Fault Diagnosis of Three Phase Induction Motor Using Neural Network Techniques

Vilas N. Ghate and Sanjay V. Dudul

Abstract-- **Fault diagnosis of induction motor is gaining importance in industry because of the need to increase reliability and to decrease possible loss of production due to machine breakdown. Due to environmental stress and many others reasons different faults occur in induction motor. Many researchers proposed different techniques for fault detection and diagnosis. However, many techniques available presently require a good deal of expertise to apply them successfully. Simpler approaches are needed which allow relatively unskilled operators to make reliable decisions without a diagnosis specialist to examine data and diagnose problems. In this paper simple, reliable and economical Neural Network (NN) based fault classifier is proposed, in which stator current is used as input signal from motor. Thirteen statistical parameters are extracted from the stator current and PCA is used to select proper input. Data is generated from the experimentation on specially designed 2 Hp, 4 pole 50 Hz. three phase induction motor. For classification, NNs like MLP, SVM and statistical classifiers based on CART and Discriminant Analysis are verified. Robustness of classifier to noise is also verified on unseen data by introducing controlled Gaussian and Uniform noise in input and output.**

Index Terms-- **Induction motor, Fault diagnosis, MLP, SVM, CART, Discriminant Analysis, PCA**

I. Introduction

INDUCTION motors play an important role as prime movers in manufacturing, process industry and transportation due to their reliability and simplicity in construction. In spite of their robustness and reliability, they do occasionally fail, and unpredicted downtime is obviously costly [1], [2] hence they required constant attention. The faults of induction motors may not only cause the interruption of product operation but also increase costs, decrease product quality and affect the safety of operators. If the lifetime of induction machines was extended, and efficiency of manufacturing lines was improved, it would lead to smaller production expenses and lower prices for the end user. In order to keep machines in good condition, some techniques i.e., fault monitoring, fault detection, and fault diagnosis have become increasingly essential [5]. The most common faults of induction motors are bearing failures, stator phase winding failures, broken rotor bar or cracked rotor end-rings and air-gap irregularities [6]. Different approaches for motors incipient fault detection and diagnosis have been successfully proposed [7]-[14]. Most of these techniques involve vibration analysis and stator current analysis because they are easy to measure and highly reliable. With the development of AI systems, expert systems based on NN, fuzzy logic, fuzzy NN, have been employed in order to assist the fault detection task for correctly interpreting the fault data [15]-[19]. The application of artificial intelligence methods, like neural networks are rather easy to develop and to perform [20]-[23]. Neural networks can be applied when the information about the process is obtained by measurements, which later can be used in the training procedures of neural nets [24]. The main advantage of such solution is to obtain on-line information about the kind and the "size" of a fault without developing very complicated mathematical models. Neural detectors can be designed using the data acquired from simulation or experimental tests [25]-[36].

The objective of this research is to develop an alternative neural network based incipient fault-detection scheme that overcome the limitations of the present schemes in the sense that, they are costly, applicable for large motors, furthermore many design parameters are requested and especially concerning to long time operating machines, these parameters cannot be available easily. As compared to existing schemes, proposed scheme is simple, accurate, reliable and economical. This research work is based on real time data and so proposed neural network based classifier demonstrates the actual feasibility in a real industrial situation. Four different neural network structures are presented in this paper with all kinds of performances and about 100% classification accuracy is achieved.

II. Fault Classification Using NN

The proposed fault detection and diagnosis scheme consists of four procedures as shown in Fig. 1:

- Data collection & acquisition
- Feature extraction
- Feature selection
- Fault classification

A. Data Collection and Data acquisition

In this paper the most common faults namely stator winding interturn short (I), rotor dynamic eccentricity (E) and both of them (B) are considered.

Vilas N.Ghate is with the Govt.College of Engineering Maharashtra, India, +9107212550074; E-mail: vng786@ rediffmail.com.

