iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37362

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Fundamentals of Neural Networks

Amey Thakur[1], Archit Konde[2]
[1, 2]*Department of Computer Engineering, University of Mumbai, Mumbai, Maharashtra, India*

Abstract: ***The purpose of this study is to familiarise the reader with the foundations of neural networks. Artificial Neural Networks (ANNs) are algorithm-based systems that are modelled after Biological Neural Networks (BNNs). Neural networks are an effort to use the human brain's information processing skills to address challenging real-world AI issues. The evolution of neural networks and their significance are briefly explored. ANNs and BNNs are contrasted, and their qualities, benefits, and disadvantages are discussed.***
The drawbacks of the perceptron model and their improvement by the sigmoid neuron and ReLU neuron are briefly discussed. In addition, we give a bird's-eye view of the different Neural Network models. We study neural networks (NNs) and highlight the different learning approaches and algorithms used in Machine Learning and Deep Learning. We also discuss different types of NNs and their applications. A brief introduction to Neuro-Fuzzy and its applications with a comprehensive review of NN technological advances is provided.
Keywords: Neural Networks (NNs), Artificial Neural Networks (ANNs), Biological Neural Networks (BNNs), Convolution Neural Networks (CNNs), Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), Feedforward Neural Network (FNN), Neural Network and Fuzzy Logic (Neuro-Fuzzy), Perceptron, Sigmoid, Gradient Descent, Rectified Linear Unit (ReLU), NN Architecture, NN Algorithms, NN Analysis, Deep Learning (DL), Machine Learning (ML).

I. INTRODUCTION

A. What are Neural Networks?

Artificial Neural Networks (ANNs) are algorithm-based systems inspired by Biological Neural Networks (BNNs). ANNs provide strong solutions to problems in several areas, including classification, prediction, filtering, optimization, pattern recognition, and function approximation. The genuine, biological nervous system is extremely complicated; artificial neural network algorithms seek to abstract this complexity and focus on what may theoretically matter most from an information-processing standpoint.

B. Why do we need Neural Networks?

The attraction of ANNs stems from the biological system's exceptional information processing features, such as nonlinearity, high parallelism, resilience, fault tolerance, learning, the capacity to handle imprecise and fuzzy information, and the ability to generalise [6][7].

C. When were Neural Networks introduced?

In the early 1940s, mathematicians Warren McCulloch and Walter Pitts developed a simple algorithm-based system to imitate human brain activity, paving the path for research into neural networks' application to artificial intelligence.

D. How do Neural Networks work?

Artificial neural networks learn to execute tasks by analysing samples and inferring rules without the need for direct involvement from a programmer. ANN-based models are empirical in nature; nonetheless, they may give practically correct answers for precisely or imprecisely defined issues, as well as phenomena that can only be understood through experimental data and field observations.

E. Where do we use Neural Networks?

Today's best commercial neural networks are so good that banks use them to process checks and post offices use them to recognise addresses. Driverless cars can recognise and respond to their surroundings in real-time, allowing them to travel safely. They do so by employing a set of algorithms known as deep neural networks, or DNNs. CNN, or Convolutional Neural Network, has become the most widely used approach in the field of face recognition. These are just a few instances of how neural networks have made our lives simpler.

©IJRASET: All Rights are Reserved

F. Analogy of BNN and ANN

In the following two respects, neural networks are similar to the human brain.

1) A neural network learns to gain knowledge.

2) The information of a neural network is stored in the intensities of inter-neuron connections known as synaptic weights.

TABLE 1
An Analogy of BNN and ANN

	
Biological Neurons	Silicon Transistors
200 Billion Neurons	Billion Bytes RAM
32 Trillion interconnections in Neurons	Trillions of Bytes on Disk
Neuron Size is 10^{-6} m	Single Transistor size is 10^{-9} m
Energy consumption is 6^{-10} joules per operation per second	Energy consumption is 10^{-16} joules per operation per second
Learning Capability	Programming Capability

G. Biological Neurons

Neurons, which are unique cells in the brain, carry information throughout the body. Each component of the neuron is crucial. Dendrites, the "hairs" that surround the cell body, transmit data. The cell body integrates data from several dendrites. The message is carried by the long, tail-like axon. From the axon terminals, information "jump" to the next neuron [1].

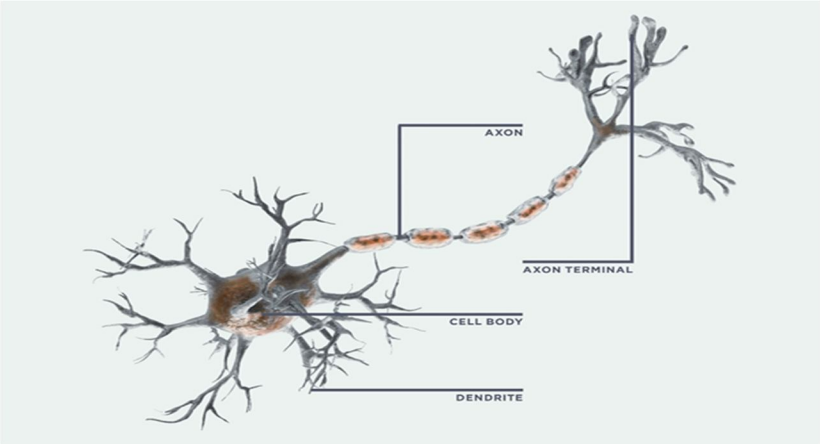


Figure 1: Biological Neuron

Your brain sends and receives information using both electrical and chemical impulses. To begin, substances known as neurotransmitters "transport" incoming signals across space between two neurons. This chemical surge causes the receiving neuron to shoot out an electrical signal, which transfers the message along with the cell. The message is then passed on to the next cell via neurotransmitters at the neuron's tip.

Neuronal connections are formed every time you learn anything new. Important brain connections become stronger with time, whereas less-used routes go away. Throughout our lifetimes, we generate new neurons and new neural connections.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

H. Biological Neural Networks

The degree to which a neural network accurately replicates a biological brain system varies. This is a major priority for some researchers; for others, the net's capacity to execute useful tasks (such as approximate or represent a function) is more essential than its biological plausibility. Despite the fact that our focus is almost entirely on the computational capabilities of neural networks, we will give a brief overview of certain biological neuronal properties that may assist us to understand some of the most essential aspects of artificial neural networks. Biological neural systems reveal traits that have specific computational advantages, in addition to providing the initial inspiration for artificial neural networks [1].

The structure of a real neuron (i.e., a brain or nerve cell) and the processing element (or artificial neuron) given in the rest of this paper are strikingly similar. In reality, the structure of a single neuron changes far less between species than the organisation of the system in which the neuron is a component.

The characteristics of real neurons indicate several essential features of the processing parts of artificial neural networks, including:

1) Many signals are received by the processing element.
2) Weight at the receiving synapse can change the signal.
3) The weighted inputs are added in the processing element.
4) The neuron transmits a single output when the conditions are right (enough input).
5) A neuron's output may be distributed to a large number of additional neurons (the axon branches).

Biological neurons also indicate the following characteristics of artificial neural networks:

a) The processing of data occurs on a local level (Despite the fact that additional modes of transmission, such as hormone activity, may offer ways to regulate the entire process,).
b) Distributed memory:
- The synapses or weights of neurons are where long-term memory is stored.
- The impulses provided by the neurons are related to short-term memory.

c) The strength of a synapse can be altered by experience.
d) Synaptic neurotransmitters can be excitatory or inhibitory.

I. Artificial Neurons

A neural network is defined by three factors:

1) Its architecture (the interconnection arrangement between neurons).
2) Its technique for calculating the weights on the connections (the training, or learning, algorithm).
3) Its activation functions.

For e.g., Perceptron - The perceptron takes in several binary inputs and produces a single binary output.

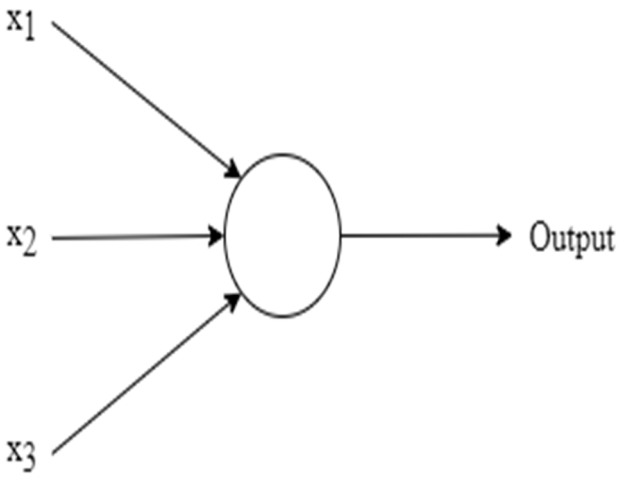


Figure 2: Perceptron

©IJRASET: All Rights are Reserved

The perceptron in the example contains three inputs, x1, x2, and x3. It might have more or fewer inputs in general. To compute the outcome, Rosenblatt developed a simple rule. He created weights, w1, w2, ..., which are actual values that reflect the significance of the various inputs to the output. The neuron's output, which is either 0 or 1, is decided by whether the weighted sum $\sum_j w_j x_j$ is less than or larger than some threshold number. A threshold, like weights, is a real number that serves as a neuronal parameter [2]. To express it more precisely:

$$\text{Output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases}$$

We may simplify this equation by expressing $\sum_j w_j x_j$ as a dot product, $w \cdot x \equiv \sum_j w_j x_j$, where w and x are vectors, whose components are the weights and inputs, respectively and shifting the threshold to the opposite side of the inequality, and replacing it with the perceptron's bias, b ≡ threshold. The perceptron rule may be modified by using the bias instead of the threshold:

$$\text{Output} = \begin{cases} 0 & \text{if } w \,.\, x + b \leq 0 \\ 1 & \text{if } w \,.\, x + b > 0 \end{cases}$$

That's how a perceptron works. That's the basic mathematical model. The perceptron may be thought of as a device that makes judgments by balancing the facts [2].

