Deep Learning and Neural Network

29 November 2023 20:30

Deep Learning

Machine Learning

Deep learning is a method in artificial intelligence (AI) that teaches computers to process data in a way that is inspired by the human brain. Deep learning models can recognize complex patterns in pictures, text, sounds, and other data to produce accurate insights and predictions.

Artificial intelligence (AI) attempts to train computers to think and learn as humans do. Deep learning technology drives many AI applications used in everyday products, such as the following:

- Digital assistants
- Voice-activated television remotes
- Fraud detection
- Automatic facial recognition

NLP

LLM

Deep learning has several use cases in automotive, aerospace, manufacturing, electronics, medical research, and other fields. These are some examples of deep learning:

- Self-driving cars use deep learning models to automatically detect road signs and pedestrians.
- Defense systems use deep learning to automatically flag areas of interest in satellite images.
- Medical image analysis uses deep learning to automatically detect cancer cells for medical diagnosis. →
- Factories use deep learning applications to automatically detect when people or objects are within an unsafe distance of machines.

Osteoarthritis

>45-50

Deep learning algorithms are neural networks that are modeled after the human brain. For example, a human brain contains millions of interconnected neurons that work together to learn and process information. Similarly, deep learning neural networks, or artificial neural networks, are made of many layers of artificial neurons that work together inside the computer.

Nucleus →

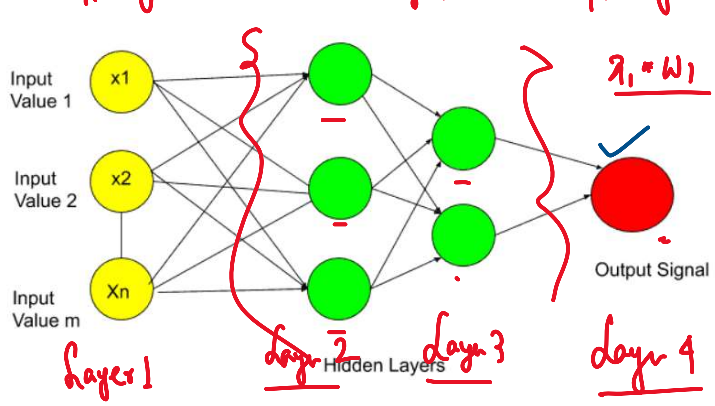
Artificial neurons are software modules called nodes, which use mathematical calculations to process data. Artificial neural networks are deep learning algorithms that use these nodes to solve complex problems.

*: DL Frameworks :-

Tensorflow, Keras, Pytorch

edges & corner of the image

I/P layer Hidden layers O/P layer



Neural Network Architecture

The components of a deep neural network are the following.

ANN – Regression & classification

Input layer

The components of a deep neural network are the following.

ANN – Regression & classification

Input layer

An artificial neural network has several nodes that input data into it. These nodes make up the input layer of the system.

Hidden layer

The input layer processes and passes the data to layers further in the neural network. These hidden layers process information at different levels, adapting their behavior as they receive new information. Deep learning networks have hundreds of hidden layers that they can use to analyze a problem from several different angles.

Output layer

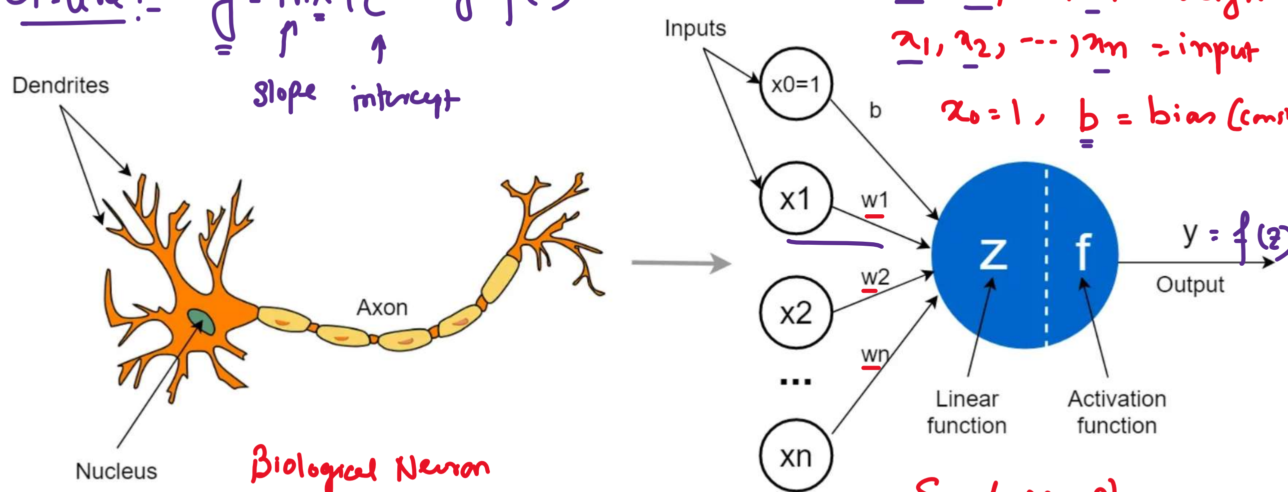
The output layer consists of the nodes that output the data. Deep learning models that output "yes" or "no" answers have only two nodes in the output layer. On the other hand, those that output a wider range of answers have more nodes.

