CAP 6635 – Artificial Intelligence

Lecture 11: How do machines learn? (Part 3)

Oge Marques, PhD

Professor

College of Engineering and Computer Science

College of Business

FAU FLORIDA ATLANTIC UNIVERSITY

@ProfessorOge

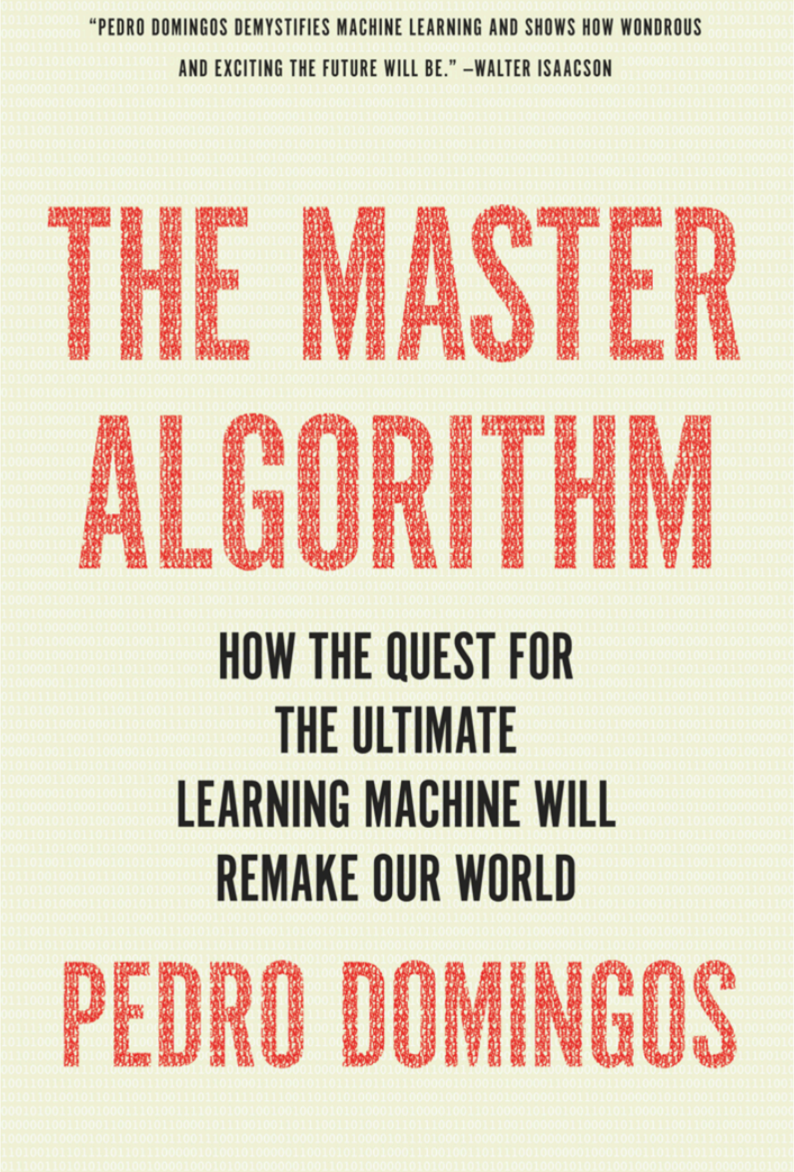
ProfessorOgeMarques

Previously on CAP 6635...

The Master Algorithm

(by Pedro Domingos)

"All knowledge—past, present, and future—can be derived from data by a single, universal learning algorithm."



The symbolists

Sidebar: David Marr (1945-1980)

- One of the most influential neuroscientists of vision.
- Thought of vision as an information-processing task.
- In his book *Vision* (1982), he distinguished three different levels of description involved in understanding complex information processing systems:
  - Computational level
  - Algorithmic level
  - Implementation level
    - An important point is that the levels can be considered independently.

Source: Palmer, S. *Vision Science*.

Marr’s computational framework

Computational theory
What is the nature of the problem to solved, what is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?

Representation and algorithm
How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation from input to output?

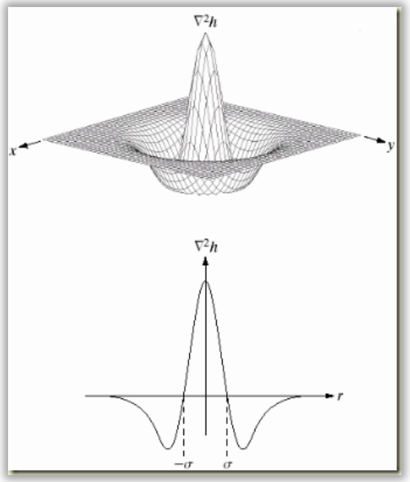
Hardware implementation
How can the representation and algorithm be realised physically?

Source: Frisby and Stone. *Seeing, 2nd Edition*.

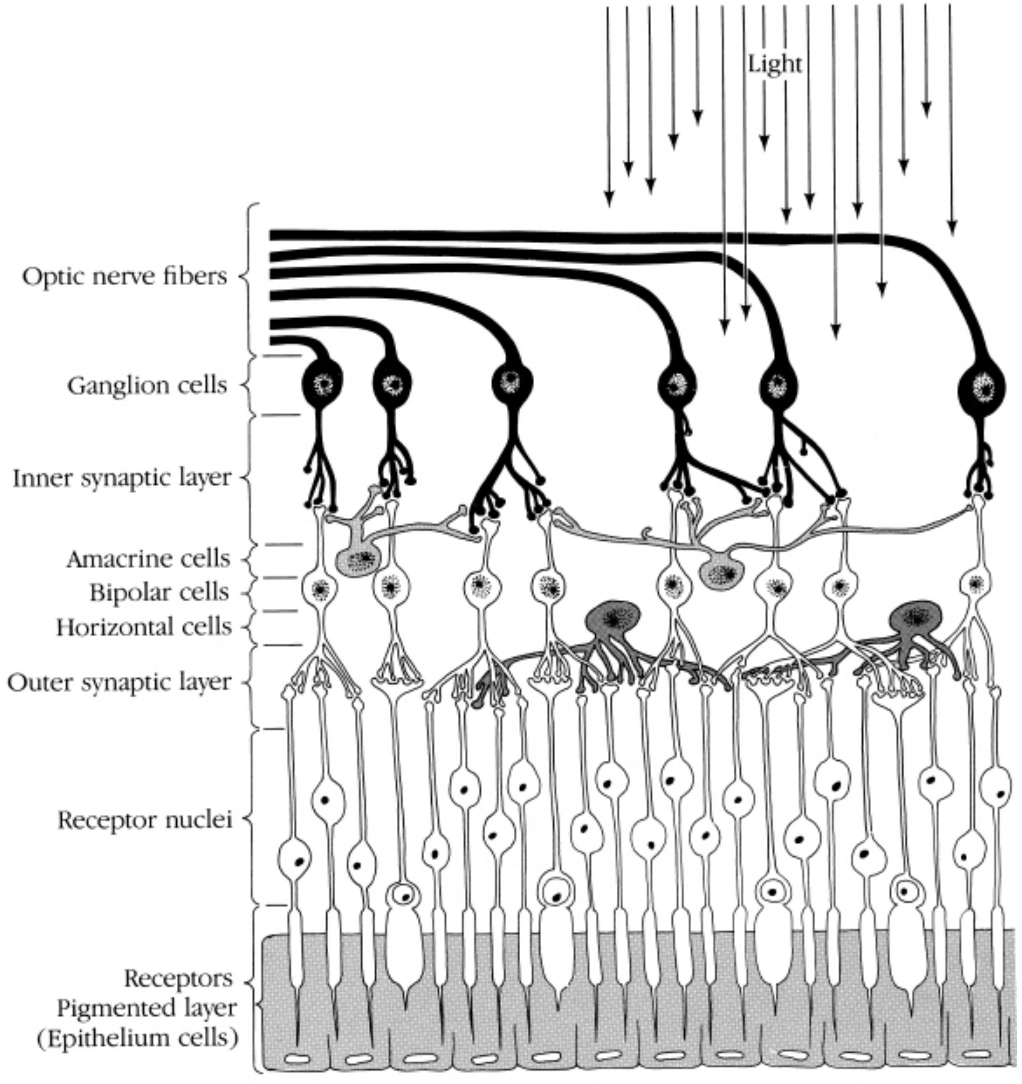
Marr’s impact

- Edge detection

$$\nabla^2 G * I(x,y),$$
$$\text{where } \nabla^2 G(r) = -\frac{1}{\pi\sigma^4}\left(1-\frac{r^2}{2\sigma^2}\right)\exp\left(\frac{-r^2}{2\sigma^2}\right)$$



Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, *207*(1167), 187-217.

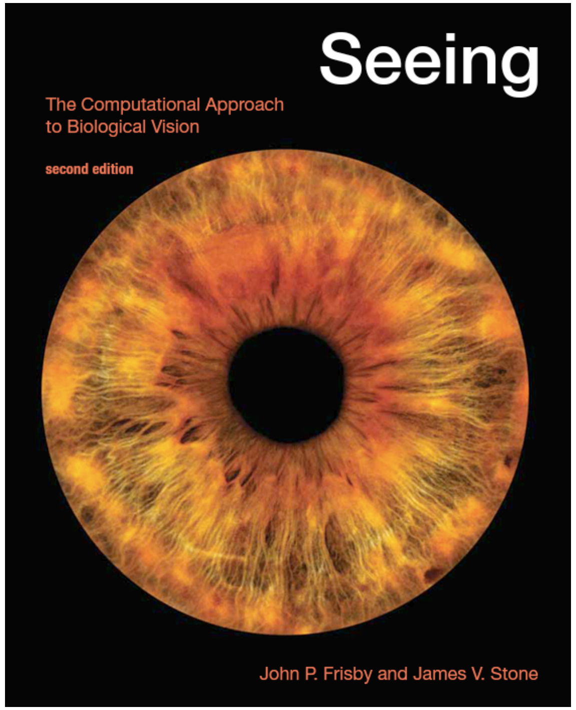


Source: Marr, D. *Vision*.

New material starts here…

Marr's impact

- Frisby and Stone's textbook is entirely grounded on Marr's computational framework



Marr’s book (Epilogue)

- Objections and clarifications
  - An excerpt (highlighted by Frisby and Stone):
    - “[…] mechanism-based approaches are genuinely dangerous. The problem is that the goal of such studies is mimicry rather than true understanding, and these studies can easily degenerate into the writing of programs that do no more than mimic in an unenlightening way some small aspect of human performance.
    - If we believe that the aim of information-processing studies is to formulate and understand particular information-processing problems, then the structure of those problems is central, not the mechanisms through which their solutions are implemented.”

Source: Marr, D. *Vision*.

Symbolic vs. subsymbolic AI

- Symbolic AI was originally inspired by mathematical logic as well as by the way people described their conscious thought processes.
- In contrast, subsymbolic approaches to AI took inspiration from neuroscience and sought to capture the sometimes-unconscious thought processes underlying what some have called *fast perception*, such as recognizing faces or identifying spoken words.
    - Subsymbolic AI programs do not contain the kind of human-understandable language we often see in symbolic solutions.
    - Instead, they are essentially a stack of equations designed to learn from data how to perform a task.

From: Melanie Mitchell. “Artificial Intelligence.”

The connectionists

Connectionists

Yann LeCun

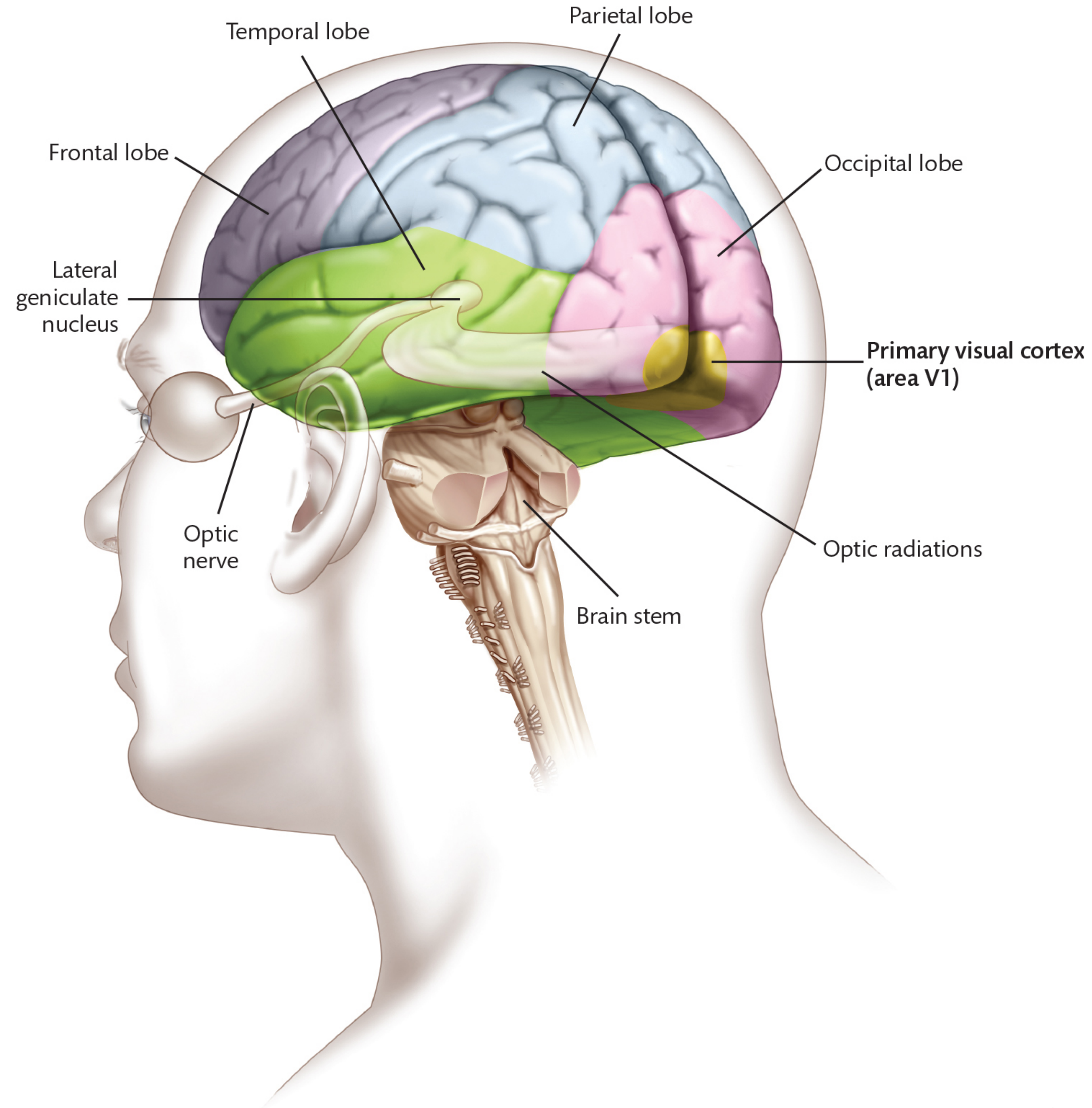
Geoff Hinton

Yoshua Bengio

Source: P. Domingos

Our source of inspiration

Our source of inspiration



Fundamental questions

What do we know about the (human) brain?