J. Artificial Neural Networks

An artificial neural network is a data-processing system that resembles biological neural networks [1][10] in terms of performance. Artificial neural networks were created as extensions of mathematical models of human cognition or neural biology, with the following assumptions:

1) Information is processed by a large number of basic components known as neurons.
2) Signals are transferred from one neuron to the next via connecting linkages.
3) Each connecting link has a weight associated with it, which doubles the signal conveyed in a conventional neural network.
4) Each neuron determines its output signal by applying an activation function (typically nonlinear) to its net input (sum of weighted input signals).

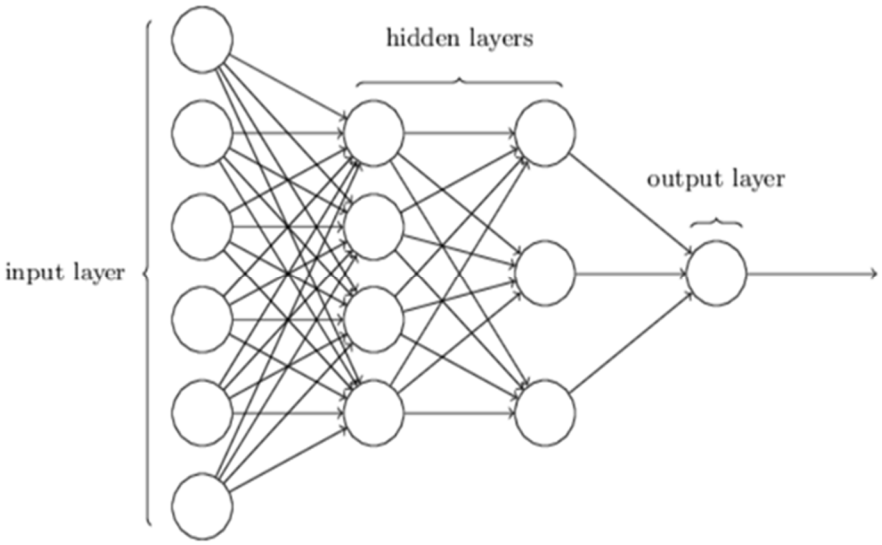


Figure 3: Artificial Neural Network

©IJRASET: All Rights are Reserved

Nodes in the hidden layer are connected to nodes in the input layer, while nodes in the output layer are connected to nodes in the input layer. Each link, like synapses in a human brain, has the ability to convey a signal to other neurons. The input layer receives data from the network. The raw data from the input layer is received and processed by the hidden layer. The acquired value is then transferred to the output layer, which will also analyse the information from the hidden layer and provide the output. The signal of each connection is a real number, and the output of each neuron is a non-linear function of the total of its inputs. The weight of neurons and connections is usually adjusted as learning progresses. The weight influences the intensity of the signal at a connection. Neurons may have a threshold that causes a signal to be delivered only if the aggregate signal surpasses that threshold [2].

TABLE 2
Characteristics of BNN and ANN

Characteristics	Biological Neural Network	Artificial Neural Network
Speed	Processes information at a slower rate. Response time is measured in milliseconds.	Information is processed at a faster rate. The response time is measured in nanoseconds.
Processing	Massively parallel processing.	Serial processing.
Size & Complexity	An extremely intricate and dense network of linked neurons of the order of 1011 neurons and 1015 interconnections.	Size and complexity are reduced. It is incapable of performing sophisticated pattern recognition tasks.
Storage	An extremely intricate and dense network of linked neurons with 1015 interconnections, including neurons on the order of 1011.	The term "replaceable information storage" refers to the practice of replacing fresh data with old data.
Fault tolerance	The fact that information storage is flexible means that new information may be added by altering the connectivity strengths without deleting existing information.	Intolerant of faults. In the event of a system failure, corrupt data cannot be recovered.
Control Mechanism	There is no unique control mechanism outside of the computational task.	Controlling computer activity is handled by a control unit.

II. IMPORTANCE OF NEURAL NETWORKS

Neural networks (NN), also known as artificial neural networks (ANN), are computer models – simply algorithms. They are a branch of machine learning. The capacity of neural networks to extract meaning from imprecise or complicated data in order to discover patterns and discern trends that are too convoluted for the human brain or other computer approaches is unrivalled. Ridesharing applications, Gmail smart sorting, and Amazon recommendations are just a few examples of how neural networks have made our lives easier.

Scientists are challenged to employ machines successfully for activities that are relatively straightforward for humans as current computers are becoming more powerful. We quickly learn to recognise the letter A or distinguish between an animal and a bird based on examples and feedback from a teacher. We may modify our replies and enhance our performance as we get more experience. Although we may ultimately be able to articulate rules that allow us to make such judgments, these rules may not always represent the method we employ. We can group similar patterns together even if we don't have an instructor. Another frequent human activity is attempting to attain a goal that entails maximising a resource (for example, time with one's family) while adhering to specific restrictions (such as the need to earn a living). Each of these problems demonstrates a challenge for which a computer solution may be considered.

©IJRASET: All Rights are Reserved

Traditional, sequential, logic-based digital computing succeeds in many areas but has struggled in others. Artificial neural networks (ANNs) were first developed some 50 years ago, with the goal of better understanding the brain and replicating some of its capabilities. Rapid advances in digital computing eclipsed early triumphs. Furthermore, promises made about the capabilities of early neural network models turned out to be overstated, throwing suspicion at the entire discipline.

Several reasons have contributed to the recent resurgence of interest in neural networks. For increasingly advanced network designs, training approaches have been devised to address the limitations of the early, basic neural nets. The simulation of brain processes is now more possible because of high-speed digital computers. The technology to create specialised hardware for neural networks is already accessible. However, although advances in conventional computing have made studying neural networks simpler, constraints in traditional computing's essentially sequential structure have prompted several new paths in neural network research. The study of biological brain networks, which are extremely parallel, may assist new methods to parallel computing. Traditional computer techniques to many sorts of issues have a high level of success, thus alternatives should be considered.

Researchers in a variety of fields are interested in neural networks for a variety of reasons. Signal processing and control theory has various applications for electrical engineers. The potential for hardware to effectively implement neural nets, as well as applications of neural nets to robotics, pique the interest of computer engineers. Neural nets offer promise for tough issues in artificial intelligence and pattern recognition, according to computer experts. Neural nets are a valuable tool for applied mathematics when it comes to modelling issues when the precise form of connections between specific variables is unknown.

There are several viewpoints on the nature of a neural network. Is it a specialised piece of computer hardware (such as a VLSI chip) or computer software, for example? We'll assume that neural networks are just mathematical representations of information processing. They provide a new way of describing relationships than Turing machines or computers with stored instructions. The availability of computer resources, either software or hardware, substantially increases the utility of the methodology, especially for big problems, as it does with other numerical approaches.

III. HISTORY OF NEURAL NETWORKS

The first neural networks were created in the early 1940s by mathematicians Warren McCulloch and Walter Pitts, who created a basic algorithm-based system to mimic human brain activity. The perceptron, a revolutionary algorithm created to handle complicated recognition tasks, was invented by Cornell University's Frank Rosenblatt in 1957, accelerating work in the field. The lack of computer capacity required to analyse vast volumes of data slowed progress in the four decades that followed. In the 2000s, computer scientists finally got what they needed, owing to increased processing power and sophisticated hardware, as well as the availability of enormous data sets from which to draw, and neural networks and AI took off, with no end in sight. Consider that 90% of internet data has been produced since 2016. Because of the Internet of Things' rapid expansion, this rate will continue to rise (IoT) [1].

IV. ANALYSIS OF NEURAL NETWORKS

A. *Architecture of Neural Network*

Neural Networks are multi-input, multi-output systems made up of artificial neurons. A Neural Network's principal function is to convert input into meaningful output. A Neural Network usually has an input and output layer, as well as one or more hidden layers. All of the neurons in a Neural Network impact each other, therefore they are all linked. The network can recognise and observe every element of the dataset in question, as well as how the various pieces of data may or may not be related to one another. This is how Neural Networks can detect highly complicated patterns in massive amounts of data.

