HUDK 4051: LEARNING ANALYTICS: PROCESS & THEORY

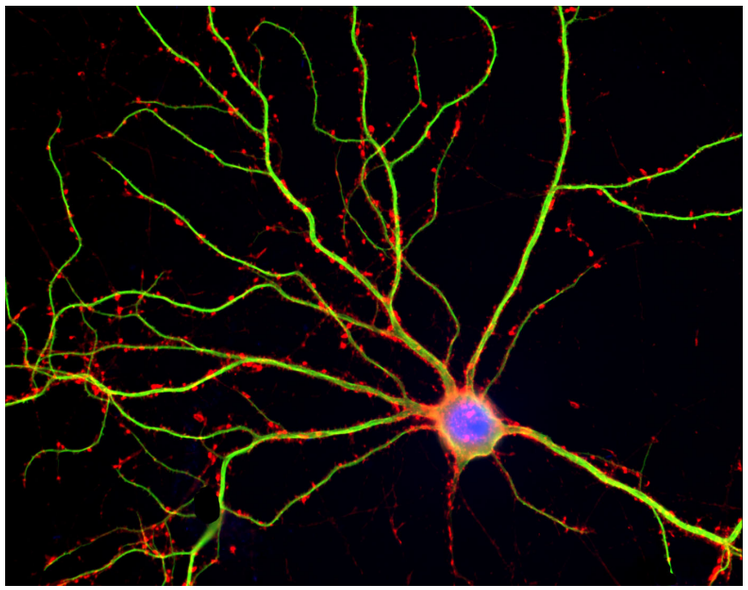
4/9/18 5:26 PM

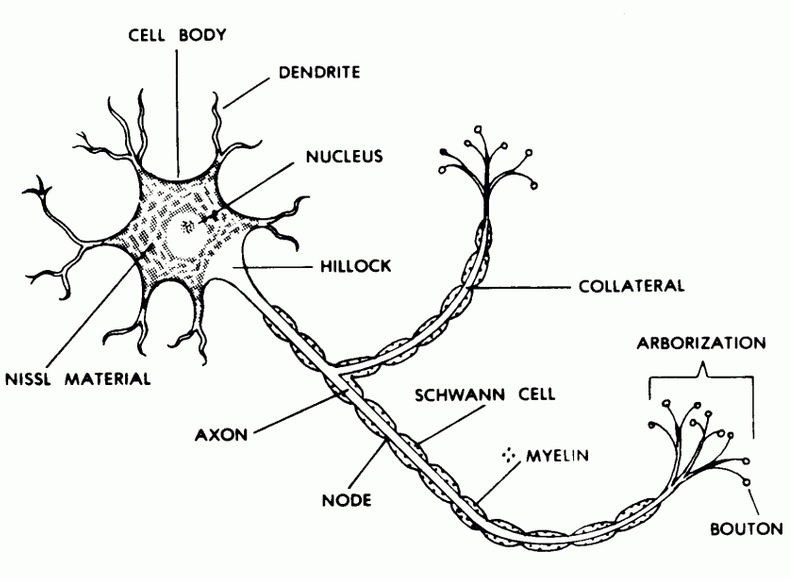
Today

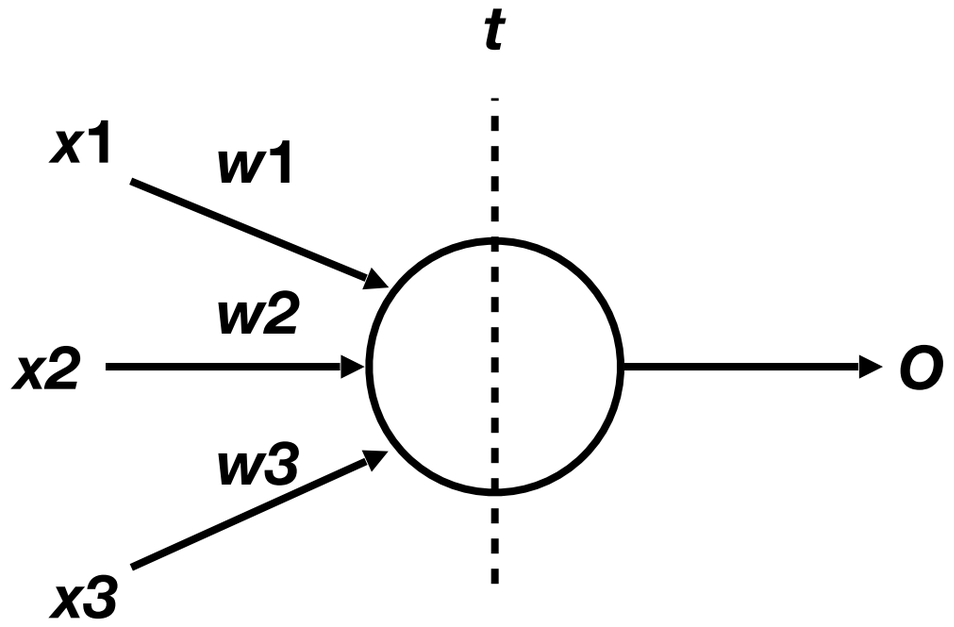
- Artificial Neural Networks
  - Perceptron
  - Sigmoid Function
  - Back propagation
- Work time