St. line :- $y = mx + c$ (slope m, intercept c) $y = f(x)$

$w_1, w_2, \ldots, w_n$ = weights

$x_1, x_2, \ldots, x_n$ = input

$x_0 = 1$, b = bias (const)



$y = f(z)$

Biological Neuron

Single perceptron

$$z = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n + b$$

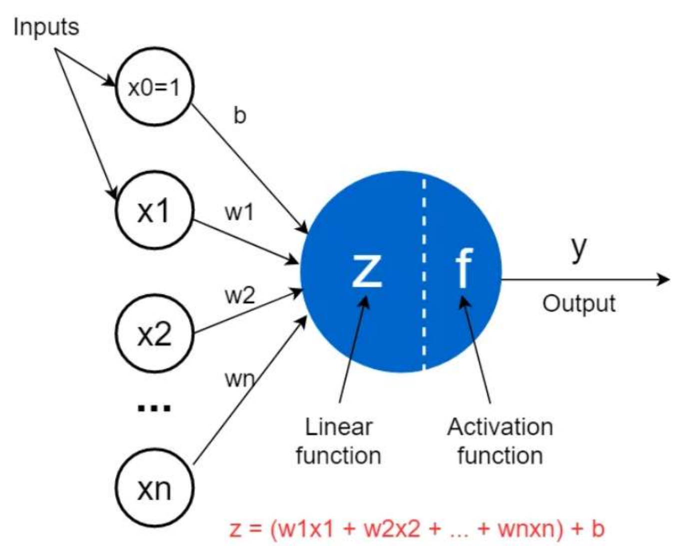
Artificial neurons (also called *Perceptrons*, *Units* or *Nodes*) are the simplest elements or building blocks in a neural network. They are inspired by biological neurons that are found in the human brain.

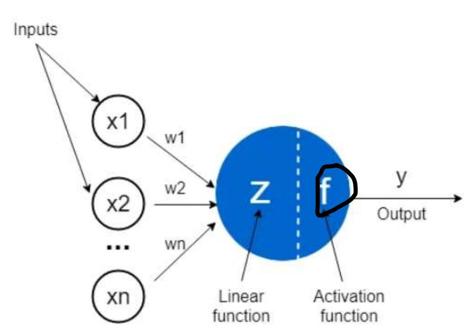
How are perceptrons inspired by biological neurons?

- A biological neuron receives its input signals from other neurons through **dendrites** (small fibers). Likewise, a perceptron receives its data from other perceptrons through **input neurons** that take numbers.
- The connection points between dendrites and biological neurons are called **synapses**. Likewise, the connections between inputs and perceptrons are called **weights**. They measure the importance level of each input.
- In a biological neuron, the ***nucleus*** produces an output signal based on the signals provided by dendrites. Likewise, the ***nucleus*** (colored in blue) in a perceptron performs some calculations based on the input values and produces an output.
- In a biological neuron, the output signal is carried away by the ***axon***. Likewise, the axon in a perceptron is the **output value** which will be the input for the next perceptrons.

The structure of a perceptron

The following image shows a detailed structure of a perceptron. In some contexts, the bias, **b** is denoted by **w0.** The input, **x0** always takes the value 1. So, **b*1 = b**.





Inside a perceptron

A perceptron usually consists of two mathematical functions.

