Intro to Machine Learning

Josh Wolfe

February 23, 2018

What problems can machine learning solve?

- Computer vision
- Audio and text analysis
- Traffic predictions
- Social media and product recommendations
- Spam filtering

Types of machine learning

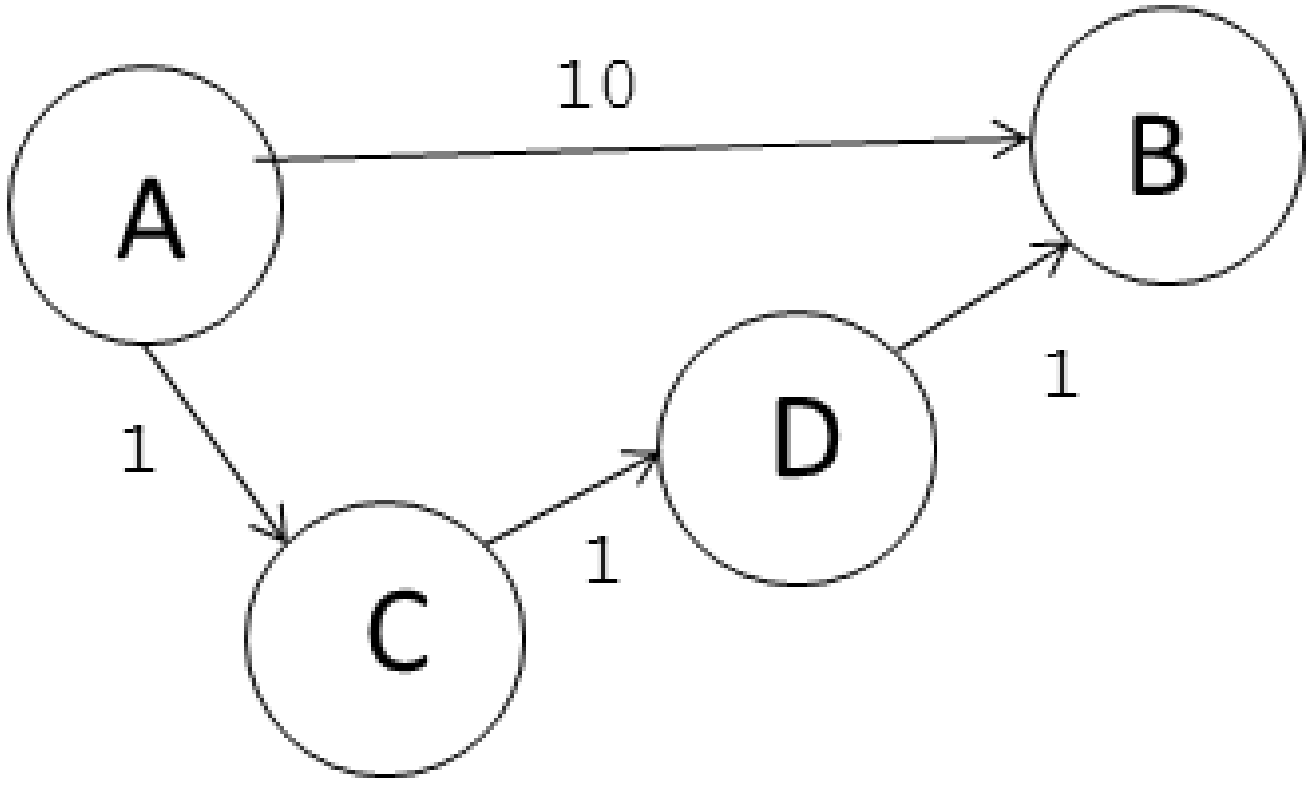
Supervised

- More commonly used
- Training data is labeled ahead of time

Unsupervised

- No labeling is provided
- Tries to find commonalities in the data
- Often used for identifying patterns in transaction data

Graph data structure



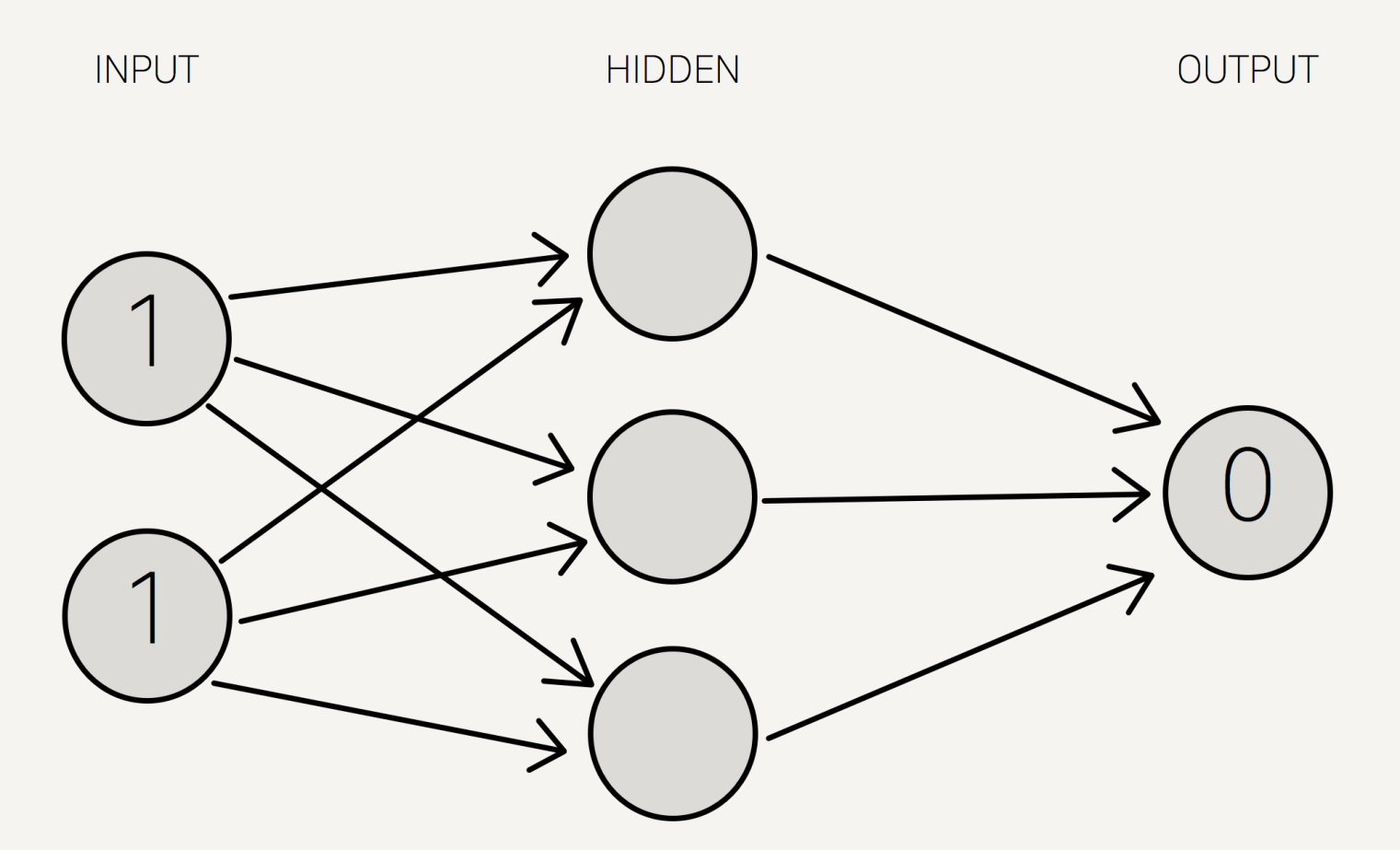
- Directed
- Acyclic
- Simple
- Weighted

