UNIT I

What is a Neural Network?

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.

Why use neural networks?

Neural networks can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. **Other advantages include:**

Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.

Self-Organisation: An ANN can create its own organisation or representation of the information it receives during learning time.

Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.

Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Artificial neural networks (ANNs) or connectionist **systems:** ANN is computing systems vaguely inspired by the biological neural networks that constitute animal brains. Such systems "learn" (i.e. progressively improve performance on tasks by considering examples, generally without task-specific programming. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any a priori knowledge about cats, e.g., that they have fur, tails, whiskers and cat-like faces. Instead, they evolve their own set of relevant characteristics from the learning material that they process.

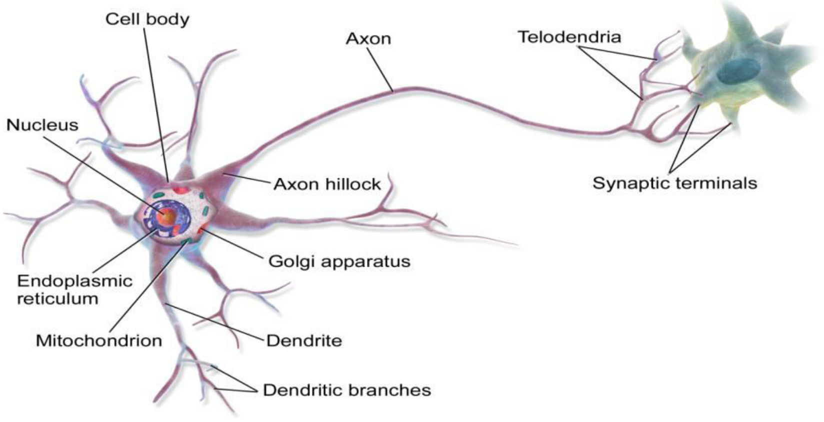
An ANN is based on a collection of connected units or nodes called artificial neurons (a simplified version of biological neurons in an animal brain). Each connection (a simplified version of a synapse) between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it and then signal artificial neurons connected to it.

In common ANN implementations, the signal at a connection between artificial neurons is a real number, and the output of each artificial neuron is calculated by a non-linear function of the sum of its inputs. Artificial neurons and connections typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that only if the aggregate signal crosses that threshold is the signal sent. Typically, artificial neurons are organized in layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first (input), to the last (output) layer, possibly after traversing the layers multiple times.

The original goal of the ANN approach was to solve problems in the same way that a human brain would. However, over time, attention focused on matching specific tasks, leading to deviations from biology. ANNs have been used on a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games and medical diagnosis.

NEURON

A neuron, also known as a neurone (British spelling) and nerve cell, is an electrically excitable cell that receives, processes, and transmits information through electrical and chemical signals. These signals between neurons occur via specialized connections called synapses. Neurons can connect to each other to form neural networks. Neurons are the primary components of the central nervous system, which includes the brain and spinal cord, and of the peripheral nervous system, which comprises the autonomic nervous system and the somatic nervous system.



Types of specialized neurons:

There are many types of specialized neurons.

Sensory neurons respond to one particular type of stimulus such as touch, sound, or light and all other stimuli affecting the cells of the sensory organs, and converts it into an electrical signal via transduction, which is then sent to the spinal cord or brain**.**

Motor neurons receive signals from the brain and spinal cord to cause everything from muscle contractions and affect glandular outputs.

Inter neurons connect neurons to other neurons within the same region of the brain or spinal cord in neural networks.

A typical neuron consists of a **cell body** (soma)**,** dendrites**, and an** axon. The term neurite is used to describe either a dendrite or an axon, particularly in its undifferentiated stage**. Dendrites** are thin structures that arise from the cell body, often extending for hundreds of micrometers and branching multiple times, giving rise to a complex "dendritic tree". **An axon** (also called a nerve fiber) is a special cellular extension (process) that arises from the cell body at a site called the axon hillock and travels for a distance, as far as 1 meter in humans or even more in other species. Most neurons receive signals via the dendrites and send out signals down the axon. Numerous axons are often bundled into fascicles that make up the nerves in the peripheral nervous system (like strands of wire make up cables). Bundles of axons in the central nervous system are called tracts. The cell body of a neuron frequently gives rise to multiple dendrites, but never to more than one axon, although the axon may branch hundreds of times before it terminates. At the majority of synapses, signals are sent from the axon of one neuron to a dendrite of another. There are, however, many exceptions to these rules: for example, neurons can lack dendrites, or have no axon, and synapses can connect an axon to another axon or a dendrite to another dendrite.

All neurons are electrically excitable, due to maintenance of voltage gradients across their membranes by means of metabolically driven ion pumps, which combine with ion channels embedded in the membrane to generate intracellular-versus-extracellular concentration differences of ions such as sodium, potassium, chloride, and calcium. Changes in the cross-membrane voltage can alter the function of voltage-dependent ion channels. If the voltage changes by a large enough amount, an all-or-none electrochemical pulse called an action potential is generated and this change in cross-membrane potential travels rapidly along the cell's axon, and activates synaptic connections with other cells when it arrives.

A neuron is a specialized type of cell found in the bodies of all eumetozoans. Only sponges and a few other simpler animals lack neurons. The features that define a neuron are electrical excitability and the presence of synapses, which are complex membrane junctions that transmit signals to other cells. The body's neurons, plus the glial cells that give them structural and metabolic support, together constitute the nervous system. In vertebrates, the majority of neurons belong to the central nervous system, but some reside in peripheral ganglia, and many sensory neurons are situated in sensory organs such as the retina and cochlea.

Parts of Neuron:

A typical neuron is divided into three parts: the soma or cell body, dendrites, and axon.

- The soma is usually compact; the axon and dendrites are filaments that extrude from it.
- Dendrites typically branch profusely, getting thinner with each branching, and extending their farthest branches a few hundred micrometers from the soma.
- The axon leaves the soma at a swelling called the axon hillock, and can extend for great distances, giving rise to hundreds of branches.

Unlike dendrites, an axon usually maintains the same diameter as it extends. The soma may give rise to numerous dendrites, but never to more than one axon. Synaptic signals from other neurons are received by the soma and dendrites; signals to other neurons are transmitted by the axon. A typical synapse, then, is a contact between the axon of one neuron and a dendrite or soma of another. Synaptic signals may be excitatory or inhibitory. If the net excitation received by a neuron over a short period of time is large enough, the neuron generates a brief pulse called an action potential, which originates at the soma and propagates rapidly along the axon, activating synapses onto other neurons as it goes.

Many neurons fit the foregoing schema in every respect, but there are also exceptions to most parts of it. There are no neurons that lack a soma, but there are neurons that lack dendrites, and others that lack an axon. Furthermore, in addition to the typical axodendritic and axosomatic synapses, there are axoaxonic (axon-to-axon) and dendrodendritic (dendrite-to-dendrite) synapses.

