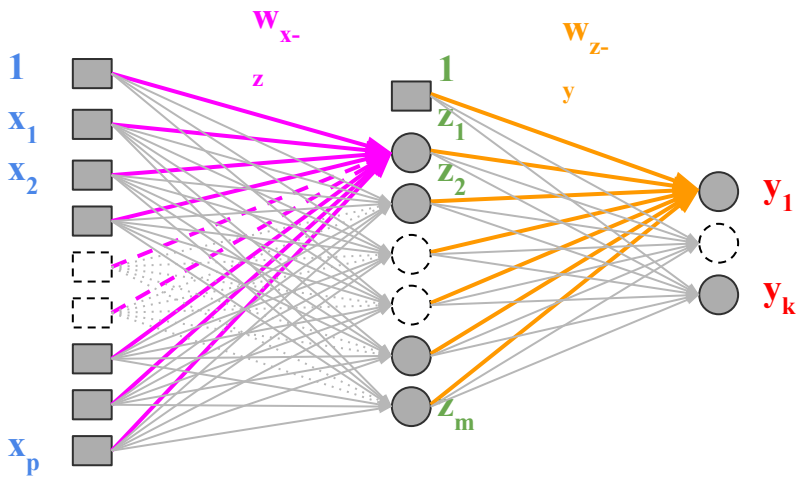
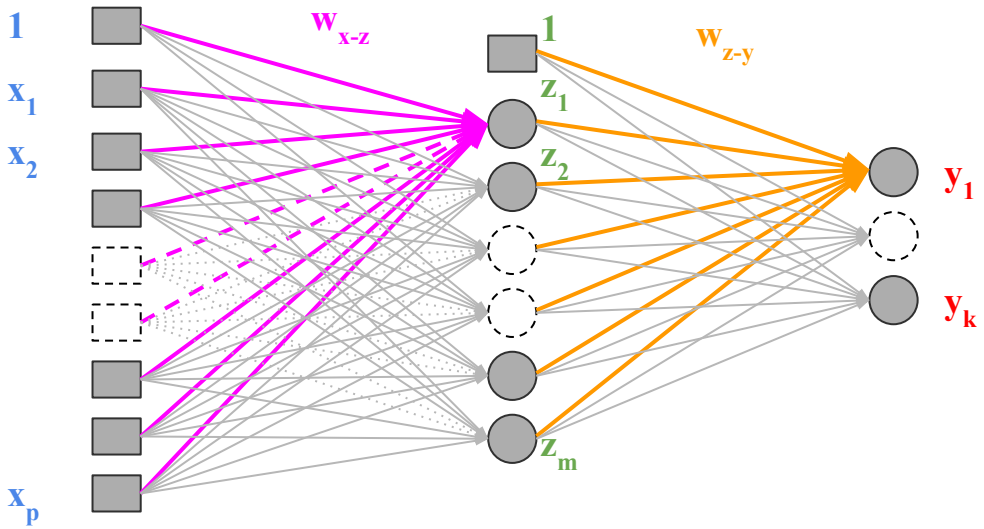
Neural Networks

