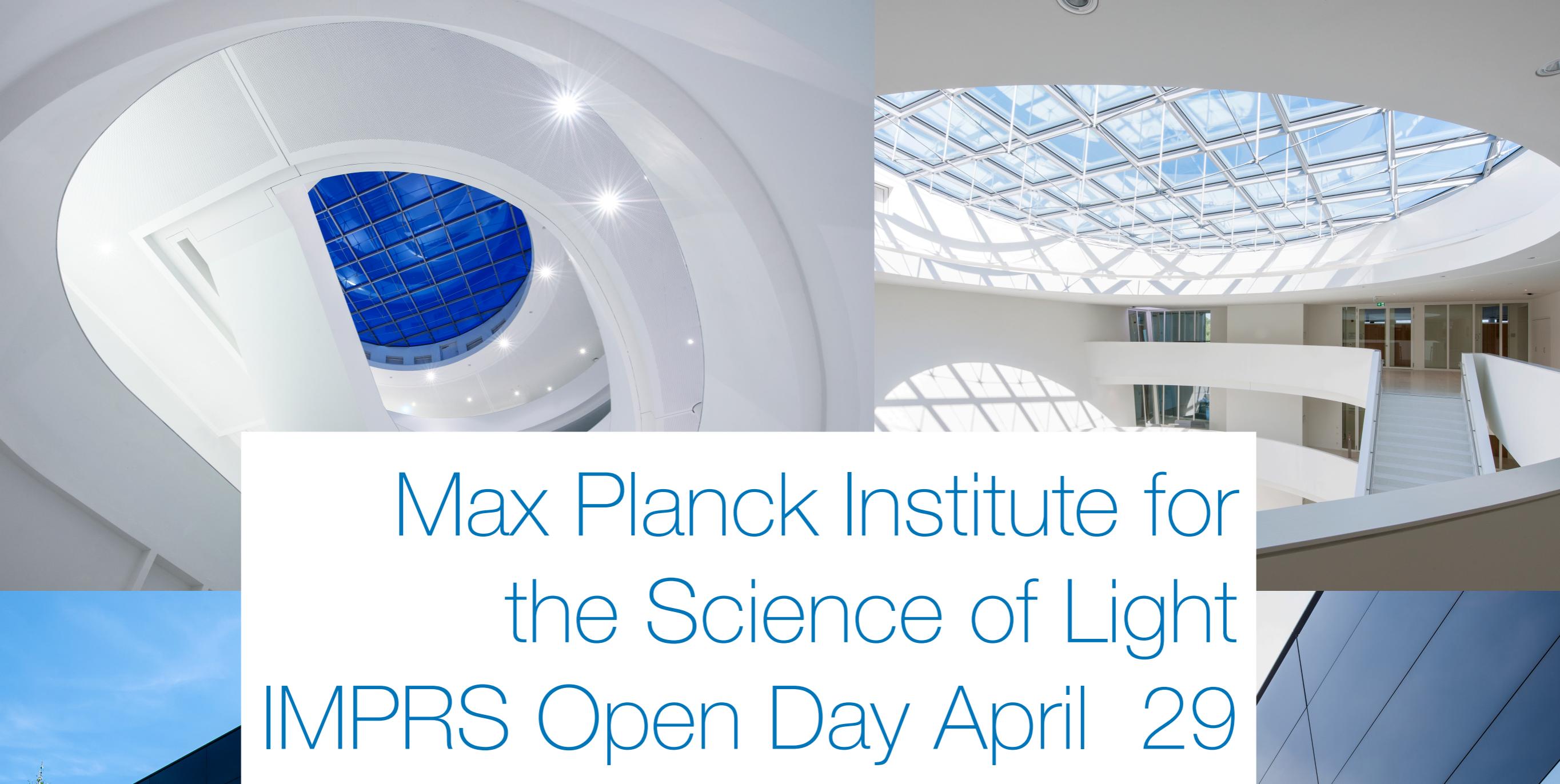


self-learning machines

ML for quantum technology



e.g. reinforcement learning for
quantum feedback

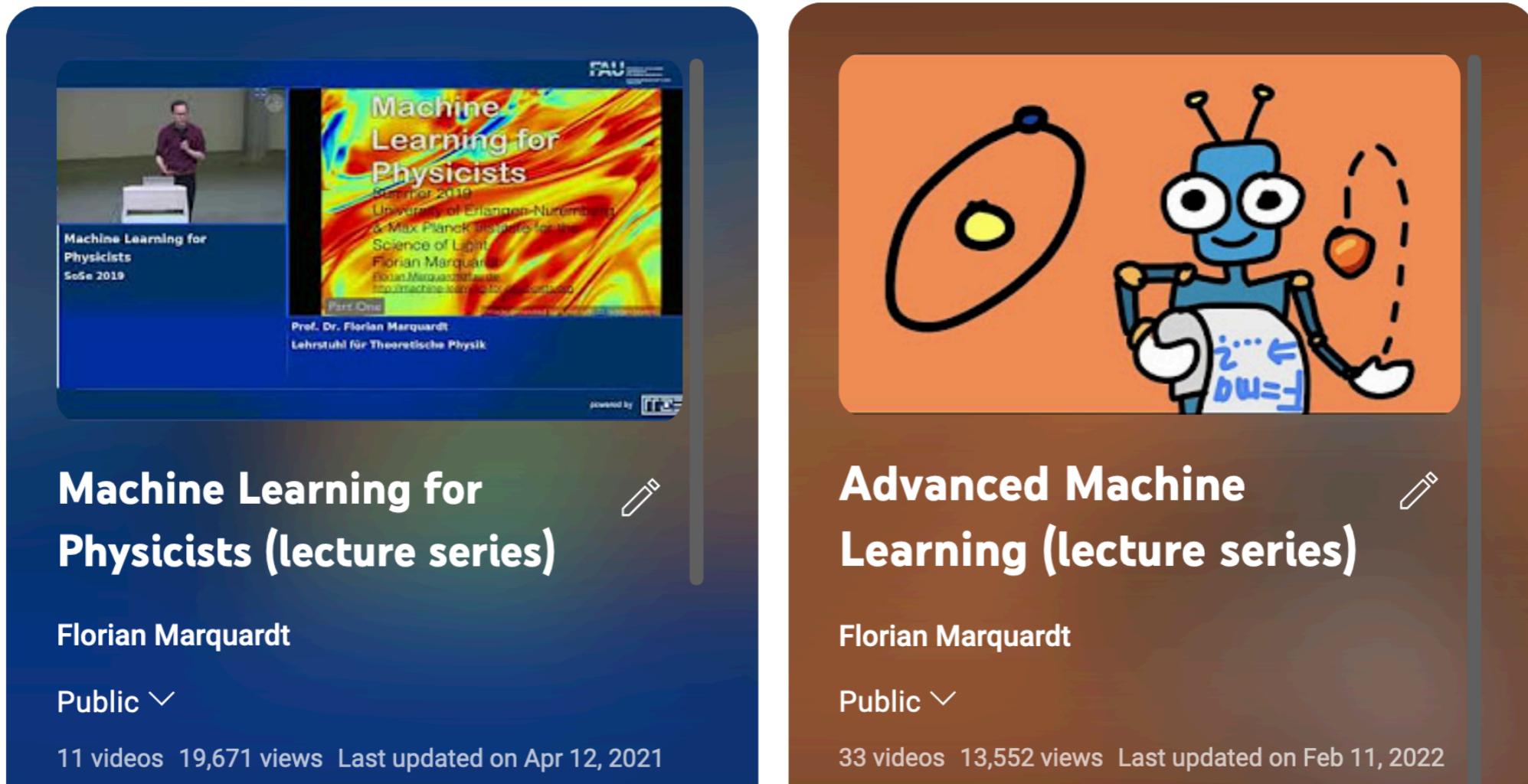


Max Planck Institute for the Science of Light

IMPRS Open Day April 29

Learn more...