Sanjay V. Dudul, is with Sant Gadge Baba Amravati University, Amravati, Maharashtra India. E-mail: svdudul@rediffmail.com

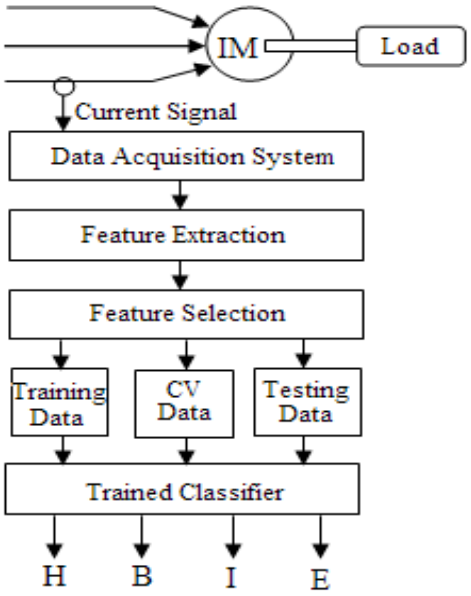


Fig. 1. General Block Diagram of proposed classifier

For experimentation and data generation the specially designed 2 HP, three phase, 4 pole, 415V, 50 Hz induction motor is selected. Experimental set up is as shown in Fig. 2

Fig. 2. Experimental Setup

The load of the motor was changed by adjusting the spring balance and belt. Three AC current probes were used to measure the stator current signals for testing the fault diagnosis system. The maximum frequency of used signal was 5 kHz and the number of sampled data was 2500. From the time waveforms of stator currents as shown in Fig.3, no conspicuous difference exists among the different conditions.

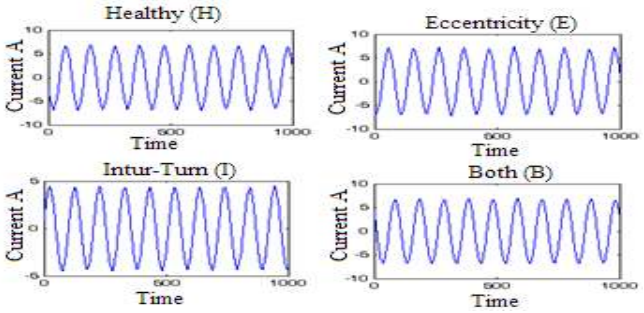


Fig. 3. Experimental Waveforms of Stator current

B. Feature Extraction

There is a need to come up with a feature extraction method to classify faults. In order 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 input feature space. Minimum set of statistics 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:

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

Where $\bar{x}$ denotes mean, x denotes median and S_x denotes the sample standard deviation. The sample coefficient of variation v_x is defined by;

$$v_x = \frac{S_x}{\bar{x}} \tag{2}$$

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

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

m_2 denotes spread about the center, m_3 refers to skewness about the center; m_4 denotes 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.

$$g_3 = \frac{m_3}{\left(\sqrt{m_2}\right)^3} \tag{4}$$

$$g_4 = \frac{m_4}{\left(\sqrt{m_2}\right)^4} \tag{5}$$

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

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

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

$$r_{jk} = \frac{c_{jk}}{S_j - S_k} \tag{7}$$

C. Feature Selection

Before a feature set is fed into a classifier, most superior features providing dominant fault-related information should be selected from the feature set, and irrelevant or redundant features must be discarded to improve the classifier

performance and avoid the curse of dimensionality. Here Principal Component Analysis (PCA) technique is used to select the most superior features from the original feature set. It is a technique of multivariate statistical analysis that can linearly or nonlinearly transform an original set of variables into a significantly smaller set of variables. It can be viewed as a classical method of multivariate statistical analysis for dimensionality reduction. Due to the fact that a small set of uncorrelated or independent variables is much easier to understand and use in further analysis than a larger set of correlated or dependent variables, this technique has been widely applied to virtually every substantive area including cluster analysis, visualization of high dimensionality data, regression, data compression and pattern recognition. In this paper, component analysis is used to extract the optimal feature and to reduce the dimension of original features. Principal Components (PCs) are computed by Pearson rule. The Fig.4 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.

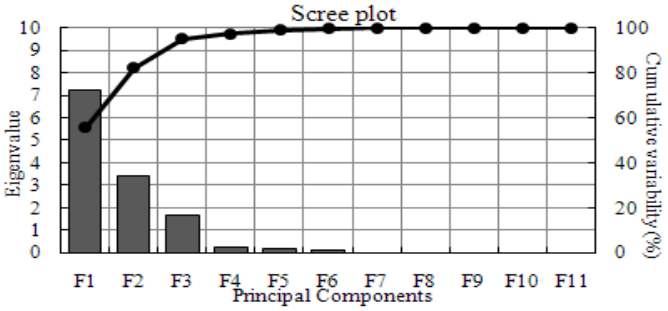


Fig. 4. Principal Component, Eigenvalues and percent variability

D. Fault Classifier

(1) MLP NN Based Classifier

Simple Multilayer Perceptron (MLP) Neural Network is proposed as a fault classifier. Four Processing Elements are used in output layer for four conditions of motor namely Healthy, Inter turn fault, Eccentricity and Both faults. Number of input Processing Elements (PEs) must be equal to that of number of inputs. To decide the number of inputs, number of PCs are given and on each input set, average minimum MSE and average classification accuracy are checked. From results as shown in Fig. 5, five PCAs are selected as inputs; hence number of PEs in input layer is five.

