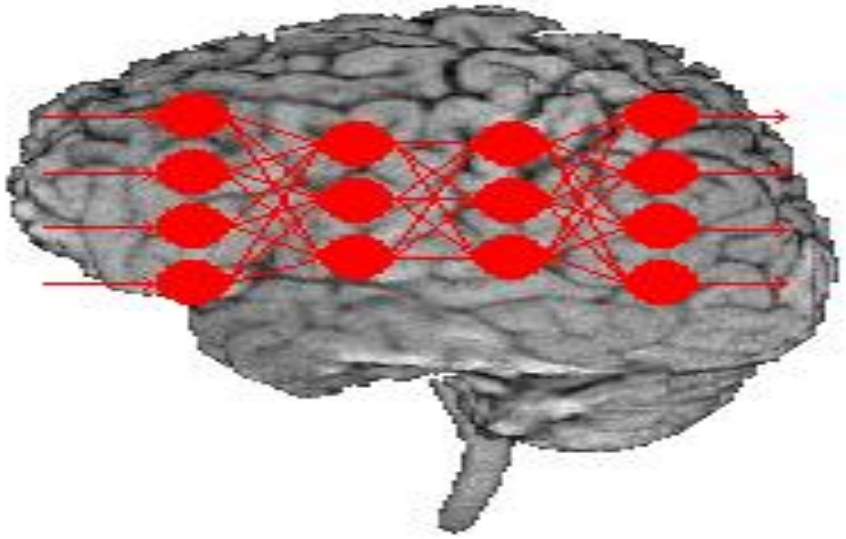
Unit 1

Introduction to Neural Network



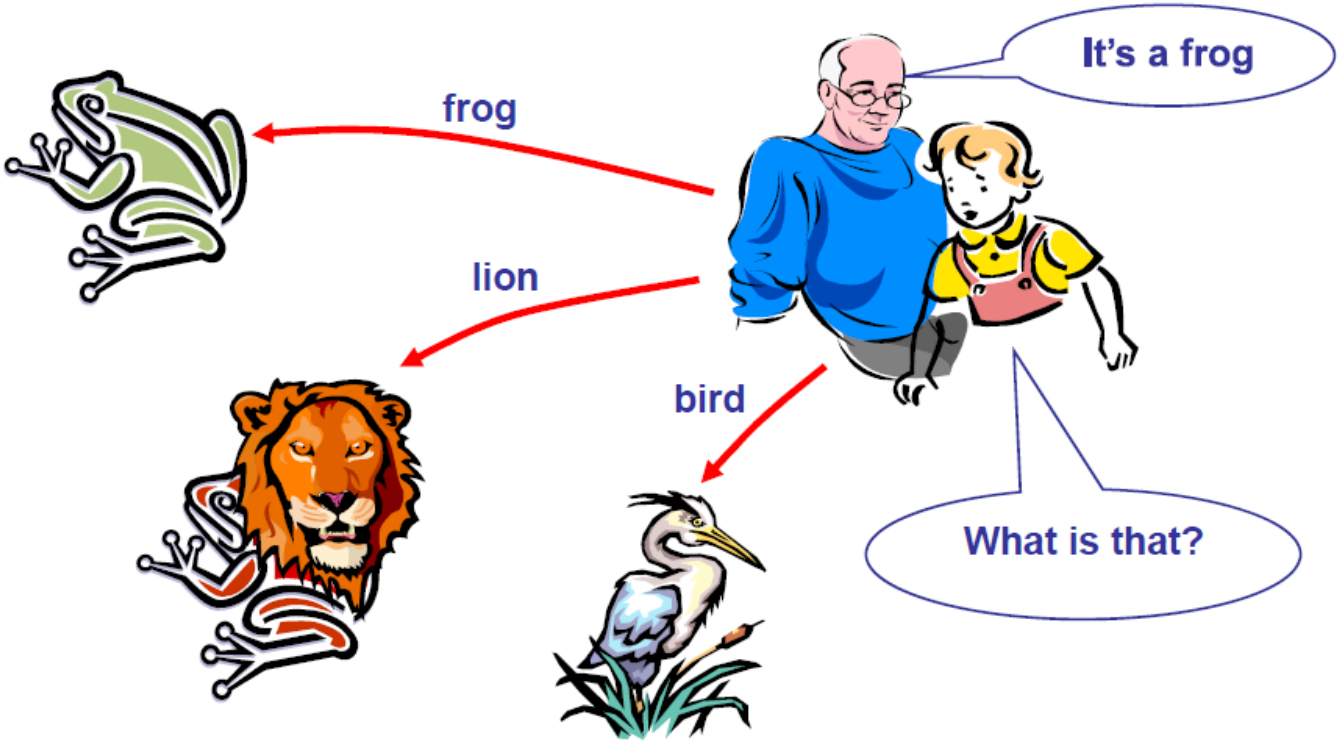
Er. Sachita Nand Mishra

Motivation

- Brain computation compared to digital computers
  - Neurons are 5-6 orders of magnitude slower than digital logic (ms vs. ns)
  - Brain has a huge number of neurons (10 billion neurons and about 60 000 billions of interconnections)
  - Brain is enormously energy efficient (10^{-16} J per operation per second vs 10^{-6} J per operation per second)
  - Brain is a very complex, nonlinear, parallel computer

The idea of ANNs..?

- NNs learn relationship between cause and effect or organize large volumes of data into orderly and informative patterns.



What is a Neural Network?

- **Neural network: I**nformation processing paradigm inspired by biological nervous systems, such as our brain
- Structure: large number of highly interconnected processing elements (neurons) working together
- Like people, they learn from experience (by example)

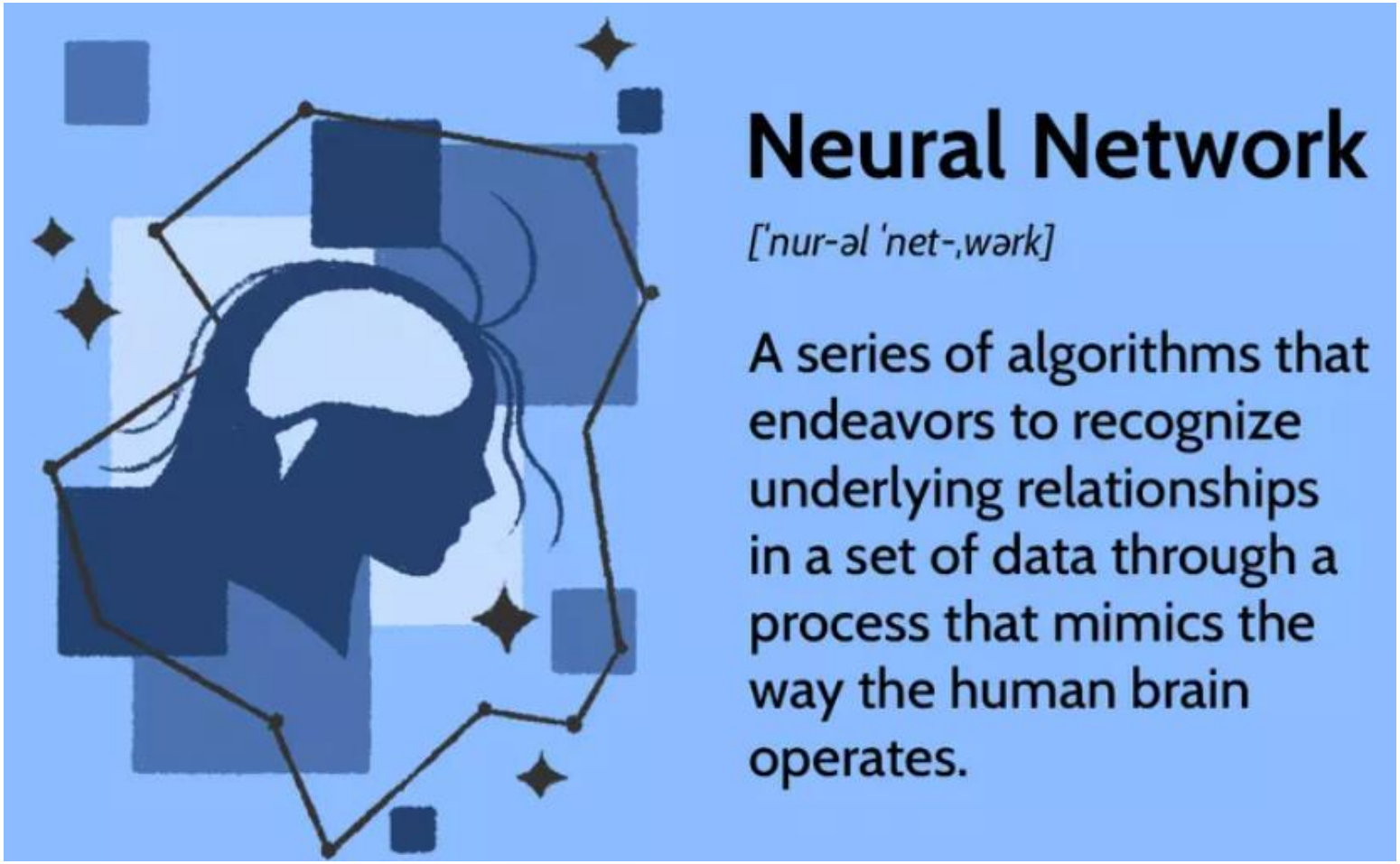
What is a Neural Network?

- Definition of ANN
- “Data processing system consisting of a large number of simple,highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex of the brain”

What is a Neural Network?

- An **artificial neural network** learning algorithm, or **neural network**, or just **neural net**, is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output, usually in another form.
- The concept of the artificial neural network was inspired by human biology and the way neurons of the human brain function together to understand inputs from human senses.

What is a Neural Network?



What is a Neural Network?

Similar to humans in that a nueral network constantly adjusts based on changing inputs or situations.

Hidden layer that processes the data

Constantly changing input data

Consistent output data used for analysis

Biological Neural Network Vs Artificial Neural Network

Characteristics	Artificial neural network	Biological neural network
Speed	Faster in processing information	Slower in processing information
Processing	In a sequential mode(One after another)	Parallel operations
Size and complexity	Small size and less complex	Big size and more complex
Fault tolerance	Not fault tolerant(information corrupted in the memory cannot be retrieved)	Tolerant
Control mechanism	Control unit monitors all the activities	There is no central control for processing information in the brain

Brain Vs Artificial Neural Network

	Brain	ANN
Speed	Few ms.	Few nano sec. massive ||el processing
Size and complexity	10^{11} neurons & 10^{15} interconnections	Depends on designer
Storage capacity	Stores information in its interconnection or in synapse. No Loss of memory	Contiguous memory locations loss of memory may happen sometimes.
Tolerance	Has fault tolerance	No fault tolerance Inf gets disrupted when interconnections are disconnected
Control mechanism	Complicated involves chemicals in biological neuron	Simpler in ANN

Why Artificial Neural Networks?

- There are two basic reasons why we are interested in building artificial neural networks (ANNs):

- **Technical viewpoint:** Some problems such as character recognition or the prediction of future states of a system require massively parallel and adaptive processing.

