

Artificial Neural Networks

From Neurons to Perceptrons

The study of ANNs has been inspired in part by the observation that biological learning systems are built of very complex webs of interconnected neurons. The perceptron was the first (very simple) mathematical model for a neuron.

It takes a vector of real-valued inputs $\vec{x} \equiv (x_1 x_2 \dots x_n)$ calculates a linear combination of these inputs, then outputs a 1 if the result is greater than some threshold and -1 otherwise. The weights are the analogon to the "strength" of a certain connection between two biological neurons.

For example a perceptron with two inputs, weights $w_1 = 0.6, w_2 = 0.6$ and threshold 1 can "calculate" AND

x_1	x_2	$x_1 \wedge x_2$	$x_1 w_1 + x_2 w_2$
0	0	0	0.0 (< 1)
0	1	0	0.6 (< 1)
1	0	0	0.6 (< 1)
1	1	1	1.2 (> 1)

For a perceptron, there's a very simple training rule (the "perceptron learning rule", Rosenblatt 1957):

$$w_i \leftarrow w_i + \Delta w_i \text{ where } \Delta w_i = \eta(t - o)x_i$$

t is the target output for the current training example, o is the output generated by the perceptron and η is the so called training rate. Unfortunately, this training rule only succeeds, when the problem space is linearly separable. Thus we need other training rules and multilayer networks.

Gradient descent and Backpropagation

A measure for the training error:

$\sum_{d \in D} (t_d - o_d)^2$ Where D is the set of training examples, t_d is the target output for training example d and o_d is the output of the linear unit.

Visualizing the Hypothesis Space

The "new" training rule: $w_i \leftarrow w_i + \Delta w_i$ where $\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$

Backpropagation in multilayer networks

- The backpropagation algorithm propagates the input forward and the errors backward through the network.
- Because the error surface for multilayer networks may contain many different local minima, gradient descent and thus backpropagation can become trapped in any of these.
- BP is highly effective in practise but it's very unlikely, that this kind of training rule is "applied" in the human brain.

Differences between ANNs and the human brain

- ANNs don't use equivalents for neuro-transmitters and hormones.
- Human brain is much more complex. Most ANNs have only three layers, the brain has far more.
- In ANNs the units are mostly all of the same type (same number of inputs, same activation function, etc.). In the human brain there're different kinds of neurons in different parts of the brain and they have a different number of inputs.
- The output of units in ANNs is always a single constant value whereas the output of a neuron is a complex time series of spikes.
- ANNs aren't "real" networks where the units have "real" distances to each other. ANNs are often simulated on computers which works sequentially not parallel.
- There's no real perception in ANNs, only training sets reduced to some features.
- ...

Other Problems

- Every task needs a different kind of network (designing ANNs like arts).
- No "one shot" learning.
- If one choose a network that is to big, it will be able to memorize all the examples by forming a large lookup table, but will not necessarily generalize well to inputs that have not been seen before.
- ...

Other Network structures / Current research

Today, there are two main categories of neural networks: acyclic or **feed-forward** networks and cyclic or **recurrent** networks.

A **feed-forward network** represents a function only of its current input; thus, it has no internal states other than the weights (often called the long-term memory of the network).

A **recurrent network** feed its outputs back into its own inputs, thus it gets a dynamical system that may reach a stable state or exhibit oscillations or even chaotic behavior. The response of a network depends on its previous state(s), which may depend on previous input. Hence recurrent networks can support a kind of short term memory (unlike the feed forward networks).

References

- Mitchell, Tom M.; Machine Learning, International Edition (1997)
Russel, Stuart and Norvig, Peter; Artificial Intelligence - A Modern Approach, Second Edition, International Edition (2003)
Garson, James; in Stanford Encyclopedia of Philosophy (1997, 2002)