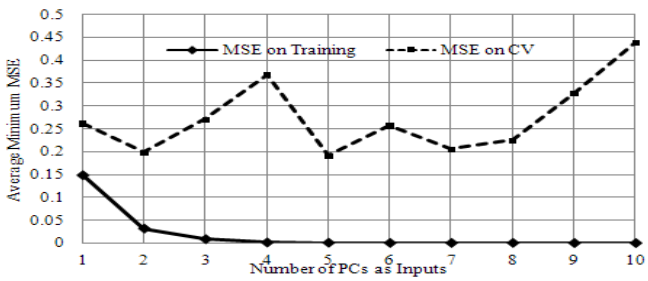


Fig. 5(a). Variation of Average MSE on training and CV with number of PCs as input

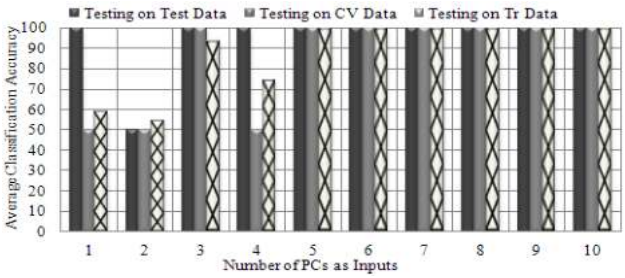


Fig. 5(b). Variation of Average Classification Accuracy on Testing on Test data, Training data and CV data with number of PCs as input

The randomized data is fed to the neural network and is retrained five times with different random weight initialization so as to remove biasing and to ensure true learning and generalization for different hidden layers. This also removes any affinity or dependence of choice of initial connection weights on the performance of NN. It is observed that MLP with a single hidden layer gives better performance. The number of Processing Elements (PEs) in the hidden layer is varied. The network is trained and minimum MSE is obtained when 5 PEs are used in hidden layer as indicated in Fig. 6.

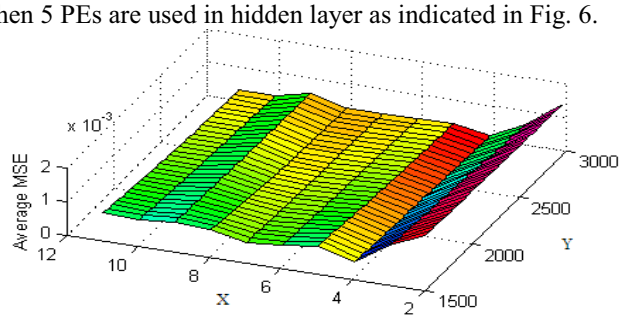


Fig. 6. Variation of Average MSE with number of PEs in Hidden Layer

Various Transfer functions, namely, Tanh, Sigmoid, Liner-tanh, Linear-sigmoid, Softmax, Bias axon, Linear axon and learning rules, namely, Momentum, Conjugate-Gradient, Quick Propagation, Delta Bar Delta, and Step are verified for training, cross validation and testing. Minimum MSE and average classification accuracy on training and CV data set are compared. It is found that Tanh transfer function and momentum learning rule give the optimum results. Stepsize and momentum of hidden layer and output layer is also varied for optimum average minimum MSE and average classification accuracy. Time elapsed per epoch per exemplar for each learning rule and transfer functions are calculated. For experimentation, desktop computer with Pentium-R-4, 3.0 GHz, 1GB RAM configuration is used. With above experimentations finally, the MLP NN classifier is designed with following specifications,

Number of Inputs: 5; Number of Hidden Layers: 01;
Number of PEs in Hidden Layer: 04;
Hidden Layer:

Transfer function: tanh	Learning Rule: Momentum
Step size: 0.6	Momentum: 0.5

Output Layer:

Transfer function: tanh	Learning Rule: Momentum
Step size: 0.1	Momentum: 0.5

Number of connection weights: 44

Training time required per epoch per exemplar: 0.0063 ms

Different datasets are formed using variable split ratios and leave-N-out cross validation technique. Leave-N-Out training is a technique that allows one to evaluate how well the model generalizes. It also is very useful for small data sets, since it allows one to use the entire data set for training and testing. The algorithm trains the network multiple times, each time omitting a different subset of the data and using that subset for testing. Proposed NN is trained and tested five times 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 forth 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. 7 and Fig. 8.

