

Induction Machine Fault Detection Using Generalized Feed Forward Neural Network

V. N. Ghate[†] and S. V. Dudul^{*}

Abstract – Industrial motors are subject to incipient faults which, if undetected, can lead to motor failure. The necessity of incipient fault detection can be justified by safety and economical reasons. The technology of artificial neural networks has been successfully used to solve the motor incipient fault detection problem. This paper develops inexpensive, reliable, and noninvasive NN based incipient fault detection scheme for small and medium sized induction motors. Detailed design procedure for achieving the optimal NN model and Principal Component Analysis for dimensionality reduction is proposed. Overall thirteen statistical parameters are used as feature space to achieve the desired classification. GFFD NN model is designed and verified for optimal performance in fault identification on experimental data set of custom designed 2 HP, three phase 50 Hz induction motor.

Keywords: Induction motor, Fault detection, GFFDNN, PCA

1. Introduction

Induction machines play an essential role in certain industries, such as manufacture, transportation, etc. They provide the core capabilities for industrial success and the maintenance of them is essential and profitable to most electrical industrial processes. A lack of coherent maintenance strategy may result in the loss of individual items of a plant, and a heavy capitalized losses burden. As it is not economical to introduce redundant backup machines, online monitoring for induction machines is important for safe operation and production quality. In order to keep machines in good condition, techniques such as fault monitoring, detection, classification, and diagnosis have become increasingly essential [1]–[3]. There are invasive and non-invasive methods for machine fault detection [4], [5], [7]. The noninvasive methods are more preferable than the invasive methods because they are based on easily accessible and inexpensive measurements to diagnose the machine conditions without disintegrating the machine structure. Recently, artificial intelligence (AI) techniques have been proposed for the noninvasive machine fault detection [4], [6], [8]. They have several advantages over the traditional model-based techniques [4], [6]. They require no detailed analysis of the different kinds of faults or modeling of the system. These AI-based techniques include expert systems, neural network, and fuzzy logic. An expert system is able to manage knowledge-based production rules that model the physical system [6]. Neural network approaches can be considered as “black-box” methods as they do not provide heuristic reasoning about the fault detection process [4], [5]. Fuzzy logic systems can heuristically implement fault de-

tection principles and heuristically interpret and analyze their results [4]. In this paper, neural network type approach is used because generalized feed forward neural network (GFFDNN) is able to provide an accurate fault diagnostic classification.

2. Feature Extraction

The main problems facing the use of ANN are the selection of the best inputs and how to choose the ANN parameters making the structure compact, and creating highly accurate networks. For the proposed system, the feature selection is also an important process since there are many features after feature extraction. Many input features require significant computational efforts to calculate, and maybe result in a low success rate. To collect data at different conditions i.e. healthy condition, under inter turn fault, Eccentricity and both i.e. inter turn and Eccentricity specially designed 2 HP, 4 pole, 415V, 50 Hz, three phase induction motor is used. Three AC current probes were used to measure the stator current signals. From the time waveforms, as shown in Fig.1, no conspicuous difference exists among the different conditions.

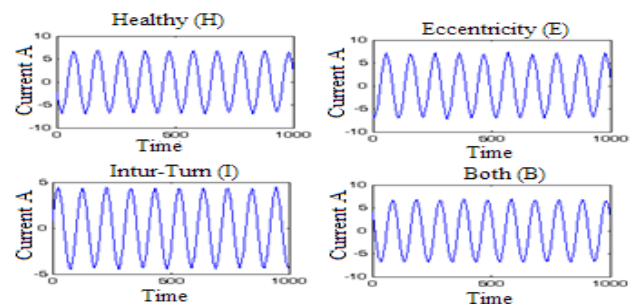


Fig. 1. Experimental Waveform for Current of Phase A for Various Conditions

[†] Corresponding Author : Dept. of Electrical Engineering, Govt. College of Engineering Amravati Maharashtra, India. (vng786@rediffmail.com)

^{*} Dept. of Applied Electronic Engineering, Sant Gadge Baba. Amravati University, Maharashtra, India. (dudulsv@rediffmail.com)

There is a need to come up with a feature extraction method to classify them. To classify the different faults the statistical parameters are used. To be precise, 'sample' statistics will be calculated for current data. Overall thirteen parameters are calculated as feature space. Minimum set of statistic to be examined includes the root mean square (RMS) of the zero mean signal (which is the standard deviation), the maximum, and minimum values the skewness coefficient and kurtosis coefficient. *Pearson's coefficient of skewness*, g_2 defined by:

$$g_2 = \frac{3(\bar{x} - \tilde{x})}{S_x} \quad (1)$$

Where \bar{x} denotes mean, \tilde{x} denotes median and S_x denotes the sample standard deviation.