YouTube @florian_marquardt_physics



Review on Machine Learning for Quantum Technologies:
Krenn, Landgraf, Fösel, F.M. Phys. Rev. A 107, 010101 (2023)
Dawid et al, arXiv:2204.04198

<https://mpl.mpg.de/divisions/marquardt-division/machine-learning-for-physics-science-and-artificial-scientific-discovery>

Today

Teaser: the power of machine learning

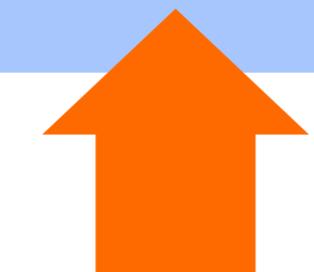
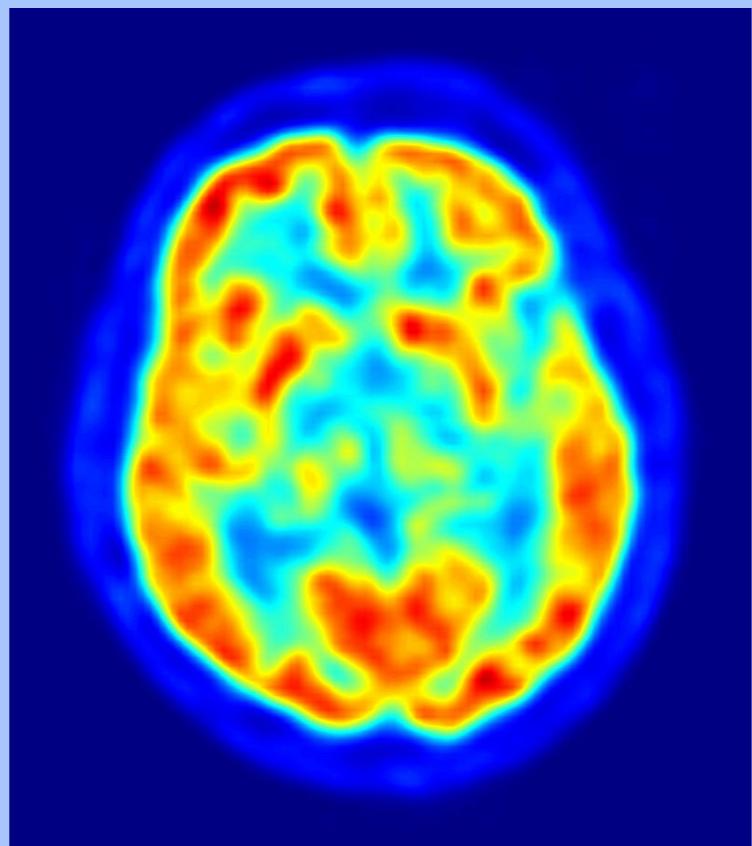
About this lecture

Website/Code

Organizational Details

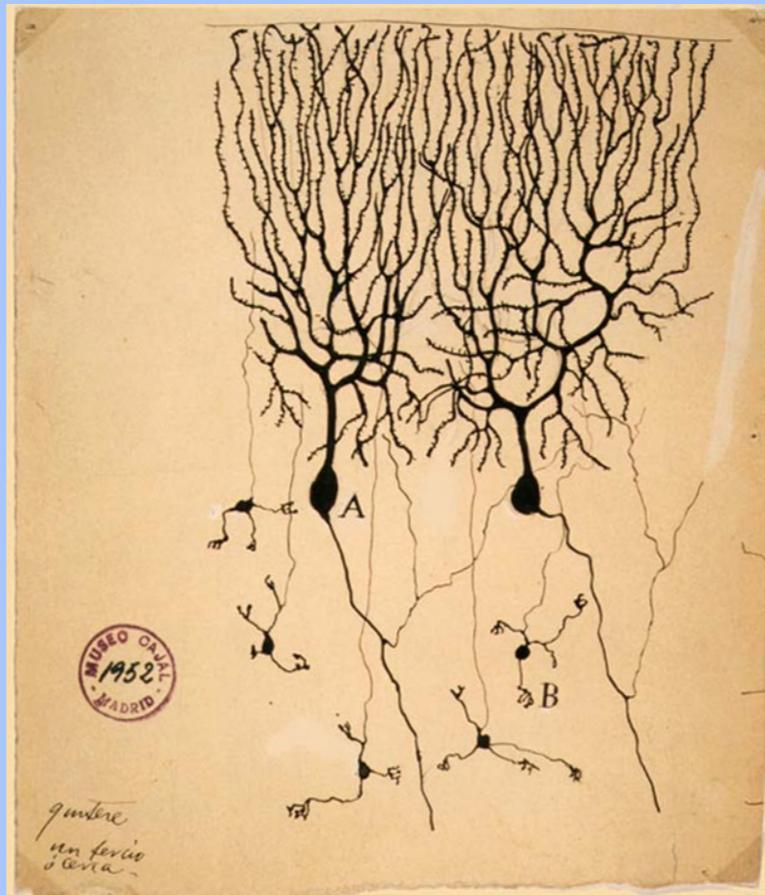
Historical remarks

OUTPUT



INPUT

OUTPUT



INPUT

(drawing by
Ramon y Cajal,
~1900)

