Improved Classification by Non Iterative and Ensemble Classifiers in Motor Fault Diagnosis

Parth Sarathi PANIGRAHY, Paramita CHATTOPADHYAY
Department of Electrical Engineering
Indian Institute of Engineering Science and Technology, Shibpur, Howrah, 711103, India
ps.panigrahy.rs2013@ee.iiests.ac.in, pc@ee.iiests.ac.in

***Abstract*—Data driven approach for multi-class fault diagnosis of induction motor using MCSA at steady state condition is a complex pattern classification problem. This investigation has exploited the built-in ensemble process of non-iterative classifiers to resolve the most challenging issues in this area, including bearing and stator fault detection. Non-iterative techniques exhibit with an average 15% of increased fault classification accuracy against their iterative counterparts. Particularly RF has shown outstanding performance even at less number of training samples and noisy feature space because of its distributive feature model. The robustness of the results, backed by the experimental verification shows that the non-iterative individual classifiers like RF is the optimum choice in the area of automatic fault diagnosis of induction motor.**

***Index Terms*—discrete wavelet transforms, fault diagnosis, feature extraction, induction motors, machine learning.**

I. Introduction

Condition monitoring of converter fed induction motors has been emerged as an active field of research due to their increasing demands in industrial applications. Among various faults associated with the induction motor, bearing (40%-50%), stator (30%-40%) and rotor (5%-10%) related mal operations are the most frequently occurred [1-3] phenomena. Most of the existing fault diagnosis schemes either use motor vibration [2, 4-9] or stator current [10-12] or both the signals [13] as raw data. It is clear from the literatures [2, 4-9] that the vibration signals are more fault informative than the stator current when dealt with an individual motor. However, the vibration based monitoring fault diagnosis method is very sensitive to noise and hence not suitable for industrial environment. On the other hand, the motor current signature analysis is considered as comparatively worthy because of its simplicity, availability of cheap and qualitative sensors and less sensitivity to noise. Motor current signature analysis (MCSA) based motor fault detection methods depend on either transient start up [14-16] or steady state currents [10-12]. It has been surveyed that the majority of research papers have primarily addressed rotor-related faults, followed by stator-related faults, and finally, with bearing faults using MCSA [1, 17, 18] based approach. The most frequently occurred faults such as bearing faults are less investigated by the research community. Identification of such defects play crucial role when a converter comes in the scenario and so more insight investigation of the problem is an essential task. Although the additional converter system improves the speed regulation and the dynamic response of the motor, it pollutes the current frequency spectrums. The characteristic frequency components for a specific faulty motor are obscured by the noise. So, the predefined rules for fault detection of induction motor become crucial to address the intrinsic harmonic impact.

The data driven approaches for fault diagnosis of induction motor are based on different state of art classifiers like multi-layer perceptron (MLP), radial basis function network (RBF), support vector machine (SVM) and k-NN etc.[8, 11-12, 19-20]. Few researches have also reported the application of decision tree or random forest (RF) as classifiers in the area of fault diagnosis of induction motor. However these machine learning techniques have mainly explored in the area of vibration based data driven approaches [2, 4-5, 7-9, 19-20]where the data pattern is relatively less complex in nature. In recent years ensemble classifiers are gaining importance in several application [21-23] for enhancing the classification accuracy. The main advantages of ensembles classifier are combined decision of several base classifiers to overcome the limitations of individual ones. The ensemble classifiers including stacking based knowledge driven process is capable to enhance the classification accuracy up to 2-4% than their individual base classifiers. This enhanced accuracy is achieved at the cost of huge computational complexity to a great extent. However ensemble classifiers have been rarely explored in condition monitoring of induction motors. In the data driven approach, raw vibration signals having complex waveform, produces good quality data in the feature space. On the other hand, raw current signal with relatively simple waveform in time domain, yields overlapping region among the various fault classes with complex data pattern in the feature space.

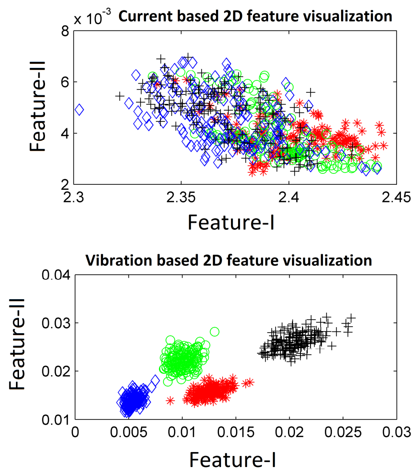

Figure 1. Feature visualization in best 2D space

This makes the crucial challenge in the area of stator

Digital Object Identifier 10.4316/AECE.2018.01012

current based fault classification of induction motors. The typical data pattern of current and vibration signals in their best 2D-feature space are presented in Fig. 1 respectively.

Keeping in mind these challenging issues in the area of fault diagnosis by stator current monitoring, the present investigation aims to solve complex data pattern of stator current in the feature space and to obtain an optimum solution for classifying the motor conditions. In this paper an exhaustive study of various advanced machine learning approaches including ensemble classifier with stacking process have been examined. The novelty of the paper lies on the intelligent use of inherent ensemble processing of non-iterative classifiers to recognize the complex pattern of the stator current based data set, without compromising with computational complexity, making the system suitable for industrial applications.

II. Machine Learning And Classifier Ensemble

Machine Learning is an important wing of data science. It manifests refinement and organization of data driven knowledge into more effective and useful way to improve the systems solving capability for a specific problem [19, 22,24]. Machine learning started its journey with primitive iterative classifiers like MLP, RBF etc. These methods target optimization of the objective functions designed heuristically. Hence they are time consuming, expensive and usually non-adaptive in nature. Selection of termination conditions is user specific and determined based on the values of the parameters involved in updating the objective functions. Hence the performance greatly depends on the initialization of the parameters.

Next generation machine learning approaches are non-iterative in nature. Non-iterative methods are nonparametric and exploit the inherent relationship within the dataset to classify the objects. They depend on the structural property within the data and therefore, easily tractable and adaptive in nature. Random Forest and k-NN classifiers belong to this category as they do not require any converging criteria. For these classifiers, the samples are selected randomly avoiding initialization process and for each case the outcomes are ensemble to prevent loss of information that may result due to random selection of the training data points.

The latest concept in the area of machine learning is ensemble classifiers. Ensemble process does not rely on single classifier rather uses several individual classifiers and combines their outputs in order to obtain an improved outcome as compared to the individual counterparts [21-23]. The individual classifiers whose outputs are combined for final decision are referred as base classifiers. The reliable and enhanced performance of ensemble method greatly depends on two important factors like diversity and combining rules of the base classifiers. The base classifiers may be iterative or non-iterative in nature.