The sample coefficient of variation, v_x is defined by;

$$v_x = \frac{S_x}{x} \quad (2)$$

The r^{th} sample moment about the sample mean for a data set is given by;

$$m_r = \frac{\sum_{i=1}^n (x_i - \bar{x})^r}{n} \quad (3)$$

m_2 denotes to spread about the center, m_3 refers to skewness about the center; m_4 denotes to how much data is massed at the center. Second, third and fourth moments are used to define the sample coefficient of skewness, g_3 and the sample coefficient of kurtosis, g_4 as follows.

$$g_3 = \frac{m_3}{(\sqrt{m_2})^3} \quad (4)$$

$$g_4 = \frac{m_4}{(\sqrt{m_2})^4} \quad (5)$$

The sample covariance between dimensions j and k is defined as;

$$c_{jk} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{(n-1)} \quad (6)$$

The ordinary correlation coefficient for dimensions j and k , r_{jk} is defined as;

$$r_{jk} = \frac{c_{jk}}{S_j - S_k} \quad (7)$$

3. Generalized Feed Forward NN

Generalized feed forward network is a generalization of the MLP such that connections can jump over one or more layers. In this network it is allowed to cross the hidden layer, i.e. output layer will get the input from the hidden layer and directly from input layer also.

Generalized feed forward Neural Network is proposed as fault classifier. Number of input Processing Elements

(PE) must be equal to that of number of input statistical parameters so 13 input Processing Elements are used in input layer. Four Processing Elements are used in output layer for four conditions of motor namely Healthy, Inter turn fault, Eccentricity and Both faults. For data processing MATLAB7.1, Neuro Solution 5.0 and XLSTAT are used. General learning algorithm used is as follows:

Initialization of Weights:

Step 1: Initialize the weights to small random values

Step 2: While stopping condition is false, do step 3-10

Step 3: For each training pair do steps 4-9

Feed forward:

Step 4: Each input unit receives the input signal x_i and transmits this signals to all units in the hidden layer

Step 5: Each hidden unit ($z_j, j=1, \dots, p$) sums its weighted input signals

$$z_{-inj} = v_{oj} + \sum_{i=1}^n x_i v_{ij} \quad (8)$$

Applying the activation function $Z_j = f(z_{inj})$ here the activation function is $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ and sends this signal to all units in output units.

Step 6: Each output unit ($y_k, k=1, \dots, m$) sums its weighted input signals ,

$$y_{-ink} = w_{ok} + \sum_{j=1}^p z_j w_{jk} + \sum_{j=1}^p x_j v_{ij} \quad (9)$$

And applies its activation function to calculate the output signals $Y_k = f(y_{-ink})$ here the activation function is

$$\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x}) \quad (10)$$

Back Propagation Error:

Step 7: Each output unit ($y_k, k=1, \dots, m$) receives a target pattern corresponding to an input pattern error information term is calculated as

$$\delta_k = (t_k - y_k) f'(y_{-ink}) \quad (11)$$

Step 8: Each hidden unit ($z_j, j=1, \dots, p$) sums its delta inputs from units in the layer above

$$\delta_{-inj} = \sum_{k=1}^m \delta_k w_{jk} \quad (12)$$

The error information term is calculated as

$$\delta_j = \delta_{-inj} f'(z_{-inj}) \quad (13)$$

Updation of weight and Biases:

Step 9: Each output unit ($y_k, k=1, \dots, m$) updates its bias and weights ($j=0, \dots, p$)

$$w_{jk}(t+1) = w_{jk}(t) + \alpha \delta_k z_j + \mu [w_{jk}(t) - w_{jk}(t-1)] \quad (14)$$

Where α is learning rate and μ is momentum factor

And each hidden unit ($z_j, j=1, \dots, p$) updates its bias and weights ($i=0, \dots, n$)

$$v_{jk}(t+1) = v_{jk}(t) + \alpha \delta_j x_i + \mu [v_{ij}(t) - v_{ij}(t-1)] \quad (15)$$

Step 10: Test the stopping condition

Selection of Error criterion:

