

History

Artificial Neural Networks,...

- Artificial Neural Networks (ANNs) are mathematical models that were developed to mimic the brain functioning.
 - Many do not believe this. We still have not understood the workings of the brain.
 - You can teach the difference between 2 and 3 with a single example to a human, but ANN (even today's deep CNN models) require millions of examples

History, McCulloch and Pitts network

1943

The first model of a neuron was invented by McCulloch (physiologists) and Pitts (logician).

The model had two inputs and a single output.

A neuron would not activate if only one of the inputs was active.

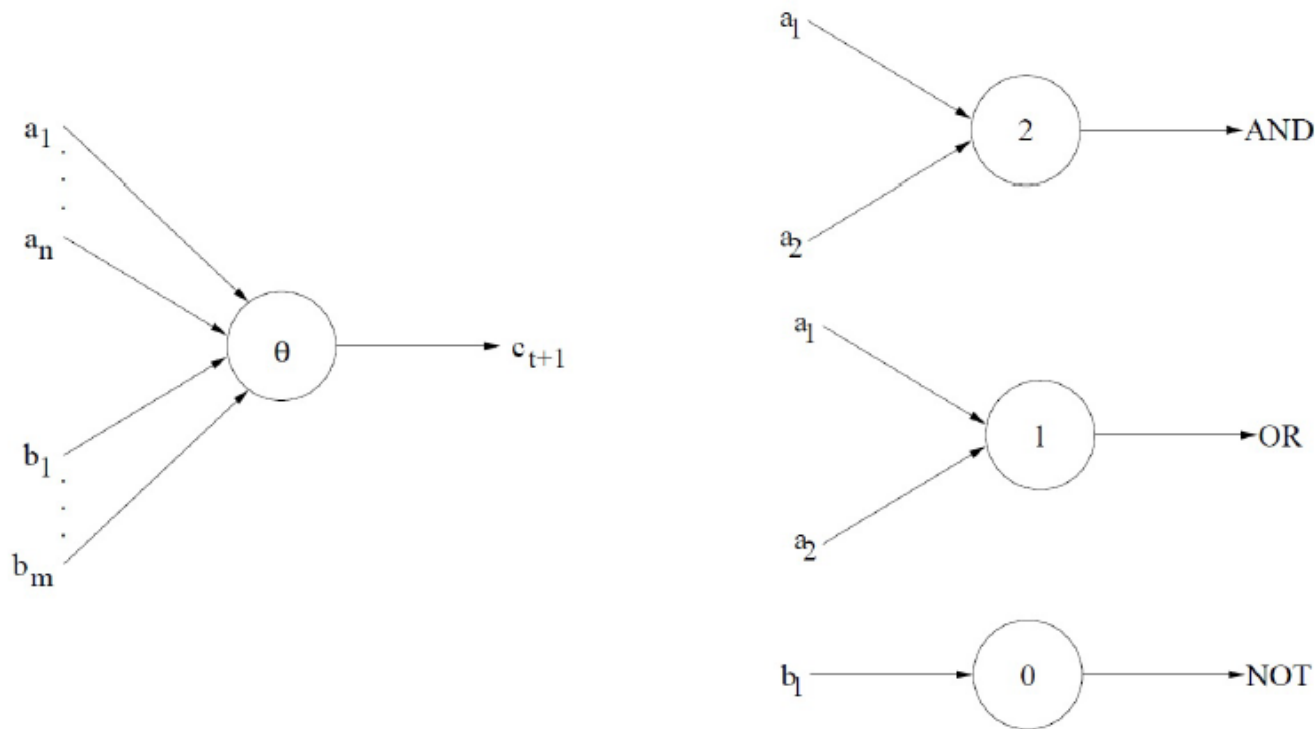
The weights for each input were equal, and the output was binary.

Until the inputs summed up to a certain threshold level, the output would remain zero.

The McCulloch and Pitts' neuron has become known today as a logic circuit.