There are various combination rules such as majority voting (MV), maximum probability (MP), average probability (AP), stacking etc. Ensemble processes like AP, MP are non-trainable, where the outputs of the base classifiers are combined through an algebraic expression. In each case, the final decision is the class that receives the largest support after applying algebraic operation on individual classifier output. In voting scheme (MV), the class assigned to a test instance is the one, suggested by most of the base level classifiers involved in the ensemble. The Voting based ensemble process selects the class that receives the largest vote contributed by the individual classifiers.

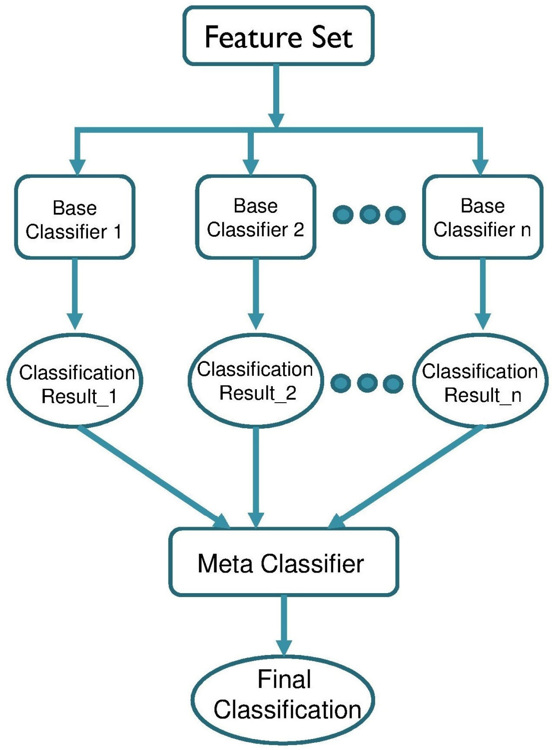

Figure 2. Stacked generalization algorithm

Unlike other combining rule, stacking based ensemble technique is an efficient knowledge driven approach where the prediction results of base classifiers are used as the input to another learning algorithm referred as meta-classifier. Hence it is a knowledge driven combination process. The pictorial representation of stacked generalization method is given in Fig. 2. The performance enhancement of the stacking based ensemble greatly depends on the meta level classifier and the features used to train the meta learner.

III. Proposed Scheme

In the present investigation, the advantages of non-iterative and ensemble classifiers, as discussed in the earlier section, have been exploited to resolve the bottlenecks of stator current based multiclass fault diagnosis of induction motor using data driven approaches. The broken rotor bar, stator inter-turn and bearing (inner and outer race defects) faults are investigated using stator current. The pipeline schematic starting from data acquisition to classification output is shown in Fig. 3. However the detail description of the data preparation including signal processing, feature extraction, feature selection etc. are explained in Sec. V and Sec. VI respectively.

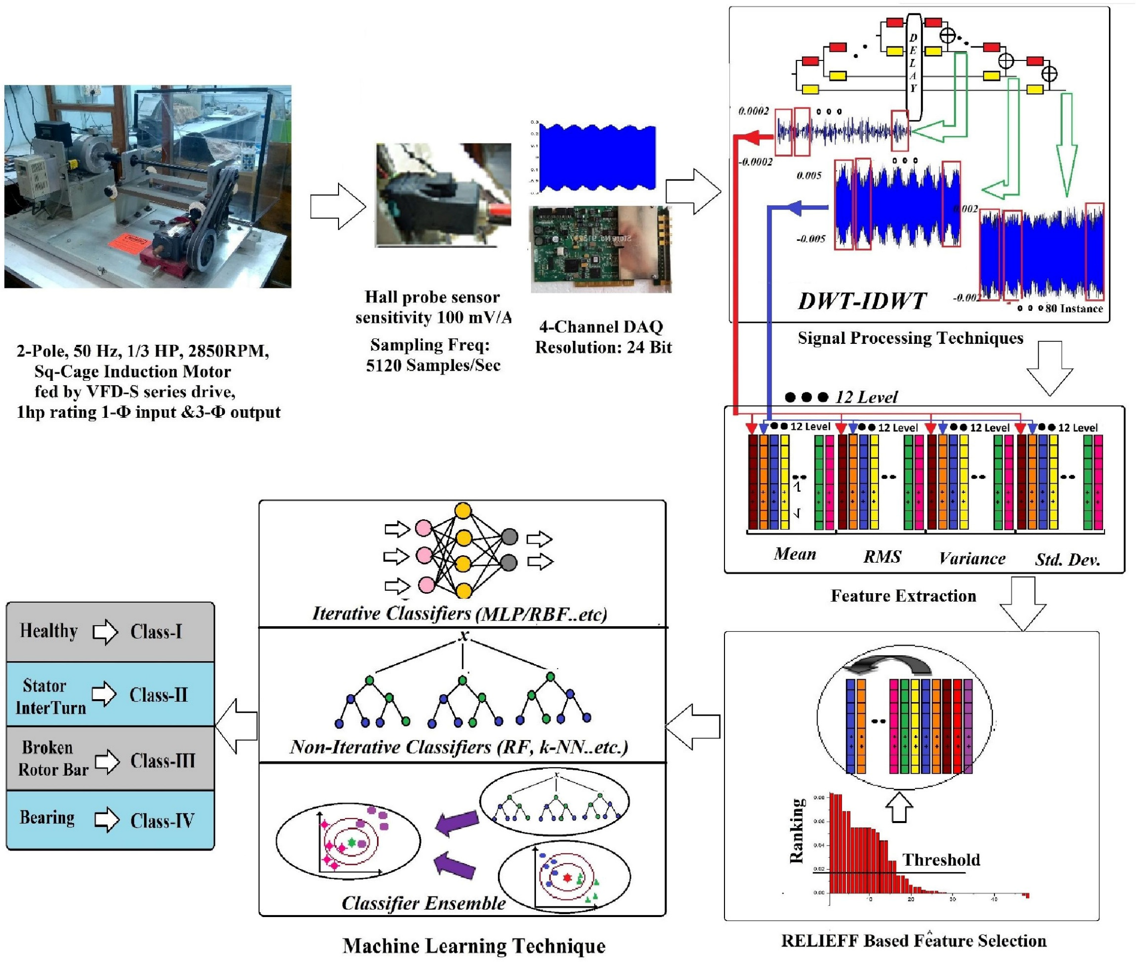

Figure 3. The schematic representation of the proposed approach

IV. Data Acquisition

The proposed data driven approach has been experimentally verified using Machinery Fault Simulator (MFS), a laboratory prototype from Spectra Quest deployed to emulate different induction motor faults [2, 22-23] like an industrial environment.