Supervised learning requires a metric, a measure of how the network is doing. Members of the ErrorCriteria family monitor the output of a network, compare it with some desired response and report any error to the appropriate learning procedure. In gradient descent learning, the metric is determined by calculating the sensitivity that a cost function has with respect to the network's output. This cost function, J , is normally positive, but should decay towards zero as the network approaches the desired response. The literature has presented several cost functions, in which p is to be define such as $p=1, 2, 3 \dots$ criterion is $L_1, L_2, L_3 \dots$

Components in the Error Criteria family are defined by a cost function of the form:

$$J(t) = \frac{1}{2} \sum_i (d_i(t) - y_i(t))^p \tag{16}$$

and error function:

$$J(t) = -(d_i(t) - y_i(t)) \tag{17}$$

Where $d(t)$ and $y(t)$ are the desired response and network's output, respectively. To select the correct error criterion various error criterion has been tested and results are shown in Fig. 2 and Fig. 3.

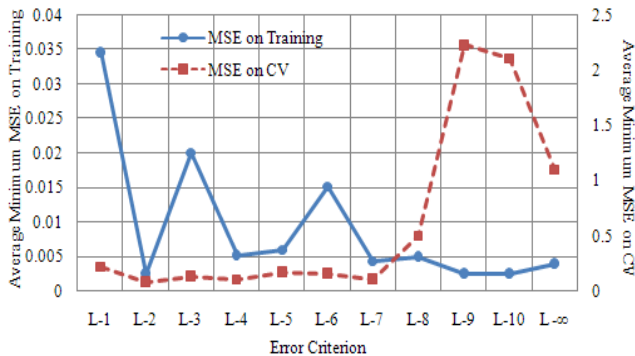


Fig. 2. Variation of Average Minimum MSE with Error Criterion

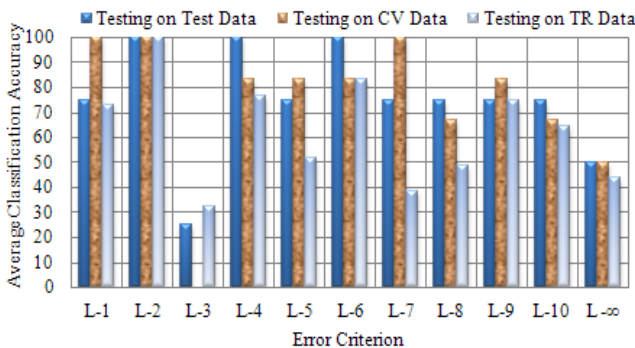


Fig. 3. Variation of Average Classification Accuracy with Error Criterion

For selection of hidden layer PEs randomize data is fed to the neural network and is retrained five times with different random weight initialization. It is observed that net-

work with single hidden layer and 5 PEs in hidden layer gives the better results as shown in Fig. 4.

Various transfer functions and learning rules namely Momentum (MOM), Conjugate-Gradient (CG), Quick Propagation (QP), Delta Bar Delta (DBD), Levenberg-Marquardt (LM) and Step (STP) are verified for training and testing the network. Average minimum MSE on training and CV data and classification accuracy on testing, CV and training data is compared in Table I and Table 2 (Appendix)

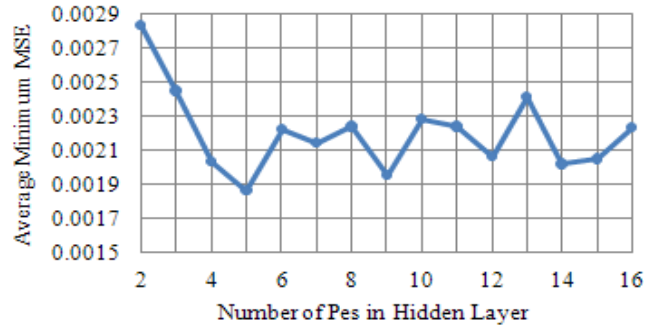


Fig. 4. Variation of Average Minimum MSE with Number of PEs in Hidden Layer

The parameters of the hidden layer and output layer i.e. stepsize and momentum are selected by comparing average minimum MSE. In Hidden layer optimum value of Step size is 0.13 and momentum is 0.6 and for output layer Step size is 0.05 and momentum is 0.08. Performance is shown in Fig. 5, Fig. 6 and Fig. 7

