

# A Concise History of Neural Networks



Jiaconda

Follow

Aug 13, 2016 · 6 min read

*"From the barren landscapes inside our personal devices come furtive anthems hummed by those digital servants who will one day be our overlords"*

## A.I. winter

The idea of neural networks began unsurprisingly as a model of how neurons in the brain function, termed 'connectionism' and used connected circuits to simulate intelligent behaviour. In 1943, portrayed with a simple electrical circuit by neurophysiologist Warren McCulloch and mathematician Walter Pitts. Donald Hebb took the idea further in his book, *The Organization of Behaviour* (1949), proposing that neural pathways strengthen over each successive use, especially between neurons that tend to fire at the same time thus beginning the long journey towards quantifying the complex processes of the brain.

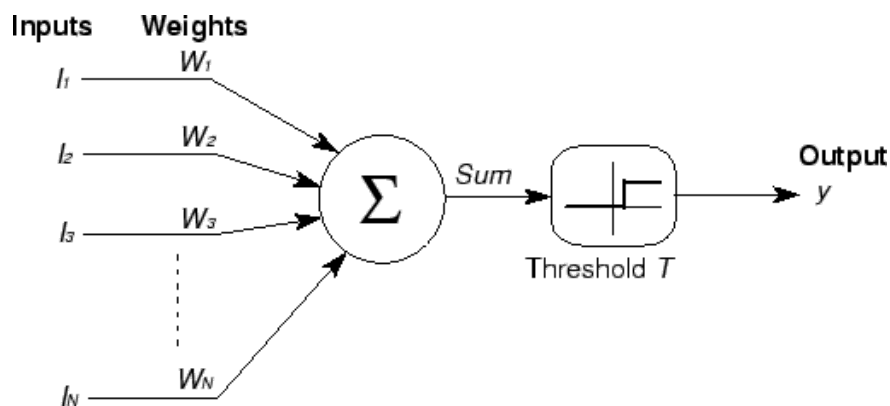
Two major concepts that are precursors to Neural Networks are

- 'Threshold Logic'—converting continuous input to discrete output
- 'Hebbian Learning'—a model of learning based on neural plasticity, proposed by Donald Hebb in his book *"The Organization of Behaviour"* often summarized by the phrase: "Cells that fire together, wire together."

both proposed in the 1940's. In 1950s, as researchers began trying to translate these networks onto computational systems, the first Hebbian network was successfully implemented at MIT in 1954.

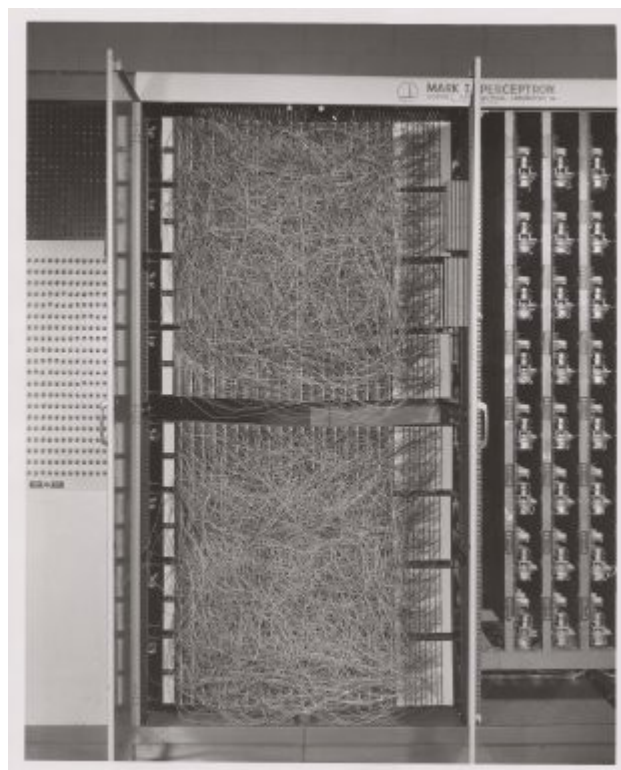
Around this time, Frank Rosenblatt, a psychologist at Cornell, was working on understanding the comparatively simpler decision systems present in the eye of a fly, which underlie and determine its flee response. In an attempt to understand and quantify this process, he proposed the idea of a Perceptron in 1958, calling it Mark I Perceptron. It was a system with a simple input output relationship, modeled on a *McCulloch-Pitts* neuron, proposed in 1943 by Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician to explain the complex

decision processes in a brain using a linear threshold gate. A McCulloch-Pitts neuron takes in inputs, takes a weighted sum and returns '0' if the result is below threshold and '1' otherwise.



