Perceptrons

Michael J. Watts

http://mike.watts.net.nz

Lecture Outline

- ANN revision
- Perceptron architecture
- Perceptron learning
- Problems with perceptrons

ANN Revision

- Artificial neurons
  - mathematical abstractions of biological neurons
  - three functions associated with each neuron
    - input
    - activation
    - output
  - McCulloch and Pitts neuron the first to be designed

ANN Revision

- Artificial Neural Networks
  - Collections of connected artificial neurons
  - Each connection has a weighting value attached to it
  - Weight can be positive or negative
  - Usually layered
  - Signals are propagated from one layer to another

Perceptron Architecture

- An ANN using McCulloch and Pitts neurons
- Proposed by Rosenblatt in 1958
- Developed to study theories about the visual system
- Layered architecture

Perceptron Architecture

- Feedforward
  - signals travel only forward
  - no recurrent connections
- Inputs are continuous
- Outputs are binary
- Used for classification

Perceptron Architecture

- Three neuron layers
  - first is a buffer, usually ignored
  - second is the input or "feature" layer
  - third is the output or "perceptron" layer

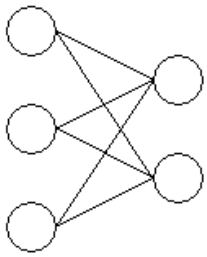
Perceptron Architecture

- Buffer neurons are arbitrarily connected to feature neurons
  - usually one-to-one
    - each buffer neuron is connected only to the corresponding feature neuron
  - can combine multiple buffer values into one feature
- Connection weight values from buffer to feature layer are fixed

Perceptron Architecture

- Feature neurons are fully connected to perceptron neurons
  - Each feature neuron is connected to every perceptron neuron
- Weights from feature to perceptron layer are variable
- One modifiable connection layer
  - single layer network

Perceptron Architecture



Perceptron Architecture

- Threshold neurons are used in the perceptron layer
- Threshold can be a step function or a linear threshold function
- Biases can be used to create the effect of a threshold

Perceptron Architecture

- Perceptron recall
  - propagate a signal from buffer to perceptron layer
  - signal elements are multiplied by the connection weight at each layer
  - output is the activation of the output (perceptron) neurons

Perceptron Learning

- Supervised learning algorithm
- Minimises classification errors

Perceptron Learning

1. propagate an input vector through the perceptron
2. calculate the error for each output neuron by subtracting the actual output from the target output: $E_j = D_j - O_j$
3. modify each weight $w_{ij}^{t+1} = w_{ij}^{t} + lr\, O_j E_j$
4. repeat steps 1-3 until the perceptron converges

Perceptron Learning

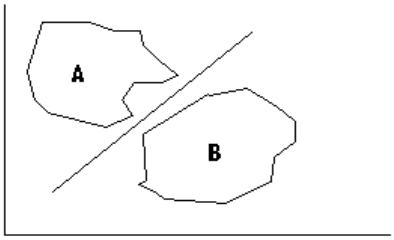
- Other error measures used
- Widrow-Huff learning rule
- Used in a system call ADALINE
  - ADAptive LINEar neuron

$$Err_j = \sum w_{ij} x_i$$

Problems with Perceptrons

- Minsky and Papert (1969)
  - Mathematically prove limitations of perceptrons
  - Caused most AI researchers of the time to return to symbolic AI
- The "Linear Separability" problem
  - Perceptrons can only distinguish between classes separated by a plane

Linear Separability



Linear Separability

- XOR problem is not linearly separable
- Truth table

X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	0

Linear Separability

Linear Separability

- This problem killed research into ANN
- ANN field was moribund for >10 years
- Reasons for this problem
  - Hyperplanes

Linear Separability

- A hyperplane is a plane of >3 dimensions
  - in 2 dimensions it is a line
- Each output neuron in a perceptron corresponds to one hyperplane
- Dimensionality of the hyperplane is determined by the number of incoming connections

Linear Separability

- The position of the hyperplane depends on the weight values
- Without a threshold the hyperplane will pass through the origin
  - causes problems with learning
- Threshold / bias allows the hyperplane to be positioned away from the origin

Linear Separability

- Learning involves positioning the hyperplane
- If the hyperplane cannot be positioned to separate all training examples, then the algorithm will never converge
  - never reach 0 errors

Advantages of Perceptrons

- Simple models
- Efficient
- Guaranteed to converge over linearly separable problems
- Easy to analyse
- Well known

Disadvantages of Perceptrons

- Cannot model non-linearly separable problems
- Some problems may require a bias or threshold parameter
- Not taken seriously anymore

Summary

- Perceptrons are the oldest neural network
- One of the first uses of the McCulloch and Pitts neuron
- Supervised learning algorithm
- Still useful for some problems
- Cannot handle non-linearly separable problems