Journal of Advanced Research in Mechanical Engineering and Technology
Volume 6, Issue 1&2 - 2019, Pg. No. 27-31
Peer Reviewed Journal

Research Article

Neural Computing for Determining the Accuracy of Ultimate Tensile Strength of Friction Stir Welded Joints by using Various Activation Functions

Akshansh Mishra

Project Scientific Officer, Center of Artificial Intelligence based Friction Stir Welding, Stir Research Technologies, Uttar Pradesh, India.

DOI: https://doi.org/10.24321/2454.8650.201905

I N F O

E-mail Id:
akshansh.frictionwelding@gmail.com
Orcid Id:
https://orcid.org/0000-0003-4939-359X

How to cite this article:
Mishra A. Neural Computing for Determining the Accuracy of Ultimate Tensile Strength of Friction Stir Welded Joints by using Various Activation Functions. *J Adv Res Mech Engi Tech* 2019; 6(1&2): 27-31.

Date of Submission: 2019-07-01
Date of Acceptance: 2019-09-10

A B S T R A C T

Activation functions in a particular Artificial Neural Network (ANN) architecture plays a vital role. It imparts non-linear properties to our Neural Networks. There is a complicated and non-linear complex functional mapping between the inputs and response variable. In our present work, we have focussed on the accuracy of the UTS of the dissimilar Friction Stir Welded joints obtained by the training and testing the Artificial Neural Network architecture on Sigmoid activation function, Rectified Linear unit (ReLu) activation function and Hyperbolic tangent activation function. Tool Rotational Speed (rpm), Welding speed (mm/min) are the inputs and Ultimate Tensile Strength (MPa) is the output in our neural network architecture.

Keywords: Artificial Neural Network, Friction Stir Welding, Activation Functions, Google Colaboratory

Introduction

Artificial Neural network has gained more publicity in order to generalize and learn. Neural Networks have a wide class consisting of data reduction models, nonlinear dynamical systems, flexible nonlinear regression and discriminant models. Basically ANN are used in three ways: as data analytic methods, as models of biological nervous systems and "intelligence" and as real-time adaptive signal processors or controllers implemented in hardware for applications such as robots.[1]

Artificial Neural Network is more like our biological neurons because they possess capability to learn by examples. For example, suppose when you were beginner in calculus, you first went through the examples of the given chapter and at last you jump to the exercise part. This same thing works with the ANN too. They learn from the previous dataset on which they are trained and after that predict the output. Learning process of the ANN falls into three categories: supervised, unsupervised and reinforcement learning. Direct comparison between the actual output of an ANN and the desired correct output, also known as the target output is done in the case of Supervised Learning while Unsupervised learning is solely based on the correlations among input data. No further information on "correct output" is available for learning process.[2]

The fundamental blocks that made up neural network is a neuron. In Deep Learning we called neuron as a perceptron. The idea of a perceptron or a single neuron is very simple. Figure 1, represents the feed forward propagation of the information through a particular model. We have defined the set of inputs from x_1 to x_m which can be seen in the Figure 1.

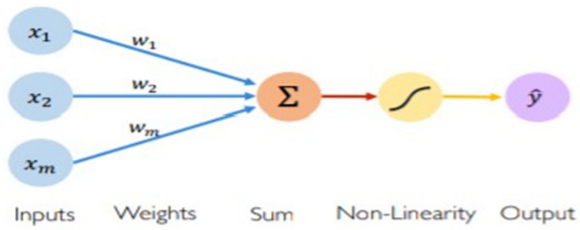

Figure 1.Feed forward propagation of the information through the perceptron

Each of the input is multiplied by its corresponding weight w_1 to w_m and all these multiplications are added up which comes together in summation. This given expression can be as follows:

$x_1w_1+x_2w_2+x_3w_3+.............x_mw_m$

We pass this weighted sum through non-linear activation function to produce a final output which we will call . We also have a bias term which allows us to shift the activation function left and right. The mathematical formula of the forward propagation in the perceptron is shown below:

Output

Linear combination of inputs

$$\hat{y} = g\left(w_0 + \sum_{i=1}^{m} x_i\ w_i\right)$$

Non-linear activation function

Bias

The purpose of activation functions is to introduce non-linearities into the network. This is extremely important in deep learning because in real life data is always non-linear in nature. Consider the Figure 2. Suppose we have to separate green from the red points.

