IUST
Iran University of Science and Technology

Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Double Sigmoid Activation Function for Fault Detection in Wind Turbine Generator using Artificial Neural Network

Noor Fazliana Fadzail*(C.A.), Samila Mat Zali* and Ernie Che Mid*

Abstract: The activation function has gained popularity in the research community since it is the most crucial component of the artificial neural network (ANN) algorithm. However, the existing activation function is unable to accurately capture the value of several parameters that are affected by the fault, especially in wind turbines (WT). Therefore, a new activation function is suggested in this paper, which is called the double sigmoid activation function to capture the value of certain parameters that are affected by the fault. The fault detection in WT with a doubly fed induction generator (DFIG) is the basis for the ANN algorithm model that is presented in this study. The ANN model was developed in different activation functions, namely linear and double sigmoid activation functions to evaluate the effectiveness of the proposed activation function. The findings indicate that the model with a double sigmoid activation function has greater accuracy than the model with a linear activation function. Moreover, the double sigmoid activation function provides an accuracy of more than 82% in the ANN algorithm. In conclusion, the simulated response demonstrates that the proposed double sigmoid activation function in the ANN model can effectively be applied in fault detection for DFIG based WT model.

Keywords: Activation Function, Fault Detection, Artificial Neural Network, Machine Learning, Doubly Fed Induction Generator, Wind Turbine.

1 Introduction

NOWADAYS, activation function in learning algorithm has witnessed the remarkable studies in various applications such as facial expression recognition [1, 2], brain tumor image classification [3] and classification of Alzheimer disease [4]. The activation function in artificial neural network (ANN) aids in learning and understanding complex and non-linear mappings between inputs and outputs [5]. There are many activation functions available namely binary step function, linear, sigmoid, tanh, ReLU, Leakly ReLU, swish and SoftMax. Each of this activation function has its own equation in producing the output of the ANN algorithm. ANNs without activation functions typically have limited performance and power, similar to linear regression models [5].

One of the components that contribute to accuracy in prediction is the activation function in ANNs. Thus, the activation function is a major area of focus for many researchers. As can be seen, Liang et al., [6] suggested recreating activation functions to increase deep learning accuracy for a variety of applications, including scientific computing and computer vision. The highest accuracy is produced by an ANN with the ideal activation function. Random weight ANN was introduced by Ertugrul, [7] to obtain the optimal activation function in linear regression. Samatin et al., [8] introduced a straightforward, novel type of activation function that can handle real-world challenges like recognition and categorization in multilayer feed-forward systems. Therefore, this proves that choosing an activation function is crucial for the ANN algorithm. Typically, trials or tuning are used to identify the ideal activation function [7].

Iranian Journal of Electrical & Electronic Engineering, 2025.
Paper first received 07 Dec. 2024 and accepted 18 Feb. 2025.
* The author is with the Faculty of Electrical Engineering & Technology, University Malaysia Perlis (UniMAP), 02600, Arau, Perlis, Malaysia.
E-mail: fazlianafadzail@unimap.edu.my
Corresponding Author: Noor Fazliana Fadzail

The most common activation function used in ANN is ReLU [9]. It was shown by Varshney and Singh [10], the ReLU activation function was generalized using a number of learnable slope parameters. Similar to Agarwal et al., [11], they modified the ReLU activation function to improve the model's ability to recognize the type and severity of disease in cucumber plants. ReLU activation function has higher accuracy than other activation functions, according to the results by Anadkat and Diwanji, [12]. This study analyzed the data to show how different activation functions affect the model's overall accuracy.

Another study examined the activation function was by Gao et al., [13]. An efficient convolutional neural network (CNN) for deep learning has been developed by Gao et al., [13]. It is used to detect and diagnose the multiple coexisting faults of the operating wind turbine (WT) gearbox. The developed machine learning model uses SoftMax function to classify the different conditions of faults using vibration signal. The mathematical operation known as SoftMax transforms an array of numbers into an array of probabilities, where the likelihood of each value in the vector is inversely proportional to its relative size.

Majority of studies in the subject of fault detection established fault detection by classifying the type of fault. As by Heo and Lee, [14], fault detection was developed using ANN with SoftMax activation function by classifying the type of fault into categories. Moreover, Leh et al., Tayeb and Asghar et al., [15 - 17] used binary classification using ANN for fault detection. Based on previous research, it is unknown which specific parameter is affected by the fault. Therefore, there is a need for a new activation function in ANN algorithm to classify the type of fault by capturing variable output value of certain parameters are that affected by the fault in ANN algorithm.

The main contributions of this work are described as follows:

(1) Double sigmoid activation function is proposed in ANN algorithm for fault detection in doubly fed induction generator (DFIG) WT.

(2) Provides accuracy of more than 82% in ANN algorithm.

The ANN algorithm model developed in this work is based on fault detection in DFIG WT. The fault detection cases are classified under internal and external stator fault of the DFIG model in the WT system. The ANN model was developed in two different activation functions (linear and double sigmoid activation function) to compare their performance. The outcome was presented based on the accuracy and RMSE value of each network in the model. Additionally, the model was developed in six distinct network configurations based on various hyperparameter values, such as the number of layers and neurons.