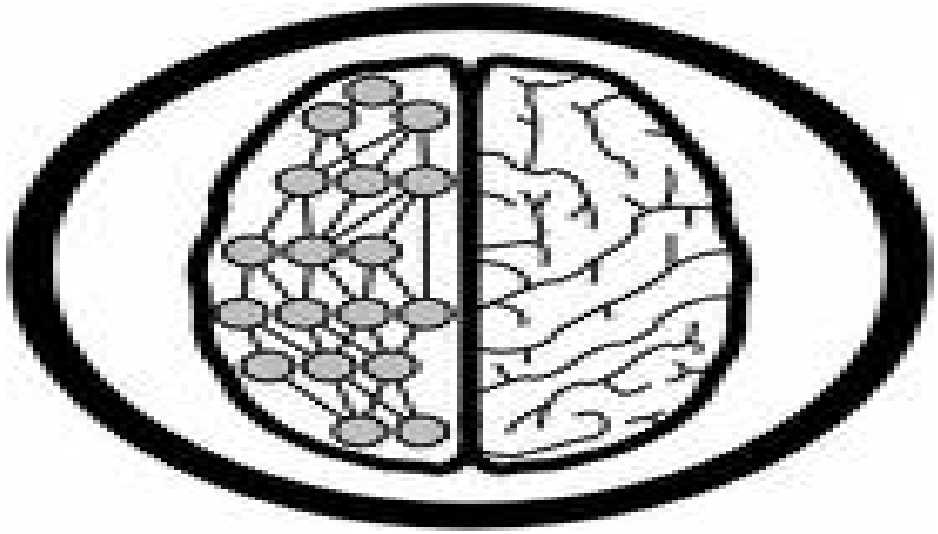
- **Biological viewpoint:** ANNs can be used to replicate and simulate components of the human (or animal) brain, thereby giving us insight into natural information processing.

Artificial Neural Networks

- The “building blocks” of neural networks are the **neurons**.
  - In technical systems, we also refer to them as **units** or **nodes**.
- Basically, each neuron
  - receives **input** from many other neurons.
  - changes its internal state (**activation**) based on the current input.
  - sends **one output signal** to many other neurons, possibly including its input neurons (recurrent network).

How do ANNs work?

- An artificial neural network (ANN) is either a hardware implementation or a computer program which strives to simulate the information processing capabilities of its biological exemplar. ANNs are typically composed of a great number of interconnected artificial neurons. The artificial neurons are simplified models of their biological counterparts.
- ANN is a technique for solving problems by constructing software that works like our brains.

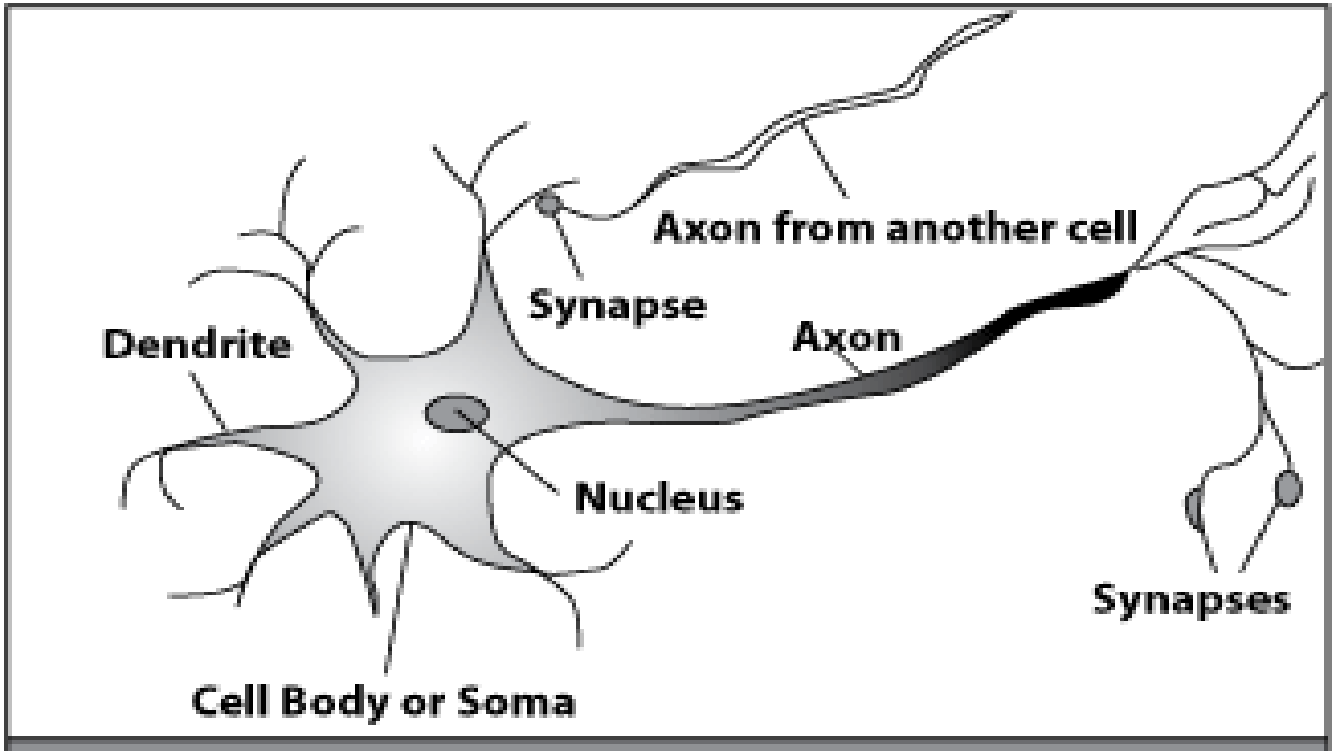


How do our brains work?

- The Brain is A massively parallel information processing system.
- Our brains are a huge network of processing elements.
- A typical brain contains a network of 10 billion neurons.

How do our brains work?

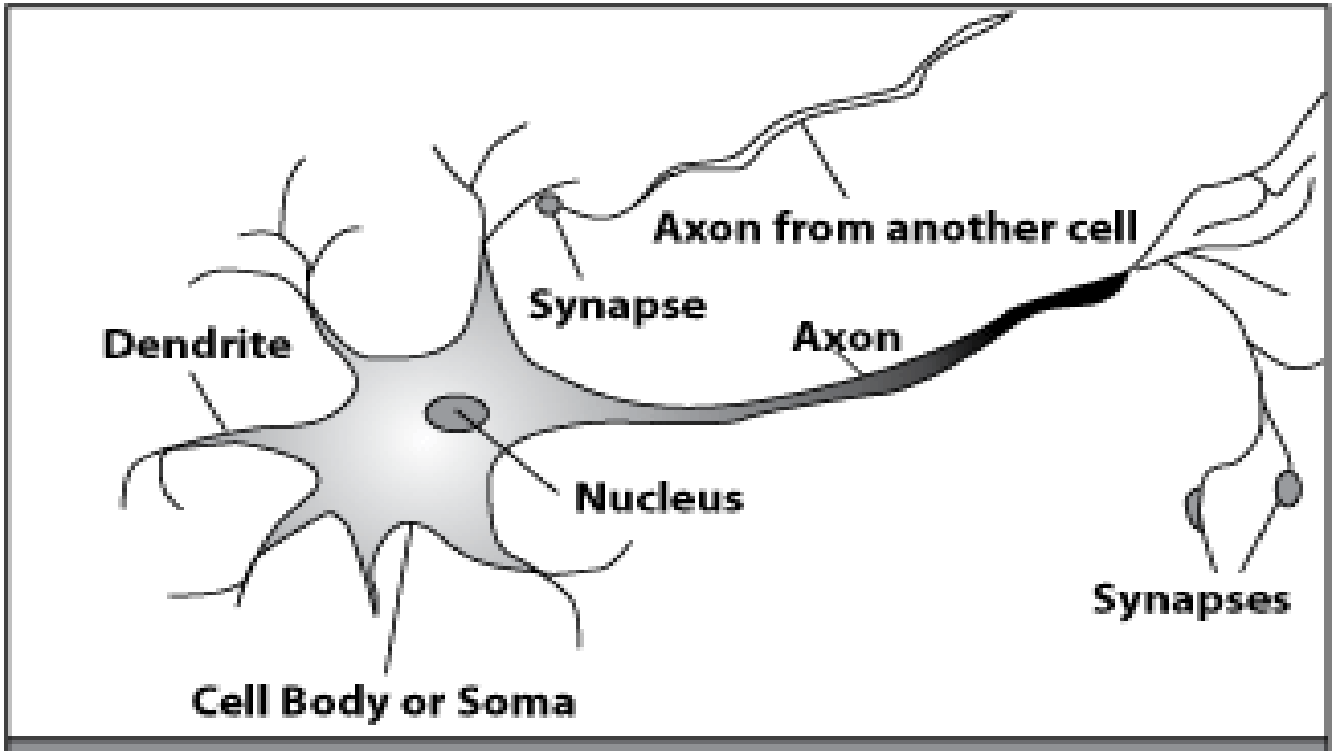
- A processing element



Dendrites: Input
Cell body: Processor
Synaptic: Link
Axon: Output

How do our brains work?

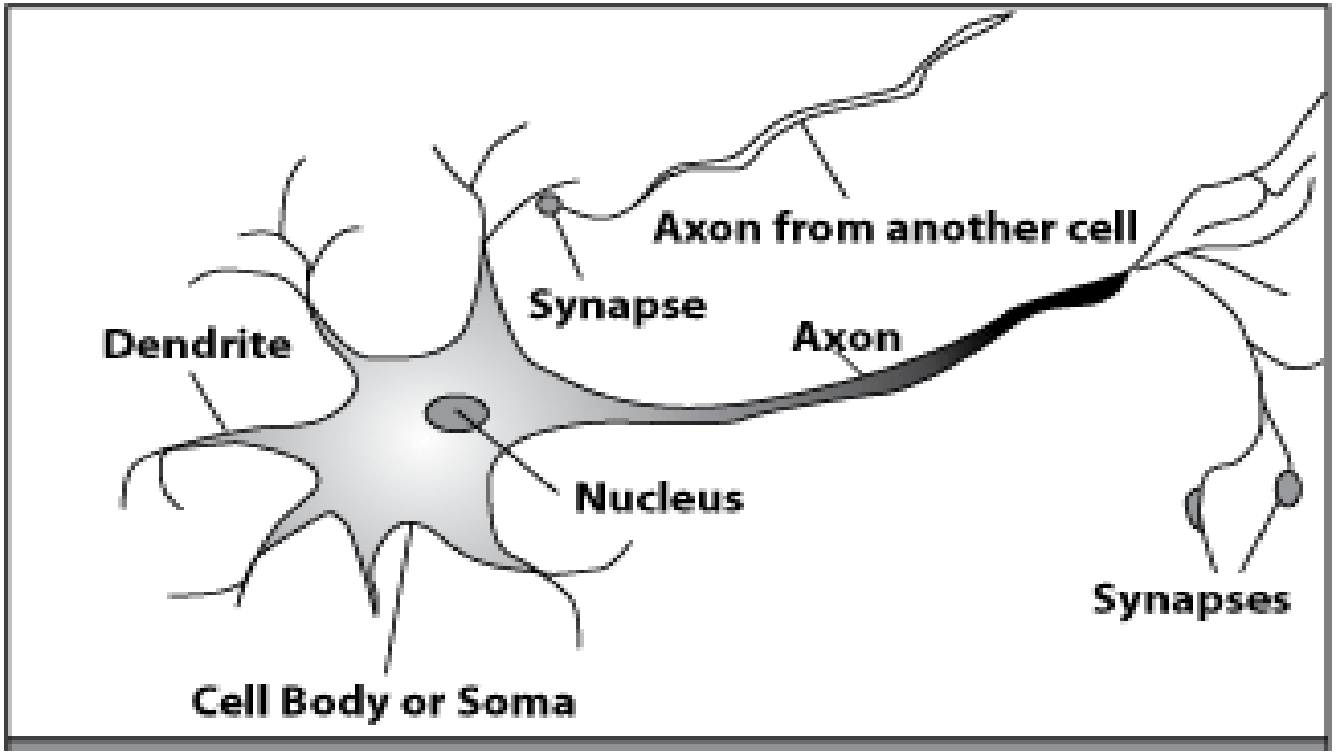
- A processing element



A neuron is connected to other neurons through about *10,000 synapses*

How do our brains work?

