8 types of neural network models explained

There are many different types of artificial neural networks, varying in complexity. They share the intended goal of mirroring the function of the human brain to solve complex problems or tasks. The structure of each type of artificial neural network in some way mirrors neurons and synapses. However, they differ in terms of complexity, use cases, and structure. Differences also include how artificial neurons are modelled within each type of artificial neural network, and the connections between each node. Other differences include how the data may flow through the artificial neural network, and the density of the nodes.

5 examples of the different types of artificial neural network include:

- Feedforward artificial neural networks
- Perceptron and Multilayer Perceptron neural networks
- Radial basis function artificial neural networks
- Recurrent neural networks
- Modular neural networks
- Long Short-Term Memory Networks
- Convolutional Neural Networks
- Generative Advesarial Networks

Feedforward artificial neural networks

As the name suggests, a Feedforward artificial neural network is when data moves in one direction between the input and output nodes. Data moves forward through layers of nodes, and won't cycle backwards through the same layers. Although there may be many different layers with many different nodes, the one-way movement of data makes Feedforward neural networks relatively simple. Feedforward artificial neural network models are mainly used for simplistic classification problems. Models will perform beyond the scope of a traditional machine learning model, but don't meet the level of abstraction found in a deep learning model.

Perceptron and Multilayer Perceptron neural networks

A perceptron is one of the earliest and simplest models of a neuron. A Perceptron model is a binary classifier, separating data into two different classifications. As a linear model it is one of the simplest examples of a type of artificial neural network.

Multilayer Perceptron artificial neural networks adds complexity and density, with the capacity for many hidden layers between the input and output layer. Each individual node on a specific layer is connected to every node on the next layer. This means Multilayer Perceptron models are fully connected networks, and can be leveraged for deep learning.

They're used for more complex problems and tasks such as complex classification or voice recognition. Because of the model's depth and complexity, processing and model maintenance can be resource and time-consuming.

Radial basis function artificial neural networks

Radial basis function neural networks usually have an input layer, a layer with radial basis function nodes with different parameters, and an output layer. Models can be used to perform classification, regression for time series, and to control systems. Radial basis functions calculate the absolute value between a centre point and a given point. In the case of classification, a radial basis function calculates the distance between an input and a learned classification. If the input is closest to a specific tag, it is classified as such.

A common use for radial basis function neural networks is in system control, such as systems that control power restoration after a power cut. The artificial neural network can understand the priority order to restoring power, prioritising repairs to the greatest number of people or core services.

Recurrent neural networks

Recurrent neural networks are powerful tools when a model is designed to process sequential data. The model will move data forward and loop it backwards to previous steps in the artificial neural network to best achieve a task and improve predictions. The layers between the input and output layers are recurrent, in that relevant information is looped back and retained. Memory of outputs from a layer is looped back to the input where it is held to improve the process for the next input.

The flow of data is similar to Feedforward artificial neural networks, but each node will retain information needed to improve each step. Because of this, models can better understand the context of an input and refine the prediction of an output. For example, a predictive text system may use memory of a previous word in a string of words to better predict the outcome of the next word. A recurrent artificial neural network would be better suited to understand the sentiment behind a whole sentence compared to more traditional machine learning models.

Recurrent neural networks are also used within sequence to sequence models, which are used for natural language processing. Two recurrent neural networks are used within these models, which consists of a simultaneous encoder and decoder. These models are used for reactive chatbots, translating language, or to summarise documents.

Modular neural networks

A Modular artificial neural network consists of a series of networks or components that work together (though independently) to achieve a task. A complex task can therefore be broken down into smaller components. If applied to data processing or the computing process, the speed of the processing will be increased as smaller components can work in tandem.

Each component network is performing a different subtask which when combined completes the overall tasks and output. This type of artificial neural network is beneficial as it can make complex processes more efficient, and can be applied to a range of environments.

Long Short-Term Memory Networks

LSTM neural networks overcome the issue of Vanishing Gradient in RNNs by adding a special memory cell that can store information for long periods of time. LSTM uses gates to define which output should be used or forgotten. It uses 3 gates: Input gate, Output gate and a Forget gate. The Input gate controls what all data should be kept in memory. The Output gate controls the data given to the next layer and the forget gate controls when to dump/forget the data not required.

Convolutional Neural Networks

When it comes to image classification, the most used neural networks are Convolution Neural Networks (CNN). CNN contain multiple convolution layers which are responsible for the extraction of important features from the image. The earlier layers are responsible for low-level details and the later layers are responsible for more high-level features.

The Convolution operation uses a custom matrix, also called as filters, to convolute over the input image and produce maps. These filters are initialized randomly and then are updated via backpropagation. One example of such a filter is the Canny Edge Detector, which is used to find the edges in any image.

After the convolution layer, there is a pooling layer which is responsible for the aggregation of the maps produced from the convolutional layer. It can be Max Pooling, Min Pooling, etc. For regularization, CNNs also include an option for adding dropout layers which drop or make certain neurons inactive to reduce overfitting and quicker convergence.

CNNs use ReLU (Rectified Linear Unit) as activation functions in the hidden layers. As the last layer, the CNNs have a fully connected dense layer and the activation function mostly as Softmax for classification, and mostly ReLU for regression.

Generative Advesarial Networks

Given training data, Generative Adversarial Networks (or simply, GANs) learn to generate new data with the same statistics as the training data. For example, if we train a GAN model on photographs, then a trained model will be able to generate new photographs that look similar to the input photographs.

A GAN contains two parts: a generator and a discriminator. The generator model creates new data while the discriminator tries to determine real data from generated data. As the generator and discriminator get better at their respective jobs, the generated data improves as a result, until it is (ideally) nearly identical in quality to the training data.