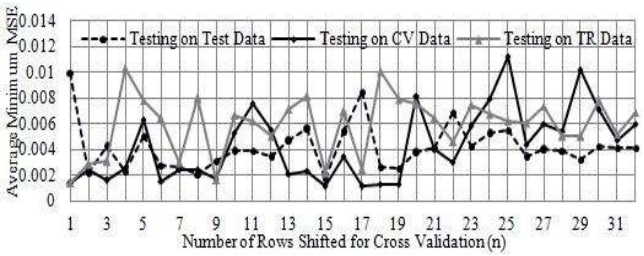


Fig. 7. Variation of Average Minimum MSE on Testing on Test data, CV data and Training data with number of rows shifted (n)

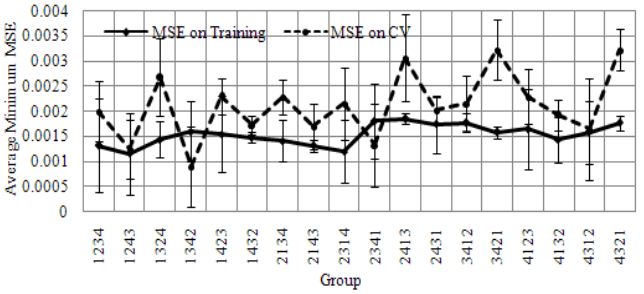


Fig. 8. Variation of Average Minimum MSE on Training and CV with various groups

(2) SVM NN Based Classifier

The support vector machine (SVM) is a new kind of classifier that is motivated by two concepts. First, transforming data into a high-dimensional space can transform complex problems (with complex decision surfaces) into simpler problems that can use linear discriminant functions. Second, SVMs are motivated by the concept of training and using only those inputs that are near the decision surface since they provide the most information about the classification. It is a kind of learning machine based on statistical learning theory. The basic idea of applying SVM to pattern classification can be stated as follows: first map the input vectors into one features space, possible in higher space, either linearly or nonlinearly, which is relevant with the kernel function. Then, within the feature space from the first step, seek an optimized linear division, that is, construct a hyperplane which separates two classes. It can be extended to multi-class. SVMs training always seek a global optimized solution and avoid over fitting, so it has ability to deal with a large number of feature.

Generalized Algorithm for the classifier:

For N dimensional space data $x_i \ (i = 1 \ldots N)$ this algorithm can be easily extended to network by substituting the inner product of patterns in the input space by the kernel function, leading to the following quadratic optimization problem:

$$J(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j G(x_i - x_j, 2\sigma^2) \quad (8)$$

Subject to

$$\sum_{i=1}^{N} d_i \alpha_i = 0 \qquad \alpha_i \geq 0, \forall i \in \{1, \ldots N\} \quad (9)$$

where $G(x, \sigma^2)$ represents a Gaussian function, N is the number of samples, α_i are a set of multipliers (one for each sample),

$$J(x_i) = d_i \left(\sum_{i=1}^{N} d_j \alpha_j G(x_i - x_j, 2\sigma^2) + b \right) \quad (10)$$

and

$$M = \min_i g(x_i) \quad (11)$$

and choose a common starting multiplier α_i, learning rate η, and a small threshold. Then, while $M >$ t, we choose a pattern x_i and calculate an update $\Delta\alpha_i = \eta(1 - g(x_i))$ and perform the update

If $\alpha_i(n) + \Delta\alpha_i > 0$

$$\alpha_i(n+1) = \alpha_i(n) + \Delta\alpha_i(n)$$
$$b(n+1) = b(n) + d_i \Delta\alpha_i \quad (12)$$