2 Methodology

Two models were developed in this study. The first model was created to identify a DFIG WT internal stator problem, while the other model was developed to investigate external stator defect in the DFIG WT. Both models were developed in python using Keras tool based on ANN algorithm. To determine whether the suggested activation function is effective, different activation functions were used namely linear and double sigmoid activation functions in the output layer of the model.

2.1 ANN algorithm

Fig. 1 illustrates an ANN model developed in this work consisting of the input, hidden, and output layers that make up this structure. Each neuron has a weight, bias, and activation function that characterise it. Activation function is one of the building blocks in ANN. Each layer contains their own activation function. An ANN without activation function is essentially a linear regression model. In this model, the hidden layers 1 and 2 are set as ReLU activation function [9, 11-12, 18], while the output layers; two different activation functions were used (linear and double sigmoid activation function).

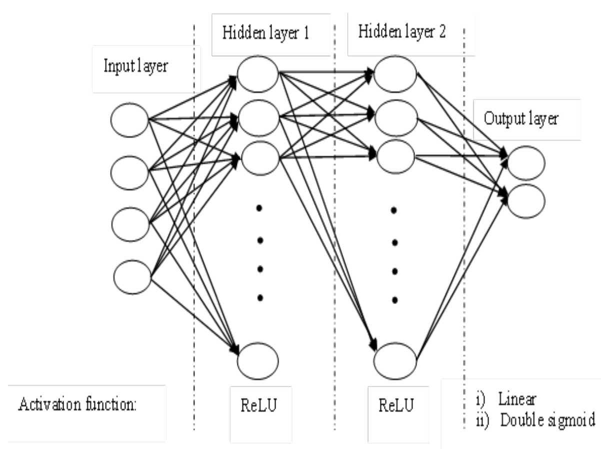

Fig 1. ANN model.

In this model, the ANN output was compared to the target value until the RMSE value was close to zero. At the same time, the model will adjust the parameter to obtain the best model. Fig. 2 depicts the procedure for obtaining the best model outcome. Non-parametric ANN models have many neurons, and more parameters are typically added as more connections are made. Parameters and hyperparameters make up an ANN. The difference between these two is that hyperparameters are fixed beforehand (for instance, the predetermined number of hidden layers), while parameters are changed during training [19].

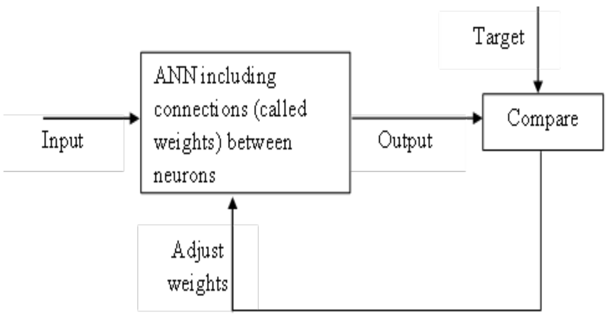

Fig 2. ANN algorithm model procedure.

Layer number and number of neurons in each hidden layer are the two primary hyperparameters that control the network's structure or architecture in this model. This hyperparameter configuration is most effectively implemented for a specific predictive modelling problem through systematic experimentation with a robust test harness. A rule of thumb suggests that a minimum of 2 layers should be used, though there are no rules that specify the minimum or maximum number of layers required to improve the performance and accuracy of neural networks [5]. Therefore, different configurations of layer and neuron numbers were used in this model, as shown in Table 1.

Table 1. Different configurations of hyperparameter value.

Networks	Number of neurons at layer 1	Number of neurons at layer 2
10	100	100
10	100	200
11	200	100
18	200	200
9	250	250
8	300	300

2.2 Double sigmoid activation function

In this work, a double sigmoid activation function was proposed in the output layer of the ANN algorithm. It is known as double sigmoid due to the double-S-shaped function. A double sigmoid is a continuous real function that has been smoothed and is differentiable [20]. Each function argument as well as each point in the domain are subject to differentiability [20]. Combining two logits in their sum or product, yields the double sigmoid function [20]. The graph for the double sigmoid activation function is illustrated in Fig. 3.

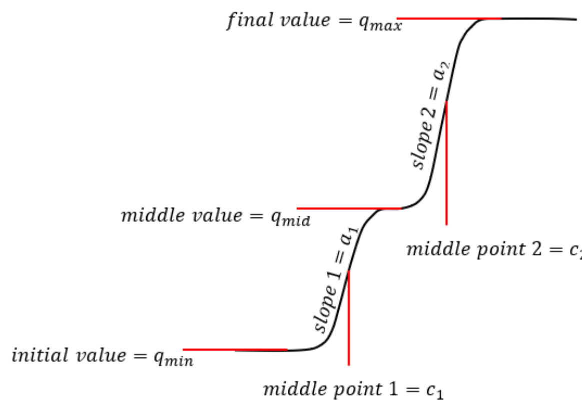

Fig 3. A graph of double sigmoid activation function [20].

