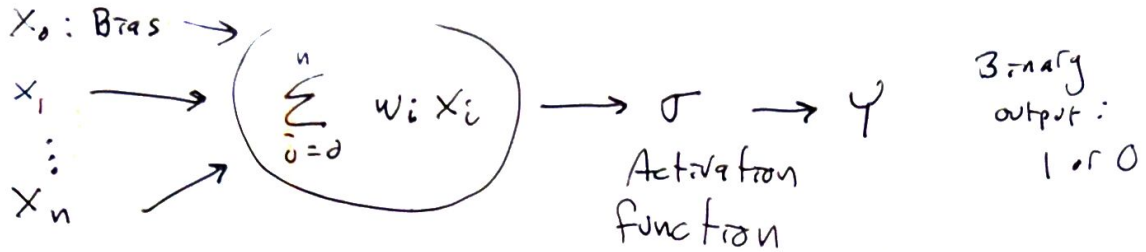
Supervised Learning: Neural Networks

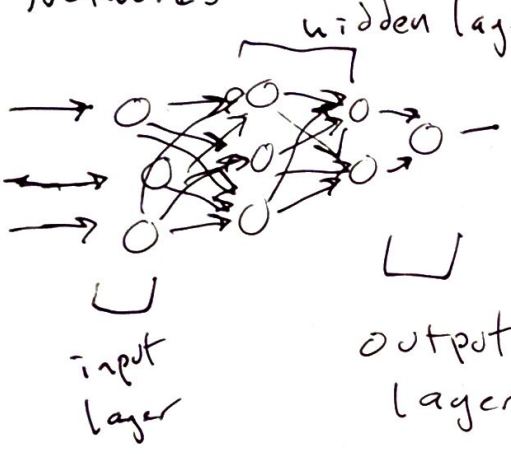
Perceptron single layer NN



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

Bias term – like the intercept in linear regression

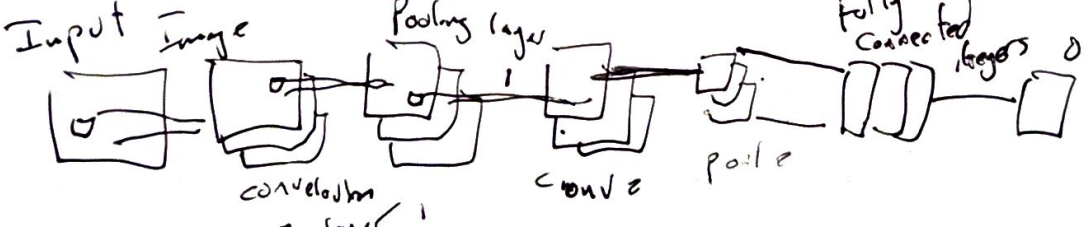
Neural Networks



- Layers of nodes connected together
- Each node is one multivariate linear function, with a univariate non linear transformation
- Trained via (stochastic) gradient descent
- Can represent any non-linear function (very expressive)

- Generally hard to interpret
- Expensive to train, fast to predict
- scikit-learn: sklearn.neural_network.MLPClassifier
- Deep learning frameworks
  MXNet, TensorFlow, Caffe, PyTorch

Convolutional Neural Networks



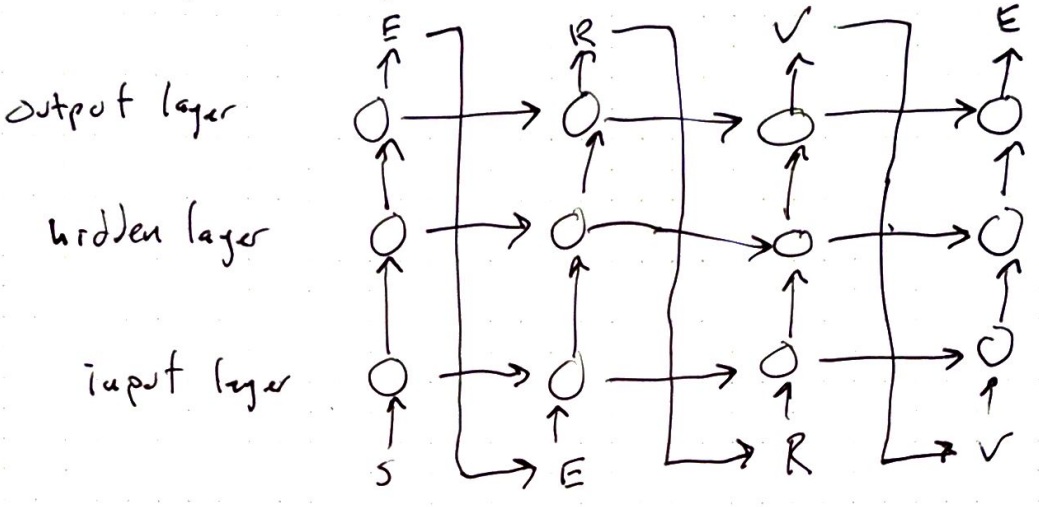
Use filters to convolve ... ie image to create the next layer

Pooling - reduce the size of the output

Take 2×2 output - pick the max or compute the average, output a scalar.

Tensor → Vector

Recurrent Neural Networks



Time series data, NLP, etc

Sequential features? → RNN

Info flow reused in propagation