The key to neural function is the synaptic signaling process, which is partly electrical and partly chemical. The electrical aspect depends on properties of the neuron's membrane. Like all animal cells, the cell body of every neuron is enclosed by a plasma membrane, a bilayer of lipidmolecules with many types of protein structures embedded in it. A lipid bilayer is a powerful electrical insulator, but in neurons, many of the protein structures embedded in the membrane are electrically active. These include ion channels that permit electrically charged ions to flow across the membrane and ion pumps that actively transport ions from one side of the membrane to the other. Most ion channels are permeable only to specific types of ions. Some ion channels are voltage gated, meaning that they can be switched between open and closed states by altering the voltage difference across the membrane. Others are chemically gated, meaning that they can be switched between open and closed states by interactions with chemicals that diffuse through the extracellular fluid. The interactions between ion channels and ion pumps produce a voltage difference across the membrane, typically a bit less than 1/10 of a volt at baseline. This voltage has two functions: first, it provides a power source for an assortment of voltage-dependent protein machinery that is embedded in the membrane; second, it provides a basis for electrical signal transmission between different parts of the membrane.

Neurons communicate by chemical and electrical synapses in a process known as neurotransmission, also called synaptic transmission. The fundamental process that triggers the release of neurotransmitters is the action potential, a propagating electrical signal that is generated by exploiting the electrically excitable membrane of the neuron. This is also known as a wave of depolarization.

Neurons are highly specialized for the processing and transmission of cellular signals. Given their diversity of functions performed in different parts of the nervous system, there is a wide variety in their shape, size, and electrochemical properties. For instance, the soma of a neuron can vary from 4 to 100 micrometers in diameter.

The soma is the body of the neuron. As it contains the nucleus, most protein synthesis occurs here. The nucleus can range from 3 to 18 micrometers in diameter.

The dendrites of a neuron are cellular extensions with many branches. This overall shape and structure is referred to metaphorically as a dendritic tree. This is where the majority of input to the neuron occurs via the dendritic spine.

The axon is a finer, cable-like projection that can extend tens, hundreds, or even tens of thousands of times the diameter of the soma in length. The axon carries nerve signalsaway from the soma (and also carries some types of information back to it). Many neurons have only one axon, but this axon may—and usually will—undergo extensive branching, enabling communication with many target cells. The part of the axon where it emerges from the soma is called the axon hillock. Besides being an anatomical structure, the axon hillock is also the part of the neuron that has the greatest density of voltage-dependent sodium channels. This makes it the most easily excited part of the neuron and the spike initiation zone for the axon: in electrophysiological terms, it has the most negative action potential threshold. While the axon and axon hillock are generally involved in information outflow, this region can also receive input from other neurons.

The axon terminal contains synapses, specialized structures where neurotransmitter chemicals are released to communicate with target neurons.

The accepted view of the neuron attributes dedicated functions to its various anatomical components. however, dendrites and axons often act in ways contrary to their so-called main function.

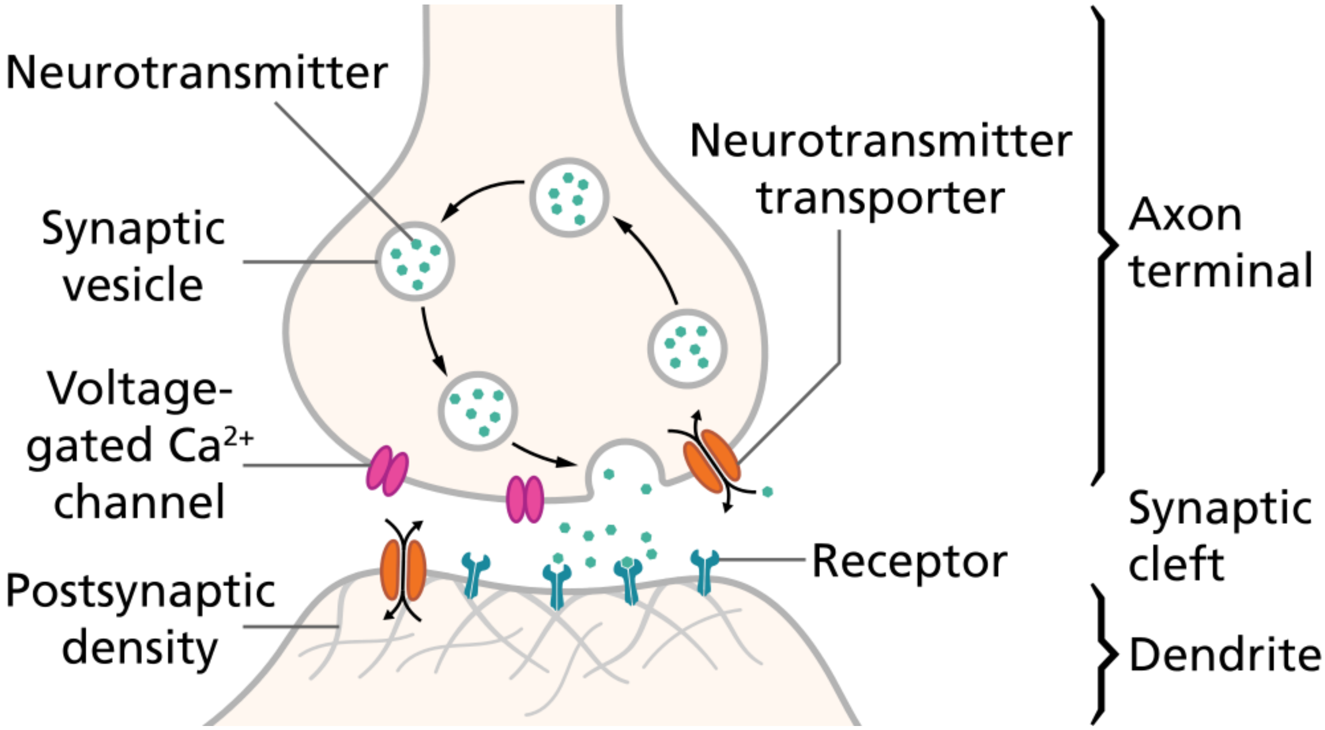
Axons and dendrites in the central nervous system are typically only about one micrometer thick, while some in the peripheral nervous system are much thicker. The soma is usually about 10–25 micrometers in diameter and often is not much larger than the cell nucleus it contains. The longest axon of a human motor neuron can be over a meter long, reaching from the base of the spine to the toes.

Sensory neurons can have axons that run from the toes to the posterior column of the spinal cord, over 1.5 meters in adults. Giraffes have single axons several meters in length running along the entire length of their necks. Much of what is known about axonal function comes from studying the squid giant axon, an ideal experimental preparation because of its relatively immense size (0.5–1 millimeters thick, several centimeters long).