Perceptron

Perceptron

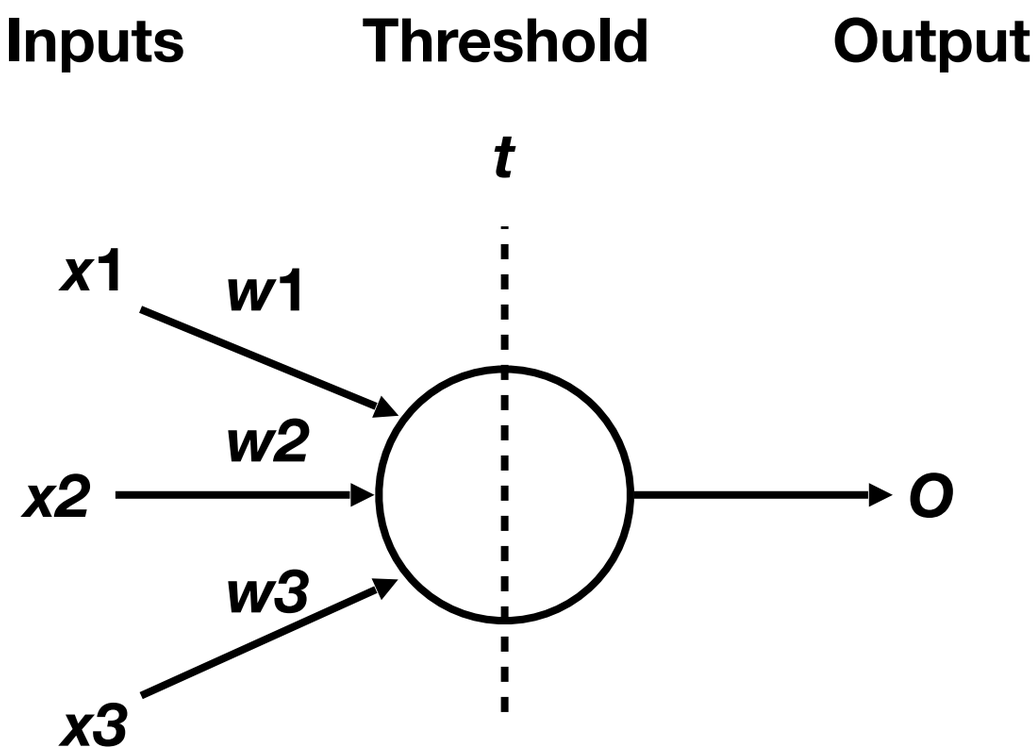




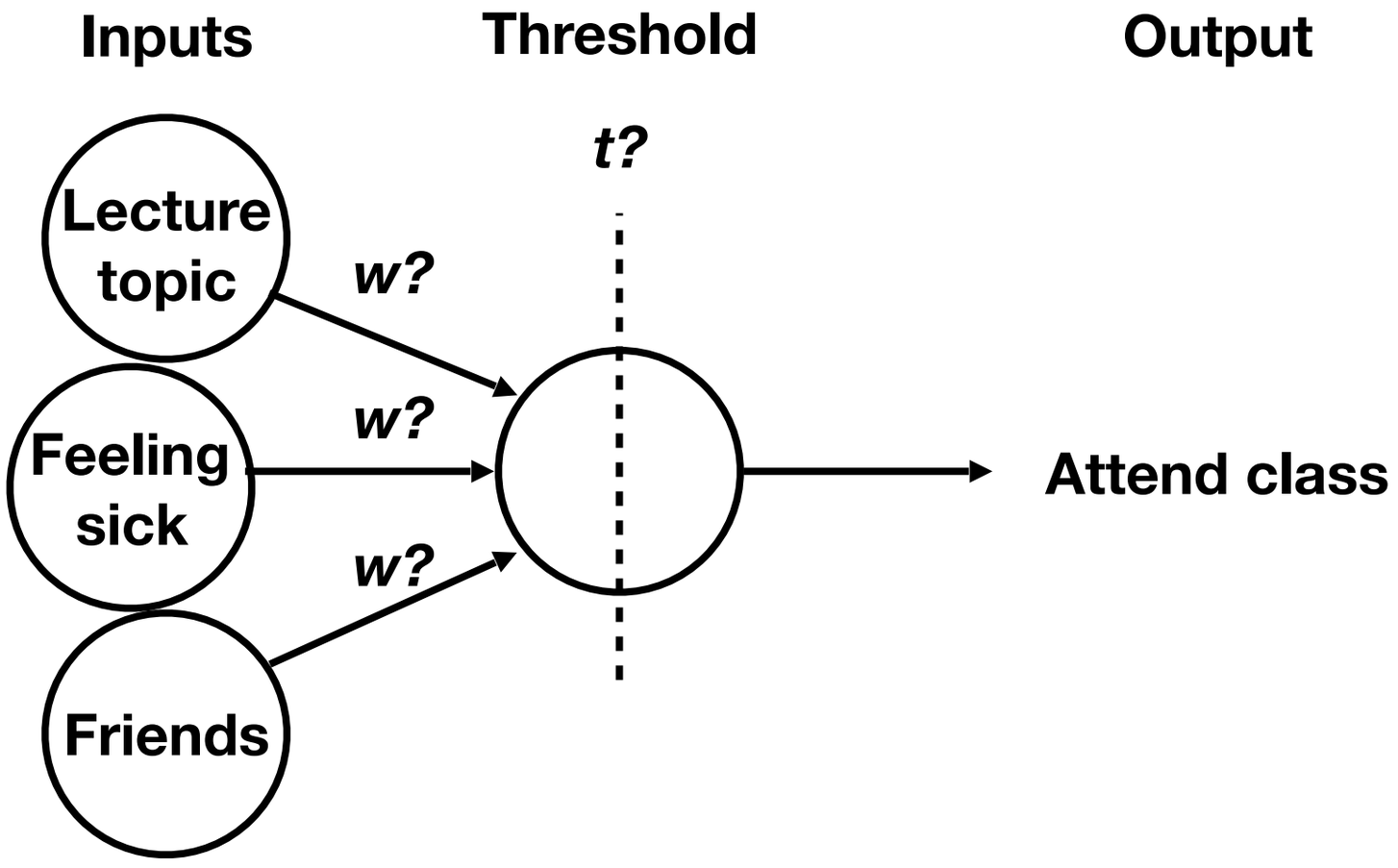


Frank Rosenblatt, 1957

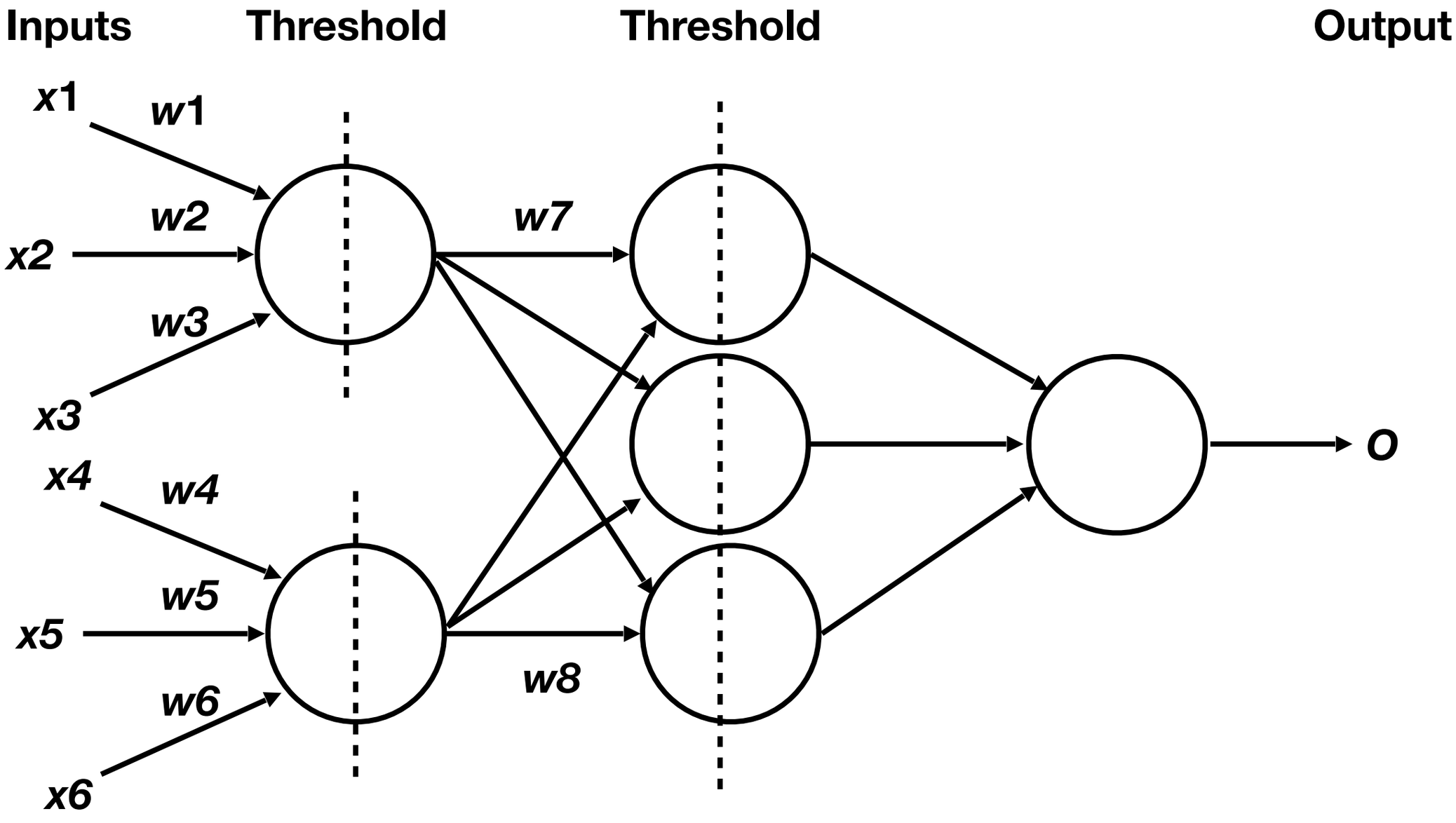
Perceptron