DSI SEA5, jf.omhover, Oct 18 2016

Based on slides from B. Mann



Parameters



Learning rate

Number of hidden layers

Number of neurons on hidden layers

Initialization of weights

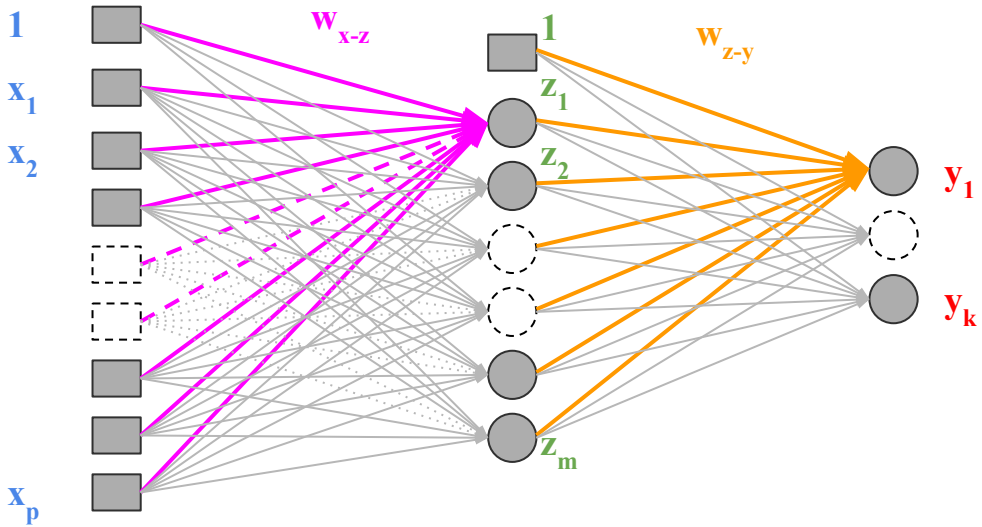
Scaling

Epoch / batches

Activation functions

...

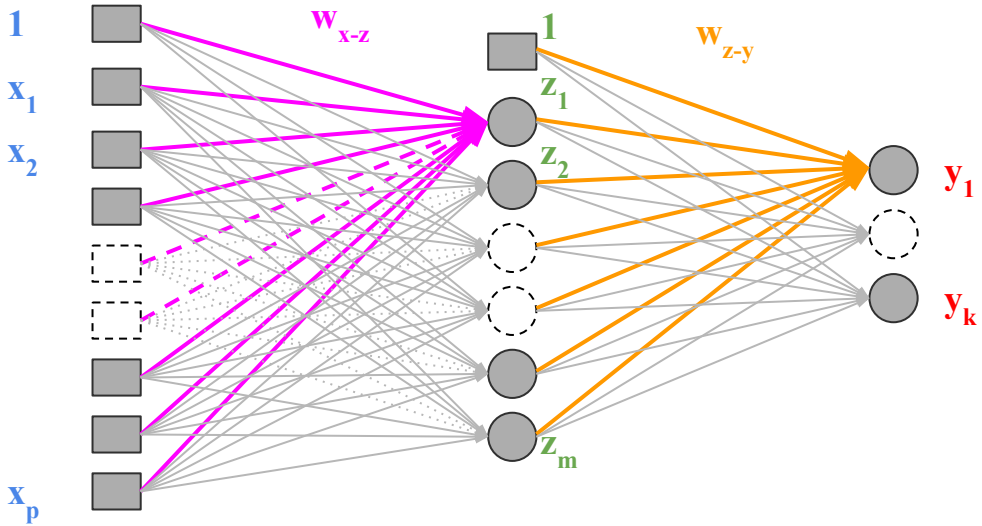
Initialization



Don’t set weights to 0

Sample weights as normal centered around 0

Activation Functions



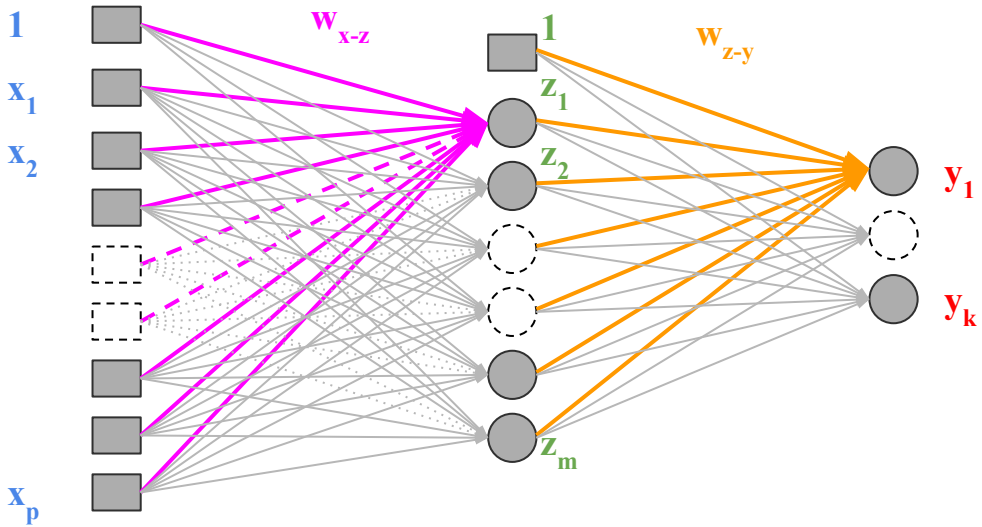
Tanh

Sigmoid

Softmax

ReLu (max(0,x))