The curve of the double sigmoid function as in Fig. 3 is determined by seven parameters (q_{min}, q_{mid}, q_{max}, a_1, a_2, c_1, c_2). An initial value, q_{min} and a final value, q_{max} correspond to the horizontal asymptotes, as well as a middle value, q_{mid} corresponds to the point at which the first increasing or decreasing phase ends and the second one begins. In addition, the last few variables relate to the two slopes, a_1 and a_2 and the two matching middle points, c_1 and c_2 have the same meaning as the logit function [20]. The double sigmoid function's equation is expressed in Eq. (1).

$$DS = q_{min} + \frac{q_{mid} - q_{min}}{1 + e^{(-a_1(\Delta t - c_1))}} + \frac{q_{max} - q_{mid}}{1 + e^{(-a_2(\Delta t - c_2))}} \tag{1}$$

where DS is a double sigmoid, $q_{min} < q_{mid} < q_{max}$, $a_1 > 0$, $a_2 > 0$, and $c_1 < c_2$ [20].

2.3 Internal stator fault

The ANN model needs to be trained and tested in order to obtain the best model. The training and testing data for internal stator fault cases emerged from the MATLAB Simulink model of the WT system with the DFIG model as described by Gagnon [21]. Internal stator fault covers inter-turn short circuit and open circuit fault. According to Wang et al., Wang et al., Zhao et al., Liu et al., Li et al., Chen et al., and Qi et al., [22-28], the current increases when there is an inter-turn short circuit fault and decreases when there is an open circuit fault. Therefore, the values of stator resistance and stator inductance are set to specific values as shown in Table 2, to achieve these characteristics. Impedance is set at 0.1 times the value of the normal condition for inter-turn short circuit faults and at 10 times the value of the normal condition for open circuit faults [24, 29].

Table 2. Parameters of the impedance value with its conditions.

Conditions	R_s (pu)	L_s (pu)
Normal	0.023	3.08
Inter-turn short circuit fault	0.0023	0.308
Open circuit fault	0.23	30.8

The simulated response of stator and rotor current of dq axis (I_{qs}, I_{ds}, I_{qr} and I_{dr}) were used as inputs in developing the ANN model. The outputs are impedance values, which include stator resistance, R_s, and stator inductance, L_s. The output of the model represents the condition of the DFIG WT. The model was developed in linear and double sigmoid activation function. The parameters of double sigmoid activation function are q_{min}=0.0023, q_{mid}=3.08, q_{max}=30.8, a_1=a_2=2, c_1=3 and c_2=9. These parameters are based on the impedance value of DFIG model in Table 2. Based on the parameters, Eq. (1) for this case becomes Eq. (2):

$$DS = 0.0023 + \frac{3.0777}{1+e^{(-2(x-3))}} + \frac{27.72}{1+e^{(-2(x-9))}} \tag{2}$$

2.4 External stator fault

In external stator fault, the model of WT based DFIG by Gagnon [21] was simulated for three conditions i.e. normal condition, loss of excitation (LOE) and external short circuit (ESC) fault. All the simulated responses were used as training and testing data.

In LOE fault, two 30 km parallel lines were connected to the DFIG trough a transformer [30]. The LOE fault caused DFIG to absorb reactive power, resulting in a drop in terminal voltage [30-35]. To achieve these characteristics, the DFIG model was simulated using the parameters listed in Table 3.

Table 3. Parameters for load model.

Parameter	Value
Nominal phase-to-phase voltage, V_n	575 Vrms
Nominal frequency, f_n	60 Hz
Active power, P	120k*6/50 W
Inductive reactive power, Q_L (positive var)	0 var
Capacitive reactive power, Q_C (negative var)	120k var

Meanwhile, in ESC fault, three-phase fault was inserted in the DFIG model. The three-phase fault occurs between phase A and the ground [33, 35]. The three-phase fault resulted in a drop in terminal voltage [35, 36]. The parameters for the three-phase fault are tabulated in Table 4.

Table 4. Parameters for three phase faults [33].

Parameter	Value
Fault resistance, R_{on}	0.001 Ω
Ground resistance, R_g	0.01 Ω
Snubber resistance, R_S	1e6 Ω

The inputs of the developed ANN model are voltage and current of stator and rotor based on dq axis (V_{qs}, V_{ds}, V_{qr}, V_{dr}, I_{qs} ,I_{ds} ,I_{qr} and I_{dr}). The outputs for the developed model are stator flux and negative sequence current [33]. Table 5 shows the outputs of the ANN model to identify the condition of the WT generator.

Table 5. Parameters for the output value with its condition [33].

