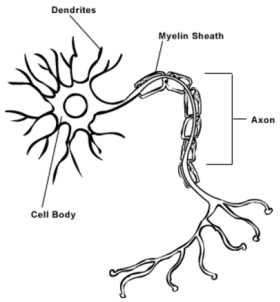
Intro to neural nets

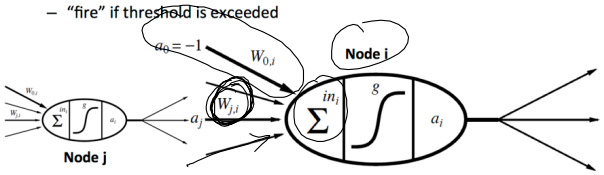
Friday, September 30, 2016 8:49 AM

Neural Networks

- Neuron
- Brain information processing emerges from networks of neurons



- McCulloch & Pitts (1943)
  - Linear combination of inputs
  - "fire" if threshold is exceeded



Nodes are neurons

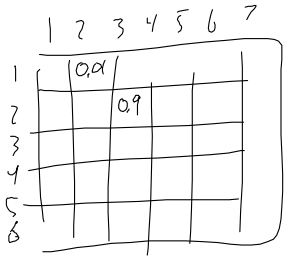
There is a link from j to i

j sends a signal of strength a_j

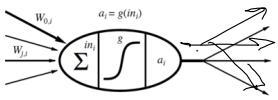
i receives it with weight $W_{j,i}$

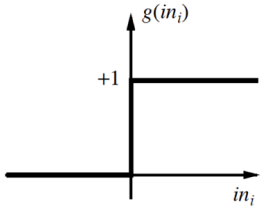
Looks like a matrix lookup

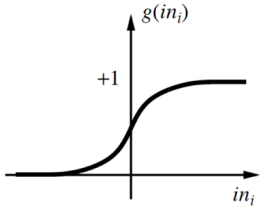
Additionally, each node has a bias, a_0, $W_{0,i}$



- Input to a node: $in_i = \sum_{j=0}^{n} W_{j,i} a_j$
- Output of a node: ~~$a_i = g(in_i)$~~
- What is g?
  - A function that computes near 1.0 when the "right" inputs are given and computers near 0.0 when the "wrong" inputs are given
  - $W_{0,i}$ sets the threshold – actual inputs must overcome bias

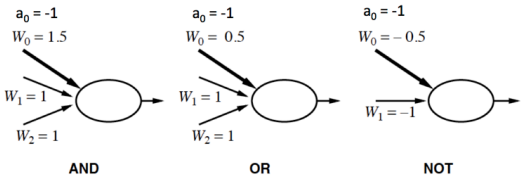






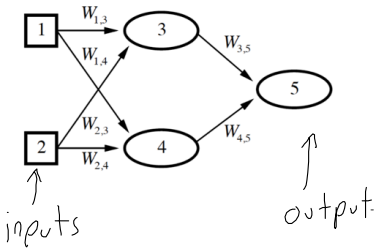
Comparison to logic

- Can replicate logic gates with nodes
- Can compute any boolean logic statement with neural network



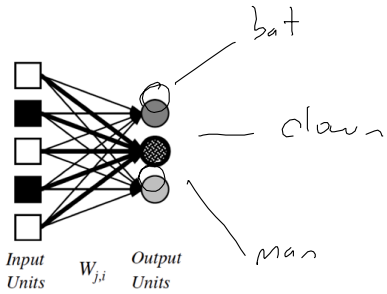
Network structures

- Feed-forward network
  - Represents a function of current inputs
  - No internal state other than weights
  - Output is the result of the function

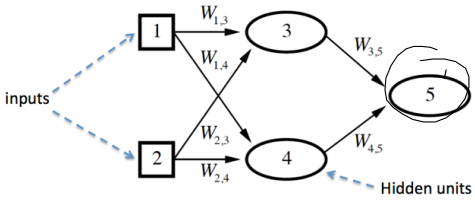


Network structures

- Multiple outputs



A simple network

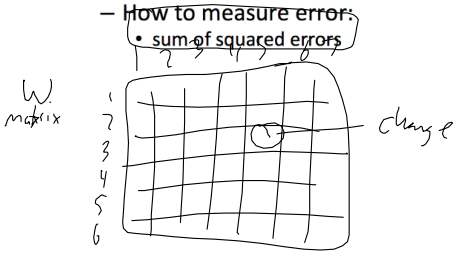


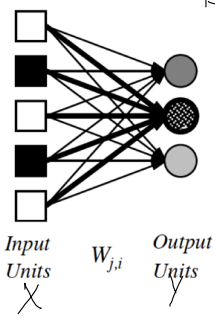
$$a_5 = g(W_{3,5} \cdot a_3 + W_{4,5} \cdot a_4)$$
$$= g(W_{3,5} \cdot g(W_{1,3} \cdot a_1 + W_{2,3} \cdot a_2) + W_{4,5} \cdot g(W_{1,4} \cdot a_1 + W_{2,4} \cdot a_2))$$

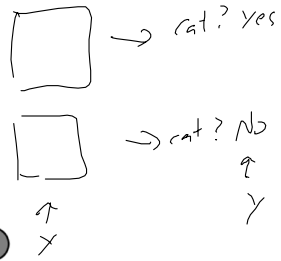
- Adjusting the weights changes the function that the network represents
- This is how learning occurs