Epoch / Batch Size



You'll often see the word epoch:

An epoch is a single sweep through all your data

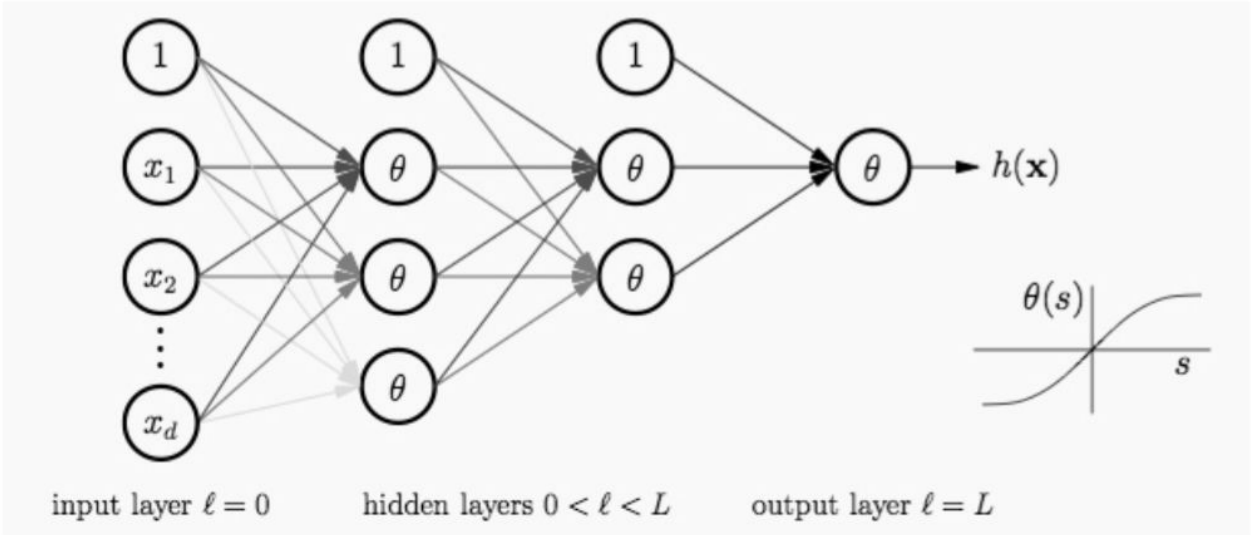
If you have 100,000 observations, and a batch size of 100, each epoch consists of 1,000 gradient descent update steps