Perceptron's linear function

This is also called the linear component of the perceptron. It is denoted by **z**. Its output is the weighted sum of the inputs plus bias unit and can be calculated as follows.

$$z = (w1.x1 + w2.x2 + \dots + wn.xn) + b$$

$(\omega_i x_i + b)$

Dataset - non-linearity

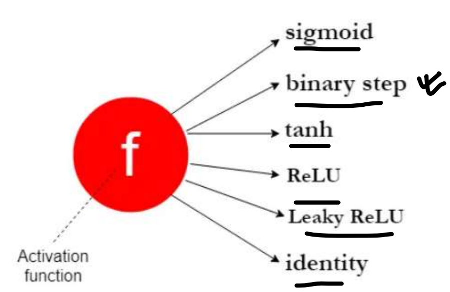
Perceptron’s non-linear (activation) function

This is also called the non-linear component of the perceptron. It is denoted by **f**. It is applied on **z** to get the output **y** based on the type of activation function we use.

$$y = f(z)$$

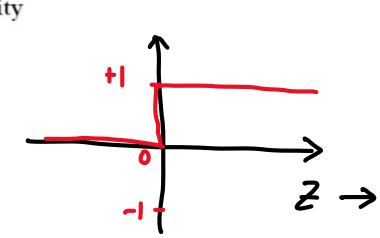
$$\underbrace{y}_{\text{Output}} = \underbrace{f}_{\substack{\text{Activation function}\\ \text{(non-linear)}}}\Big(\underbrace{(w1.x1 + w2.x2 + \dots + wn.xn)}_{\substack{\text{Weighted sum of inputs}\\ \text{(linear combination)}}} + \underbrace{b}_{\text{Bias}}\Big)$$

The function **f** can be a different type of activation function.



What does it mean by “firing a neuron”?

$$\text{binary step}(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z \le 0 \end{cases}$$



Binary step AF

z = (w1.x1 + w2.x2 + ... + wn.xn) + b

z = (weighted sum of inputs) + bias

bias (b) = -2

z > 0

(weighted sum of inputs) + bias > 0

⇒ (weighted sum of inputs) > -bias

When the bias is -2 in our example,

(weighted sum of inputs) > -(-2)

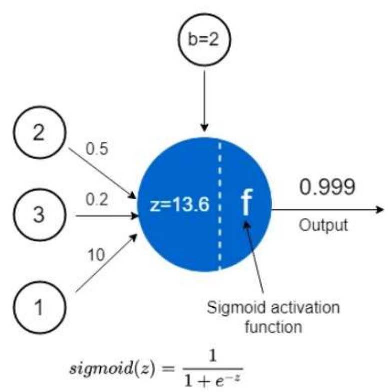
(weighted sum of inputs) > 2

```
x1=2, x2=3 and x3=1
```

```
w1=0.5, w2=0.2 and w3=10
```

```
b=2
```

$$sigmoid(z) = \frac{1}{1+e^{-z}}$$



What is deep learning in the context of machine learning?

Deep learning is a subset of machine learning. Deep learning algorithms emerged in an attempt to make traditional machine learning techniques more efficient. Traditional machine learning methods require significant human effort to train the software. For example, in animal image recognition, you need to do the following:

ML way

- Manually label hundreds of thousands of animal images.
- Make the machine learning algorithms process those images.
- Test those algorithms on a set of unknown images.
- Identify why some results are inaccurate.
- Improve the dataset by labeling new images to improve result accuracy.

epoch = ⇐

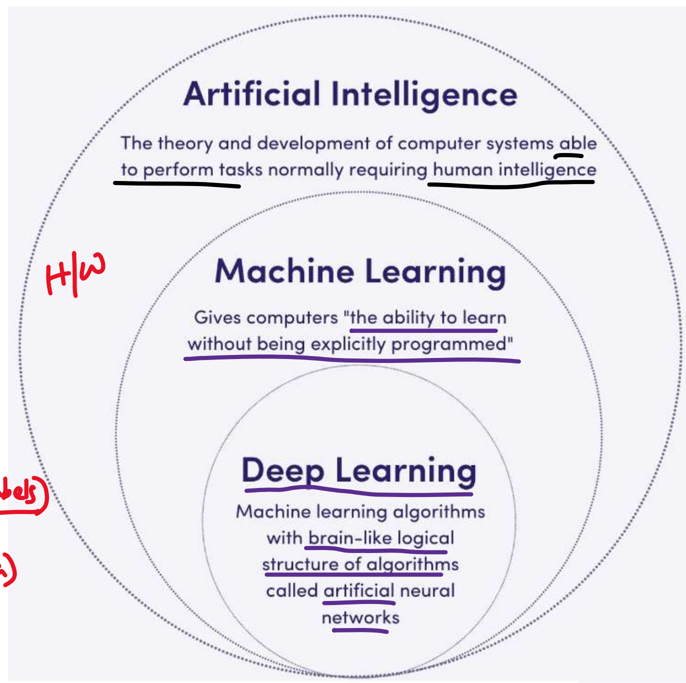
This process is called supervised learning. In supervised learning, result accuracy improves only when you have a broad and sufficiently varied dataset. For instance, the algorithm might accurately identify black cats but not white cats because the training dataset had more images of black cats. In that case, you would need to label more white cat images and train the machine learning models once again.

