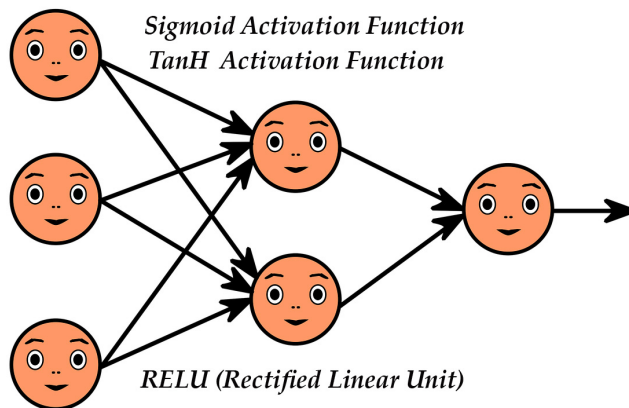




Comparison of Sigmoid, Tanh and ReLU Activation Functions

Posted on August 19, 2020 (<https://www.aitude.com/comparison-of-sigmoid-tanh-and-relu-activation-functions/>) - by Sandeep Kumar (<https://www.aitude.com/author/aitudes/>) - in ANN (<https://www.aitude.com/category/ann/>), Artificial Intelligence (<https://www.aitude.com/category/artificial-intelligence/>), Deep Learning (<https://www.aitude.com/category/deep-learning/>), Machine Learning (<https://www.aitude.com/category/machine-learning/>)



Introduction

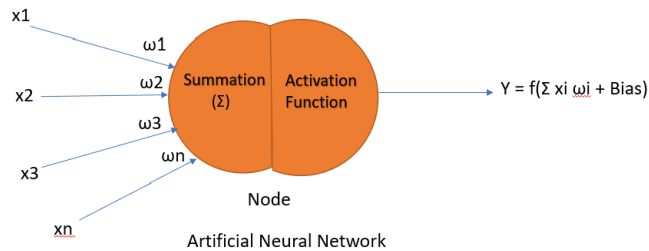
In Artificial Neural network (https://en.wikipedia.org/wiki/Artificial_neural_network) (ANN), activation functions are the most informative ingredient of Deep Learning which is fundamentally used for to determine the output of the deep learning models. In this blog, we will discuss the working of the ANN and different types of the Activation functions like Sigmoid, Tanh and ReLU (Rectified Linear Unit) in a very easy manner.

What is Artificial Neuron Network (ANN)?

ANN is the part of the Deep learning (<https://www.aitude.com/category/deep-learning/>) where we will learn about the artificial neurons. To understand this we have to understand about the working of the neurons in the proper way.

In biology we understand that the neurons are used to accept the information of a signals sensed by the organs and these organs sends the sensed data to our brain so our brain can take the appropriate decisions based on the sensed organs. According to the sensed results our brain do some operations, calculations and give some appropriate answer/output. This output is followed by our sensed organs.

If we want to implement all these brain functionality artificially then this type of network is known as the Artificial Neural Network where we will take the single node(which is a replica of the Neuron) and partition it into further two parts. First part is known as Summation and second one is considered as a function is known as the **Activation Function**.



Summation

This summation is used to collect all the neural signals along with their weights. For example first neuron signal is x_1 and their **weight** is ω_1 so the first neuron signal would be $x_1 \omega_1$. Similarly we will calculate the neural values for the second neuron, third neuron and so on. At last we will take the sum of all the neurons. So the total summation weight is calculated as

$$x_1\omega_1 + x_2\omega_2 + x_3\omega_3 + \dots + x_n\omega_n$$

Activation Function

Activation function is used to generate or define a particular output for a given node based on the input is getting provided. That means we will apply the activation function on the summation results.

$$Y = f(\sum x_i \omega_i + \text{Bias})$$

Activation Functions are multiple types like Linear Activation Function, Heaviside Activation Function, Sigmoid Function, Tanh function and RELU activation Function.

In deep learning, Activation functions are the very important part of any neural network because it is able to perform a very complicated and critical work like object detection, image classification, language translation, etc. which are necessary to address by using an activation function. We can't imagine to perform these tasks without using a deep learning.

But in this blog we will put more concentration on the Sigmoid Activation Function, Tanh Activation Function and Relu (Rectified Linear Unit) Activation Function because these Activation Functions are mostly used in ANN (Artificial Neural Network) and deep learning.

Sigmoid Activation Function

Sigmoid function is known as the logistic function which helps to normalize the output of any input in the range between 0 to 1. The main purpose of the activation function is to maintain the output or predicted value in the particular range, which makes the good efficiency and accuracy of the model.