Simple example - The XOR operation

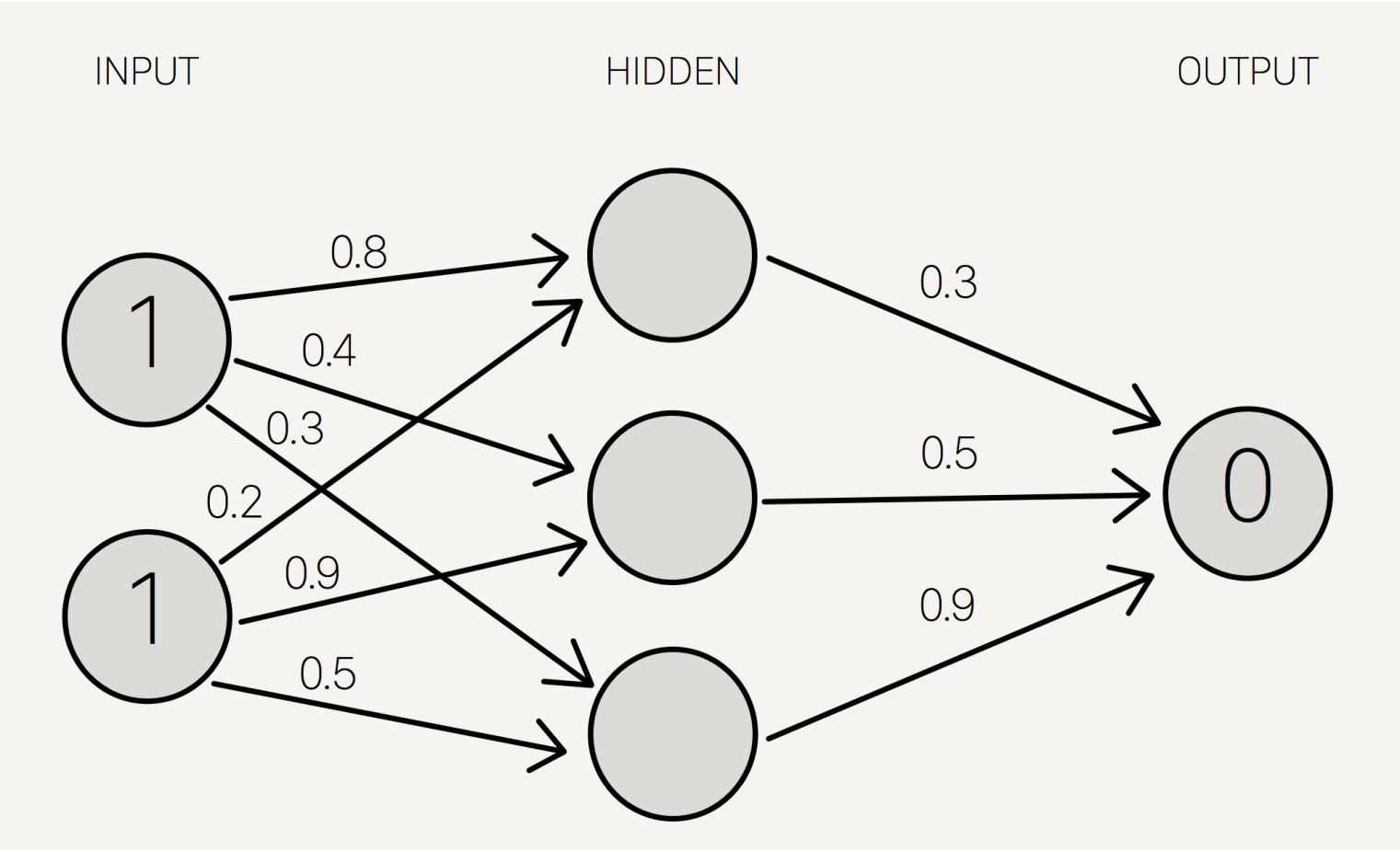
A	B	Output
0	0	0
0	1	1
1	0	1
1	1	0

- Has well defined inputs and outputs
- Very simple

XOR as a graph



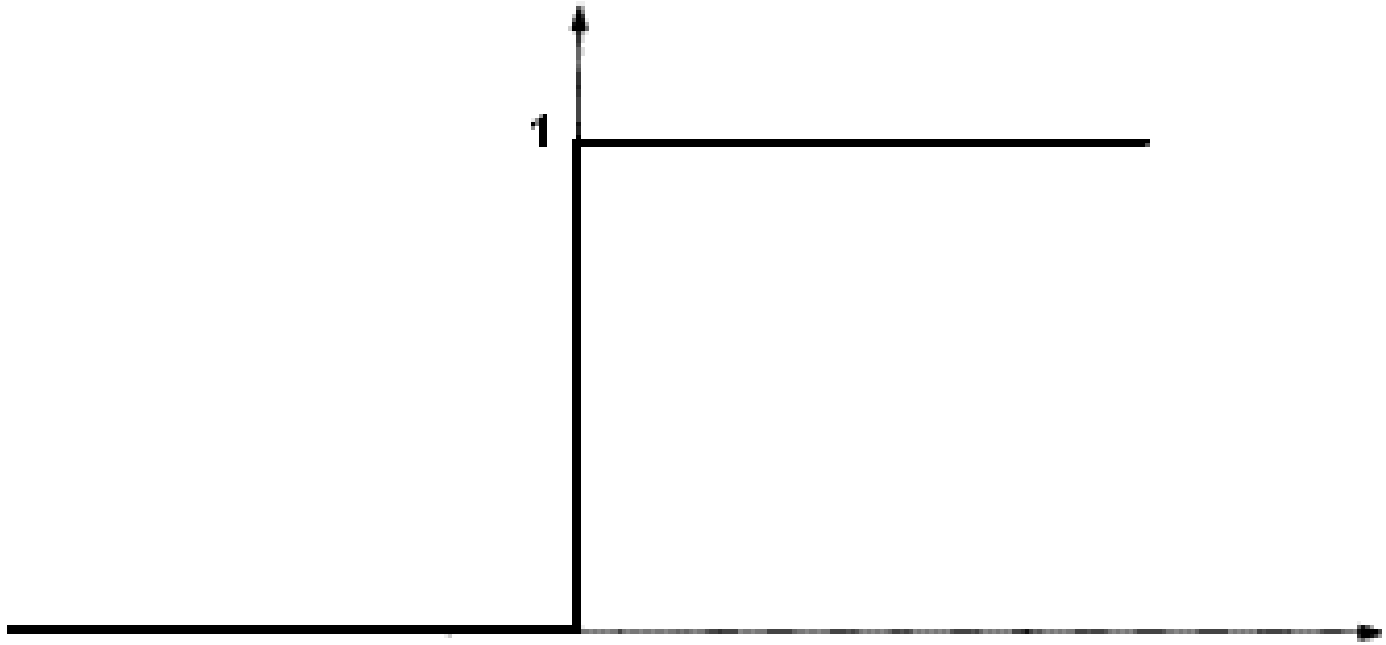
Graph weightings