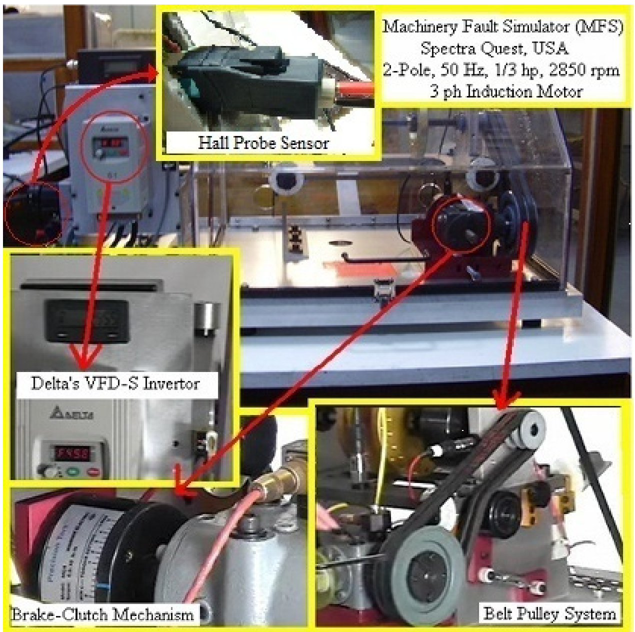

Figure 4. The experimental set up for motor fault analysis

Four 3 phase induction motors (bearing fault, stator inter-turn fault, broken rotor bar fault) of identical rating 1/3 H.P., 190 V, 50 Hz, 2980 rpm are used for the investigation. The motors are driven by a static induction motor drive of 1 hp rating. The steady state current signatures are captured at frequency of 5.12 kHz using Hall probe sensors (Model PR 30ACV, sensitivity 100mV/A). The whole experimental setup is shown in Fig. 4.

The data acquisition system has been realized using the National Instruments (NI) based hardware and its software has been developed in LabView for deterministic data acquisition. The input signal on each channel is buffered, conditioned, and then sampled by a 24 bit Delta- Sigma ADC with maximum sampling frequency of 102kS/s. One analog prefilter is used before the ADC to take care of the aliasing effect which may appear in the low frequency range.

V. Data Processing And Feature Extraction

In the present work, discrete wavelet transform (DWT)-inverse DWT(IDWT) has been used as the signal processing technique[11, 14-15, 19]. DWT is a continuous process of orthogonal decomposition of the signal information by using appropriate low and high pass filters. The decomposition process can be represented in terms of dilation equations of selected basis function, where the coefficient sequence represents the impulse response of filter bank. Therefore, the relationship of scaling function Φ(t) and wavelet function ψ(t) with filters can be written as Eq. (1) and Eq. (2) respectively.

$$\varphi_L = \sum_k h(k) 2^{\left(\frac{L+1}{2}\right)} \varphi\left(2^{L+1} t - k\right) \quad (1)$$

$$\psi_L = \sum_k g(k) 2^{\left(\frac{L+1}{2}\right)} \varphi\left(2^{L+1} t - k\right) \quad (2)$$

where, h(k) and g(k) are the impulse response of low and high pass filters respectively and these filters are called as analysis filters.

$$G_0(z) H_0(-z) + G_1(z) H_1(-z) = 0 \quad (3)$$

$$G_0(z) H_0(z) + G_1(z) H_1(z) = k(D) \quad (4)$$

$$k(D) = 2C_0 Z^{-D} \quad (5)$$

In the whole process of DWT-IDWT, the utilized filters satisfy two prime conditions given in Eq. (3) - Eq. (5) aiming cancellation of aliasing and perfect reconstruction respectively. $H_0(z)$, $H_1(z)$ are respectively z-domain response of low and high pass filters at analysis section and similarly, $G_0(z)$, $G_1(z)$ are at synthesis section. C_0 and D are constant and delay respectively. The stator current data samples are processed through 12-Level DWT-IDWT architecture where filtered data at different levels are used for feature extraction. In this study db10 based mother wavelet has been found effective in shape matching and used for de-composition of current signal. A typical 12 level decomposition of the raw stator current signal has presented in Fig. 5.

Keeping in mind, the importance of shape matching aspect of mother wavelet to extract embryonic fault features from the different sub bands, db 10 mother wavelet has been picked up for analyzing stator current signal. The sampling frequency f_{samp} of the stator current signal has been set to 5.12 kHz. The related fault frequency components [1] and the associated sub bands are furnished in Table I. The detail calculations of the characteristics fault frequencies are offered in Appendix-A.

In any data driven approach, feature extraction has a vital role. The standard statistical features like R.M.S, Variance, Mean and Standard Deviation are extracted from different sub bands of DWT-IDWT ranging 0.0625 Hz- 2560 Hz by exploiting its time frequency preservation quality. These features are computed for every 2048 data points from each of the filtered detail signal. The corresponding time span covering 2048 data points is 0.4 s ($2048/f_{samp}$) which represents one instance and all features are computed for 80 such independent instances. This time span is intentionally taken so as to cover at least one cycle of the minimum fault frequency as shown in Table I. Thus total 48 features (12-Level*4 Features) are extracted. Further to reduce the dimensionality and computation complexity, RELIEFF based feature selection technique is applied over 48 features where 24 best features are selected and used as attributes of the classifiers to categorize various motor conditions. Therefore, each feature pool in this work is denoted by informative 1600 data sets or instances (80 instances × 5 motor load × 4 motor cases) which are accountable by varying loading conditions and motor states.

