What is an Artificial Neural Network (ANN)?

Based on operation of the brain. Decision making is more important than crunching numbers.

Based on simultaneous (parallel), rather than serial computation.

Built from neural elements (cells) that are modeled after living neurons.

Architecture:

Cells are assembled into layers

Layers are assembled into networks

Output of one layer feeds inputs of the next layer.

The inputs to the network are “features” of an image

The outputs of the network are categories.

The network may or may not have feedback.

The network “learns” by modifying the connections from one layer to the next.

Network signals may be binary or continuous, depending on the application.

May be implemented in hardware or software.

Applications of ANNs.

Pattern recognition

Visual (faces, handwriting)

Auditory (voice)

Chemical (agricultural, industrial monitoring)

Mechanical (vibrations, forces)

Optimization

Image processing (denoising, medicine)

Control (robotics, vehicles)

Prediction

Complex, nonlinear relationships.

Operating principles

Processing occurs on a cellular level

Neurons are linked via connective pathways

Pathways have variable weights (training)

Cells have variable I/O (activation) functions

Cells have variable thresholds.

Fan-in is limitless

Fan-out is limitless

Architecture is limitless

Exhibit “robustness”.

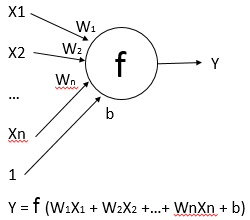
Biological neurons have all of the same features, plus:

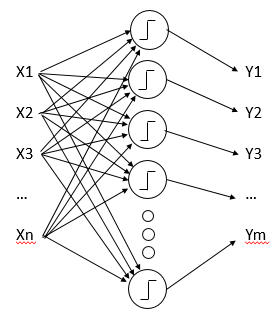
Global controls (hormones)

Self-training capability.

Short-term and long-term memory

May be excited or inhibited.

A typical cell

A typical layer

A history of ANNs

Neurons have been studied for over a century-known to be electrical devices. In 1911, it was found that the brain was made up of a collection of interconnected cells. The human brain has about 10 billion cells and 60 trillion synapses. Neurons have a time delay of about 1 mS, and are 10 billion times more energy efficient than electronic gates.

Since then, the brain has been studied on many levels:

Molecules-chemical messengers

Synapses-the basis for communication and learning

Connectivity and local circuitry

Neural systems (sensory, processing, and motor)

Artificial neural networks were born in 1943.

McCullouch & Pitts

Inputs are time-variant, discrete values.

Weights are constant, positive or negative.

Thresholds are constant, positive or negative.

Cell outputs are only 0 or 1 (like a logic gate).

Each cell has a time delay ().

What can you do with McCullouch and Pitts networks?

Combinatorial logic

Sequential logic

Wiener (1948) “Cybernetics” envisioned neural elements for signal processing, communication,

and control.

Hebb (1948) “The Organization of Behavior”, introduced synaptic modification.

Learning rules WNEW = WOLD +  X Y

Rochester et al implemented Hebbian learning in software and showed that inhibition is needed

for training as well as excitation.

Ashby (1952) introduced the idea that adaptive behavior is learned, and not inborn. Learned

behavior takes place for better functioning. There are dynamics and stability of learning.

John von Neumann (1958) deeper exploration of learning rules, redundancy, and robustness.

Went on to become the father of modern computers.

Minsky (1954) theorized ANNs in his dissertation, later applied them to AI and computation.

Gabor (1954) applied neural principles to adaptive filters.

Taylor (1956) introduced associative memory.

Rosenblatt (1958) formulated the Perceptron-a neural element that is more similar to a neuron

than a McCullouch and Pitts cell. This formed the basis for modern ANNs.

Weights are adjustable by training that takes place over many cycles.

Widely accepted until Minsky and Papert proposed the X-OR problem in 1969.

Widrow and Hoff (1960) introduced the Adaline and the notion of input (feature) spaces and

error-based learning, the Generalized Delta Rule GDR. They showed how multilayered

networks of Perceptrons (Madalines) could solve the X-OR problem in 1962.

Amari (1967) introduced training by stochastic gradient descent.

Kohonen (1972) introduced the notion of Associative Memory, how the brain stores information

as opposed to how digital computers store information. Gave rise to Feature Maps, self

organization, speech recognition, and constrained optimization.

Then, 1n 1969, the field of ANNs went into a 10-year recession. They still couldn’t solve the X-OR problem with a single cell, and computers weren’t powerful enough to apply them to solving nontrivial problems. Then, with the development of more powerful computers, they came back in the 1980’s.

Grossberg and Carpenter (1980) came up with Adaptive Resonance Theory (ART), whereby a

network goes into a resonant state when it correctly identifies a pattern.

John Hopfield (1982) proposed an error (energy) function as a guide for training. It supported

recurrent architecture, symmetrical synapses, dynamics, stability, and attractors. The

time course of training is in itself a form of memory.

Kirkpatrick et al (1983) introduced simulated annealing and the effects of initial conditions.

Barto et al (1986) used Reinforcement Learning for control problems, eg backing up a semi,

inverted pendulum, pushing a caster with only one degree of feedback.

Rummelhart et al (1986) introduced backpropagation training, and the applications opened up.

Broohhead et al (1988) introduced Radial Basis Functions (RBF) for clustering of images in input

Space.

Mead (1989) wrote Analog VLSI and Neural Systems for modeling of sensory processes.

Since then, ANNs have abounded, mostly simulated in software on high-performance machines that offer enough storage capacity, processing power, and user interface.