Randomly chosen between 0 and 1 based on a gaussian distribution

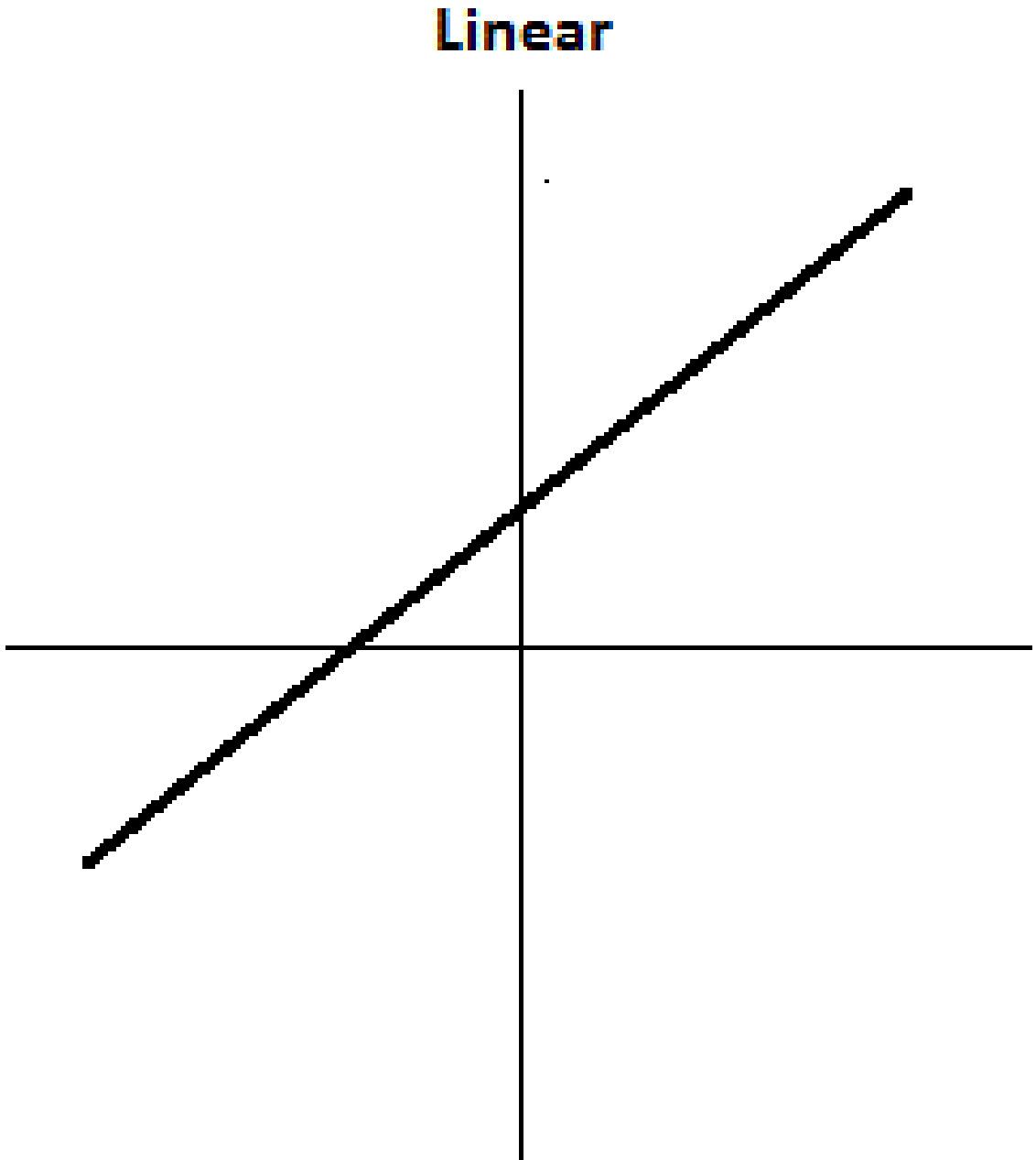
Activation functions

- The step function



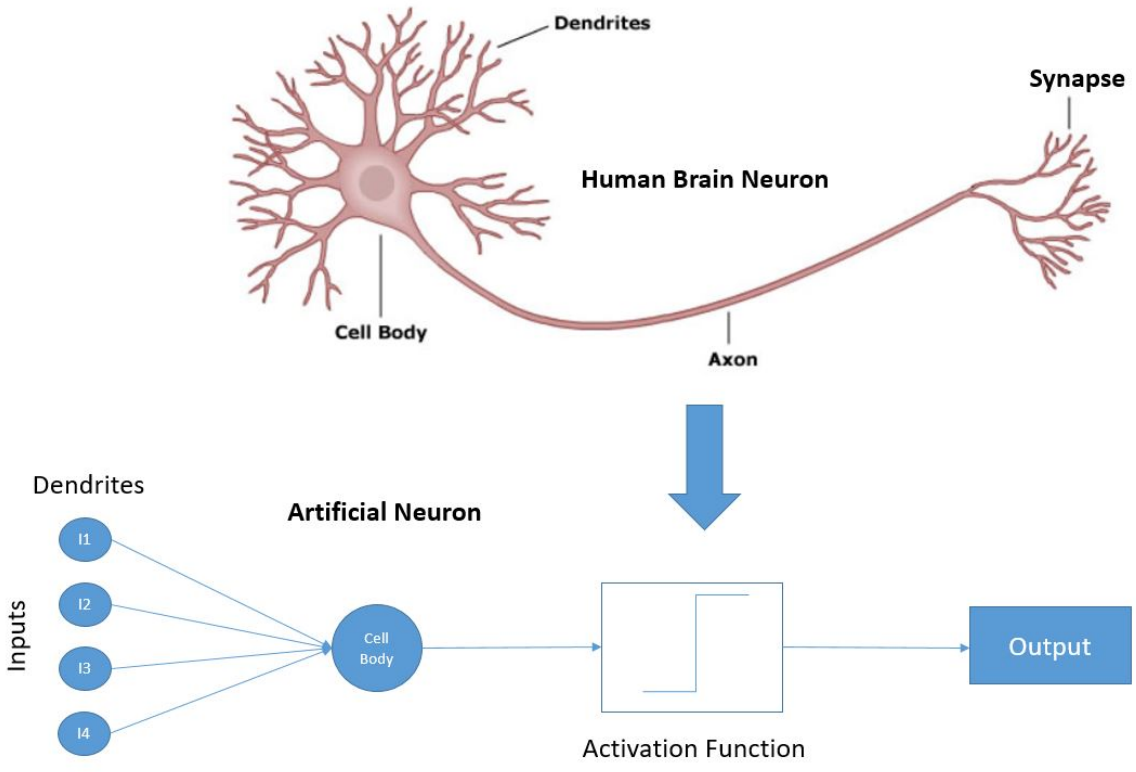
Activation functions cont.

- Linear function



Human neurual biology

