Understanding Neural Networks: From Activation Function To Back Propagation

Farhad Malik
Nov 26, 2018 · 8 min read ★

This article aims to provide an overview of neural networks. It outlines fundamental concepts of following key areas:
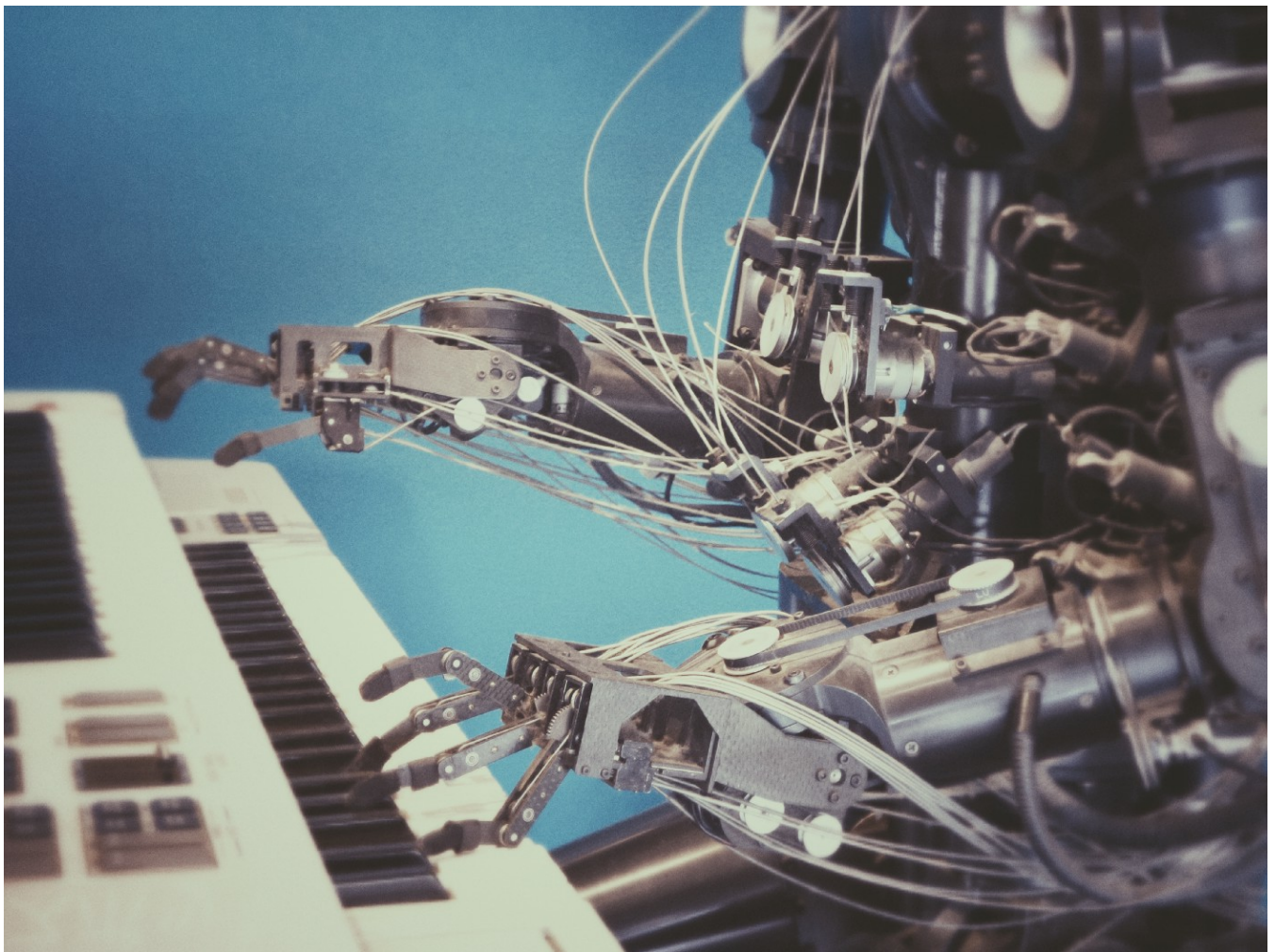
- What Are Neural Networks?