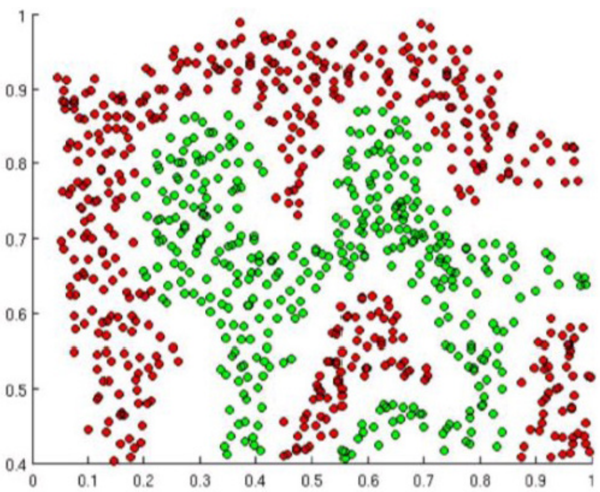

Figure 2.Graph consisting of Red and Green points

Many will think it is easy but what if only single line as shown in the Figure 3 is to be used to do it. It seems to be impossible because linear activation functions produce linear decisions no matter the network size.

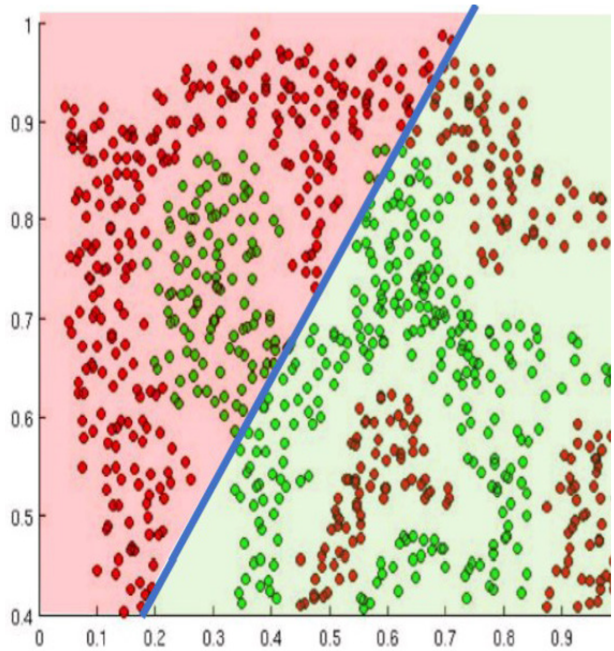

Figure 3.Single line drawn between the collection of Green and Red points

When we introduce non-linear activation function as shown in the figure 4 then it can allows us to approximate arbitrarily complex decisions boundaries in the feature space. So this is exactly which makes neural network powerful for practise.

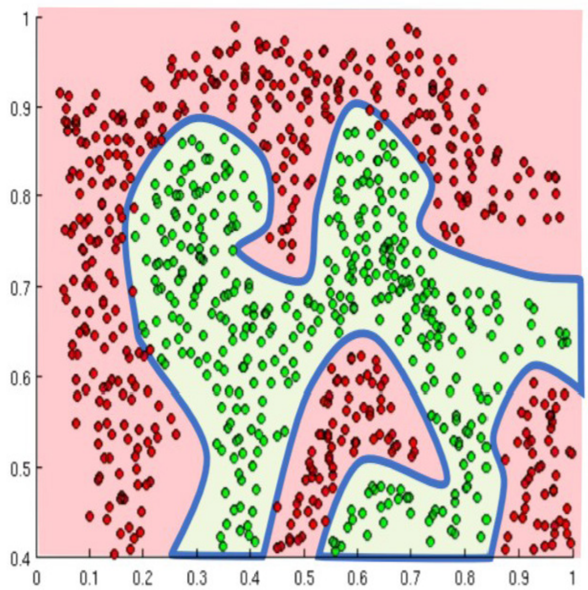

Figure 4.Non-linear activation function introduced in the graph

There are main three activation functions which are shown in the Figure 5.