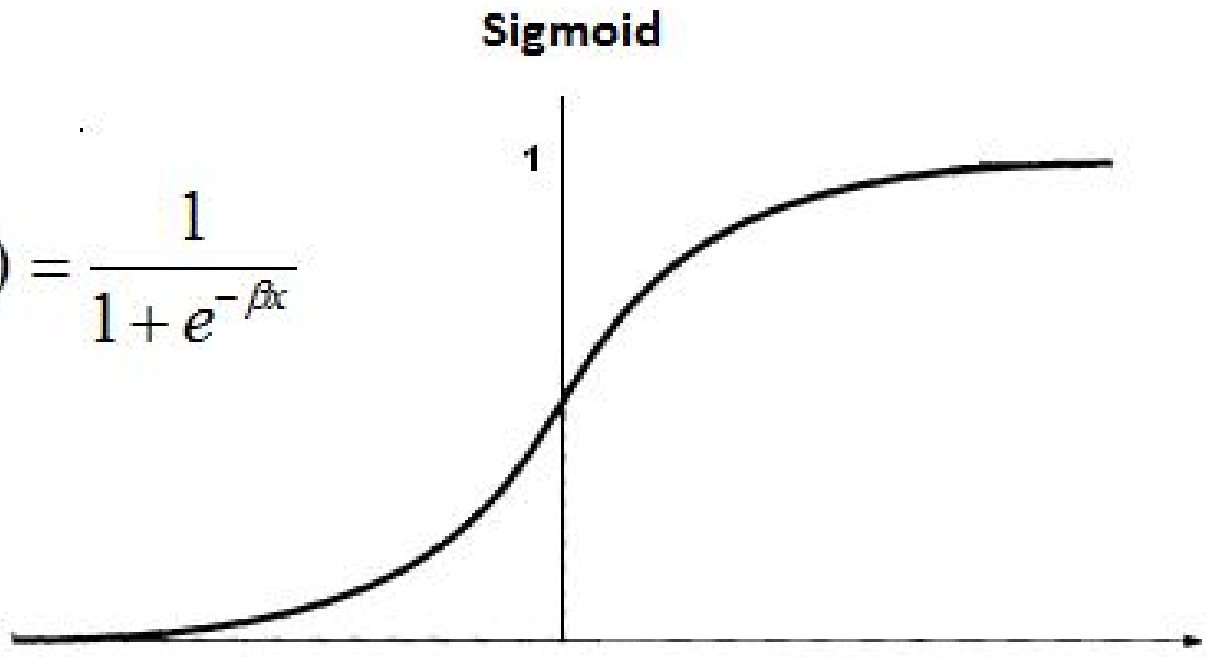
- Machine learning is inspired by the structure of our brains
- Each node in our graph represents a neuron
- Our activation functions represent a neuron firing
- Neurons don't fire in a step-wise or linear manner