- What Are The Main Components Of Neural Networks?

- How Do Neural Networks Work?

- What Is An Activation Function?

- What Is Back Propagation?

- What Are The Different Types Of Neural Networks?

If you want to know how machine learning works in general then please read "Machine Learning In 8 Minutes" article.

Please Read FinTechExplained Disclaimer.

What Are Neural Networks?

Artificial Neural Networks (ANN) concept has been inspired by biological neural network.

First and foremost, neural network is a concept. It is not a machine or a physical box.

In a biologic neural network, multiple neurons work together, receive input signals, process information and fire an output signal.

## Biological Neuron Vs Artificial Neuron

Biological neurons are grouped in various layers and transmit updated signals. These signals contain information that can help us in determining patterns, identifying images, calculating numbers and making informed decisions throughout our life.

Neural Networks Learn From Past Experiences

Biological neural network is constantly learning and updating its knowledge and understanding of the environment based on experiences that it encountered.

Artificial intelligence (AI) neural network

Artificial intelligence (AI) neural network is based on the same biological neural network model.
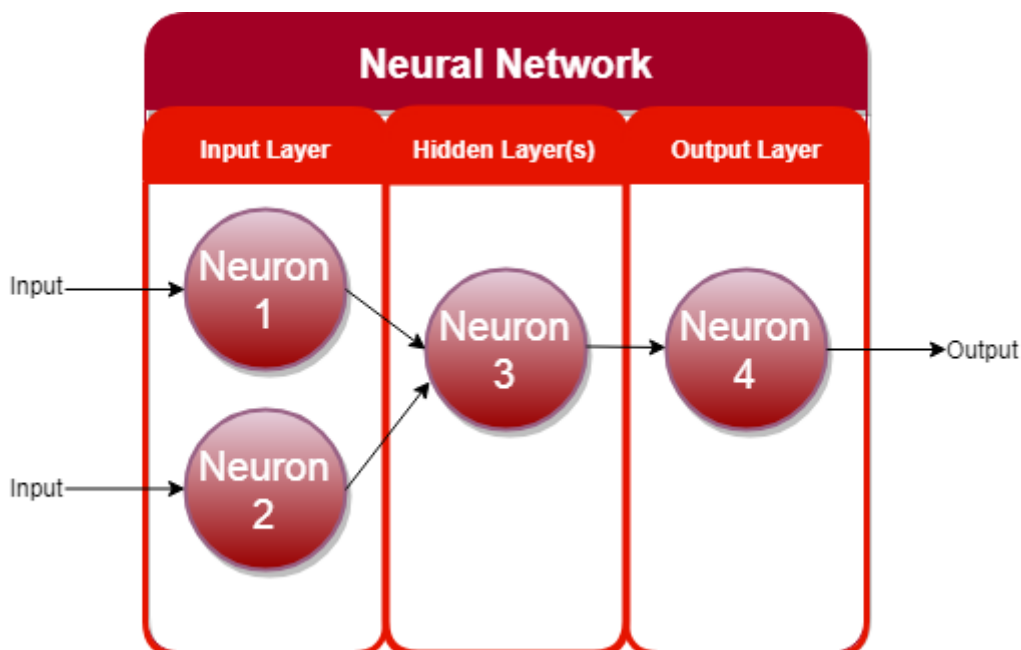
Although the underlying concept is the same as biological networks but think of AI neural network as a group of mathematical algorithms that produce output from the input data.

These algorithms can be packaged together to produce desired results.

In an artificial intelligence neural network, multiple algorithms work together to perform calculations on input data to compute an output. These outputs can also help neural network to learn and improve their accuracy.