Conditions	Stator flux (pu)	Negative sequence current (pu)
Normal	1.01<flux<1.02	Insignificant <0.05
LOE fault	Flux<0.5	Insignificant <0.05
ESC fault	0.9<flux<1.08	Large >0.05

The model as shown in Fig. 1 was developed in linear and double sigmoid activation function. The parameters of double sigmoid activation function are q_{min}=0.001, q_{mid}=0.5, q_{max}=1.08, a_1=a_2=2, c_1=3 and c_2=9. These parameters are based on the output values of DFIG model in Table 5. Based on the parameters, Eq. (1) for this case become Eq. (3):

$$DS = 0.001 + \frac{0.499}{1+e^{(-2(x-3))}} + \frac{0.58}{1+e^{(-2(x-9))}} \tag{3}$$

3 Results and discussion

The results are presented in two parts, which are internal and external stator fault. The accuracy and the RMSE value of the developed model were calculated based on Eq. (4) and Eq. (5) to the determine the model's performance.

$$Accuracy(\%) = \frac{min(x, \hat{x})}{max(x, \hat{x})} \times 100 \tag{4}$$

$$RMSE = \sqrt{\frac{\sum (x - \hat{x})^2}{n}} \tag{5}$$

where x is the actual value, $\hat{x}$ is the predicted value, and n is the sample size.

3.1 Internal stator fault

In internal stator fault cases, the model was tested with three conditions which are normal condition, inter-turn short circuit fault and open circuit fault. Tables 6 to Table 8 show the accuracy and the RMSE value of the output impedance value (R_s and L_s) of the model. In Table 6, the accuracy and RMSE between linear and double sigmoid activation functions in normal condition for each six networks based on Table 1 was presented. The accuracy of R_s in linear activation function was higher than 90% except for network 4, which had the lowest accuracy of 67.61%. Meanwhile, the accuracy for R_s in double sigmoid activation function showed good accuracy as the accuracy was higher than 94.51% for all networks. In L_s output, both activation functions show good accuracy because all the networks have accuracy greater than 98%. The RMSE values for all outputs were low and nearly zero in all networks.

Table 7 demonstrates the results for inter-turn short circuit fault. Based on Table 7, the result for double sigmoid activation function was more acceptable compared to linear activation function. It can be clearly seen in the accuracy of R_s and L_s in double sigmoid where the accuracy was higher than 96% for both output in network 3 and above. Moreover, the RMSE value in double sigmoid activation function was low compared to linear activation function. In linear activation function, the accuracy for R_s was low which is below than 78.69% for all networks. However, the accuracy for L_s was high.

The accuracy and the RMSE for open circuit fault are demonstrated in Table 8. The result for this case shows good performance as the accuracy in R_s and L_s was high for both activation functions. Besides, the RMSE was low as near to zero for all networks and all outputs.

Table 6. The accuracy and RMSE values in normal condition.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	R_s	L_s	R_s	L_s	R_s	L_s	R_s	L_s
1	94.01	98.92	0.00192	0.04325	97.68	99.86	0.00069	0.00555
2	95.95	99.26	0.00127	0.02938	95.00	99.62	0.00147	0.01456
3	96.34	99.03	0.00112	0.03992	97.30	99.89	0.00083	0.00415
4	67.61	99.18	0.01111	0.03160	**98.37**	**99.95**	**0.00050**	**0.00195**
5	97.35	99.67	0.00080	0.01925	94.51	99.59	0.00157	0.01555
6	97.30	99.43	0.00080	0.02324	97.59	99.96	0.00068	0.00163

Table 7. The accuracy and RMSE values in inter-turn short circuit fault.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	R_s	L_s	R_s	L_s	R_s	L_s	R_s	L_s
1	57.263	79.26	0.00220	0.08523	98.27	0.75	0.00006	0.30570
2	63.60	87.26	0.00139	0.05017	77.30	80.05	0.00122	0.08439
3	60.64	86.86	0.00176	0.05343	99.91	98.77	0.00000	0.00503
4	13.57	85.15	0.01508	0.05621	**99.49**	**98.91**	**0.00001**	**0.00404**
5	74.35	94.85	0.00082	0.02035	99.59	96.56	0.00001	0.01382
6	78.69	93.59	0.00083	0.02441	99.66	98.98	0.00001	0.00333

Table 8. The accuracy and RMSE values in open circuit fault.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	R_s	L_s	R_s	L_s	R_s	L_s	R_s	L_s
1	99.28	99.84	0.00194	0.06164	99.52	99.99	0.00139	0.00119
2	99.64	99.88	0.00100	0.04687	98.82	99.95	0.00347	0.01548
3	99.52	99.85	0.00137	0.05791	99.61	99.99	0.00113	0.00025
4	93.52	99.87	0.01599	0.05078	**99.63**	**99.99**	**0.00105**	**0.00019**
5	99.75	99.95	0.00075	0.02064	98.99	99.99	0.00299	0.00065
6	99.77	99.93	0.00069	0.02832	98.70	99.99	0.00310	0.00004

Based on these results, double sigmoid activation function was more acceptable compared to linear activation function in internal stator fault detection. Moreover, network 4, which has 200 neurons at layer 1 and 200 neurons at layer 2, exhibits higher accuracy and the lowest RMSE. Therefore, the output response of the model was presented based on network 4. Fig. 4 to Fig. 6 show the output response for internal stator fault case. The response was illustrated in three responses, which are the actual response and the predicted response (linear and double sigmoid activation function). The actual response was the response obtained from MATLAB Simulink, while the predicted response was obtained from the developed ANN model. Based on the figure, all the responses from the developed model in double sigmoid activation function matched the actual response. Thus, it shows that the developed model with double sigmoid activation function can detect the fault in DFIG model accurately.