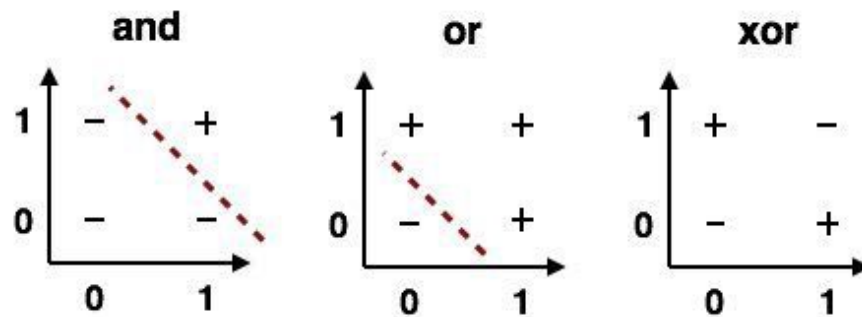
A McCulloch-Pitts neuron

The beauty of Mark I Perceptron lay in the fact that its weights would be 'learnt' through successively passed inputs, while minimizing the difference between desired and actual output.



First known implementation of a Mark I Perceptron. The machine was connected to a camera that used 20×20 cadmium sulfide photocells to produce a 400-pixel image. The main visible feature is a patchboard that allowed experimentation with different combinations of input features. To the right of that are arrays of potentiometers that implemented the adaptive weights.[wiki]

A major drawback? This perceptron could only learn to separate linearly separable classes, making the simple but non-linear exclusive-or circuit an insurmountable barrier.



Despite the messy and somewhat dis-satisfactory advent of the use of Machine Learning to quantify decision systems apart from the brain, today's artificial neural networks are nothing more than several layers of these perceptrons.

Things started moving quickly for neural networks around this time and in 1959 at Stanford, Bernard Widrow and Marcian Hoff developed the first neural network successfully applied to a real world problem. These systems were named ADALINE and MADALINE after their use of Multiple ADaptive LINear Elements, the latter of which was specifically designed to eliminate noise in phone lines and still remains in use today(!). These artificial neurons however were different from perceptrons in what they returned as output, which in this case was the weighted input.

As is the case with every small improvement in AI technology in history ever, these early successes gave rise to an increasing hype on the ability and potential of neural networks while researchers were coming up against one roadblock after another. At the peak of the hype around these 'Thinking Machines', NYtimes published [this](#) article on the potential of neural networks while this video was released around the same time,

## The Thinking Machine (Artificial Intelligenc...



As with the several ‘close calls’ before, we were still nowhere close to spewing conscious man-made beings as much as we’ve always loved to believe (or fear, depends how you look at it). One of the problems that arose was with the impractically long runtimes required for running these networks given that this was the 60s apart from its inability to learn simple boolean exclusive-or circuits.