OUTPUT



**Artificial
Neural Network**

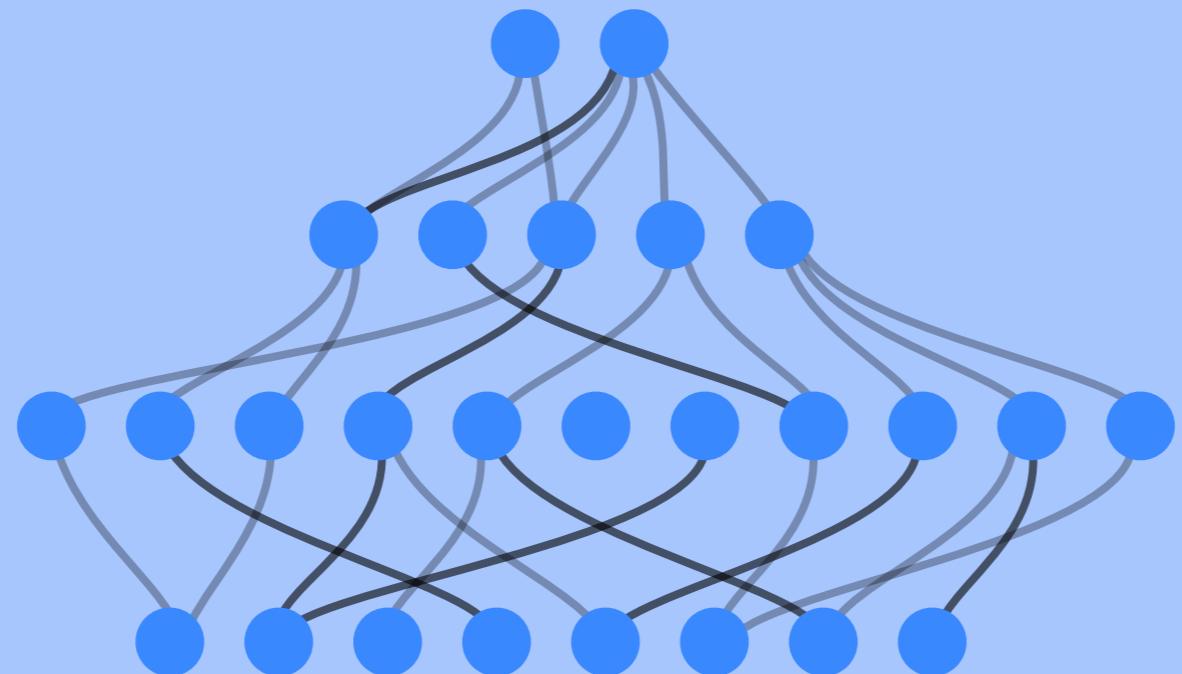


INPUT

OUTPUT



output layer

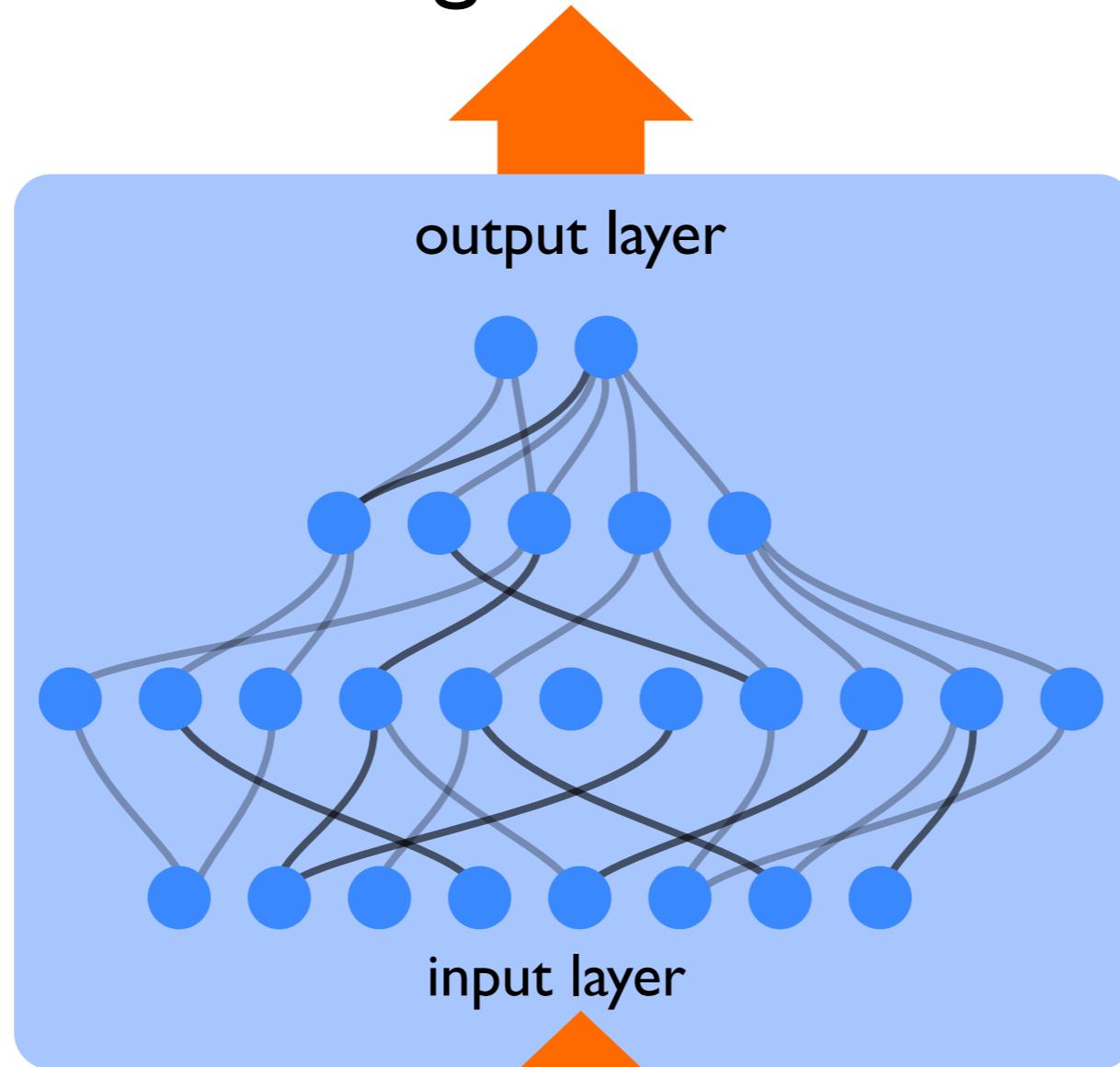


input layer



INPUT

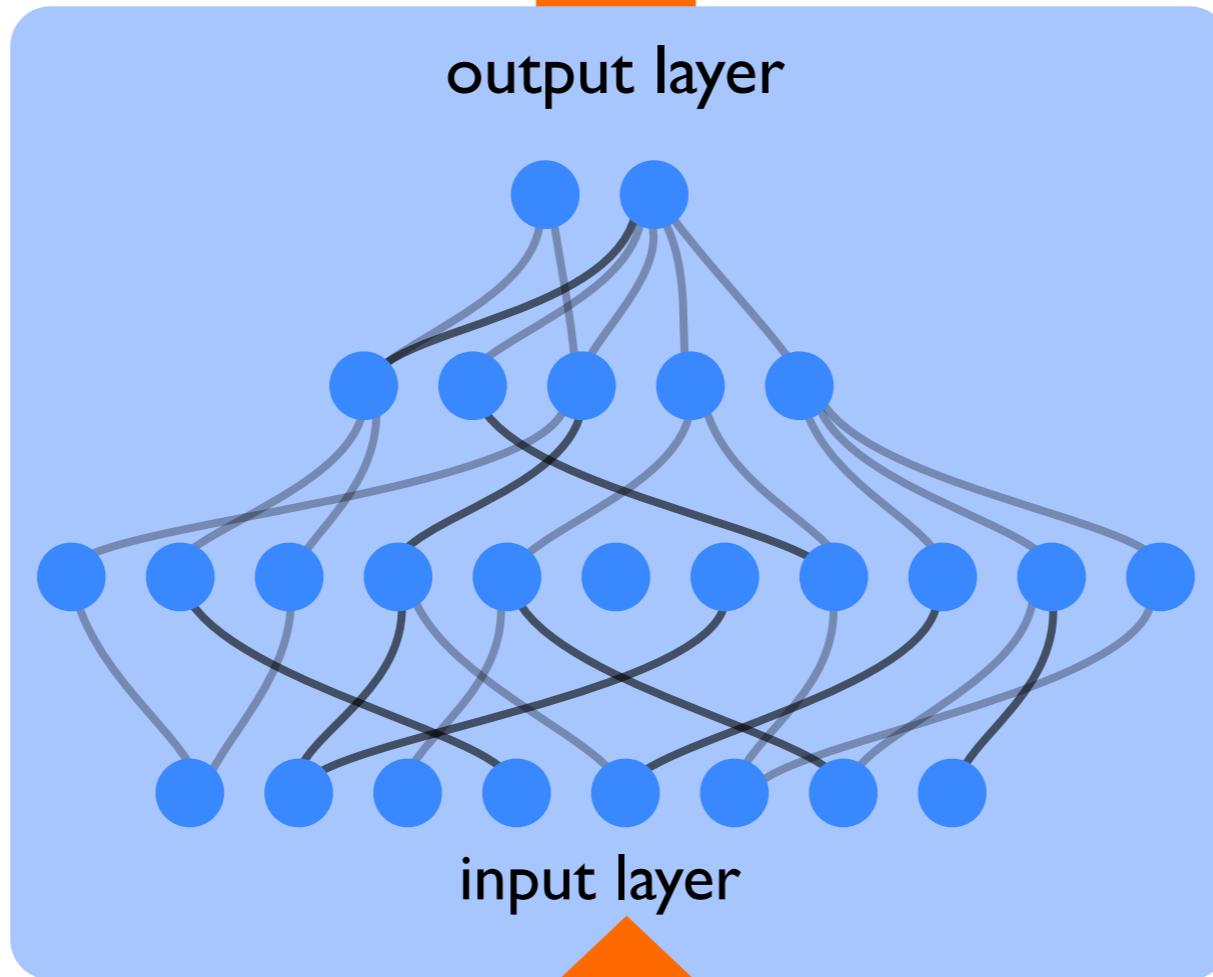
“light bulb”



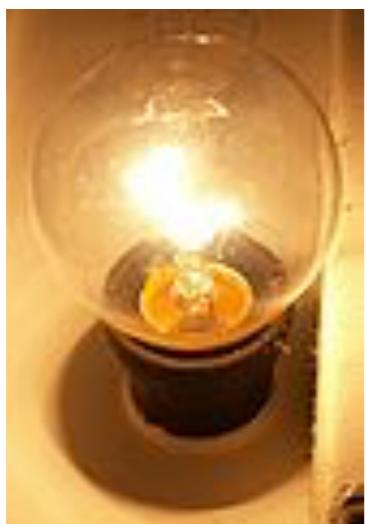
(this particular picture
has never been seen
before!)

(Picture:Wikimedia Commons)

“light bulb”