ISSN: 2454-8650
DOI: https://doi.org/10.24321/2454.8650.201905

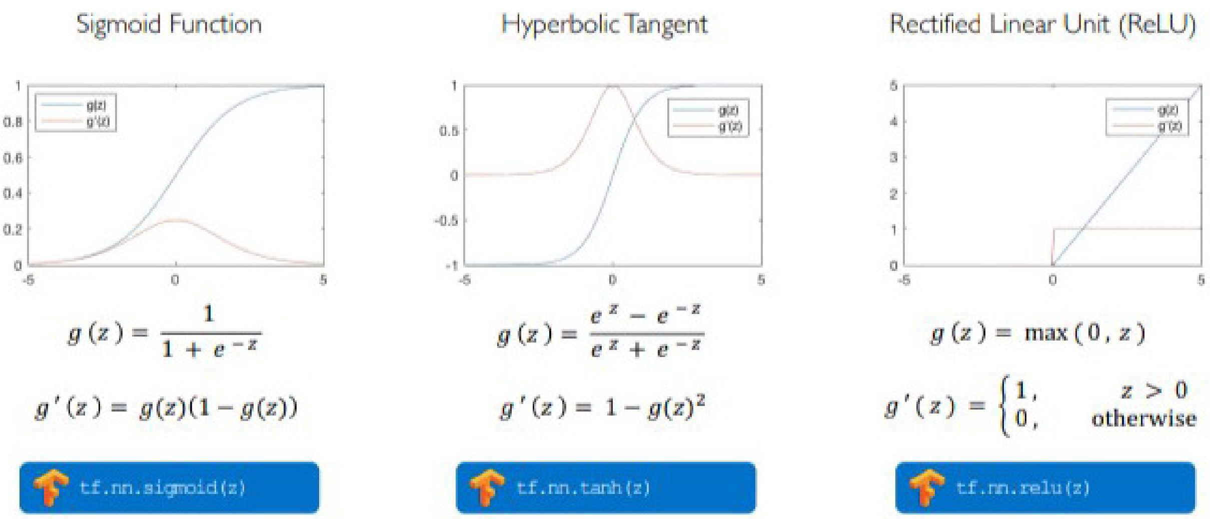

Figure 5.Three main types of activation functions

Sigmoid Activation function is mainly used for modelling probability. It collapses the input to be between 0 and 1. Since the probabilities are modelled between 0 and 1 so this is the perfect activation function at the end of the neural network to predict the probability distribution at the end. Another option is the ReLu function which can be seen in the Figure 2. This function is an extremely simple one to compute, it is piece-wise linear and is very popular because it is simple to compute. But this function has a non- linearity at z=0. So at z<0 this function is 0 and at z>0 it is just equal to the input. Due to its non-linearity it is capable to capture all properties of activation function while extremely being simple to compute.

Artificial Neural Network possesses unique ability for pattern clustering and pattern recognition. So it is very promising to take advantage of NN in prediction application.[3] Due to this reason application of Artificial Neural Network is employed in Friction Stir Welding process. Friction Stir Welding Process is a solid state joining process which is mainly used for joining dissimilar alloys.[4] Anand et al[5] with the help of ANN illustrated the correlation between the input and output responses of the friction welding of Incoloy 800H. For optimization purpose five different training algorithms were used to train ANN for both forward and reverse mapping and ANN tuned force approach. It was observed that the ANN model with genetic algorithm provides good ability to forecast the friction welding process parameters to weld Incoloy 800H.Fratini et al[6] properly trained an artifiical neural network model linked to the finite element model of the process to predict the local values of the average grain size.

In our study we have used the experimental dataset of Ghazaly et al[7] who studied the optimization of Friction Stir Welding Parameters of Al 6061 and Al 7075 Using Gray Rational Analysis. The main objective of our case study is to obtain Mean Square Error between the actual Ultimate Tensile Strength and predicted Ultimate Tensile Strength by using Sigmoid, Rectified Linear unit (ReLu) activation function and Hyperbolic Tangent Activation function separately.