Neural networks are trained with a range of inputs and their expected outputs. Neural networks then compute the output, compare it with the expected output and continuously update itself to improve the results if required.

Neural networks can learn themselves

Over time, the output is used to improve the accuracy of neural network model. Neural networks can help machines identify patterns, images and forecast time series data.

Textual information is usually encoded into numbers (binary) and each bit is passed to a single neuron.

What Are The Main Components Of Neural Network?

Neural network is composed of following main components:

Neurons: Set of functions

They take in an input and produce an output. A number of neurons are grouped into layers. All neurons within the same group perform similar type of function.

> To explain, input neurons receive inputs, process it and pass it to the neurons in next successive layer. Hidden neurons take outputs from input neurons, compute new outputs and pass them to successive layers.

In a 3 layer neural network, neurons in hidden layer pass outputs as inputs to the neurons in output layer. Output neurons take input from predecessor neurons and output results.

Layers: Grouping of neurons

Contain neurons and help pass information around. There are at minimum two layers in a neural network: Input and Output layer.

> We can have a large number of layers in a complex neural network.

The layers, other than the input and output layers, are known as hidden layers.

Weights & Biases: Numerical values

These are essentially variables in the model that are updated to improve network's accuracy. A weight is applied to input of each of the neuron to compute an output.

Neural networks update these weights on continuous basis. Hence there is a feedback loop implemented in most neural networks.

Biases are also numerical values which are added once weights are applied to inputs. Hence weights and biases make neural networks self-learning algorithms.

Think of weight as input importance of a neuron.

Activation Function: Mathematical algorithms applied to outputs

Essentially activation functions *smooth or normalise* the output before it is passed on to the next or previous neurons in the chain. These functions help neural networks learn and improve themselves.

Neural Networks Is A Machine Learning Concept Modeled On Biological Brain
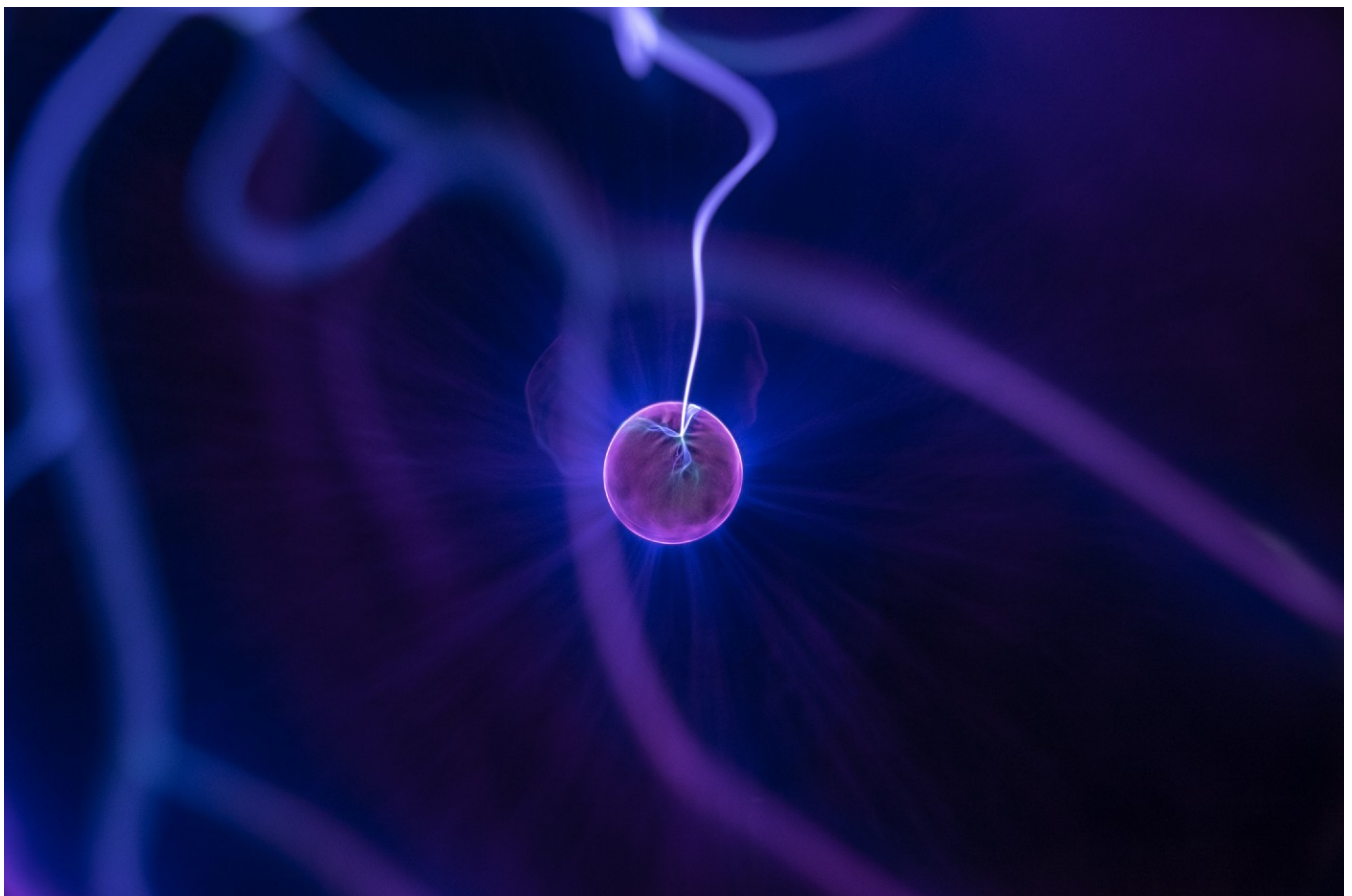
How Do Neural Networks Work?