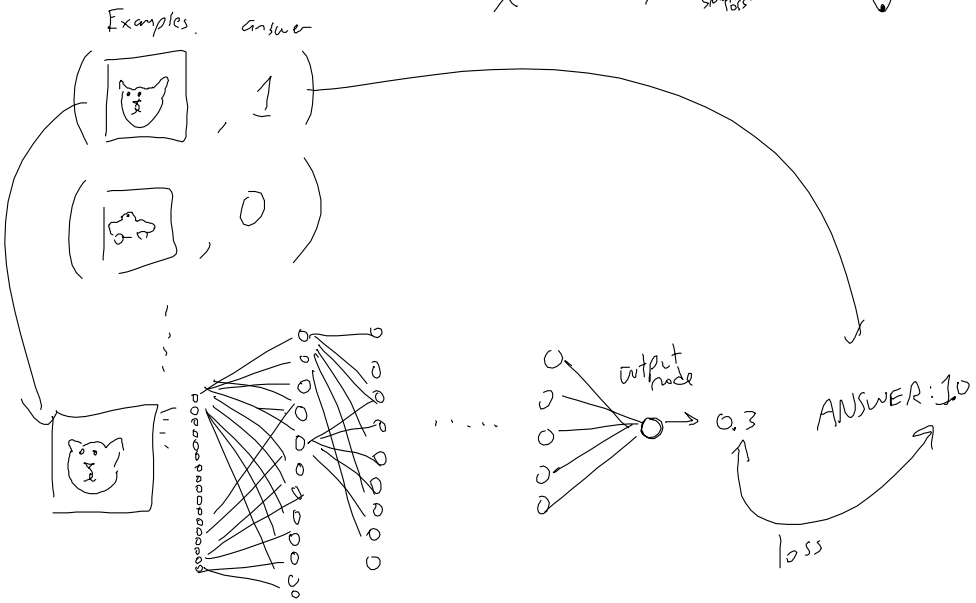
Perceptron learning

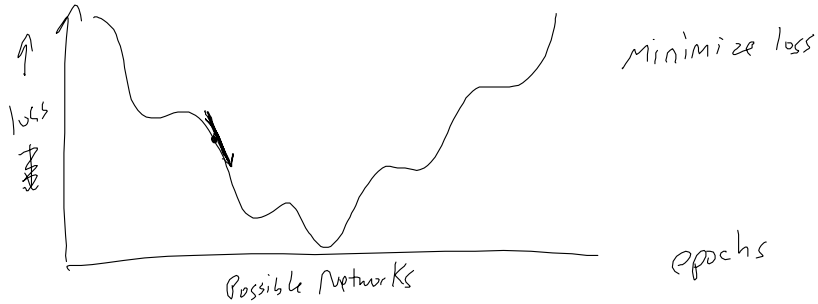
- Want the network to learn to replicate some function
- Adjust the weights of the network to minimize error on the training set
  - Optimization search in weight space
  - How to measure error:
    - sum of squared errors











Minimize loss

continuous problem ⇒ gradient descent

backpropagation

PROS:
- Approx lots of fns
- accurate

CONS:
- Need a lot of data
- Need a lot of compute horsepower

Deep Learning

# of layer

2+

Image Rec ⟶ 100+ layers

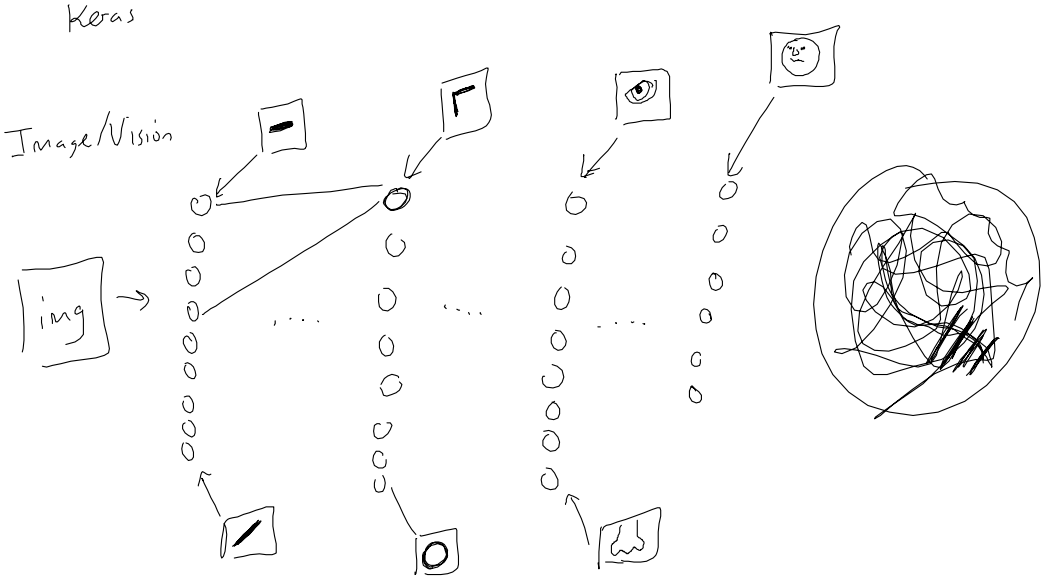
NN are parallizable

Tensorflow

Theano

Keras

Image/Vision



Recurrent Neural Nets (RNN) — time sequences