How much do we know about the (human) brain?

How do we learn about the (human) brain?

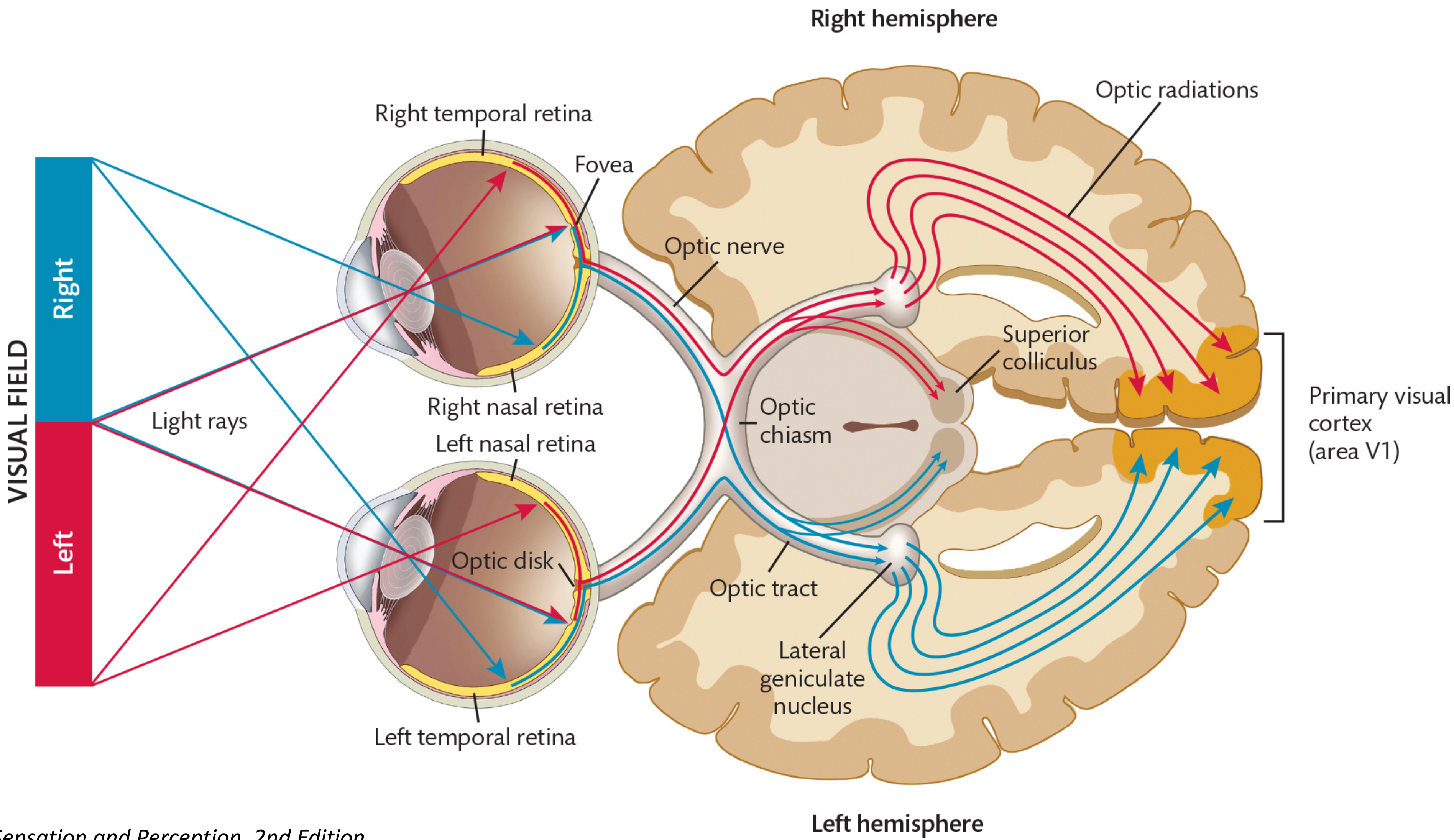
What about **other animals' brains**?

What can we achieve by **reverse engineering** the brain?

What do we know about the (human) brain?

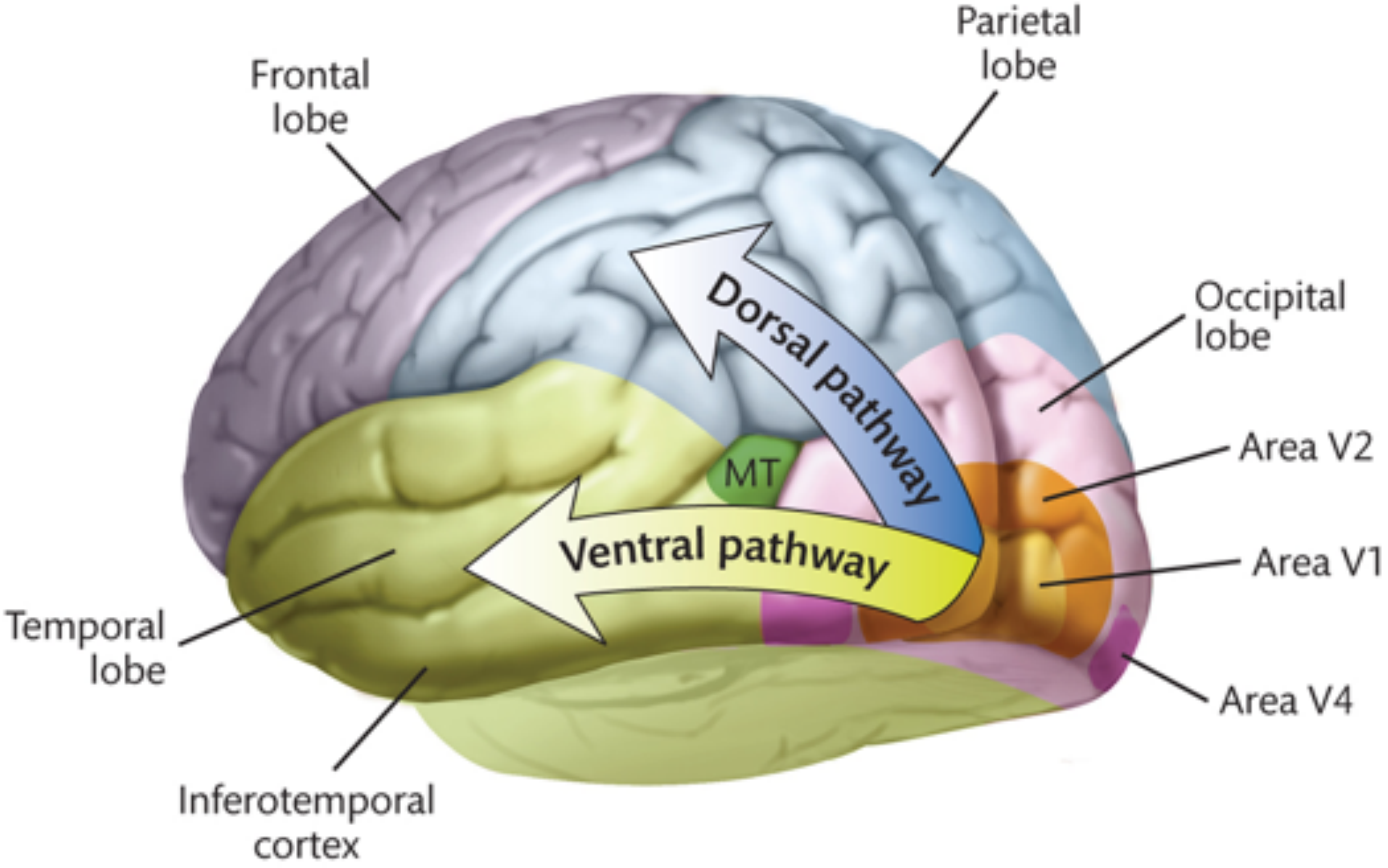
- Some mappings between cause and effect (and the underlying physical / chemical / electrical processes)
- Some processes
- Some pathways
- Some specialized regions
- Some "general properties"
- …

Main pathways from retina to the brain



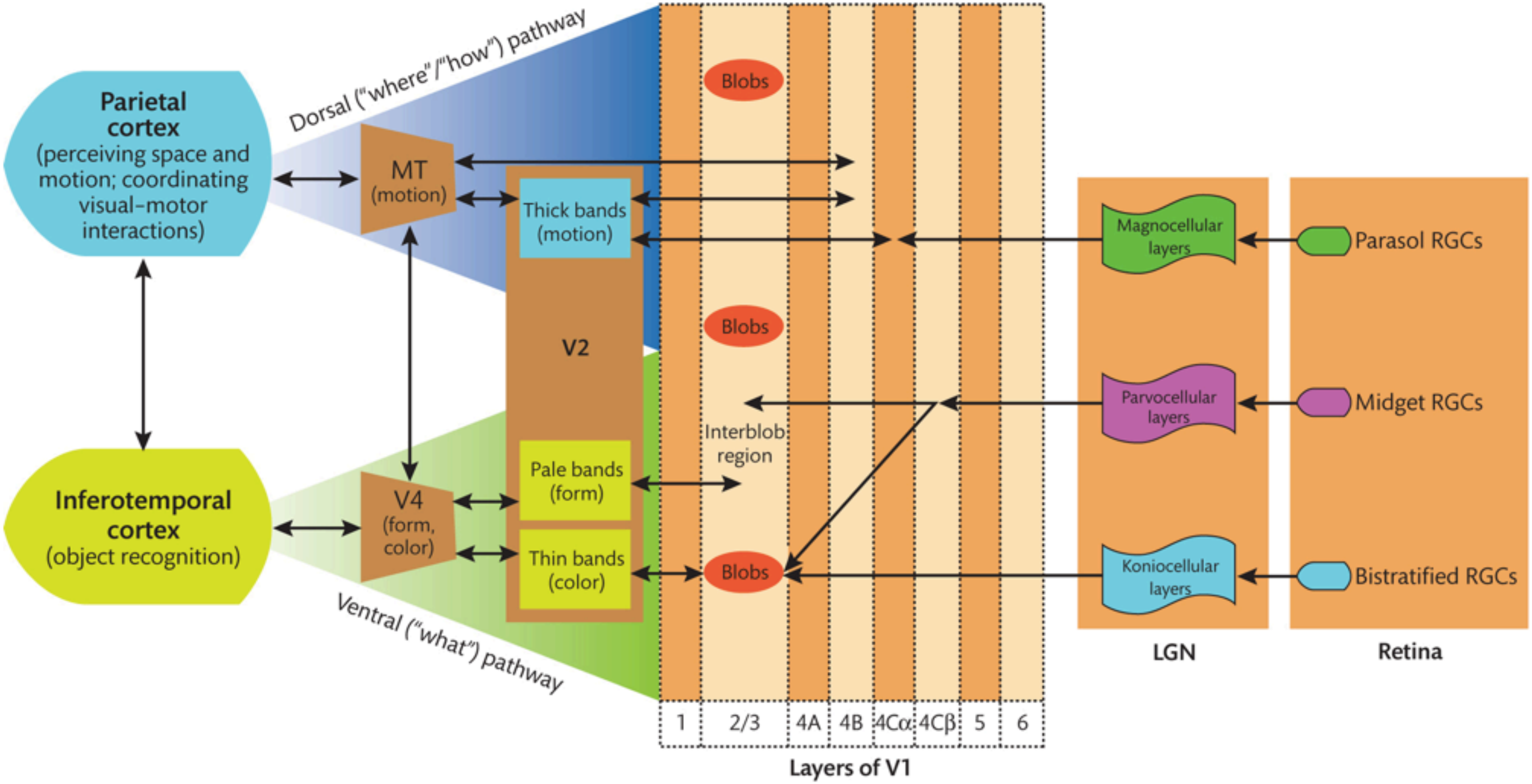
Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