Experimental Procedure

Ghazaly et al[7] in his case study used Al alloys of AA6061 and AA7075 to produce dissimilar joints by the Friction Stir Welding (FSW) process. The experimental setup is shown in the Figure 6.

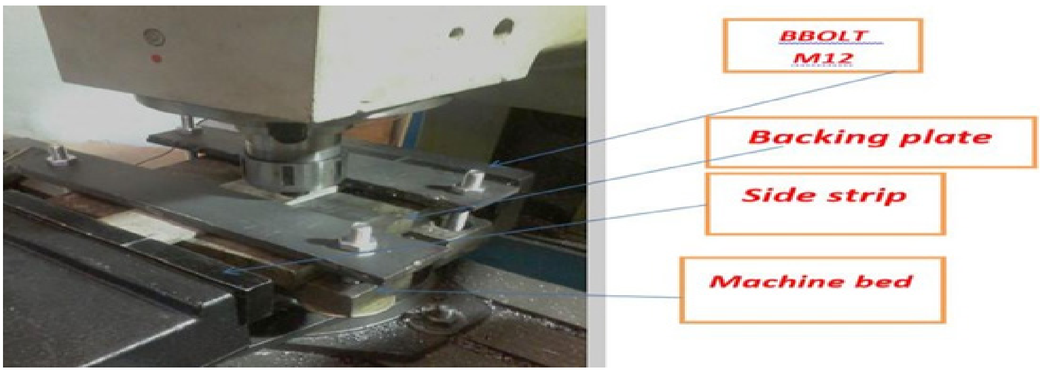

Figure 6.Experimental Setup of Friction Stir Welding process

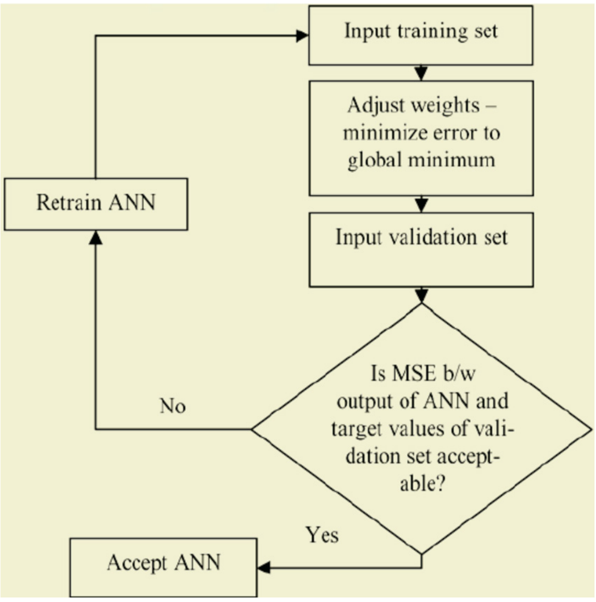

Figure 7.Flow chart of the working of the Artificial Neural Network

The two alloys used in the experimental study were plates 6 mm thick, 100 mm length and 50 mm width. Tensile test samples of the Friction Stir Welded dissimilar alloy plates were prepared as per ASTM E8 standard and transverse tensile properties like UTS, YS and E of the Friction Stir Welded joints were assessed by the computerized test machine.

In our case study we used Google Colaboratoryas a platform for executing our Python codes. Firstly, the required libraries were imported at the beginning of the coding. The further process is represented by the flowchart shown in the Figure 7. In this case study Epoch is equal to 1000 and Adam is used as an Optimizer.

Result and Discussions

The experimental results obtained are tabulated in the Table 1. TRS stands for Tool Rotational Speed, WS stands for Welding Speed, YS stands for Yield Strength, UTS stands for Ultimate Tensile Strength and E stands for Elongation %.