History, McCulloch and Pitts network (MPN)

1943

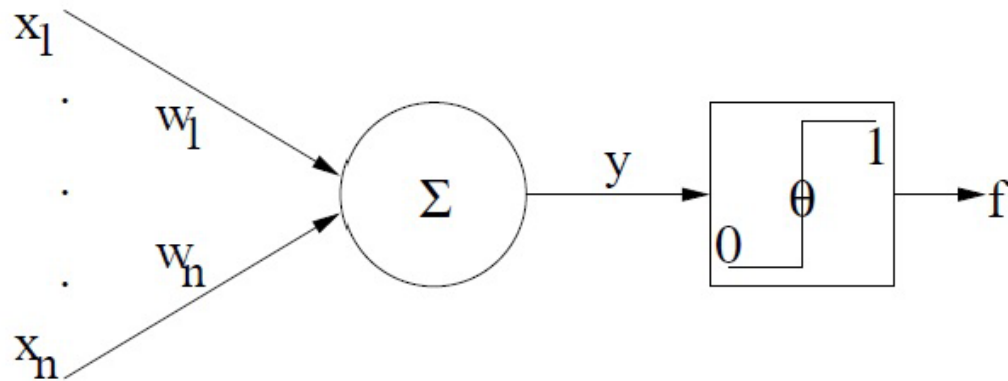


logic functions can be modeled by a network of MP-neurons

History, Perceptron

1958

The perceptron was developed by Rosenblatt (physiologist).



Perceptron, the dream

1958

Rosenblatt randomly connected the perceptrons and changed the weights in order to achieve "learning."

Based on Rosenblatt's statements in a press conference in 1958, The New York Times reported the perceptron to be
'the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.'

MPN vs Perceptron

Apparently McCulloch and Pitts' neuron is a better model for the electrochemical process inside the neuron than the perceptron.

But perceptron is the basis and building block for the modern neural networks.

History, optimization

1960

Widrow and Hoff proposed a method for adjusting the weights. They introduced a gradient search method based on minimizing the error squared (Least Mean Squares).

In the 1960's, there were many articles promising robots that could think.

It seems there was a general belief that perceptrons could solve any problem.

History, shattered dream

1969

Minsky and Papert published their book *Perceptrons*. The book shows that perceptrons could only solve linearly separable problems.

They showed that it is not possible for perceptron to learn an XOR function.

After *Perceptrons* was published, researchers lost interest in perceptron and neural networks.

1969

Arthur E. Bryson and Yu-Chi Ho described proposed Backpropagation as a multi-stage dynamic system optimization method. (Bryson, A.E.; W.F. Denham; S.E. Dreyfus. Optimal programming problems with inequality constraints. I: Necessary conditions for extremal solutions. AIAA J. 1, 11 (1963) 2544-2550)

1972

Stephen Grossberg proposed networks capable of learning XOR function.

1974

Backpropagation was reinvented / applied in the context of neural networks by Paul Werbos, David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams.

Back propagation allowed perceptrons to be trained in a multilayer configuration.

1980s

The field of artificial neural network research experienced a resurgence.

2000s

Neural network fell out of favor partly due to BP limitations.

Backpropagation Limitations

It requires labeled training data.

It is very slow in networks with multiple layers (doesn't scale well).

It can converge to poor local minima.

Around this downfall, SVMs emerged which do not have the above limitations.