Functional areas and pathways



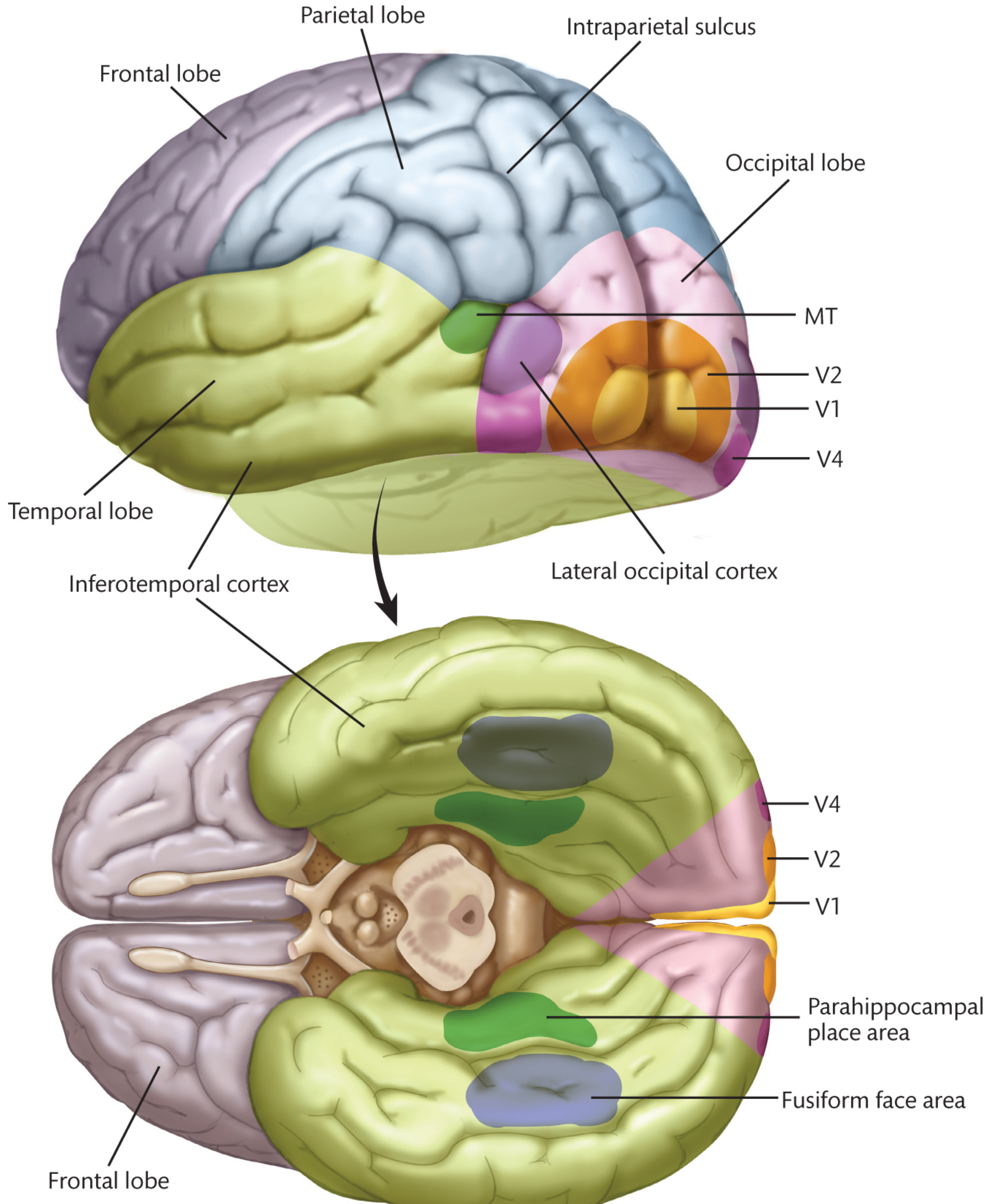
Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

Functional areas and pathways



Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

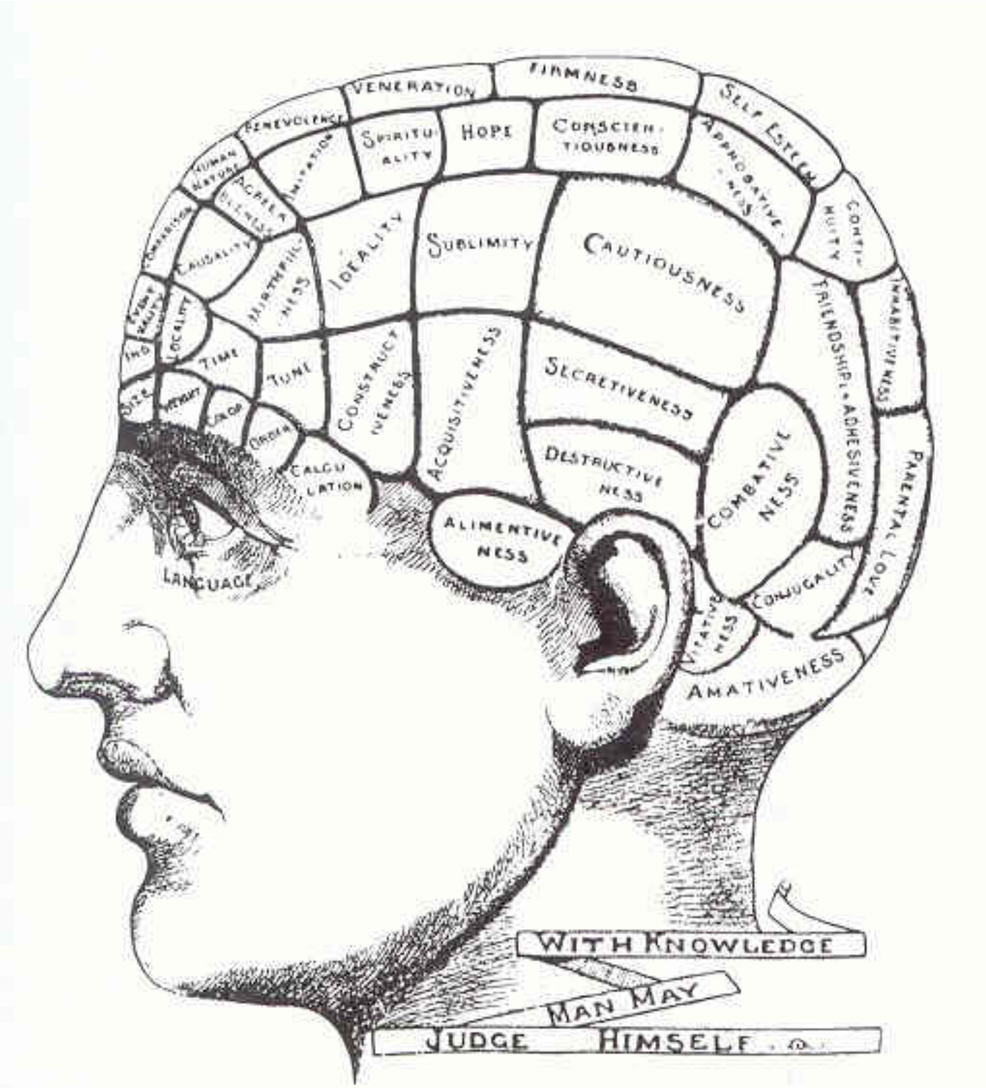
Functional areas and pathways



Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

Caution

- Don’t take functional areas too far…
  - Plasticity
  - Phrenology ☺



Source: Palmer, S. *Vision Science*.

Fundamental questions

- **What do we know** about the (human) brain?
- **How much do we know** about the (human) brain?
- **How do we learn** about the (human) brain?
- What about **other animals' brains**?
- What can we achieve by **reverse engineering** the brain?

How much do we know about the (human) brain?

- Ask a (neuro)scientist (or several)!

“We know very little about the brain. We know about connections, but we don’t know how information is processed."

Stanford MEDICINE

Lu Chen

PROFESSOR OF NEUROSURGERY AND OF PSYCHIATRY AND BEHAVIORAL SCIENCES

How much do we know about the (human) brain?

- Ask a Nobel winner neuroscientist!

"Mapping the components of the brain is far more complex than mapping the human genome."

Stanford MEDICINE

Thomas Sudhof

AVRAM GOLDSTEIN PROFESSOR IN THE SCHOOL OF MEDICINE AND PROFESSOR, BY COURTESY, OF NEUROLOGY AND OF PSYCHIATRY AND BEHAVIORAL SCIENCES

Molecular & Cellular Physiology

Fundamental questions

- **What do we know** about the (human) brain?
- **How much do we know** about the (human) brain?
- **How do we learn** about the (human) brain?
- What about **other animals' brains**?
- What can we achieve by **reverse engineering** the brain?

How scientists learn about human vision

- Patients with brain damage or eye conditions
- Direct access to the brain
  - Single-cell recording (*in animals*)
  - Modern brain imaging and activity recording devices (fMRI)
- Controlled experiments
  - Calibrated monitors and rooms
  - Eye-tracking devices
  - Psychophysics

Fundamental questions

What do we know about the (human) brain?

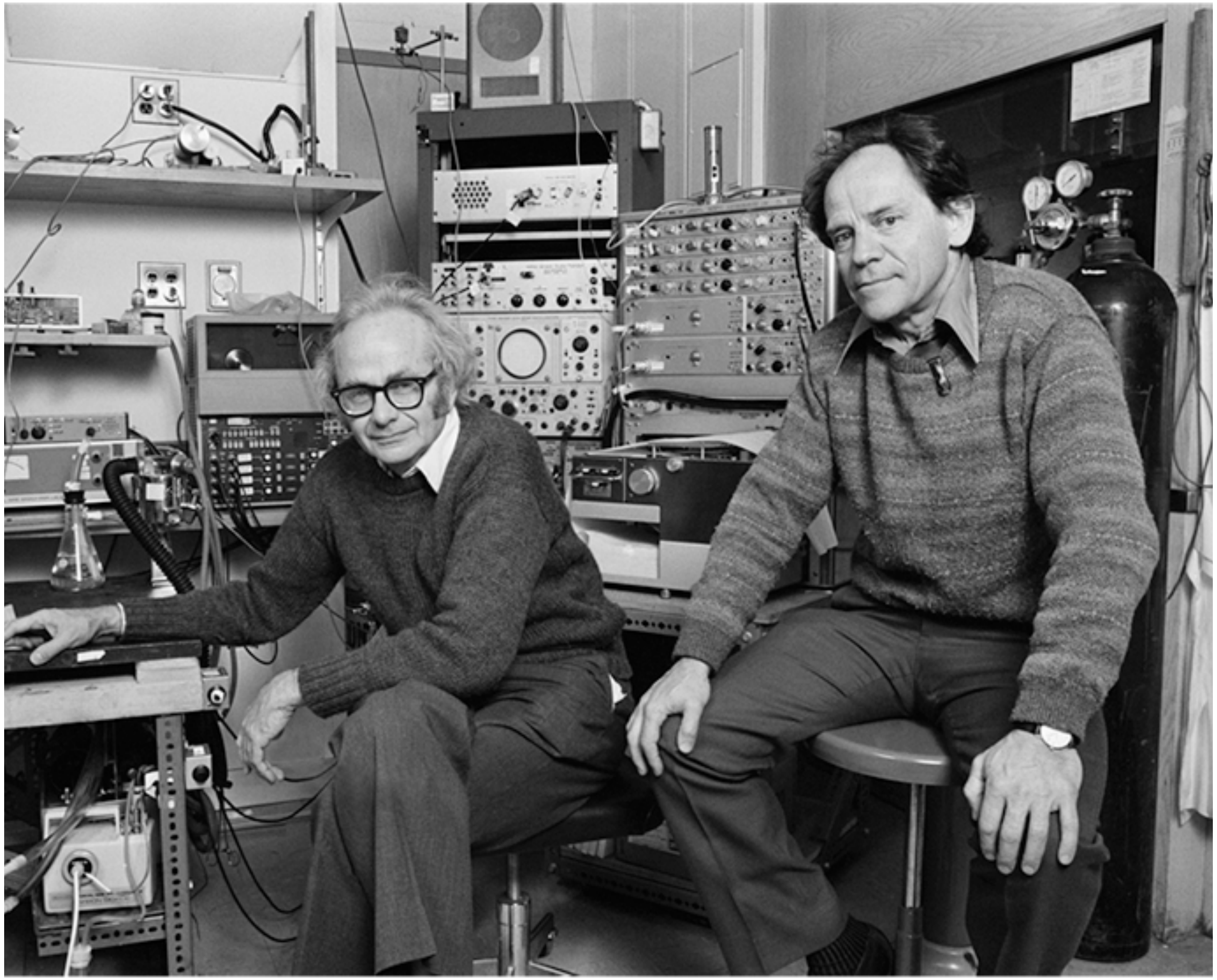
How much do we know about the (human) brain?

How do we learn about the (human) brain?

What about **other animals' brains**?

What can we achieve by **reverse engineering** the brain?

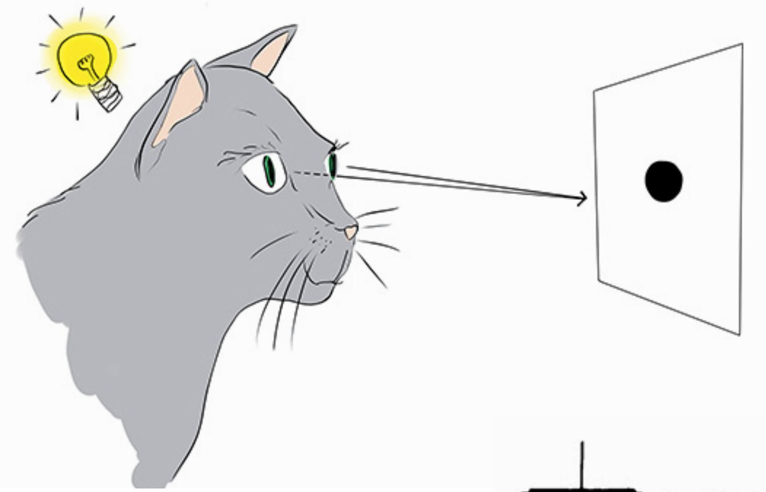
Primary Visual Cortex (Area V1)

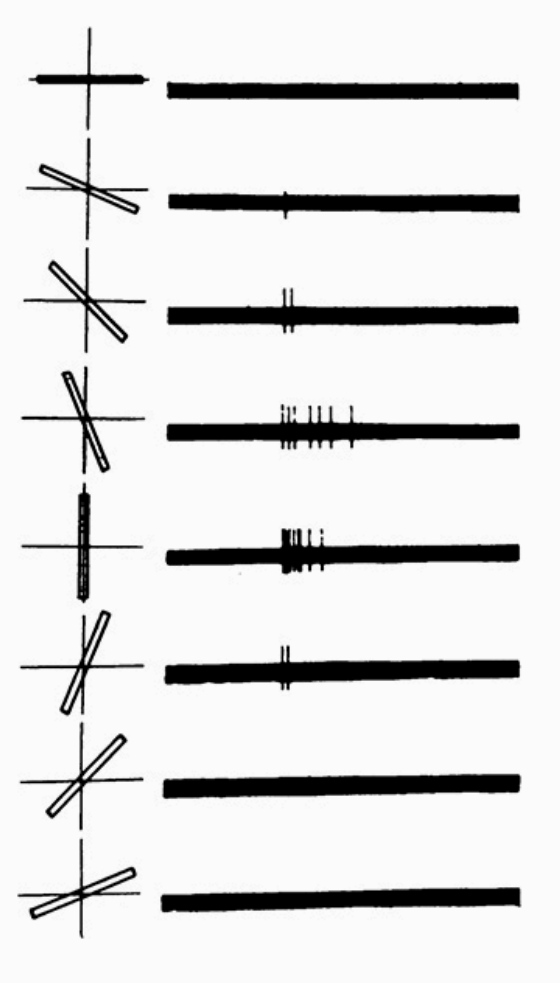


(left) David H. Hubel (1926–2013) and (right) Torsten Wiesel (b. 1924)