Fully differentiated neurons are permanently postmitotic, however, research starting around 2002 shows that additional neurons throughout the brain can originate from neural stem cells through the process of neurogenesis. These are found throughout the brain, but are particularly concentrated in the subventricular zone and subgranular zone.

Synapse:

In the nervous system, a **synapse** is a structure that permits a neuron (or nerve cell) to pass an electrical or chemical signal to another neuron or to the target efferent cell.



Santiago Ramón y Cajal proposed that neurons are not continuous throughout the body, yet still communicate with each other, an idea known as the neuron doctrine. The word "synapse" – from the Greek synapsis(συνάψις), meaning "conjunction", in turn from συνάπτεὶν (συν ("together") and ἅπτειν ("to fasten")) – was introduced in 1897 by the English neurophysiologist Charles Sherrington in Michael Foster's Textbook of Physiology. Sherrington struggled to find a good term that emphasized a union between two separate elements, and the actual term "synapse" was suggested by the English classical scholar Arthur Woollgar Verrall, a friend of Michael Foster. Some authors generalize the concept of the synapse to include the communication from a neuron to any other cell type, such as to a motor cell, although such non-neuronal contacts may be referred to as junctions (a historically older term).

Synapses are essential to neuronal function. Neurons are cells that are specialized to pass signals to individual target cells, and synapses are the means by which they do so. At a synapse, the plasma membrane of the signal-passing neuron (the presynaptic neuron) comes into close apposition with the membrane of the target (postsynaptic) cell. Both the presynaptic and postsynaptic sites contain extensive arrays of a molecular machinery that link the two membranes together and carry out the signaling process. In many synapses, the presynaptic part is located on an axon and the postsynaptic part is located on a dendrite or soma. Astrocytes also exchange information with the synaptic neurons, responding to synaptic activity and, in turn, regulating neurotransmission. Synapses (at least chemical synapses) are stabilized in position by synaptic adhesion molecules (SAMs) projecting from both the pre- and post-synaptic neuron and sticking together where they overlap; SAMs may also assist in the generation and functioning of synapses.

The human nervous system:

In humans and other vertebrates, the nervous system can be broadly divided into two sections: **the central nervous system and the peripheral nervous system.**

The central nervous system (CNS) consists of the brain and the spinal cord. It is in the CNS that all of the analysis of information takes place.

The peripheral nervous system (PNS), which consists of the neurons and parts of neurons found outside of the CNS, includes sensory neurons and motor neurons. Sensory neurons bring signals into the CNS, and motor neurons carry signals out of the CNS.

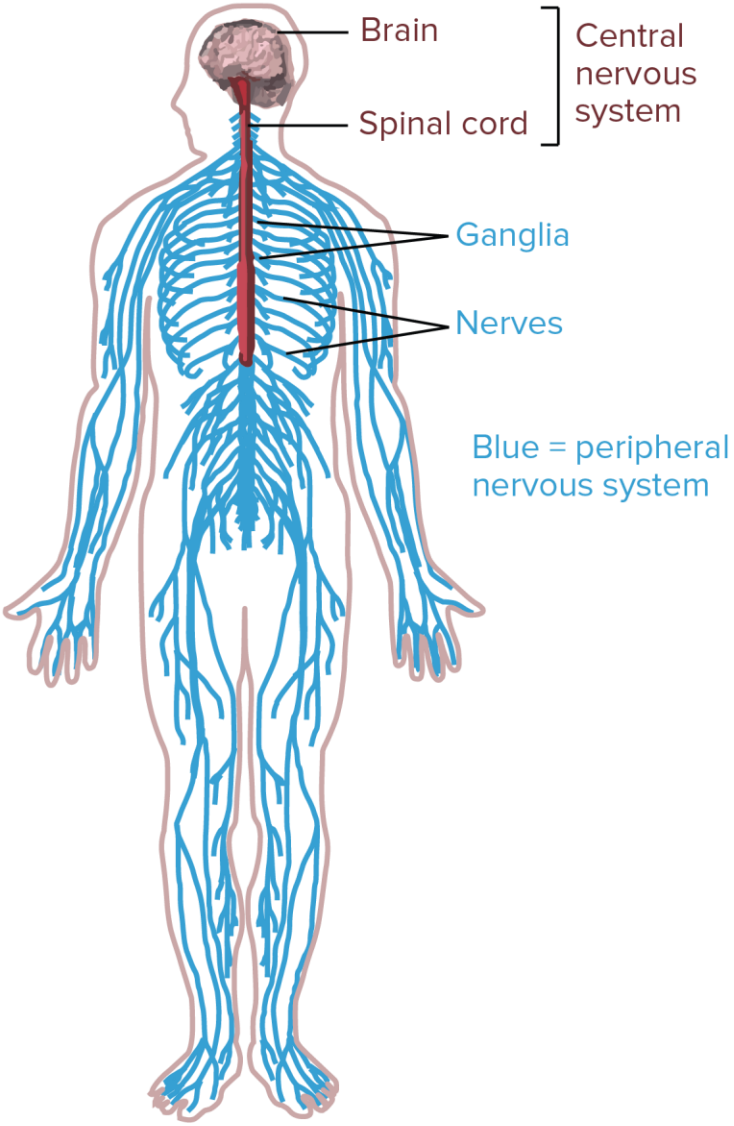


Diagram of the human nervous system.

Central nervous system: portions of the nervous system in the brain and spinal cord.

Peripheral nervous system: portions of the nervous system outside the brain and spinal cord.

Also marked on the diagram are ganglia, clusters of cell bodies bodies in the PNS, and nerves, bundles of axons that travel along the same route. The marked ganglia are located near, but not in, the spinal cord. The marked nerves are spinal nerves.

The cell bodies of some PNS neurons, such as the motor neurons that control skeletal muscle (the type of muscle found in your arm or leg), are located in the CNS. These motor neurons have long extensions (axons) that run from the CNS all the way to the muscles they connect with (innervate). The cell bodies of other PNS neurons, such as the sensory neurons that provide information about touch, position, pain, and temperature, are located outside of the CNS, where they are found in clusters known as ganglia.

The axons of peripheral neurons that travel a common route are bundled together to form nerves.

Classes of neurons

Based on their roles, the neurons found in the human nervous system can be divided into three classes: sensory neurons, motor neurons, and interneurons.

Sensory neurons

Sensory neurons get information about what's going on inside and outside of the body and bring that information into the CNS so it can be processed. For instance, if you picked up a hot coal, sensory neurons with endings in your fingertips would convey the information to your CNS that it was really hot.

Motor neurons

Motor neurons get information from other neurons and convey commands to your muscles, organs and glands. For instance, if you picked up a hot coal, it motor neurons innervating the muscles in your fingers would cause your hand to let go.