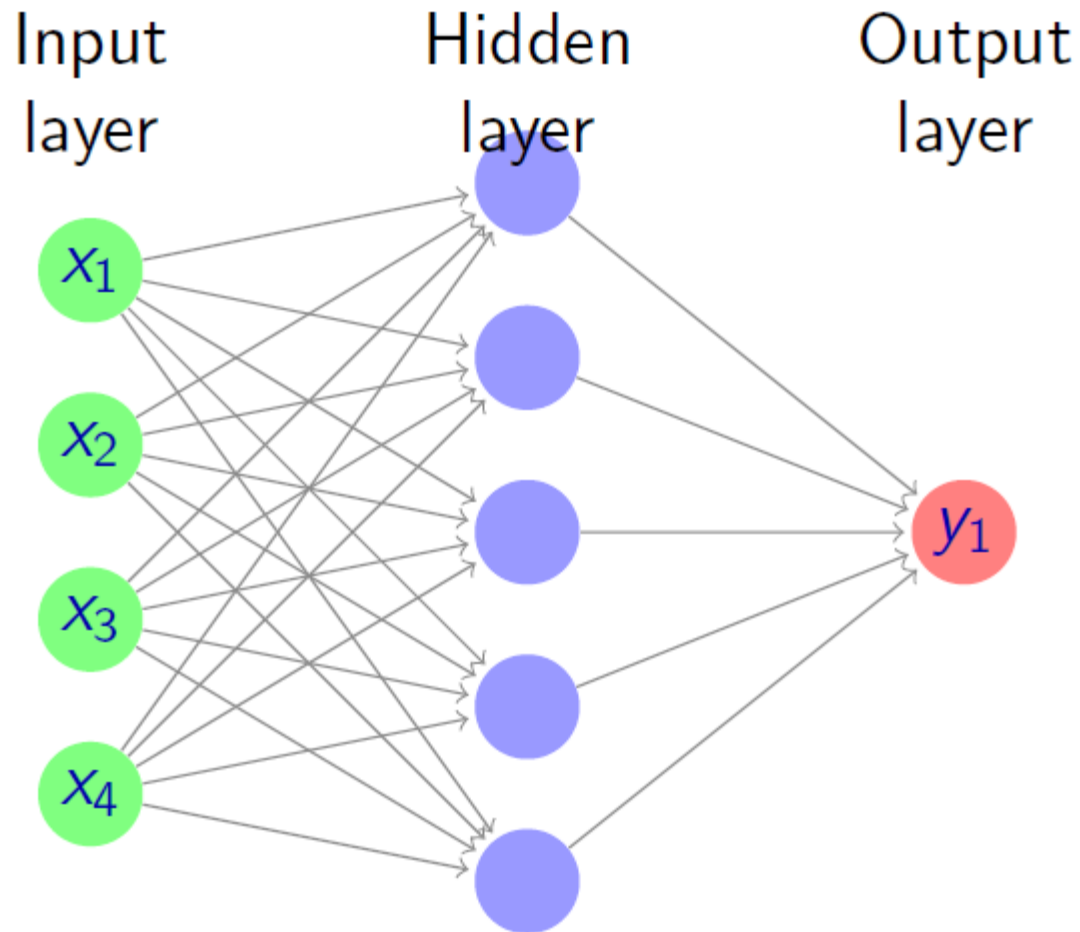
but has returned again in the 2010s, now able to train much larger networks using huge modern computing power such as GPUs. For example, in 2013 top speech recognisers now use backpropagation-trained neural networks.

- Some researchers say that invention of Neural networks before SVMs was a historical mistake.
 - History of ANNs is multimodal. Between two modes of this, SVM appeared !!!