And if $\alpha_i(n) + \Delta\alpha_i \leq 0$

$$\alpha_i(n+1) = \alpha_i(n)$$
$$b(n+1) = b(n) \quad (13)$$

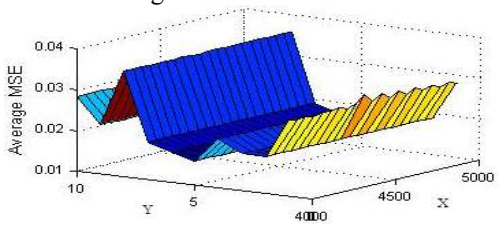
After adaptation only some of the α_i are different from zero (called the support vectors). They correspond to the samples that are closest to the boundary between classes. This algorithm can be considered the "on-line" version of the quadratic optimization approach utilized for SVMs, and it can find the same solutions as Vapnik's original algorithm for SVMs. It is easy to implement the kernel Adatron algorithm since $g(x_i)$ can be computed locally to each multiplier, provided that the desired response is available in the input file. In fact, the expression for $g(x_i)$ resembles the multiplication of an error with an activation, so it can be included in the framework of neural network learning. The Adatron algorithm essentially prunes the RBF network so that its output for testing is given by,

$$f(x) = \text{sgn}\left(\sum_{i \in \text{support vectors}}^{N} d_i \alpha_i G(x_i - x_i, 2\sigma^2) - b \right) \quad (14)$$

And cost function in error criterion is

$$J(t) = \frac{1}{2}\sum_{i=1}(d_i(t) - (\tanh(y_i(t)))^2 \qquad (15)$$

Number of PCs as input and step size is selected by checking the average minimum MSE and average classification accuracy; results are shown in Fig 9.



X-No. of PCs as input Y-No. of Epochs

Fig. 9(a). Variation of Average MSE on training and CV with number of PCs as input

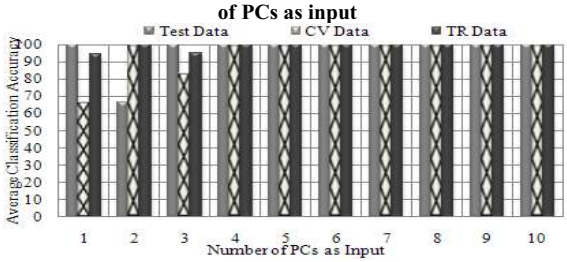


Fig. 9(b). Variation of Average Classification Accuracy on Testing on Test data, Training data and CV data with number of PCs as input

Finaly the SVM based classifier is designed with following specifications,

Number of Inputs: 5; Step Size: 0.7

Time required per epoch per exemplar: 0.693 ms

Number of connection weights: 264

Designed classifier is trained and tested using the similar datasets and results are as shown in Fig. 10 and Fig. 11

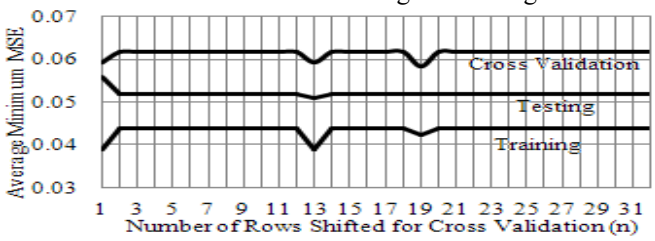


Fig. 10. Variation of Average Minimum MSE on Testing on Test data,CV data and Training data with number of rows shifted (n)

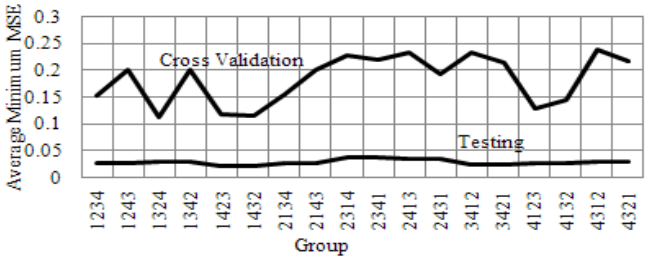
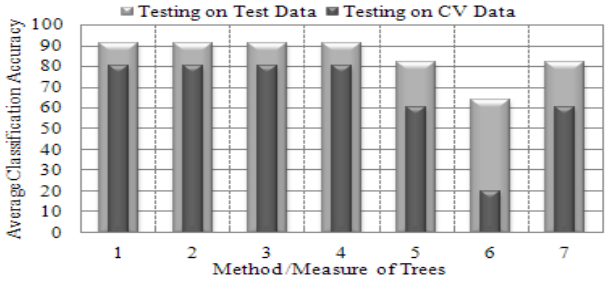


Fig. 11. Variation of Average Minimum MSE on Training and CV with various groups

(3) Classification and Regression Trees (CART)

CART induces strictly binary trees through a process of binary recursively partitioning of feature space of a data set. The trees produced by CART also consist of internal nodes (with two children) and terminal nodes or leaf nodes (without children). Each internal node is associated with a decision function to indicate which node to visit next, whilst each terminal node shows the output of a given input vector that leads the visit to this node. CART extensively builds the tree by using the data set of already classified instances which is called training set, and then prunes the tree back based on a minimum cost-complexity principle. The first phase is called tree building, and the other is tree pruning. Classification tree is developed using XLSTAT-2009. Various methods, measures and maximum tree depth are checked and results are shown in Fig. 12. It is observed that optimum average classification accuracy on testing on test data and CV data is found to be 90.91 and 80 percent, respectively.