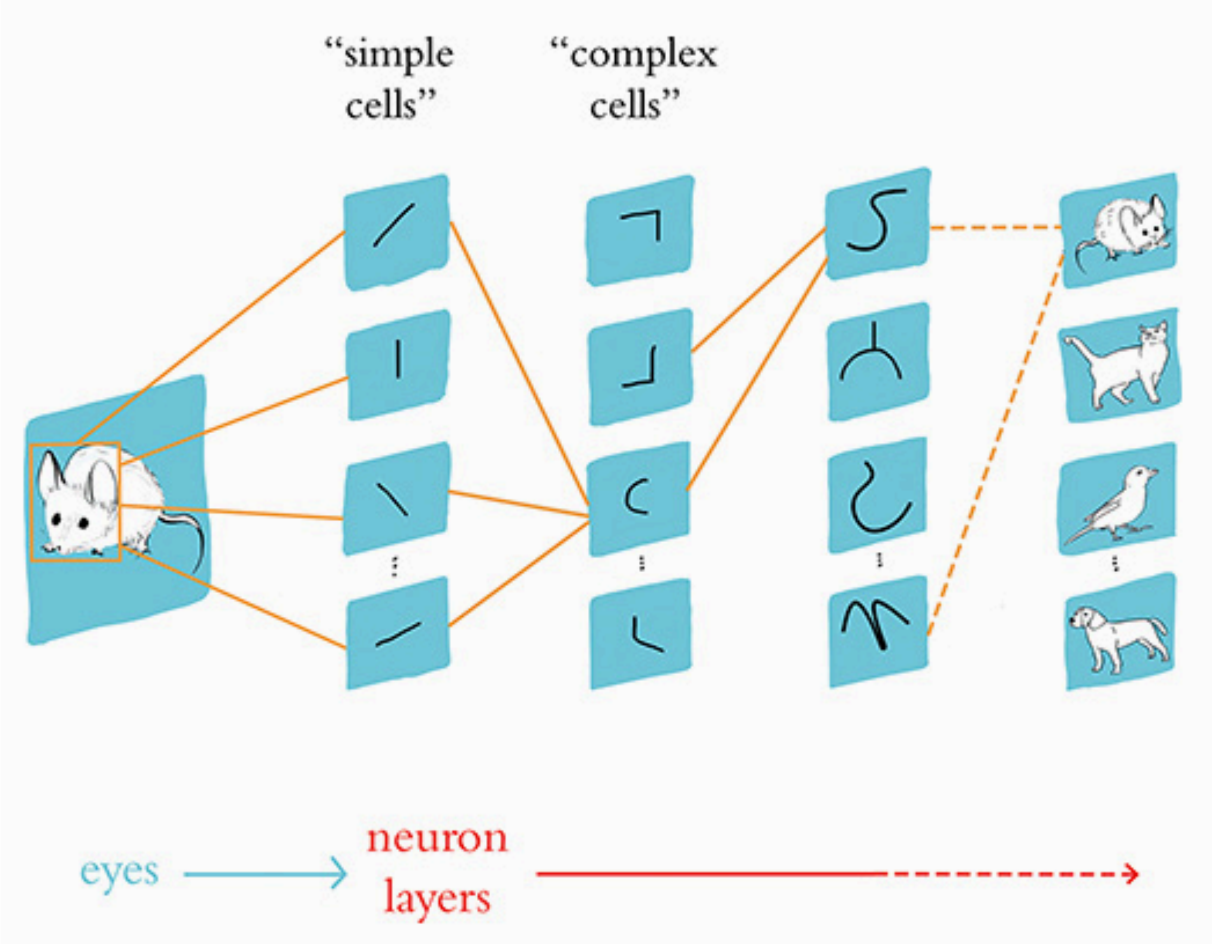
Source: Yantis, Steven. *Sensation and Perception, 2nd Edition*.

- Hubel and Wiesel used a light projector to present slides to anesthetized cats while they recorded the activity of neurons in the cats' primary visual cortex.
- In the experiments, electrical recording equipment was implanted within the cat's skull. Here we use a lightbulb to represent neuron activation.
- In the beginning, nothing happened.
- Then…



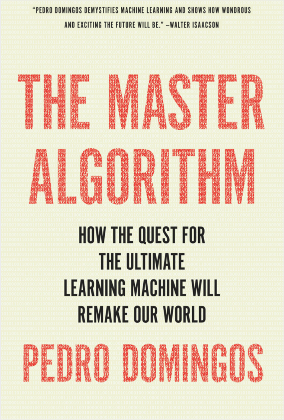


Source: "Deep Learning Illustrated"


"simple cells"
"complex cells"
eyes
neuron layers

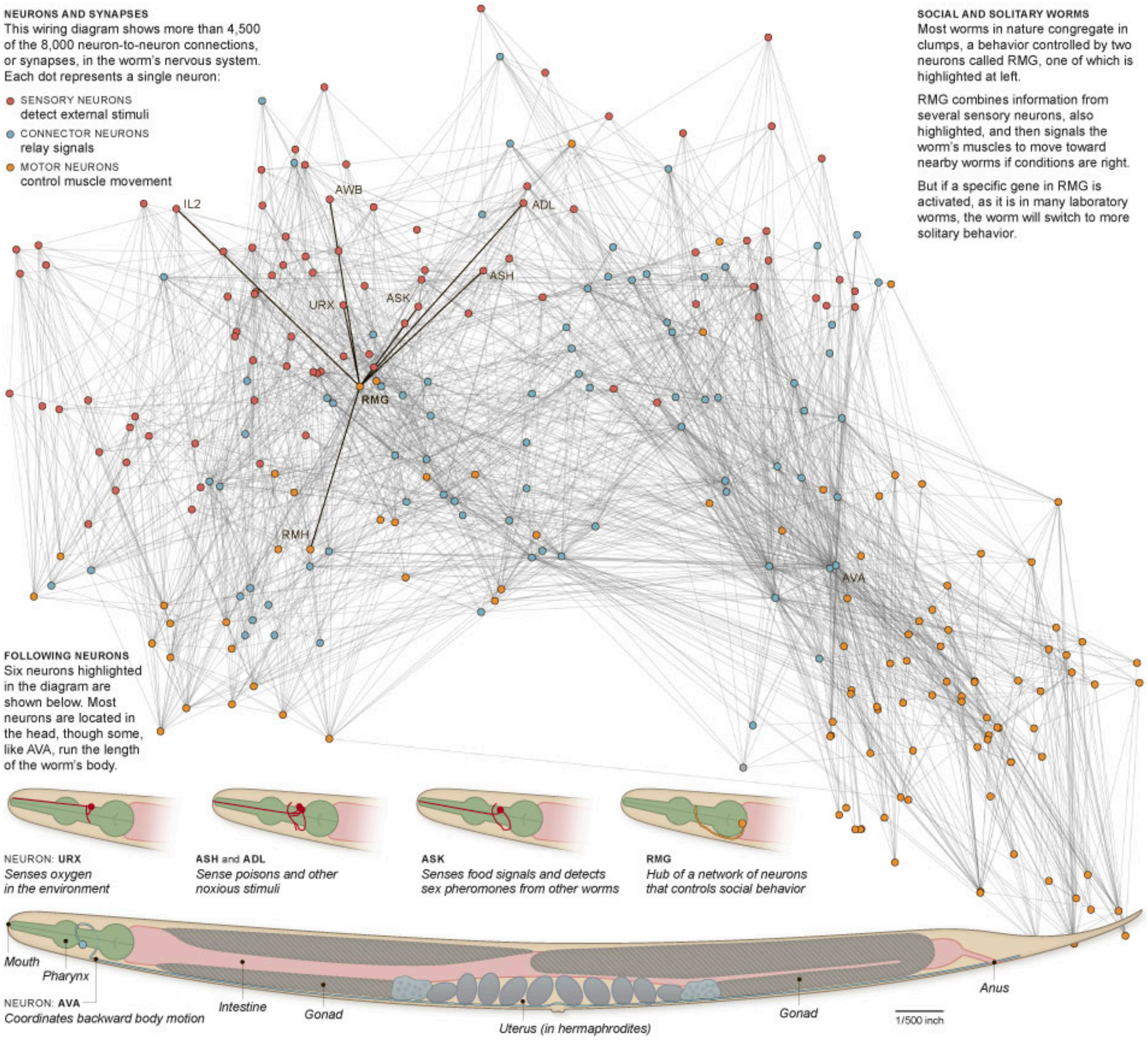
What about other animals' brains?

- "Even if we had a complete map of the brain, we would still be at a loss to figure out what it does."
- The nervous system of the *C. elegans* worm consists of only 302 neurons and was completely mapped in 1986, but we still have only a fragmentary understanding of what it does.



The New York Times

Science



Fundamental questions

What do we know about the (human) brain?

How much do we know about the (human) brain?

How do we learn about the (human) brain?

What about **other animals' brains**?

What can we achieve by **reverse engineering** the brain?

What can we achieve by reverse engineering the brain?

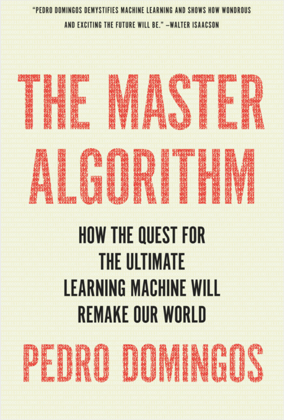
How many brain processes have been successfully modeled so far?

Should “biological plausibility” be a goal (of brain models)?

When are “short cuts” acceptable?

What can we achieve by reverse engineering the brain?

- We need higher-level concepts to make sense of the morass of low-level details, weeding out the ones that are specific to wetware or just quirks of evolution.
- We don’t build airplanes by reverse engineering feathers, and airplanes don’t flap their wings.
  - Rather, airplane designs are based on the principles of aerodynamics, which all flying objects must obey.
- We still do not understand those analogous principles of thought.



Learning from "first principles"

- ***Elon Musk***
- *"Well, I do think there's a good framework for thinking. It is physics. You know, the sort of first principles reasoning. Generally I think there are — what I mean by that is, boil things down to their fundamental truths and reason up from there, as opposed to reasoning by analogy."*
- *"It is important to view knowledge as sort of semantic tree. Make sure you understand the fundamental principles, i.e. the trunk and big branches, before you get into the leaves/details or there is nothing for them to hang on to."*

Sidebar

Selected questions about human vision
(and how close we are to answering them)

20 Years of Learning About Vision: Questions Answered, Questions Unanswered, and Questions Not Yet Asked

Bruno A. Olshausen

B.A. Olshausen (✉)
Helen Wills Neuroscience Institute and School of Optometry, University of California—Berkeley, Berkeley, CA, USA

Redwood Center for Theoretical Neuroscience, UC Berkeley, 575A Evans Hall, MC 3198, Berkeley, CA 94720-3198, USA
e-mail: baolshausen@berkeley.edu

J.M. Bower (ed.), *20 Years of Computational Neuroscience*, Springer Series in Computational Neuroscience 9, DOI 10.1007/978-1-4614-1424-7_12,

Summary of (Olshausen 2013)

- **Abstract**
  - A review of the progress that computational neuroscience has made over the past 20 years in understanding how vision works.

  - Perhaps the most important advance we have made is in gaining <u>*a deeper appreciation of the magnitude of the problem*</u> before us.

  - While there has been steady progress in our understanding […] we are still confronted with <u>*profound mysteries about how visual systems work*</u>.
    - These are not just mysteries about biology, but also about the *general principles* that enable vision in any system whether it be biological or machine.
    - The biggest mysteries are likely to be ones we are not currently aware of, and that bearing this in mind is important as it encourages a more exploratory, as opposed to strictly hypothesis-driven, approach.

Summary of (Olshausen 2013)

- Vision, though a seemingly simple act, presents us with *profound computational problems*.
  - <u>Even stating what these problems are has proven to be a challenge</u>.

  - One might hope that we could gain insight from studying biological vision systems, but this approach is plagued with its own problems: nervous systems are composed of many tiny, interacting devices that are difficult to penetrate.
    - *The closer one looks, the more complexity one is confronted with*.
  - The solutions nature has devised will not reveal themselves easily, but as we shall see <u>the situation is not hopeless</u>.