Interneurons

Interneurons, which are found only in the CNS, connect one neuron to another. They receive information from other neurons (either sensory neurons or interneurons) and transmit information to other neurons (either motor neurons or interneurons).

For instance, if you picked up a hot coal, the signal from the sensory neurons in your fingertips would travel to interneurons in your spinal cord. Some of these interneurons would signal to the motor neurons controlling your finger muscles (causing you to let go), while others would transmit the signal up the spinal cord to neurons in the brain, where it would be perceived as pain.

Interneurons are the most numerous class of neurons and are involved in processing information, both in simple reflex circuits (like those triggered by hot objects) and in more complex circuits in the brain. It would be combinations of interneurons in your brain that would allow you to draw the conclusion that things that looked like hot coals weren't good to pick up, and, hopefully, retain that information for future reference.

The basic functions of a neuron: If you think about the roles of the three classes of neurons, you can make the generalization that all neurons have three basic functions. These are to:

- Receive signals (or information).
- Integrate incoming signals (to determine whether or not the information should be passed along).
- Communicate signals to target cells (other neurons or muscles or glands).

These neuronal functions are reflected in the anatomy of the neuron.

Anatomy of a neuron

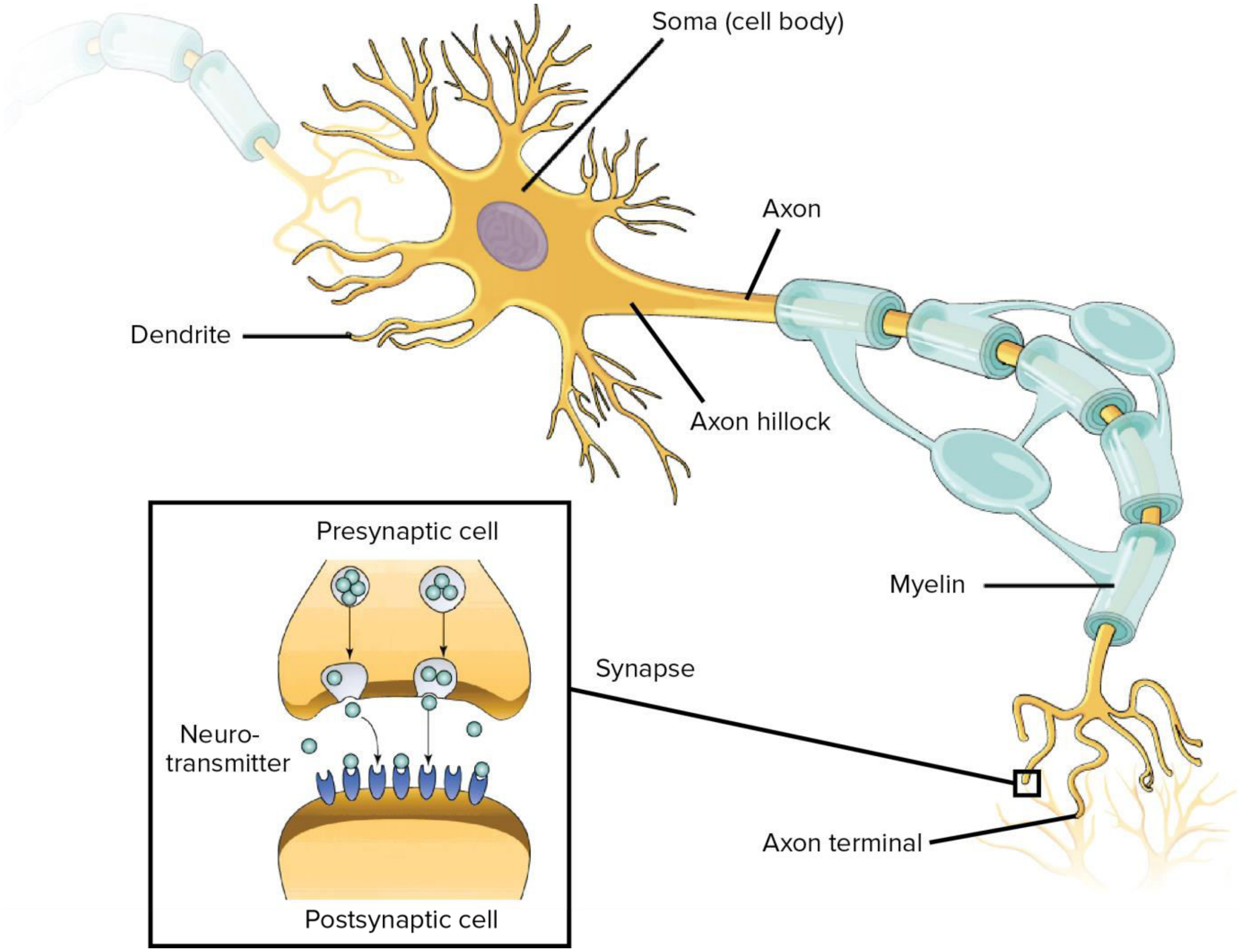
Neurons, like other cells, have a cell body (called the soma). The nucleus of the neuron is found in the soma. Neurons need to produce a lot of proteins, and most neuronal proteins are synthesized in the soma as well.

Various processes (appendages or protrusions) extend from the cell body. These include many short, branching processes, known as dendrites, and a separate process that is typically longer than the dendrites, known as the axon.

Dendrites

The first two neuronal functions, receiving and processing incoming information, generally take place in the dendrites and cell body. Incoming signals can be either excitatory – which means they tend to make the neuron fire (generate an electrical impulse) – or inhibitory – which means that they tend to keep the neuron from firing.

Most neurons receive many input signals throughout their dendritic trees. A single neuron may have more than one set of dendrites, and may receive many thousands of input signals. Whether or not a neuron is excited into firing an impulse depends on the sum of all of the excitatory and inhibitory signals it receives. If the neuron does end up firing, the nerve impulse, or action potential, is conducted down the axon.



Structure of a neuron.

At one end of the cell body (and indeed, around most of its periphery) are many small, branching protrusions called dendrites. Extending from the other end of the cell body at a location called the axon hillock is the axon, a long, thin, tube-like protrusion. The axon is wrapped in myelin, which ensheathes some sections but leaves sections of the axon bare between the sheathed portions.

At its far end, the axon splits up into many axon terminal. Each forms a synapse with a dendrite or cell body of another neuron. The cell to which the axon terminal belongs (sending cell) is called the presynaptic cell, while the cell to which the dendrite or cell body belongs (receiving cell) is called the postsynaptic cell. There is a space between the two cells, across which they communicate. When an action potential arrives at the axon terminal, it triggers the release of molecules of neurotransmitter from the presynaptic cell. These diffuse to to the other side of the synapse and bind to receptors on the membrane of the postsynaptic cell.

Axons: Axons differ from dendrites in several ways.

The dendrites tend to taper and are often covered with little bumps called spines. In contrast, the axon tends to stay the same diameter for most of its length and doesn't have spines.