The flow of information in a Neural Network occurs in two ways -

1) *Feedforward Networks:* The signals in this model only go in one way, to the output layer. With zero or many hidden layers, feedforward networks have an input layer and a single output layer. Pattern recognition makes extensive use of them.
2) *Feedback Networks:* The recurrent or interactive networks in this approach handle the series of inputs using their internal state (memory). Signals can go in both ways through the network's loops (hidden layers). They're commonly employed in jobs that require a sequence of events to happen in a certain order.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

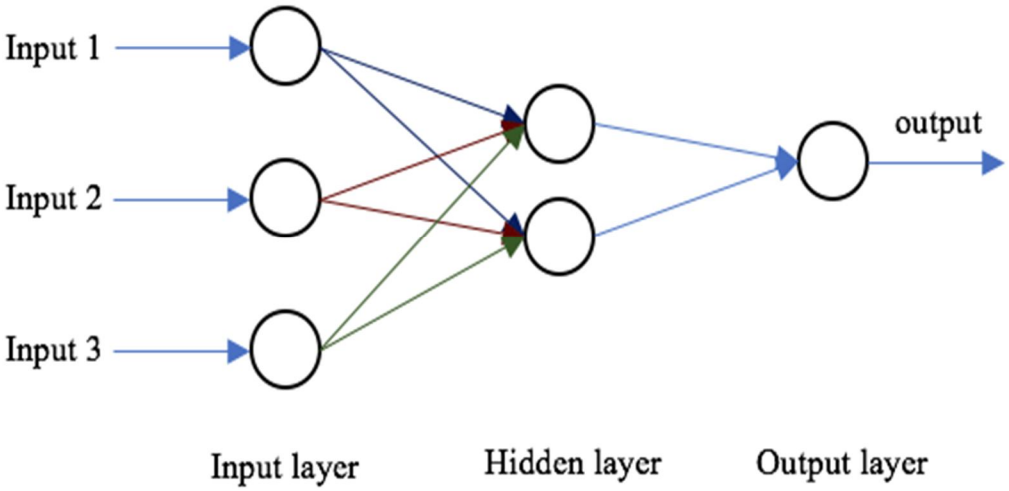


Figure 4: Architecture of Neural Network

The three layers of a neural network are: The input layer comes first. The input neurons that send data to the hidden layer are housed in this layer. The input data is computed in the hidden layer, and the output is sent to the output layer. Weight, activation, and cost functions are all part of it.

The weight of the link between neurons is measured in numerical numbers. The neural network's learning ability is determined by the weight between neurons. The weight between the neuron's changes throughout the learning of artificial neural networks. Weights are allocated at random in the beginning.

The "activation function" is utilised to normalise the neuron's output. The mathematical equations that calculate the neural network's output are known as activation functions. The data is converted with a 0 mean and a 1 standard deviation. This is known as standardisation. Each neuron has a unique activation function. Without mathematical thinking, it is impossible to comprehend. It also aids in the normalisation of the output in the 0 to 1 or -1 to 1 range. The transfer function is another name for an activation function.

B. Characteristics of Neural Network

With characteristics of Neural Networks, the capacity to solve problems in a human-like manner [8] and to apply that ability to large datasets neural networks have the following characteristics -

1) Adaptive Learning: Neural networks, like humans, represent non-linear and complicated interactions and build on prior knowledge through adaptive learning. Software that teaches arithmetic and language arts, for example, employs adaptive learning.
2) Fault Tolerance: Neural networks can fill in the gaps when important portions of a network are lost or absent. This skill is particularly valuable in space travel, where electronic gadget failure is a constant risk.
3) Prognosis: The capacity of neural networks to forecast based on models has several uses, including weather and transportation.
4) Real-Time Response: As with self-driving vehicles and drone navigation, neural networks may give real-time responses.
5) Self-Organization: Neural networks are especially suited for organising the complex visual issues posed by medical image analysis due to their capacity to cluster and categorise large volumes of data.

C. Features of Neural Network

Neural Networks are used to accomplish a variety of tasks -

1) *Associating:* Pattern recognition may be taught to neural networks. When you present an unknown version of a pattern, the network correlates it with the most similar version in its memory and switches to that version.
2) *Classification:* Neural Networks classify patterns or information into predetermined categories.
3) *Clustering:* They categorise the data by identifying a unique characteristic without having any prior knowledge of the data.
4) *Prediction:* They get the anticipated result from the provided input.

©IJRASET: All Rights are Reserved

D. Learning Techniques of Neural Network

Neural Network Learning Techniques -

1) *Supervised Learning:* The training data is fed into the network, and the intended output is known, thus the weights are changed until production produces the required value.
2) *Unsupervised Learning:* Train the network whose output is known using the input data. The network classifies the input data and changes the weights by extracting features from the input data.
3) *Reinforcement Learning:* In this case, the output value is unknown, but the network offers feedback on whether the output is correct or incorrect. It is called Semi-Supervised Learning.
4) *Offline Learning:* Only after the training data has been given to the network are the weight vector and threshold adjustments done. Batch learning is another name for it.
5) *Online Learning:* After submitting each training sample to the network, the weight and threshold are adjusted.

E. Learning Rules in Neural Network

Mathematical reasoning is used in the learning rule. It promotes a Neural Network to benefit from current situations in order to improve its efficiency and performance. The brain's neuronal structure is altered during the learning process. Their activity determines whether the quality of its synaptic connections improves or deteriorates. Neural network learning rules -

1) Correlation learning rule: It is analogous to supervised learning.
2) Delta learning rule: A change in a node's sympatric weight equals the mistake multiplied by the input.
3) Hebbian learning rule: It defines how to adjust the weights of a system's nodes.
4) Perceptron learning rule: The network begins learning by assigning a random value to each load.

F. Advantages of Neural Network

1) A neural network is capable of tasks that a linear programme is incapable of performing.
2) When a component of the neural network fails, the parallel nature of the network remains unaffected.
3) A neural network learns without the need for reprogramming.
4) It may be used for any purpose.
5) It is not difficult to carry out.

G. Limitations of Neural Network

1) To function, the neural network requires training.
2) A neural network's architecture differs from that of a microprocessor. Emulation is so required.
3) Large neural networks need a lot of processing time.

V. NEURAL NETWORKS AND ARTIFICIAL INTELLIGENCE

In certain quarters, neural networks are synonymous with artificial intelligence. Others regard them as a "brute force" method, characterised by a lack of intellect because they begin with a blank slate and pound their way through to an accurate model. According to this view, neural networks are successful yet inefficient in their modelling approach since they do not make assumptions about the functional relationships between output and input.

The top AI research organisations are pushing the discipline's frontiers by training increasingly bigger neural networks. Using brute force is effective.

It is a required, though not sufficient, the prerequisite for AI advances. The development of more generic AI by OpenAI stresses a brute force approach, which has proven effective with well-known models such as GPT-3.

Algorithms such as Hinton's capsule networks require many fewer samples of data to arrive at the right model; hence, the present approach has the potential to overcome deep learning's brute force inefficiencies.

While neural networks are helpful as a function approximator, translating inputs to outputs in many perceptual tasks, they may be coupled with other AI approaches to accomplish more complicated tasks to obtain a broader intelligence. Deep reinforcement learning, for example, embeds neural networks within a reinforcement learning framework, mapping actions to rewards to achieve goals. Examples are victories by Deepmind in video games and the board games Go by AlphaGo.

©IJRASET: All Rights are Reserved

VI. LEARNING AND ALGORITHM

Today's most advanced commercial neural networks include millions of inputs and billions of connections between them. We couldn't reasonably hope to compute the weights and biases for such large-scale decision-making models by hand. So, what we'd want to do is create algorithms that will allow our neural network to automatically infer rules from instances (training set) and learn from them in order to fine-tune our network's weights and biases for making accurate judgments. This tuning occurs in reaction to environmental inputs and without the participation of a programmer. These learning techniques allow neural networks to learn to solve problems by themselves.



Figure 5: Example

A. How Can We devise Such Algorithms?

Assume we have a perceptron network that we want to utilise to learn how to solve an issue. The raw pixel data from the dog image above, for example, might be one of the network's inputs. In addition, we'd want the network to learn weights and biases so that the network's output accurately identifies the picture as a dog. To see how learning may work, consider making a modest adjustment to one of the network's weights (or biases). What we want is for this little adjustment in weight to result in only a minor change in network output. If it were true that a little change in weight (or bias) generates just a minor change in output, we could utilise this knowledge to tweak the weights and biases to get our network to behave more as we want it to. Then we'd do it again, adjusting the weights and biases to get better and better results. The network would be gaining knowledge.

The issue is that this does not occur when our network comprises perceptrons. In reality, a slight change in the weights or bias of any one perceptron in the network might occasionally cause its output to fully flip, say from 0 to 1. That switch might then force the remainder of the network's behaviour to entirely alter in some extremely complex way.

We can solve this difficulty by developing a new form of artificial neuron known as a sigmoid neuron. Sigmoid neurons are perceptrons that have been adjusted so that tiny changes in their weights and biases produce only minor changes in their output. This is the critical fact that will enable a network of sigmoid neurons to learn.