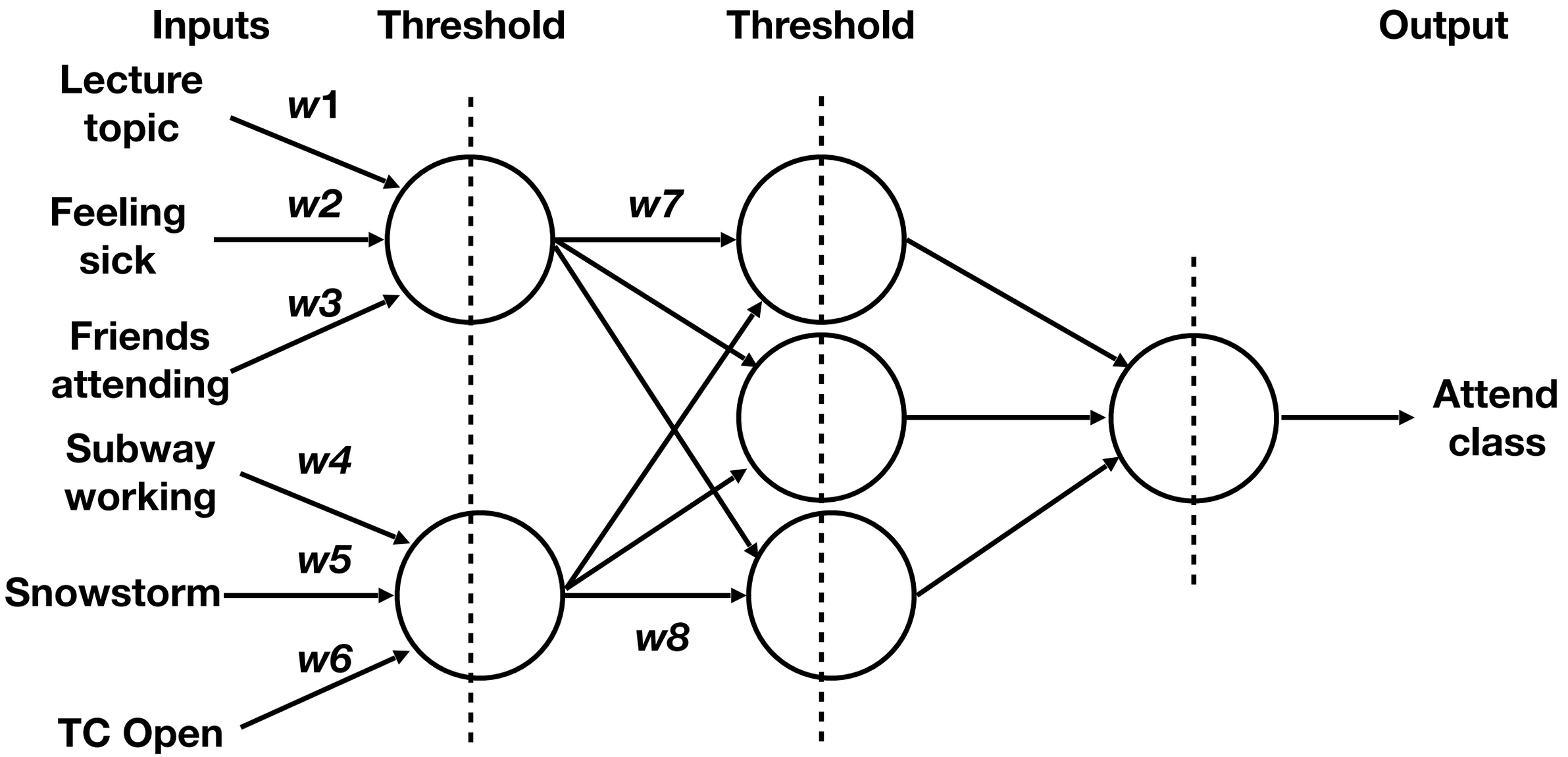
Perceptron


Inputs
Threshold
Output
t?
Lecture topic
w?
Feeling sick
w?
w?
Friends
Attend class

Perceptron


Inputs
Threshold
Threshold
Output
x1
w1
w2