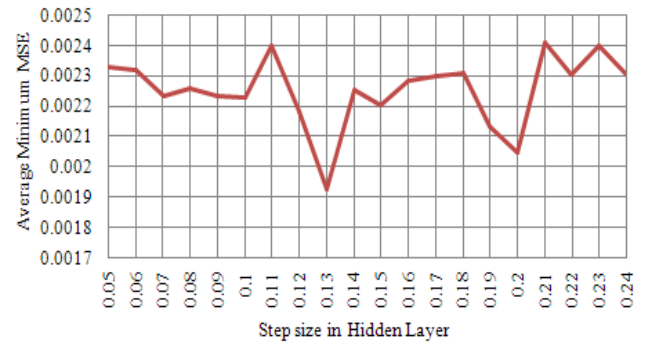


Fig. 5. Variation of Average Minimum MSE with Step size in Hidden Layer

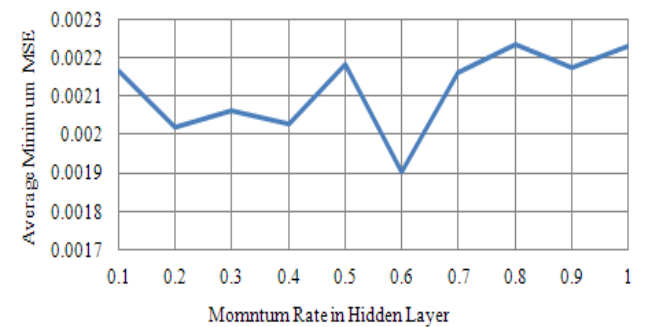


Fig. 6. Variation of Average Minimum MSE with Momentum rate in Hidden Layer

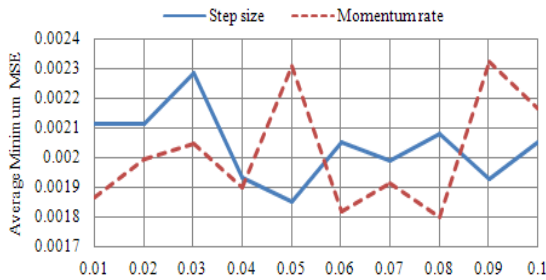


Fig. 7. Variation of Average Minimum MSE with Step size and Momentum rate in output Layer

From above experimentation, selected parameters for GFFD-NN are given below.

GFFD-NN (13-5-4), Number of epochs = 5000, Exemplars for training = 70%,

Exemplars for cross validation = 15%,

Exemplars for Testing = 15%

Number of Hidden Layers: 01

T.F.: Tanh Learning Rule: Momentum

Step size: 0.13 Momentum: 0.6

Output Layer:

T.F.: Tanh Learning Rule: Momentum

Step size: 0.05 Momentum: 0.08

Different datasets are formed using variable split ratios and leave-N-out cross validation technique. Proposed NN is trained on various datasets and later validated carefully so as to ensure that its performance does not depend on specific data partitioning scheme. The performance of the NN should be consistently optimal over all the datasets with respect to MSE and classification accuracy. To check the learning ability and classification accuracy the total data is divided in four groups. First two groups (50% data)

are tagged as Training data and third and fourth group (each 25%) is tagged for Cross Validation and Testing (1234:1,2-TR, 3-CV, 4-Test). Similar 18 combinations are prepared and network is trained and tested for each group. Results are shown in Fig. 8 and Fig. 9.

4. Dimensionality Reduction Using Principal Component Analysis

One problem appears after the feature extraction. There are too many input features that would require a significant computational efforts to calculate, and may result in low accuracy of the monitoring and fault diagnosis. The potential improvements which can be achieved by first mapping the data into a space of lower dimensionality. Reduction in dimensionality of the input space and hence the network can be achieved by Principal Component Analysis (PCA). PCA is performed by Pearson rule. The Fig. 10 is related to a mathematical object, the eigenvalues, which reflect the quality of the projection from the 13-dimensional to a lower number of dimensions.