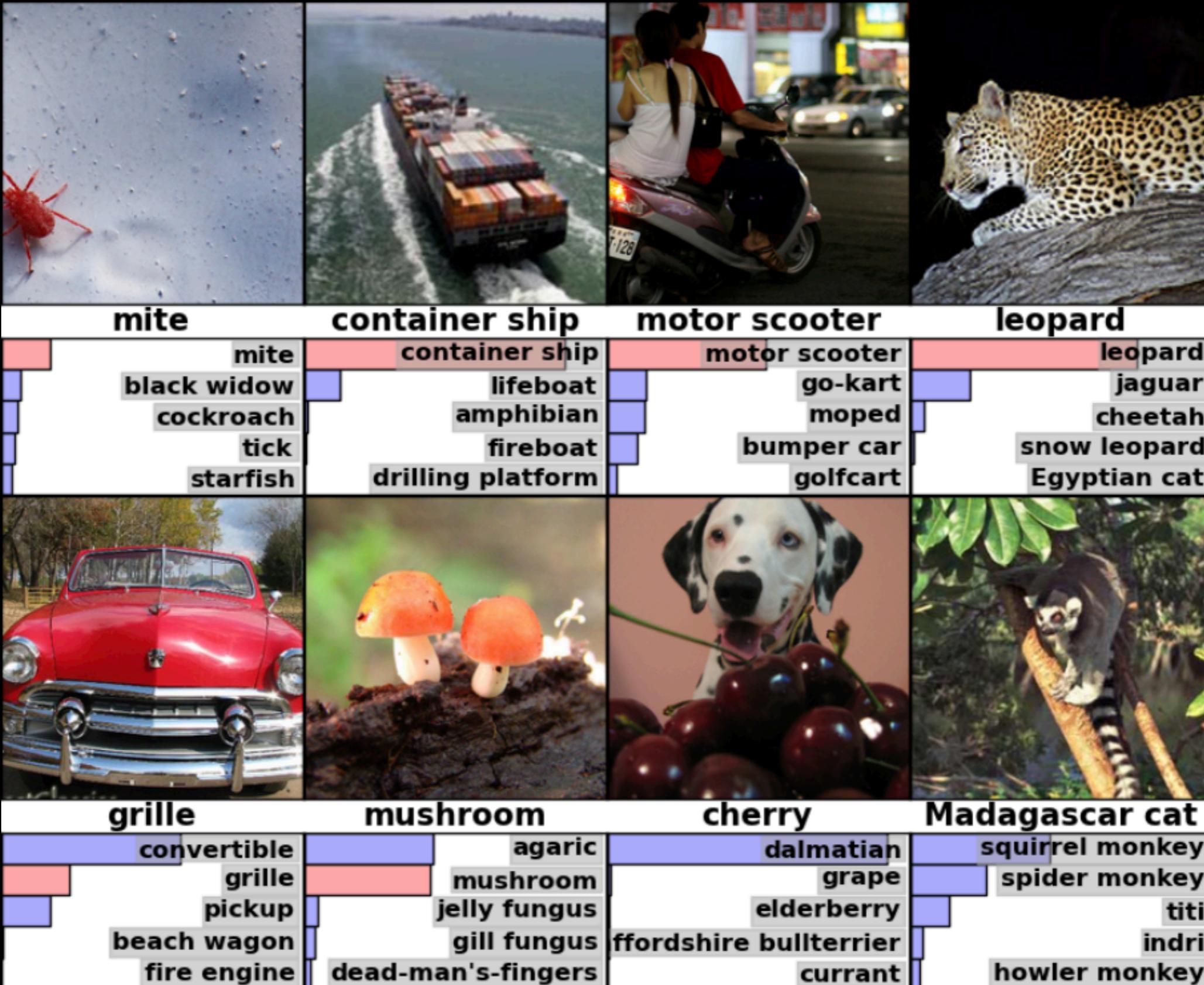


(training images)



(Picture:Wikimedia Commons)

Image classification: ImageNet Challenge (2012)



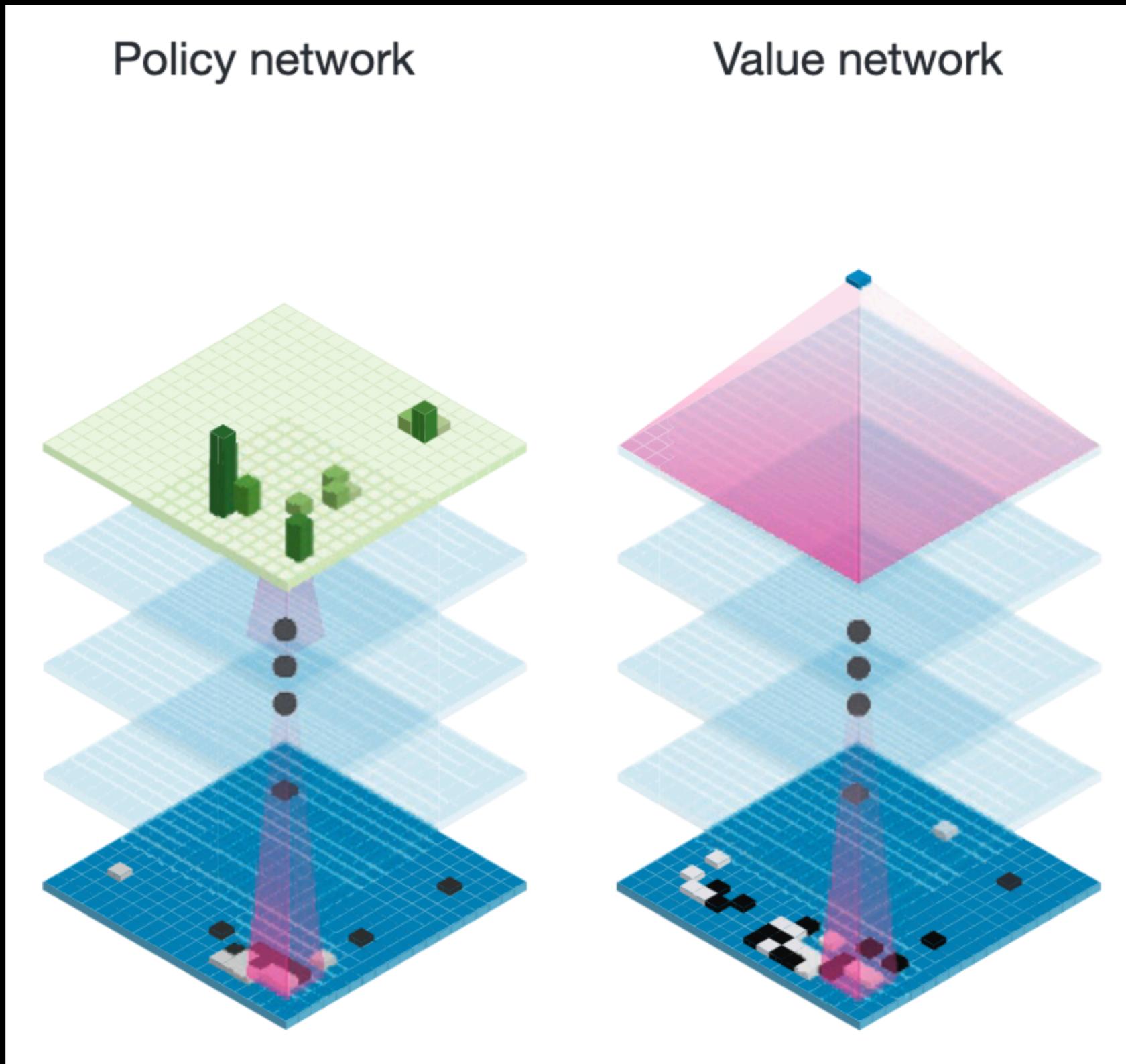


Google's AlphaGo Defeats Chinese Go Master in Win for A.I.

'Like A God,' Google A.I. Beats Human Champ Of Notoriously Complex Go Game

May 23, 2017 · 1:38 PM ET

Reinforcement Learning: AlphaGo (2016)



Silver, ..., Hassabis Nature 2016

Text Generation: GPT-3 (2020)

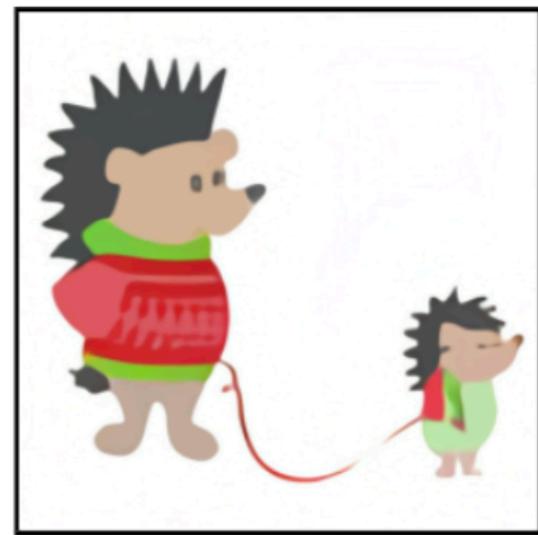
Title: Star's Tux Promise Draws Megyn Kelly's Sarcasm

Subtitle: Joaquin Phoenix pledged to not change for each awards event

Article: A year ago, Joaquin Phoenix made headlines when he appeared on the red carpet at the Golden Globes wearing a tuxedo with a paper bag over his head that read, "I am a shape-shifter. I can't change the world. I can only change myself." It was a promise to not change to fit into the Hollywood mold: "I think that's a really special thing, to not change yourself. I think it's a really special thing to say, `This is what's inside of me, I'm proud of it, and I'm not going to be ashamed because of the way that someone else thinks I should be.'" Now, it's the Oscars, and Phoenix is at it again. But this time, his publicist is saying he'll be wearing a tux no matter what.

Megyn Kelly was not impressed, and she let him have it on The Tonight Show. "You know, I feel like, I feel like you could have worn the tux," she says. "But you're saying you're a shape-shifter. I don't know if you can change your tux, but you can change your mind. You can change your mind. You can change your mind." Phoenix says he did, but it didn't stick. "I was like, `Okay, I'm going to wear a tuxedo to this thing.' And then I thought, `I don't want to wear a tuxedo to this thing.'" Kelly goes on to encourage him to change his mind again, but Phoenix says it's too late: "I'm committed to wearing this."

Generative Models: text-to-image (DALL-E, 2021)



(a) a tapir made of accordion.
a tapir with the texture of an
accordion.

(b) an illustration of a baby
hedgehog in a christmas
sweater walking a dog

About this lecture

"machine learning...":

...well, really mostly deep neural networks

"...for physicists":

- or other backgrounds (engineering, applied math, ...)
- rely on typical math level after 1st year physics undergraduate
- no deep programming background expected
- point out/use some connections to physics
- sometimes use physics examples in tutorials
- last few lectures: applications to science/physics

What you need to know already...

The chain rule of differentiation

Basic vector analysis: gradient

Basic linear algebra,
e.g. matrix-vector multiplication

Basics of programming
(in any programming language;
we will use python)

Contents

How do neural networks look like?
How do you train them efficiently?