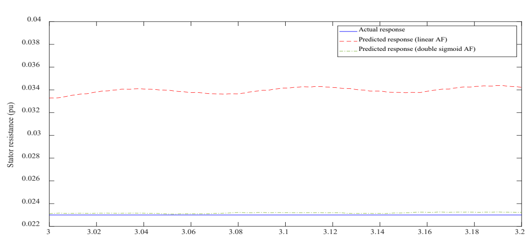

Fig 4. The output response for R_s in normal condition.

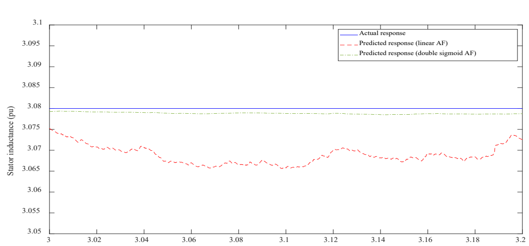

Fig. 5 The output response for L_s in normal condition.

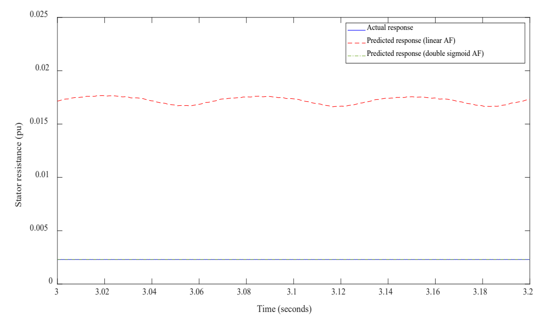

Fig. 6 The output response for R_s in inter-turn short circuit fault.

3.2 External stator fault

For external stator fault cases, the model was tested with three conditions which are normal condition, LOE and ESC faults. Table 9 to Table 11 show the accuracy and the RMSE of output value (stator flux and negative sequence current) of the ANN model. Table 9 depicts the result for normal condition, while Table 10 presents the result for LOE fault. Based on both tables, the accuracy for negative sequence current in linear activation function showed lowest accuracy compared to double sigmoid activation function. Meanwhile, the accuracy for flux shows high accuracy (99%) for all networks in linear and double sigmoid activation functions.

The results for accuracy and RMSE values in ESC fault are illustrated in Table 11. In this case, the result shows good performance as the accuracies for both activation functions were higher than 96% for all networks and all outputs. Moreover, the RMSE value was low, close to zero.

Table 9. The accuracy and RMSE in normal condition.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current
1	99.76	3.59	0.00301	0.00315	99.61	59.23	0.00522	0.00557
2	99.77	13.70	0.00299	0.00373	99.44	80.09	0.00690	0.00075
3	99.32	22.19	0.00759	0.00362	99.50	80.29	0.00656	0.00070
4	99.74	13.30	0.00319	0.00237	**99.74**	**84.97**	**0.00315**	**0.00023**
5	99.60	8.15	0.00486	0.00214	99.53	84.64	0.00590	0.00026
6	99.71	17.40	0.00342	0.00161	99.56	84.18	0.00490	0.00033

Table 10. The accuracy and RMSE in LOE fault.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current
1	99.57	8.24	0.00255	0.00247	99.41	11.12	0.00353	0.00730
2	99.74	4.95	0.00160	0.00210	99.51	78.59	0.00283	0.00032
3	99.45	9.92	0.00317	0.00330	99.64	57.52	0.00224	0.00075
4	99.58	25.85	0.00234	0.00201	**99.72**	**82.39**	**0.00162**	**0.00023**
5	99.10	3.07	0.00448	0.00144	99.81	81.32	0.00115	0.00024
6	99.57	14.06	0.00228	0.00157	99.77	77.99	0.00139	0.00028

Table 11. The accuracy and RMSE in ESC fault.

Network	Linear Activation Function				Double Sigmoid Activation Function			
	Accuracy (%)		RMSE		Accuracy (%)		RMSE	
	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current	Stator flux	Negative sequence current
1	99.26	98.62	0.00870	0.00523	99.01	96.00	0.01195	0.01466
2	99.35	98.60	0.00763	0.00500	99.22	98.85	0.00983	0.00377
3	99.55	98.96	0.00574	0.00398	99.29	98.14	0.00850	0.00623
4	99.51	99.33	0.00609	0.00271	**99.04**	**99.61**	**0.01091**	**0.00214**
5	99.56	99.12	0.00552	0.00316	99.40	99.38	0.00741	0.00315
6	99.54	99.13	0.00570	0.00299	99.47	99.17	0.00657	0.00253

Based on these results, the double sigmoid activation function was selected as the best activation function compared to linear activation function. It is due to the capability of the developed model with double sigmoid activation function to produce highest accuracy in negative sequence current output compared to linear activation function which produces lowest value of accuracy. Moreover, the accuracy in ESC fault with double sigmoid activation function obtained the highest accuracy of above 96% in all networks. Network 4 (200/200) was selected as the best model because of the high accuracy and low RMSE value. Due to this, network 4 was chosen to produce the output response. Fig. 7 to Fig. 9 show the output responses for external stator fault. Based on the figure, the output response from double sigmoid activation function was able to capture the actual response accurately. Therefore, it shows that the ANN model with double sigmoid activation was able to detect external stator fault in DFIG WT precisely.