x2
w7
w3
x3
x4
w4
w5
x5
w8
w6
x6
O

Perceptron


Inputs
Threshold
Threshold
Output
Lecture topic
Feeling sick
Friends attending
Subway working
Snowstorm
TC Open
w1
w2
w3
w4
w5
w6
w7
w8
Attend class

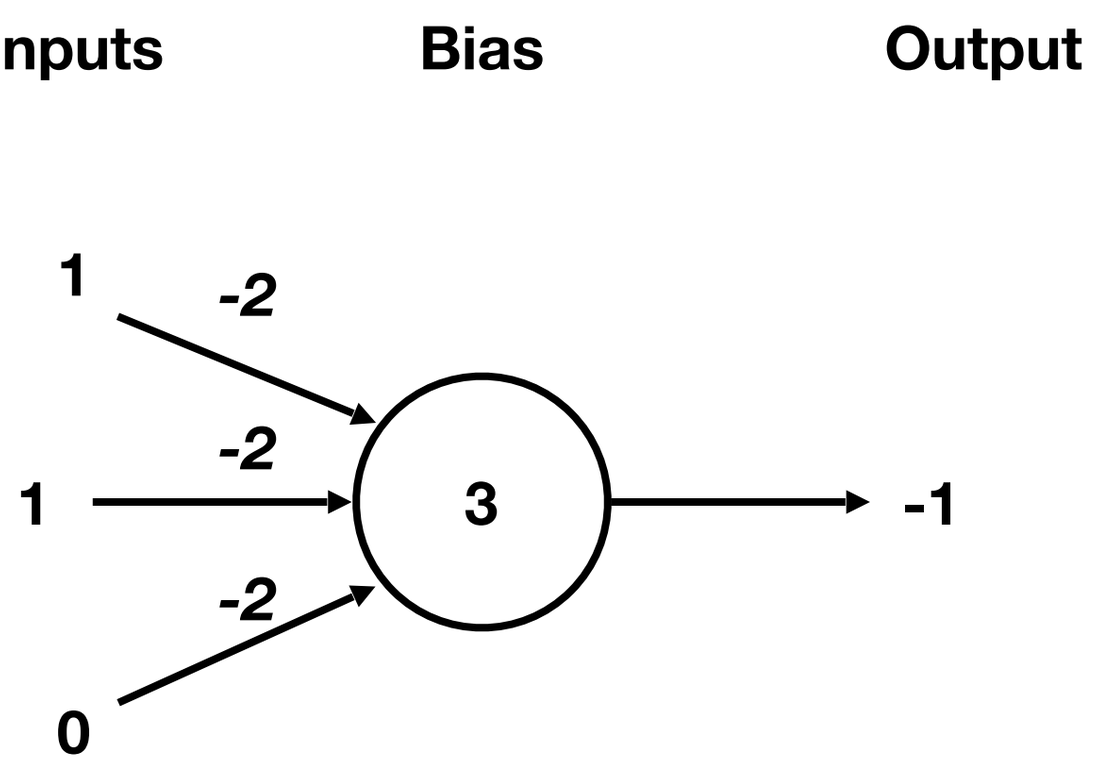
Logic

- From the perceptron we can create a NAND gate
- From a NAND gate can create all other logic units (AND, NOR, etc.)
- See Nielson 2016*

