

What are neural networks?

It is a process used in machine learning solve artificial intelligence problems. It involves an input and an output layer, with one or more layers in between. There are multiple nodes on each layer, and data is categorised into dimensions in each node. Each node is assigned a numbered weight which connects them to the next layer of nodes, which are each assigned with a bias. The weights and biases are numbers assigned to them, which are then summed to produce a value as the output, thus categorises them based on their value (usually between 0 and 1). In machine learning, the weights and biases are changed so that the resulting value will be the correct output. The input and output numbers are range between 0 and 1, and a threshold value is defined as a cutoff level when making the conclusions. The accuracy of the output can be changed based on the set learning rate and the momentum. The former relates to the number of predictors set onto the models, and the momentum relates to the training process to ensure the output values are correct.

Historically, there only 2 or 3 layers used to compute simple data, which is why it did not take off — it was too time consuming to compute simple data using neural networks. However, the resurgence of neural networks may be credited to their use in the development of video games. The modern use of neural networks involve 10-50 layers which allow for the solving of more complex problems. Examples of the use of neural networks include image recognition, speech recognition, or risk management.

<https://www.bmc.com/blogs/neural-network-introduction/>