Table 1.Experimental Results

Exp. No	TRS (rpm)	WS (mm/min.)	YS (MPa)	UTS (MPa)	E %
1	400	10	137.20	145.51	3.52
2	400	20	135.10	143.28	3.21
3	400	30	132.44	140.46	3.00
4	400	40	130.27	138.16	2.73
5	600	10	146.70	161.14	4.22
6	600	20	144.50	158.70	3.85
7	600	30	141.61	155.55	3.60
8	600	40	139.30	153.01	3.27
9	800	10	156.80	184.10	5.92
10	800	20	154.50	181.40	4.49
11	800	30	151.36	177.70	4.20
12	800	40	148.88	174.80	3.82
13	1000	10	167.30	203.30	5.63
14	1000	20	164.80	204.25	5.13
15	1000	30	161.50	198.20	5.80
16	1000	40	158.85	145.50	4.36

The above dataset is firstly converted into csv file so that it can be imported during the Python Programming. Neural Network architecture used in our case study is 2-4-8-1. In the Neural Network Architecture TRS and WS are the inputs while UTS is the output. The heat map obtained between the Ultimate Tensile Strength (MPa), Tool Rotational Speed (rpm) and Welding Speed (mm/ min) is shown in the Figure 8.

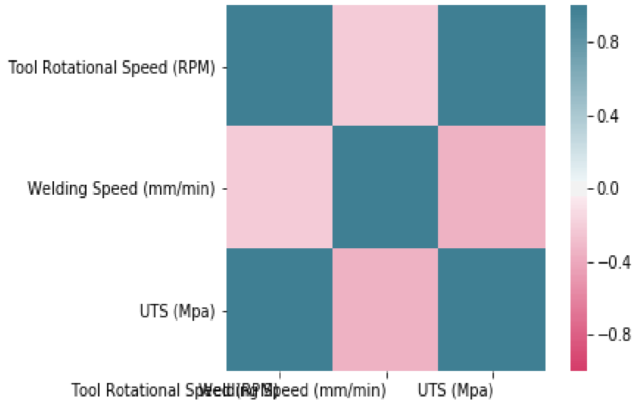

Figure 8.Heat Map of the Neural Network Architecture used in the case study

MSE obtained by the Sigmoid activation function, Hyperbolic Tangent activation function and ReLu activation function are 22.26, 11.74 and 14.05.

Conclusion

Previous research show that in deep learning the ReLU has become the activation function of choice because the math is much simpler from sigmoid activation functions such as tanh or logit, especially if we have many layers. To assign weights using backpropagation, we normally calculate the gradient of the loss function and apply the chain rule for hidden layers, meaning we need the derivative of the activation functions. ReLU is a ramp function where we have a flat part where the derivative is 0 and a skewed part where the derivative is 1. This makes the math really easy. If we use the hyperbolic tangent you might run into the fading gradient problem, meaning if x is smaller than -2 or bigger than 2, the derivative gets really small and our

ISSN: 2454-8650
DOI: https://doi.org/10.24321/2454.8650.201905

network might not converge, or we might end up having a dead neuron that does not fire anymore.

But in our case study hyperbolic tangent activation function us resulting more accurate value than ReLu activation function. So we can conclude that accuracy of various activation functions depend upon the model's architecture, the hyperparameters and the features that we are attempting to capture. Ideally, we utilize the ReLU function on our base models but we can always try out others if we are not able to reach an optimal result.

References

1. Sarle WS. Neural networks and statistical models. 1994.
2. Yao X. Evolving artificial neural networks. *Proceedings of the IEEE* 1999; 87(9): 1423-1447.
3. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art. *International journal of forecasting* 1998; 14(1): 35-62.
4. Mishra RS, Ma ZY. Friction stir welding and processing. Materials science and engineering: R: reports, 2005; 50(1-2): 1-78
5. Anand K, Barik BK, Tamilmannan K et al. Artificial neural network modeling studies to predict the friction welding process parameters of Incoloy 800H joints. *Engineering Science and Technology, an International Journal* 2015; 18(3): 394-407.
6. Fratini L, Buffa G, Palmeri D. Using a neural network for predicting the average grain size in friction stir welding processes. *Computers & Structures* 2009; 87(17-18): 1166-1174.
7. Al-Ghazaly SB, Al-shafaie SH. Optimization of Friction Stir Welding Parameters of Al 6061 and Al 7075 Using GRA. *JUBES* 2018; 26(5): 147.

ISSN: 2454-8650
DOI: https://doi.org/10.24321/2454.8650.201905