Using the results of Principal Component Analysis, as shown in Fig. 11 and Fig. 12, dimensions of GFFD-NN can be reduced. Number of inputs is reduced to five. By similar experimentations, the optimum GFFD-NN classifier is designed with the following changes;

Number of Inputs = 5, Number of PE in Hidden Layer = 5, Number of epochs = 5000, Exemplars for training = 70%,

Exemplars for cross validation = 15%,

Exemplars for Testing = 15%,

Number of Hidden Layers: 01

T.F.: Tanh

Learning Rule: Momentum

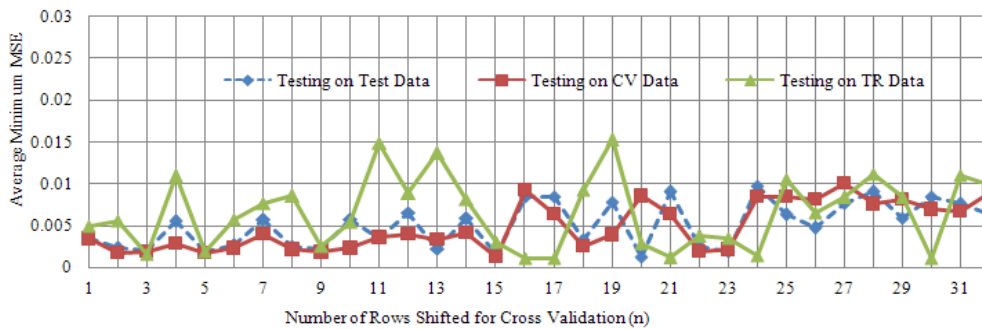


Fig. 8. Variation of average MSE with Test on Testing, CV and Training dataset with CV rows shifted (n)

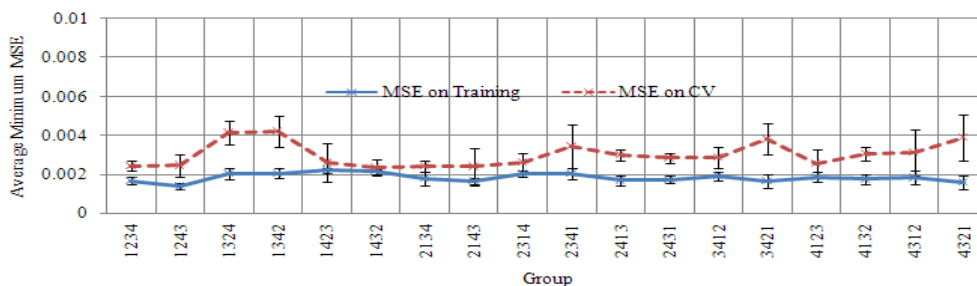


Fig. 9. Variation of average Minimum MSE with Training and CV with Different Groups of Dataset

Step size: 0.09 Momentum: 0.6
 Output Layer :
 T.F.: Tanh Learning Rule: Momentum
 Step size: 0.05 Momentum: 0.08

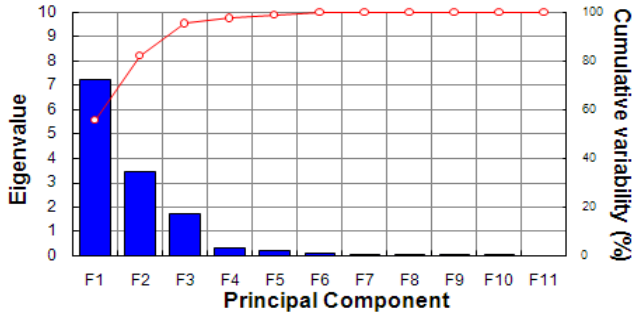


Fig. 10. Principal Component, Eigenvalues and percent variability

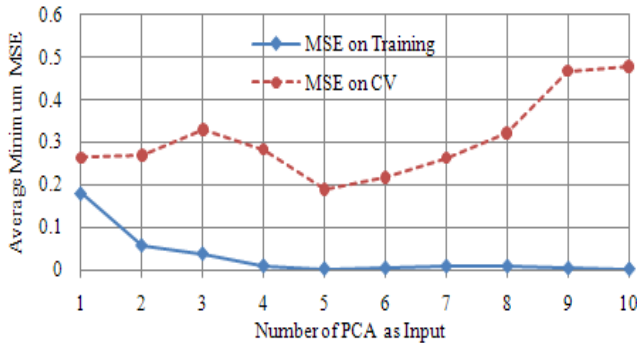


Fig. 11. Variation of Average minimum MSE of training and CV on number of PCs as inputs

Training and testing results for new model is as shown in Fig.13 to Fig.14.