Bias

- Another way to describe the threshold
- Negative threshold
- More convenient for notation
- Describes how easy it is to get make the perceptron “fire”

Notation



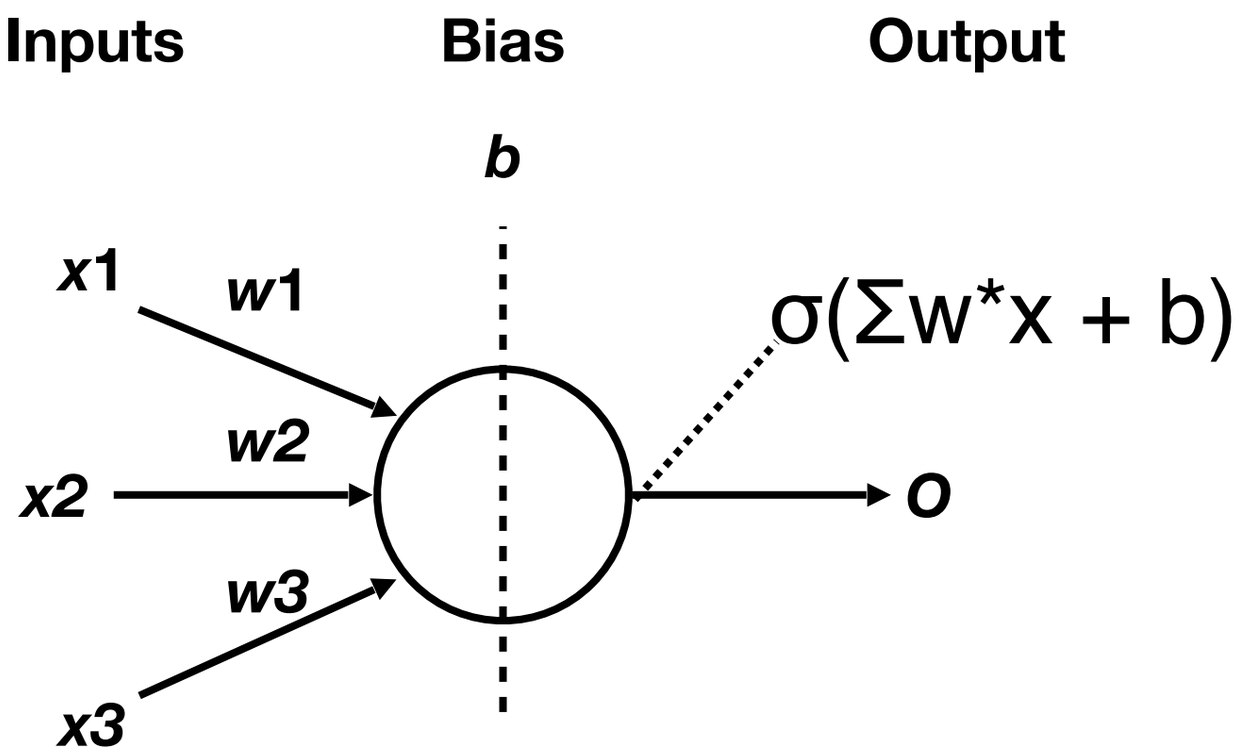
(1)*-2 + (1)*-2 + (0)*-2 + 3 = -1

Sigmoid Neurons

Sigmoid Neurons

- Want to build a learning algorithm
- Could change b or w
- BUT that will cause very large changes
- Network will never “fix”
- Solution: “smooth” the output