Architectures

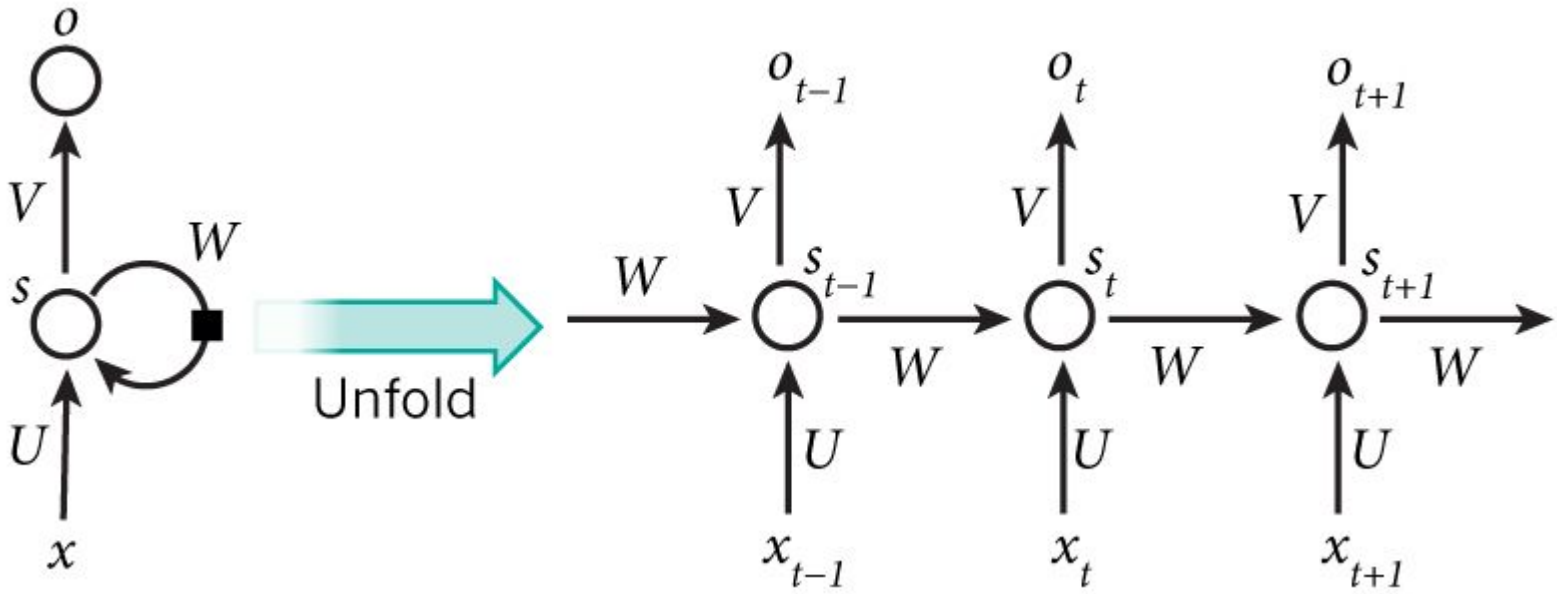
Fully connected networks

This is the type of neural network you've already seen.

- Each layer is *fully connected* to the next
- No missing edges between nodes



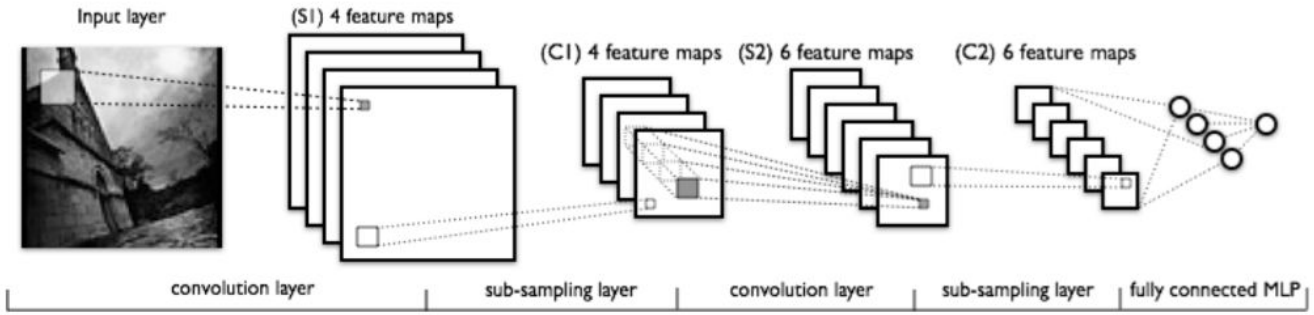
Recurrent Neural Networks



Convolutional Neural Networks

A type of NN modeled after the human eye.

- Not fully connected
- Used mainly for image classification
  - State of the art
- Employs *convolution layers*
  - Each node only "sees" a subset of the previous layer's nodes
  - Applies *convolutions* (a sort of filter) to each sub-image to "look for" certain patterns or shapes (which are learned)



Neural Networks

DSI SEA5, jf.omhover, Oct 18 2016

Based on slides from B. Mann