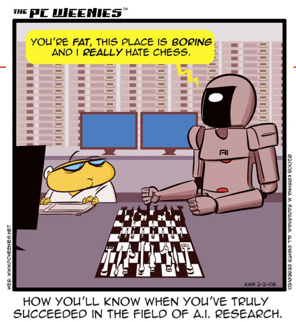


http://www.bbspot.com/comics/PC-Weenies/2008/02/3248.php

1

Intro to AI

CS51A
David Kauchak
Spring 2025

Adapted from notes from:
Sara Owsley Sood

2

Admin

Assignment 6 due Friday

Lab this week

Extra mentor hours on Friday (TBA)

Ethics presentations:

Over the past assignments, you have read a number of articles that have exposed you to potential pitfalls of AI technology. For the *Tuesday 3/25* class, you will give a 3–4 minute presentation in a team of 2–3 students on a topic of your choice that revolves around the intersection of AI and ethics.

3

AI is a huge field

What is AI (artificial intelligence)…

4

AI is a huge field

What is AI…

One definition:

"Building programs that enable computers to do what humans can do."

For example:
read, walk around, drive, play games, solve problems, learn, have conversations…

5

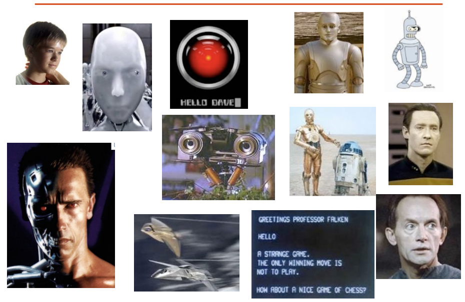
A broader definition

"Building programs that enable computers to do do *intelligent* things"

	human	rational
thinking	**Think like a human** Cognitive Modeling	**Think rationally** Logic-based Systems
acting	**Act like a human** Turing Test	**Act rationally** Rational Agents

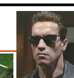
6

How is AI viewed in popular media?



7

What challenges are there?



8

What challenges are there?

Perception
- perceive the environment via sensors

Computer vision (perception via images/video)
- process visual information
- object identification, face recognition, motion tracking

Natural language processing and generation
- speech recognition, language understanding
- language translation, speech generation, summarization

9

What challenges are there?

Knowledge representation
- encode known information
- water is wet, the sun is hot, Dave is a person, …

Learning
- learn from environment
- What type of feedback? (supervised vs. unsupervised vs. reinforcement vs …)

Reasoning/problem solving
- achieve goals, solve problems
- planning
- How do you make an omelet? I'm carrying an umbrella and it's raining… will I get wet?

Robotics
- How can computers interact with the physical world?

10

What can we currently do?

11

What can we currently do?

Understand spoken language?

speech recognition is really good, if:
- restricted vocabulary
- specific speaker with training

Gotten quite good in the last few years and shows up in lots of places:
- Mobile: Siri, Ok Google, etc.
- Home assistants: Alexa, Google Home

What does the spoken language actually mean (language understanding)?
- much harder problem!
- LLMs are getting better at this, though they don't explicit represent the "meaning"

12

What can we currently do?

Speak?

Understandable, but you wouldn't confuse it for a person

Can do accents, intonations, etc.

Better with restricted vocabulary

Loquendo

- https://www.nuance.com/omni-channel-customer-engagement/voice-and-ivr/text-to-speech.html
- Dealing with facial expression is challenging


Kismet (MIT)

13

What can we currently do?

Drive a car?

14

What can we currently do?

Drive a car?

Freeway driving is relatively straightforward

Off-road a bit harder

- see DARPA grand challenges (2004, 2005)

And urban driving is even trickier

- Waymo
- Tesla
- Uber

15

What can we currently do?

Drive a car?

Many driver assist technologies:

- Automatic breaking
- Automatic pedestrian detection
- Lane drift avoidance
- "smart" cruise control
- Blind spot warning
- …

17

What can we currently do?

Identify emotion?

This is hard!