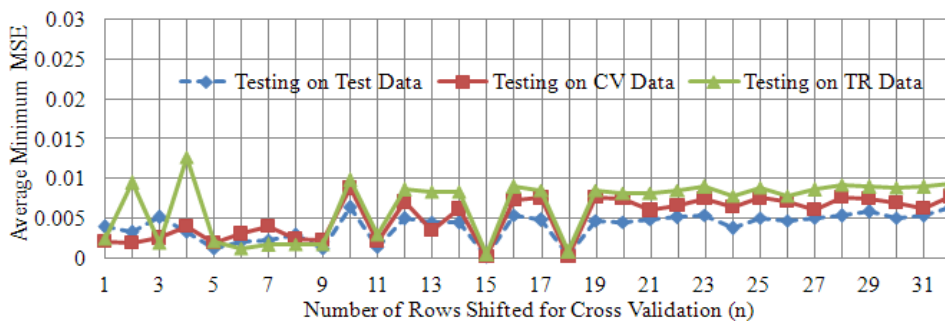


Fig. 13. Variation of average MSE with Test on Testing, CV and Training dataset with CV rows shifted (n)

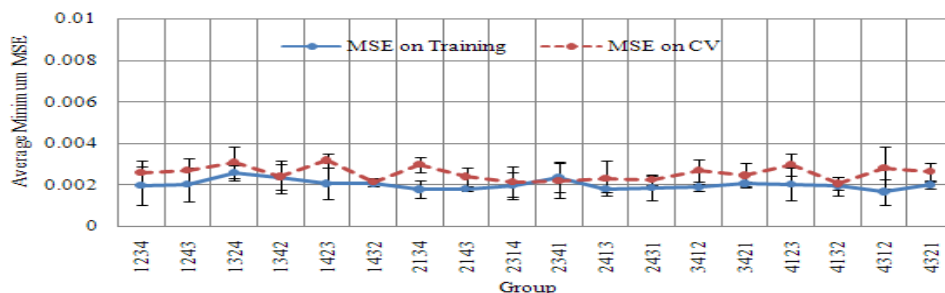


Fig. 14. Variation of average Minimum MSE with Training and CV with group of Dataset

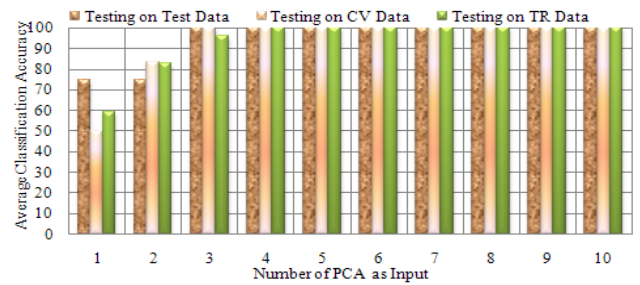


Fig. 12. Variation of average classification accuracy of Test, CV and TR data on number of PCs as inputs

5. Results and Discussion

In this paper, the authors evaluated the performance of the developed GFFD NN based classifier for detection of four conditions of three phase induction motor and examined the results. After completion of the training, the learned network is able to detect different types of faults. Various learning rules and transfer functions are investigated for different number of hidden layers and processing elements in hidden layer. It is observed that Momentum learning rule and tanh transfer function gives the optimal results. By performing Principal Component Analysis, number of inputs is reduced from 13 to 5 and thus dimension reduction is achieved. From the analysis, it is seen that designed network works as an elegant classifier for fault diagnosis of three phase induction motor, in the sense that, average MSE on testing and cross validation samples is consistently observed as reasonably low such as 0.05653 and 0.04624, respectively. In addition, average classification accuracy on testing as well as cross validation instances is obtained as 98.5% and 95.04%, respectively indicating a reasonable classification.

APPENDIX

TABLE 1. VARIATION OF AVERAGE MINIMUM MSE AND AVERAGE CLASSIFICATION ACCURACY WITH TRANSFER FUNCTIONS

TF	Average Minimum MSE		Average Classification Accuracy		
	Training	CV	Test	CV	TR
Tanh	0.002060616	0.116781132	100	100	100
Linear-Tanh	0.015755657	0.164229098	75	100	75
Linear-Sigmoid	0.366815041	0.362634201	75	66.6667	65.625
Sigmoid	0.356406224	0.364738836	50	66.6667	52.23214
Linear Axon	0.405	0.405	50	0	45.53571
Bias Axon	0.30375	0.470100566	21.875	50	21.875
Axon	0.128023849	0.543661969	50	83.3333	46.42857
Softmax	0.070957095	0.173723905	100	83.3333	86.16071

