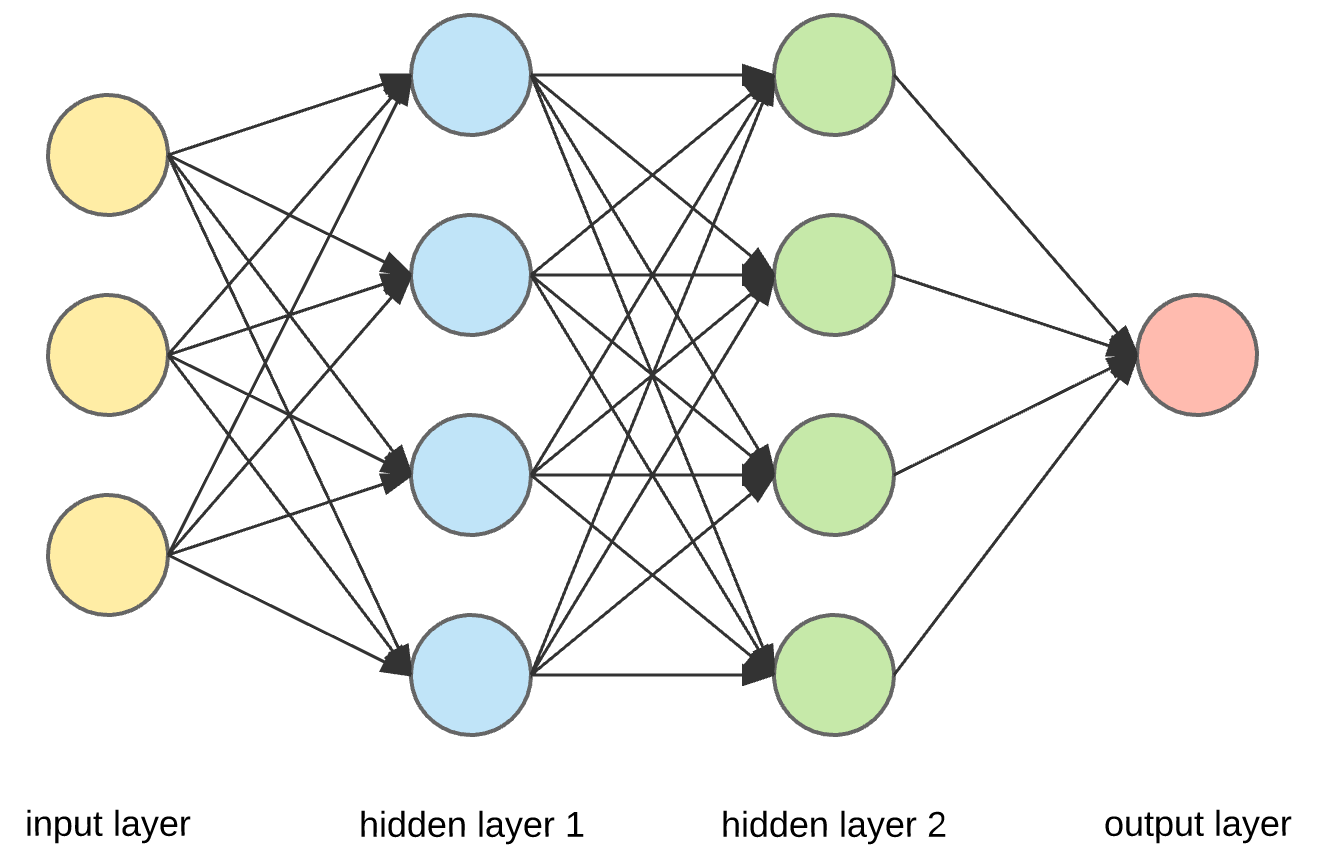
What is a neural network?

A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain. It creates an adaptive system that computers use to learn from their mistakes and improve continuously. Thus, artificial neural networks attempt to solve complicated problems, like summarizing documents or recognizing faces, with greater accuracy.



Why are neural networks important?

Neural networks can help computers make intelligent decisions with limited human assistance. This is because they can learn and model the relationships between input and output data that are nonlinear and complex. For instance, they can do the following tasks.

**Make generalizations and inferences**

Neural networks can comprehend unstructured data and make general observations without explicit training. For instance, they can recognize that two different input sentences have a similar meaning:

* Can you tell me how to make the payment?
* How do I transfer money?

A neural network would know that both sentences mean the same thing. Or it would be able to broadly recognize that Baxter Road is a place, but Baxter Smith is a person’s name.

What are neural networks used for?

Neural networks have several use cases across many industries, such as the following:

* Medical diagnosis by medical image classification
* Targeted marketing by social network filtering and behavioral data analysis
* Chemical compound identification

We give four of the important applications of neural networks below.

**Computer vision**

**Speech recognition**

**Natural language processing**

**Recommendation engines**

How do neural networks work?

The human brain is the inspiration behind neural network architecture. Human brain cells, called neurons, form a complex, highly interconnected network and send electrical signals to each other to help humans process information. Similarly, an artificial neural network is made of artificial neurons that work together to solve a problem. Artificial neurons are software modules, called nodes, and artificial neural networks are software programs or algorithms that, at their core, use computing systems to solve mathematical calculations.

**Simple neural network architecture**

A basic neural network has interconnected artificial neurons in three layers:

**Input Layer**

Information from the outside world enters the artificial neural network from the input layer. Input nodes process the data, analyze or categorize it, and pass it on to the next layer.

**Hidden Layer**

Hidden layers take their input from the input layer or other hidden layers. Artificial neural networks can have a large number of hidden layers. Each hidden layer analyzes the output from the previous layer, processes it further, and passes it on to the next layer.

**Output Layer**

The output layer gives the final result of all the data processing by the artificial neural network. It can have single or multiple nodes. For instance, if we have a binary (yes/no) classification problem, the output layer will have one output node, which will give the result as 1 or 0. However, if we have a multi-class classification problem, the output layer might consist of more than one output node.

What are the types of neural networks?

**Feedforward neural networks**

**Backpropagation algorithm**

**Convolutional neural networks**

**Machine learning vs. deep learning**

Traditional machine learning methods require human input for the machine learning software to work sufficiently well. A data scientist manually determines the set of relevant features that the software must analyze. This limits the software’s ability, which makes it tedious to create and manage.

On the other hand, in deep learning, the data scientist gives only raw data to the software. The deep learning network derives the features by itself and learns more independently. It can analyze unstructured datasets like text documents, identify which data attributes to prioritize, and solve more complex problems.

For example, if you were training a machine learning software to identify an image of a pet correctly, you would need to take these steps:

* Find and label thousands of pet images, like cats, dogs, horses, hamsters, parrots, and so on, manually.
* Tell the machine learning software what features to look for so it can identify the image using elimination. For instance, it might count the number of legs, then check for eye shape, ear shape, tail, fur, and so on.
* Manually assess and change the labeled datasets to improve the software’s accuracy. For example, if your training set has too many pictures of black cats, the software will correctly identify a black cat but not a white one.
* In deep learning, however, the neural networks would process all the images and automatically determine that they need to analyze the number of legs and the face shape first, then look at the tails last to correctly identify the animal in the image.