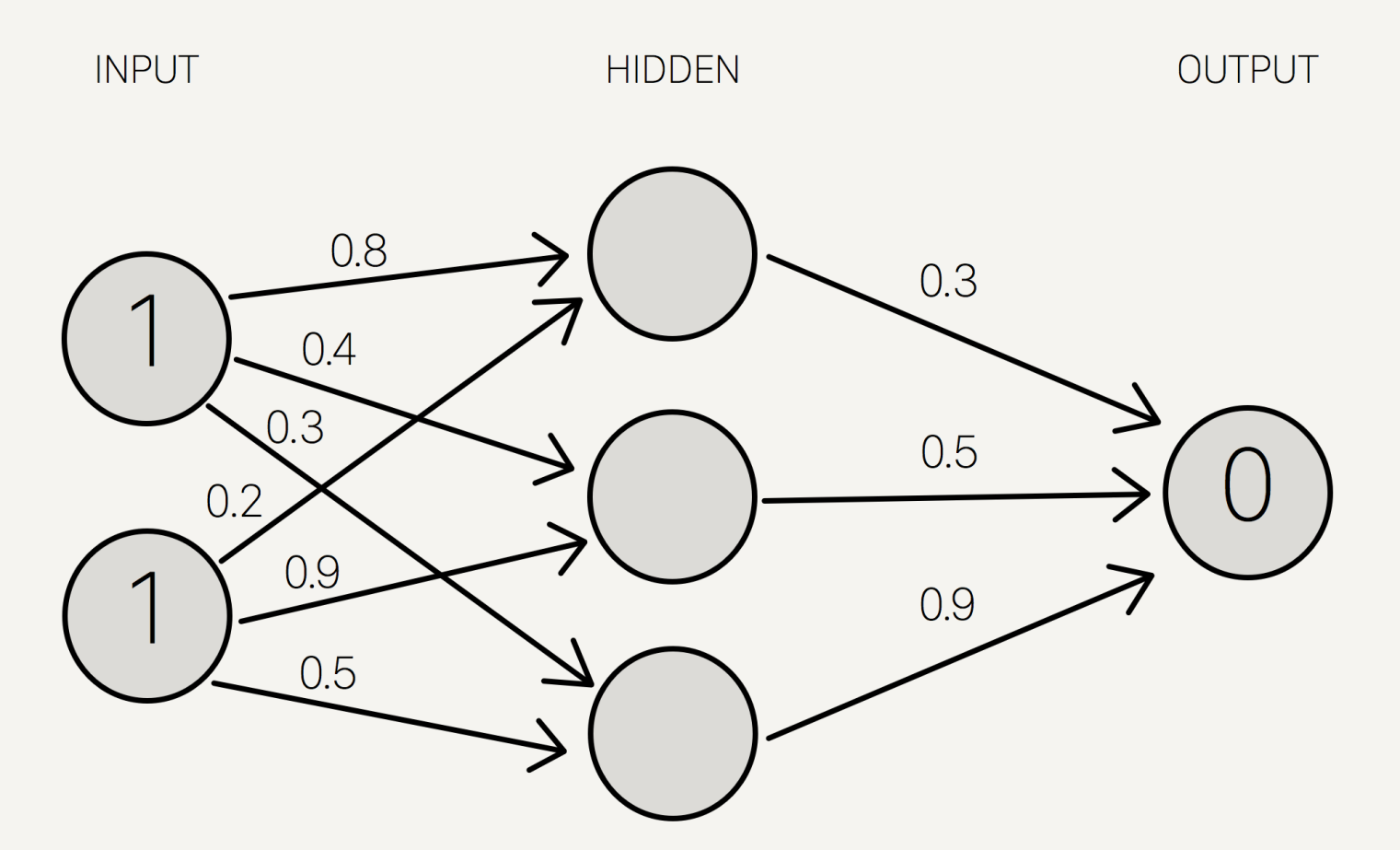
Activation functions cont.