Some success in text
- movie reviews (assignment 7!)
- blogs
- twitter
- dealing with sarcasm is hard

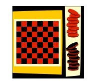
Some success with faces
- strongly biased by training data
- works best when exaggerated

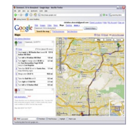
18

What can we currently do?

Reasoning?

Success on small sub-problems





General purpose reasoning is harder
- Wolfram Alpha
- OpenCyc

19

What can we currently do?

Walk?

Robots have had a variety of locomotion methods

Walking with legs, is challenging
- Differing terrains, stairs, running, ramps, etc.

Getting better every year
- https://www.youtube.com/watch?v=8dFTc4W8wm0

20

When will I have my robot helper?

What can we currently do?

21

What can we currently do?



22

What can we currently do?

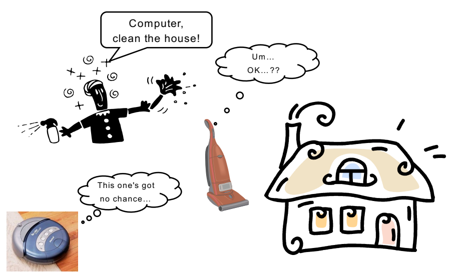
Fold a pile of towels?

UC Berkeley towel folding robot:

http://www.youtube.com/watch?v=gy5g33S0Gzo

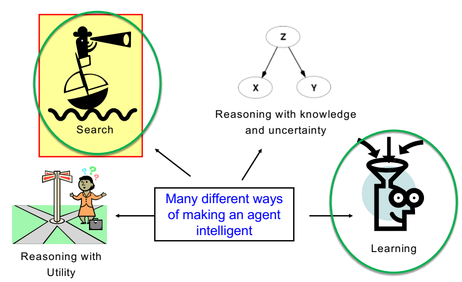
23

How do we make computers "intelligent?"



24

Fundamental problem of AI

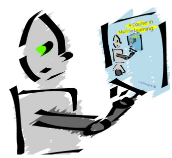


25

Machine Learning is…

Machine learning is about predicting the future based on the past.

-- Hal Daume III

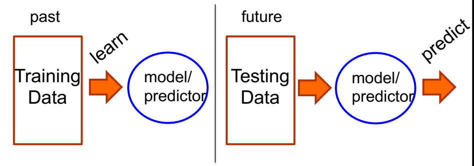


26

Machine Learning is…

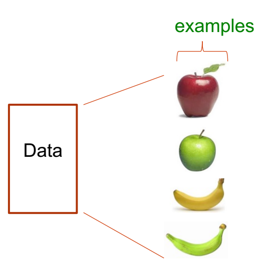
Machine learning is about predicting the future based on the past.

