Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks

Arnaud WATUSADISI MAVAKALA

AMMI, AIMS-Senegal

August 27, 2021

Overview

1. INTRODUCTION
   - Motivation
   - Objectives
2. SIMULATION
   - Activation functions
   - Comparison with different activation functions
3. CONCLUSION

Overview

1. INTRODUCTION
   - Motivation
   - Objectives
2. SIMULATION
   - Activation functions
   - Comparison with different activation functions
3. CONCLUSION

INTRODUCTION

Motivation

1. The activation function is used to transform the activation level of a unit (neuron) into an output signal.
2. Hidden units in neural networks need activation functions to introduce nonlinearity into the networks.
3. Nonlinearity makes multilayer networks so efficient.

Overview

INTRODUCTION

Objectives

To analyze the performance of generalized MLP architectures that have a backpropagation algorithm using different activation functions for the hidden and output layer neurons.

Overview

Uni-Polar Sigmoid Function

Activation function of Uni-polar sigmoid function is given as follows:

$$g(x) = \frac{1}{1 + e^{-x}}$$

Because it is easy to distinguish, which can interestingly minimize the computational capacity for training.

Uni-Polar Sigmoid Function

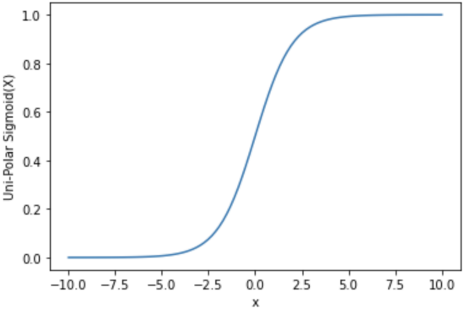


Figure: Uni-Polar Sigmoid Function

Bi-polar sigmoid function

Activation function of Bi-polar sigmoid function is given by

$$g(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

This type of activation function is well suited for applications that produce output values in the range [-1, 1].

Bi-polar sigmoid function

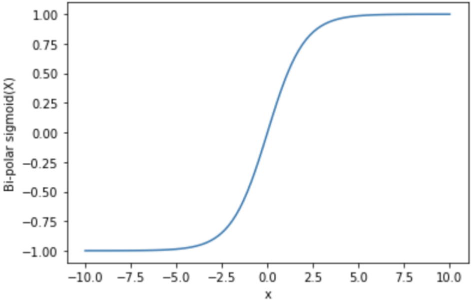


Figure: Bi-polar sigmoid function

Hyperbolic Tangent Function

Activation function of Hyperbolic Tangent Function is given by :

$$tanh(x) = \frac{sinh(x)}{cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

The hyperbolic tangent function is similar to the sigmoid function. Its range outputs are between -1 and 1.

Hyperbolic Tangent Function

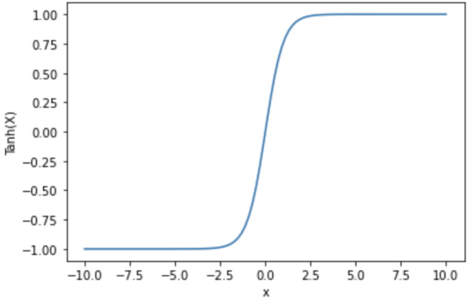


Figure: Hyperbolic Tangent Function

Radial Basis Function

A radial basis function (RBF) is a real-valued function whose value depends only on the distance to the origin, so that

$$g(x) = g(\|x\|)$$

or alternatively on the distance from some other point c, called a center, so that

$$g(x) = g(\|x - c\|)$$

Radial Basis Function

RBFs are generally used to construct approximations of functions of the form

$$y(x) = \sum_{i=1}^{N} W_i g(\|x - c_i\|)$$

The RBF can also be interpreted as a type of fairly simple single-layer artificial neural network called a radial basis function network, with the radial basis functions acting as the activation functions of the network.

Radial Basis Function

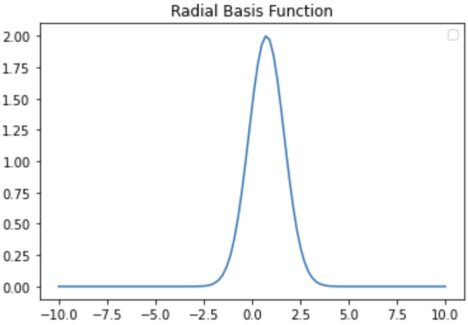


Figure: Radial Basis Function

Conic Section Function

CSF takes a parameter that determines the value of the angle of the function and can be defined as follows:

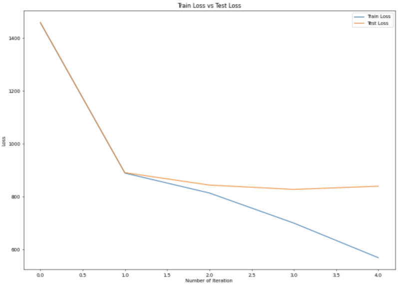
$$f(x) = \sum_{i=1}^{N+1} (a_i - c_i) w_i - cos w_i (\|a_i - c_i\|)$$

Where a_i is the input coefficient, c_i is the center, w_i is the weight in the multilayer perceptron (MLP).

Overview

Comparison with different activation functions

The different activation functions depend on different numbers of iterations to compare their performance using the same data.



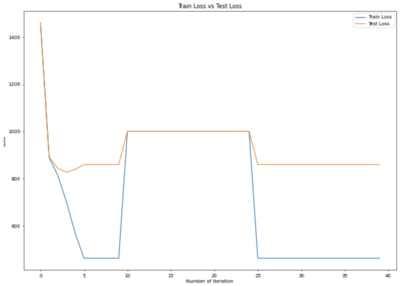


Figure: Uni-Polar Sigmoid and Bi-polar sigmoid

Comparison with different activation functions

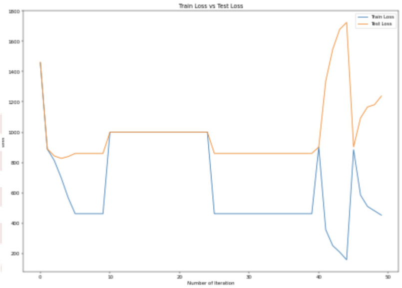


Figure: Tanh

CONCLUSION

We used five classical monotonic differentiable activation functions for MLP architecture evolution with generalized Delta rule learning. After comparing their performances, the simulation results show that the Tanh (hyperbolic tangent) function provides better recognition accuracy than the other functions. We claimed that the Tanh activation function can be used in the vast majority of MLP applications as a good choice to achieve high accuracy.

REFERENCES

- T. Poggio, F. Girosi, “A theory of networks for approximation and learning”. A.I. Memo No.1140, Artificial Intelligence, Laboratory, Massachusetts Institute of Technology, 1989
- B DasGupta, G. Schnitger, “The Power of Approximating: A Comparison of Activation Functions”. In Giles, C. L., Hanson, S. J., and Cowan, J. D., editors, Advances in Neural Information Processing Systems, 5, pp. 615-622, San Mateo, CA. Morgan Kaufmann Publishers, 1993
- B. Ciocoiu, “Hybrid Feedforward Neural Networks for Solving Classification Problems”. Neural Processing Letters, 16(1):81-91, 2002

Thank you for your attention