Summary of (Olshausen 2013)

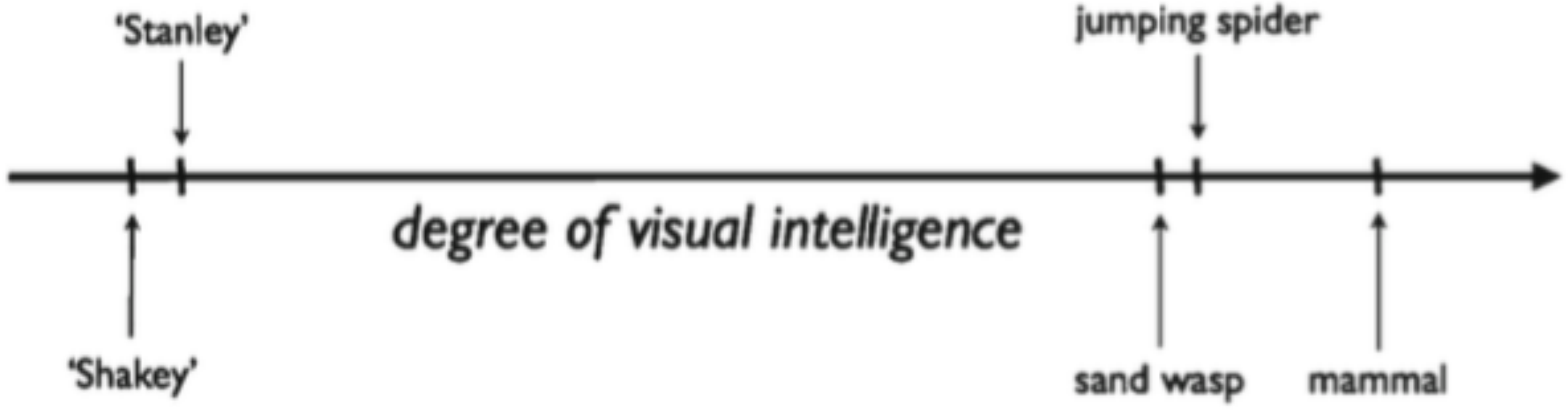
- **Questions Answered**
  - Tiling in the Retina
  - The Relation Between Natural Image Statistics and Neural Coding
  - The Nature of Intermediate-Level Vision
  - Functional Organization of Human Visual Cortex
  - *How to Infer Scene Geometry from Multiple Views*
    - *SIFT and other keypoint detectors and invariant feature descriptors*
    - *Photosynth*

Summary of (Olshausen 2013)

- **On the interplay between human and computer vision**
  - There has long been a productive interchange of ideas between the fields of computer vision and biological vision.
  - I believe the advances in multiple-view geometry tell us something important about vision, and that they open the door to a new area of investigation in visual neuroscience:
    <u>*How do animals assimilate the many views they obtain of their environment into a unified representation of the 3D scene*</u>?

Summary of (Olshausen 2013)

- **Questions Unanswered**
  - How Is Sophisticated Vision Possible in Tiny Nervous Systems?
  - How Do Cortical Microcircuits Contribute to Vision?
  - How Does Feedback Contribute to Vision?
  - What Is the Role of Neuronal Oscillations in Visual Processing?
  - *How to Build Robust, Autonomous Vision Systems?*



Summary of (Olshausen 2013)

- **Questions Not Yet Asked**
  - *The Need for Exploratory Approaches*
    - No one should feel ashamed to report a complete, unfiltered set of findings without a story to envelop them. After all, one person's untidy finding may provide the missing piece in another person's theory.
  - *Learning About Vision by Building Autonomous Systems*
    - There is very little that neuroscience per se has taught us about the principles of vision.
    - Computer vision to the rescue?

(Olshausen 2013): my comments

- These remarks are from almost 10 years ago
- We were on the verge of the deep learning era in Computer Vision
- Self-driving vehicles have come a long way since then!
- The lists / topics may have shifted a bit…
  - … but they are far from being an empty set

https://braininitiative.nih.gov/

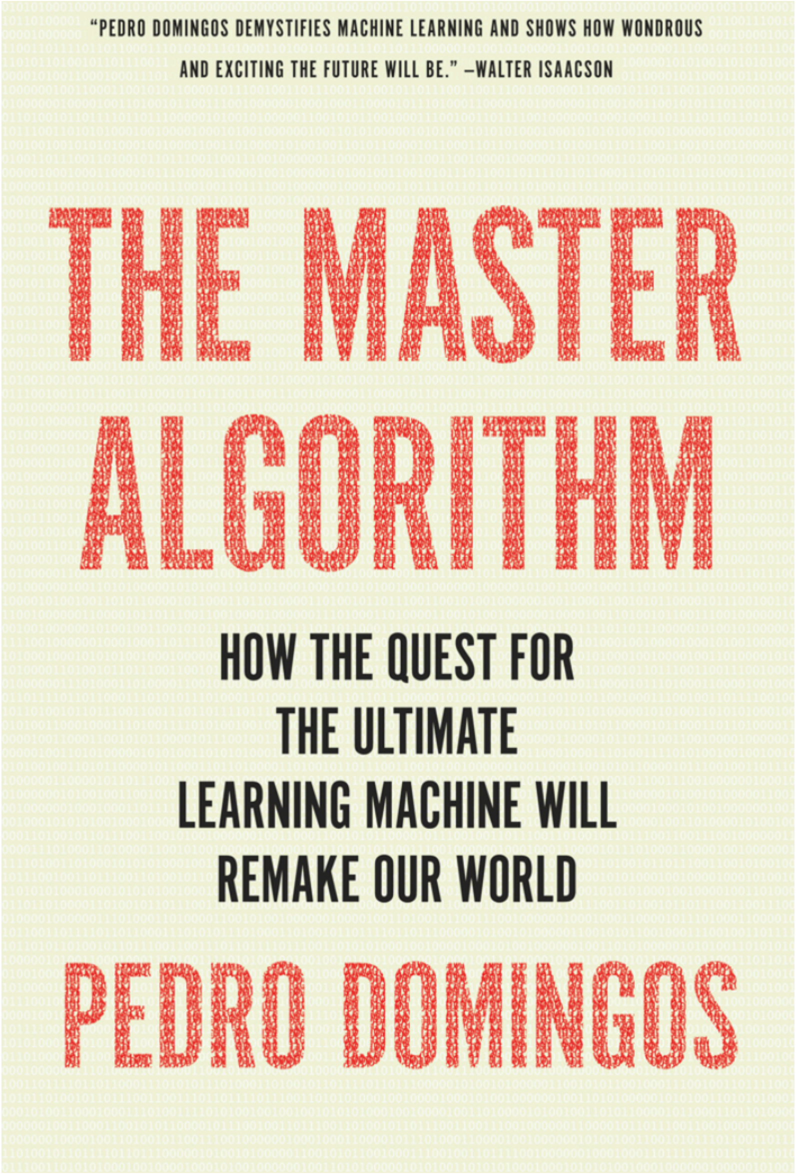
NIH The BRAIN Initiative®

The Brain Research through Advancing Innovative Neurotechnologies® (BRAIN) Initiative is aimed at revolutionizing our understanding of the human brain.

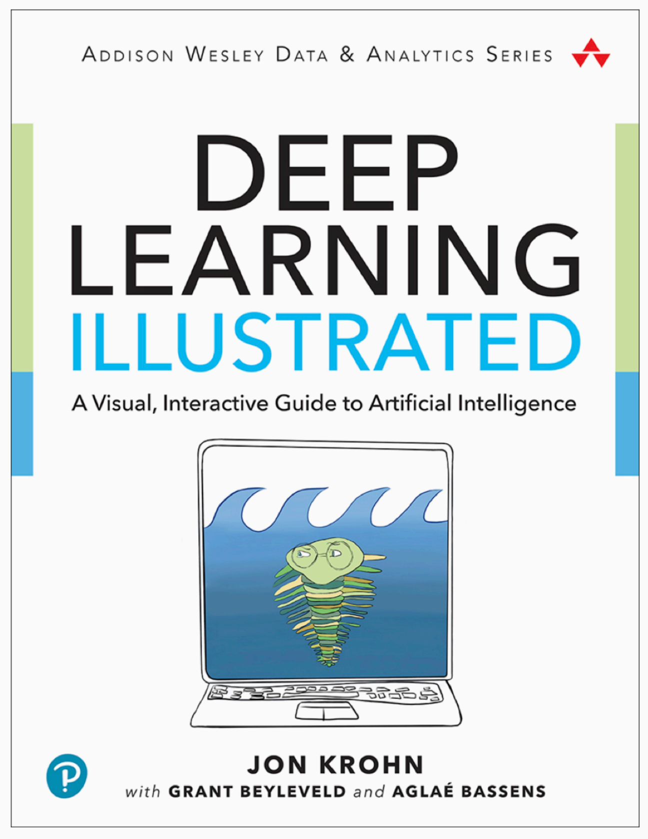
- Cell Type
- Circuit Diagrams
- Monitor Neural Activity
- Interventional Tools
- Theory & Data Analysis Tools
- Human Neuroscience
- Integrated Approaches

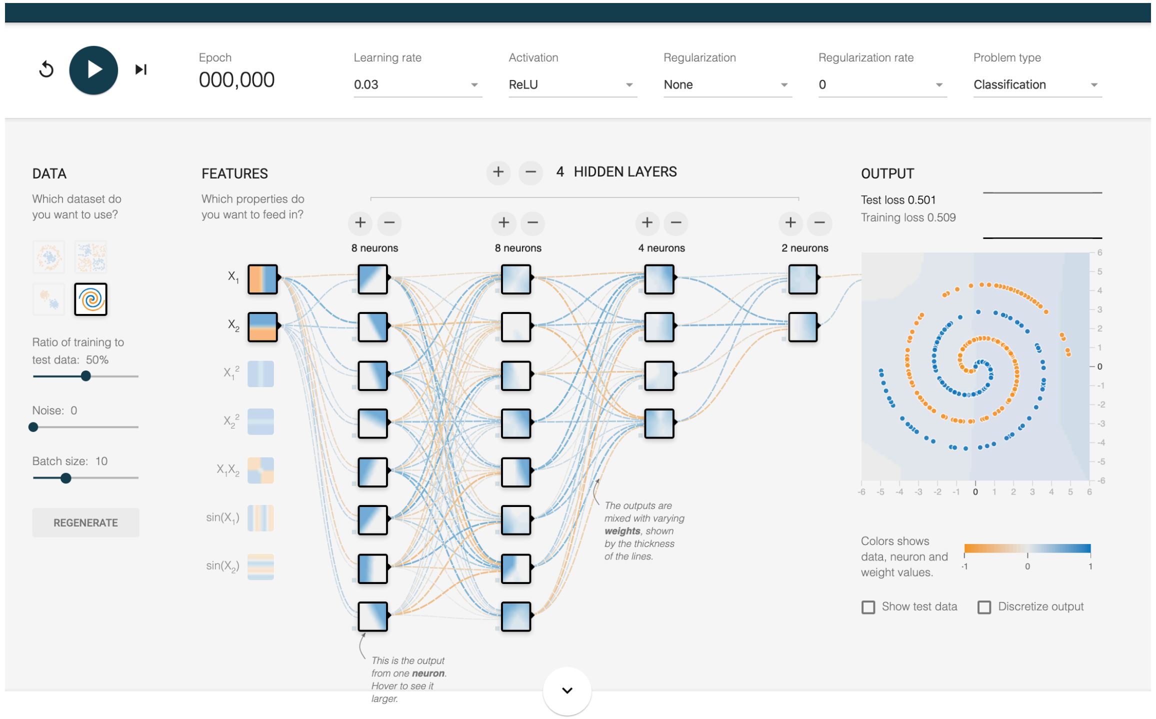
The Brain Research through Advancing Innovative Neurotechnologies® (BRAIN) Initiative is aimed at revolutionizing our understanding of the human brain. By accelerating the development and application of innovative technologies, researchers will be able to produce a revolutionary new dynamic picture of the brain that, for the first time, shows how individual cells and complex neural circuits interact in both time and space.

Back to the book…



… with “special guests”





Hebb’s rule

- The cornerstone of connectionism.
  - Indeed, the field derives its name from the belief that knowledge is stored in the connections between neurons.
- Donald Hebb, a Canadian psychologist, stated it this way in his 1949 book *The Organization of Behavior*:
  - “When an axon of cell A is near enough cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”
  - It’s often paraphrased as “Neurons that fire together wire together.”

Symbolist vs. Connectionist Learning

- In symbolist learning, there is a one-to-one correspondence between symbols and the concepts they represent.
- In contrast, connectionist representations are distributed: each concept is represented by many neurons, and each neuron participates in representing many different concepts.
- Neurons that excite one another form what Hebb called a *cell assembly*.
- Concepts and memories are represented in the brain by cell assemblies.

Symbolist vs. Connectionist Learning

- Another difference between symbolist and connectionist learning is that the former is sequential, while the latter is parallel.

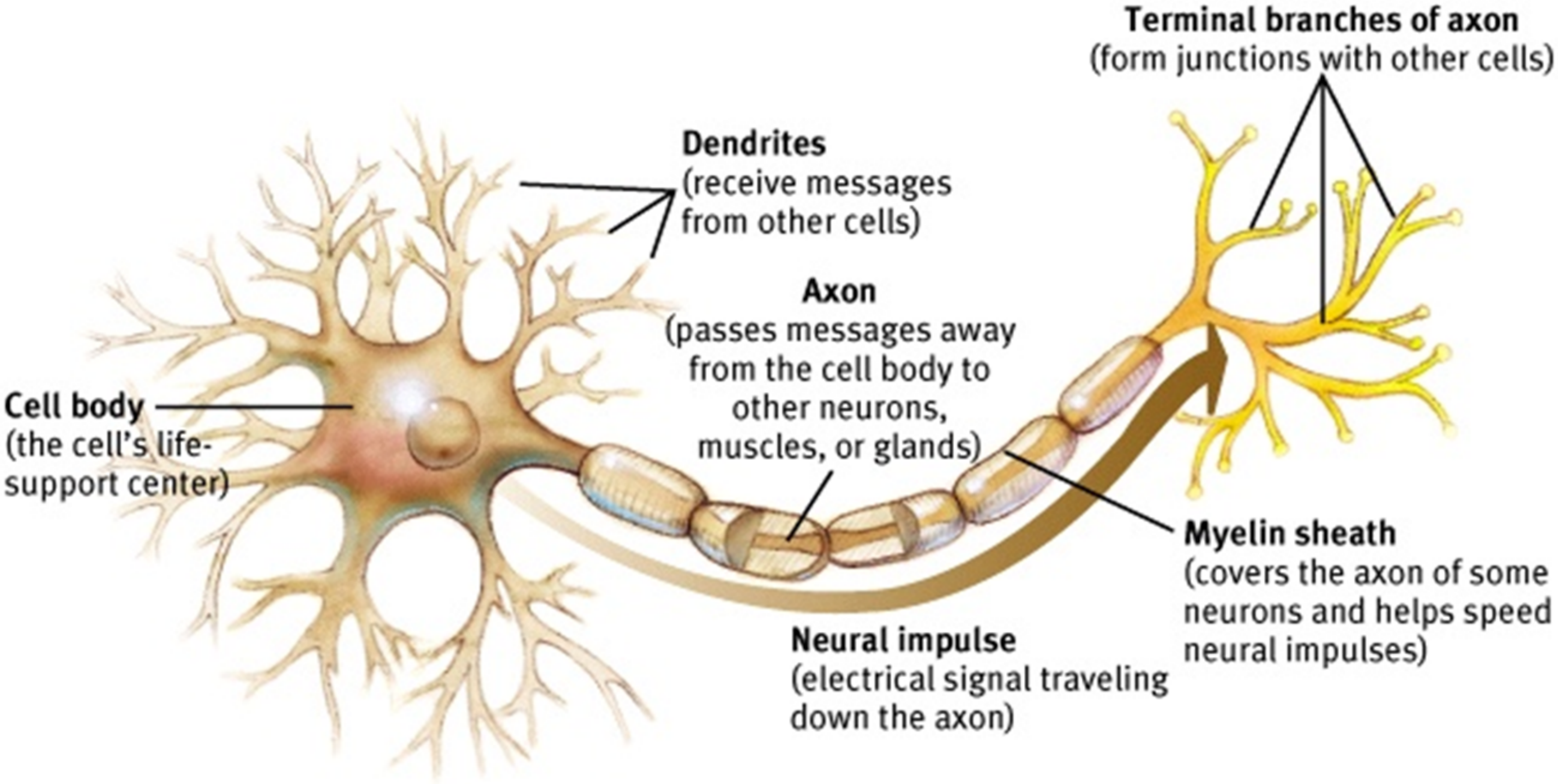
- Brains can perform a large number of computations in parallel, with billions of neurons working at the same time; but each of those computations is slow, because neurons can fire at best a thousand times per second.