-- Hal Daume III



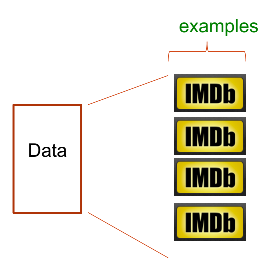
27

Data



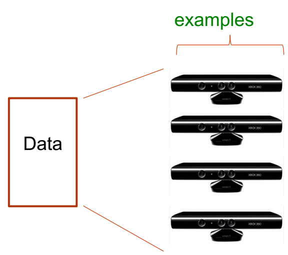
28

Data



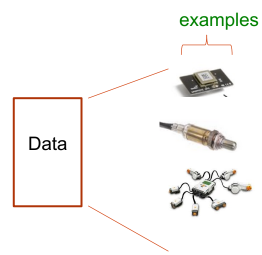
29

Data



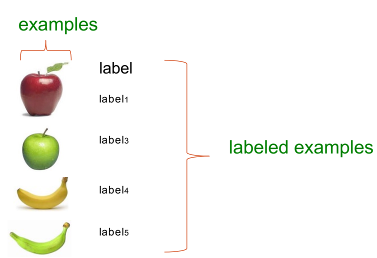
30

Data



31

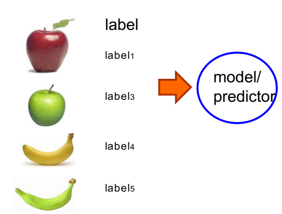
Supervised learning



Supervised learning: given labeled examples

32

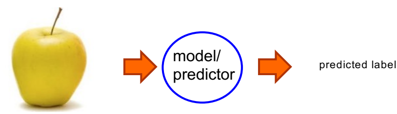
Supervised learning



Supervised learning: learn a model

33

Supervised learning



Supervised learning: learn to predict new example

34

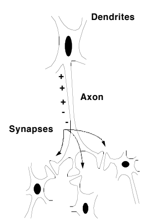
Neural Networks

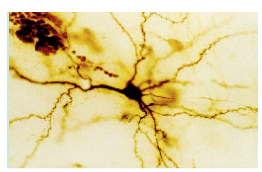
Neural Networks try to mimic the structure and function of our nervous system

People like biologically motivated approaches

35

Our Nervous System



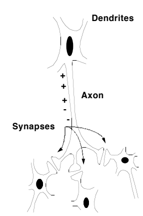


Neuron

What do you know?

36

Our nervous system: the computer science view



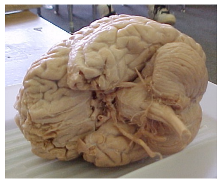
the human brain is a large collection of interconnected neurons

a NEURON is a brain cell

- they collect, process, and disseminate electrical signals
- they are connected via synapses
- they FIRE depending on the conditions of the neighboring neurons

37

Our nervous system



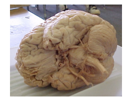
The human brain

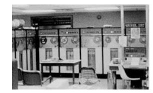
- contains ~10^{11} (100 billion) neurons
- each neuron is connected to ~10^4 (10,000) other neurons
- Neurons can fire as fast as 10^{-3} seconds

How does this compare to a computer?

38

Humans vs. Machines

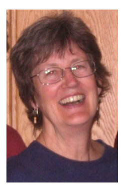




10^{11} neurons	10^{10} transistors
10^{11} neurons	10^{11} bits of ram/memory
10^{14} synapses	10^{13} bits on disk
10^{-3} "cycle" time	10^{-9} cycle time

39

Brains are still pretty fast



Who is this?

40

Brains are still pretty fast

If you were me, you'd be able to identify this person in 10^{-1} (1/10) s!

Given a neuron firing time of 10^{-3} s, how many neurons in sequence could fire in this time?

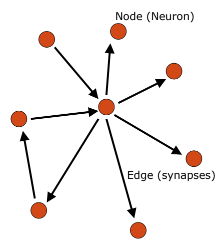
- A few hundred

What are possible explanations?

- either neurons are performing some very complicated computations
- brain is taking advantage of the **massive** parallelization (remember, neurons are connected ~10,000 other neurons)

41

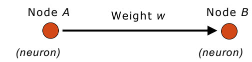
Artificial Neural Networks



our approximation

42

Neural networks



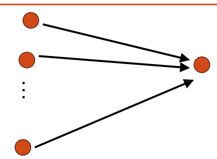
W is the strength of signal sent between A and B.

If *A* fires and *w* is **positive**, then *A* **stimulates** *B*.

If *A fires* and *w* is **negative**, then *A* **inhibits** *B*.

43

Neural networks



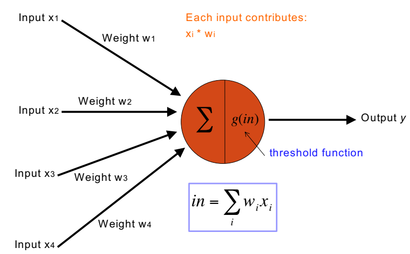
A given neuron has many, many connecting, input neurons

