Artificial intelligence is the theory and practice of how computers think; namely, it is the development of getting computers to perform tasks. As one might expect, there are a plethora of branches of artificial intelligence, like machine learning, expert systems, natural language processors and so on. One branch this paper will touch on is deep learning; in particular, Artificial Neural Networks.

Artificial Neural Networks show an analogous and computational model of how biological neurons work. This model is represented using a discrete set of objects called nodes, which connect to one another with edges. As biological neurons are the source for human learning, so too is artificial neurons used for machines' learning.

A simple model of an artificial neural network is the Perceptron. The perceptron's purpose is to take input and produce output. It consists of input values, synaptic weights, a bias, and an activation function. The way it works is as follows: suppose we have inputs $x_0, x_1, x_2, ...x_n$, where each are assigned a given weight, call them $c_0, c_1, c_2, ...c_n$. Then naturally, $x_0c_0 + x_1c_1 + \cdots + x_nc_n = \sum x_ic_i$. Here, the weights represent how strong a given node is.

Given the result of our partial sum, we may now construct the activation function, which is $\phi(\sum x_ic_i)$. An activation function helps transforms an input signal to an output signal. Note, this idea is analogous to how a dendrite helps neurons accept input signals, processes it through a myelin sheat, and then uses axons to produce output.

The activation function is largely important for its role in deciding whether or not a neuron should be activated. Mathematically, the function computes the weighted sum with the bias, or $\phi(\sum x_i c_i) = b + x_0 c_0 + x_1 c_1 + \cdots + x_n c_n$, where b is the bias. The bias value brings the activation function up or down.

The theoretical need for an activation function is apparent. That is, if we didn't have it, then the machine does not achieve an output signal. Yet, its practical need is cardinal in that without it, we no longer have a model that can learn and represent data, such as sound, images, and text.

Note, the activation function should be a non-linear function. The reason is so that our model can handle more complex tasks, giving it more power. The idea here is that we want a model that can learn anything we throw at it. A non-linear function helps accomplish this.

It's worthwhile to mention that there are several activation functions. One such function is the Threshold Activation Function, which is a step function that activates neurons and sends signals onto the next layer. Another function is the Sigmoid Activation Function, which is a logistic function that predicts the probability as an output. The Hyperbolic Tangent Function is slightly better than the Sigmoid function in that particular values keep their features after they've been transformed (e.g. negative values map to negative values, etc.). The most used activation function is the Rectified Linear Units (ReLu), and its ability to approximate values more accurately than the other functions is what makes its use standard in the industry.

To summarize, the artificial neural network takes a set of inputs that go through the weighted synapses. The values will be analyzed, where an activation function is then applied. The results of this function is then produced as output. One more thing we must account for is what happens when our model is not producing accurate results.

In order for the neural network to learn, it must compute another function called the Cost Function. The cost function is an error value and depicts the difference between the actual value and output value. The lower the error value is, the more accurate the model will be. Getting the Cost Function down as small as possible requires a process known as Back-propagation. This process allows the machine to go through the entire model again to adjust the weight and threshold values for the next set of inputs.

Furthermore, artificial neural networks have a huge range of commercial applications across several fields of study. One very important commercial application for artificial neural networks is the use of them for medical diagnosis. Use of artificial neural networks in diagnosing medical conditions not only provides a faster diagnostic process, but this technology is also able to prevent misdiagnosis in some cases. These ANN's have the ability to take input data such as clinical symptoms, outputs of imaging devices, standard measurements, and various types of biochemical data. Using all of these as inputs, the adaptive learning algorithms are able to perform some pretty remarkable diagnosis. This technology is already being used all over the world when it comes to cancer diagnosis, particularly lung cancer, prostate cancer, and colorectal cancer. Artificial neural networks even have the ability to determine the severity of a cancer using only cell shape. This technology has broad applications that in the future could be very essential in the medical field.