Biological Neuron
↓ Mathematical Description

Neural Network/
Artificial NN

① Supervised (data, labels)
② Unsupervised (data)

H/W



Data (Seat, Age, KM, ...)
⇓
I/p → Model → Program
O/p →
⇓
(New price)

✓ Prediction

✓ Data < Training / Test

⇐ Model Training

Machine Learning	Deep Learning
Apply statistical algorithms to learn the hidden patterns and relationships in the dataset.	Uses artificial neural network architecture to learn the hidden patterns and relationships in the dataset.
Can work on the smaller amount of dataset	Requires the larger volume of dataset compared to machine learning
Better for the low-label task.	Better for complex task like image processing, natural language processing, etc.
Takes less time to train the model. CPU	Takes more time to train the model. GPU, TPU,
A model is created by relevant features which are manually extracted from images to detect an object in the image.	Relevant features are automatically extracted from images. It is an end-to-end learning process.
Less complex and easy to interpret the result.	More complex, it works like the black box interpretations of the result are not easy.
It can work on the CPU or requires less computing power as compared to deep learning.	It requires a high-performance computer with GPU.

CNN (image)
NLP (text)
RNN (text, time dependent)

CPU

epoch,
batch size

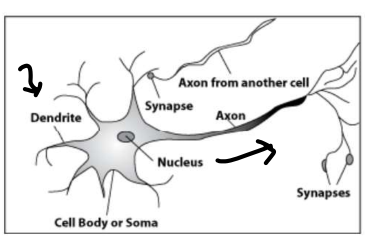
Origin of Neural Network

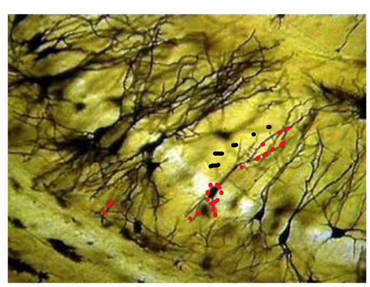
- Human brain has many incredible characteristics such as massive parallelism, distributed representation and computation, learning ability, generalization ability, adaptivity, which seems simple but is really complicated.
- It has been always a dream for computer scientist to create a computer which could solve complex perceptual problems this fast.
- ANN models was an effort to apply the same method as human brain uses to solve perceptual problems.
- Three periods of development for ANN:

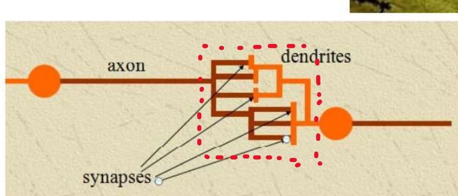
- **1940**: Mcculloch and Pitts: Initial works
- **1960**: Rosenblatt: perceptron convergence theorem
  Minsky and Papert: work showing the limitations of a simple perceptron
- **1980**: Hopfield/Werbos and Rumelhart: Hopfield's energy approach/back-propagation learning algorithm

Biological Neural Network

Single perceptron

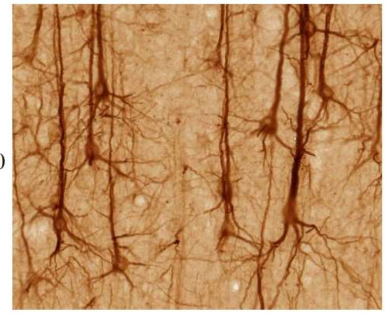






Biological Neural Network

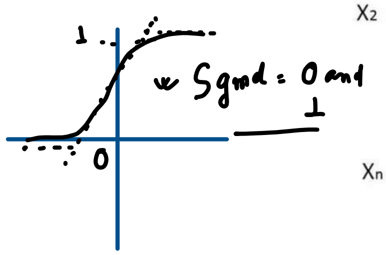
- When a signal reaches a synapse: Certain chemicals called neurotransmitters are released.
- Process of learning: The synapse effectiveness can be adjusted by signal passing through.
- Cerebral cortex :a large flat sheet of neurons about 2 to 3 mm thick and 2200 cm , 10^11 neurons
- Duration of impulses between neurons: milliseconds and the amount of information sent is also small(few bits)
- Critical information are not transmitted directly , but stored in interconnections
- The term Connectionist model initiated from this idea.