The axon arises from the cell body at a specialized area called the axon hillock. In motor neurons and interneurons, it's at the axon hillock that the action potential is initiated.

Finally, many axons are covered with a special insulating substance called myelin, which helps them convey the nerve impulse rapidly. Myelin is never found on dendrites.

Towards its end, the axon splits up into many branches and develops bulbous swellings known as axon terminals (or nerve terminals). These axon terminals make connections on target cells.

Synapses: Neuron-to-neuron connections are made onto the dendrites and cell bodies of other neurons. These connections, known as synapses, are the sites at which information is carried from the first neuron, the presynaptic neuron, to the target neuron (the postsynaptic neuron). The synaptic connections between neurons and skeletal muscle cells are generally called neuromuscular junctions, and the connections between neurons and smooth muscle cells or glands are known as neuroeffector junctions.

At most synapses and junctions, information is transmitted in the form of chemical messengers called neurotransmitters. When an action potential travels down an axon and reaches the axon terminal, it triggers the release of neurotransmitter from the presynaptic cell. Neurotransmitter molecules cross the synapse and bind to membrane receptors on the postsynaptic cell, conveying an excitatory or inhibitory signal.

Thus, the third basic neuronal function – communicating information to target cells – is carried out by the axon and the axon terminals. Just as a single neuron may receive inputs from many presynaptic neurons, it may also make synaptic connections on numerous postsynaptic neurons via different axon terminals.

Neurons form networks

A single neuron can’t do very much by itself, and nervous system function depends on groups of neurons that work together. Individual neurons connect to other neurons to stimulate or inhibit their activity, forming circuits that can process incoming information and carry out a response. Neuronal circuits can be very simple, and composed of only a few neurons, or they can involve more complex neuronal networks.

The knee-jerk reflex

The simplest neuronal circuits are those that underlie muscle stretch responses, such as the knee-jerk reflex that occurs when someone hits the tendon below your knee (the patellar tendon) with a hammer. Tapping on that tendon stretches the quadriceps muscle of the thigh, stimulating the sensory neurons that innervate it to fire.

Axons from these sensory neurons extend to the spinal cord, where they connect to the motor neurons that establish connections with (innervate) the quadriceps. The sensory neurons send an excitatory signal to the motor neurons, causing them to fire too. The motor neurons, in turn, stimulate the quadriceps to contract, straightening the knee. In the knee-jerk reflex, the sensory neurons from a particular muscle connect directly to the motor neurons that innervate that same muscle, causing it to contract after it has been stretched.

Simplified diagram of neural circuits involved in the knee-jerk reflex. When the patellar tendon is tapped, the quadriceps muscle on the front of the thigh is stretched, activating a sensory neuron that wraps around a muscle cell. The sensory neuron's axon extends all the way into the spinal cord, where it synapses on two targets:

Motor neuron innervating the quadriceps muscle. The sensory neuron activates the motor neuron, causing the quadriceps muscle to contract.

Interneuron: The sensory neuron activates the interneuron. However, this interneuron is itself inhibitory, and the target it inhibits is a motor neuron traveling to the hamstring muscle on the back of the thigh. Thus, the activation of the sensory neuron serves to inhibit contraction in the hamstring muscle.

Sensory neurons from the quadriceps are also part of a circuit that causes relaxation of the hamstring, the muscle that antagonizes (opposes) the quadriceps. It wouldn't make sense for the sensory neurons of the quadriceps to activate the motor neurons of the hamstring, because that would make the hamstring contract, making it harder for the quadriceps to contract. Instead, the sensory neurons of the quadriceps connect to the motor neurons of the hamstring indirectly, through an inhibitory interneuron. Activation of the interneuron causes inhibition of the motor neurons that innervate the hamstring, making the hamstring muscle relax.

The sensory neurons of the quadriceps don't just participate in this reflex circuit. Instead, they also send messages to the brain, letting you know that someone tapped your tendon with a hammer and perhaps causing a response. (“Why did you do that?”) Although spinal cord circuits can mediate very simple behaviors like the knee jerk reflex, the ability to consciously perceive sensory stimuli – along with all of the higher functions of the nervous system – depends on the more complex neuronal networks found in the brain.

Glial cells

At the beginning of this article, we said that the nervous system was made up of two types of cells, neurons and glia, with the neurons acting as the basic functional unit of the nervous system and the glia playing a supporting role. Just as the supporting actors are essential to the success of a movie, the glia are essential to nervous system function. Indeed, there are many more glial cells in the brain than there are neurons.

There are four main types of glial cells in the adult vertebrate nervous system. Three of these, astrocytes, oligodendrocytes, and microglia, are found only in the central nervous system (CNS). The fourth, the Schwann cells, are found only in the peripheral nervous system (PNS).

Types of glia and their functions

Astrocytes are the most numerous type of glial cell. In fact, they are the most numerous cells in the brain! Astrocytes come in different types and have a variety of functions. They help regulate blood flow in the brain, maintain the composition of the fluid that surrounds neurons, and regulate communication between neurons at the synapse. During development, astrocytes help neurons find their way to their destinations and contribute to the formation of the blood-brain barrier, which helps isolate the brain from potentially toxic substances in the blood.

Microglia are related to the macrophages of the immune system and act as scavengers to remove dead cells and other debris.

The oligodendrocytes of the CNS and the Schwann cells of the PNS share a similar function. Both of these types of glial cells produce myelin, the insulating substance that forms a sheath around the axons of many neurons. Myelin dramatically increases the speed with which an action potential travels down the axon, and it plays a crucial role in nervous system function.

Left panel: Glia of the central nervous system. Astrocytes extend their "feet" (projections) onto the cell bodies of neurons, while oligodendrocytes form the myelin sheaths around the axons of neurons. Microglial cells hang around in the interstices, scavenging dead cells and debris. Ependymal cells line the ventricles of the brain and have projections (on the non-ventricle side of the ependymal layer) that link up with the "feet" of the astrocytes.

Right panel: Glia of the peripheral nervous system. The cell body of a sensory neuron in a ganglion is covered with a layer of satellite glial cells. Schwann cells myelinate the single process extending from the cell body, as well as the two processes produced by the splitting of that single process (one of which will have axon terminals at its end, and the other of which will have dendrites at its end.

Other types of glia (in addition to the four main types) include satellite glial cells and ependymal cells.

Satellite glial cells cover the cell bodies of neurons in PNS ganglia. Satellite glial cells are thought to support the function of the neurons and might act as a protective barrier, but their role is still not well-understood.

Ependymal cells, which line the ventricles of the brain and the central canal of the spinal cord, have hairlike cilia that beat to promote circulation of the cerebrospinal fluid found inside the ventricles and spinal canal.