The concept of neural network is based on three main steps:

1. For each neuron in a layer, multiply input to weight.

2. Then for each layer, sum all input x weights of neurons together.

3. Finally, apply activation function on the output to compute new output.
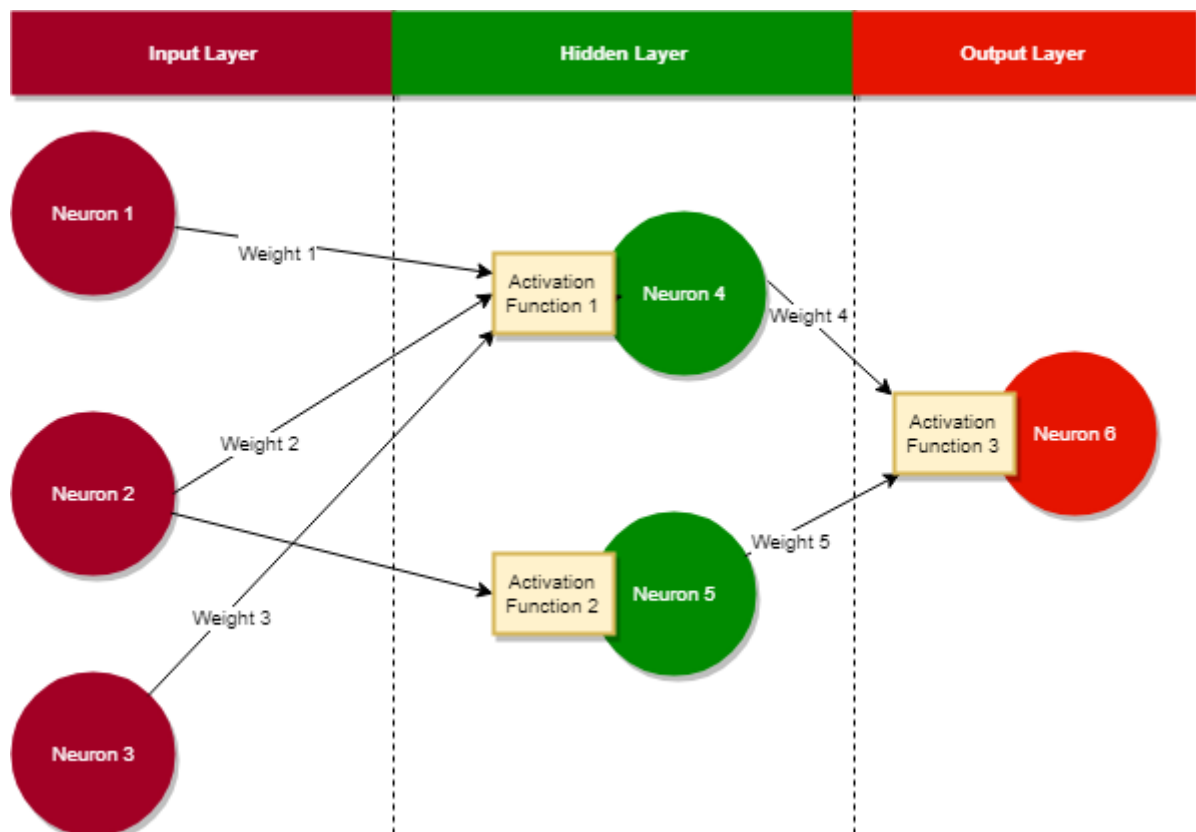
Remember the word: **S.IW.A**