1: Chaid-Pearson 2: Chaid-Likelihood
3: ExChaid-Pearson 4: ExChaid-Likelihood
5: C&RT-Gini 6: C&RT-Twoing 7: Quest

Fig. 12(a). Variation of Average Classification Accuracy on Testing on Test data and CV data with Method and Measure of Trees

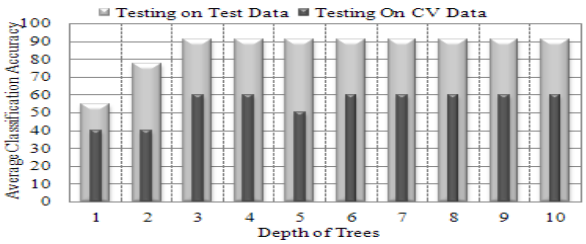
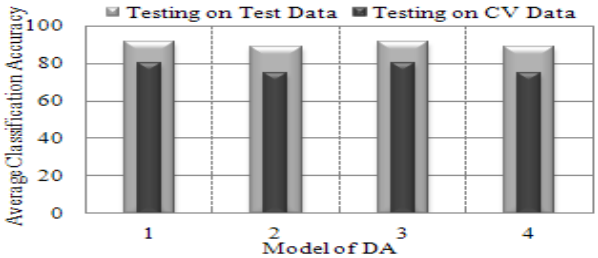


Fig. 12(b). Variation of Average Classification Accuracy on Testing on Test data and CV data with Depth of Trees

(4) Discriminant Analysis

Discriminant analysis is a technique for classifying a set of observations into predefined classes. The purpose is to determine the class of an observation based on a set of variables known as predictors or input variables. The model is built based on a set of observations for which the classes are known. This set of observations is sometimes referred to as the training set. Based on the training set , the technique constructs a set of linear functions of the predictors, known as discriminant functions, such that $L = b_1x_1 + b_2x_2 + \ldots + b_nx_n + c$, where the *b's* are discriminant coefficients, the *x's* are the input variables or predictors and c is a constant. These discriminant functions are used to predict the class of a new observation with unknown class. For a *k* class problem *k* discriminant functions are constructed. Given a new observation, all the *k* discriminant functions are evaluated and the observation is assigned to class *i* if the i^{th} discriminant function has the highest value. Discriminant analysis is done using XLSTAT-2009. Various models are checked and results

are shown in Fig. 13. It is observed that optimum average classification accuracy on testing on test data and CV data is found to be 91.77 and 80 percent, respectively.



1: Stepwise-forward 2: Stepwise Backward
3: Forward 4: Backward

Fig. 13. Variation of Average Classification Accuracy on Testing on Test data and CV data with Model of DA

III. NOISE SUSTAINABILITY OF CLASSIFIER

Since the proposed classifier is to be used in real time, where measurement noise is anticipated, it is necessary to check the robustness of classifier to noise. To check the robustness, Uniform and Gaussian noise with mean value zero and variance varies from 1 to 20 % is introduced in input and output and average classification accuracy on testing data i.e. unseen data is checked. It is seen that SVM based classifier is the most robust classifier in the sense that it can sustain both uniform and Gaussian noise with 14% and 20% variance in input and output, respectively. Results are as shown in Table I

G - Gaussian Noise

U - Uniform Noise

TABLE I

EFFECT OF NOISE ON AVERAGE CLASSIFICATION ACCURACY WHEN CLASSIFIER TESTED ON TESTING DATA

NN-Model	MLP				SVM			
Noise in	Input		Output		Input		Output	
% Variance	Average Classification Accuracy on Testing on Testing Data i.e. unseen Data							
Type of Noise	G	U	G	U	G	U	G	U
1	100	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100	100
3	100	100	100	100	100	100	66	100
4	100	100	100	100	66.7	100	100	100
5	100	100	100	100	66.7	66.7	100	100
6	100	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100	100
10	100	100	100	100	100	66.7	100	100
11	75	100	100	100	100	100	100	100
12	100	100	100	100	100	100	100	100
13	100	100	100	100	66.7	66.7	100	100
14	100	100	100	100	100	100	100	100
15	100	75	100	100	66.7	66.7	100	100
16	75	100	50	50	66.7	66.7	100	100
17	75	75	75	75	66.7	66.7	100	66.7
18	100	100	75	50	33.3	33.3	100	100
19	100	100	75	75	100	100	100	66.7
20	100	62	75	50	33.3	33.3	100	100