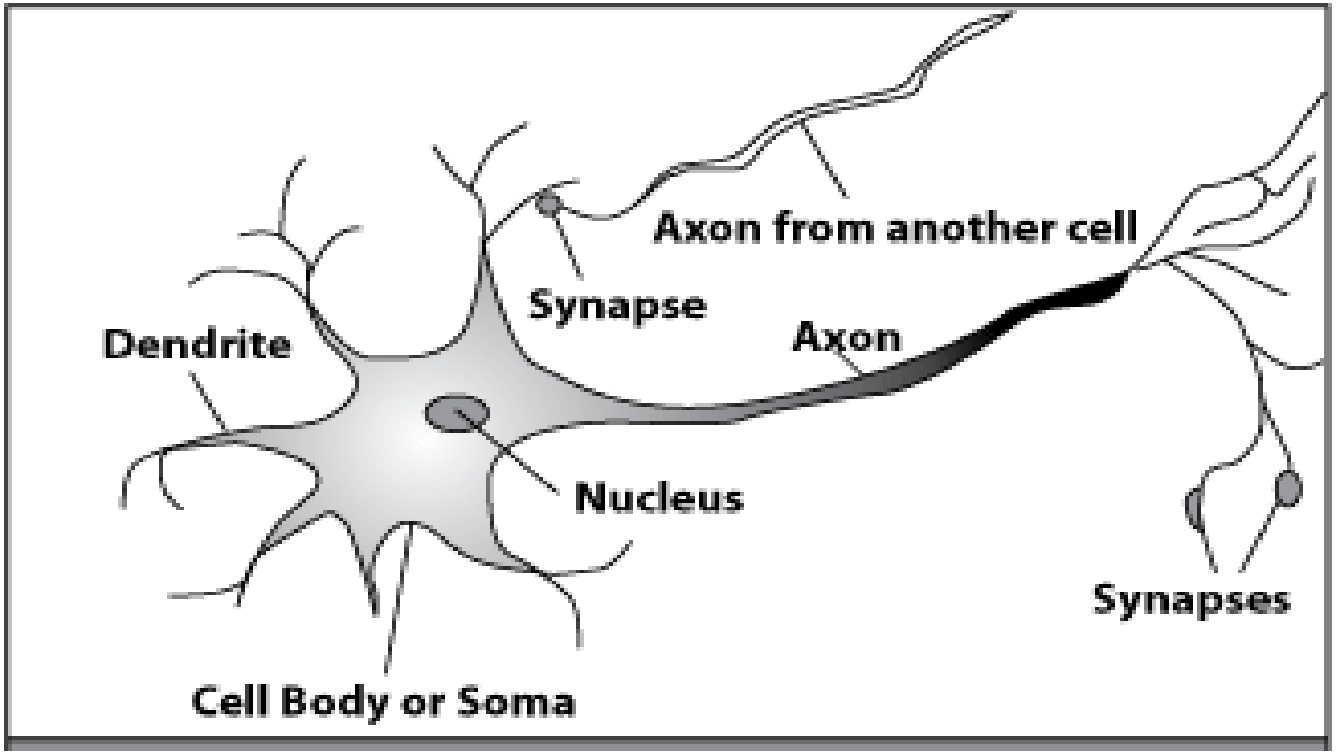
- A processing element



A neuron receives input from other neurons. Inputs are combined.

How do our brains work?

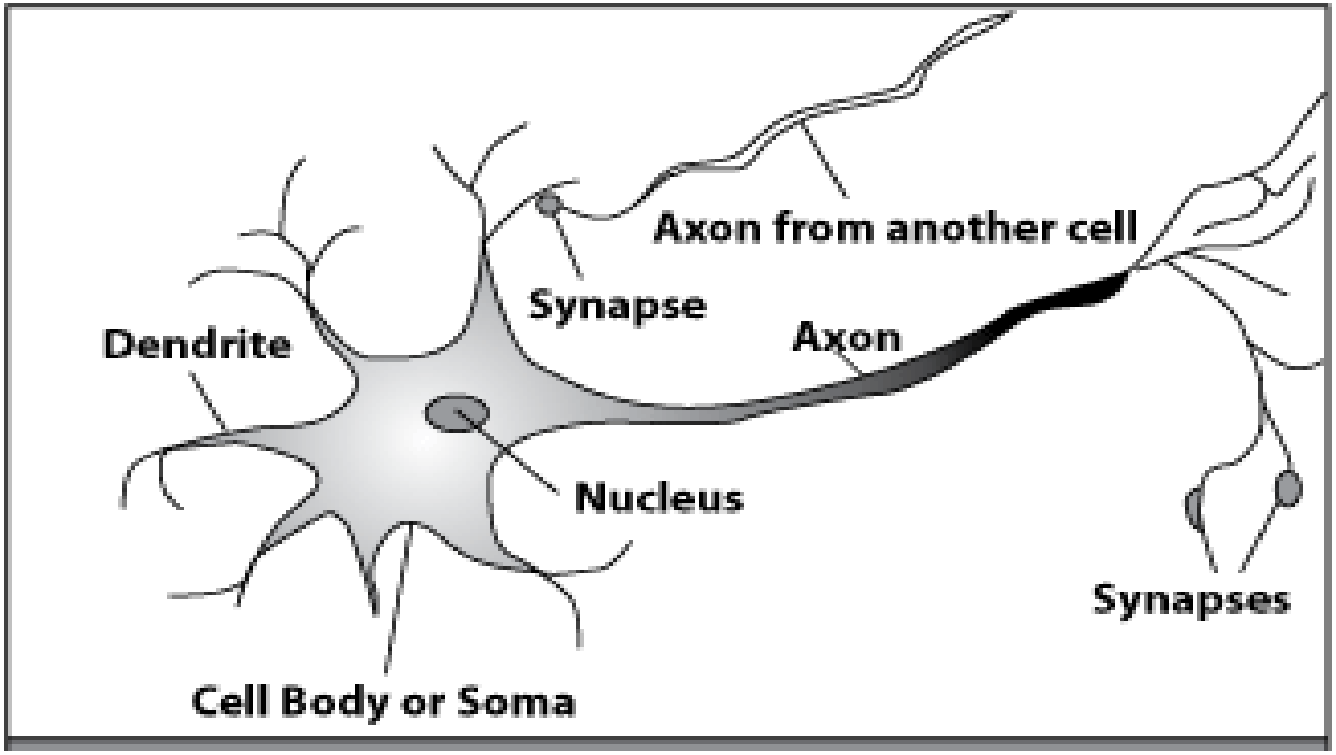
- A processing element



Once input exceeds a critical level, the neuron discharges a spike - an electrical pulse that travels from the body, down the axon, to the next neuron(s)

How do our brains work?

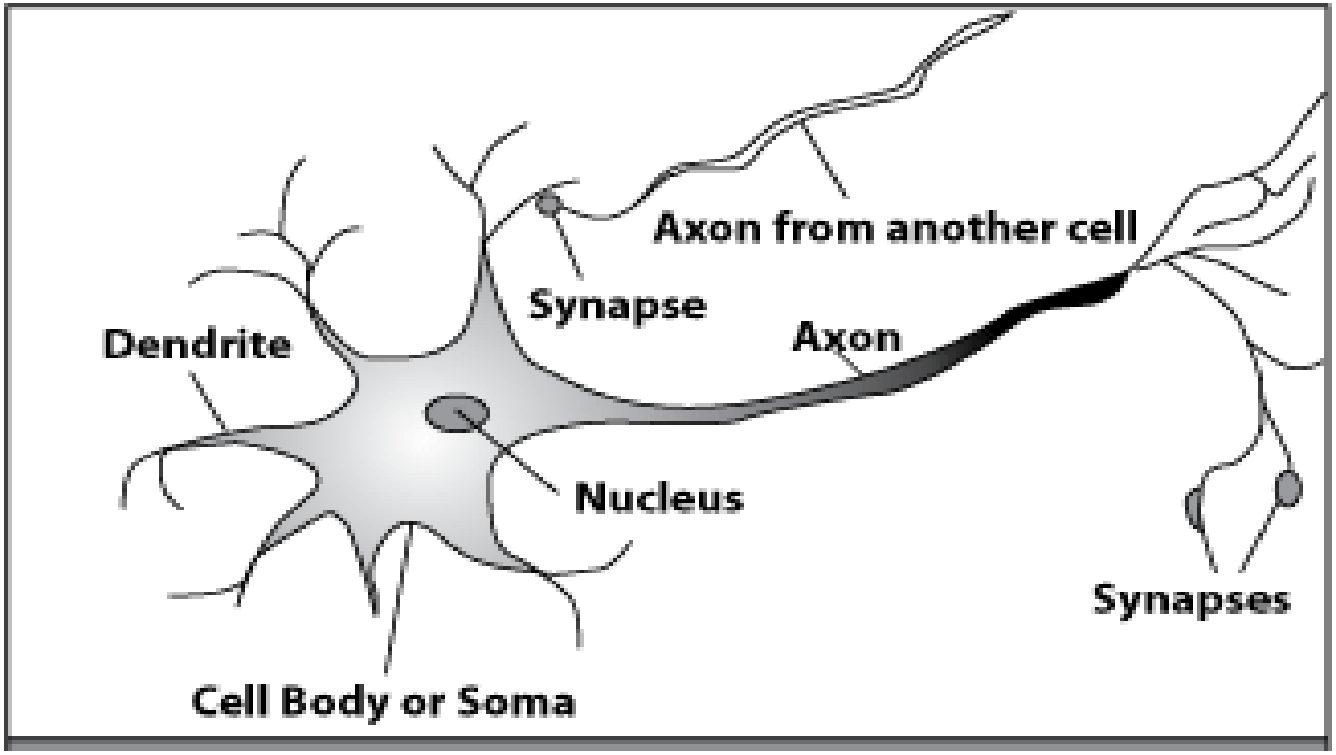
- A processing element



The axon endings almost touch the dendrites or cell body of the next neuron.

How do our brains work?