- Sigmoid function

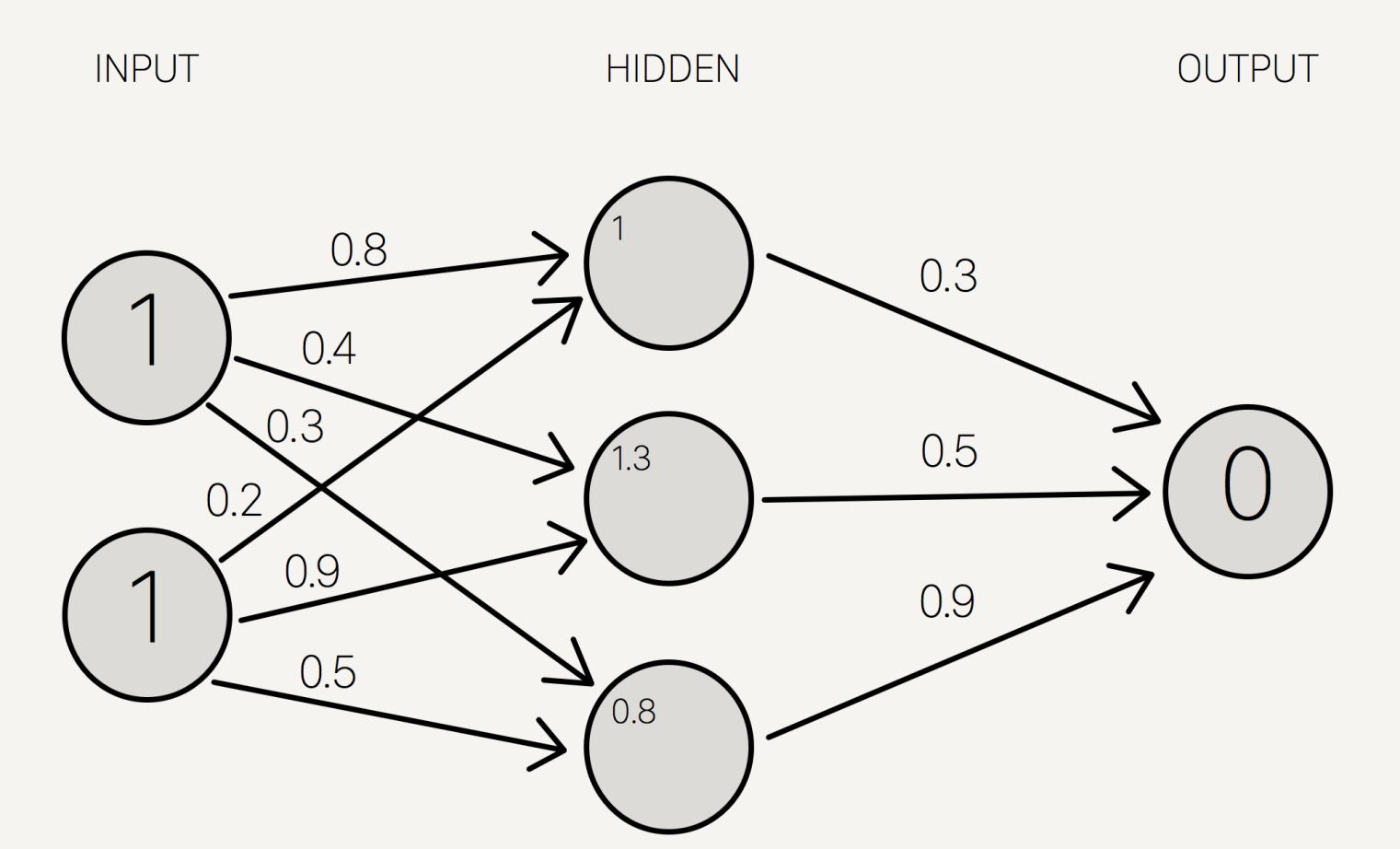
$$f(x) = \frac{1}{1 + e^{-\beta x}}$$



Back to XOR

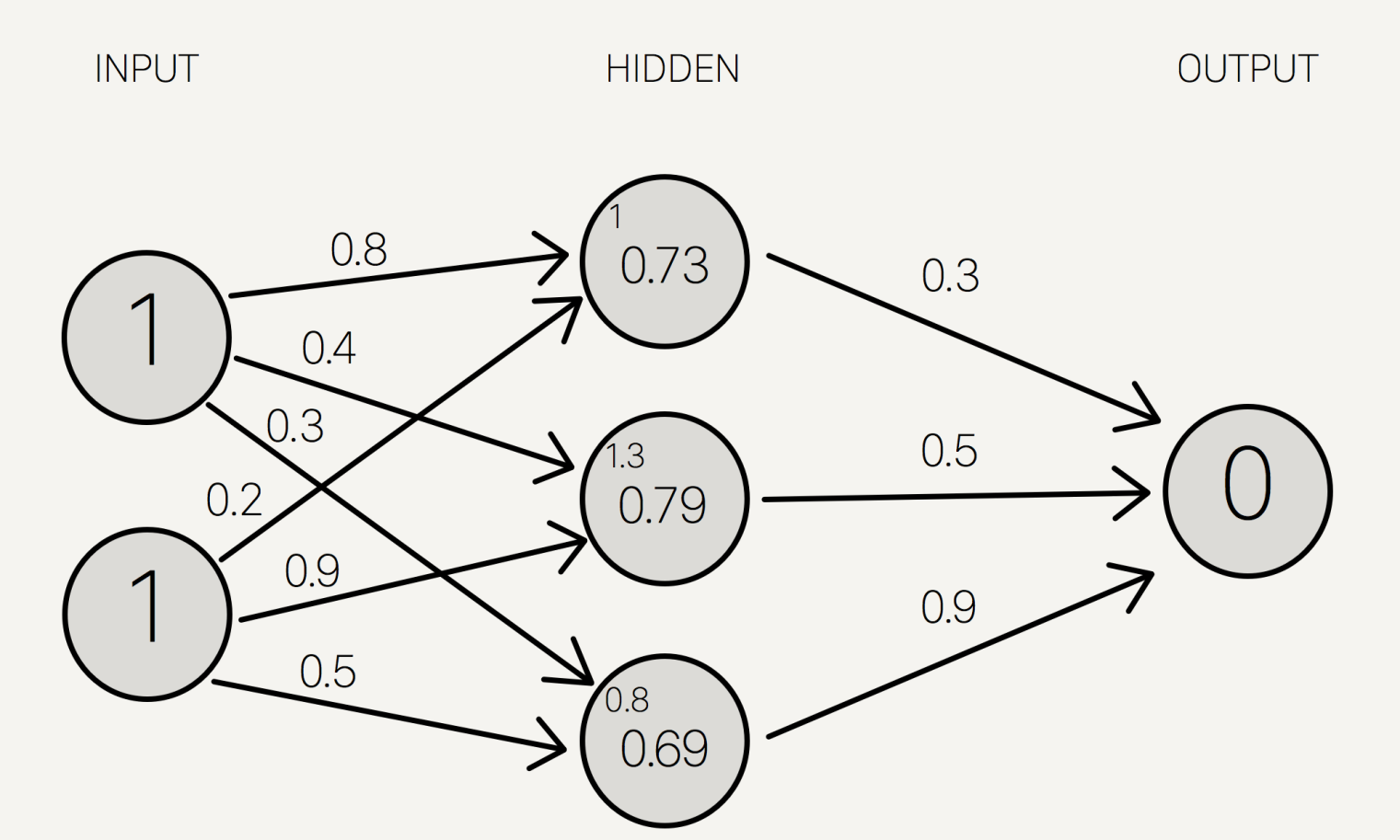


Back to XOR cont.