COMPUTATIONAL

ANN Overview: [illegible] MODEL FOR ARTIFICIAL NEURON

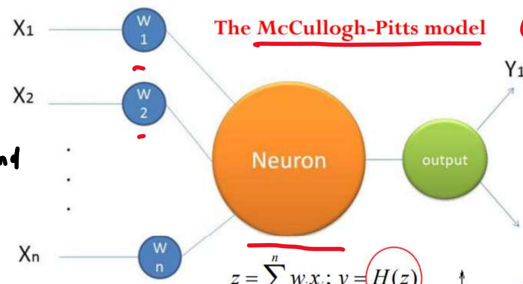
Sigmoi..



log (wx+b)

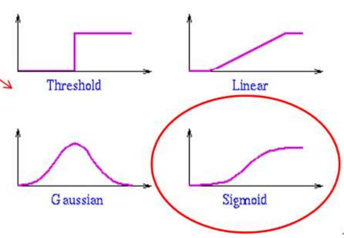
Linear Regression

The McCullogh-Pitts model (1940)



$$z = \sum_{i=1}^{n} w_i x_i ;\ y = H(z)$$

$y = f(z) = H(z)$



ReLU

Sigmoid

Wires : axon & dendrites
Connection weights: Synapse
Threshold function: activity in soma

ANN Overview: Network Architecture

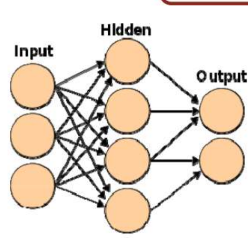
Connection patterns

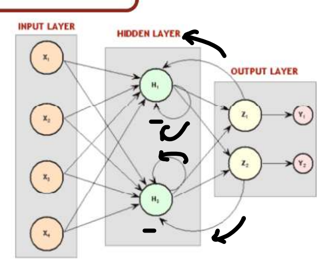
- Feed-Forward
- Recurrent

(Backpropagation Algorithm)

CNN

images

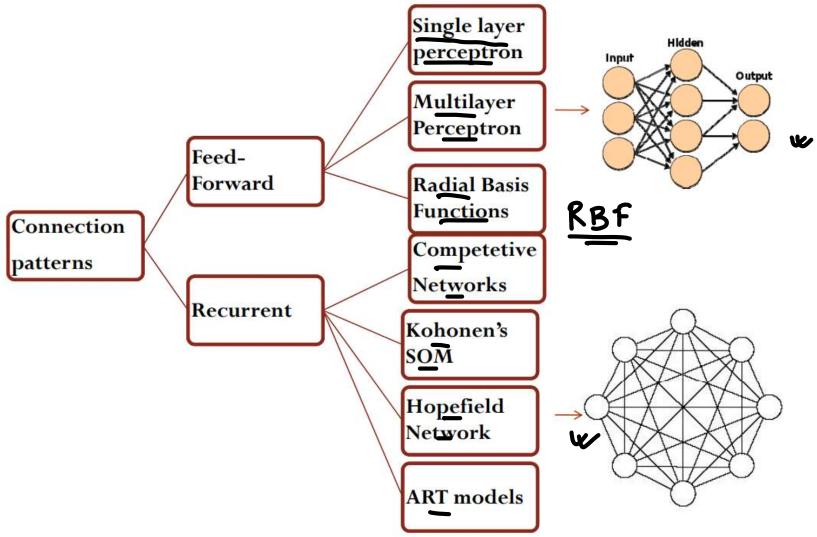




NN having feedback

RNN

Time Series



LSTM

Learning

- **What is the learning process in ANN?**
  - updating network architecture and connection weights so that network can efficiently perform a task

initialize weight → random value

- **What is the source of learning for ANN?**
  - Available training patterns
  - The ability of ANN to automatically learn from examples or input-out put relations (generalization ability)
- **How to design a Learning process?**
  - Knowing about available information
  - Having a model from environment: Learning Paradigm
  - Figuring out the update process of weights: Learning rules
  - Identifying a procedure to adjust weights by learning rules: Learning algorithm

Learning Paradigm

1. Supervised
   - The correct answer is provided for the network for every input pattern
   - Weights are adjusted regarding the correct answer
   - In reinforcement learning only a critique of correct answer is provided
2. Unsupervised
   - Does not need the correct output
   - The system itself recognize the correlation and organize patterns into categories accordingly
3. Hybrid
   - A combination of supervised and unsupervised
   - Some of the weights are provided with correct output while the others are automatically corrected.