How do you recognize images?
How do you learn a compact representation of
data, without providing any extra information?
How do you analyze time series and sentences?

How do you discover strategies from scratch
without a teacher?

Modern applications in science/physics

Some questions about your background...

Course website

<https://pad.gwdg.de/s/c3nYFbDzz#>

F CHANGED 3 HOURS AGO

FREELY ▾

Machine Learning for Physicists 2025

Lecture Series by Florian Marquardt

This is the website for the lecture series on machine learning (April 24-July 24, 2025). Here we will collect the course overview and links to the code etc.

This lecture series takes place every Thursday 16:00, at the Max Planck Institute for the Science of Light (Leuchs-Russell auditorium A1.500). An associated tutorial series will be scheduled in the first meeting on April 24.

Exam timing to be determined

DEEP LEARNING BASICS LINK MAP



BASIC STRUCTURE OF NEURAL NETWORKS

Lecture 1

Review 89 Hinton

EXPRESS ARBITRARY FUNCTIONS WITH 1 HIDDEN LAYER

89 Cybenko

COST FUNCTION AND STOCHASTIC GRADIENT DESCENT



Lecture 2

Review 89 Hinton

BACKPROPAGATION



76 Linnainmaa

85 Rumelhart et al

IMAGE RECOGNITION



Lecture 3

DISCRETE LABELS, SOFTMAX, CATEGORICAL CROSS-ENTROPY



0123456789

BASIC STRUCTURE OF NEURAL NETWORKS

Lecture 1

Review 89 Hinton

EXPRESS ARBITRARY FUNCTIONS WITH 1 HIDDEN LAYER

89 Cybenko

COST FUNCTION AND STOCHASTIC GRADIENT DESCENT

Lecture 2

Review 89 Hinton

BACKPROPAGATION



76 Linnainmaa

85 Rumelhart et al

CONVOLUTIONAL NEURAL NETWORKS



Lecture 4

80 Fukushima

89 LeCun et al

12 Krizhevsky et al

AUTOENCODERS

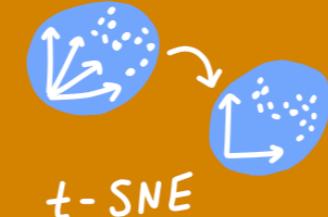
89 Baldi et al 06 Hinton et al

Review 12 Bengio et al



PRINCIPAL COMPONENT ANALYSIS

OTHER DIMENSIONALITY-REDUCING TECHNIQUES

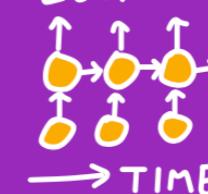


Lecture 5

08 van der Maaten et al

RECURRENT NEURAL NETWORKS

LSTM



95 Hochreiter et al

WORD VECTORS

13 Mikolov et al

REINFORCEMENT LEARNING



Lecture 6

Lecture 7

Lecture 8

92 Williams

92 Watkins et al

BOLTZMANN MACHINE



Lecture 9

85 Sejnowski

LECTURE NOTES

Lecture Notes: 20 Marquardt



BOOKS & TUTORIALS

Book: 16 Goodfellow et al

Book: 15 Nielsen



localhost

File Edit View Run Kernel Tabs Settings Help

2DFeatureExampl X LiquidDropModel- X 01_MachineLearni X

Code Python 3

$$y_j^{\text{out}} = f\left(\sum_k w_{jk} y_k^{\text{in}} + b_k\right)$$

where w is the weight matrix, b is the bias vector, and f would be the activation function (e.g. the sigmoid here), which is applied independently for each j .

[35]:

```
N0=3 # input layer size
N1=2 # output layer size

# initialize random weights: array dimensions
w=random.uniform(low=-1,high=+1,size=(N1,N0))
# initialize random biases: N1 vector
b=random.uniform(low=-1,high=+1,size=N1)
```

[17]:

```
# input values
y_in=array([0.2,0.4,-0.1])
```

Output View

P(y) for relax 20, 1E4 samples, nbin 200

Output View

P(y) for relax 20, 1E4 samples, nbin 60

Output View

relax steps = 20; 10000 samples

Output View

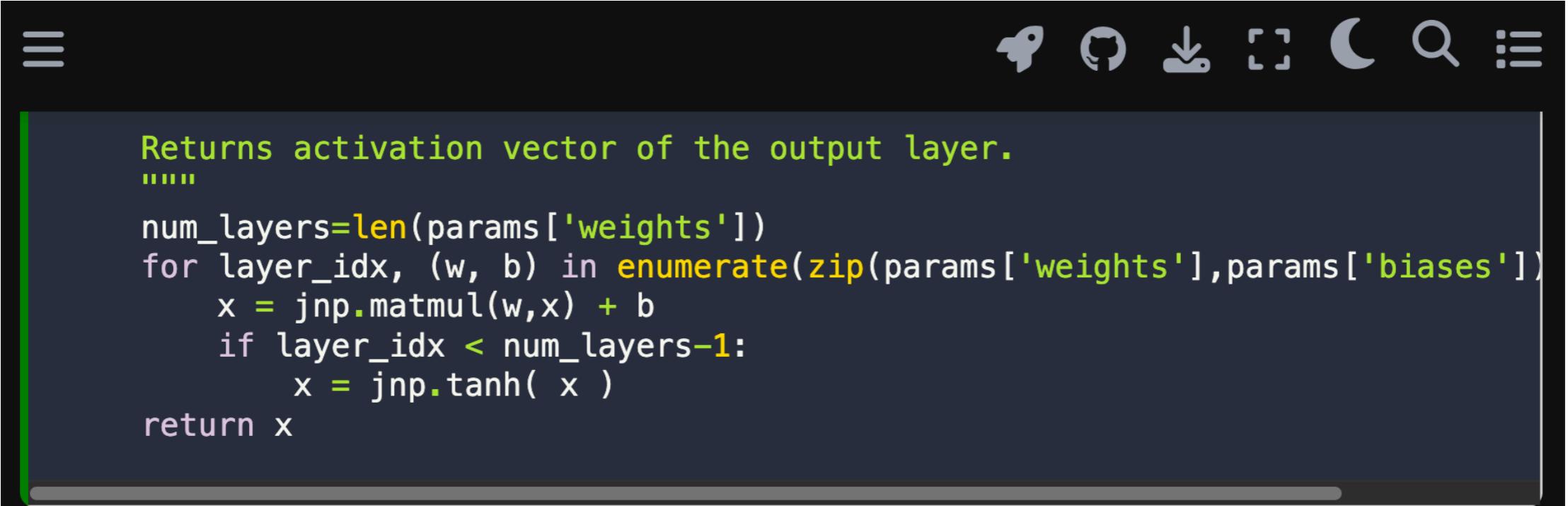
relax steps = 20; 1000 samples

0 3 5 Python 3 | Mode: Ln 4. Col 01 MachineLearning Basics NeuralNetworksPython.ipynb



A few lines of code to set up and
train a neural network

[https://florianmarquardt.github.io/
MachineLearningThreeEasyLessons/intro.html](https://florianmarquardt.github.io/MachineLearningThreeEasyLessons/intro.html)



```
Returns activation vector of the output layer.  
"""  
    num_layers=len(params['weights'])  
    for layer_idx, (w, b) in enumerate(zip(params['weights'],params['biases'])):  
        x = jnp.matmul(w,x) + b  
        if layer_idx < num_layers-1:  
            x = jnp.tanh( x )  
    return x
```

It has been shown already in the 1980s that a network of this form can approximate any smooth high-dimensional function to arbitrary precision with sufficiently many neurons, even with a single hidden layer.