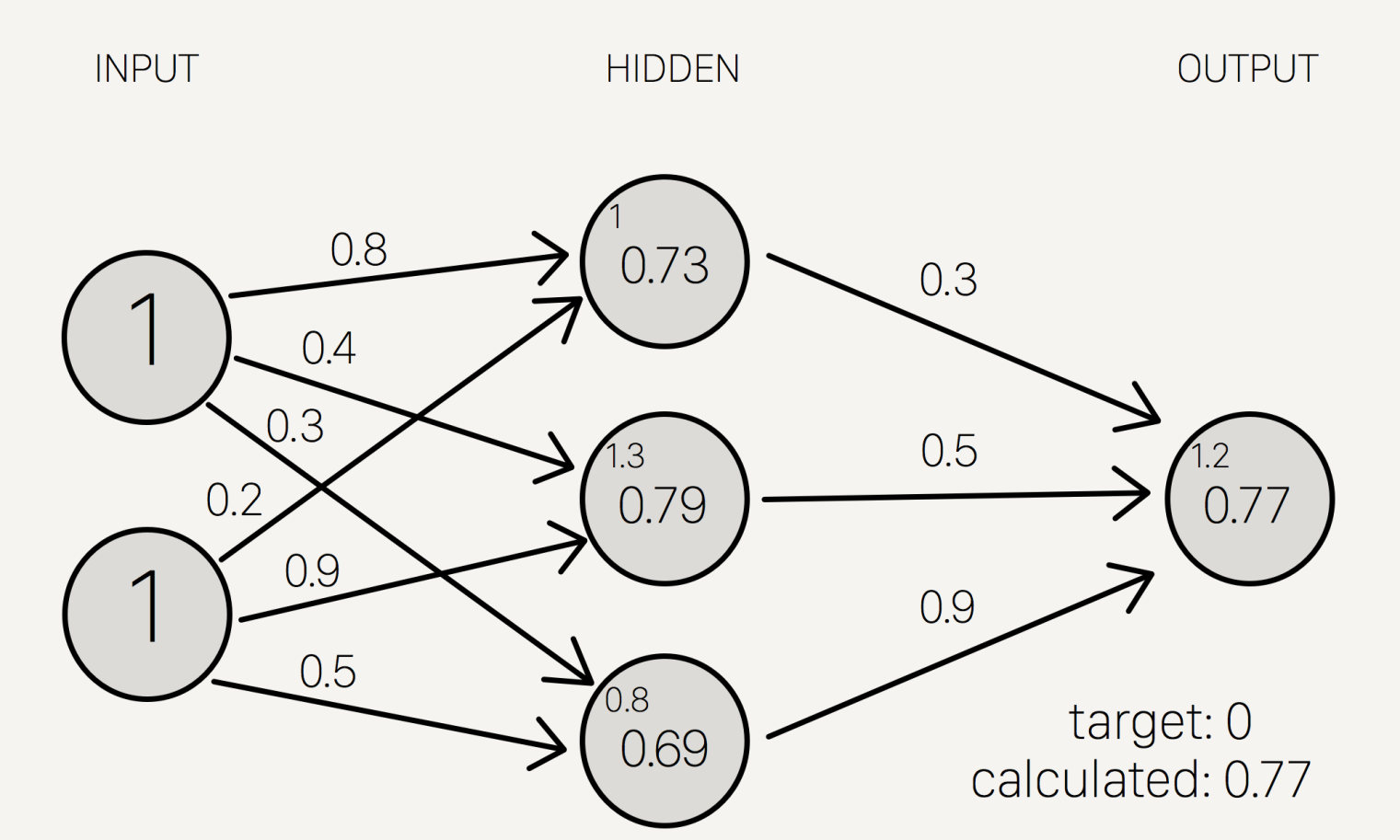
- Sum our input values multiplied by their weights

Back to XOR cont.



- Apply our activation function

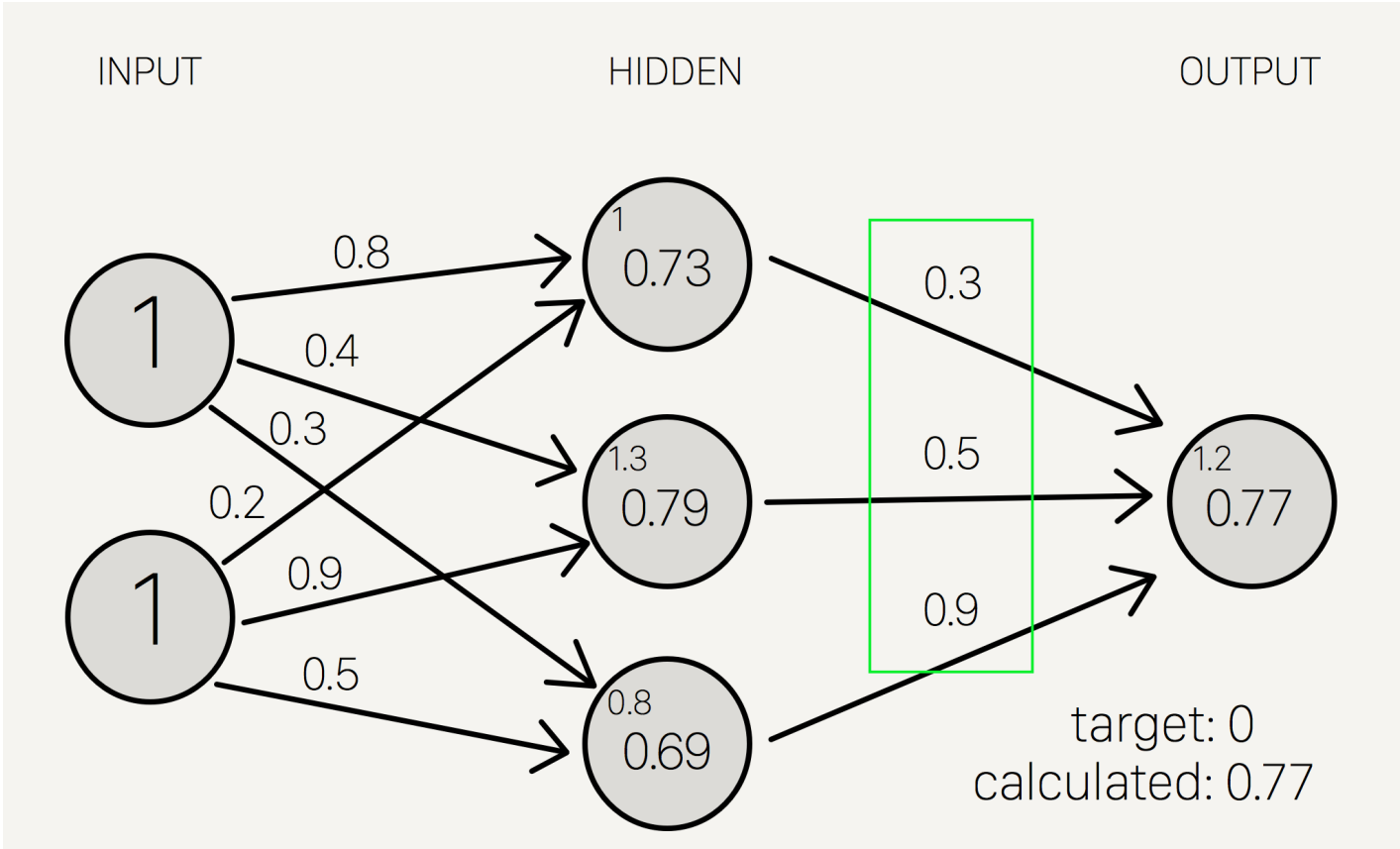
Back to XOR cont.