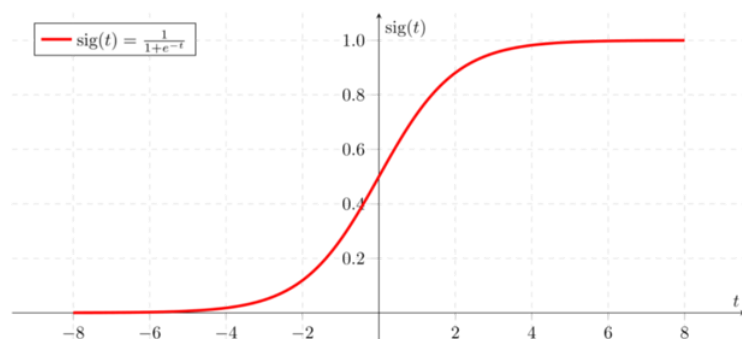


fig: sigmoid function

Equation of the sigmoid activation function is given by:

$$y = 1/(1+e^{(-x)})$$

Range: 0 to 1

Here Y can be anything for a neuron between range -infinity to +infinity. So, we have to bound our output to get the desired prediction or generalized results.

The major drawback of the sigmoid activation function is to create a vanishing gradient problem.

- This is the Non zero Centered Activation Function
- The model Learning rate is slow
- Create a Vanishing gradient problem.

Vanishing Gradient Problem

Vanishing gradient problem mostly occurs during the backpropagation when the value of the weights are changed. To understand the problem we will increase the value of the input values in the activation function, At that time we will notice that the predicted output is available on the range of the selected activation function and maintain the threshold value.

For the sigmoid function, the range is from 0 to 1. We know that the maximum threshold value is 1 and the minimum value is 0. So when we increase the input values, the predicted output must lie near to the upper threshold value which is 1. So the predicted output must be less than or near to the 1.

We again increasing the input value and the output comes on the max threshold value and lies there. When the neuron outputs are very small for example ($-1 < \text{output} < 1$), the patterns are created during the optimization will be smaller and smaller towards the upper layers. This causes them to make the learning process very slow, and make them converge to their optimum and this problem is known as the Vanishing Gradient Problem.

$$Y = \text{Activation function}(\sum(\text{weights} * \text{input} + \text{bias}))$$

In the nutshell, a neural network is a very dominant method and technology in machine learning which mimics how a brain perceives and operates.

Hyperbolic Tangent Activation Function

Tanh Activation function is superior then the Sigmoid Activation function because the range of this activation function is higher than the sigmoid activation function. This is the major difference between the Sigmoid and Tanh activation function (<https://www.aitude.com/comparison-of-sigmoid-tanh-and-relu-activation-functions/>). Rest functionality is the same as the sigmoid function like both can be used on the feed-forward network.

Range : -1 to 1

Equation can be created by:

$$y = \tanh(x)$$

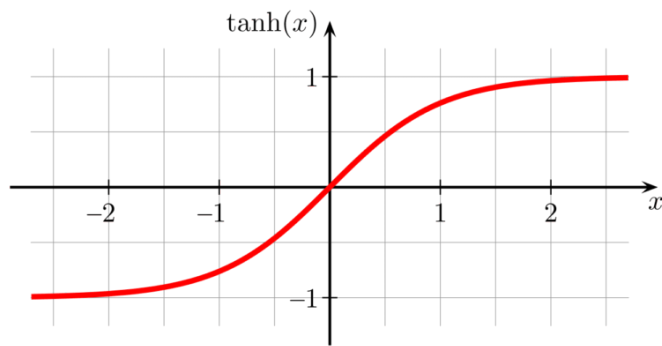


fig: Hyperbolic Tangent Activation function

Advantage of TanH Activation function

Here negative values are also considered whereas in the sigmoid minimum range is 0 but in Tanh, the minimum range is -1. This is why the Tanh activation function is also known as the **zero centered activation function**.

Disadvantage

Also facing the same issue of Vanishing Gradient Problem like a sigmoid function.

ReLu (Rectified Linear Unit) Activation Function

ReLu ([https://en.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks))) is the best and most advanced activation function right now compared to the sigmoid and TanH because all the drawbacks like Vanishing Gradient Problem is completely removed in this activation function which makes this activation function more advanced compare to other activation function.