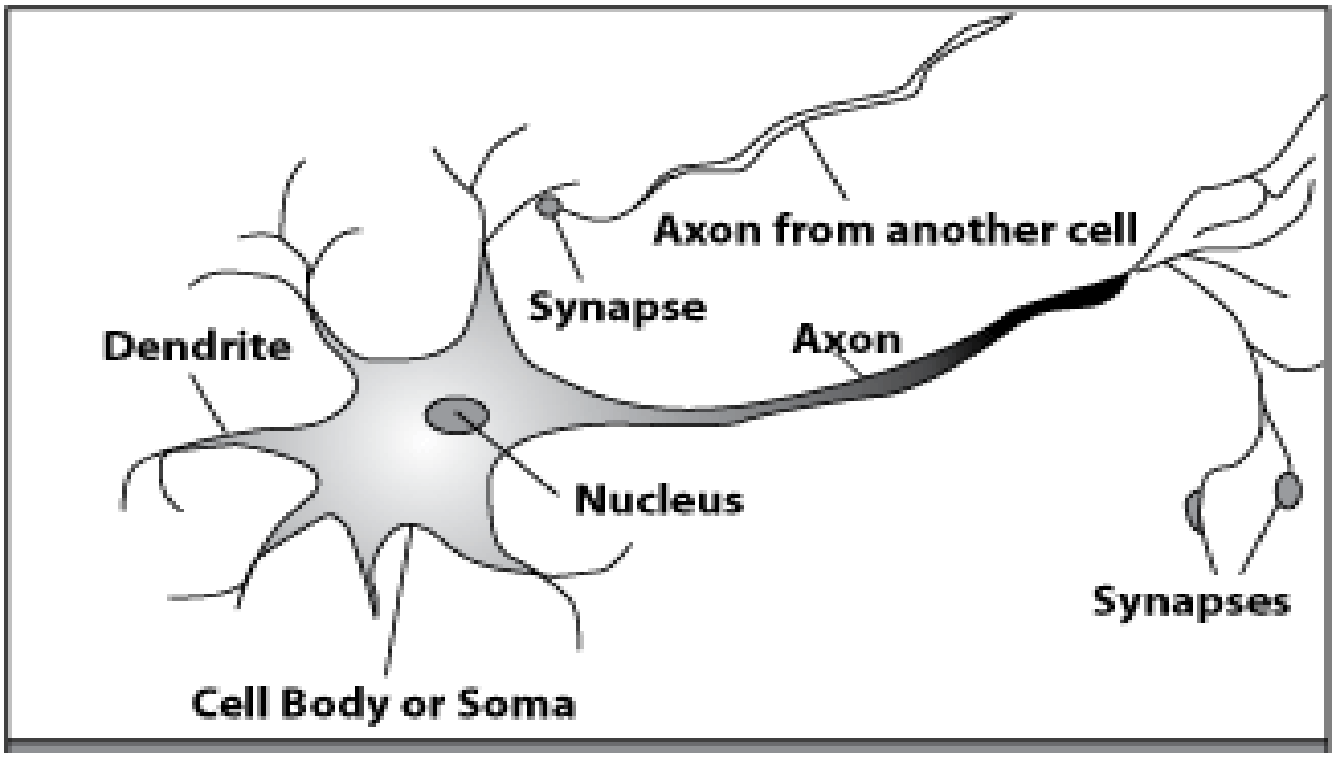
- A processing element



Transmission of an electrical signal from one neuron to the next is effected by neurotransmitters.

How do our brains work?

- A processing element

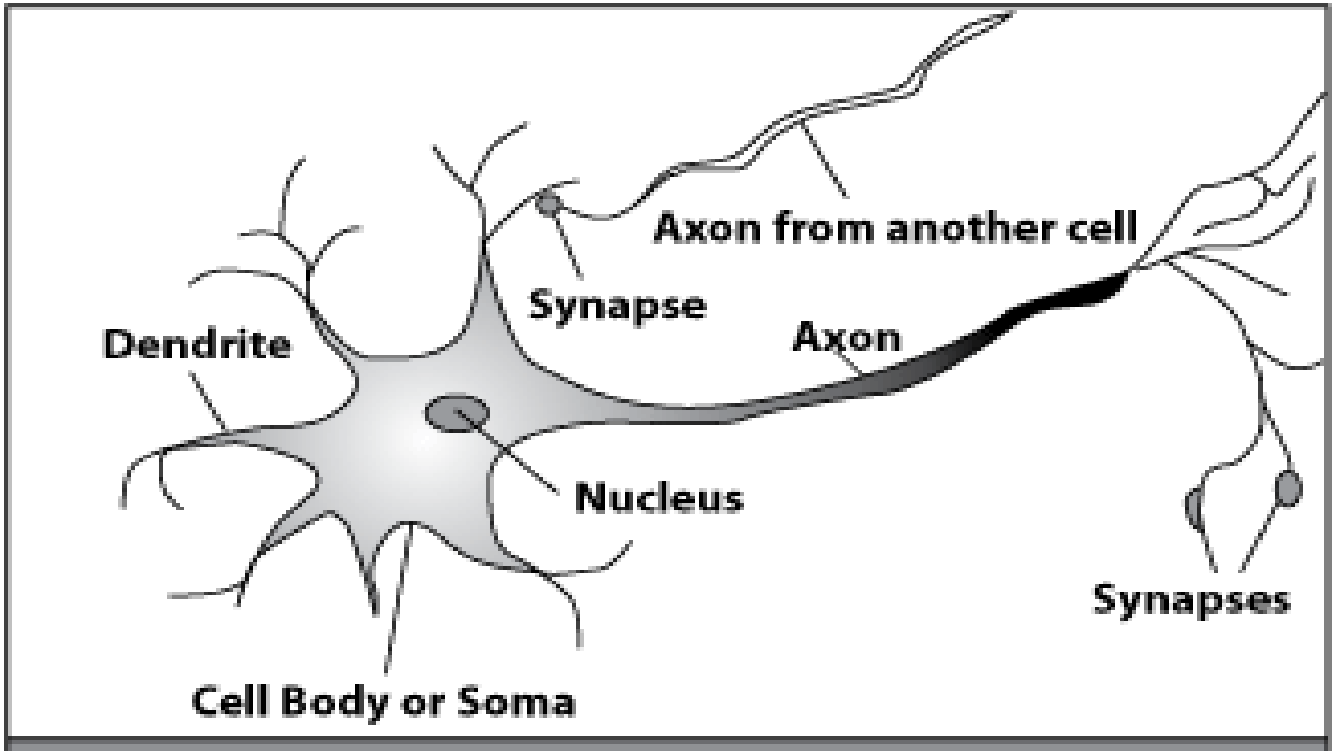


Neurotransmitters are chemicals which are released from the first neuron and which bind to the

Second.

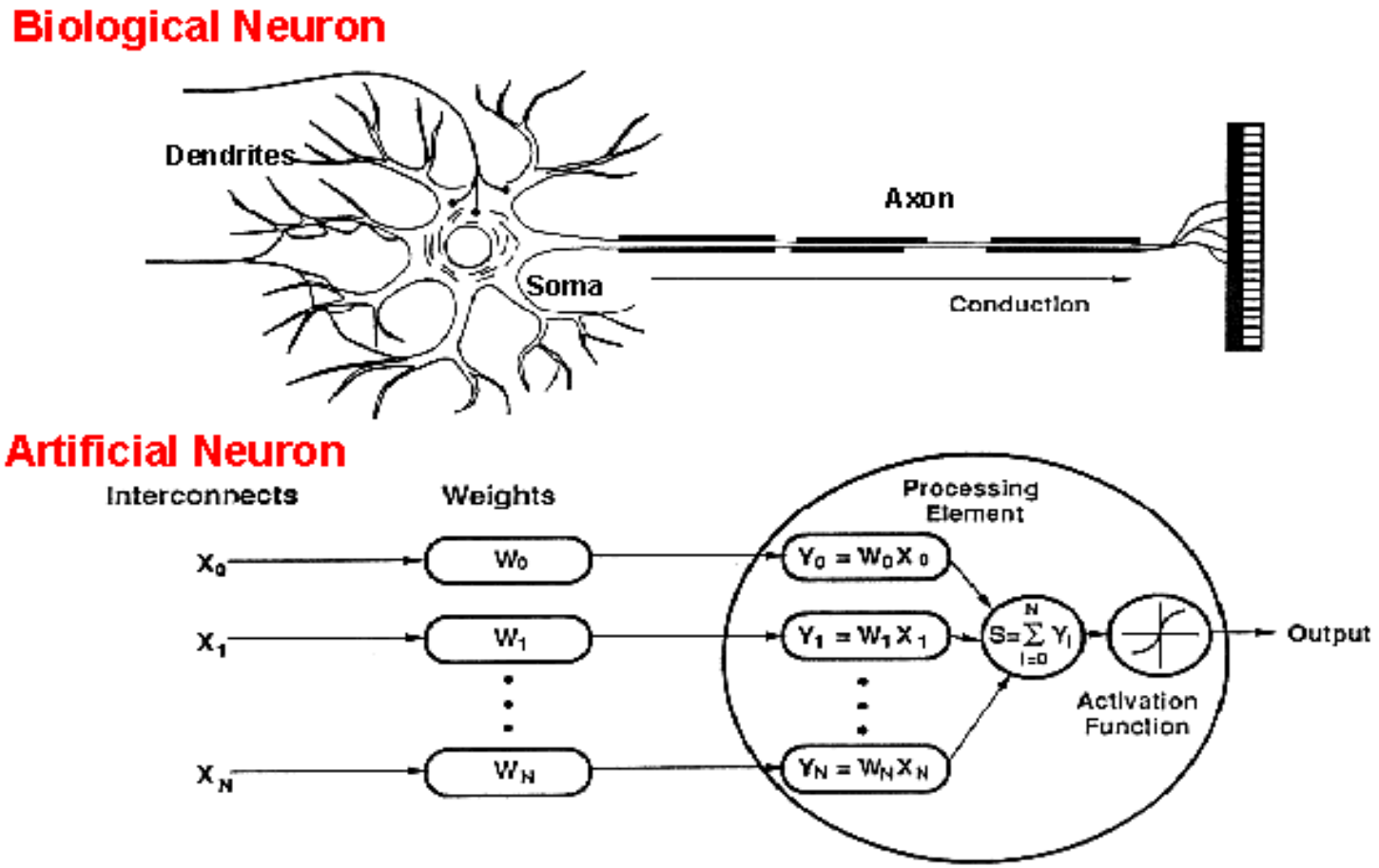
How do our brains work?

- A processing element



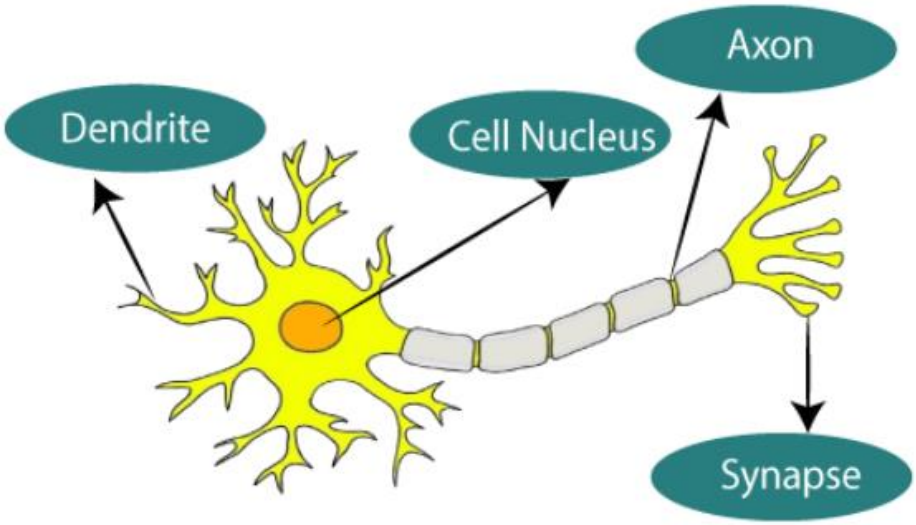
This link is called a synapse. The strength of the signal that reaches the next neuron depends on factors such as the amount of neurotransmitter available.