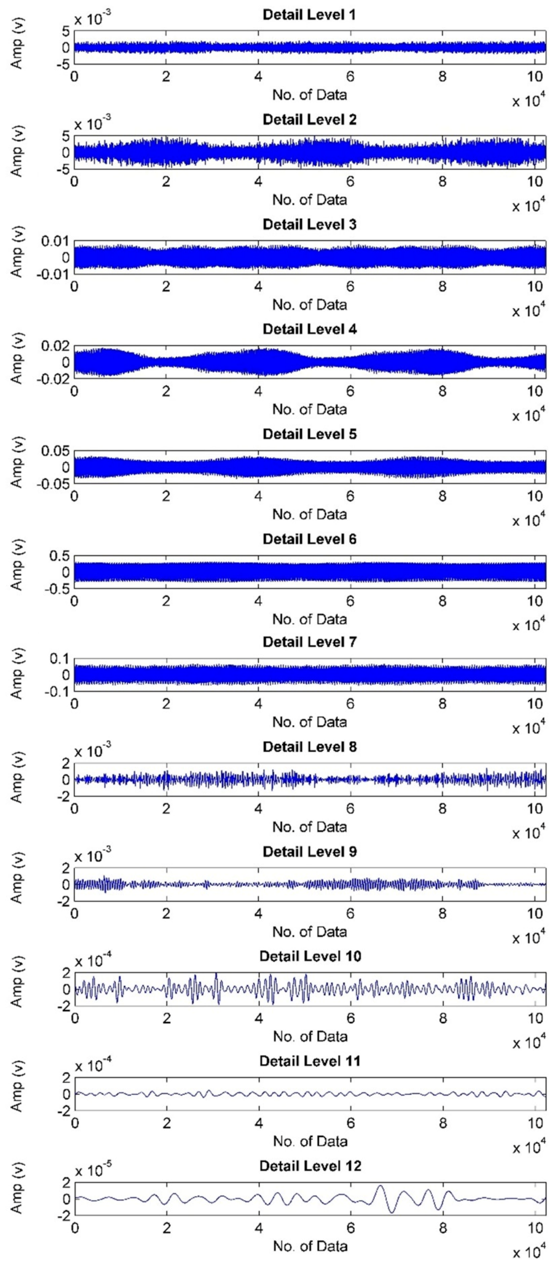

Figure 5. Reconstructed details of 12-level DWT-IDWT

The typical pattern of the data in two dimensional feature spaces is furnished in Fig 6. To measure the compactness and separation ability of the fault classes, Davies–Bouldin index (DBI) was calculated, which yields a value of 69.8530. Such a high value of DB index means uncluttered overlapping data, which is also reflected from the pictorial representation of the data distribution as in Fig. 6.

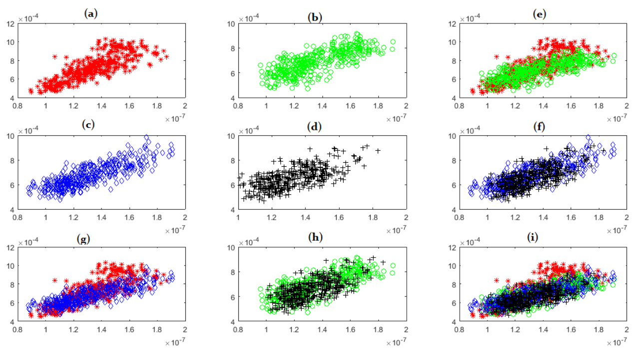

Figure 6. Typical data distribution of stator current in 2-D feature space, a.

Healthy(HLT), b. Stator Fault (STR), c. Broken Rotor Bar Fault (BRB), d. Bearing Fault (BRG) e. HLT and STR overlapping, f. BRB and BRG overlapping, g. HLT and BRB overlapping, h. STR and BRG overlapping, i. HLT, STR, BRB and BRG overlapping

TABLE I. FAULT FREQUENCIES AND DWT-IDWT SUB BANDS

Decomposed Level	Detail Bandwidth (Hz)	Fault characteristic frequencies (Hz)
1	[1280 2560]	-
2	[640 1280]	*f(brgir)*=655 & 755
3	[320 640]	*f(brgor)*= 339.8, 382 & 482; *f(brgir)*= 420 & 520
4	[160 320]	*f(brgor)*= 194.9,239.8; *f(brgir)*= 185 *f(str)*= 197.5
5	[80 160]	*f(brgor)*= 94; *f(str)*= 97.5 *f(str)*= 102.5
6	[40 80]	*f(brb)*=45 and 55 ; *f(brb)*=40 and 60 ; *f(brb)*=35 and 65
7	[20 40]	-
8	[10 20]	-
9	[5 10]	-
10	[2.5 5]	-
11	[1.25 2.5]	*f(str)*= 2.5
12	[0.625 1.25]	-

VI. FEATURE SELECTION

Feature selection is a dimensionality reduction technique widely used for data mining and knowledge discovery. It allows elimination of irrelevant/redundant features, while retaining the underplaying discriminative features. Feature selection techniques not only reduces the dimensionality and computation complexity, it also improves the quality of data and hence performance of the machine learning system. In the present investigation RELIEFF based non-heuristic feature selection technique has been used. It sorts the features and assigns ranks based on their importance. Noise tolerance and immunity to the feature interaction properties of RELIEFF based feature selection techniques [26, 27] have attracted the investigators to apply this method in the present study.

The flowchart of the feature selection algorithm is presented in Fig. 7. The original feature pool with 48 features is ranked in accordance to their significance using RELIEFF based feature selection technique. Relevant weights or relevancy of the features ranging from 0.1 to -0.004 are furnished in Fig. 8. Different threshold weights are used in the feature selection process in this range with decremented steps of 0.001. Thus, the performance of several feature subsets determined by setting various threshold values are examined based on the outcome of the classifiers. Finally, an optimum feature set of 24 features has been picked up using the threshold weight of 0.005 without compromising any classification performances.