IV. RESULTS AND DISCUSSION

In this paper, the authors evaluated the performance of the developed ANN based classifiers for detection of four fault conditions of three phase induction motor and examined the results. MLP NN, and SVM are optimally designed and after completion of the training, the learned network is tested to detect different types of faults. For MLP NN 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 in hidden and output layer. Similarly step size is varied in SVM and 0.7 step size is found to be optimum. From the analysis, it is seen that support vector machine (SVM) based classifier 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 0.0591 and 0.0619, respectively. In addition, average classification accuracy on testing as well as cross validation instances is obtained as 99.61% and 98.72%, respectively indicating a reasonable classification. This might suggest that some of the features selected randomly contain too much fault unrelated information and there is a high degree of overlap between the values of these features of these four classes. These features would confuse the classifier and therefore, might cause significant performance degradation. These confirm our idea that the proposed feature selection method based on the PCA can select the most superior features from the original feature set, and therefore, is a powerful feature selection method. Also proposed classifier is enough robust to the noise, in the sense that classifier gives satisfactory results for Uniform and Gaussian noise with 14% variance in input and with 20% variance in output. Comparative results are shown in Fig.14 and Table II.

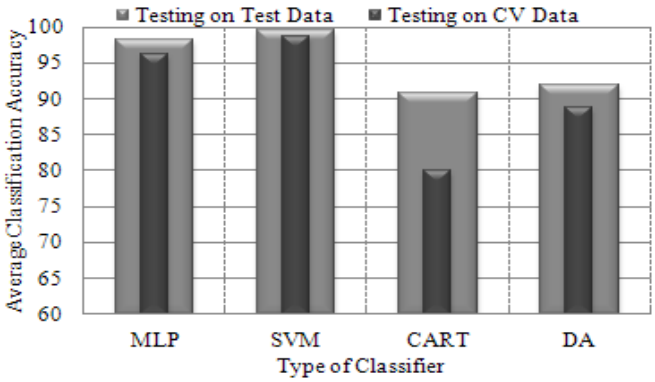


Fig. 14. Comparative analysis of various classifier w.r.t. Average classification accuracy.

REFERENCES

[1] Yang, B. S., Jeong, S. K., Oh, Y. M., & Tan, A. " Case-based reasoning system with Petri nets for induction motor fault diagnosis", *Expert Systems with Applications,* 27(2), 301–311,2004.

[2] 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.

[3] 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.

[4] A. H. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Ind. Appl. Mag.*, vol. 14, no. 1, pp. 29–36, Jan./Feb. 2008.

[5] W. T. Thomson, D. Rankin, and D. G. Dorrell, "On-line current monitoring to diagnose airgap eccentricity in large threephase induction motors—industrial case histories verify the predictions," *IEEE Transactions on Energy Conversion*, vol. 14, no. 4, pp. 1372–1378, 1999.

[6] B. Akin, U. Orguner, H. A. Toliyat, and M. Rayner, "Low order PWM inverter harmonics contributions to the inverter-fed induction machine fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 610–619, Feb. 2008.

[7] Acosta, G. G., Verucchi, C. J., & Gelso, E. R. "A current monitoring system for diagnosis electrical failures in induction motors". Mechanical Systems and Signal Processing, 20, 953–965,2006

[8] Benbouzid, M. E. H., & Nejjari, H., "A simple fuzzy logic approach for induction motors stator condition monitoring", *Proceedings of the IEEE IEMDC-2001* (pp. 634–639), 2001.

[9] Casimir, R., Boutleux, E., Clerc, G., & Yahoui, A., "The use of feature selection and nearest neighbors rule for faults diagnosis in induction motors." *Engineering Applications of Artificial Intelligence*, 19, 169–177, 2006.

[10] Satish, B., & Samar, N. D. R. , "A fuzzy approach for diagnosis and prognosis of bearing faults in induction motors", *IEEE Power Engineering Society General Meeting*, 3, 2291–2294,2005.