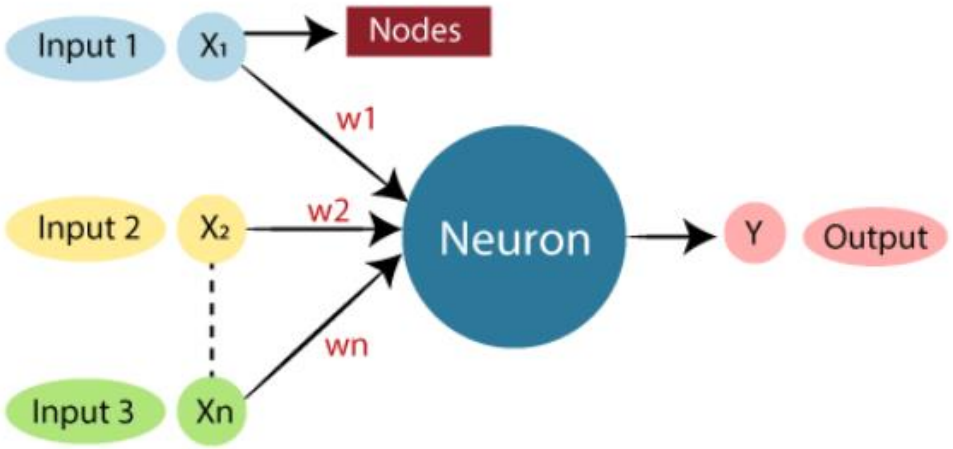
How do ANNs work?



An artificial neuron is an imitation of a human neuron

Relation between biological Neural Network and Artificial neural Network

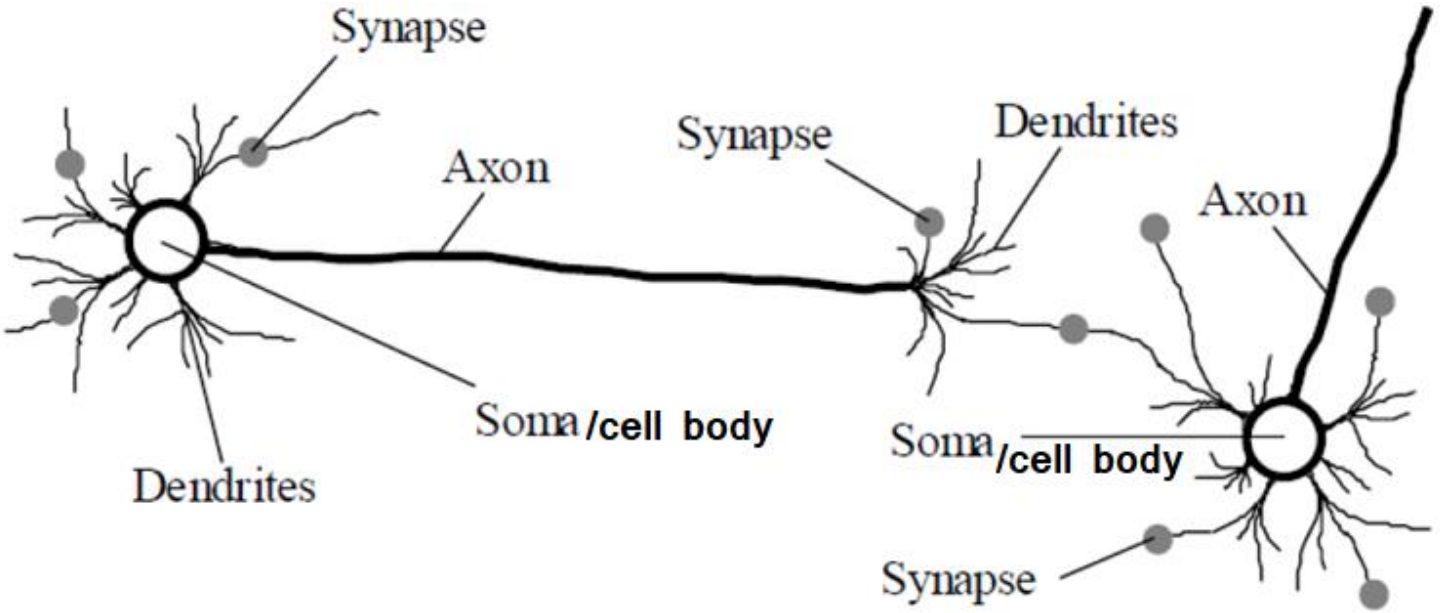




Biological Neural Network.

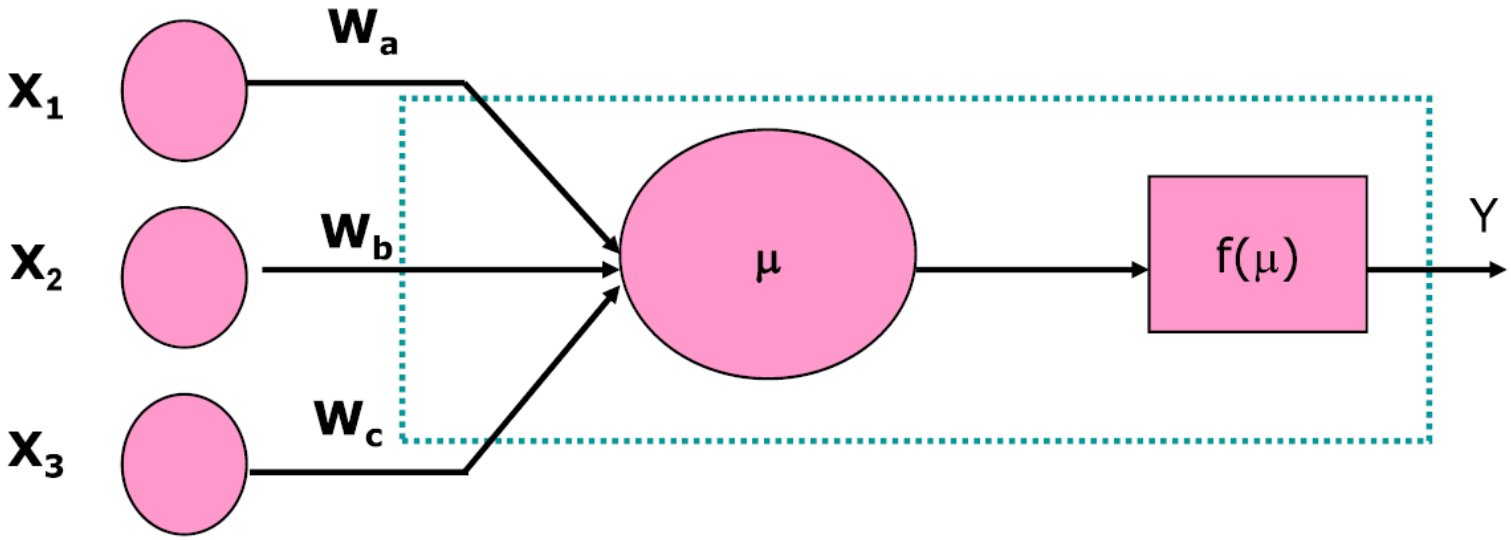
Artificial Neural Network

Biological Neuron



Biological Neural Network	*Artificial Neural Network*
Soma/Cell body	Neuron
Dendrite	Input
Axon	Output
Synapse	Weight

Model of Artificial Neuron



Input units

Connection weights

Summing function

computation

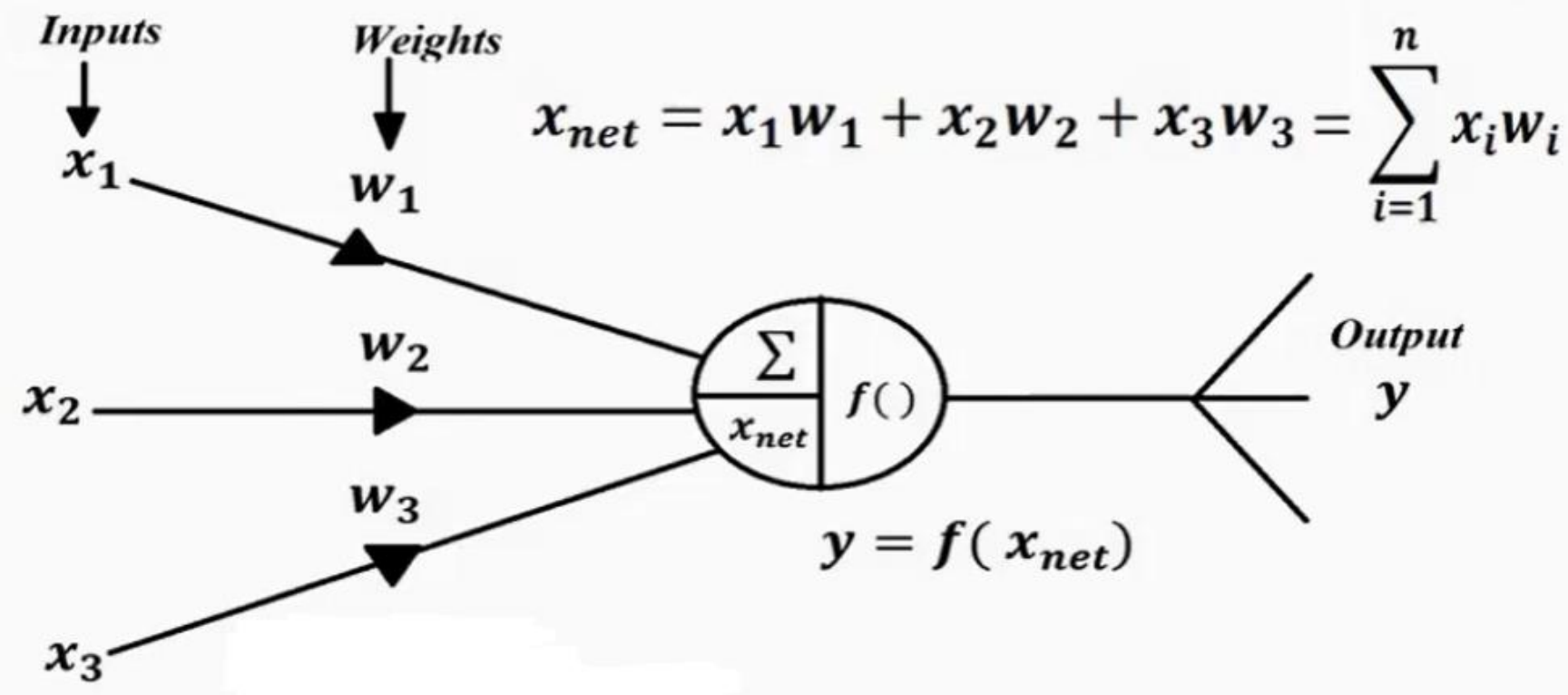
(dendrite)

(synapse)

(soma)

(axon)

Artificial Neuron



Artificial Neuron

- A neural net consists of a large number of simple processing elements called neurons, units, cells or nodes.
- Each neuron is connected to other neurons by means of directed communication links, each with associated weight.
- The weight represent information being used by the net to solve a problem.