🔔 Outlook

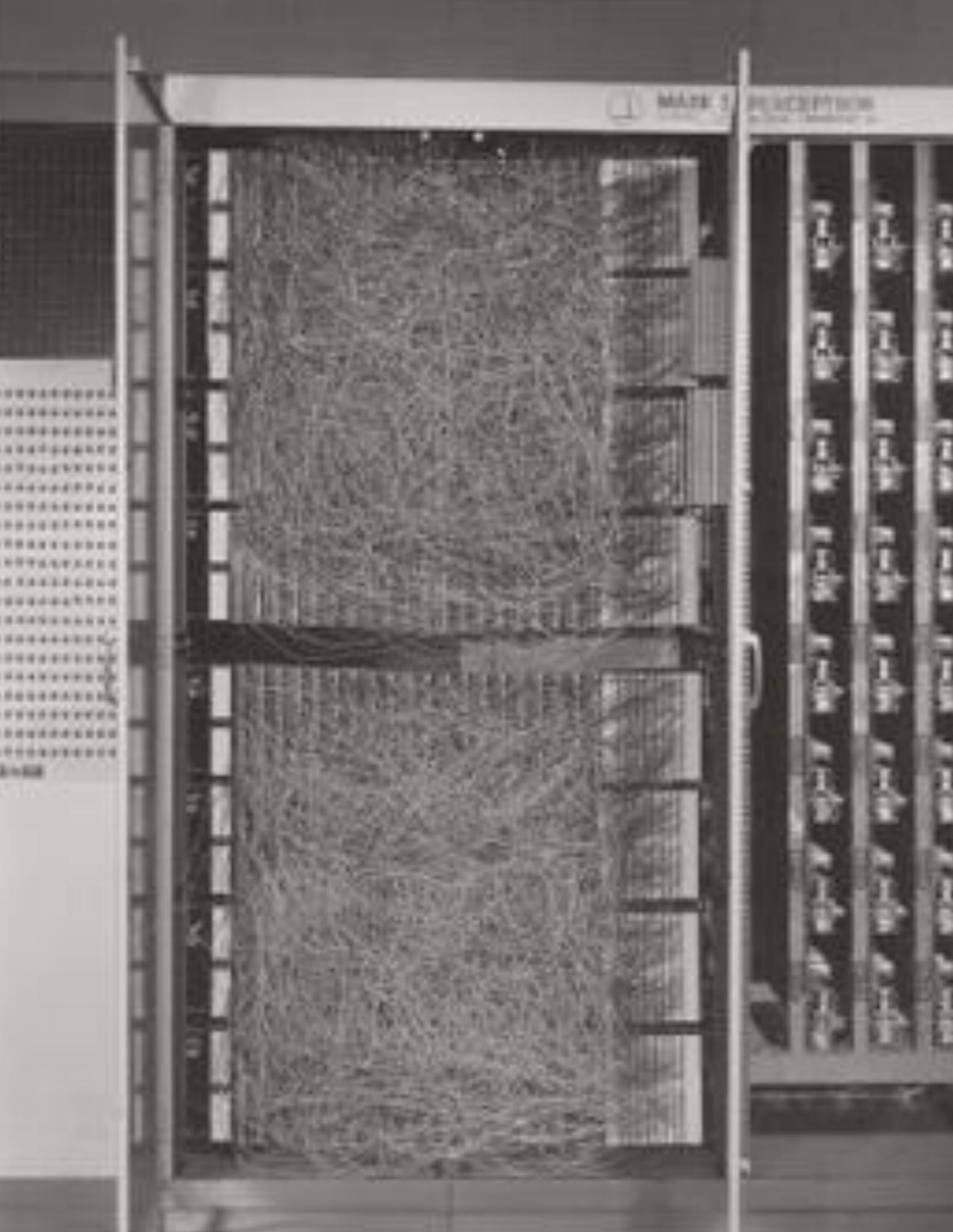
This neural network structure ("fully connected", "multi-layer perceptron") covers a wide variety of applications and is the 'default'. In principle, one can

[↑ Back to top](#)

Tutorials: 2h / week

Select time slot!

Homework problems: will suggest some homework to try (typically not be discussed in-depth in tutorials)



How it all
started...

...in the
1950s



I.—COMPUTING MACHINERY AND INTELLIGENCE

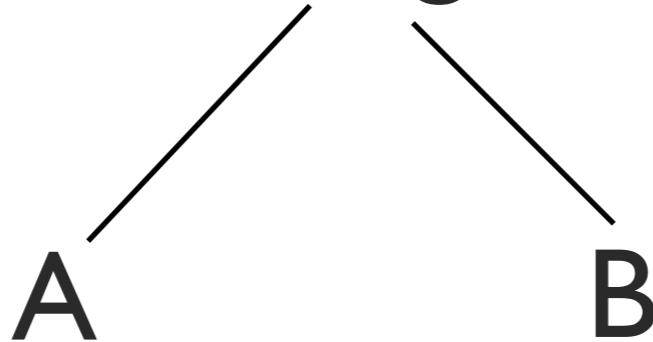
BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, ‘Can machines think?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’. The definitions might be framed so as to reflect so far as

Alan Turing, "Computing Machinery and Intelligence" (Mind, 1950):

interrogator



...

We now ask the question, ‘What will happen when a machine takes the part of A in this game?’ Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’

I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning.

Alan Turing, "Computing Machinery and Intelligence" (1950):

F

Alice puts a gold coin inside a wallet. She puts the wallet into a suitcase and hands the suitcase to Bob. Bob delivers the suitcase to Charly. Who has the gold coin now?



Based on the information given, it is most likely that Charly has the suitcase that contains the wallet with the gold coin. However, we cannot be completely certain without additional information. It is possible that Bob could have taken the gold coin out of the wallet before delivering the suitcase to Charly.



F

Good. Let's say Charly now has the gold coin. He wants to use it to buy an ice cream for Alice. Unfortunately, the ice cream vendor does not accept gold coins. What can Charly do?



If the ice cream vendor does not accept gold coins as a form of payment, Charly will need to exchange the gold coin for a more widely accepted form of currency, such as cash or a credit card. There are a few options he could consider:



1. **Find a local coin dealer:** Charly could look for a coin dealer in the area who may be interested in buying the gold coin from him. The dealer may be able to offer a fair price for the coin, which Charly could then use to purchase the ice cream with cash.
2. **Use a gold exchange service:** There are various online and offline gold exchange services that can provide a fair price for gold coins. Charly could use one of these services to

F

Are you a human?



No, I am not a human. I am an artificial intelligence language model created by OpenAI, designed to understand and respond to natural language queries and tasks. My responses are generated based on patterns and associations learned from analyzing vast amounts of text data, rather than through human thought processes.



We also wish to allow the possibility that an engineer or team of engineers may construct a machine which works, but whose manner of operation cannot be satisfactorily described by its constructors because they have applied a method which is largely experimental.

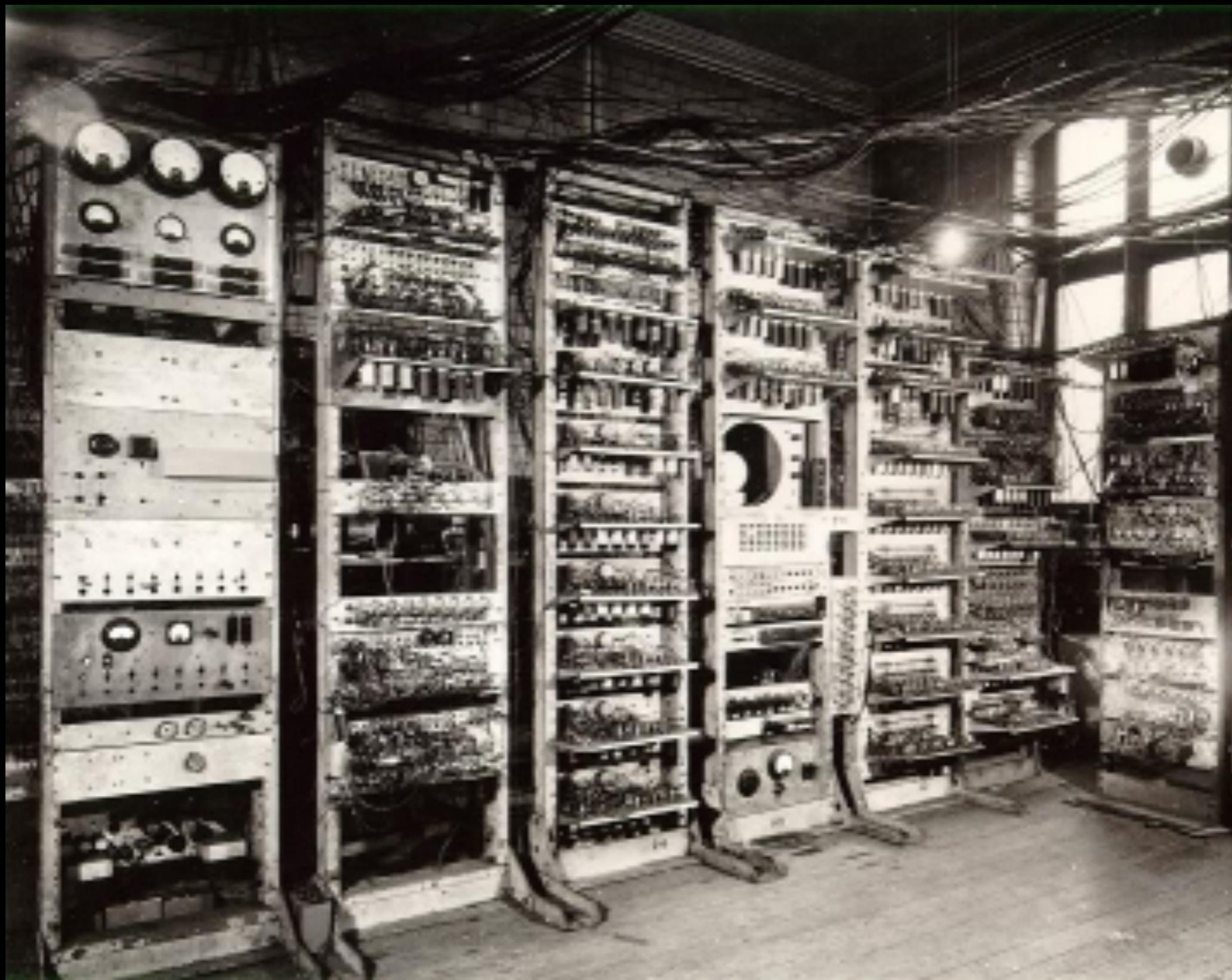
interpreting neural networks!

