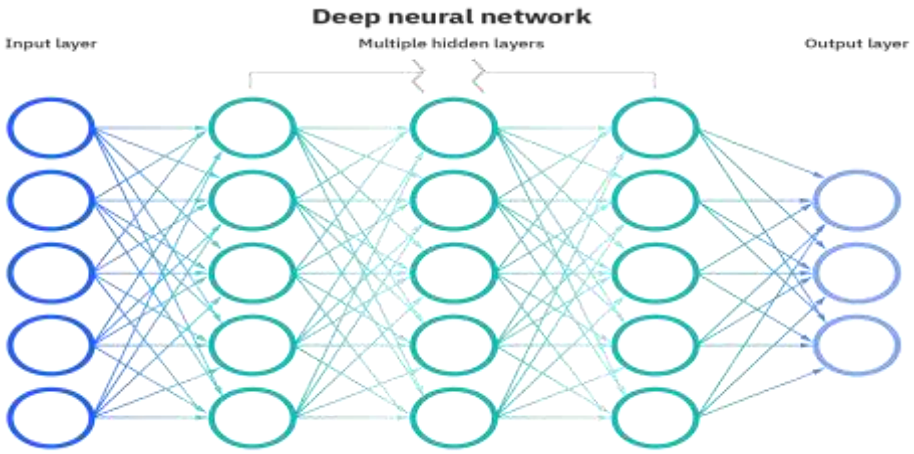
CMPE 407 Spring

Neural Networks

IBM def: Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.

Artificial neural networks (ANNs) are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.



How do neural networks work?

Think of each individual node as its own linear regression model, composed of input data, weights, a bias (or threshold), and an output. The formula would look something like this:

Mathematical formula used to determine summation

$\sum w_i x_i + bias = w_1x_1 + w_2x_2 + w_3x_3 + bias$

Mathematical formula used to determine the output

$output = f(x) = 1 \text{ if } \sum w_1x_1 + b >= 0; 0 \text{ if } \sum w_1x_1 + b < 0$

Once an input layer is determined, weights are assigned. These weights help determine the importance of any given variable, with larger ones contributing more significantly to the output compared to other inputs. All inputs are then multiplied by their respective weights and then summed. Afterward, the output is passed through an activation function, which

determines the output. If that output exceeds a given threshold, it “fires” (or activates) the node, passing data to the next layer in the network. This results in the output of one node becoming in the input of the next node. This process of passing data from one layer to the next layer defines this neural network as a feedforward network.

i represents the index of the sample,

y-hat is the predicted outcome,

y is the actual value, and

m is the number of samples.

Mathematical formula used to determine cost function

$Cost\ Function = MSE = 1/2m \sum 129\_(i=1)\hat{}m\blacksquare(y\hat{}((i)\ ) - y\hat{}((i)\ )\ )\hat{}2$

Types of neural networks

Neural networks can be classified into different types, which are used for different purposes. While this isn’t a comprehensive list of types, the below would be representative of the most common types of neural networks that you’ll come across for its common use cases:

The perceptron is the oldest neural network, created by Frank Rosenblatt in 1958. It has a single neuron and is the simplest form of a neural network:

Feedforward neural networks, or multi-layer perceptrons (MLPs). They are comprised of an input layer, a hidden layer or layers, and an output layer. While these neural networks are also commonly referred to as MLPs, it’s important to note that they are actually comprised of sigmoid neurons, not perceptrons, as most real-world problems are nonlinear. -> NLP

Convolutional neural networks (CNNs) are similar to feedforward networks, but they’re usually utilized for image recognition, pattern recognition, and/or computer vision.

Recurrent neural networks (RNNs) are identified by their feedback loops. These learning algorithms are primarily leveraged when using time-series data to make predictions about future outcomes, such as stock market predictions or sales forecasting.

DeepAI Definition

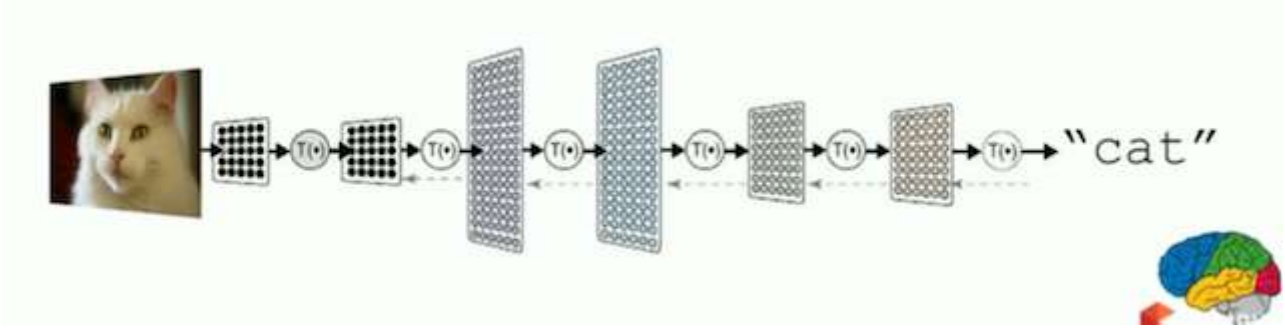
What is a Neural Network?

An artificial neural network learning algorithm, or neural network, or just neural net, is a computational learning system that uses a network of functions to understand and translate a

data input of one form into a desired output, usually in another form. The concept of the artificial neural network was inspired by human biology and the way neurons of the human brain function together to understand inputs from human senses.

How Does a Neural Network Work?

Machine learning algorithms that use neural networks generally do not need to be programmed with specific rules that define what to expect from the input. The neural net learning algorithm instead learns from processing many labeled examples (i.e. data with with "answers") that are supplied during training and using this answer key to learn what characteristics of the input are needed to construct the correct output. Once a sufficient number of examples have been processed, the neural network can begin to process new, unseen inputs and successfully return accurate results. The more examples and variety of inputs the program sees, the more accurate the results typically become because the program learns with experience.



(image is taken from a Google Tech Talk by Jeff Dean at Campus Seoul on March 7, 2016)

Some common applications of neural networks today, include image/pattern recognition, self driving vehicle trajectory prediction, facial recognition, data mining, email spam filtering, medical diagnosis, and cancer research.

BMC Blogs

Handwriting recognition with Neural Networks

Handwriting and facial recognition using neural networks does the same thing, meaning making a series of binary decisions. This is because any image can be broken down into its smallest object, the pixel. In the case of handwriting, like shown below, each pixel is either black (1) or white (meaning empty, or 0).

Neural Network Training

With artificial intelligence, we train the neural network by varying the weights x1, x2, x3, … , xn and the bias b. That is to say, we vary the inputs to minimize the loss function. That is no different than simple linear regression.

With simple linear regression, the loss function is the distance between the observed value z and the predicted value p, or z – p. With neural networks we use something more complicated

called the stochastic gradient descent, which is not necessary to be understood.It will suffice to say that it is basically the same thing. But finding the minimum value in some function with thousands of input variables is hard to achieve, so the stochastic gradient descent first takes a guess and then works from there.

References :

https://www.ibm.com/cloud/learn/neural-networks

https://deepai.org/machine-learning-glossary-and-terms/neural-network

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