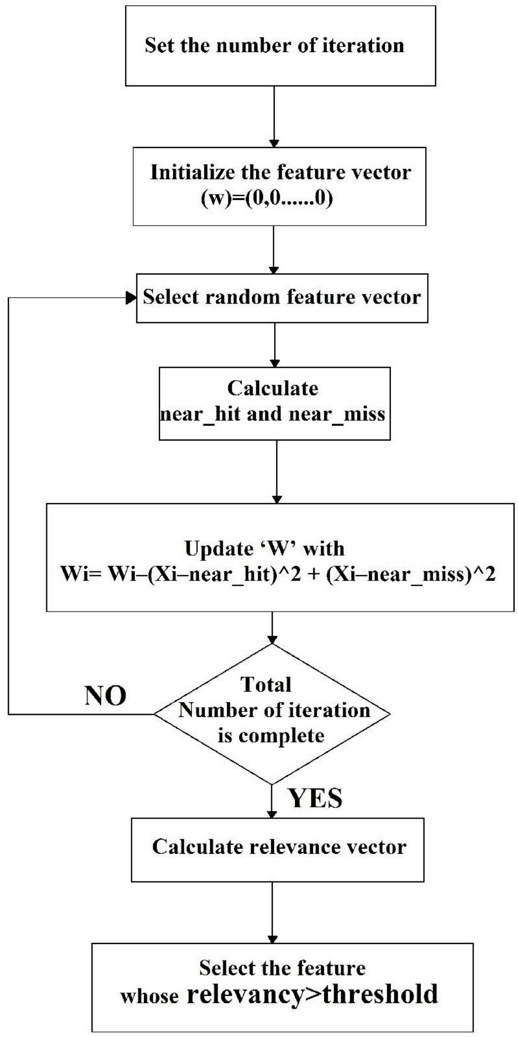

Figure 7. Flow chart of RELIFF feature selection technique

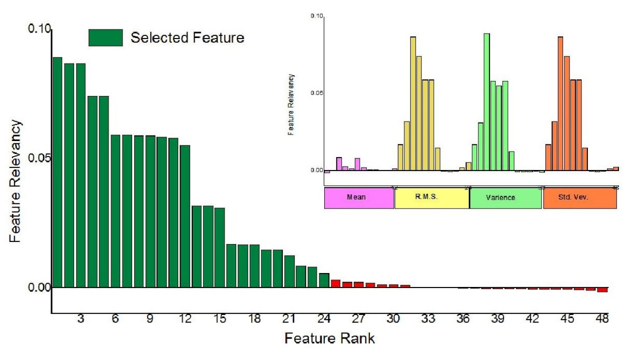

Figure 8. Relevant weights and the ranked features

The data distribution in two dimensional best feature spaces is presented in Fig. 9. DB index of selected data set is 28.7471. It indicates, the clustering quality of the data is better as compared with Fig. 6, due to the contribution of relevant and informative features only.

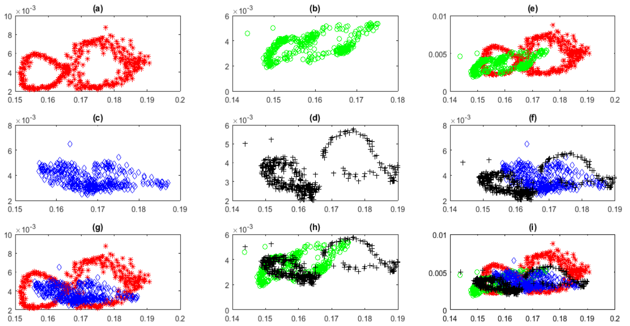

Figure 9. Data distribution of stator current in best 2-D feature space, a. Healthy(HLT), b. Stator Fault (STR), c. Broken Rotor Bar Fault (BRB), d. Bearing Fault (BRG), e. HLT and STR overlapping, f. BRB and BRG overlapping, g. HLT and BRB overlapping, h. STR and BRG overlapping, i. HLT, STR, BRB and BRG overlapping

The feature subset of 24 features and their allied DWT-IDWT decomposition levels are presented in Table II. Some noteworthy observations have been made such as: (a) Most of the features have been selected from decomposition levels: D2-D6, and these levels are exactly matching with the fault frequency as highlighted in Table I. (b) RMS and Standard deviation have been found as the most significant statistical features for the proposed fault classification problem.

TABLE II. THE MOST RELEVANT 24 FEATURES

Statistical Features From Decomposed Detail Level (D) And Ranks(R)							
Mean		**RMS**		**Variance**		**Standard Deviation**	
D	*R*	*D*	*R*	*D*	*R*	*D*	*R*
D3	R21	D1	R15	D2	R24	D1	R16
D6	R20	D2	R13	D3	R1	D2	R14
		D3	R2	D4	R10	D3	R3
		D4	R5	D5	R12	D4	R4
		D5	R8	D6	R11	D5	R6
		D6	R7			D6	R9
		D7	R23			D7	R22
		D11	R19			D12	R18
		D12	R17				
2 Attributes 8.33%		9 Attributes 37.5 %		5 Attributes 20.83 %		8 Attributes 33.33%	

VII. PERFORMANCE ANALYSIS

The performances of diverse classifiers with different machine learning algorithms have been evaluated for detecting motor faults using the experimental date set obtained from the stator current signals. In iterative category MLP, RBF, NvB and in non-iterative category; SVM, k-NN, RF have been explored. The effects of ensemble classifiers in iterative and non-iterative categories are rigorously studied. The popular ensemble process of voting based combining rules like AP, MV PP and MP along with knowledge driven staking method have been attempted here.

The parameters of the different classifiers have been tuned and the optimum classification accuracies of the individual classifiers with original 48 feature sets are furnished in the Table III whereas Table IV depicts the performance the classifiers after removing the irrelevant features by applying RELIEFF based feature selection technique as discussed in Sec. VI. The performances are evaluated using 10-fold cross validation [25] of WEKA Software.

TABLE III. PERFORMANCE OF INDIVIDUAL CLASSIFIER (48 FEATURES)

Motor Condition	**Accuracy(%)_Iterative**			**Accuracy(%)_Non-Iterative**		
	MLP	*RBF*	*NvB*	*k-NN*	*SVM*	*RF*
Healthy(HLT)	88.3	89	71.2	97.2	96.5	98.2
Stator(STR)	71.5	65	40.2	87	87.5	91
Rotor (BRB)	85.5	79.5	94	78.5	90.5	86.7
Bearing(BRG)	82	88	66.2	94.2	90.7	89.7
Average	81.8	80.4	67.9	89.2	91.3	93.1

TABLE IV. PERFORMANCE OF INDIVIDUAL CLASSIFIER (24 FEATURES)

Motor Condition	**Accuracy(%)_Iterative**			**Accuracy(%)_Non-Iterative**		
	MLP	*RBF*	*NvB*	*k-NN*	*SVM*	*RF*
Healthy(HLT)	88	88.2	74.7	98.2	98	98.5
Stator(STR)	76.2	67	45.2	90	92.7	92.2
Rotor (BRB)	85.7	88.7	86.7	82.7	92.7	94.2
Bearing(BRG)	79.5	89	76.7	95.2	93.2	90.5
Average	82.3	83.2	70.8	91.6	94.2	93.8

VIII. IMPACT OF NON-ITERATIVE CLASSIFIERS

The classification results shown in Table III unveil the impact of various machine learning algorithms where the massive discrimination has been observed. The individual performance of SVM, k-NN and RF using non iterative based error optimization technique outperform over iterative category. The classification accuracies of this category is really appreciable which are approximately 10% - 20% more, as compared to its iterative counterpart. It is important to note that, all fault classes including healthy condition have been classified more uniformly by SVM, k-NN and RF proving their robustness and satisfying the demand of any classification problem.