If a neuron is stimulated enough, then it also fires

How much stimulation is required is determined by its **threshold**

44

A Single Neuron/Perceptron



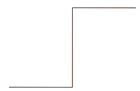
$$in = \sum_i w_i x_i$$

45

Possible threshold functions

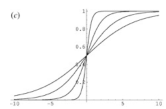
hard threshold

$$g(x) = \begin{cases} 1 & \text{if } x \geq threshold \\ 0 & otherwise \end{cases}$$



sigmoid

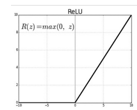
$$g(x) = \frac{1}{1 + e^{-ax}}$$



46

Many other activation functions

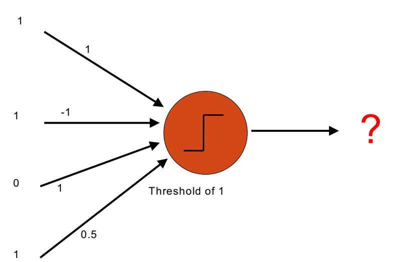
Rectified Linear Unit



Softmax (for probabilities)

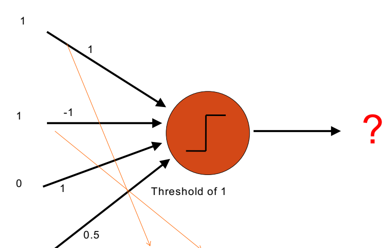
47

A Single Neuron/Perceptron



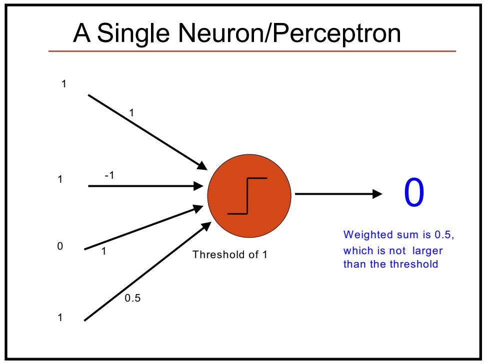
48

A Single Neuron/Perceptron

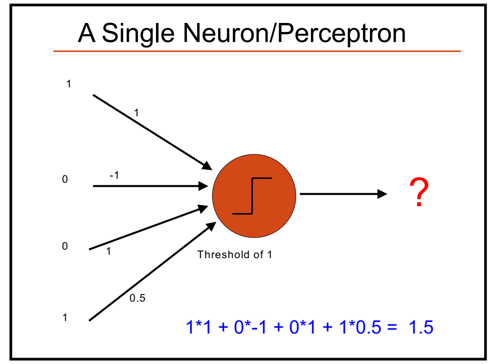


1*1 + 1*-1 + 0*1 + 1*0.5 = 0.5

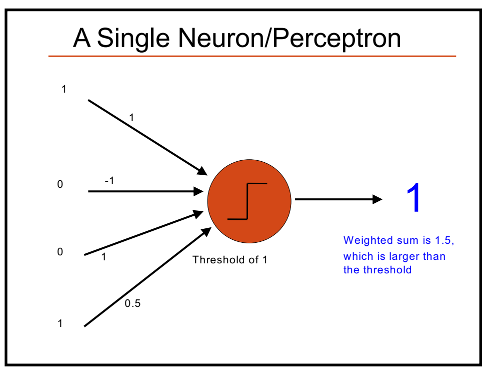
49


A Single Neuron/Perceptron
1
1
1
-1
0
1
1
0.5
Threshold of 1
0
Weighted sum is 0.5,
which is not larger
than the threshold

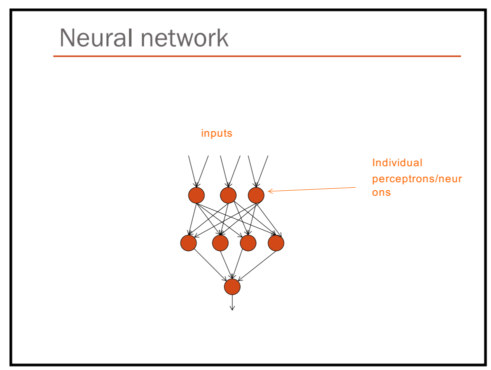
50


A Single Neuron/Perceptron
1
1
0
-1
0
1
1
0.5
Threshold of 1
?
1*1 + 0*-1 + 0*1 + 1*0.5 = 1.5