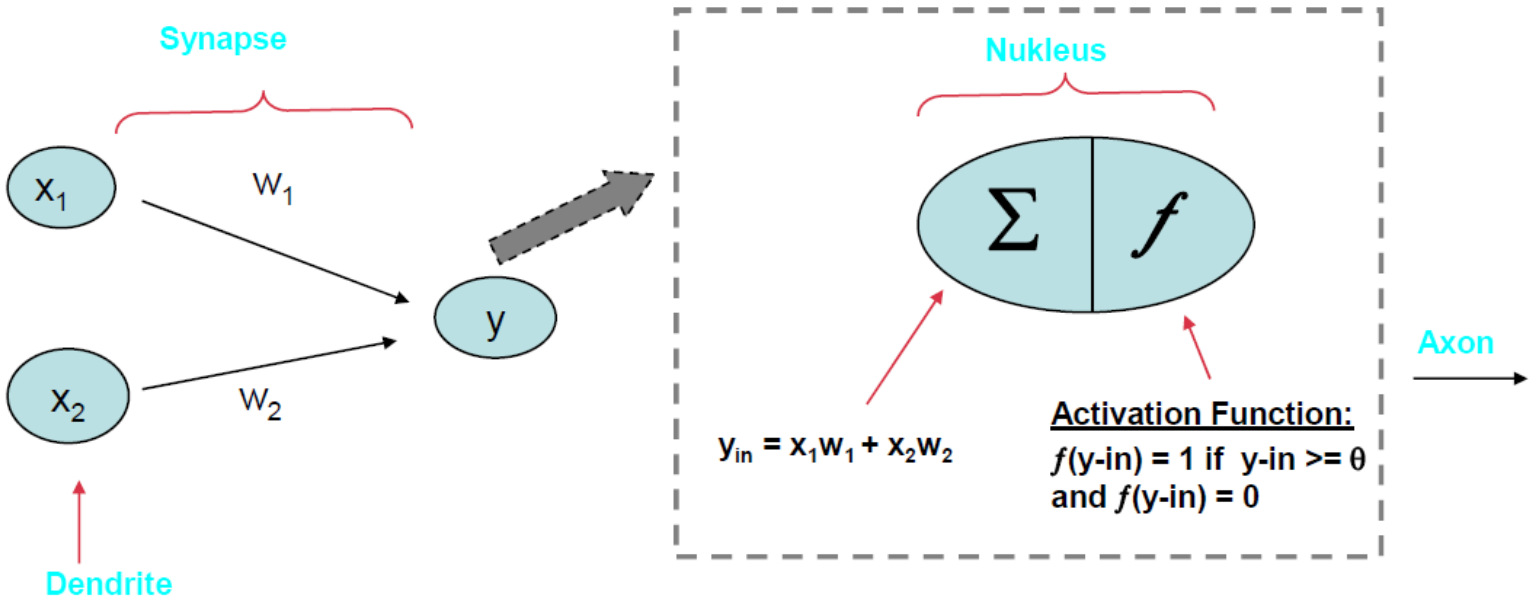
Artificial Neuron

- Each neuron has an internal state, called its activation or activity level, which is a function of the inputs it has received.
- Typically, a neuron sends its activation as a signal to several other neurons.
- It is important to note that a neuron can send only one signal at a time, although that signal is broadcast to several other neurons.

Artificial Neuron

- Neural networks are configured for a specific application, such as pattern recognition or data classification, through a learning process
- In a biological system, learning involves adjustments to the synaptic connections between neurons

Artificial Neural Network



- A neuron receives input, determines the strength or the weight of the input, calculates the total weighted input, and compares the total weighted with a value (threshold)
- The value is in the range of 0 and 1
- If the total weighted input greater than or equal the threshold value, the neuron will produce the output, and if the total weighted input less than the threshold value, no output will be produced

History

- The year 1943 is often considered the initial year in the development of artificial neural systems. McCulloch and Pitts (1943) outlined the first formal model of an elementary computing neuron. The model included all necessary elements to perform logic operations, and thus it could function as an arithmetic logic computing element.
- The implementation of its compact electronic model, however, was not technologically feasible during the era of bulky vacuum tubes.
- The formal neuron model was not widely adopted for the vacuum tube computing hardware description, and the model never became technically significant.
- However, the McCulloch and Pitts neuron model laid the groundwork for future developments.

History

- Influential researchers of that time suggested that research in design of brain-like processing might be interesting.
- To quote John von Neumann's (1958) observations on the "brain language":
- We have now accumulated sufficient evidence to see that whatever language the central nervous system is using, it is characterized by less logical and arithmetical depth than what we are normally used to.
- The following is an obvious example of this: the retina of the human eye performs a considerable reorganization of the visual image as perceived by the eye.
- Now this reorganization is effected on the retina, or to be more precise, at the point of entry of the optic nerve by means of the successive synapses only, i.e. in terms of three consecutive logical steps.

History

- Donald Hebb (1949) first proposed a learning scheme for updating neuron's connections that we now refer to as the ***Hebbian learning rule.***
- He stated that the information can be stored in connections, and postulated the learning technique that had a profound impact on future developments in this field.
- Hebb's learning rule made primary contributions to neural networks theory.

History

- During the 1950s, the first neurocomputers were built and tested (Minsky 1954).
- They adapted connections automatically.
- During this stage, the neuron-like element called a ***perceptron*** was invented by Frank Rosenblatt in 1958.
- It was a trainable machine capable of learning to classify certain patterns by modifying connections to the threshold elements (Rosenblatt 1958).
- The idea caught the imagination of engineers and scientists and laid the groundwork for the basic machine learning algorithms that we still use today.

History

- In the early 1960s a device called ***ADALINE*** (for ADAptive LINEar combiner) was introduced, and a new, powerful learning rule called the ***Widrow-Hoff learning rule*** was developed by Bernard Widrow and Marcian Hoff (1960, 1962).
- The rule minimized the summed square error during training involving pattern classification.
- Early applications of ADALINE and its extensions to MADALINE (for Many ADALINES) include pattern recognition, weather forecasting, and adaptive controls.
- The monograph on learning machines by Nils Nilsson (1965) clearly summarized many of the developments of that time.

History

- Despite the successes and enthusiasm of the early and mid-1960s, the existing machine learning theorems of that time were too weak to support more complex computational problems.
- Although the bottlenecks were exactly identified in Nilsson's work and the neural network architectures called ***layered networks*** were also known, no efficient learning schemes existed at that time that would circumvent the formidable obstacles.
- Neural network research entered into the stagnation phase.
- Another reason that contributed to this research slowdown at that time was the relatively modest computational resources available then.

History

- The final episode of this era was the publication of a book by Marvin Minsky and Seymour Papert (1969) that gave more doubt as to the layered learning networks' potential.
- The stated limitations of the perceptron-class networks were made public; however, the challenge was not answered until the mid-1980s.
- The discovery of successful extensions of neural network knowledge had to wait until 1986

History

- During the period from 1965 to 1984, further pioneering work was accomplished by a handful of researchers.
- The study of learning in networks of threshold elements and of the mathematical theory of neural networks was pursued by Sun-Ichi Amari (1972,1977).
- Also in Japan, Kunihiko Fukushima developed a class of neural network architectures known as ***neocognitrons*** (Fukushima and Miyaka 1980).
- The neocognitron is a model for visual pattern recognition and is concerned with biological plausibility.

✓ The network emulates the retinal images and processes them using two-dimensional layer of neurons.

History

- Associative memory research has been pursued by, among others, Tuevo Kohonen in Finland (1977, 1982, 1984, 1988) and James A. Anderson (Anderson et al. 1977).
- Unsupervised learning networks were developed for feature mapping into regular arrays of neurons (Kohonen 1982).
- During the period from 1982 until 1986, several seminal publications were published that significantly furthered the potential of neural networks.
- The era of renaissance started with John Hopfield (1982, 1984) introducing a recurrent neural network architecture for associative memories. His papers formulated computational properties of a fully connected network of units

History

- Another revitalization of the field came from the publication in 1986 of two volumes on parallel distributed processing, edited by James McClelland and David Rumelhart (1986).
- The new learning rules and other concepts introduced in this work have removed one of the most essential network training barriers that grounded the mainstream efforts of the mid-1960s.
- The publication by Mc-Clelland and Rumelhart opened a new era for the once-underestimated computing potential of layered networks

History

- Although the mathematical framework for the new training scheme of layered networks was discovered in 1974 by Paul Werbos, it went largely unnoticed at that time (Werbos 1974).
- According to the most recent statement (Dreyfus 1990), the first authors of the optimization approach for multilayer feedforward systems were Bryson (Bryson and Ho 1969) and Kelley (Kelley 1969) who obtained a gradient solution for multistage network training.

History

- Beginning in 1986-87, many new neural networks research programs were initiated.
- The intensity of research in the neurocomputing discipline can be measured by a quickly growing number of conferences and journals devoted to the field.
- In addition to many edited volumes that contain collections of papers, several books have already appeared.
- The list of applications that can be solved by neural networks has expanded from small test-size examples to large practical tasks.
- Very-large-scale integrated neural network chips have been fabricated science.
- Although neurocomputing has had an interesting history, the field is still in its early stages of development.

History: Summary

- 1943 McCulloch-Pitts neurons
- 1949 Hebb's law
- 1958 Perceptron (Rosenblatt)
- 1960 Adaline, better learning rule (Widrow, Huff)
- 1969 Limitations (Minsky, Papert)
- 1972 Kohonen nets, associative memory

History: Summary

- 1977 Brain State in a Box (Anderson)
- 1982 Hopfield net, constraint satisfaction
- 1985 ART (Carpenter, Grossfield)
- 1986 Backpropagation (Rumelhart, Hinton, McClelland)
- 1988 Neocognitron, character recognition (Fukushima)

Application of artificial neural networks in medicine

- The use of neural networks in medicine, normally is linked to disease diagnostics systems.
- However, neural networks are not only able to recognize examples, but maintain very important information.
- For this reason, one of the main areas of application of neural networks is the interpretation of medical data

Application of artificial neural networks in medicine

- Neural networks are nonlinear systems, which make it possible to classify the data better than linear methods.
- A distinctive feature of neural networks is that they are not programmed - do not use any rule for diagnosis, but are trained to do so in the examples.
- Regarding this, neural networks aren’t similar to expert systems, the development of which in the 70s was held after a temporary "victory" of artificial intelligence approach against memory modeling, recognition and synthesis models, which was based on the study of neural organization of human brain.

Application of artificial neural networks in medicine

- Diagnostics of diseases is broad and challenging area.
- Its task is to detect a disease that patient with the symptoms have.
- This process is very complicated, because not all disease's symptoms are specific to only one disease and often the symptoms are overlapping.
- Errors caused by human factor are not rare in this process.
- To eliminate human error, in modern medicine, different technologies are used nowadays. Some of them are clinical decision support systems.
- These are interactive computer programs that assist the doctors at diagnostic the patient's diseases. Using information about a patient's condition in the mathematical model the probable diagnosis can be determined.
- These mathematical models are based on statistical distributions, regression models and artificial intelligence

Application of artificial neural networks in medicine

- Application of neural networks for diagnosing diabetes disease in its early stages.
- Nowadays, diabetes is considered one of the most prevalent diseases in the world.
- According to the World Health Organization, around 30 million people of various ages and breeds suffer from various forms of diabetes1 .
- Diabetes is not a consequence of any particular organ pathology, but in general displayed by the metabolic disorder.
- His symptoms occur in organs or organ systems, which are more vulnerable to this process.
- Clinical signs of diabetes depend on the type of disease, gender, age, level of insulin, arterial pressure and many other factors.