Sigmoid neurons may receive any input between 0 and 1, for example, 0.420, and their output is provided by (w. x + b) where σ is known as the sigmoid function [2]. As a result, the output of a sigmoid neuron with inputs x1, x2, ..., weights w1, w2, ..., and bias b equals $(z) \equiv 1 / 1 + e^{-z}$

However, because it is difficult to train, neural networks seldom utilise this sigmoid function nowadays. As a result, various functions were explored, and one function in particular that performs pretty well and is easier to train is ReLU, which stands for Rectified Linear Units.

The rectified linear activation function, or ReLU for short, is a piecewise linear function that will output the input directly if it is positive, else it will output zero. Because it is faster to train and typically produces better results, it has become the default activation function for many types of neural networks. It is provided by f(x) = max (0, x), where x is the neuron's input.

©IJRASET: All Rights are Reserved

B. Gradient Descent

Gradient descent is the term given to a frequently used optimization algorithm that modifies weights based on the error they produce. The gradient is another term for slope, and slope, in its normal form on an x-y graph, illustrates how two variables relate to one another: increase overrun, money change over time change, and so on. In this example, the slope we are interested in reflects the connection between the network's error and a single weight; that is, how the error varies when the weight is modified.

To be more specific, whatever weight will result in the least amount of error? Which one represents the signals in the incoming data accurately and converts them to the right classification? Who can hear the word "nose" in an input image and recognise that it should be labelled as a face rather than a frying pan?

As a neural network learns, it gradually changes numerous weights in order to accurately translate signals to meaning. The connection between network Error and each of those weights is a derivative, dE/dw, which indicates the degree to which a little change in weight produces a small change in error.

Each weight is only one element in a deep network with multiple transformations; the weight signal goes via activations and sums over numerous layers. As a result, we employ the chain rule of calculus to traverse the network's activations and outputs in order to determine the weight in question and its link to the total inaccuracy.

In calculus, the chain rule states that –

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

The relationship between the net's error and a single weight in a feedforward network is as follows -

$$\frac{dError}{dweight} = \frac{dError}{dactivation} * \frac{dactivation}{dweight}$$

To put it another way, given two variables, Error and weight, which are mediated by a third variable, activation, through which the weight is passed, you can calculate how a change in weight affects a change in Error by first calculating how a change in activation affects a change in Error, and then how a change in weight affects a change in activation.

VII. LEARNING STRATEGIES

A. Neural Network for Machine Learning

1) Back Propagation Network (supervised classification).
2) Deep Neural Networks (unsupervised clustering).
3) Hopfield Network (for pattern association).
4) Multilayer Perceptron (supervised classification).

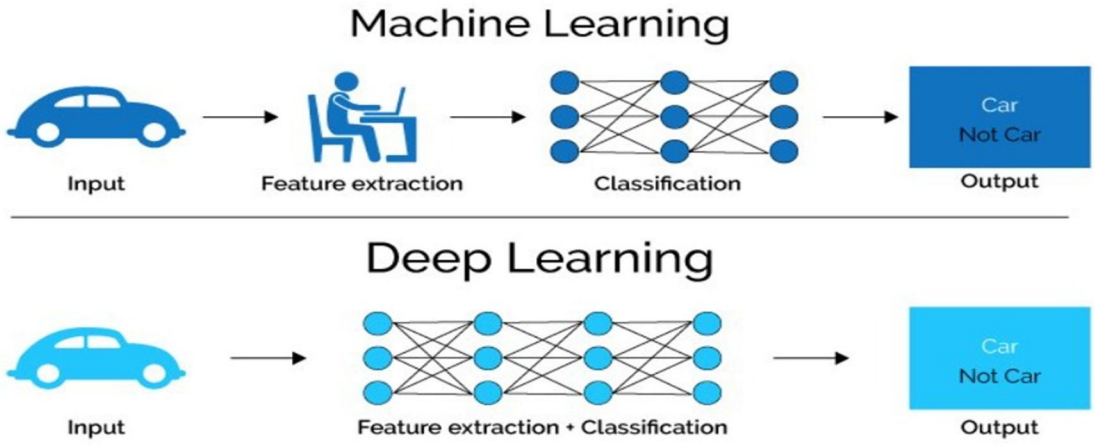


Figure 6: Machine Learning and Deep Learning Strategies

©IJRASET: All Rights are Reserved

B. Neural Network for Deep Learning

Deep Learning employs topologies similar to those used in Neural Networks.

1) Convolutional deep belief networks.
2) Convolutional neural networks.
3) Deep Boltzmann machines.
4) Deep belief networks.
5) Feed-forward neural networks.
6) Recurrent neural network.
7) Multi-layer perceptrons (MLP).
8) Recursive neural networks.
9) Self-Organizing Maps.
10) Stacked de-noising auto-encoders.

C. What Distinguishes Deep Learning from Neural Networks?

While it was suggested in the description of neural networks, it is worth emphasising again. The phrase deep learning refers to the amount of layers in a neural network. A deep learning algorithm may be defined as a neural network with more than three layers, which includes the inputs and outputs.

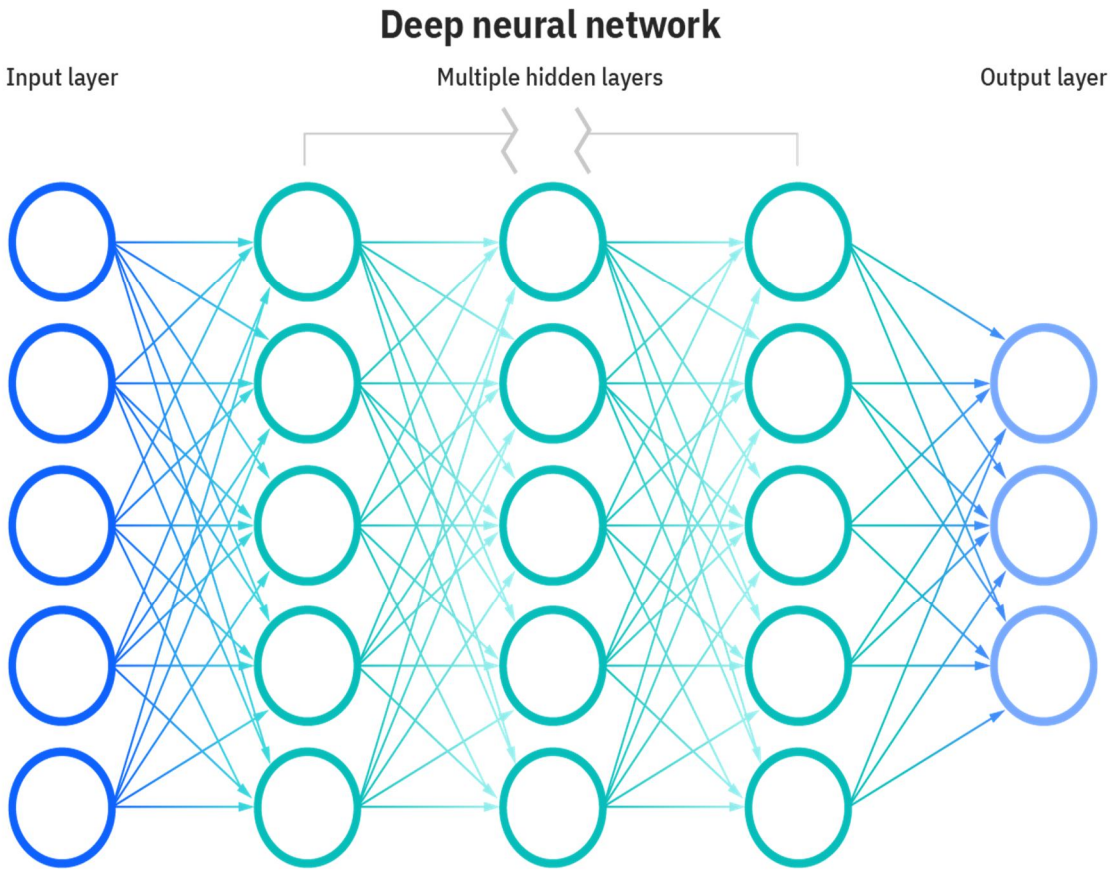


Figure 7: Deep Neural Network

The majority of deep neural networks are feed-forward, which means they only flow in one way, from input to output. However, you may train your model using backpropagation, which involves moving in a reverse way from output to input. Backpropagation enables us to quantify and assign the error associated with each neuron, allowing us to correctly adapt and fit the algorithm.