Artificial neuron:

An artificial neuron is a mathematical function conceived as a model of biological neurons, a neural network. Artificial neurons are elementary units in an artificial neural network. The artificial neuron receives one or more inputs (representing excitatory postsynaptic potentials and inhibitory postsynaptic potentials at neural dendrites) and sums them to produce an output (or activation, representing a neuron's action potential which is transmitted along its axon). Usually each input is separately weighted, and the sum is passed through a non-linear function known as an activation function or transfer function[clarification needed]. The transfer functions usually have a sigmoid shape, but they may also take the form of other non-linear functions, piecewise linear functions, or step functions. They are also often monotonically increasing, continuous, differentiable and bounded. The thresholding function has inspired building logic gates referred to as threshold logic; applicable to building logic circuits resembling brain processing.

Neural networks are parallel computing devices, which is basically an attempt to make a computer model of the brain. The main objective is to develop a system to perform various computational tasks faster than the traditional systems. These tasks include pattern recognition and classification, approximation, optimization, and data clustering.

What is Artificial Neural Network?

Artificial Neural Network (ANN) is an efficient computing system whose central theme is borrowed from the analogy of biological neural networks. ANNs are also named as “artificial neural systems,” or “parallel distributed processing systems,” or “connectionist systems.” ANN acquires a large collection of units that are interconnected in some pattern to allow communication between the units. These units, also referred to as nodes or neurons, are simple processors which operate in parallel.

Every neuron is connected with other neuron through a connection link. Each connection link is associated with a weight that has information about the input signal. This is the most useful information for neurons to solve a particular problem because the weight usually excites or inhibits the signal that is being communicated. Each neuron has an internal state, which is called an activation signal. Output signals, which are produced after combining the input signals and activation rule, may be sent to other units.

A Brief History of ANN:

The history of ANN can be divided into the following three eras −

ANN during 1940s to 1960s

Some key developments of this era are as follows −

1943 − It has been assumed that the concept of neural network started with the work of physiologist, Warren McCulloch, and mathematician, Walter Pitts, when in 1943 they modeled a simple neural network using electrical circuits in order to describe how neurons in the brain might work.

1949 − Donald Hebb's book, The Organization of Behavior, put forth the fact that repeated activation of one neuron by another increases its strength each time they are used.

1956 − An associative memory network was introduced by Taylor.

1958 − A learning method for McCulloch and Pitts neuron model named Perceptron was invented by Rosenblatt.

1960 − Bernard Widrow and Marcian Hoff developed models called "ADALINE" and "MADALINE."

ANN during 1960s to 1980s

Some key developments of this era are as follows −

1961 − Rosenblatt made an unsuccessful attempt but proposed the "backpropagation" scheme for multilayer networks.

1964 − Taylor constructed a winner-take-all circuit with inhibitions among output units.

1969 − Multilayer perceptron (MLP) was invented by Minsky and Papert.

1971 − Kohonen developed Associative memories.

1976 − Stephen Grossberg and Gail Carpenter developed Adaptive resonance theory.

ANN from 1980s till Present

Some key developments of this era are as follows −

1982 − The major development was Hopfield's Energy approach.

1985 − Boltzmann machine was developed by Ackley, Hinton, and Sejnowski.

1986 − Rumelhart, Hinton, and Williams introduced Generalised Delta Rule.

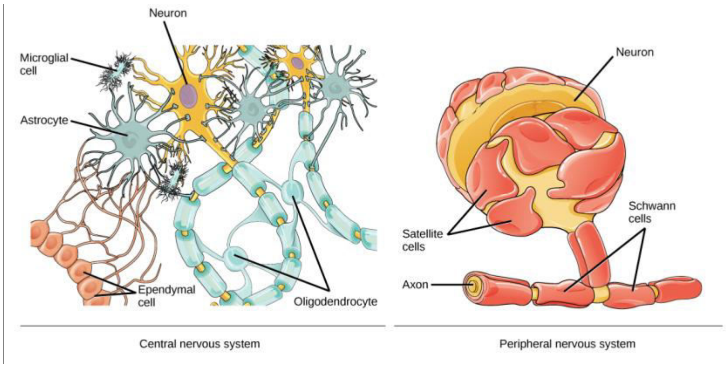
1988 − Kosko developed Binary Associative Memory (BAM) and also gave the concept of Fuzzy Logic in ANN.

The historical review shows that significant progress has been made in this field. Neural network based chips are emerging and applications to complex problems are being developed. Surely, today is a period of transition for neural network technology.

Biological Neuron

A nerve cell (neuron) is a special biological cell that processes information. According to an estimation, there are huge number of neurons, approximately 1011 with numerous interconnections, approximately 1015.

Schematic Diagram



Working of a Biological Neuron

As shown in the above diagram, a typical neuron consists of the following four parts with the help of which we can explain its working −

Dendrites − They are tree-like branches, responsible for receiving the information from other neurons it is connected to. In other sense, we can say that they are like the ears of neuron.

Soma − It is the cell body of the neuron and is responsible for processing of information, they have received from dendrites.

Axon − It is just like a cable through which neurons send the information.

Synapses − It is the connection between the axon and other neuron dendrites.

ANN versus BNN

Before taking a look at the differences between Artificial Neural Network (ANN) and Biological Neural Network (BNN), let us take a look at the similarities based on the terminology between these two.

Biological Neural	Artificial Neural

Network (BNN)	Network (ANN)
Soma	**Node**
Dendrites	**Input**
Synapse	**Weights or Interconnections**
Axon	**Output**

The following table shows the comparison between ANN and BNN based on some criteria mentioned.

Criteria	BNN	ANN
Processing	**Massively parallel, slow but superior than ANN**	**Massively parallel, fast but inferior than BNN**
Size	**1011 neurons and 1015interconnections**	**102 to 104 nodes (mainly depends on the type of application and network designer)**
Learning	**They can tolerate ambiguity**	**Very precise, structured and formatted data is required to tolerate ambiguity**
Fault tolerance	**Performance degrades with even partial damage**	**It is capable of robust performance, hence has the potential to be fault tolerant**
Storage	**Stores the information in**	**Stores the information in**

capacity	**the synapse**	**continuous memory locations**

Model of Artificial Neural Network

Processing of ANN depends upon the following three building blocks −

- Network Topology
- Adjustments of Weights or Learning
- Activation Functions

Network Topology

A network topology is the arrangement of a network along with its nodes and connecting lines. According to the topology, ANN can be classified as the following kinds −

Feedforward Network

It is a non-recurrent network having processing units/nodes in layers and all the nodes in a layer are connected with the nodes of the previous layers. The connection has different weights upon them. There is no feedback loop means the signal can only flow in one direction, from input to output. It may be divided into the following two types −

Single layer feedforward network − The concept is of feedforward ANN having only one weighted layer. In other words, we can say the input layer is fully connected to the output layer.

Multilayer feedforward network − The concept is of feedforward ANN having more than one weighted layer. As this network has one or more layers between the input and the output layer, it is called hidden layers.