Operation of iterative classifier may encumbered by the criteria like minimum error, maximum number of iteration etc. Setting of these predefined criteria is very crucial to obtain a desire output. Also the iteration can stuck at local minima. This circumstance ultimately leads to a poor accuracy. In contrary non-iterative classifier do not require such criteria. Therefore the accuracy offered by the non-iterative classifier is comparatively higher than its iterative counterpart. However the main reasons for enhanced performance and highly adaptive behavior of non-iterative classifiers lie on the uses of local information of training set data near decision boundary and collective outputs of distributed feature sets. On the other hand iterative classifier gives decision based on single bulky feature set, by avoiding significant boundary data information, which may diminish the overall accuracy level.

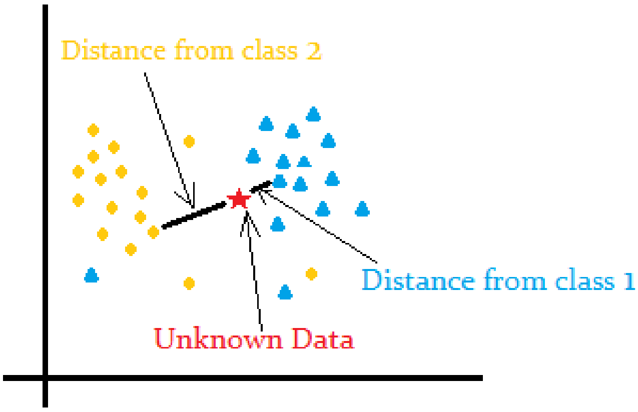

Figure 10. k-NN classification algorithm

For example, k-NN is a non-parametric classifier. A pictorial algorithm of k-NN is shown in Fig. 10. The input consists of the k number of (a positive integer, typically small) closest training examples in the feature space. An object is classified by the majority vote of its neighbors. Finally the object is assigned to the class most common among its k nearest neighbors. For complex data set under study, collective and more contributions of the nearer objects, than the distant ones, is the key factor for excellent performance of k-NN. The optimum classification results furnished in Table III and Table IV, has been obtained for k-NN when, k =5.

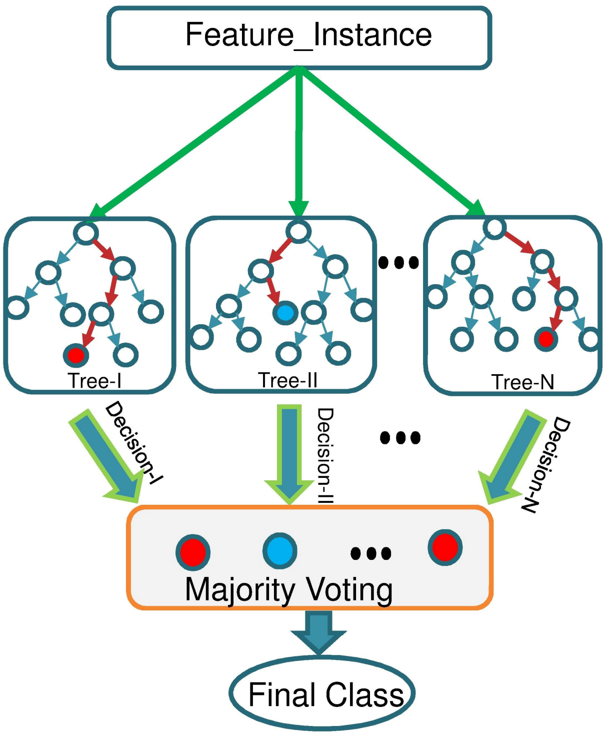

Figure 11. Random Forest algorithm

Random forest (RF) classifier was developed with the belief that the majority must be granted. Random forest learning algorithm decomposes a bulky decision tree into large number of small decision trees, called "forest", trained with randomly selected feature subset of same training data. The algorithm is presented in Fig. 11. The main motive of creating "forest" is to reduce the variance. If a single decision tree is developed with large number attributes then the tree becomes outsized, complicated and greatly affected by the irrelevant features.

RF provides the final class level of the object based on the majority decision of all the trees. So RF is very much suited for high dimensional features. Thus the built-in ensemble process of RF have helped to perform better over other classifiers including SVM and KNN, particularly for less number of training sample as in Fig. 12.

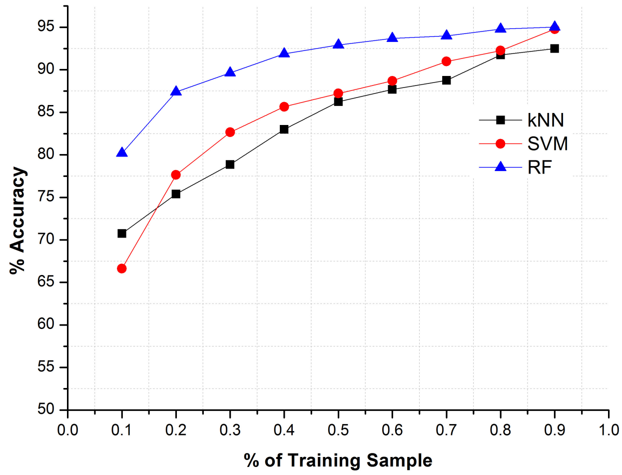

Figure 12. Performance of non-iterative classifiers with different % training samples

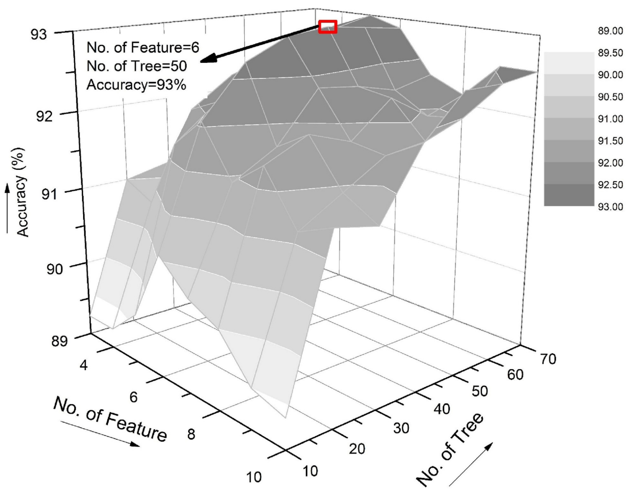

Figure 13. Feature-Tree based performance of random forest classifier

With reference to the Fig. 13, the best performances have been achieved for RF when number of trees is 50 and number of features is 6. Due to removal of irrelevant features, the classification accuracy of all the classifiers except RF, as presented in Table IV, have been improved after applying feature selection technique. This is because of the inherent feature selection process of RF. The impacts of ensemble classifiers both in, iterative and non-iterative categories are rigorously studied in the next section.