*S: Sum Of, **IW**: Inputs x Weights, **A**: Apply Activation function*

Activation Function(Sum Of Inputs X Weights)

Understanding The Process

To elaborate, each neuron takes in an input as shown in the image below.

Inputs are fed into neuron 1, neuron 2 and neuron 3 as they belong to the Input Layer.

- Each neuron has a weight associated with it. When an input enters a neuron, the weight on the neuron is multiplied to the input. For instance, *weight 1* will be

applied to the input of *Neuron 1*. If *weight 1* is 0.8 and input is 1 then 0.8 will be computed from Neuron 1:

```
1 * 0.8 = 0.8
```

- Sum of weight * inputs of neurons in a layer is calculated. As an example, the calculated value on the hidden layer in the image will be:

```
(Weight 4 x Input To Neuron 4) + (Weight 5 x Input To Neuron 5)
```

- Finally an activation function is applied. Output calculated by the neurons becomes input to the activation function which then computes a new output. This output can flow back or to the neurons in the next layer.

Assume, the activation function is:

```
If (input > 1) Then 0 Else 1
```

The output from activation function is then fed to the subsequent layers.
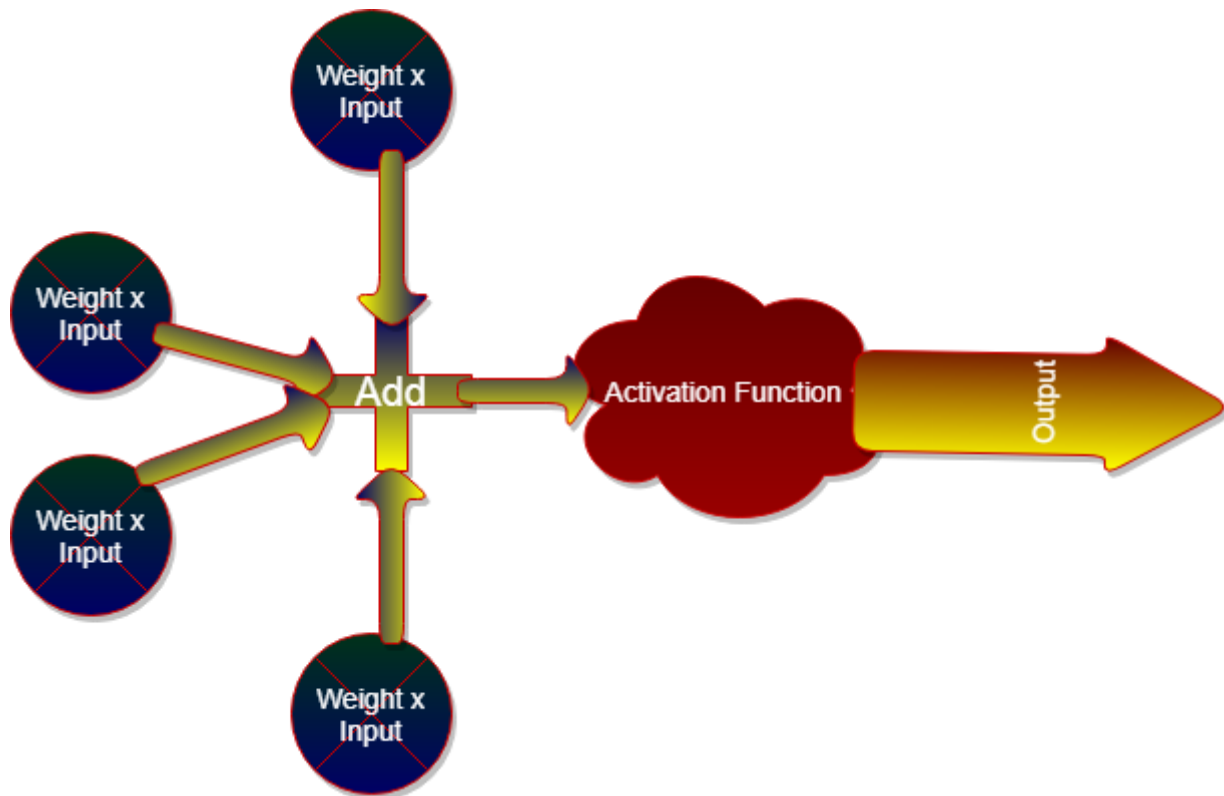
What Is An Activation Function?

As the name implies, activation function is a mathematical formula (algorithm) that is **activated** under certain circumstances. When neurons compute weighted sum of inputs, they are passed to the activation function which checks if the computed value is above the required threshold.

> If the computed value is above the required threshold then the activation function is activated and an output is computed.

This output is then passed on to the next or previous layers (dependent on the complexity of the network) which can help neural networks alter weights on their

neurons.

Activation functions introduce non-linearity to the neural networks which is required to solve complex problems.