TABLE 2. VARIATION OF AVERAGE MINIMUM MSE AND AVERAGE CLASSIFICATION ACCURACY WITH TRANSFER FUNCTIONS

LR	Average Minimum MSE		Average Classification Accuracy		
	Training	CV	Test	CV	TR
MOM	0.002060616	0.116781132	100	100	100
DBD	0.106742522	0.2265947	50	100	64.28571
QP	0.002592499	0.220618682	100	83.3333	100
CG	0.099116255	0.255298026	100	83.3333	100
STP	0.003019834	0.188494211	100	83.3333	100
LM	0.008325295	0.316297565	75	100	75

References

- [1] R. Isermann, "Supervision, fault-detection and fault-diagnosis methods-An introduction," *Control Eng. Practice*, vol. 5, no. 5, pp.639-652, 1997.
- [2] S. Leohardt and M. Ayoubi, "Methods of fault diagnosis," *Control Eng.Practice*, vol. 5, no. 5, pp. 683-692, 1997.
- [3] R. Patton, P. Frank, and R. Clark, *Fault Diagnosis in Dynamic Systems,Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [4] M.-Y Chow, *Methodologies of using Neural Network and Fuzzy Logic for Motor Incipient Fault Detection*, Singapore: World Scientific, 1997.
- [5] M. -Y. Chow, R. N. Sharpe, and J. C. Hung, "On the application and design consideration of artificial neural network fault detectors," *IEEETrans. Ind. Electron.*, vol. 40, pp. 181-198, Apr. 1993.
- [6] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "Recent developments of induction motor drives fault diagnosis using AI techniques," *IEEE Trans. Ind. Electron.*, vol. 47, pp. 994-1004, Oct. 2000.
- [7] G. K. Singh and S. A. S. Al Kazzaz, "Induction machine drive condition monitoring and diagnostic research - a survey," *Electric Power Systems Research*, vol. 64, no. 2, pp. 145-158, 2003.
- [8] Jarmo Ilonen, Joni-Kristian Kamarainen, "Diagnosis Tool for Motor Condition Monitoring" *IEEE Transactions Industry Applications*, vol. 41, no. 4, pp 963-971 JULY/AUGUST 2005
- [9] Onel I Y , El Hachemi Benbouzid M., "Induction Motor Bearing Failure Detection and Diagnosis: Park and Concordia Transform Approaches Comparative Study" *IEEE Trans on Mechatronics*, vol. 13, pp 257-262 April 2008
- [10] Tian Han, Bo-Suk Yang, "Fault Diagnosis System of Induction Motors Based on Neural Network and Genetic Algorithm Using Stator Current Signals" *Hindawi Publishing Corporation International Journal of Rotating Machinery Volume 2006*, pp. 1-13 -2006
- [11] M. S. Ballal , Z. J. Khan , H. M. Suryawanshi and R. L. Sonolikar "Adaptive neural fuzzy inference system for the detection of inter-turn insulation and bearing wear faults in induction motor," *IEEE Trans. Ind. Electron.*, vol. 54, pp. 250, Feb. 2007
- [12] M. Knight and S. P. Bertani "Mechanical fault detection in a medium-sized induction motor using stator current monitoring," *IEEE Trans. Energy Convers.*, vol. 20, pp. 753, Dec. 2005



V.N.Ghate received the B.E. degree in Electrical Engineering from the Amravati University Amravati India, in 1990 and M.E. from Shivaji University Kolhapur India in 1993.

From 1994 to 2001 he was associated with Department of Electrical Engineering Government Polytechnic, Technical Education Government of Maharashtra, India.

From 2002 he is associated with Department of Electrical Engineering Government College of Engineering, Technical Education Government of Maharashtra, India

Prof. Ghate is member of IE(I),IETE(India), ISI and ISTE (India).

His area of interest is Neural Network, Electrical Machines and Drives.



S. V. Dudul was born on 28th of August 1964. He did his B.E., M.E. and Ph.D. in 1986 (Nagpur University, Nagpur), 1989 (Shivaji University Kolhapur) and 2003 (Amravati University, Amravati), respectively. He received his B.E. and M.E. degrees with Distinctions and Gold medals.

Since January 1995, he has been a faculty of Sant Gadge Baba Amravati University of Amravati.