Simulating a brain

To simulate a brain, we need more than Hebb's rule, however; we need to understand how the brain is built.

The next step is to turn it into an algorithm.

A Neuron



Source: P. Domingos

Biological neuron

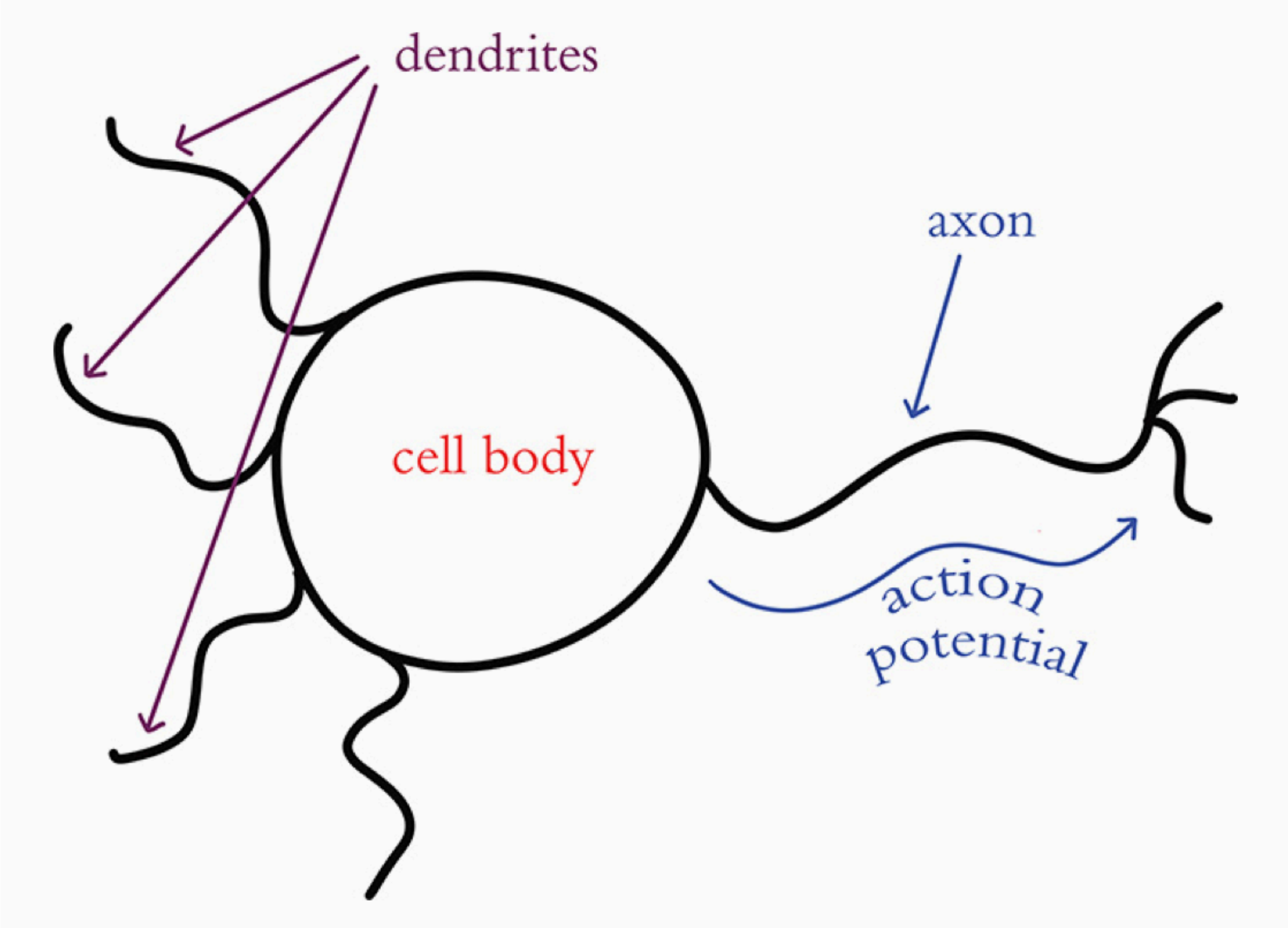
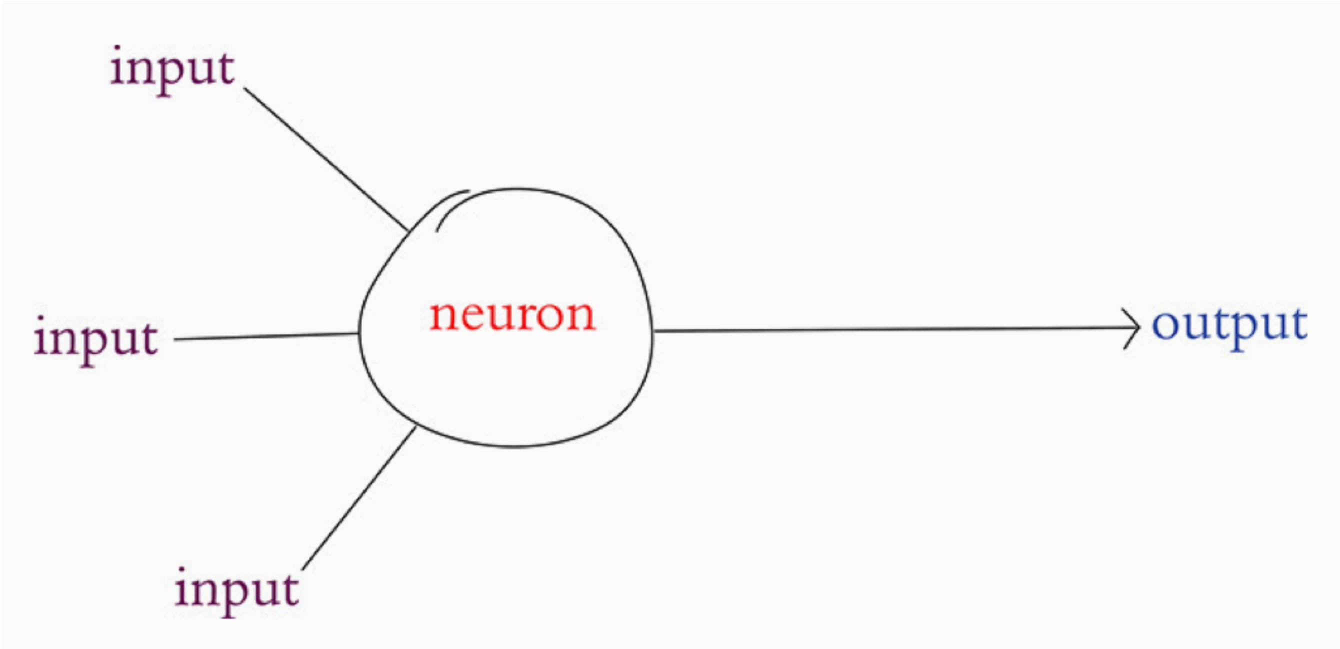


Figure from: “Deep Learning Illustrated” (Krohn, 2020)

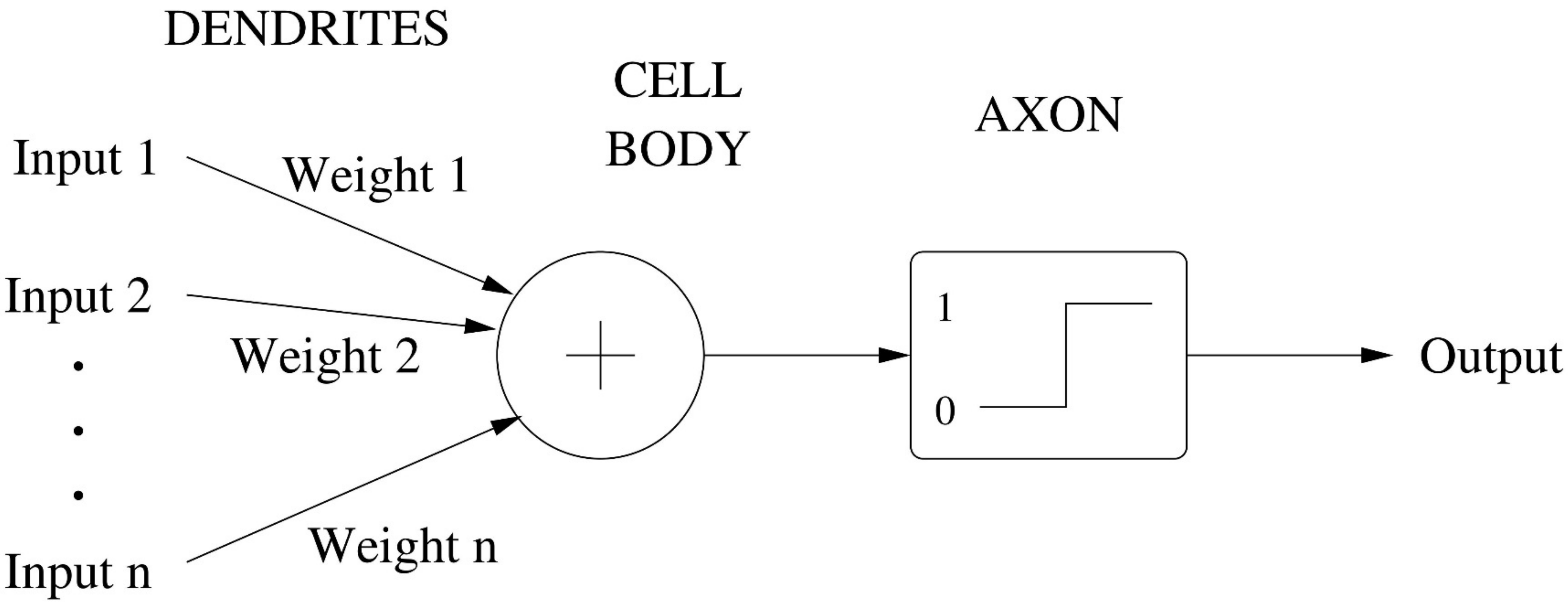
The perceptron



"Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and the organization in the brain. Psychological Review, 65, 386–408."

Figure from: "Deep Learning Illustrated" (Krohn, 2020)

An Artificial Neuron (Perceptron)



Source: P. Domingos

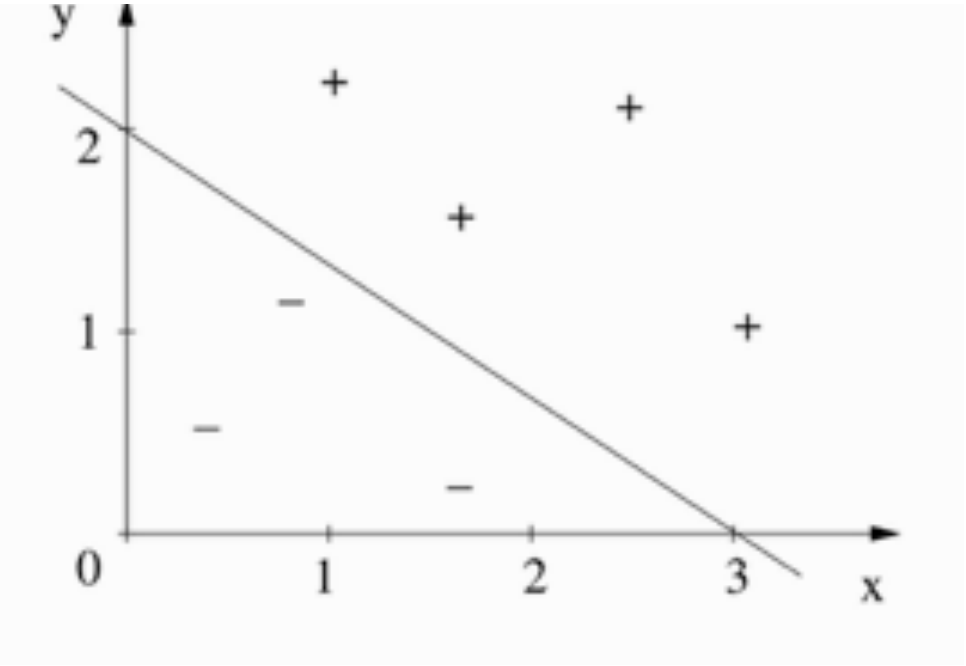
The Perceptron

- In a perceptron, a positive weight represents an *excitatory* connection, and a negative weight an *inhibitory* one. T
- he perceptron outputs 1 if the weighted sum of its inputs is above <u>threshold</u>, and 0 if it’s below.
- By varying the weights and threshold, we can change the function that the perceptron computes.
- This ignores a lot of the details of how neurons work, of course, but we want to keep things as simple as possible; our goal is to develop a general-purpose learning algorithm, not to build a realistic model of the brain.