> If we plot non-linear outputs that the activation functions produce, we will get a curvature. The slope of the curve is used to compute the gradient. Gradient helps us in understanding rate of change and relationships of variables.

From the relationships, algorithms are optimised and weights are updated.

Types Of Activation Functions

There are a large number of activation functions, such as:

- **Sigmoid**: 1/1 + exp(x) which produces a S-shaped curve. Although it is non-linear in nature but it does not capture slight changes within inputs hence variations in inputs will yield similar results.

- **Hyperbolic Tangent Functions (Tanh)**: (1- (exp(-2(x))/(1 + exp(-2x)). It is a superior function when compared to Sigmoid. However it does not capture relationships better and is slower at converging.

- **Rectified Linear Units (ReLu)**: This function converges faster, optimises and produces the objective value quicker. It is by far the most popular activation function used within the hidden layers.

- **Softmax**: Used in output layer because it reduces dimensions and can represent categorical distribution.

What Is Back Propagation?

Back propagation concept helps neural networks to improve their accuracy.

In traditional software application, a number of functions are coded. These functions take in inputs and produce an output. The inputs are not used to update the instructions.

However neural networks are artificially intelligent. They can learn and improve themselves.

When neural networks are trained, a range of inputs are passed along with corresponding expected output. Activation functions then produce an output from the set of inputs.

Back Propagation: Helps Neural Network Learn

When the actual result is different than the expected result then the weights applied to neurons are updated. Sometimes the expected and actual results are within the error threshold and neural network is considered optimal.