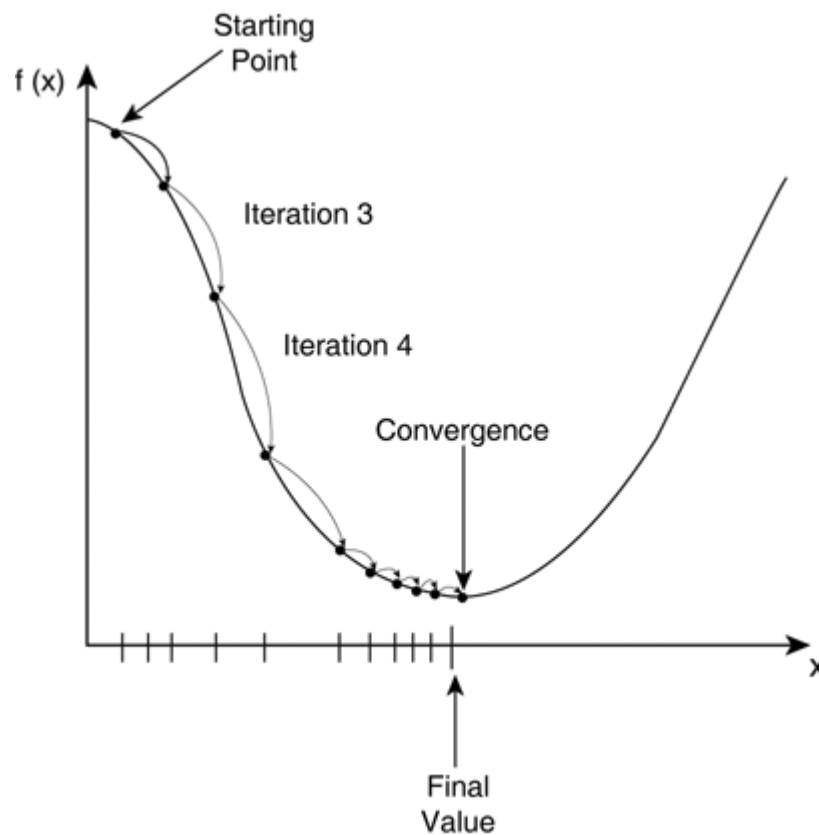
All this came to an end in 1969 with the publication of a book “Perceptrons” by Marvin Minsky, founder of the MIT AI Lab, and Seymour Papert, director of the lab. The book conclusively argued that the Rosenblatt’s single perception approach to neural networks could not be translated effectively into multi-layered neural networks. To evaluate the correct relative values of the weights of the neurons spread across layers based on the final output would take several if not infinite number of iterations and would take a very long time to compute.

Minsky in his text laid out these and other problems with Neural Nets and effectively led the larger scientific community and most importantly the funding establishments to the conclusion that further research in this direction was to lead nowhere. The effect of this text was powerful and dried up funding to an extent that, for the next 10–12 years, no-one at the largest research institutions at the time and thereby the smaller ones too, would take on any project that had that had the doomed Neural Networks as its premise. The age now famously referred to as ‘the AI winter’ had begun.

The thawing of this decade long winter began in 1982 at the National Academy of Sciences when Jon Hopfield presented his paper on what came to be known as Hopfield Net while the same year at the US-Japan conference on Cooperative/Competitive Neural Networks, Japan announced its intention to begin its fifth generation effort on Neural Networks. This got the funding to start flowing again from the coffers of a nation afraid to be left behind. Soon, the American Institute of Physics, in 1985 established a “Neural Networks in Computing” annual meeting followed by the first International Conference on Neural Networks by the Institute of Electrical and Electronic Engineers (IEEE) in 1987.

It was however a major re-discovery of a concept already in existence since the 60's that helped neural nets get out of its premature grave. Backpropagation, a method devised by researchers since the 60's and continuously developed on well into the AI winter, was an intuition based method that attributed reducing significance to each event as one went farther back in the chain of events. The first person to see their potential for neural nets and solving the question of how that would be translated for MLP's was Paul Werbos who inspired partly by its application to the human mind and Freud's work on backward flow of credit assignment, wrote a PhD thesis expounding their importance. However this work wasn't noticed by anyone in the community until Parker published a report on his work at M.I.T. in 1985. It was only after being re-re-discovered by Rumelhart, Hinton, and Williams and republished in a clear and detailed framework that the technique took over the community by storm. The same authors also addressed the specifics drawbacks laid out by Minsky in his 1969 publication in a later text.

Backpropagation along with Gradient Descent forms the backbone and powerhouse of neural networks. While Gradient Descent constantly updates and moves the weights and bias towards the minimum of the cost function, backpropagation evaluates the gradient of the cost w.r.t. weights and biases, the magnitude and direction of which is used by gradient descent to evaluate the size and direction of the corrections to weights and bias parameters.



A simple visual description of the movement towards the minima of a 2d function. The step-size of jump is determined by the value of the gradient at each point.

And thus by 1990's, Neural networks were definitely back, this time truly catching the imagination of the world and finally coming to par with, if not overtaking, its expectations. Yet again, we are asking the same questions of AI, and projecting onto it, our all too human fears, and yet again we are farther than we think from bowing in deference to our digital overlords.