©IJRASET: All Rights are Reserved

VIII. ARCHITECTURE OF NEURAL NETWORKS

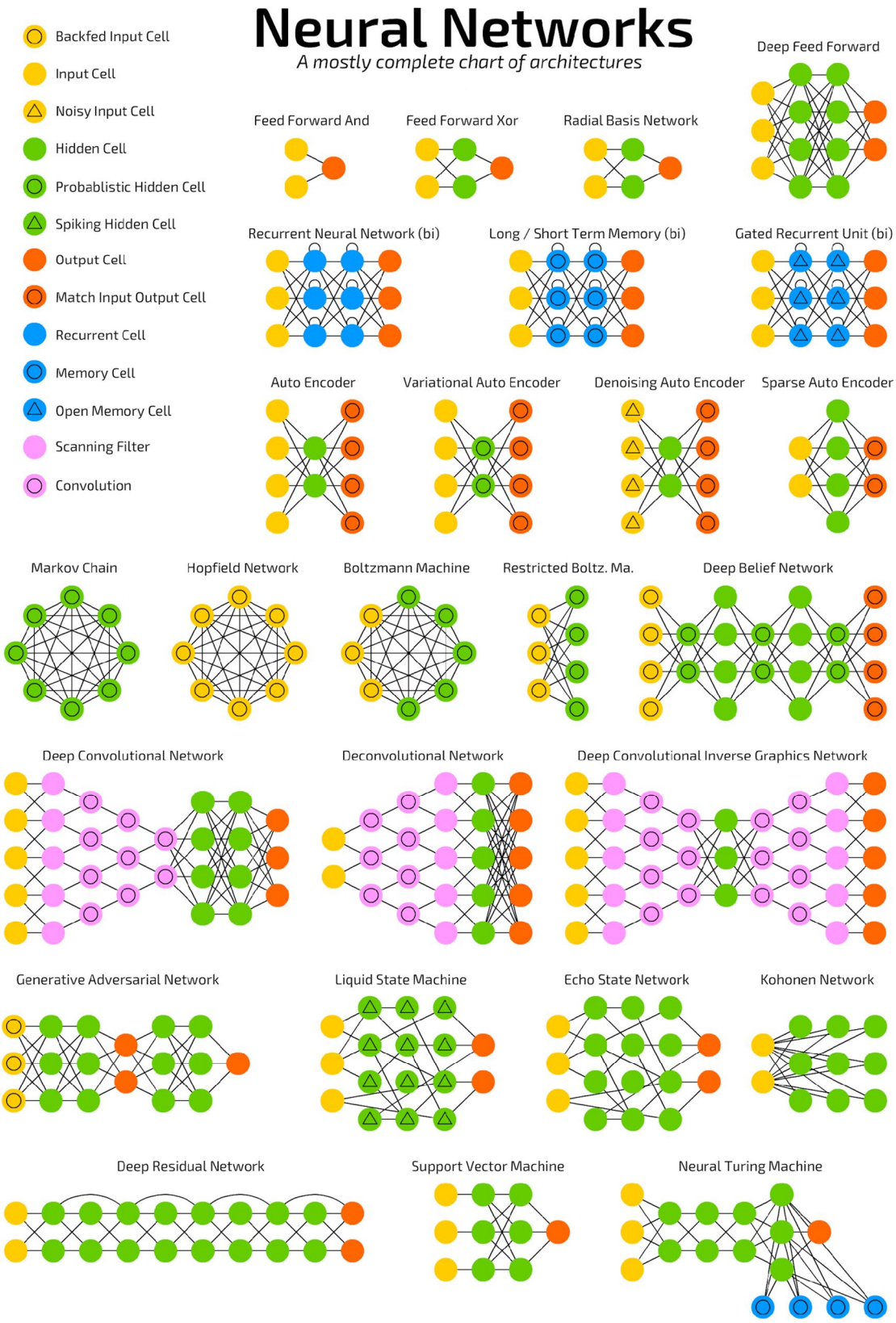


Figure 8: Neural Network Architectures

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

IX. INTRODUCTION TO NEURAL NETWORK ALGORITHMS

A. Autoencoder (AE)

AEs are generally used to decrease the amount of random variables that the system must consider so that it can learn a representation for a piece of data and, as a result, process generative data models.

B. Bidirectional Recurrent Neural Network (BRNN)

A BRNN's objective is to expand the network's information inputs by linking two hidden, directionally opposed layers to the same output. The output layer can acquire information from both past and future states by using BRNNs.

C. Boltzmann Machine (BM)

This method, which is a recurrent neural network, is capable of learning internal representations and can describe and solve difficult coupled problems.

D. Convolutional Neural Network (CNN)

CNNs are a feed-forward neural network designed to reduce pre-processing and are most often used to evaluate visual images.

E. Deconvolutional Neural Network (DNN)

DNNs allow for the development of hierarchical picture representations without the need for supervision. To add more sophisticated characteristics to a picture, each level of the hierarchy groups information from the previous level.

F. Deep Belief Network (DBN)

A DBN may learn to rebuild its inputs probabilistically by employing layers as feature detectors after being trained with an unsupervised collection of samples. You may train a DBN to do supervised classifications using this method.

G. Deep Convolutional Inverse Graphics Network (DCIGN)

A DCIGN model is designed to develop an interpretable representation of pictures that the system divides into three-dimensional scene structure elements including illumination changes and depth rotations. Many layers of convolutional and deconvolutional operators are used in a DCIGN.

H. Deep Residual Network (DRN)

DRN’s aid in the management of complex deep learning models and workloads. A DRN avoids results from degrading by having several layers.

I. Denoising Autoencoder (DAE)

DAEs are used to recover data from damaged inputs by forcing the hidden layer to acquire more resilient features. As a consequence, the output is a refined version of the original data.

J. Echo State Network (ESN)

A random, massive, fixed recurrent neural network is used in an ESN, and each node gets a nonlinear response signal. To achieve learning flexibility, the algorithm establishes and assigns weights and connections at random.

K. Extreme Learning Machine (ELM)

This approach creates a linear model by learning hidden node output weightings in one step. ELMs are capable of generalisation and learning at a far quicker rate than backpropagation networks.

L. Feed Forward Neural Network (FF or FFNN) and Perceptron (P)

These are the fundamental neural network algorithms. A feedforward neural network is an artificial neural network with node connections that create a cycle; a perceptron is a binary function with just two outcomes (up/down; yes/no, 0/1).

©IJRASET: All Rights are Reserved

M. Gated Recurrent Unit (GRU)

To accomplish machine learning tasks like grouping and remembering, GRUs employ connections established by node sequences. The control of model information flow allows GRUs to modify outputs.

N. Generative Adversarial Network (GAN)

The discriminative and generative neural networks are fought against each other in this system. In order to mimic high-level intellectual activities, it is necessary to discriminate between actual and synthetic results.

O. Hopfield Network (HN)

This type of recurrent artificial neural network has binary threshold nodes and is an associative memory system. HNs provide a paradigm for understanding human memory since they are designed to converge to a local minimum.

P. Kohonen Network (KN)

A KN is a two-dimensional map that organises a problem space. Self-organizing maps (SOMs) vary from conventional problem-solving techniques in that they employ competitive learning rather than error-correction learning.

Q. Liquid State Machine (LSM)

An LSM adds the concept of time as an element to third-generation machine learning. Because LSMs maintain memory throughout processing, they activate spatiotemporal neural networks. LSMs are used in physics and computational neuroscience.

R. Long/Short-Term Memory (LSTM)

In prediction issues involving sequence, LSTM is capable of learning or remembering order dependency. A cell, an input gate, an output gate, and a forget gate are all contained in an LSTM unit. Values are stored in cells for as long as they are needed. Through LSTM connections, each unit controls value flow. In difficult problem areas such as speech recognition and machine translation, this sequencing capacity is critical.

S. Markov Chain (MC)

An MC is a mathematical process that explains a series of potential occurrences in which the probability of each event is solely determined by the state achieved in the preceding event. Type-word predictions and Google PageRank are two examples of how to use them.

T. Neural Turing Machine (NTM)

An NTM, which is based on data scientist Alan Turing's work from the mid-twentieth century, conducts computations and expands the capabilities of neural networks by connecting with external memory. NTM is used in robots and is seen as one of the ways to create an artificial human brain by developers.

U. Radial Basis Function Networks (RBF nets)

RBF nets are used by programmers to model data that depicts an underlying trend or function. Using bell curves or non-linear classifiers, RBF networks learn to mimic the underlying trend. Simple linear classifiers that function on lower-dimensional vectors analyse less thoroughly than non-linear classifiers. These networks are used in system control and time series forecasting.

V. Recurrent Neural Network (RNN)

RNNs are memory-based models that simulate sequential interactions. An RNN calculates a new memory or hidden state based on the current input and the previous memory state at each time step. Composing music, controlling robots, and recognising human actions are just a few examples of applications.

W. Restricted Boltzmann Machine (RBM)

In an unsupervised context, an RBM is a probabilistic graphical model. The connections between binary neurons in each of these levels, as well as the visible and hidden layers, make up an RBM. Filtering, feature learning, and classification are all things that RBNs can help with. Risk identification, as well as commercial and economic assessments, are examples of possible applications.

©IJRASET: All Rights are Reserved

X. Support Vector Machine (SVM)

An SVM algorithm creates a model that allocates fresh instances to one of two categories based on training example sets that are relevant to one of two potential categories. The model then depicts the instances as mapped points in space, separating the different category examples by the largest gap feasible. The system then maps new instances into that same space, predicting which category they belong to based on which side of the divide they are on. Face recognition and bioinformatics are two examples of applications.

Y. Variational Autoencoder (VAE)

A VAE is a sort of neural network that assists in the creation of complicated models from data sets. An autoencoder, in general, is a deep learning network that uses backpropagation to attempt to rebuild a model or match target outputs to supplied inputs. In the fields of image creation and reinforcement learning, a VAE produces state-of-the-art machine learning outcomes.