Application of artificial neural networks in medicine

- Neural network technologies are designed to solve many difficult tasks, starting from formulation, among which many medical problems.
- This is related to the fact that to the researchers are often given a large number of factual materials, for which there is no mathematical model.
- Artificial neural networks models have shown good results for diagnose nerve disorders, Parkinson's disease and Huntington diagnosis.
- Multilayer perceptron models are used for predicting the risk of occurrence of osteoporosis.
- For Hepatitis B diagnostic is used generalized mathematical regression.
- One of the most useful tools for solving the tasks above are artificial neural networks - a powerful simulation method of phenomena and processes at the same time very flexible
- Modern neural networks are a combination of devices and programs designed for models and specialized equipment for solving a wide range of diagnostics tasks by applying a set of algorithms theory of object recognition
- The distinguishing feature of neural networks is the ability to be trained in a specific field.
- Applied to medical topics, experimental data represent a set of input parameters of the object (in our case health parameters).

Application of artificial neural networks in medicine

- Training network neural is an interactive process, at the entrance of which net finds secret nonlinear relationships between input parameters and final diagnosis, also the optimal combination of weight coefficients of neurons, joining the neighbour layers for which the error of definite sample class goes to minimum
- As one of the advantages of neural networks we should consider their relative simplicity, non-linearity, working with incomplete information, the ability to train on concrete examples.
- During the training process, as initial input we will give parameters together with the diagnosis, characterizing these parameters. To train a neural network we must have a sufficient number of examples to adapt the system to a certain level of credibility

Application of artificial neural networks in medicine

- Trained ANNs approach the functionality of small biological neural cluster in a very fundamental manner.
- They are the digitized model of biological brain and can detect complex nonlinear relationships between dependent as well as independent variables in a data where human brain may fail to detect.
- Nowadays, ANNs are widely used for medical applications in various disciplines of medicine especially in cardiology.
- ANNs have been extensively applied in diagnosis, electronic signal analysis, medical image analysis and radiology.

Application of artificial neural networks in medicine

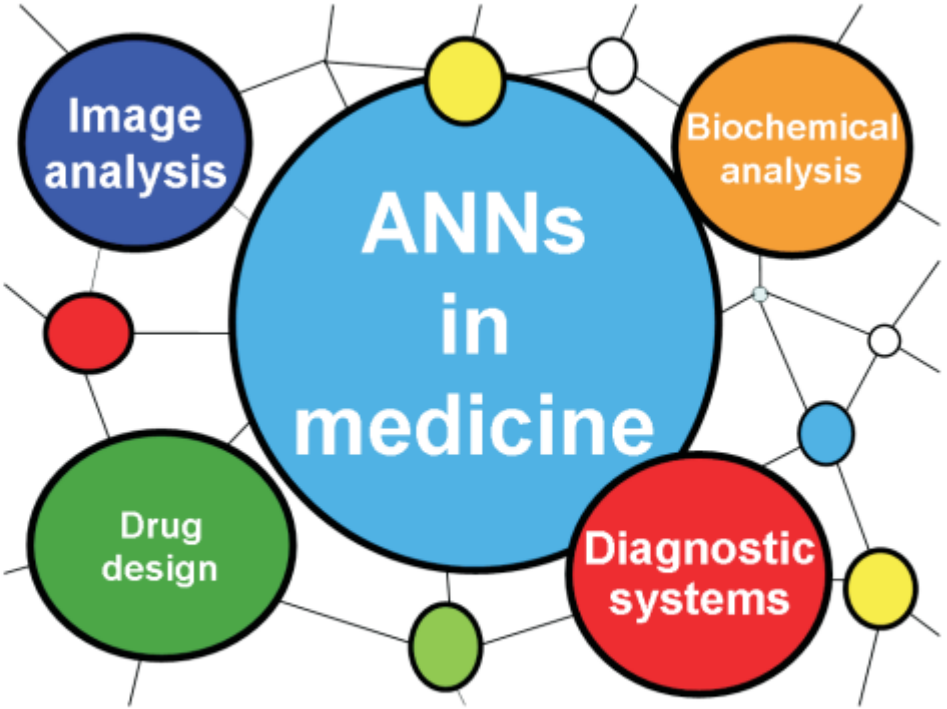


Fig. 1. **Overview of the main applications of artificial neural networks in medicine.**

Application of artificial neural networks in medicine

- Due to the substantial plasticity of input data, ANNs have proven useful in the analysis of blood and urine samples of diabetic patients, diagnosis of tuberculosis, leukemia classification, analysis of complicated effusion samples, and image analysis of radiographs or even living tissue.

Neural network for diseases classification

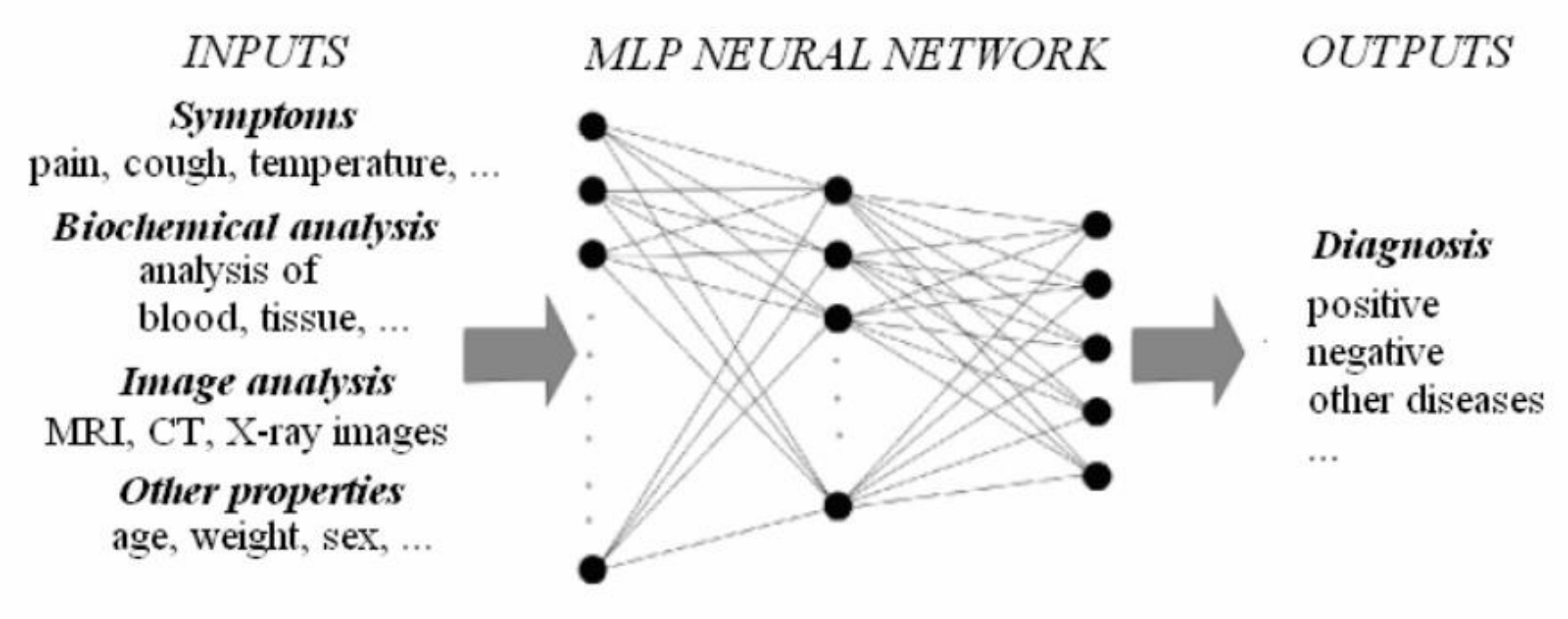


Figure Medical diagnostic using MLP neural network

Neural Network for diseases classification

Patient code	MEDICAL DATA			DIAGNOSIS
1	$data_{1,1}$ …	$data_{1,i}$ …	$data_{1,m}$	POSITIVE
2	$data_{2,1}$ …	$data_{2,i}$ …	$data_{2,m}$	POSITIVE
3	$data_{3,1}$ …	$data_{3,i}$ …	$data_{3,m}$	POSITIVE
…	……	……	……	……
k	$data_{k,1}$ …	$data_{k,i}$ …	$data_{k,m}$	NEGATIVE
***k*+1**	$data_{k+1,1}$ …	$data_{k+1,i}$ …	$data_{k+1,m}$	NEGATIVE
…	……	……	……	……
n	$data_{n,1}$ …	$data_{n,i}$ …	$data_{n,m}$	NEGATIVE

Fig. **Example of training database structure.** Each row refers to a different patient labeled with a numerical code. The element $data_{k,i}$ refers to the i-th medical data (symptom, laboratory data, etc.) of the k-th patient.

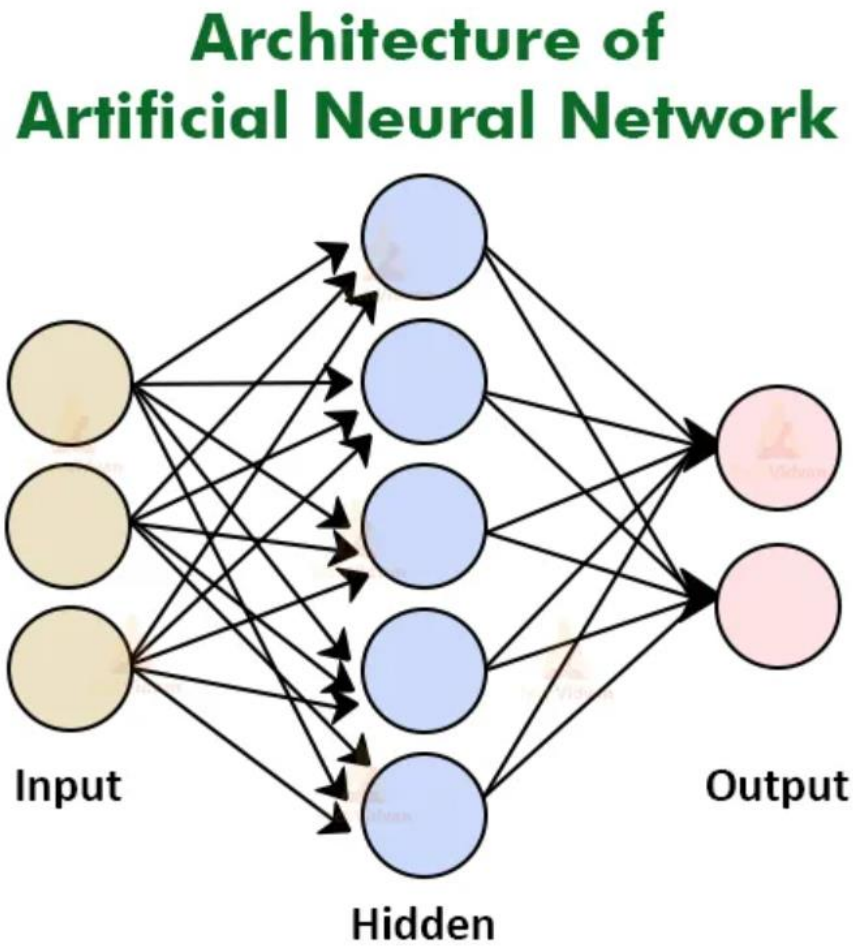
Neural Network Medical Diagnosis (Breast Cancer)

- Artificial Neural Network helps to diagnosing breast cancer.
- Breast cancer is a widespread type of cancer (for example in the UK, it's the most common cancer).
- As with any disease, it's vital to detect it as soon as possible to achieve successful treatment.
- The diagnosis of breast cancer is performed by a pathologist.
- To detect cancer, a pathologist would conduct a laboratory procedure or biopsy.
- First, a pathologist collects samples of tissues from the breast region.
- Then, he analyses the images under a microscope and classifies them as cancerous or noncancerous.
- But images can be classified automatically. Some of the recent computer-aided diagnosis methods rely on pattern recognition and artificial neural networks.
- Their approach is based on the determination of nuclei regions on the images and then using these regions into the algorithm that performs classification, or classifier.

