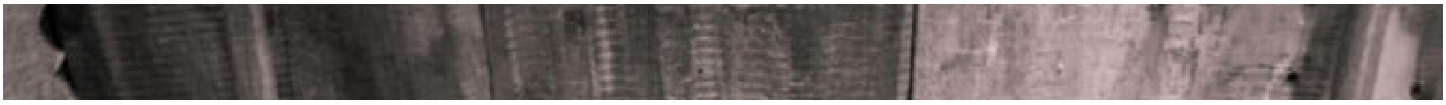

Doug Rose

The Hidden Layers of a Neural Network

Click here

Doug Rose

Author | Agility | Artificial Intelligence | Data Ethics

+ Follow

Published Jul 29, 2021

In my previous post Artificial Neural Networks, I explain what an artificial neural network (or simply a neural network) is and what it does. I also point out that what enables a neural network to perform its magic is the layering of neurons. A neural network consists of three layers of neurons — an input layer, one or more hidden layers, and an output layer.

The input and output layers are self-explanatory. The input layer receives data from the outside world and passes it to the hidden layer(s) for processing. The output layer receives the processed data from the hidden layer(s) and coveys it in some way to the outside world.

Yet, what goes on in the hidden layer(s) is more mysterious.

The Purpose of the Hidden Layer

Early neural networks lacked a hidden layer. As a result, they were able to solve only linear problems. For example, suppose you needed a neural network to distinguish cats from dogs. A neural network without a hidden layer could perform this task. It could create a linear model like the one shown below and classify all input that characterizes a cat on one side of the line and all input that characterizes a dog on the other.

However, if you had a more complex problem, such as distinguishing different breeds of dogs, this linear neural network would fail the test. You would need several layers to examine the various characteristics of each breed.

What Goes on in the Hidden Layers?

Suppose you have a neural network that can identify a dog's breed simply by "looking" at a picture of a dog. A neural network capable of learning to perform this task could be structured in many different ways, but consider the following (admittedly oversimplified) example of a neural network with several layers, each containing 20 neurons (or nodes).

When you feed a picture of a dog into this fictional neural network, the input layer creates a map of the pixels that comprise the image, recording their positions and grayscale values (zero for black, one for white, and between zero and one for different shades of gray). It then passes this map along to the 20 neurons that comprise the first hidden layer.

The 20 neurons in the first hidden layer look for patterns in the map that identify certain features. One neuron may identify the size of the dog; another, its overall shape; another, its eyes; another, its ears; another, its tail; and so forth. The first hidden layer then passes its results along to the 20 neurons in the second hidden layer.

The neurons in the second hidden layer are responsible for associating the patterns found in the first layer with features of the different breeds. The neurons in this layer may assign a percentage to reflect the probability that a certain feature in the image corresponds to different breeds. For example, based solely on the ears in the image, the breed is 20% likely to be a Doberman, 30% likely to be a poodle, and 50% likely to be a Labrador retriever. The second hidden layer passes its results along to the third hidden layer.

The neurons in the third hidden layer compile and analyze the results from the second hidden layer and, based on the collective probabilities of the dog being a certain breed, determine what that breed is most likely to be. This final determination is then delivered to the output layer, which presents the neural network's determination.

The Well-Connected Neurons

While the example I presented focuses on the layers of the neural network and the neurons (nodes) that comprise those layers, the connections between the neurons play a very important role in how the neural network learns and performs its task.

Every neuron in one layer is connected to every neuron in its neighboring layer. In the example I presented, that's 400 connections between each layer. The strength of each connection can be dialed up or down to change the relative importance of input from one neuron to another. For example, each neuron in the first hidden layer can dial up or down its connection with each neuron in the input layer to determine what it needs to focus on in the image, just as you might focus on different parts of an image.

When the neural network is being trained with a set of test data, it is given the answers — it is shown a picture of each breed and labeled with the name of the breed. During this training session, the neural network makes adjustments within the nodes and between the nodes (the connections). As the neural network is fed more and more images of dogs, it fine-tunes its connections and makes other adjustments to improve its accuracy over time.

Again, this example is oversimplified, but it gives you a general idea of how artificial neural networks operate. The key points to keep in mind are that artificial neural networks contain far more connections than they contain neurons, and that they learn by making adjustments within and between neurons.

More articles by this author

See all

How to Overcome Key Challenges with Artificial Intelligence

How to Overcome Key Challenges with…
Aug 9, 2021

How Organizations Can Approach Artificial Intelligence

How Organizations Can Approach Artificial…
Aug 9, 2021

Solving Clustering Problem with a Neural Network

Solving Clustering Problems with a Neur…
Aug 9, 2021