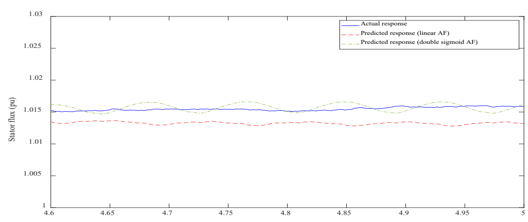

Fig. 7 The output response for stator flux in normal condition.

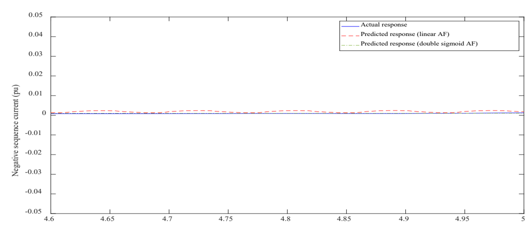

Fig. 8 The output response for negative sequence current in normal condition.

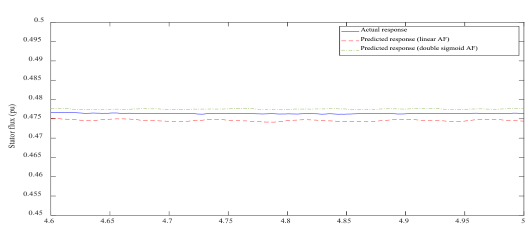

Fig. 9 The output response for stator flux in LOE fault.

4 Conclusions

This work presents double sigmoid activation function for fault detection in DFIG model-based WT using ANN algorithm. The proposed double sigmoid activation function's effectiveness was tested and evaluated using the internal and external stator fault. An ANN model using Keras tool was used to create the fault detection model. This model was developed in linear and double sigmoid activation functions to evaluate the performance of the proposed activation function. The ANN model needs to be tested and trained using a variety of data sets. Thus, the training and testing data used in this study comes from the simulated WT system - based DFIG responses in MATLAB Simulink. The ANN model was eventually developed in Python. Stator and rotor currents are employed as the input and impedance values as the output in the internal stator failure model. The impedance values classified the condition of WT i.e. normal, an inter-turn short circuit fault, and open circuit fault. Meanwhile, in external stator fault the stator and rotor currents and voltage are used as the input and stator flux and negative sequence current as the output. The stator flux and negative sequence current values indicate the conditions of the DFIG, either normal, LOE or ESC fault. Using the accuracy and RMSE value, the model's performance can be assessed. The outcomes clearly show that the accuracy in double sigmoid activation function was higher than linear activation function. The accuracy for double sigmoid was more than 82% for all cases. Moreover, the RMSE value also was low in double sigmoid activation function. Even while linear activation functions are straightforward and simple to solve, their complexity is constrained, and they are unable to learn and recognise complex mappings from data [5]. It was shown in the output response of model with linear activation function that was not able to capture the actual response accurately. Moreover, the accuracy for negative sequence current in external stator fault cases was low for certain network in linear activation function. Therefore, linear activation functions are ideal where interpretability is required and for simple tasks [5]. Besides, the best configuration model for both cases is hidden layer 1=200 nodes and hidden layer 2=200 nodes. It has been demonstrated that higher depth does appear to improve generalisation for a wide range of activities [37]. The output response for the double sigmoid action function can accurately represent the actual response and has the lowest RMSE and highest accuracy. Thus, it is proven that the proposed double sigmoid activation function can precisely recognise the internal and external stator faults in the DFIG WT system. Additionally, the constructed ANN model was quite simple because it simply predicted the model output using stator and rotor currents and voltage.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

All authors contributed equally to the conception, design, and methodology of the study. They collaboratively conducted the experiments, analysed the data, and interpreted the results. All authors were actively involved in drafting, reviewing, and revising the manuscript. Each author has approved the final version of the manuscript and agrees to be accountable for the content of the work.

Acknowledgment

The authors acknowledge the financial support provided by the Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis (UniMAP) under FKTE research fund.

References

[1] Y. Wang, Y. Li, Y. Song, and X. Rong, "The influence of the activation function in a convolution neural network model of facial expression recognition," *Applied Sciences*, vol. 10, no. 5, p. 1897, 2020, doi: 10.3390/app10051897.

[2] H. Jun, L. Shuai, S. Jinming, L. Yue, W. Jingwei and J. Peng, "Facial Expression Recognition Based on VGGNet Convolutional Neural Network," *2018 Chinese Automation Congress (CAC)*, Xi'an, China, 2018, pp. 4146-4151, doi: 10.1109/CAC.2018.8623238.