Biological Neuron Vs Artificial Neuron

- Some major differences between them are biological neural network does parallel processing whereas the Artificial neural network does series processing also in the former one processing is slower (in millisecond) while in the latter one processing is faster (in a nanosecond).

Architecture of ANN



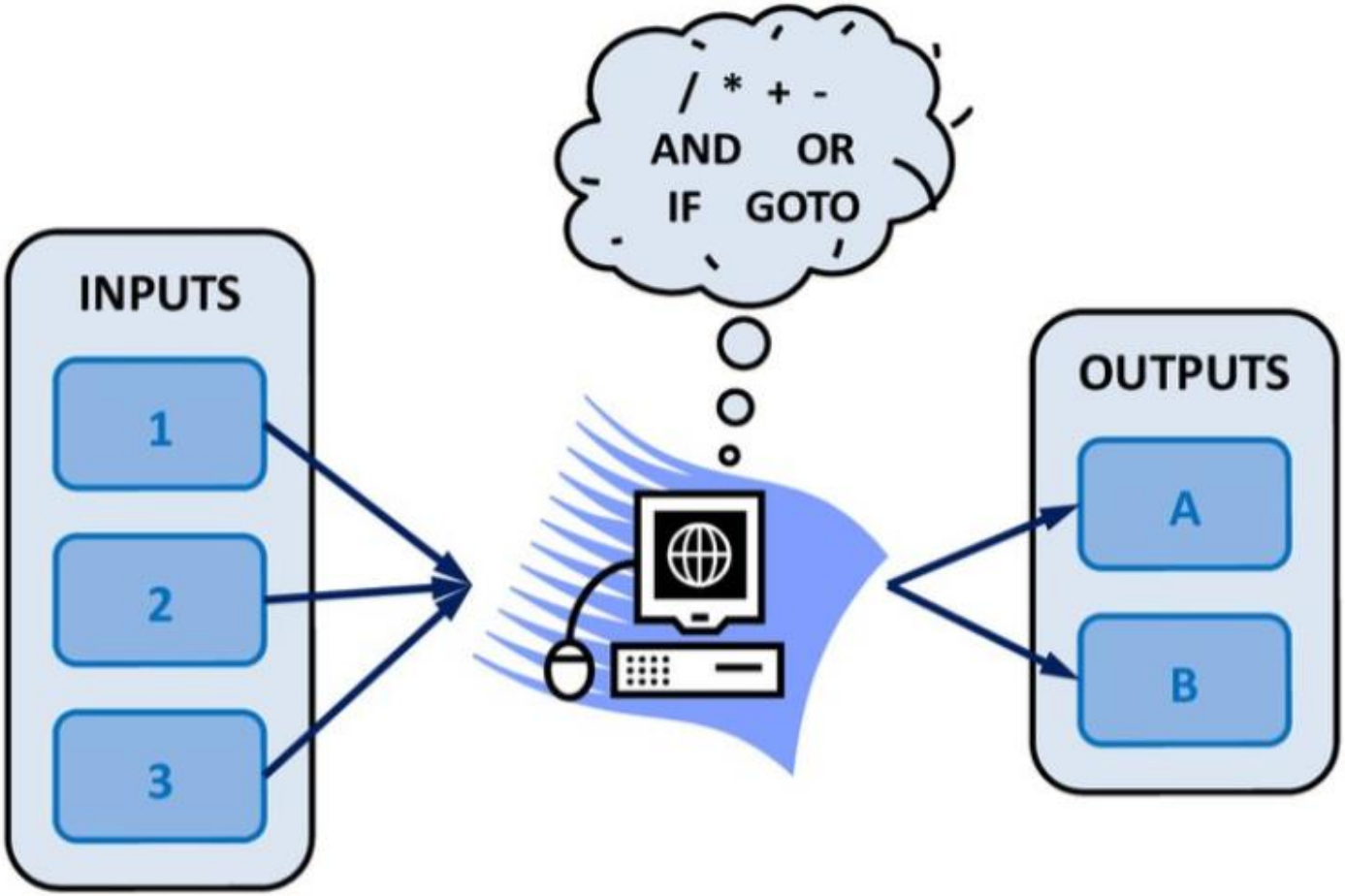
Architecture of ANN

- A neural network consists of three layers.
- The first layer is the input layer.
  - ✓ It contains the input neurons that send information to the hidden layer.
- The hidden layer performs the computations on input data and transfers the output to the output layer.
  - ✓ It includes weight, activation function, cost function.
  - ✓ The connection between neurons is known as weight, which is the numerical values.
  - ✓ The weight between neurons determines the learning ability of the neural network.
  - ✓ During the learning of artificial neural networks, weight between the neuron changes.
  - ✓ Initial weights are set randomly.
- The third layer is the output layer.

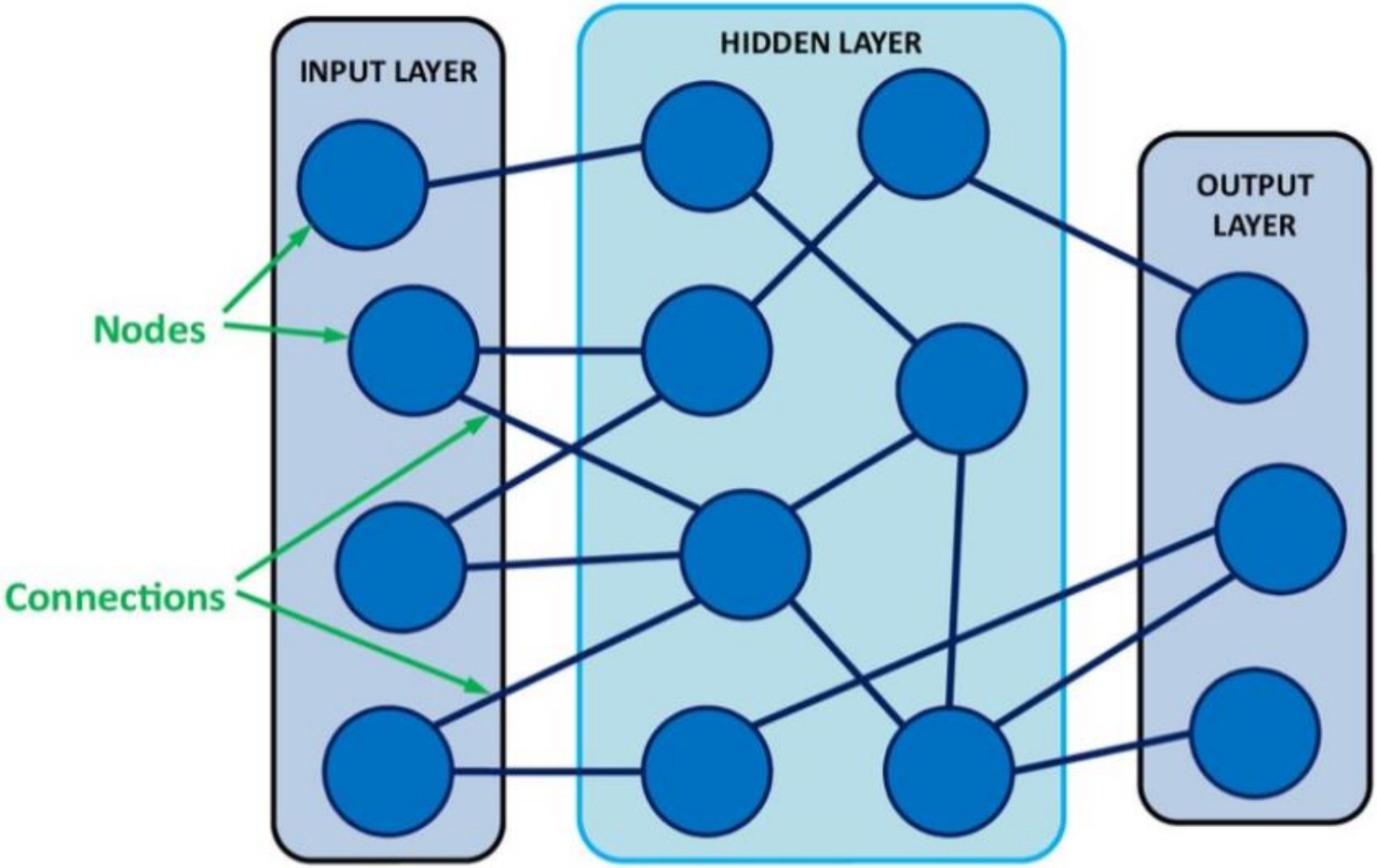
Architecture of ANN

- To standardize the output from the neuron, the “activation function” is used.
- Activation functions are the mathematical equations that calculate the output of the neural network. Standardization refers to the transformation of data to have mean 0 and standard deviation 1.
- Each neuron has its activation function. It is difficult to understand without mathematical reasoning. It also helps to normalize the output in a range between 0 to 1 or -1 to 1.
- An activation function is also known as the transfer function.

Conventional Computer



Neural Network as a Computer Model



Neural Network Vs Conventional Computer Model

COMPUTERS	ANN
• Algorithmic approach	• Learning approach
• They are necessarily programmed	• Not programmed for specific tasks
• Work on predefined set of instructions	• Used in decision making
• Operations are predictable	• Operation is unpredictable

Neural Network Vs Conventional Computer Model

"Standard" Computers	Neural Networks
• one CPU	highly parallel processing
•fast processing units	slow processing units
•reliable units	unreliable units
•static infrastructure	dynamic infrastructure

Advantages of neural networks over conventional computers

- Artificial neural networks may function through images, images, and concepts.
- While conventional computers function logically with a set of rules and calculations.
- Neural networks are inspired by the structure of a biological neural network in the human brain.
- Let’s discuss some important points on the advantages of neural networks over conventional computers.

Advantages of neural networks over conventional computers

- Neural networks have the ability to learn by themselves and produced the output that is not limited to the input provided to them.
- The input is stored in its own networks instead of the database.
- Hence, data loss does not change the way it operates.
- The neural network performs multiple tasks without affecting system performance.
- If any information is missing, the network can detect the fault and still produce the output.