

This blog provides an introduction to the basic concepts of neural networks and deep learning. At the very basics, Artificial Neural Networks (ANNs) or simply "neural networks" are computer models which are inspired by the structure of the human brain. The components of a neural network involve nodes called 'neurons' that are connected in some manner and are bifurcated into layers (1 input layer, any number of hidden layers, and an output layer). A neural network is a series of interconnected nodes that create a multilayered system: the input layer is used to input data, the hidden layers are used for processing, and the output layer is the final output or result. Deep learning refers to a neural network that has hidden layers in multiple levels of depth, meaning they can learn to represent more complex patterns in the data.

How it works is each neuron accepts inputs, which are multiplied by their associated "weights" to signify importance. The weighted inputs are summed, and a 'bias' is added to the total before the sum value is inputted into an activation function. The activation function is significant in this process, as it is what adds non-linearity to the input data and therefore the ANN can learn from a complex data set, rather than a simple linear function.

The document nicely represents four key activation functions:

Threshold functions - are essentially an off switch

Sigmoid functions - which produce a smooth output from 0-1, can be very useful if there are prediction probabilities.

ReLU (Rectified Linear Unit) - a timely, efficient function that has become the most popular of many functions.

The Hyperbolic Tangent (tanh) - a similar function to sigmoid, but which maps output from -1 to 1.

Put simply, a neural network learns by means of trial and error because it minimizes prediction errors by adjusting its weights and biases. The incredible ability of these systems to leverage that structure, layered connections, and non-linear activation functions is what enables incredible feats in topics like image recognition and natural language processing.