[11] Widodo, A., Yang, B. S., & Han, T., "Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors", *Expert Systems with Applications*, 32(2), 299–312, 2007.

[12] Ye, Z., Sadeghian, A., & Wu, B., "Mechanical fault diagnosis for induction motor with variable speed drives using adaptive neuro-fuzzy inference system. *Electric Power Systems Research*, 76, 742–752, 2006.

[13] Czeslaw T Kowalski & Teresa Orlowska-Kowalska , "Neural networks application for induction motor faults diagnosis", *Mathematics and Computers in Simulation ,63(3-5),* 435-448, 2003.

[14] M. Blodt, P. Granjon, B. Raison, and G. Rostaing, "Models for bearing damage detection in induction motors using stator current monitoring," *IEEE Trans. Ind. Electron.*, vol. 55, no. 4, pp. 1813–1822, Apr. 2008.

[15] J. Faiz, B. M. Ebrahimi, B. Akin, and H. A. Toliyat, "Finite-element transient analysis of induction motors under mixed eccentricity fault," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 66–74, Jan. 2008.

[16] M. S. Ballal , Z. J. Khan , H. M. Suryawanshi & 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.*, 54, 250. 2007.

[17] Van Tung Tran, Bo-Suk Yang, Myung-Suck Oh & Andy Chit Chiow Tan , "Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference", *Expert Systems with Applications*, *36(2),* 1840-1849,2009.

[18] F. Zidani, D. Diallo, M. E. H. Benbouzid, and R. Nait-Said, "A fuzzy based approach for the diagnosis of fault modes in a voltage-fed PWM inverter induction motor drive," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 586–593, Feb. 2008.

[19] J. Cusido, L. Romeral, J. A. Ortega, J. A. Rosero, and A. Garcia Espinosa, "Fault detection in induction machines using power spectral density in wavelet decomposition," *IEEE Trans. Ind .Electron.*, vol. 55, no. 2, pp. 633–643, Feb. 2008.

[20] M. Riera-Guasp, J. A. Antonino-Daviu, J. Roger-Folch, and M. P. Molina Palomares, "The use of the wavelet approximation signal as a tool for the diagnosis of rotor bar failures," *IEEE Trans. Ind. Appl.*, vol. 44, no. 3, pp. 716–726, May/Jun. 2008.

[21] Bo-Suk Yang, Tian Han & Yong-Su Kim (2004), "Integration of ART-Kohonen neural network and case-based reasoning for intelligent fault diagnosis" *Expert Systems with Applications*, *26(3),* 387-395,2004.

[22] 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.

[23] 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,* 1-13.2006.

[24] M. Blodt, D. Bonacci, J. Regnier, M. Chabert, and J. Faucher, "On-line monitoring of mechanical faults in variable-speed induction motor drives using the Wigner distribution," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 522–533, Feb. 2008.

[25] N. Arthur & J. Penman , "Induction Machine Condition Monitoring with Higher Order Spectra," *IEEE Trans. on Ind. Electronics, 47(5),* 1031-1041,2000.

[26] Yaguo Lei, Zhengjia He & Yanyang Zi, "A new approach to intelligent fault diagnosis of rotating machinery",*Expert Systems with Applications*, *35(4),* 1593-1600,2008.

[27] I. E. Alguindigue, A. L. Buczak, and R. E. Uhrig, "Monitoring and diagnosis of rolling element bearings using artificial neural networks," *IEEE Trans. Ind. Electron.*, vol. 40, pp. 209–217, Apr. 1993.

TABLE II
COMPARATIVE RESULTS OF NN BASED CLASSIFIERS
T - Time required per epoch per exemplar in ms W- Number of connection weights

N-N Model	Performance	Testing on Test Data				Testing on CV Data				T	W
		Max. Observed	Min. Observed	Average	SD	Max. Observed	Min. Observed	Average	SD		
MLP	MSE	0.207157	0.001773	0.046353	0.0655	0.126737	0.001135	0.02961	0.0413	0.0063	44
	Percent Correctness	100	83.33333	98.25	4.8423	100	83.33333	96.2222	6.0484		
	Percent Correctness	100	75	97.25	5.8690	100	83.33333	96.55	6.3382		
SVM	MSE	0.09926	0.05086	0.05915	0.011	0.094134	0.05454	0.06192	0.007	0.693	264
	Percent Correctness	100	88.88	99.611	1.944	100	88.88	98.722	3.514		