X. TYPES OF NEURAL NETWORKS

A. Convolutional Neural Network

A convolutional neural network (CNN) is based on a multilayer perceptron variant. A CNN may have one or more convolutional layers. These levels might be fully linked or pooled. The convolutional layer performs a convolutional operation on the input before sending the output to the next layer. Because of this convolutional process, the network may be considerably deeper while requiring many fewer parameters. Convolutional neural networks do exceptionally well in image and video recognition, natural language processing, and recommender systems as a result of this capacity. Convolutional neural networks do well in semantic parsing and paraphrase identification as well. They are also used in image categorization and signal processing [4][5].

CNN's are also utilised in image analysis and identification in agriculture, where weather characteristics gathered from satellites such as LSAT are used to forecast the growth and production of a plot of land. A Convolutional Neural Network is seen in the figure below.

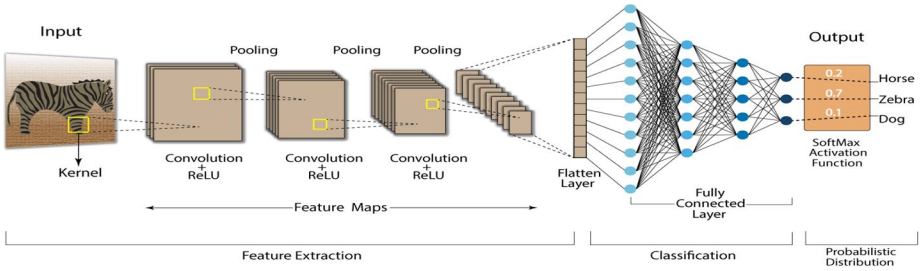


Figure 9: Convolutional Neural Network

B. Feedforward Neural Network – Artificial Neuron

One of the most basic forms of artificial neural networks. The data in a feedforward neural network is routed through the various input nodes until it reaches the output node. In other words, data travels in just one path from the first layer to the output node. This is often referred to as a front propagating wave, and it is typically produced by employing a classifying activation function.

There is no backpropagation in this form of the neural network, and data only flows in one direction [16].

A feedforward neural network might contain a single layer or many layers. The sum of the products of the inputs and their weights is determined in a feedforward neural network. The output is then fed this. An example of a single layer feedforward neural network is shown below.

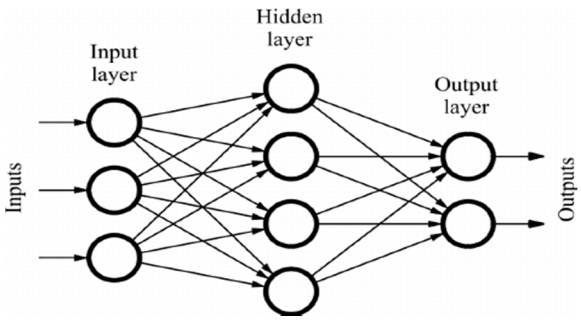


Figure 10: Feedforward Neural Network

©IJRASET: All Rights are Reserved

Face recognition and computer vision are two applications that make use of feedforward neural networks. This is due to the difficulty in classifying the target classes in these applications. A basic feedforward neural network is well-suited to dealing with noisy data.

C. Recurrent Neural Network (RNN) – Long Short-Term Memory (LSTM)

A Recurrent Neural Network (RNN) is a form of artificial neural network in which the output of a certain layer is stored and fed back into the input. This aids in predicting the layer's result. The first layer is generated in the same way as it is in a feedforward network. That is, with the product of the weights and characteristics. The recurrent neural network process, on the other hand, begins at later levels. Each node remembers some information from the previous time step from one-time step to the next. In other words, while computing and performing operations, each node function as a memory cell. The neural network starts with front propagation as normal but remembers the information it may require later. If the forecast is erroneous, the system self-learns and uses backpropagation to produce the right prediction. In text-to-speech conversion technology, this sort of neural network is extremely successful. Here's an example of a recurrent neural network [15].

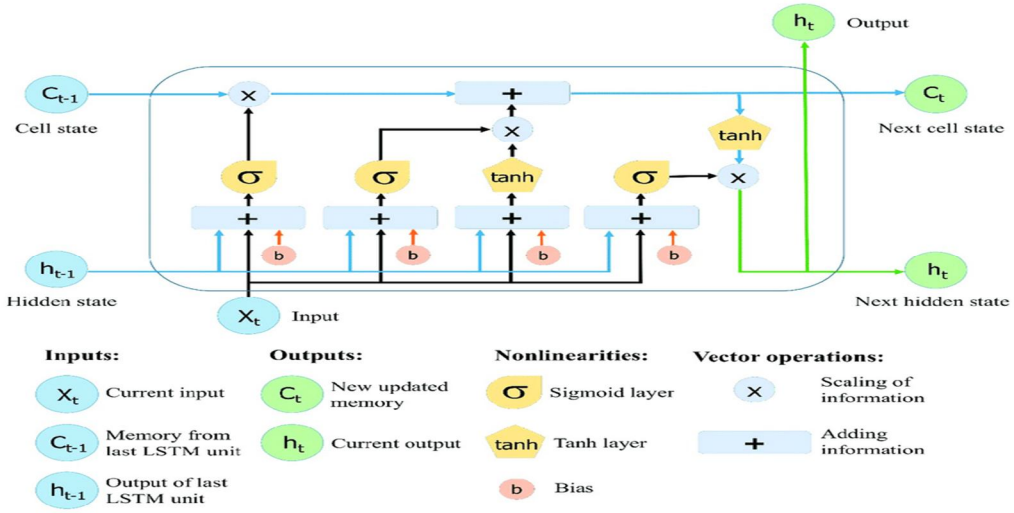


Figure 11: Recurrent Neural Network (RNN) - Long Short-Term Memory (LSTM)

D. Radial Basis Function Neural Network

A radial basis function takes into account the distance of any point from the centre. These neural networks are composed of two layers. The characteristics are merged with the radial basis function in the inner layer. The output of these characteristics is then used to calculate the identical result in the following time step [18]. A radial basis function neural network is seen in the diagram below.

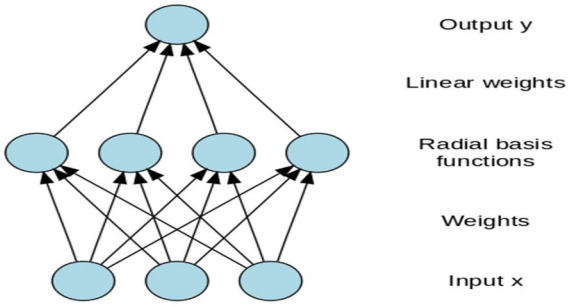


Figure 12: Radial Basis Function Neural Network

In power restoration systems, the radial basis function neural network is extensively utilised. Power systems have grown in size and complexity in recent decades. This raises the possibility of a blackout. This neural network is used in power restoration systems to restore electricity as quickly as feasible.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

E. Multilayer Perceptron

A multilayer perceptron is made up of three or more layers. It is employed in the classification of data that cannot be separated linearly. It is a form of a fully connected artificial neural network. This is due to the fact that every node in one layer is linked to every node in the next tier [20].

A nonlinear activation function is used by a multilayer perceptron (mainly hyperbolic tangent or logistic function). A multilayer perceptron is seen below.

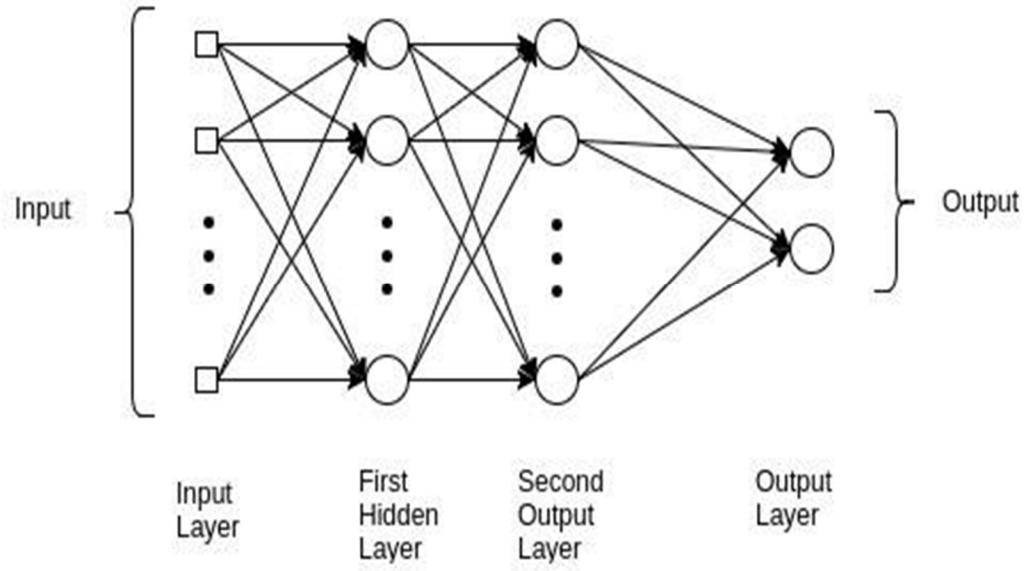


Figure 13: Multilayer Perceptron

F. Modular Neural Network

A modular neural network is made up of several networks that work independently and accomplish sub-tasks. During the computing process, the various networks do not interact or notify one another. They operate individually to get the desired result.