TABLE V. PERFORMANCE ANALYSIS OF ITERATIVE BASED ENSEMBLE CLASSIFIER IN %

Ensemble Classifier	MLP+RBF		MLP+NvB		RBF+NvB		MLP+RBF+NvB	
AP	85.3	HLT-87.2	79.0	HLT-82.6	79	HLT-81.7	83.8	HLT-86.1
		STR-85.3		STR-78.2		STR-76.9		STR-81.2
		BRB-82.5		BRB-78.6		BRB-80.2		BRB-82.3
		BRG-86.4		BRG-76.8		BRG-77.2		BRG-85.9
PP	85.3	HLT-87.1	80.0	HLT-81.2	80.3	HLT-80.6	83.6	HLT-85.6
		STR-86.2		STR-78.6		STR-78.3		STR-81.4
		BRB-83.6		BRB-77.8		BRB-81.4		BRB-82.2
		BRG-84.3		BRG-82.6		BRG-81.2		BRG-85.3
MV	85.3	HLT-85.6	79.0	HLT-79.4	79	HLT-80.6	83.7	HLT-84
		STR-85.1		STR-78.5		STR-78.2		STR-86.1
		BRB-86.6		BRB-78.3		BRB-81.9		BRB-81.5
		BRG-84.2		BRG-80		BRG-76.3		BRG-83.4
MP	85.3	HLT-84.7	78	HLT-80.1	78.5	HLT-81.1	80.1	HLT-79.8
		STR-85.1		STR-76.8		STR-78.2		STR-81.5
		BRB-85.6		BRB-77.3		BRB-78.5		BRB-78.7
		BRG-84.2		BRG-77.8		BRG-76.5		BRG-80.4
STACK	84.8	HLT-86.8	84.6	HLT-86.2	83	HLT-85.6	85.5	HLT-87.8
		STR-83.1		STR-83.1		STR-78.7		STR-84.5
		BRB-85.4		BRB-82.9		BRB-82.3		BRB-83.9
		BRG-84.2		BRG-86.3		BRG-85.4		BRG-85.8

TABLE VI. PERFORMANCE ANALYSIS OF NON-ITERATIVE BASED ENSEMBLE CLASSIFIER IN %

Ensemble Classifier	k-NN+SVM		k-NN+RF		SVM+RF		k-NN+SVM+RF	
AP	94.1	HLT-95.8	94.9	HLT-96.5	94.1	HLT-95.2	94.8	HLT-97.4
		STR-94.8		STR-94.2		STR-95.3		STR-94.1
		BRB-92.4		BRB-93.7		BRB-93		BRB-93.8
		BRG-93.7		BRG-95.3		BRG-93.2		BRG-94.2
PP	94.1	HLT-96.9	94.8	HLT-96.4	94.1	HLT-94.2	94.1	HLT-94.6
		STR-93.5		STR-94.1		STR-95.2		STR-91.3
		BRB-93.7		BRB-93.7		BRB-93.1		BRB-95.5
		BRG-94.6		BRG-95.3		BRG-94.2		BRG-95.3
MV	94.1	HLT-95.8	94.9	HLT-95.6	94.1	HLT-96.2	95.3	HLT-97.7
		STR-94.8		STR-95.1		STR-94.3		STR-92.5
		BRB-92.4		BRB-94.5		BRB-95.2		BRB-94.9
		BRG-93.7		BRG-94.5		BRG-91		BRG-96.1
MP	94.1	HLT-95.8	94.7	HLT-95.6	94.1	HLT-94.9	94.1	HLT-94.9
		STR-94.9		STR-94.5		STR-93		STR-93.6
		BRB-92.4		BRB-93.9		BRB-93.8		BRB-92.7
		BRG-93.6		BRG-95		BRG-95		BRG-96.1
STACK	94.1	HLT-94.4	95.3	HLT-97.1	95.4	HLT-97.6	95.6	HLT-97.6
		STR-94		STR-95.5		STR-94.8		STR-95.7
		BRB-94.7		BRB-94.7		BRB-92.7		BRB-94.6
		BRG-95.6		BRG-93.9		BRG-96.6		BRG-94.6

IX. IMPACT OF ENSEMBLE CLASSIFIERS

The overall cross validation accuracy [25] of iterative classifiers as in Table III and IV are good looking, but few classes are poorly classified as highlighted in these Tables. This is predominant particularly for the bearing and stator faults. Moreover the classification results for each motor condition vary from classifier to classifier. Hence in this investigation, ensemble algorithm is dealt to handle such diverged classification performances and aimed to develop a robust classifier model. The popular voting based combining rule of AP, MV, PP and MP are used in both the cases and the obtained results are furnished in Table V and Table VI. The results obtained for ensemble process of iterative classifiers depict that each motor fault including healthy state has been more uniformly classified although the accuracy level could not improve in every cases as compared to the superior base classifier. The algebraic and voting based approaches are quite application specific and could not be generalized for the guaranteed improvement in classification accuracy. The main reasons for low accuracies of voting based ensemble classifiers are un-scaled scores of base learners due to use of different learner models. On contrary, RF also uses an ensemble process by combining the learning models of same family, with randomly exploited features. Thus bootstrapping which requires identical distributed subsamples work nicely for RF but not for voting based ensemble classifier. In this scenario, knowledge driven stacking methodology using meta-learner may be helpful. The same 24 features are used in the stacking based ensemble process and the obtained results for iterative and non-iterative classifiers are presented in Table V and Table VI.

Different meta-learners are dealt in the ensemble process and results are investigated and among those, only most accurate results are presented. The acquired accuracy level using stacked generalization method outperforms over other combining rules as shown in Fig. 14. An exhaustive studies have been made with all possible blending among iterative and non-iterative classifiers, yet no significant improvements in results are noticed using ensemble process as compared to their base non-iterative classifier. However the best two combinations among them are shown in Fig. 14.

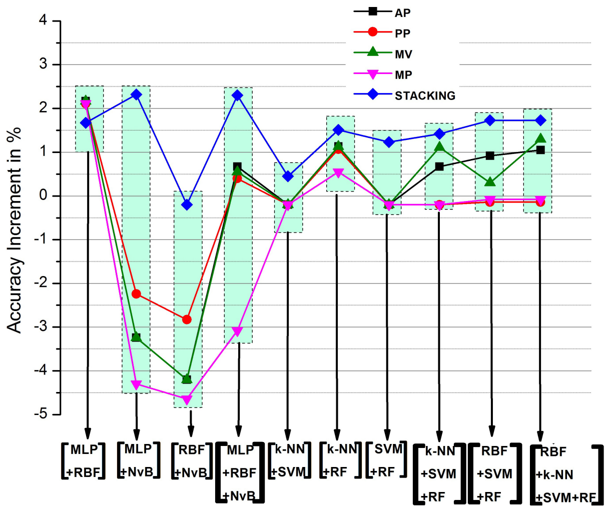

Figure 14. Improvement of accuracy for different ensemble methods

The out performance of stacking methodology is more prominent in case of non-iterative classifiers and these enhanced accuracies have been achieved at the cost of huge computational burden. The average model building time for various machine learning algorithm is presented in Fig. 15.