[3] Z. Huang, X. Du, L. Chen, Y. Li, M. Liu, Y. Chou, and L. Jin, "Convolutional neural network based on complex networks for brain tumor image classification with a modified activation function," *IEEE Access*, vol. 8, pp. 89281–89290, 2020, doi: 10.1109/ACCESS.2020.2993618.

[4] S. H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng, "Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling," Journal of Medical Systems, vol. 42, pp. 1–11, 2018, doi: 10.1007/s10916-018-0932-7.

[5] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *Towards Data Science*, vol. 6, no. 12, pp. 310–316, 2017.

[6] S. Liang, L. Lyu, C. Wang, and H. Yang, "Reproducing activation function for deep learning," *arXiv preprint*, arXiv:2101.04844, 2021.

[7] Ö. F. Ertuğrul, "A novel type of activation function in artificial neural networks: Trained activation function," *Neural Networks*, vol. 99, pp. 148–157, 2018, doi: 10.1016/j.neunet.2018.01.007.

[8] A. N. Samatin Njikam and H. Zhao, "A novel activation function for multilayer feed-forward neural networks," *Applied Intelligence*, vol. 45, pp. 75–82, 2016, doi: 10.1007/s10489-015-0744-0.

[9] S. Sarkar, S. Agrawal, T. Baker, P. K. R. Maddikunta, and T. R. Gadekallu, "Catalysis of neural activation functions: Adaptive feed-forward training for big data applications," *Applied Intelligence*, vol. 52, no. 12, pp. 13364–13383, 2022, doi: 10.1007/s10489-021-03082-y.

[10] M. Varshney and P. Singh, "Optimizing nonlinear activation function for convolutional neural networks," *Signal, Image and Video Processing*, vol. 15, no. 6, pp. 1323–1330, 2021, doi: 10.1007/s11760-021-01863-z.

[11] M. Agarwal, S. Gupta, and K. K. Biswas, "A new Conv2D model with modified ReLU activation function for identification of disease type and severity in cucumber plant," *Sustainable Computing: Informatics and Systems*, vol. 30, p. 100473, 2021, doi: 10.1016/j.suscom.2020.100473.

[12] K. D. Anadkat and H. M. Diwanji, "Effect of activation function in speech emotion recognition on the RAVDESS dataset," *Reliability: Theory & Applications*, vol. 16, no. 3 (63), pp. 228–236, 2021.

[13] Q. Gao, X. Wu, J. Guo, H. Zhou, and W. Ruan, "Machine-learning-based intelligent mechanical fault detection and diagnosis of wind turbines," *Mathematical Problems in Engineering*, vol. 2021, no. 1, p. 9915084, 2021, doi: 10.1155/2021/9915084.

[14] S. Heo and J. H. Lee, "Fault detection and classification using artificial neural networks," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 470–475, 2018, doi: 10.1016/j.ifacol.2018.09.380.

[15] N. A. M. Leh, F. M. Zain, Z. Muhammad, S. A. Hamid and A. D. Rosli, "Fault Detection Method Using ANN for Power Transmission Line," *2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, Malaysia, 2020, pp. 79-84, doi: 10.1109/ICCSCE50387.2020.9204921

[16] E. B. M. Tayeb, "Faults detection in power systems using artificial neural network," *American Journal of Engineering Research*, vol. 2, no. 6, pp. 69–75, 2013.

[17] F. Asghar, M. Talha, and S. H. Kim, "Neural network-based fault detection and diagnosis system for three-phase inverter in variable speed drive with induction motor," *Journal of Control*

Science and Engineering, vol. 2016, no. 1, p. 1286318, 2016, doi: 10.1155/2016/1286318.

[18] M. Varshney and P. Singh, "Optimizing nonlinear activation function for convolutional neural networks," *Signal, Image and Video Processing*, vol. 15, no. 6, pp. 1323–1330, 2021, doi: 10.1007/s11760-021-01863-z.

[19] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," *Renewable Energy*, vol. 133, pp. 620–635, 2019, doi: 10.1016/j.renene.2018.10.047.

[20] F. Baione, D. Biancalana, and P. De Angelis, "An application of sigmoid and double-sigmoid functions for dynamic policyholder behaviour," *Decisions in Economics and Finance*, vol. 44, pp. 5–22, 2021, doi: 10.1007/s10203-020-00279-7.

[21] R. Gagnon, "Wind Farm-DFIG detailed model (MATLAB/Simulink)," MATLAB/Simulink, 2021.

[22] C. Wang, X. Liu, and Z. Chen, "Incipient stator insulation fault detection of permanent magnet synchronous wind generators based on Hilbert–Huang transformation," *IEEE Transactions on Magnetics*, vol. 50, no. 11, pp. 1–4, 2014, doi: 10.1109/TMAG.2014.2318207.