51

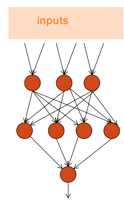

A Single Neuron/Perceptron
1
1
0
-1
0
1
1
0.5
Threshold of 1
1
Weighted sum is 1.5,
which is larger than
the threshold

52


Neural network
inputs
Individual
perceptrons/neur
ons

53

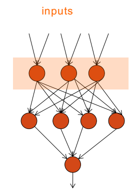
Neural network



some inputs are provided/entered

54

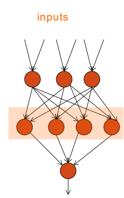
Neural network



each perceptron computes and calculates an answer

55

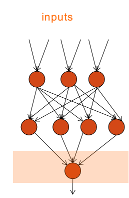
Neural network



those answers become inputs for the next level

56

Neural network

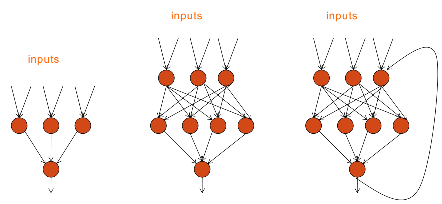


finally get the answer after all levels compute

57

Neural network

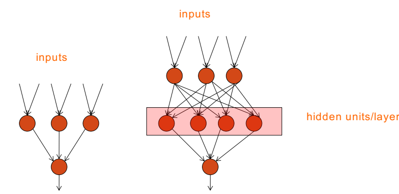
Different kinds/characteristics of networks



How are these different?

58

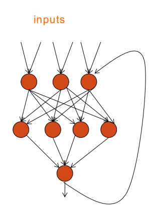
Neural network



Feed forward networks

59

Neural network



Recurrent network

Output is fed back to input

Can support memory!

How?

60

History of Neural Networks

McCulloch and Pitts (1943) – introduced model of artificial neurons and suggested they could learn

Hebb (1949) – Simple updating rule for learning

Rosenblatt (1962) - the *perceptron* model

Minsky and Papert (1969) – wrote *Perceptrons*

Bryson and Ho (1969, but largely ignored until 1980s--Rosenblatt) – invented back-propagation learning for multilayer networks

61

Training the perceptron

First wave in neural networks in the 1960's

Single neuron

Trainable: its threshold and input weights can be modified

If the neuron doesn't give the desired output, then it has made a mistake

Input weights and threshold can be changed according to a learning algorithm

62

Examples - boolean operators

AND – if all inputs are 1, return 1, otherwise return 0

OR – if at least one input is 1, return 1, otherwise return 0

NOT – return the opposite of the input

XOR – if exactly one input is 1, then return 1, otherwise return 0

63

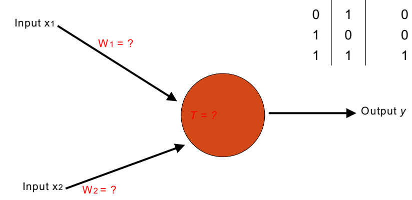
AND

x_1	x_2	x_1 **and** x_2
0	0	0
0	1	0
1	0	0
1	1	1

64

AND

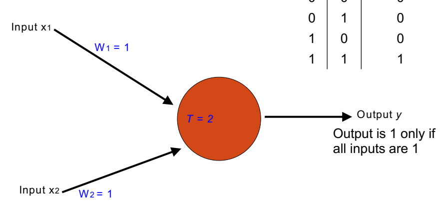
x_1	x_2	x_1 **and** x_2
0	0	0
0	1	0
1	0	0
1	1	1



65

AND

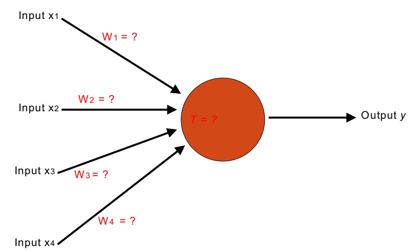
x_1	x_2	x_1 **and** x_2
0	0	0
0	1	0
1	0	0
1	1	1