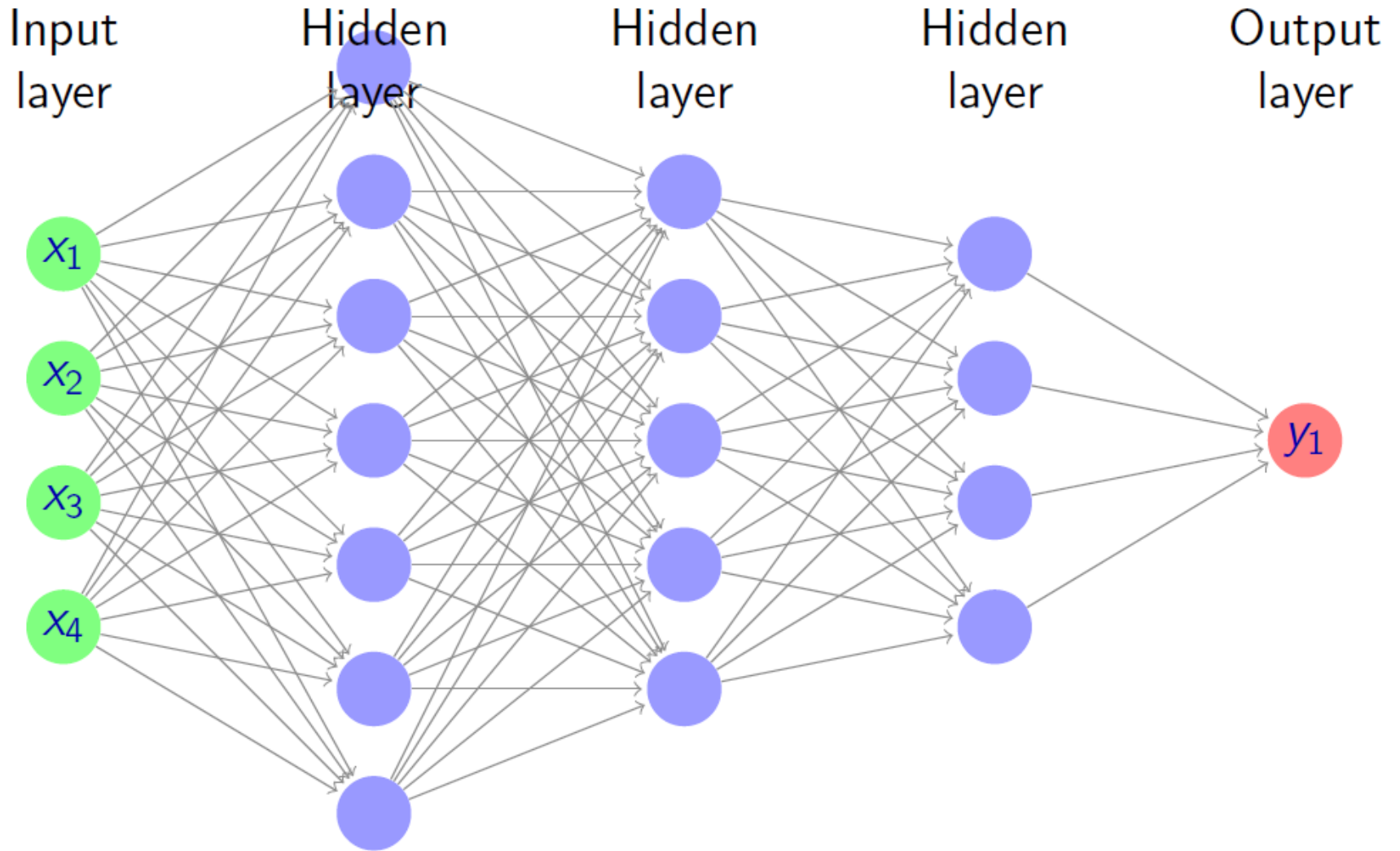
Local minima is bad..

- But in high dimensional space (input is from a high dim. space, like images or so) the things are somewhat better.
- Many local minima can be equally good (apparently as good as global minima... debatable, but this is one of the beliefs for the success of the Deep networks).

Feedforward Neuralnetwork



Feedforward Deep Networks



Feedforward Deep Networks

- Feedforward deep networks, a.k.a. multilayer perceptrons (MLPs), are parametric functions composed of several parametric functions.
- Each layer of the network defines one of these sub-functions.
- Each layer (sub-function) has multiple inputs and multiple outputs.
- Each layer composed of many units (scalar output of the layer).
- We sometimes refer to each unit as a feature.
- Each unit is usually a simple transformation of its input.
- The entire network can be very complex.