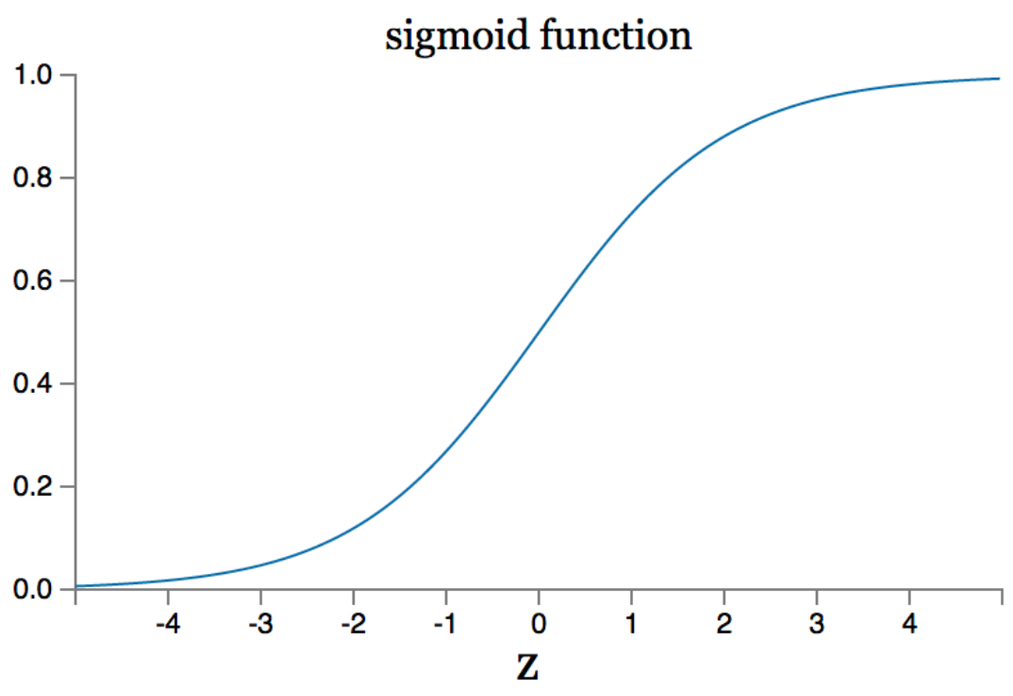
Sigmoid Neurons



Sigmoid Neurons

- Sigmoid function "smooths" the output
- Makes changing w *and* b less sudden and more predictable
- Could use lots of other functions…

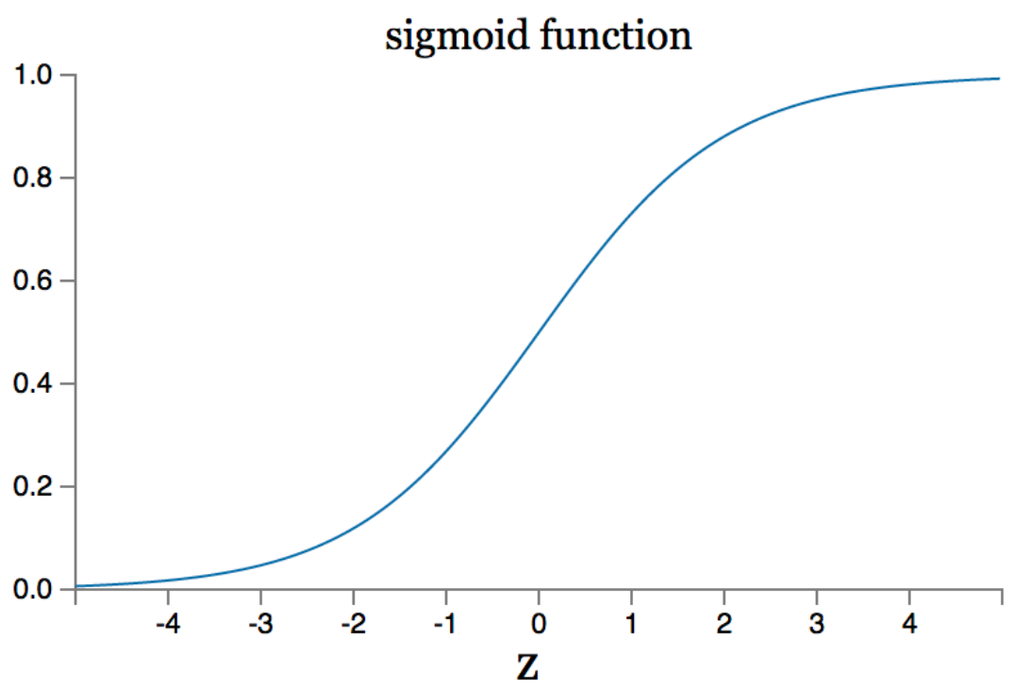
$$\sigma(z) \equiv \frac{1}{1+e^{-z}}.$$