[23] L. Wang, Y. Li, and J. Li, "Diagnosis of inter-turn short circuit of synchronous generator rotor winding based on Volterra kernel identification," *Energies*, vol. 11, no. 10, p. 2524, 2018, doi: 10.3390/en11102524.

[24] H. Zhao, W. Hao, H. Wu, Z. Yang, and X. Shi, "Modeling and simulation of aircraft main generator stator winding faults," *Journal of Computational Methods in Sciences and Engineering*, vol. 17, no. 4, pp. 691–704, 2017, doi: 10.3233/JCM-170753

[25] Y. Liu, R. Qu, J. Wang, H. Fang, X. Zhang, and H. Chen, "Influences of generator parameters on fault current and torque in a large-scale superconducting wind generator," *IEEE Transactions on Applied Superconductivity*, vol. 25, no. 6, pp. 1–9, 2015, doi: 10.1109/TASC.2015.2493126.

[26] J. Li, J. Liu, and Y. Chen, "A fault warning for inter-turn short circuit of excitation winding of synchronous generator based on GRU-CNN," *Global Energy Interconnection*, vol. 5, no. 2, pp. 236–248, 2022, doi: 10.1016/j.gloei.2022.04.020.

[27] X. Chen, P. Qin, Y. Chen, J. Zhao, W. Li, Y. Mao, and T. Zhao, "Inter-turn short circuit fault diagnosis of PMSM," *Electronics*, vol. 11, no. 10, p. 1576, 2022, doi: 10.3390/electronics11101576.

[28] Y. Qi, M. Zafarani, B. Akin, and S. E. Fedigan, "Analysis and detection of inter-turn short-circuit fault through extended self-commissioning," *IEEE Transactions on Industry Applications*, vol. 53, no. 3, pp. 2730–2739, 2016, doi: 10.1109/TIA.2016.2626264

[29] W. Yang, P. J. Tavner, and R. Court, "An online technique for condition monitoring the induction generators used in wind and marine turbines," *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 103–112, 2013, doi: 10.1016/j.ymssp.2012.03.002.

[30] M. J. Abbasi and H. Yaghobi, "Loss of excitation detection in doubly fed induction generator by voltage and reactive power rate," *Iranian Journal of Electrical and Electronic Engineering*, vol. 12, no. 4, pp. 270–280, 2016, doi: 10.22068/IJEEE.12.4.270.

[31] M. Amini, M. Davarpanah, and M. Sanaye-Pasand, "A novel approach to detect the synchronous generator loss of excitation," *IEEE Transactions on Power Delivery*, vol. 30, no. 3, pp. 1429–1438, 2014, doi: 10.1109/TPWRD.2014.2370763.

[32] Ö. Usta, M. H. Musa, M. Bayrak, and M. A. Redfern, "A new relaying algorithm to detect loss of excitation of synchronous generators," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 15, no. 3, pp. 339–349, 2007.

[33] H. Yaghobi and H. Mortazavi, "A novel method to prevent incorrect operation of synchronous generator loss of excitation relay during and after different external faults," *International Transactions on Electrical Energy Systems*, vol. 25, no. 9, pp. 1717–1735, 2015, doi: 10.1002/etep.1922.

[34] M. Abedini, M. Sanaye-Pasand, and M. Davarpanah, "Flux linkage estimation based loss of excitation relay for synchronous generator," *IET Generation, Transmission & Distribution*, vol. 11, no. 1, pp. 280–288, 2017, doi: 10.1049/iet-gtd.2016.1009.

[35] S. Soued, H. S. Ramadan, and M. Becherif, "Effect of doubly fed induction generator on transient stability analysis under fault conditions," *Energy Procedia*, vol. 162, pp. 315–324, 2019, doi: 10.1016/j.egypro.2019.04.033.

[36] M. Raghavendra Rao and E. Baby, "Effect of three-phase short circuit faults on the performance of wind farms employing doubly-fed induction generators," International Journal of Applied Engineering Research, vol. 13, no. 24, pp. 16913–

16918, 2018.

[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

Noor Fazliana Fadzail received her PhD at University Malaysia Perlis in 2023. Her research focuses on fault detection and renewable energy. Currently, she serves as a Senior Lecturer at the Faculty of Electrical Engineering & Technology, University Malaysia Perlis, Malaysia., where she continues to contribute in academic and research.

Samila Mat Zali received the B.Eng and M.Sc. degrees from the National University of Malaysia, Malaysia, in 1999 and 2002, respectively, and the Ph.D. degree from The University of Manchester, Manchester, U.K., in 2012, all in electrical engineering. She is currently a Associate Profesor with the Faculty of Electrical Engineering and Technology, Universiti Malaysia Perlis, Malaysia.

Ernie Che Mid received the B.Eng.and M.Eng.Sc. degrees in electrical engineering from Universiti Malaya, in 2006 and 2009, respectively. She obtained the Ph.D.degree in University of College London, 2019. She is currently working as senior lecturer at Faculty of Electrical Engineering &Technology, Universiti Malaysia Perlis, Malaysia. Her current research interests include fault detection, parameter estimation and control and optimization algorithm.