The Perceptron: example

- The boundary is the set of points where the weighted sum exactly equals the threshold, and a weighted sum is a linear function.
- For example, if the weights are 2 for x and 3 for y and the threshold is 6, the boundary is defined by the equation

$$2x + 3y = 6$$



The general equation for artificial neurons

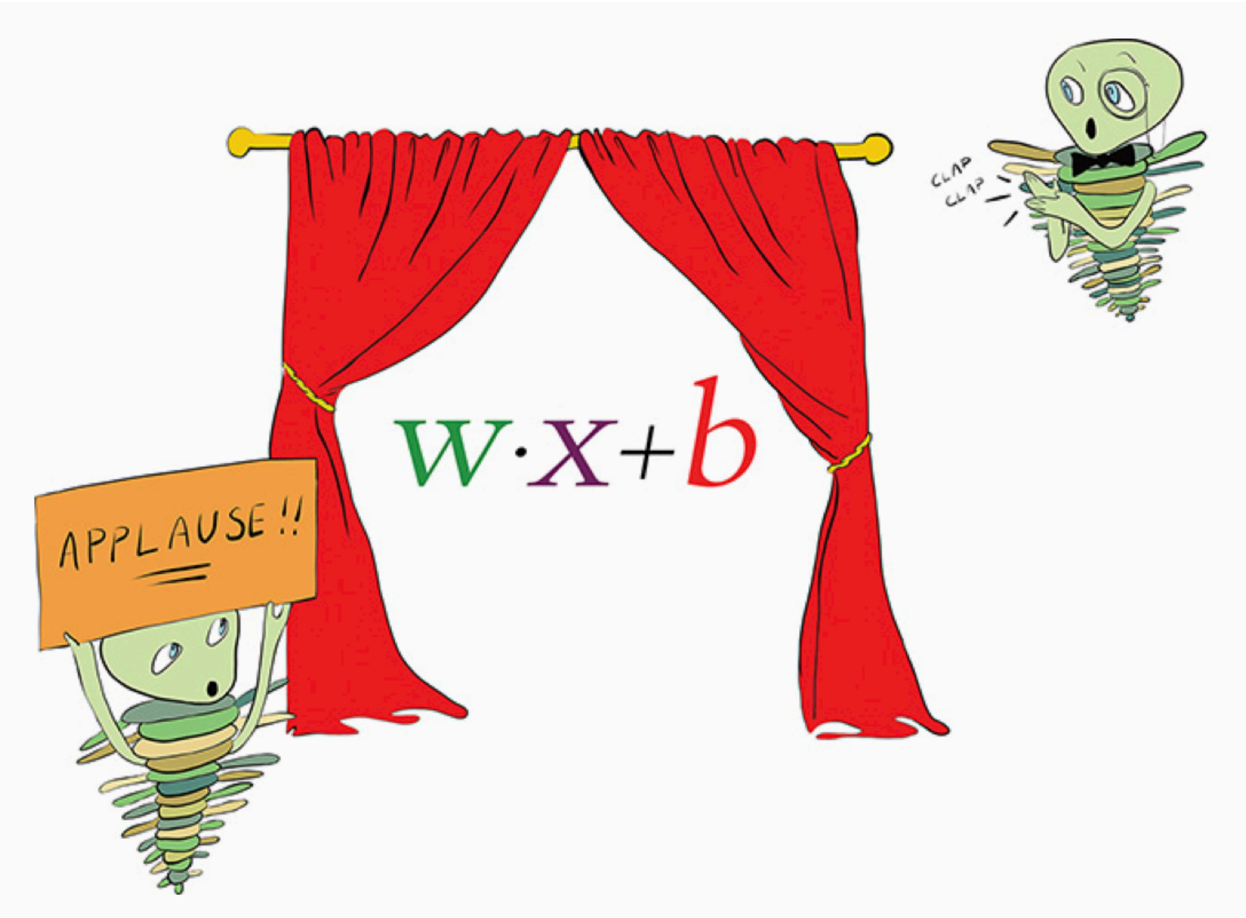
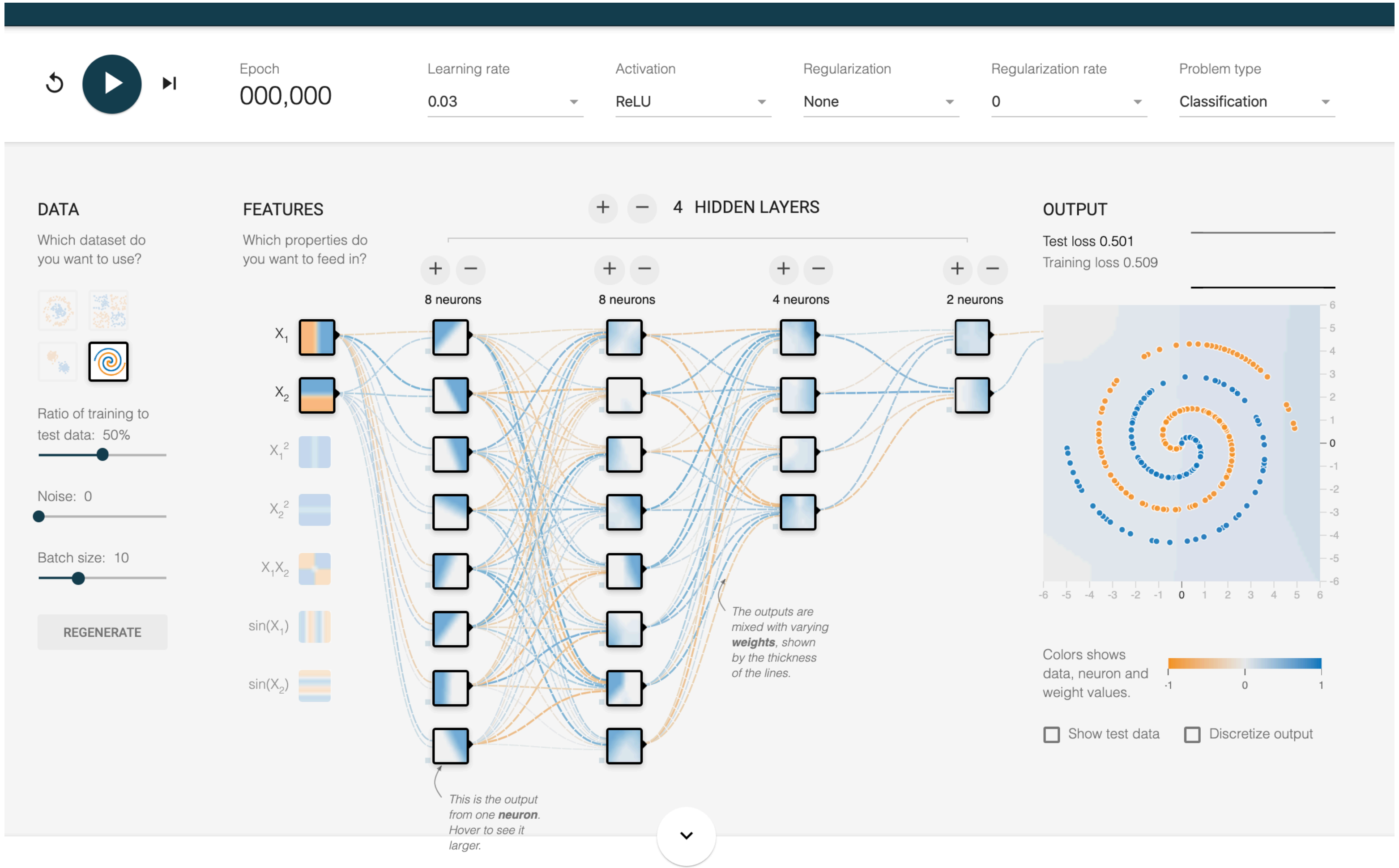


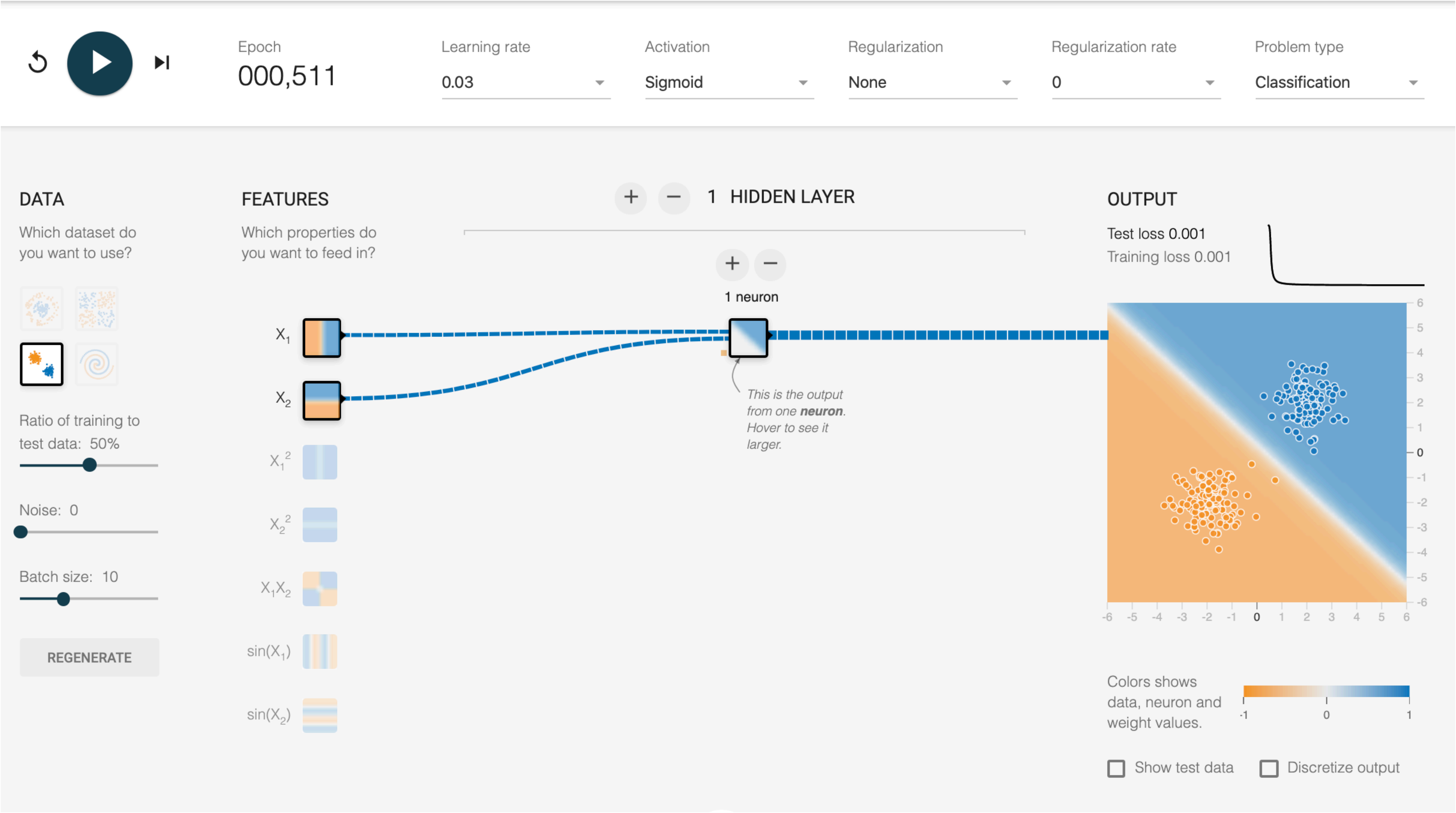
Figure from: “Deep Learning Illustrated” (Krohn, 2020)

TensorFlow Playground

https://playground.tensorflow.org/



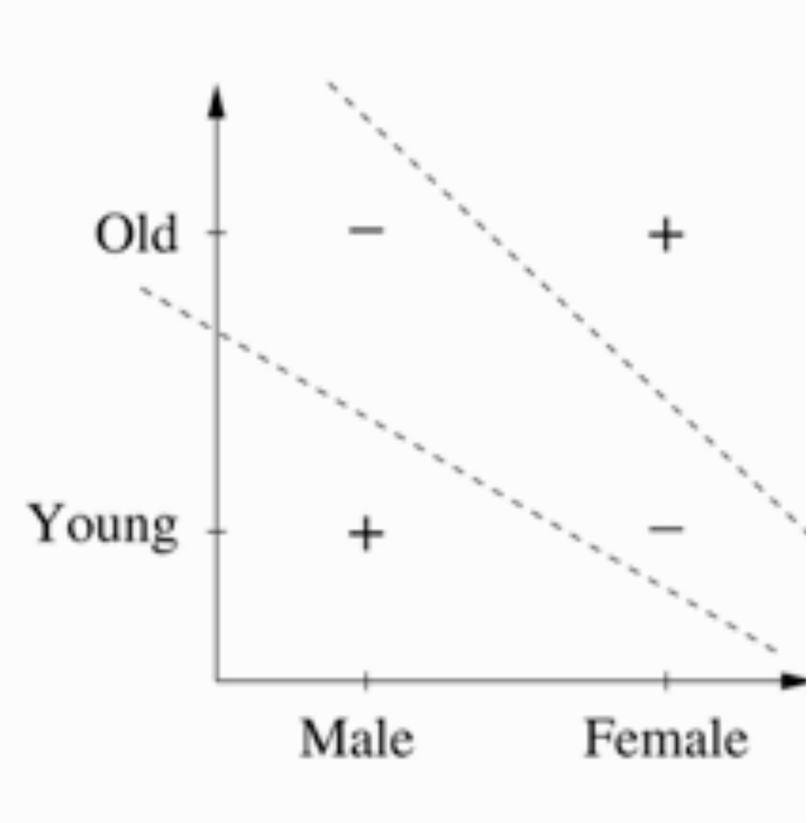
Playground time!



The XOR problem

- The problem with XOR is that there is no straight line capable of separating the positive from the negative examples.

- Since perceptrons can only learn linear boundaries, they can't learn XOR. And if they can't do even that, they're not a very good model of how the brain learns, or a viable candidate for the Master Algorithm.