Alan Turing, "Computing Machinery and Intelligence" (1950)

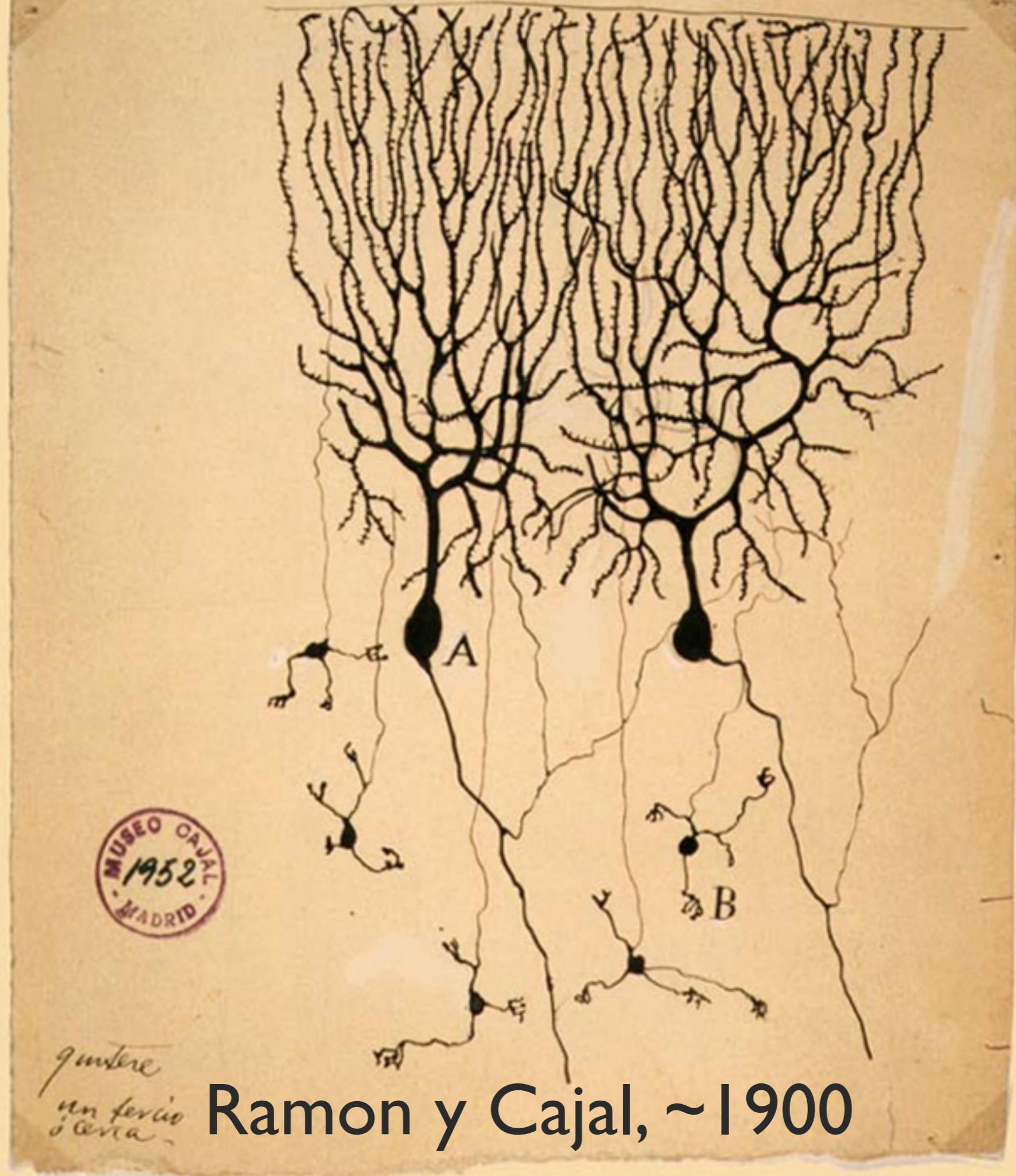
The machine has to be so constructed that events which shortly preceded the occurrence of a punishment-signal are unlikely to be repeated, whereas a reward-signal increased the probability of repetition of the events which led up to it.

"reinforcement learning"

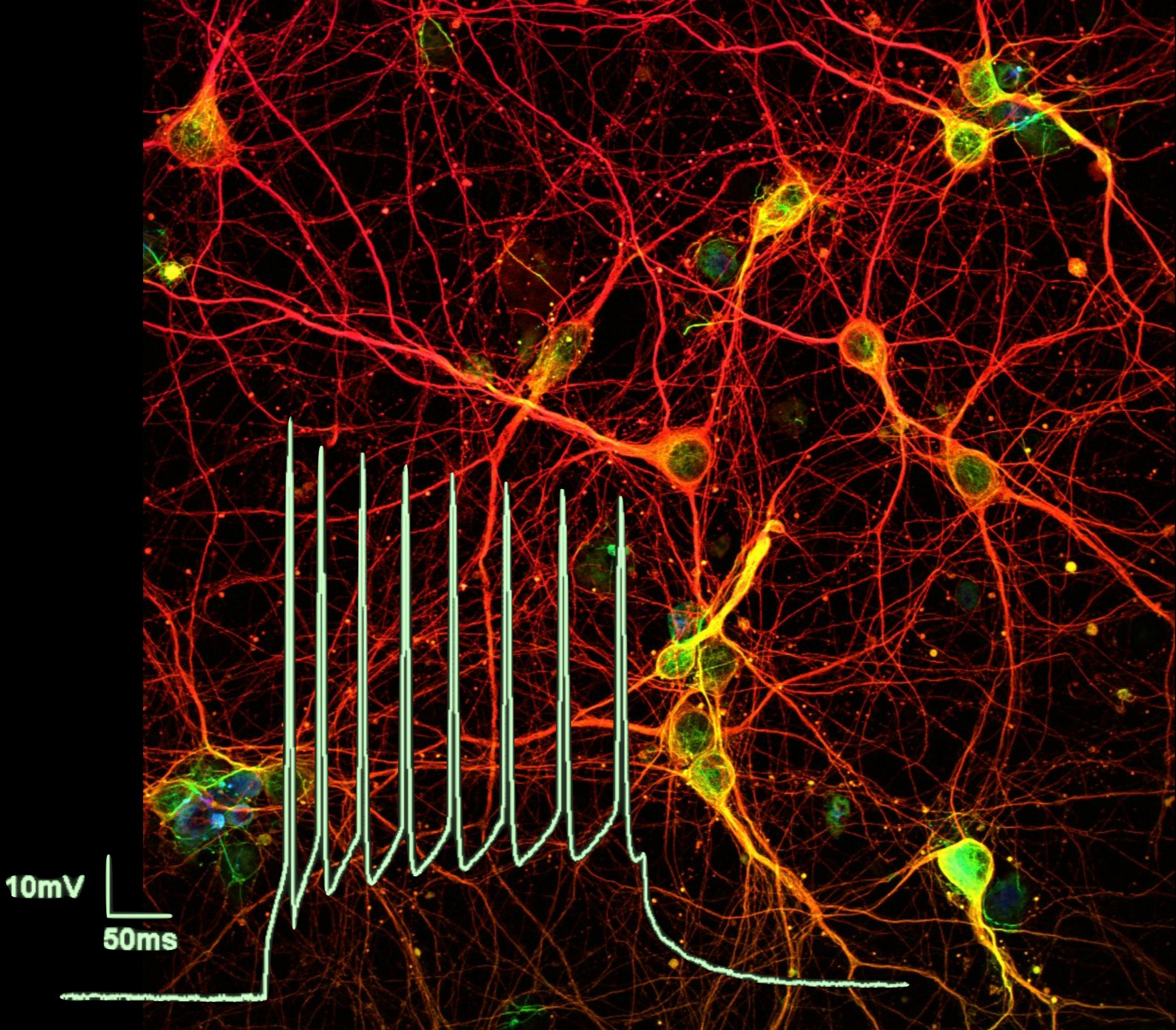
Alan Turing, "Computing Machinery and Intelligence" (1950)



"Manchester Mark I" (image:Wikipedia)



Ramon y Cajal, ~1900



neurons:Wikimedia Commons/ALol88; trace:Wikimedia



CORNELL AERONAUTICAL LABORATORY, INC.

BUFFALO, N. Y.