Range: 0 to infinity

Equation can be created by:

$$\{ x \text{ if } x \geq 0$$

$$0 \text{ if } x < 0 \}$$

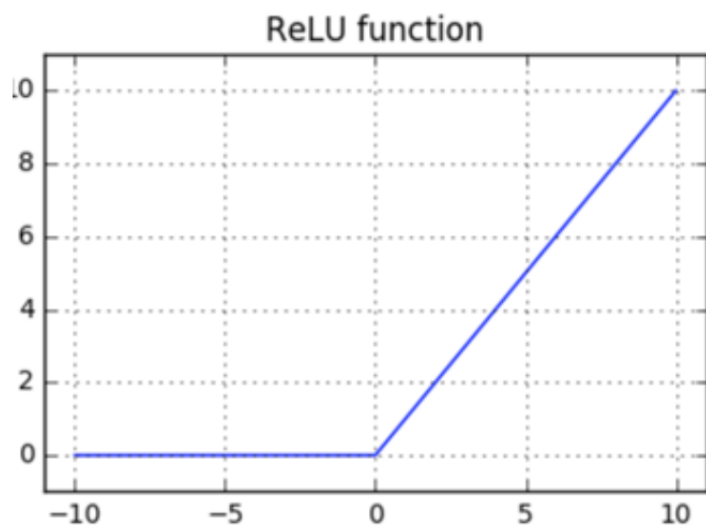


fig: ReLu Activation function

Advantage of ReLu:

- Here all the negative values are converted into the 0 so there are no negative values are available.
- Maximum Threshold values are Infinity, so there is no issue of Vanishing Gradient problem so the output prediction accuracy and there efficiency is maximum.
- Speed is fast compare to other activation function

Comparison

ACTIVATION FUNCTION	SIGMOID	TANH	RELU
Range	0 to 1	-1 to 1	0 to Infinity
Vanishing Gradient Problem	Yes	Yes	No
Nature	Non Linear	Non Linear	Linear
Zero Centered Activation Function	No	Yes	No
Symmetric Function	No	Yes	No
Equation	$y = 1/(1+e^{(-x)})$	$y = \tanh(x)$	$\{ \begin{array}{l} x \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{array} \}$
Model Accuracy	Good	Very Good	Excellent

Conclusion

Here we have discussed the working of the Artificial Neural Network and understand the functionality of sigmoid, hyperbolic Tangent (TanH), and ReLu (Rectified Linear Unit) Activation function and we also compared in a very easy manner. Now you have enough idea about the Activation functions

Tags: activation function / ANN / artificial intelligence / artificial neural network / deep learning / machine learning / RELU / Sigmoid Activation Function / TanH activation function

Understand Types of Environments in Artificial Intelligence

Older Post (<https://www.aitude.com/understand-types-of-environments-in-artificial-intelligence/>)

Build an Image Dataset in TensorFlow

Newer Post (<https://www.aitude.com/build-an-image-dataset-in-tensorflow/>)



Categories

ANN (<https://www.aitude.com/category/ann/>) (1)

Artificial Intelligence (<https://www.aitude.com/category/artificial-intelligence/>) (4)

Chatbots (<https://www.aitude.com/category/chatbots/>) (7)

Computer Vision (<https://www.aitude.com/category/computer-vision/>) (2)

Deep Learning (<https://www.aitude.com/category/deep-learning/>) (7)

Dialogflow (<https://www.aitude.com/category/dialogflow/>) (2)

Machine Learning (<https://www.aitude.com/category/machine-learning/>) (69)

Natural Language Processing (<https://www.aitude.com/category/natural-language-processing/>) (5)

Node.js (<https://www.aitude.com/category/node-js/>) (6)

TensorFlow (<https://www.aitude.com/category/tensorflow/>) (7)

Uncategorized (<https://www.aitude.com/category/uncategorized/>) (1)

Tags

activation function / algorithms / ANN / artificial intelligence / artificial neural network / bagging / boosting / chatbot / classification / computer vision / confusion matrix / decision tree / deep learning / dialogflow / echelon form / environments / evaluation metrics / express / facebook messenger / GANs / generative adversarial network / generative models / git / heroku / kaggle / knn / machine learning / matrix / mean stack / microsoft azure / model / mongodb / nlp / nodejs / random forest / regression / RELU / Sigmoid Activation Function / supervised learning / svm / TanH activation function / tensorflow / unsupervised learning / web server / web service



AI Development company in India.



(<https://facebook.com/flippercode>)
(<https://twitter.com/flippercode>)

Latest Tweets

Our Profile

- > Kaggle (<https://www.kaggle.com/aitude>)
- > Github (<http://github.com/aitude/>)
- > Parent Company (<https://www.flippercode.com>)

Image Gallery

(<https://www.aitude.com/wp-content/uploads/2020/01/Chat-Bot-Messenger-compressor.png>)

(<https://www.aitude.com/wp-content/uploads/2020/01/Computer-Vision-compressor.png>)

(<https://www.aitude.com/wp-content/uploads/2020/01/Robotics-Engineering-compressor.png>)

(<https://www.aitude.com/wp-content/uploads/2020/01/Artificial-Intelligence-compressor.png>)

(<https://www.aitude.com/wp-content/uploads/2020/01/Data-Science-compressor.png>)

(<https://www.aitude.com/wp-content/uploads/2020/01/Neural-Network-compressor.png>)

AITUDE 2022 | All Rights Reserved