Inputs are either 0 or 1

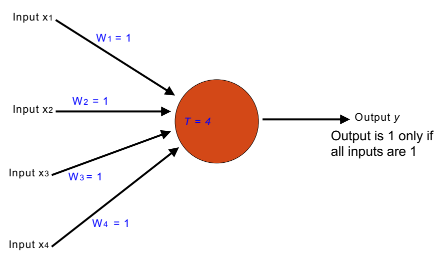
66

AND



67

AND



Inputs are either 0 or 1

68

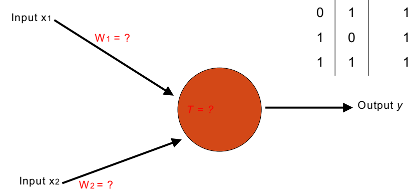
OR

x_1	x_2	x_1 **or** x_2
0	0	0
0	1	1
1	0	1
1	1	1

69

OR

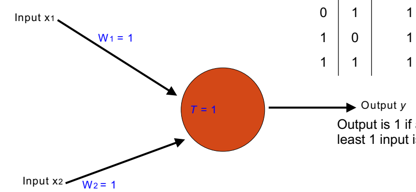
x_1	x_2	x_1 **or** x_2
0	0	0
0	1	1
1	0	1
1	1	1



70

OR

x_1	x_2	x_1 **or** x_2
0	0	0
0	1	1
1	0	1
1	1	1

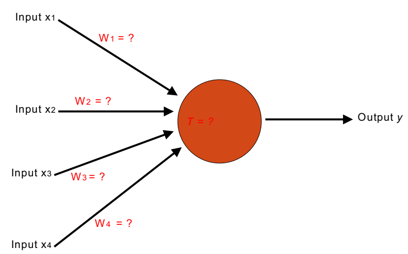


Output is 1 if at least 1 input is 1

Inputs are either 0 or 1

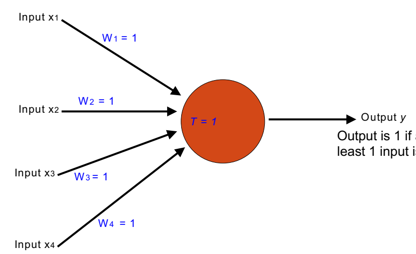
71

OR



72

OR



Output is 1 if at least 1 input is 1

Inputs are either 0 or 1

73

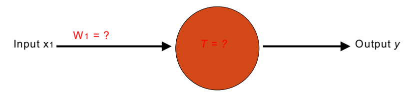
NOT

x_1	**not** x_1
0	1
1	0

74

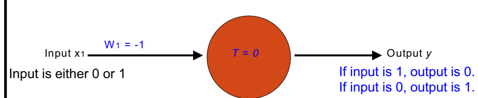
NOT

x_1	**not** x_1
0	1
1	0



75

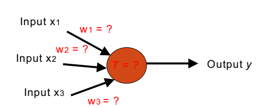
NOT



76

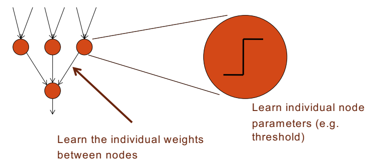
How about…

x_1	x_2	x_3	x_1 **and** x_2
0	0	0	1
0	1	0	0
1	0	0	1
1	1	0	0
0	0	1	1
0	1	1	1
1	0	1	1
1	1	1	0



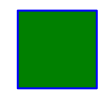
77

Training neural networks



78

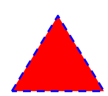
Positive or negative?



NEGATIVE

79

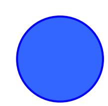
Positive or negative?



NEGATIVE

80

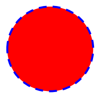
Positive or negative?



POSITIVE

81

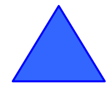
Positive or negative?



NEGATIVE

82

Positive or negative?