Feedback Network:

As the name suggests, a feedback network has feedback paths, which means the signal can flow in both directions using loops. This makes it a non-linear dynamic system, which changes continuously until it reaches a state of equilibrium. It may be divided into the following types −

Recurrent networks − They are feedback networks with closed loops. Following are the two types of recurrent networks.

Fully recurrent network − It is the simplest neural network architecture because all nodes are connected to all other nodes and each node works as both input and output.

Jordan network − It is a closed loop network in which the output will go to the input again as feedback as shown in the following diagram.

Adjustments of Weights or Learning

Learning, in artificial neural network, is the method of modifying the weights of connections between the neurons of a specified network. Learning in ANN can be classified into three categories namely supervised learning, unsupervised learning, and reinforcement learning.

Supervised Learning

As the name suggests, this type of learning is done under the supervision of a teacher. This learning process is dependent.

During the training of ANN under supervised learning, the input vector is presented to the network, which will give an output vector. This output vector is compared with the desired output vector. An error signal is generated, if there is a difference between the actual output and the desired output vector. On the basis of this error signal, the weights are adjusted until the actual output is matched with the desired output.

Unsupervised Learning

As the name suggests, this type of learning is done without the supervision of a teacher. This learning process is independent.

During the training of ANN under unsupervised learning, the input vectors of similar type are combined to form clusters. When a new input pattern is applied, then the neural network gives an output response indicating the class to which the input pattern belongs.

There is no feedback from the environment as to what should be the desired output and if it is correct or incorrect. Hence, in this type of learning, the network itself must discover the patterns and features from the input data, and the relation for the input data over the output.

Reinforcement Learning

As the name suggests, this type of learning is used to reinforce or strengthen the network over some critic information. This learning process is similar to supervised learning, however we might have very less information.

During the training of network under reinforcement learning, the network receives some feedback from the environment. This makes it somewhat similar to supervised learning. However, the feedback obtained here is evaluative not instructive, which means there is no teacher as in supervised learning. After receiving the feedback, the network performs adjustments of the weights to get better critic information in future.

Activation Functions

It may be defined as the extra force or effort applied over the input to obtain an exact output. In ANN, we can also apply activation functions over the input to get the exact output. Followings are some activation functions of interest −

Linear Activation Function

It is also called the identity function as it performs no input editing. It can be defined as −

$$F(x)\:=\:x$$

Sigmoid Activation Function

It is of two type as follows −

Binary sigmoidal function − This activation function performs input editing between 0 and 1. It is positive in nature. It is always bounded, which means its output cannot be less than 0 and more than 1. It is also strictly increasing in nature, which means more the input higher would be the output.

Bipolar sigmoidal function − This activation function performs input editing between -1 and 1. It can be positive or negative in nature. It is always bounded, which means its output cannot be less than -1 and more than 1. It is also strictly increasing in nature like sigmoid function.

What Is Learning in ANN?

Basically, learning means to do and adapt the change in itself as and when there is a change in environment. ANN is a complex system or more precisely we can say that it is a complex adaptive system, which can change its internal structure based on the information passing through it.

Why Is It important?

Being a complex adaptive system, learning in ANN implies that a processing unit is capable of changing its input/output behavior due to the change in environment. The importance of learning in ANN increases because of the fixed activation function as well as the input/output vector, when a particular network is constructed. Now to change the input/output behavior, we need to adjust the weights.

Classification

It may be defined as the process of learning to distinguish the data of samples into different classes by finding common features between the samples of the same classes. For example, to perform training of ANN, we have some training samples with unique features, and to perform its testing we have some testing samples with other unique features. Classification is an example of supervised learning.

Neural Network Learning Rules

We know that, during ANN learning, to change the input/output behavior, we need to adjust the weights. Hence, a method is required with the help of which the weights can be modified. These methods are called Learning rules, which are simply algorithms or equations. Following are some learning rules for the neural network −

Hebbian Learning Rule

This rule, one of the oldest and simplest, was introduced by Donald Hebb in his book The Organization of Behavior in 1949. It is a kind of feed-forward, unsupervised learning.

Basic Concept − This rule is based on a proposal given by Hebb, who wrote −

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”

From the above postulate, we can conclude that the connections between two neurons might be strengthened if the neurons fire at the same time and might weaken if they fire at different times.

Mathematical Formulation − According to Hebbian learning rule, following is the formula to increase the weight of connection at every time step.

$$\Delta w_{ji}(t)\:=\:\alpha x_{i}(t).y_{j}(t)$$

Here, $\Delta w_{ji}(t)$ = increment by which the weight of connection increases at time step t

α = the positive and constant learning rate

$x_{i}(t)$ = the input value from pre-synaptic neuron at time step t

$y_{i}(t)$ = the output of pre-synaptic neuron at same time step t

Perceptron Learning Rule

This rule is an error correcting the supervised learning algorithm of single layer feedforward networks with linear activation function, introduced by Rosenblatt.

Basic Concept − As being supervised in nature, to calculate the error, there would be a comparison between the desired/target output and the actual output. If there is any difference found, then a change must be made to the weights of connection.

Basic Concept of Competitive Learning Rule − As said earlier, there will be a competition among the output nodes. Hence, the main concept is that during training, the output unit with the highest activation to a given input pattern, will be declared the winner. This rule is also called Winner-takes-all because only the winning neuron is updated and the rest of the neurons are left unchanged.

Supervised learning:

As the name suggests, supervised learning takes place under the supervision of a teacher. This learning process is dependent. During the training of ANN under supervised learning, the input vector is presented to the network, which will produce an output vector. This output vector is compared with the desired/target output vector. An error signal is generated if there is a difference between the actual output and the desired/target output vector. On the basis of this error signal, the weights would be adjusted until the actual output is matched with the desired output.

Perceptron

Developed by Frank Rosenblatt by using McCulloch and Pitts model, perceptron is the basic operational unit of artificial neural networks. It employs supervised learning rule and is able to classify the data into two classes.

Operational characteristics of the perceptron: It consists of a single neuron with an arbitrary number of inputs along with adjustable weights, but the output of the neuron is 1 or 0 depending upon the threshold. It also consists of a bias whose weight is always 1

Adaptive Linear Neuron (Adaline)

Adaline which stands for Adaptive Linear Neuron, is a network having a single linear unit. It was developed by Widrow and Hoff in 1960. Some important points about Adaline are as follows −

It uses bipolar activation function.

It uses delta rule for training to minimize the Mean-Squared Error (MSE) between the actual output and the desired/target output.

The weights and the bias are adjustable.

Architecture

The basic structure of Adaline is similar to perceptron having an extra feedback loop with the help of which the actual output is compared with the desired/target output. After comparison on the basis of training algorithm, the weights and bias will be updated.

