Artificial Neural Network

Revanth reddy Bandaru Date:05-14-21

Table of Contents:

Executive Summary:

A neural network is a data processing system consisting of a large number of simple, highly interconnected processing elements in an architecture inspired by the structure of the cerebral cortex portion of the brain. Hence, neural networks are often capable of doing things which humans or animals do well but which conventional computers often do poorly. Neural networks have emerged in the past few years as an area of unusual opportunity for research, development and application to a variety of real-world problems. Indeed, neural networks exhibit characteristics and capabilities not provided by any other technology. Examples include reading Japanese Kanji characters and human handwriting, reading typewritten text, compensating for alignment errors in robots, interpreting very "noisy" signals (e.g. electrocardiograms), modeling complex systems that cannot be modelled mathematically, and predicting whether proposed loans will be good or fail. This report presents a brief tutorial on neural networks and finding a best artificial neural network (ANN) model for the given input and output.

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. It is the foundation of artificial intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards.

ANNs (Artificial Neural Network) is at the very core of Deep Learning an advanced version of Machine Learning techniques. ANNs are versatile, adaptive, and scalable, making them appropriate to tackle large datasets and highly complex Machine Learning tasks such as image classification (e.g., Google Images), speech recognition (e.g., Apple's Siri), video recommendation (e.g., YouTube), or analyzing sentiments among customers (e.g. Twitter Sentiment Analyzer)

Problem definition:

Artificial Neural Network is a branch of Artificial Intelligence that adopts the workings of the human brain in processing a combination of stimuli into an output.

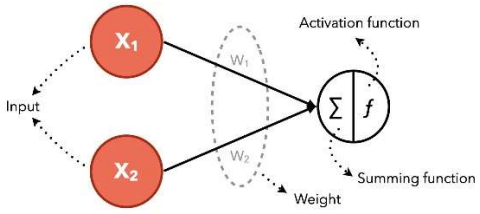
An important part of ANN is Neurons. Like the human brain consisting of many brain cells, ANN also consists of a collection of neurons that are interconnected. Inside the neuron, there is the stimulus receiver, stimulus processor, and the output part. The output produces information that can later be transmitted to other neurons.

ANN is very popular because it can solve complex problems that are difficult to be solved by other methods. This method also became the forerunner to the development of Deep Learning, which is one of the sub-fields of Machine Learning that focuses on Artificial Neural Networks that have multiple and deep layer structures.

Component of ANN:

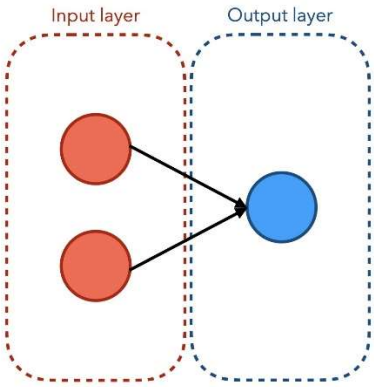
As explained in the previous section, ANN has one part called neurons. In general, ANN has three neurons, namely input neuron, neuron hidden, and output neuron.

In hidden neuron and output neurons, there is a function used to generate output from the previous neuron. This function is called the activation function. The output of neurons can be passed on which will later become an input for further neurons.

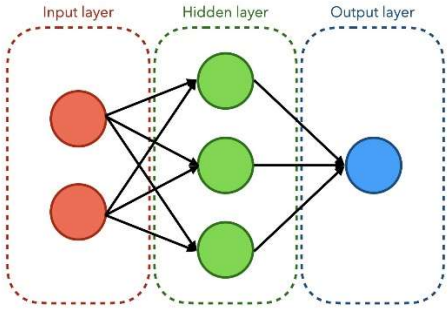


The output of ANN is a value that represents the prediction of a variable being sought. The next important part of ANN is the weight. Weights are used as input multipliers that will enter a neuron. Each neuron that is connected to each other will have their own weights. Each neuron can be added biased, which is a constant number added to the neuron input.

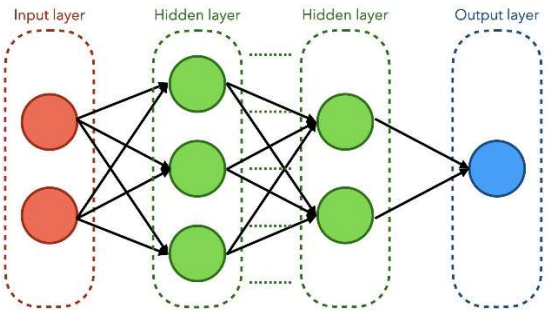
In addition, there is also a layer. A layer consists of one or more neurons. There are three layers in ANN, namely the input layer, hidden layer, and output layer. If the structure only has input and output layers, it is called the single-layer perceptron.