- Repeat for the next layer

Teaching the machine aka Back propagation

- 0.77 is a fair bit off from 0
- We need to adjust the weightings in our graph
- We'll work our way backwards

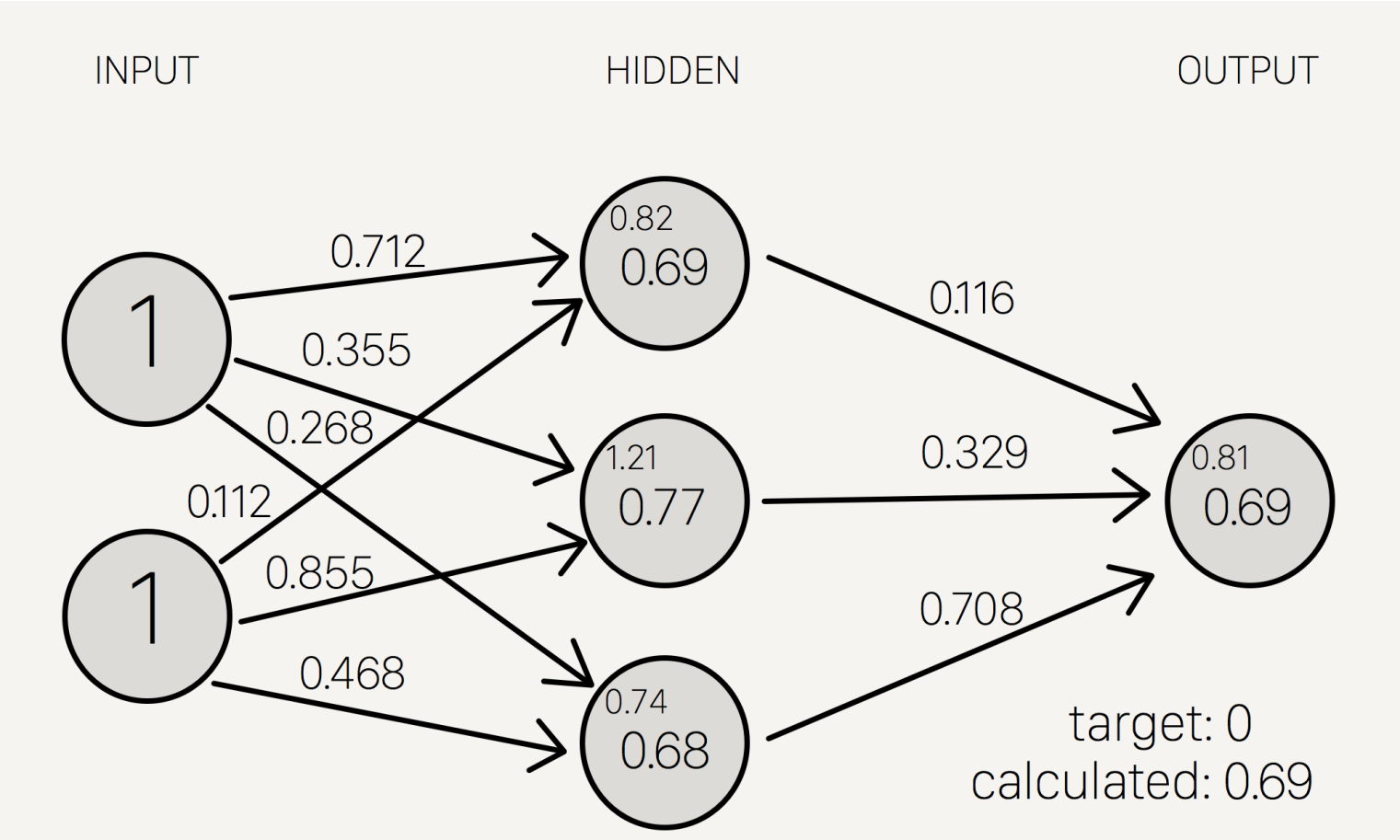


Teaching the machine aka Back propagation cont.

1. We calculate our "Loss function" i.e. exactly how wrong we were using something like the sum of square errors of all the layers.
2. We use a bit of calculus to minimize our error function.
3. We adjust our weights in the direction that helps minimize error.

Old	Adjusted
0.8	0.712
0.4	0.3548
0.3	0.2681
0.2	0.112
0.9	0.8548
0.5	0.4681
0.3	0.1162
0.5	0.329
0.9	0.708

Testing with adjusted weights



- We've improved our guess!
- We're still pretty far off from our correct output

Training

- In this simple example we only have four possible inputs so it shouldn't take too long before our network starts making good guesses.
- For more complicated problems you might run thousands of different inputs.

Resources

- Tensorflow
  - Made by the Google Brain research team
  - One of the most popular ML libraries
  - Python library
  - tensorflow/tensorflow on GitHub
- Caffe2
  - Made by the Facebook research team
  - Python & C++ libraries available
  - caffe2/caffe2 on GitHub
- Keras
  - Open source project
  - Wrapper for Tensorflow
  - Designed to be as simple to use as possible to allow for rapid prototypes and experimentation