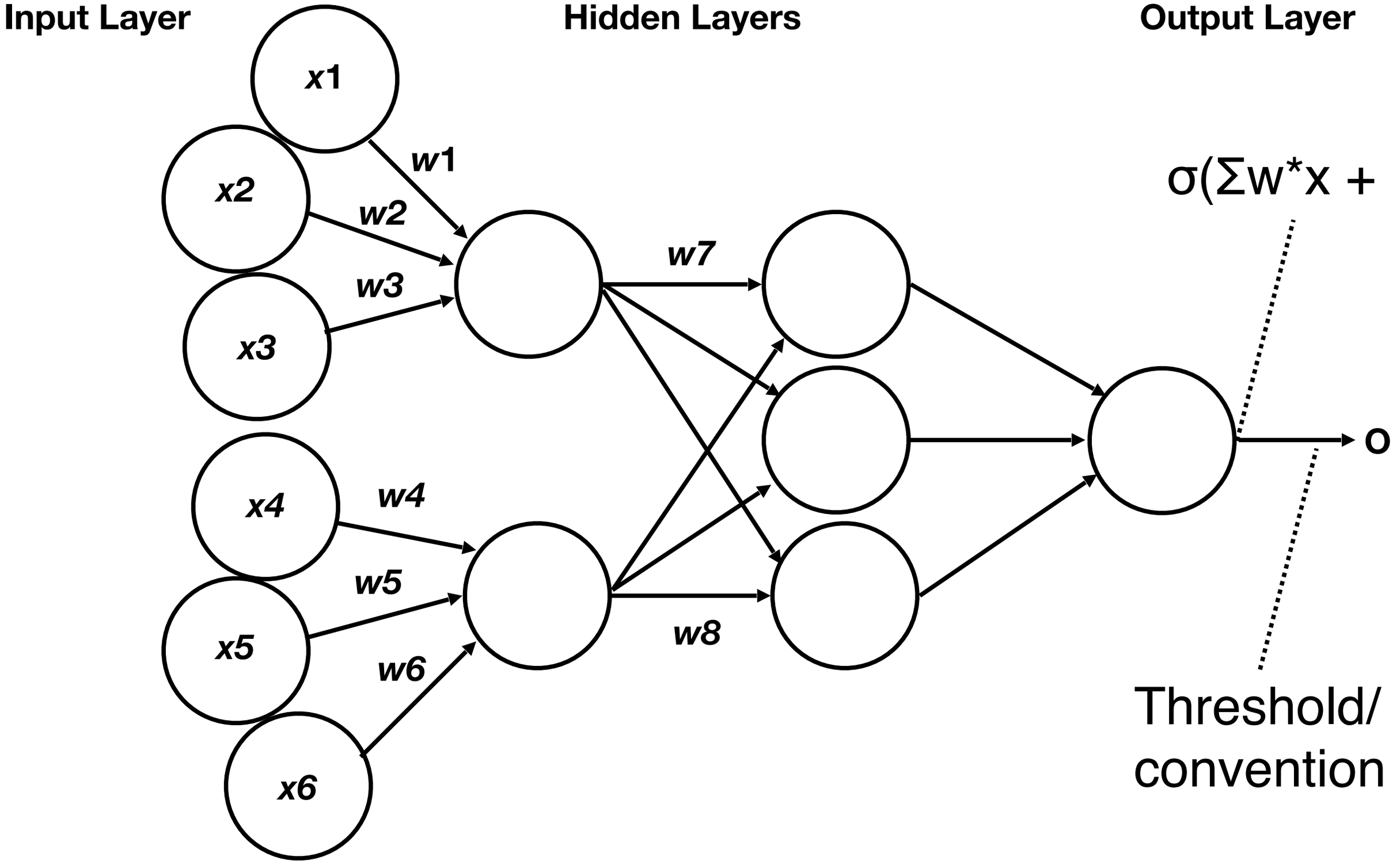
Sigmoid Neurons

- Perceptrons have 0/1 output
- Sigmoid neurons have 0 - 1 output (eg. 0.1, 0.6778, etc.)
- How to interpret sigmoid neuron output?

$$\sigma(z) \equiv \frac{1}{1+e^{-z}}.$$



Complete Feedfoward Network



How many Hidden Layers?

- No foolproof method
- The only method is really trial and error
- Heuristics:
  - Theory based starting point?
  - Number of inputs and outputs?

Exercise

Back Propagation

- Need a way to minimize error
- Error is defined by a cost function
- Then we imagine error as a surface that needs to be “searched” for the minimum