As a result, by dividing a big and complicated computing process into separate components, a large and complex computational process may be completed much quicker. Because the networks are not interacting or even linked to one other, the calculation speed improves. Here's a graphic depiction of a Modular Neural Network [22].

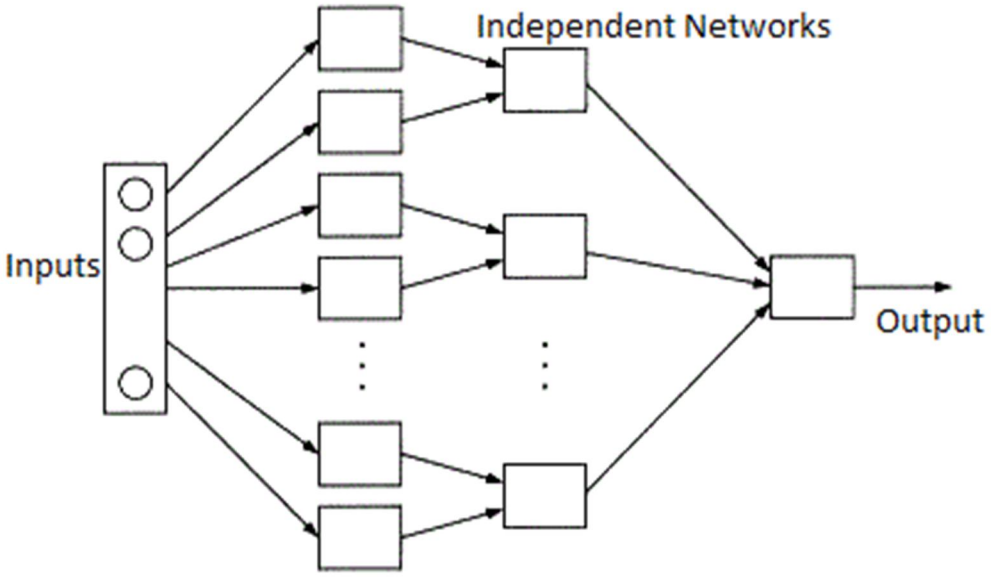


Figure 14: Modular Neural Network

©IJRASET: All Rights are Reserved

G. Sequence-To-Sequence Models

Two recurrent neural networks comprise a sequence to sequence model. The input is processed by an encoder, while the output is processed by a decoder. The settings used by the encoder and decoder might be the same or different. This approach is particularly useful when the length of the input data differs from the length of the output data. Chatbots, machine translation, and question answering systems are the most common applications for sequence-to-sequence models [19].

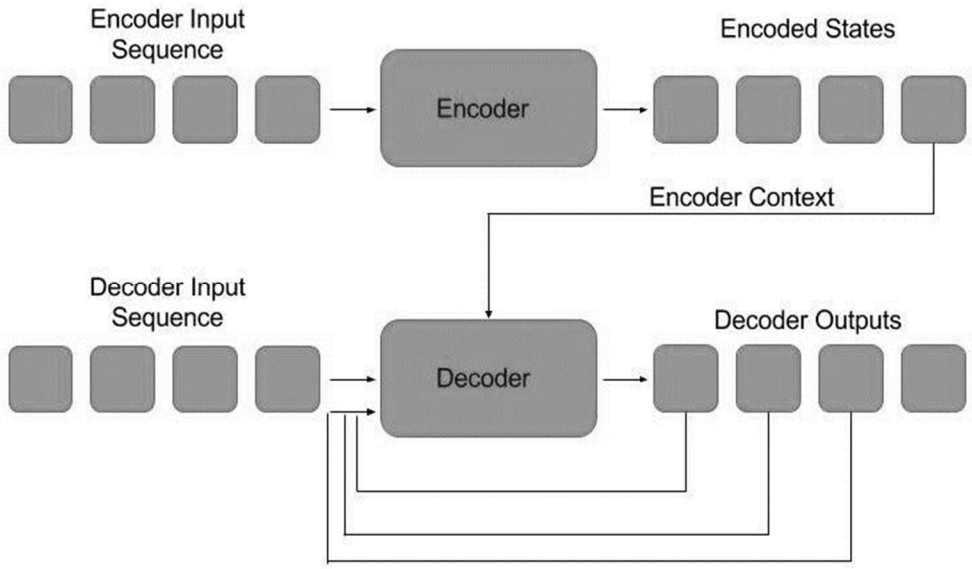


Figure 15: Sequence-To-Sequence Models

There are several forms of artificial neural networks that function in various ways to accomplish various results. The most significant aspect of neural networks is that they are meant to function similarly to neurons in the brain. As a consequence, they are meant to learn and develop more like data and usage increase. Unlike standard machine learning algorithms, which tend to plateau at a certain point, neural networks may genuinely expand as additional data and use is added. As a result, many experts think that various forms of neural networks will serve as the foundation for next-generation Artificial Intelligence.

XI. REAL-LIFE APPLICATIONS OF NEURAL NETWORKS

Both in terms of development and application, neural network research is a very multidisciplinary area. The range of neural networks' usefulness is demonstrated by a quick sampling of some of the areas in which they are now being used [24]. The examples vary from commercial triumphs to active research fields with bright future prospects.

1) *Handwriting Recognition* – The concept of handwriting recognition has gained a lot of traction. This is due to the growing popularity of portable devices such as the Palm Pilot. As a result, neural networks may be used to detect handwritten characters.
2) *Image Compression* – Neural networks receive and analyse large quantities of data all at once. As a result, they're useful for picture compression. With the growth of the Internet and the increasing use of pictures on websites, utilising neural networks for image compression is a viable option.
3) *Stock Exchange Prediction* – The stock market's day-to-day operations are extremely complex. Many factors influence whether a stock will rise or fall in value on any given day. As a result, neural networks can quickly evaluate and sort a large amount of data. As a result, they may be used to forecast stock prices.
4) *Travelling Salesman Problem* – The travelling salesman issue may also be solved using neural networks. However, this is merely an assumption to a certain degree.
5) *Medicine, Electronic Nose, Security, and Loan Applications* - These are some applications that are in their proof-of-concept stage, with the acception of a neural network that will decide whether or not to give a loan, something that has previously been utilised more successfully than many individuals.
6) *Miscellaneous Applications* - These are some very interesting applications of neural networks.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

XII. NEURAL NETWORKS AND FUZZY LOGIC (NEURO-FUZZY)

Fuzzy logic is a type of logic that was designed to convey the degree of truthiness by assigning values between 0 and 1, as opposed to classic Boolean logic that represented 0 and 1. One thing unites fuzzy logic with neural networks. They can be used to tackle pattern recognition difficulties and other non-mathematical challenges. Neuro-fuzzy systems are systems that combine fuzzy logic with neural networks. To perform better, these hybrid systems might combine the benefits of neural networks with fuzzy logic. Fuzzy logic and neural networks have been combined to be used in the following applications:

1) Automotive engineering
2) Applicant screening of jobs
3) Control of crane
4) Monitoring of glaucoma

Neural Networks Learning Algorithms are combined with fuzzy reasoning in a hybrid (neuro-fuzzy) model. The neural network decides on parameter values, whereas fuzzy logic governs if-then rules [23].

XIII. FUTURE OF NEURAL NETWORKS

1) *Fuzzy Logic Integration:* Fuzzy logic acknowledges more than just true and false values; it also considers relative notions such as somewhat, sometimes, and generally. Fuzzy logic and neural networks are utilised in a range of applications, such as job screening, auto-engineering, crane control, and glaucoma monitoring. Future neural network applications will rely heavily on fuzzy logic.
2) *Improvement of Existing Technologies:* Neural networks have barely scratched the surface of what they can achieve, thanks to new software and hardware, as well as existing neural network technologies and the enhanced processing capacity of neurosynaptic designs. Better problem-solving and training approaches that are quicker, simpler, and more human-like have a wide range of commercial applications.
3) *Pulsed Neural Networks:* Mammalian biological brain networks link and interact by pulsing and employ the timing of pulses to convey information and conduct calculations, according to recent neurobiological experiment findings. This acknowledgement has greatly increased major research, including theoretical studies, model creation, neurobiological modelling, and hardware deployment, all targeted at making computing more like how our brains work.
4) *Robotics:* Numerous predictions have been made concerning robots that will be able to feel, perceive, and anticipate what will happen in the environment around them. From the Terminator film series to Blade Runner and Westworld, these prophecies feature dismal interpretations of the future. However, according to Futurist Richard Yonck, we still have a long way to go before robots replace us: "While these robots are learning in a limited sense, it's a fairly big leap to claim they're 'thinking.'" Before these systems can genuinely think in a fluid, non-brittle fashion, a lot of things must happen. One of the important things I discuss in my book is our ability to develop and act on self-determined values in real-time, which we do hundreds of times each day. These systems will fail if conditions change outside of a predetermined domain without this."
5) *Specialized Hardware:* There is now a hardware development explosion underway to speed up and decrease the cost of neural networks, machine learning, and deep learning. Established businesses and start-ups are competing to build better processors and graphics processing units, but the main storey is the rapid development of neural network processing units (NNPUs) and other AI-specific hardware, dubbed neurosynaptic architectures. Because they work more like a biological brain than a typical computer's core, neurosynaptic chips are critical to the advancement of AI. IBM is a pioneer in the creation of neurosynaptic processors, thanks to its Brain Power technology. Unlike traditional chips, which run indefinitely, Brain Power's chips are event-driven and only run when they are needed. Memory, processing, and communication are all integrated into the technology.