The XOR problem

- Problem: a perceptron models only a single neuron's learning.
- Possible solution: add a hidden layer.

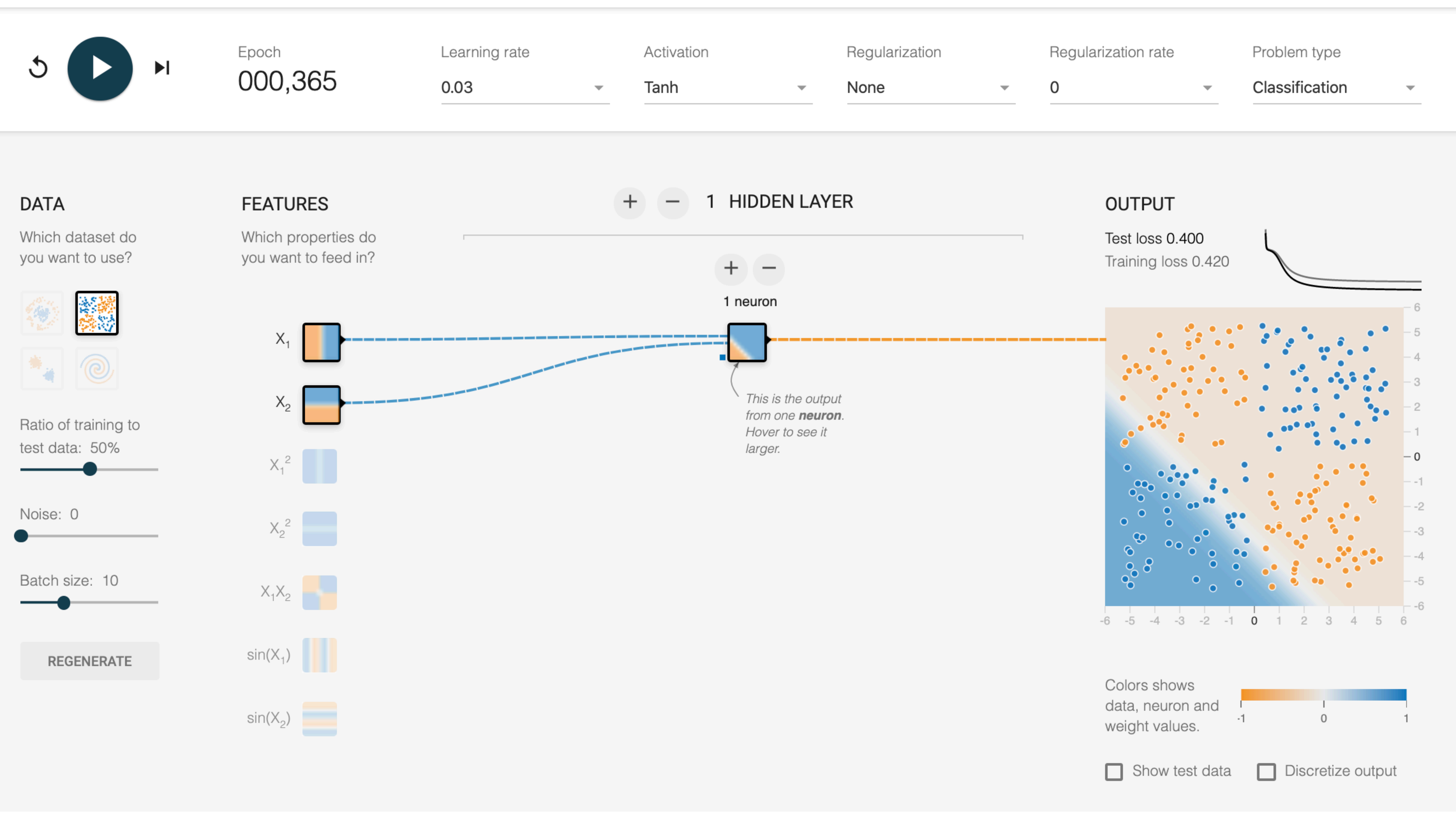
The problem is that there's no clear way to change the weights of the neurons in the "hidden" layers to reduce the errors made by the ones in the output layer.

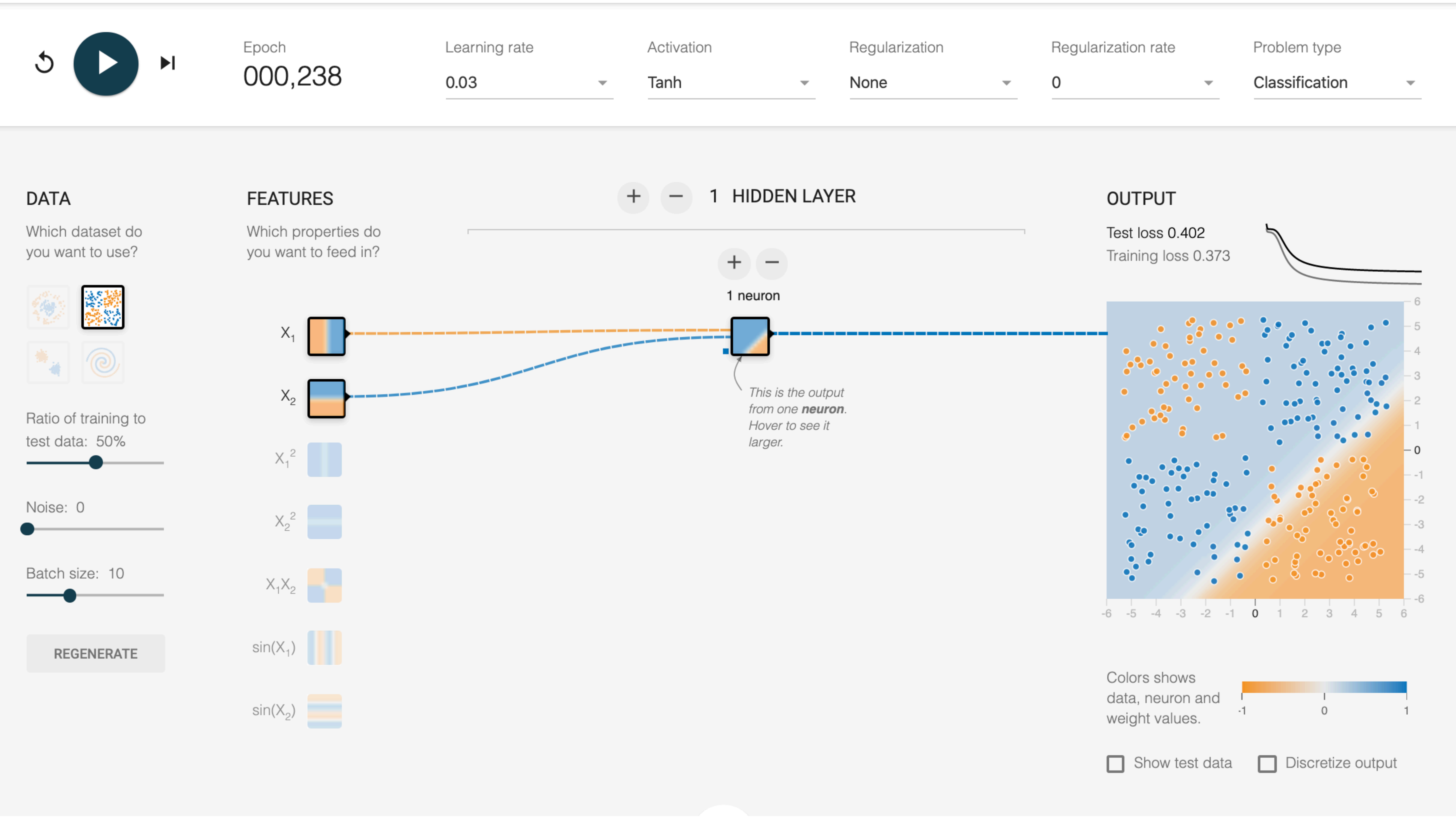
Every hidden neuron influences the output via multiple paths, and every error has a thousand fathers.

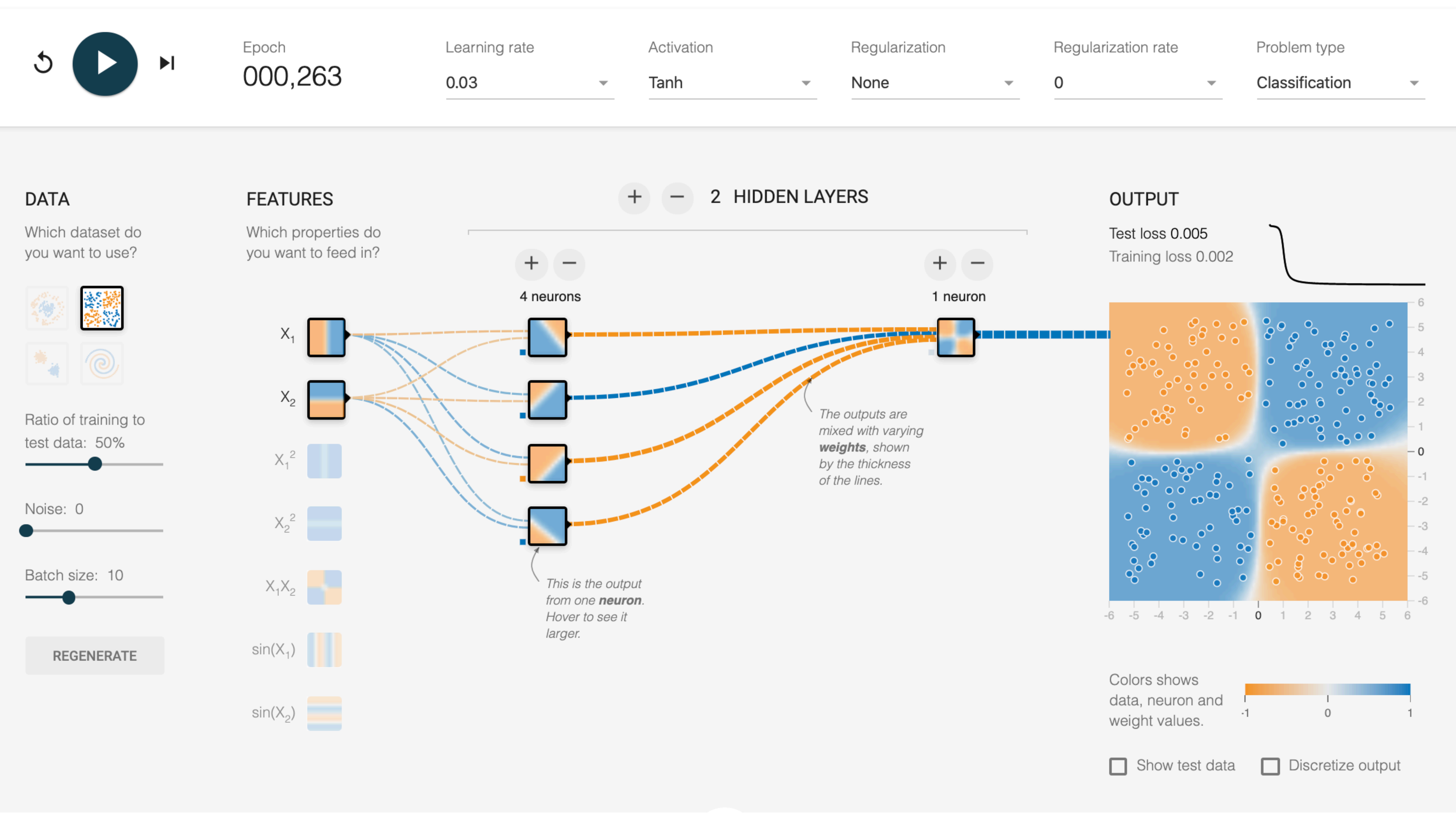
Who do you blame? Or, conversely, who gets the credit for correct outputs? This credit-assignment problem shows up whenever we try to learn a complex model and is one of the central problems in machine learning.

Playground time!

- Simulate the XOR problem in TensorFlow Playground.
- Search (by trial-and-error) for a solution (there are many!).
- Record your steps and findings.
- Write down your conclusions and questions.


Epoch
000,365
Learning rate
0.03
Activation
Tanh
Regularization
None
Regularization rate
0
Problem type
Classification
DATA
Which dataset do you want to use?
Ratio of training to test data: 50%
Noise: 0
Batch size: 10
REGENERATE
FEATURES
Which properties do you want to feed in?
X1
X2
X1^2
X2^2
X1X2
sin(X1)
sin(X2)
1 HIDDEN LAYER
1 neuron
This is the output from one neuron. Hover to see it larger.
OUTPUT
Test loss 0.400
Training loss 0.420
Colors shows data, neuron and weight values.
-1
0
1
Show test data
Discretize output


Epoch
000,238
Learning rate
0.03
Activation
Tanh
Regularization
None
Regularization rate
0
Problem type
Classification
DATA
Which dataset do you want to use?
Ratio of training to test data: 50%
Noise: 0
Batch size: 10
REGENERATE
FEATURES
Which properties do you want to feed in?
X1
X2
X1^2
X2^2
X1X2
sin(X1)
sin(X2)
1 HIDDEN LAYER
1 neuron
This is the output from one neuron. Hover to see it larger.
OUTPUT
Test loss 0.402
Training loss 0.373
Colors shows data, neuron and weight values.
-1
0
1
Show test data
Discretize output


Epoch
000,263
Learning rate
0.03
Activation
Tanh
Regularization
None
Regularization rate
0
Problem type
Classification
DATA
Which dataset do you want to use?
Ratio of training to test data: 50%
Noise: 0
Batch size: 10
REGENERATE
FEATURES
Which properties do you want to feed in?
X1
X2
X1^2
X2^2
X1X2
sin(X1)
sin(X2)
2 HIDDEN LAYERS
4 neurons
1 neuron
The outputs are mixed with varying weights, shown by the thickness of the lines.
This is the output from one neuron. Hover to see it larger.
OUTPUT
Test loss 0.005
Training loss 0.002
Colors shows data, neuron and weight values.
-1
0
1
Show test data
Discretize output