However occasionally, the expected output is different than the actual output. As a consequence, information is fed back into the network and the weights and biases are enhanced. This process is recursive in nature and is known as back propagation.

Back Propagation Process Makes Algorithms Self-Learning

Thus, back propagation makes neural networks intelligent and self-improving.
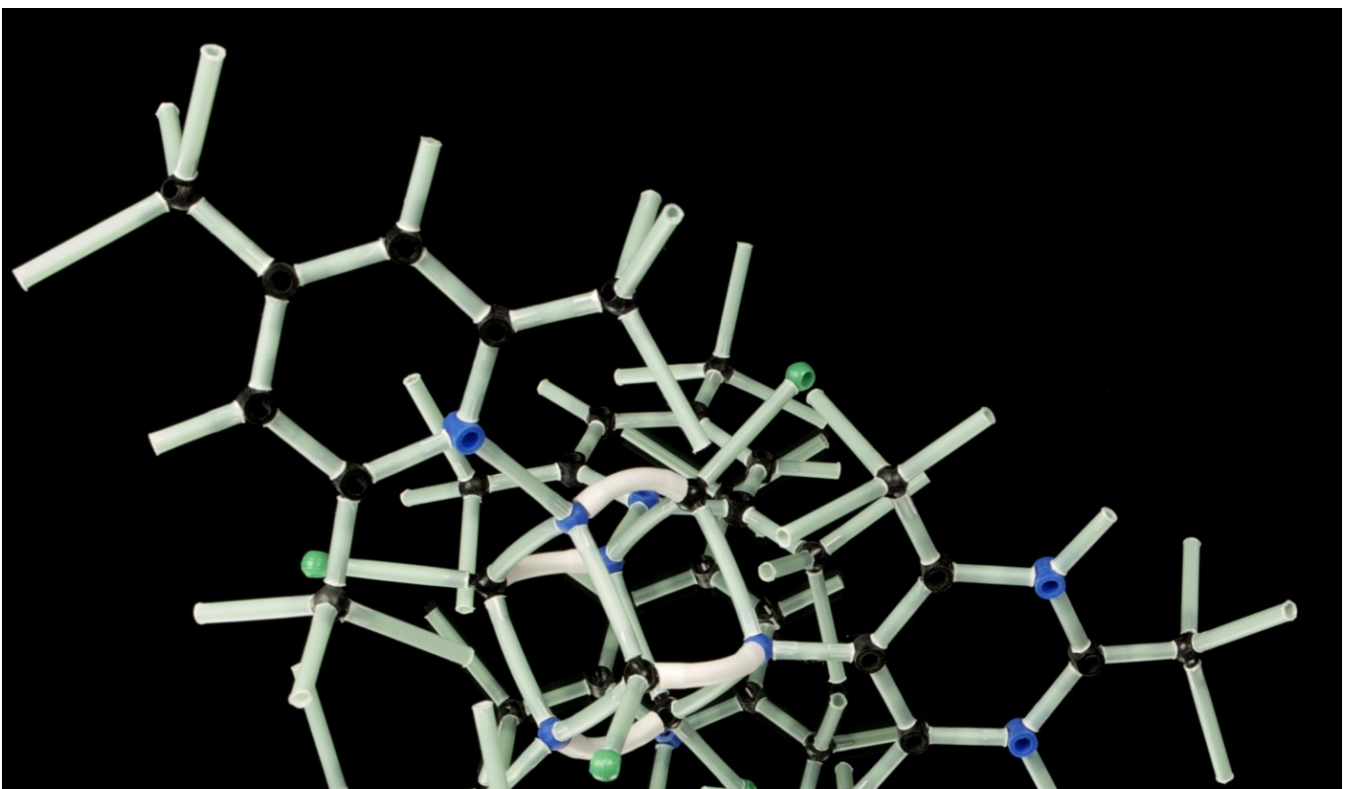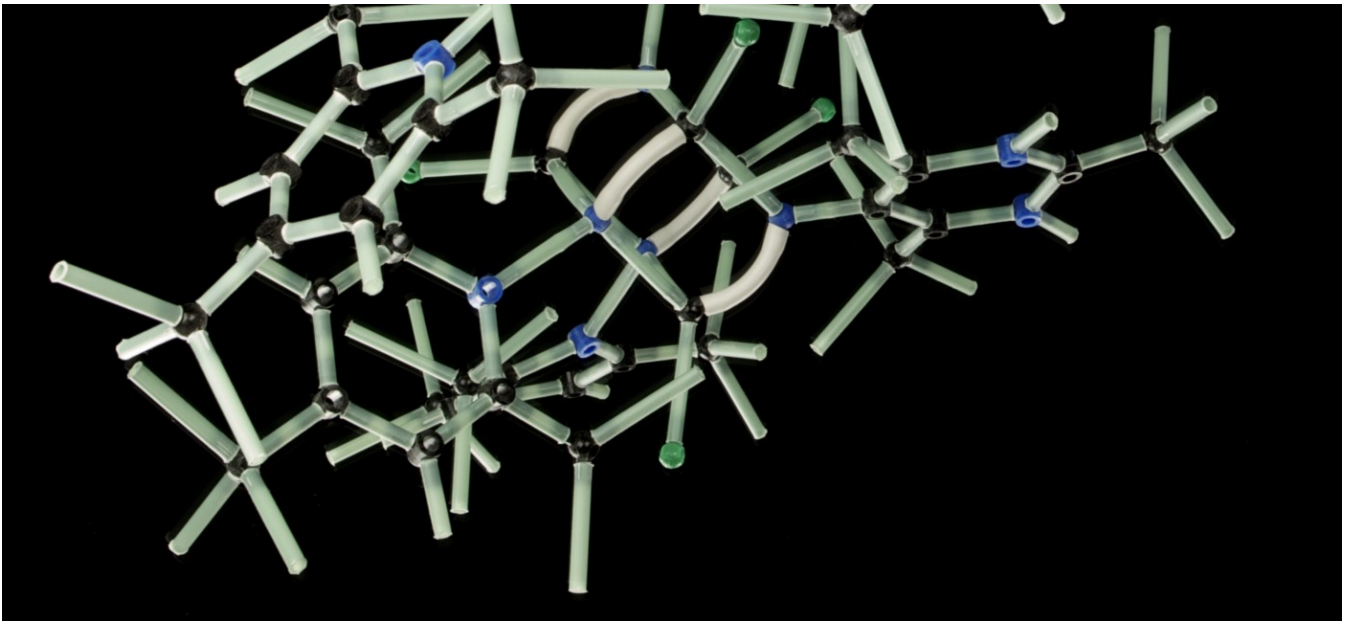
What Are The Different Types Of Neural Networks?

There are different types of neural networks. Two of the most popular neural networks are:

1. **Recurrent Neural Network (RNN):**

These are specialised neural networks that use the context of the inputs when computing the output. Output is dependent on the inputs and the previously computed outputs.

RNN can work with varying lengths of input and outputs and require a large quantity of data.

Thus RNN are suitable for applications where historic information is important. These networks help us in predicting time series in trading applications and forecasting words in chat-bot applications.

2. Convolution Neural Network (CNN):

These networks rely on convolution filters (numerical matrices). Filters are applied to the inputs before the inputs are passed to the neurons.

CNNs are useful in image processing and forecasting.

Summary

This article introduced readers to the concept of neural networks.

Moreover, it provided an overview of components that make neural networks artificially intelligent.

Lastly, the article introduced two famous types of neural networks.

If you want to know how machine learning works in general then please read "Machine Learning In 8 Minutes" article.

Please let me know if you have any feedback.

Machine Learning   Data Science   Neural Networks   Fintech   Deep Learning

About   Help   Legal

Get the Medium app