Meanwhile, if it has at least 1 layer (hidden layer) between the input and output layers, then it is called Multilayer Perceptron.



If the Multilayer Perceptron has many hidden layers, it is called a Deep Neural Network. In this article we only focused on Multilayer Perceptron (1 hidden layer).



Working:

As explained earlier, ANN can make predictions and classifications by studying the data patterns that have been provided. The more data provided, the more patterns can be learned by this

method. In general, there are two main processes in the Neural Network that are commonly used, namely feedforward and backpropagation.

Feedforward is an algorithm for calculating output values based on input values, while backpropagation is an algorithm for training neural networks (changing weights) based on errors obtained from output values.

The learning process in a neural network begins with initiating the value of the weight, learning rate, and epoch. We also have to determine the number of input neurons, hidden neurons, and output neurons. Epoch is an iteration in ANN. We can determine the number of iterations in the learning process, if it reaches a certain iteration, the learning process will stop. We can also determine the minimum error value to stop the learning process.

Learning rate is a value to determine the speed of learning, the greater the value of learning rate, the faster the machine is learning, but it will be difficult to get optimum results, conversely, the smaller the learning rate the slower the machine is learning, but the greater the possibility of getting optimum results.

Feed Forward:

The feedforward process is focused on finding the output of the system that has been built with predetermined parameters. The results of these outputs will be compared with predetermined targets. Let's look at the following illustration.

The illustration above is an example of using a single layer perceptron. x1 and x2 are the input data we use, while w1 and w2 are weighting. Usually, the weight value is determined randomly [-1…1]. The process for calculating the output is as follows.

1. Multiply the weights by the input of the neurons then add them up.
2. Enter the results of the multiplication of weights by the input into the activation function

For example, we have the values of x1 and x2 respectively 1 and 0, and the values of w1 and w2 are 0.25 and 0.75, respectively. Then the calculation is like this x1 * w1 + x2 * w2 = 1 * 0.25 + 0 * 0.75 = 0.25. Then, the value is entered into the activation function. The activation function that we will use here is the sigmoid function. By using the sigmoid function the result is 0.56.

$$\sigma = \frac{1}{1+ e^{-x}}$$

So, what if the structure consists of 3 layers, namely the input layer, hidden layer, and output layer? Basically, the same. The value of the activation function above will be the value for the hidden layer which will be forwarded to the output layer.

After we get the value for the output layer, the next process is to get the error value from the output layer. This error value will later become a parameter to change the weight value. To calculate errors there are also many methods such as MSE (Mean Squared Error), SSE (Sum of Squared Error), etc.

Backpropagation:

Backpropagation focuses on updating and evaluating the parameters (weights) used in feedforward. The weight value will be changed according to the error value obtained. Error value obtained from Target value — output value.

Weight renewal usually uses an optimization algorithm. The optimization algorithm used in ANN varies, such as Adam (Adaptive Moment Estimation), RMSProp, Gradient Descent, etc.

After we update the previous weight, we have completed one iteration. The next step is to enter the second iteration by repeating the feedforward process. This process will continue to repeat until the stop conditions that we specify before.

After the training process is deemed sufficient, the ANN model can be used for the classification process. The way to do classification is similar to the feedforward process where the data entered will be processed on neurons in ANN by means of weight multiplication. the difference in weight multiplication here with feedforward is, the weight used is the best weight obtained in the training process. When the process is in the output layer, the output obtained is a decimal number for each neuron which will be transformed into a predictive class label from the system.

Analysis approach:

To implement the same problem space using a neural network, we need to create a neuron based structure. Before jumping into the architecture, let's take a look at some of the components of a Neural Network.

1. **The Input Layer** — Represents the input variables plus the bias term. Hence if there are n input variables, the size of the input layer is n + 1, where + 1 is the bias term
2. **The Hidden Layer/ Layers** — These signify neurons where all mathematical calculations are done. Note a given neural network can have more than one neuron in a hidden layer or multiple hidden layers as well
3. **The Activation Function** — Converts the output of a given layer before passing on the information to consecutive layers. **Activation functions** are mathematical equations that determine the output of a given **neural network**. The is a part of each **neuron** in the **hidden layers** and determines output relevant for prediction
4. **The Output Layer** — The final "output prediction" of the network
5. **Forward Propagation** — Calculating the output of each iteration from the input layer to the output layer
6. **Backward Propagation** — Calculates revised weights (w1, w2, w3, and b1) after each forward propagation by analyzing the derivative of the cost function used to optimize the model output
7. **Learning Rate** — Determined the percentage change attributed to each weight and bias term after every backward propagation, i.e. controls the speed at which the model learns information about the data

