ML Week 0xJ2-2 ANN

0.1 Perceptron

Controversy

- Led to AI stagnating, though not immediately nor only because of this
- Minsky and Papert book showed impossible to learn XOR
- Often believed (incorrectly) that they claimed same result for multi-layer perceptron
- ANN research resurgence in the 1980's

Algorithm

- Talk about binary classifier
- Talk about bias term meaning
- Talk about weights meaning
- Talk about decision boundary
- Algorithm doesn't terminate if not linearly separable
- Perceptron is the simplest instance of a feedforward neural network
- Initialize weights randomly (or to zero)
- For each input, compute output
- Update weights by adding α if correct

Convergence

- If not linearly separable, don't even get an approximate solution
- If linearly separable, then upper bound on number of times weights updated
- Solution quality not guaranteed
- "Perceptron of optimal stability" now known as SVM

0.2 Overview, Neurons, and deep learning

Review

- WTF is deep learning? ($\times 2$)
- Supervised, reinforcement, unsupervised
- Supervised = regression, classification ($\times 3$)
- Labels are expensive ($\times 3$)

Deep nets

- Architectures
- Applications
- Data sets ($\times 2$)
- Hardware ($\times 5$)
- Brains ($\times 2$)
- Neurons ($\times 3$ each)
  - linear
  - binary threshold
  - rectified linear
  - sigmoid
  - stochastic binary

Example: handwriting ($\times 6$)

Architectures ($\times 10$)

History ($\times 4$)

Perceptron ($\times 5$)

Progress ($\times 5$)

Hinton ($\times 5$)

2011 (before, then since)

Basic ideas ($\times 3$)

0.3 Restricted Boltzmann Machines

Convolution:

$$f(t) * g(t) = (f * g)(t) := \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

Examples (gif's):

- Sliding windows
- Rolling mean