XIV. CONCLUSIONS

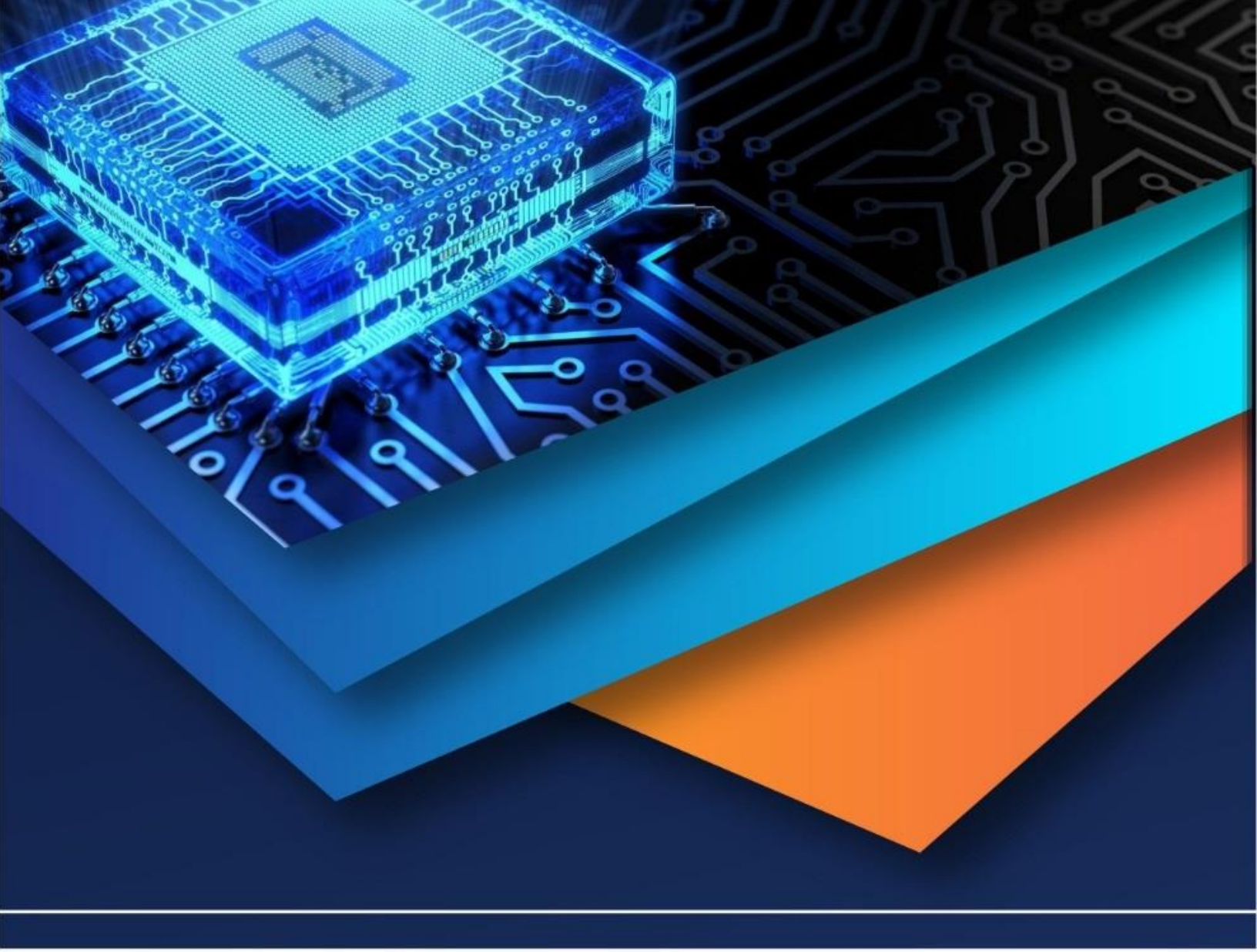
Neural networks is a broad topic. Many data scientists are primarily concerned with neural network approaches. We addressed the fundamental topics in this paper. Neural networks perform well on certain types of problems, such as image recognition. The neural network techniques are quite computationally intensive. They necessitate extremely efficient computer devices. There is a lot of fascinating research taking place right now in the field of neural networks. The artificial neural network (ANN) is a very helpful model that may be used in problem-solving and machine learning. The Neural Network has a lot to offer the computing world. As a result of their capacity to learn by example, they are extremely adaptable and strong. To make the greatest use of the ANN for various issues, it is necessary to grasp both the potential and the limitations of the Neural Network.

©IJRASET: All Rights are Reserved

REFERENCES

[1] Fausett, Laurene V. Fundamentals of neural networks: architectures, algorithms and applications. Pearson Education India, 2006.
[2] Nielsen, Michael, Neural Networks and Deep Learning: Perceptron, Architecture, Algorithms, 2019
[3] Yegnanarayana, Bayya. Artificial neural networks. PHI Learning Pvt. Ltd., 2009.
[4] Thakur, Amey, Hasan Rizvi, and Mega Satish. "White-Box Cartoonization Using An Extended GAN Framework." arXiv preprint arXiv:2107.04551 (2021)
[5] Thakur, Amey, Hasan Rizvi, and Mega Satish. "WHITE-BOX CARTOONIZATION USING AN EXTENDED GAN FRAMEWORK.", Volume 5, Issue 12, International Journal of Engineering Applied Sciences and Technology (IJEAST), Page No: 294-298, 2021.
[6] Hecht-Nielsen, R., Neurocomputing.:Why do we need Neural Networks? Addison-Wesley, Reading, MA, 1990.
[7] Schalkoff, R.J. Artificial Neural Networks:Why do we need Neural Networks?. McGraw-Hill, New York, 1997.
[8] Mehrotra, Kishan, Chilukuri K. Mohan, and Sanjay Ranka. Elements of artificial neural networks. MIT press, 1997.
[9] Graves, Alex. "Generating sequences with recurrent neural networks." arXiv preprint arXiv:1308.0850 (2013).
[10] Hassoun, Mohamad H. Fundamentals of artificial neural networks. MIT press, 1995.
[11] Hopfield, John J. "Neural networks and physical systems with emergent collective computational abilities." Proceedings of the national academy of sciences 79, no. 8 (1982): 2554-2558.
[12] Cheng, Bing, and D. Michael Titterington. "Neural networks: A review from a statistical perspective." Statistical science (1994): 2-30.
[13] Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom. "A convolutional neural network for modelling sentences." arXiv preprint arXiv:1404.2188 (2014).
[14] Lawrence, Steve, C. Lee Giles, Ah Chung Tsoi, and Andrew D. Back. "Face recognition: A convolutional neural-network approach." IEEE transactions on neural networks 8, no. 1 (1997): 98-113.
[15] Sherstinsky, Alex. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network." Physica D: Nonlinear Phenomena 404 (2020): 132306.
[16] Fine, Terrence L. Feedforward neural network methodology. Springer Science & Business Media, 2006.
[17] Le, Xuan Hien & Ho, Hung & Lee, Giha & Jung, Sungho. (2019). Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water. 11. 1387. 10.3390/w11071387.
[18] Vt, S. Elanayar, and Yung C. Shin. "Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems." IEEE transactions on neural networks 5, no. 4 (1994): 594-603.
[19] Chiu, Chung-Cheng, Tara N. Sainath, Yonghui Wu, Rohit Prabhavalkar, Patrick Nguyen, Zhifeng Chen, Anjuli Kannan et al. "State-of-the-art speech recognition with sequence-to-sequence models." In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4774-4778. IEEE, 2018.
[20] Pal, Sankar K., and Sushmita Mitra. "Multilayer perceptron, fuzzy sets, classification." (1992).
[21] Gardner, Matt W., and S. R. Dorling. "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences." Atmospheric environment 32, no. 14-15 (1998): 2627-2636.
[22] Happel, Bart LM, and Jacob MJ Murre. "Design and evolution of modular neural network architectures." Neural networks 7, no. 6-7 (1994): 985-1004.
[23] Lin, Chin-Teng, and CS George Lee. Neural fuzzy systems: a neuro-fuzzy synergism to intelligent systems. Prentice hall, 1996.
[24] Widrow, Bernard, David E. Rumelhart, and Michael A. Lehr. "Neural networks: applications in industry, business and science." Communications of the ACM 37, no. 3 (1994): 93-106.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)