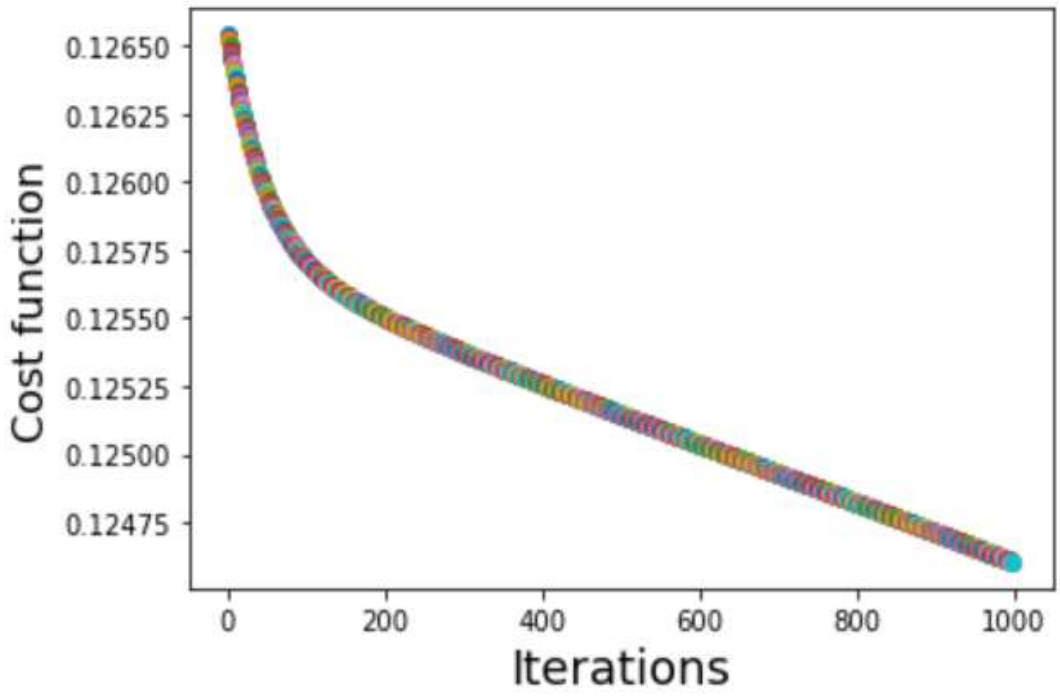
The idea of an ANN algorithm is to determine the optimum values of weights and bias terms such that the error function is as minimum as possible. Hence

- **Step 1** — Weights w1, w2, w3, and b1 are initialized with some random values
- **Step 2** — The hidden layers use the given weights and input variables to calculate the probability score. Activation functions (e.g. sigmoid) are applied if applicable
- **Step 3** — The cost function is put to use and the gradient descent algorithm repeats to calculate the updated weights. Updated weights are determined as weight (w) = w-alpha*(dJ/dw), where alpha is the learning rate and dJ/dw is the change in cost function when weight changes. Similarly bias (b) = b-alpha*(dJ/db)
- **Step 4** — Steps 2 and 3 are repeated till the cost function minimizes

Results:

To predict the ANN model for the given input X and output y, I am using the Feed Forward, back propagation and update parameter methods in python with the following parameters linear parameter is 0, learning rate is 0.1 and no of iterations as 1000 and no of hidden layers are 2. The dimensions for the first hidden layer is 3 and the dimensions for the second hidden layer is 2. The values X, y, W1, W2, b1, b2 as given as inputs where W1 and W2 are the weights of the hidden layers and b1 and b2 are bias, these are added to the hidden layers in each loop.

For every iteration, the methods Feed Forward, back propagation and update parameters are done simultaneously. The inputs are given to Feed Forward where the cost function is calculated. The Cost Function is defined as the difference between the estimated output and output obtained. The output from the Feed Forward is given to back propagation method and the output is given to update parameters, where the parameters are updated and given as input for the next iteration, the Iterations will continue unit a constant cost function is obtained. When the cost function is minimum and constant, the model is said to best fit model and the parameters are called Artificial neural networks. The above-mentioned process is done in python using NumPy and matplotlib libraries and plotted a graph between iterations versus cost function and obtained the minimum cost function using the gradient descent method.



Conclusion:

From the given data, the best fit ANN model is found and from the above result cost function is reduced to 0.1246. By this we can conclude that the desired output is obtained.