Step 8 − Test for the stopping condition, which will happen when there is no change in weight or the highest weight change occurred during training is smaller than the specified tolerance.

Multiple Adaptive Linear Neuron (Madaline)

Madaline which stands for Multiple Adaptive Linear Neuron, is a network which consists of many Adalines in parallel. It will have a single output unit. Some important points about Madaline are as follows −

It is just like a multilayer perceptron, where Adaline will act as a hidden unit between the input and the Madaline layer.

The weights and the bias between the input and Adaline layers, as in we see in the Adaline architecture, are adjustable.

The Adaline and Madaline layers have fixed weights and bias of 1.

Training can be done with the help of Delta rule.

Architecture

The architecture of Madaline consists of "n" neurons of the input layer, "m"neurons of the Adaline layer, and 1 neuron of the Madaline layer. The Adaline layer can be considered as the hidden layer as it is between the input layer and the output layer, i.e. the Madaline layer.

Back Propagation Neural Networks

Back Propagation Neural (BPN) is a multilayer neural network consisting of the input layer, at least one hidden layer and output layer. As its name suggests, back propagating will take place in this network. The error which is calculated at the output layer, by comparing the target output and the actual output, will be propagated back towards the input layer.

Architecture

As shown in the diagram, the architecture of BPN has three interconnected layers having weights on them. The hidden layer as well as the output layer also has bias, whose weight is always 1, on them. As is clear from the diagram, the working of BPN is in two phases. One phase sends the signal from the input layer to the output layer, and the other phase back propagates the error from the output layer to the input layer.

Training Algorithm

For training, BPN will use binary sigmoid activation function. The training of BPN will have the following three phases.

Phase 1 − Feed Forward Phase

Phase 2 − Back Propagation of error

Phase 3 − Updating of weights

All these steps will be concluded in the algorithm as follows

Step 1 − Initialize the following to start the training −

Weights

Learning rate α

For easy calculation and simplicity, take some small random values.

Step 2 − Continue step 3-11 when the stopping condition is not true.

Step 3 − Continue step 4-10 for every training pair.

Phase 1

Step 4 − Each input unit receives input signal xi and sends it to the hidden unit for all i = 1 to n

Step 5 − Calculate the net input at the hidden unit using the following relation −

Generalized Delta Learning Rule

Delta rule works only for the output layer. On the other hand, generalized delta rule, also called as back-propagation rule, is a way of creating the desired values of the hidden layer.

Unsupervised learning:

As the name suggests, this type of learning is done without the supervision of a teacher. This learning process is independent. During the training of ANN under unsupervised learning, the input vectors of similar type are combined to form clusters. When a new input pattern is applied, then the neural network gives an output response indicating the class to which input pattern belongs. In this, there would be no feedback from the environment as to what should be the desired output and whether it is correct or incorrect. Hence, in this type of learning the network itself must discover the patterns, features from the input data and the relation for the input data over the output.

Winner-Takes-All Networks

These kinds of networks are based on the competitive learning rule and will use the strategy where it chooses the neuron with the greatest total inputs as a winner. The connections between the output neurons show the competition between them and one of them would be ‘ON’ which means it would be the winner and others would be ‘OFF’.

Following are some of the networks based on this simple concept using unsupervised learning.

Hamming Network

In most of the neural networks using unsupervised learning, it is essential to compute the distance and perform comparisons. This kind of network is Hamming network, where for every given input vectors, it would be clustered into different groups. Following are some important features of Hamming Networks −

Lippmann started working on Hamming networks in 1987.

It is a single layer network.

The inputs can be either binary {0, 1} of bipolar {-1, 1}.

The weights of the net are calculated by the exemplar vectors.

It is a fixed weight network which means the weights would remain the same even during training.

Max Net

This is also a fixed weight network, which serves as a subnet for selecting the node having the highest input. All the nodes are fully interconnected and there exists symmetrical weights in all these weighted interconnections.

Architecture

It uses the mechanism which is an iterative process and each node receives inhibitory inputs from all other nodes through connections. The single node whose value is maximum would be active or winner and the activations of all other nodes would be inactive. Max Net uses identity activation function with $$f(x)\:=\:\begin{cases}x & if\:x > 0\\0 & if\:x \leq 0\end{cases}$$

The task of this net is accomplished by the self-excitation weight of +1 and mutual inhibition magnitude, which is set like [0 < ε < $\frac{1}{m}$] where “m” is the total number of the nodes.

Competitive Learning in ANN

It is concerned with unsupervised training in which the output nodes try to compete with each other to represent the input pattern. To understand this learning rule we will have to understand competitive net which is explained as follows −

Basic Concept of Competitive Network

This network is just like a single layer feed-forward network having feedback connection between the outputs. The connections between the outputs are inhibitory type, which is shown by dotted lines, which means the competitors never support themselves.

Basic Concept of Competitive Learning Rule

As said earlier, there would be competition among the output nodes so the main concept is - during training, the output unit that has the highest activation to a given input pattern, will be declared the winner. This rule is also called Winner-takes-all because only the winning neuron is updated and the rest of the neurons are left unchange**d.**

Mathematical Formulation

Following are the three important factors for mathematical formulation of this learning rule −

Condition to be a winner

Suppose if a neuron yk wants to be the winner, then there would be the following condition

$$y_{k}\:=\:\begin{cases}1 & if\:v_{k} > v_{j}\:for\:all\:\:j,\:j\:\neq\:k\\0 & otherwise\end{cases}$$

It means that if any neuron, say, yk wants to win, then its induced local field (the output of the summation unit), say vk, must be the largest among all the other neurons in the network.

Condition of the sum total of weight

Another constraint over the competitive learning rule is the sum total of weights to a particular output neuron is going to be 1. For example, if we consider neuron k then

$$\displaystyle\sum\limits_{k} w_{kj}\:=\:1\:\:\:\:for\:all\:\:k$$

Change of weight for the winner

If a neuron does not respond to the input pattern, then no learning takes place in that neuron. However, if a particular neuron wins, then the corresponding weights are adjusted as follows −

$$\Delta w_{kj}\:=\:\begin{cases}-\alpha(x_{j}\:-\:w_{kj}), & if\:neuron\:k\:wins\\0 & if\:neuron\:k\:losses\end{cases}$$

Here α is the learning rate.

This clearly shows that we are favoring the winning neuron by adjusting its weight and if a neuron is lost, then we need not bother to re-adjust its weight.

0\end{cases}$$

Calculations in C-cell

The net input of C-layer is

$$C\:=\:\displaystyle\sum\limits_i s_{i}x_{i}$$

Here, si is the output from S-cell and xi is the fixed weight from S-cell to C-cell.

The final output is as follows −

$$C_{out}\:=\:\begin{cases}\frac{C}{a+C}, & if\:C > 0\\0, & otherwise\end{cases}$$

Here ‘a’ is the parameter that depends on the performance of the network.