REPORT NO. 85-460-1

THE PERCEPTRON

A PERCEIVING AND RECOGNIZING AUTOMATON

(PROJECT PARA)

January, 1957

Prepared by: Frank Rosenblatt

Frank Rosenblatt,
Project Engineer

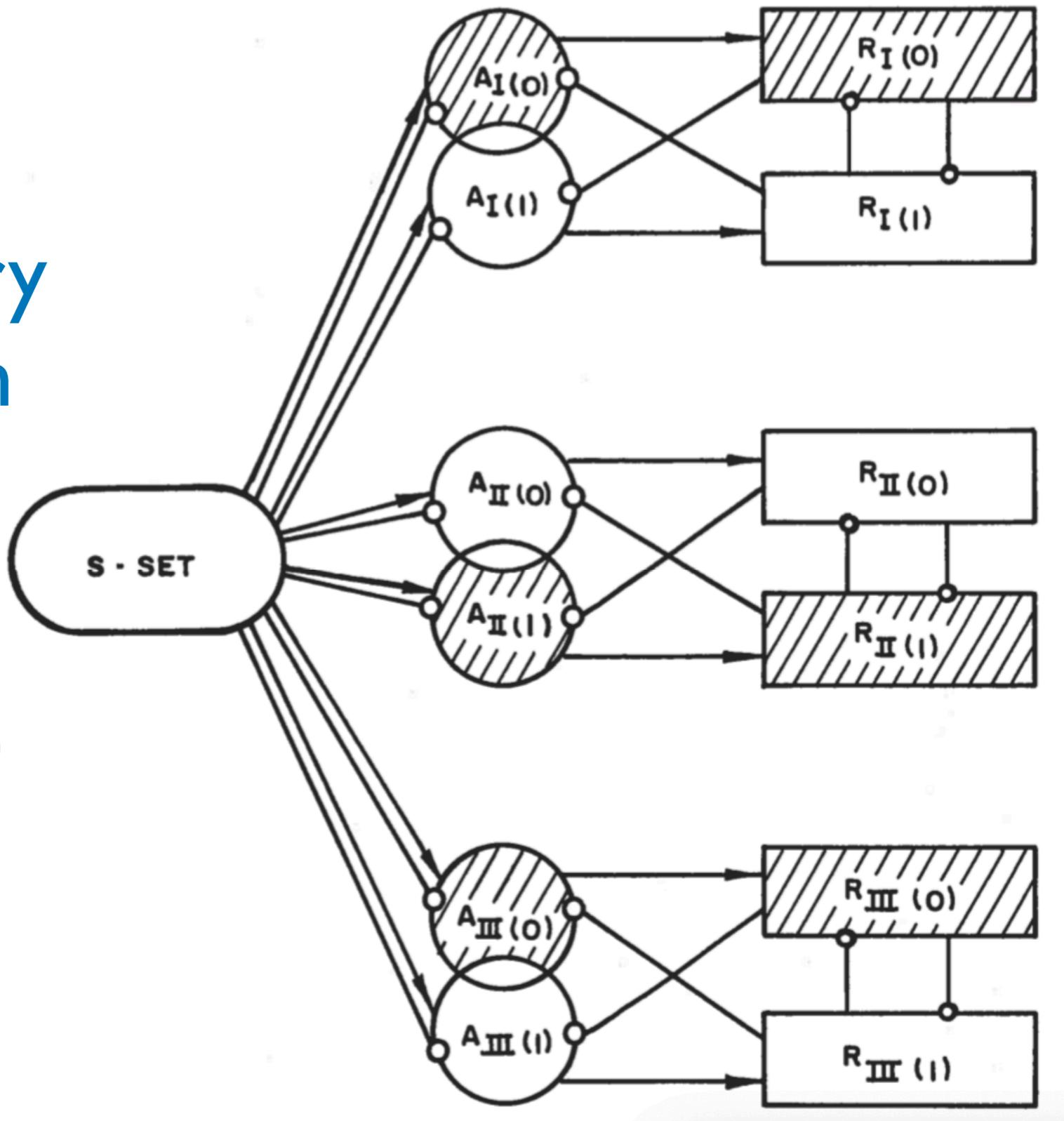
A primary requirement of such a system is that it must be able to recognize complex patterns of information which are phenomenally similar, or are experientially related -- a process which corresponds to the psychological phenomena of "association" and "stimulus generalization". The system must recognize the "same" object in different orientations, sizes, colors, or transformations, and against a variety of different backgrounds. The recognition of "similar" forms can be carried out, to a certain extent, by analytic procedures on a digital or analog computer, but it is hard to conceive of a general analytic program which would, for example, recognize the form of a man seen from any angle, and in any posture or position, without actually storing a large library of reference figures against which the percept could be compared. In general, identities of this sort must be learned, or acquired from experience, and if the system is

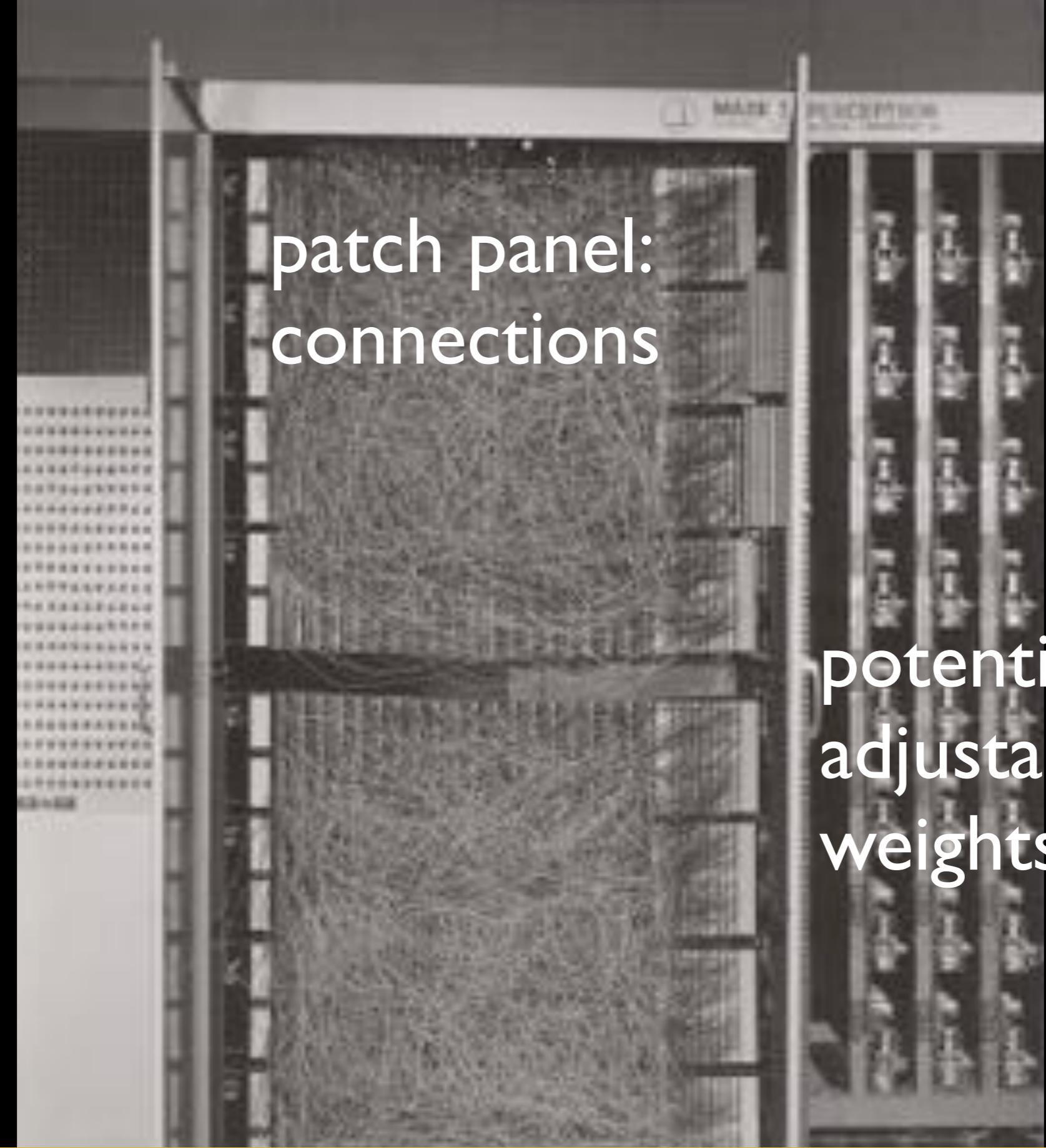
Association System

Sensory System

(input)

Response System
(output)





"Mark I Perceptron" (Rosenblatt 1960)

NEW NAVY DEVICE LEARNS BY DOING

**Psychologist Shows Embryo
of Computer Designed to
Read and Grow Wiser**

New York Times, July 8, 1958

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 [...] said Perceptrons might be fired to the planets as mechanical space explorers.

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.

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. 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.