POSITIVE

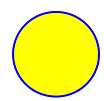
83

Positive or negative?

POSITIVE

84

Positive or negative?



NEGATIVE

85

Positive or negative?

POSITIVE

86

A method to the madness

blue = positive

yellow triangles = positive

all others negative

How did you figure this out (or some of it)?

87

Perceptron learning

A few missing details, but not much more than this

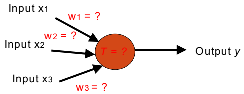
Keeps adjusting weights as long as it makes mistakes

If the training data is linearly separable the perceptron learning algorithm is guaranteed to converge to the "correct" solution (where it gets all examples right)

88

Training neural networks

x_1	x_2	x_3	x_1 **and** x_2
0	0	0	1
0	1	0	0
1	0	0	1
1	1	0	0
0	0	1	1
0	1	1	1
1	0	1	1
1	1	1	0



1. start with some initial weights and thresholds
2. show examples repeatedly to NN
3. update weights/thresholds by comparing NN output to actual output

89

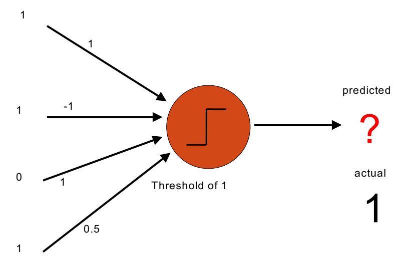
Perceptron learning algorithm

repeat until you get all examples right:

- for each "training" example:
  - calculate current prediction on example
  - if *wrong*:
    - update weights and threshold towards getting this example correct

90

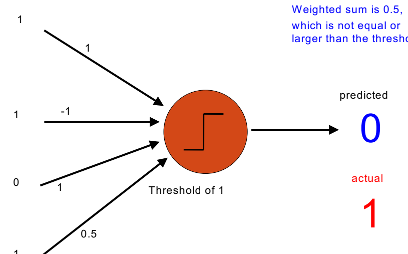
Perceptron learning



91

Perceptron learning

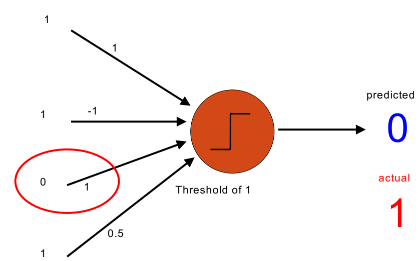
Weighted sum is 0.5, which is not equal or larger than the threshold



What could we adjust to make it right?

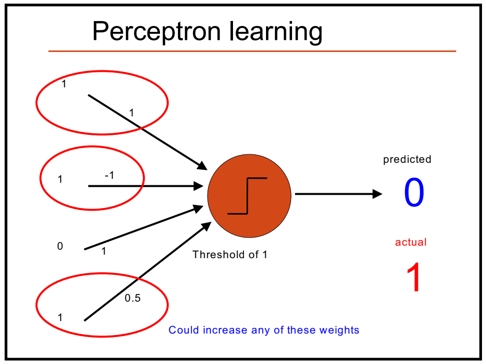
92

Perceptron learning

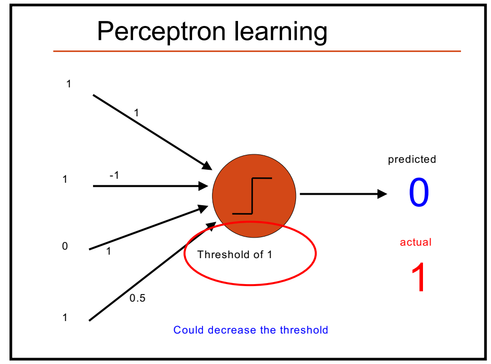


This weight doesn't matter, so don't change

93


Perceptron learning
1
1
1
-1
0
1
1
0.5
Threshold of 1
predicted
0
actual
1
Could increase any of these weights

94


Perceptron learning
1
1
1
-1
0
1
1
0.5
Threshold of 1
predicted
0
actual
1
Could decrease the threshold

95