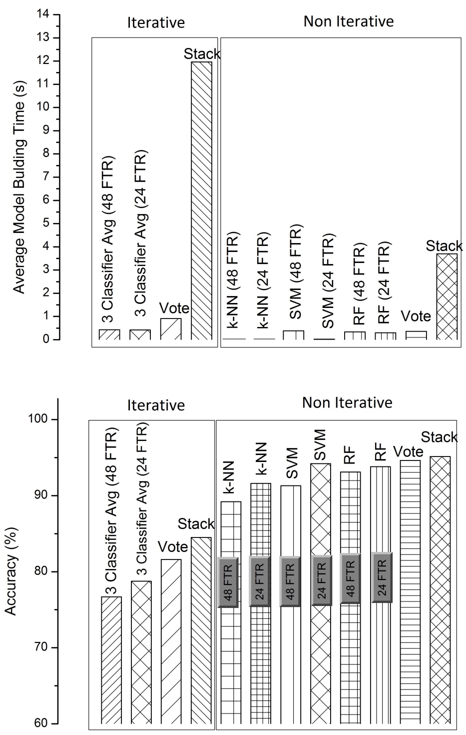

Figure 15. Complexity analysis of different machine learning approach

By comparing the Fig. 14 and Fig. 15, it is clear that non iterative classifiers like KNN, SVM and RF are better choice to handle the complex data pattern of current based fault diagnosis of induction motor. The inherent ensemble process of these classifiers as discussed in Sec. VIII is the main reason behind it.

X. CONCLUSION

An exhaustive experimental study on various machine learning techniques have been explored here in the light of complex pattern classification problem associated with fault diagnosis of induction motor using steady state stator current. In any data driven approach, classifiers and features are two crucial components to build up a robust model for practical applications. Keeping in mind the complexity of the data pattern in the present study, more powerful non iterative machine learning algorithms have been addressed and they yield decent classification accuracies. Due to built-in ensemble process of non-iterative classifiers, the cross validation accuracies have been achieved by 10-15% more than their iterative counterparts. Particularly RF classifier has shown extra ordinary learning ability for very less number of training samples even in the noisy feature space because of its distributive feature model. Hence RF is the optimum choice for industrial application of stator current based data driven fault diagnosis of induction motor. However in future to achieve a high level of generalization, deep learning based self-feature synthesis would be our motivation instead of hand crafted statistical feature based feature engineering.

APPENDIX A

In the present work, stator inter-turn fault, broken rotor bar fault and bearing fault (inner and outer race defects) are investigated using stator current. The related fault frequency components [1] are given in Eq (6)-Eq (8) respectively.

$$f_{str} = f_s(k_s \pm n(1-s)/p) \tag{6}$$

$$f_{brb} = f_s\left[1 \pm 2k_b s\right] \tag{7}$$

$$f_{brg(or/ir)} = | f_s \pm m.f_v | \tag{8}$$

where n = 1,2,3…(2p-1), k_s= 1,3, k_b=1,2,3.. m= 1,2,3…, p= No. of pole pair, s= slip in %, f_s= Supply frequency to the motor. Further f_v is the characteristic vibration frequency in Hz and dependent on the bearing design parameter and fault location as shown in Eq (9) and Eq (10) where N is number of ball in bearing, b_d is ball diameter, p_d is ball pitch diameter and β is contact angle with races.

$$f_{v(or)} = \left(N/2\right) f_r\left[1 - b_d cos\beta / d_p\right] \tag{9}$$

$$f_{v(ir)} = \left(N/2\right) f_r\left[1 + b_d cos\beta / d_p\right] \tag{10}$$

Using these parameters for the applied faulted bearing (N= 8, b_d=0.312, p_d=1.319 and β=9.08^0), the outer and inner race defect frequencies are *3.052*f_r* and *4.948*f_r* respectively where f_r is the rotor speed in rps. The motors are of 2 pole (p=1) having rated slip of 5% and operated at 50 Hz supply frequency (f_s).

Different fault frequencies are evaluated using Eq. (6) - Eq. (8) of Appendix-A and only few initial components are furnished in Table I of Sec. V. Other higher order fault components are not considered because of their feeble strengths. These evaluated fault frequencies are approximate values, as these are computed at rated slip of 0.05 only and the location can vary in small margin as per exact value of the slip.

ACKNOWLEDGMENT

We gratefully acknowledge the assistance, rendered by Deepjyoti Santra (Post Graduate student-2017, Dept. of E.E.) of IIEST Shibpur, for his thoughtful views.

REFERENCES

[1] S. Nandi, H. A. Toliyat, X. Li, "Condition monitoring and fault diagnosis of electrical motors-a review," IEEE trans. energy convers, vol. 20, pp. 719-729, 2005. doi:10.1109/TEC.2005.847955

[2] P. Konar, P. Chattopadhyay, "Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs)," Appl. Soft Computing, vol. 11, pp. 4203-4211, 2011. doi:10.1016/j.asoc.2011.03.014

[3] S. Sridhar, K. Uma Rao, S. Jade, "Detection of broken rotor bar fault in induction motor at various load conditions using wavelet transforms," IEEE Int. Conf. Recent Developments in Control, Automation and Power Engineering, 2015, pp. 77–82. doi:10.1109/RDCAPE.2015.7281373

[4] J. Seshadrinath, B. Singh, B. K. Panigrahi, "Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets," IEEE Trans. Power Electronics, vol. 29, pp. 936-945, 2014. doi:10.1109/TPEL.2013.2257869

[5] P. A. Delgado-Arredondo, A. Garcia-Perez, D. Morinigo-Sotelo, "Comparative Study of Time-Frequency Decomposition Techniques for Fault Detection in Induction Motors Using Vibration Analysis during Startup Transient," Shock and Vibration, vol. 2015, pp. 14, 2015. doi:10.1155/2015/708034

[6] E. Cabal-Yepez, M. Valtierra-Rodriguez, R.J. Romero-Troncoso, A. Garcia-Perez, R.A. Osornio-Rios, H. Miranda-Vidales, R. Alvarez-Salas, "FPGA-based entropy neural processor for online detection of multiple combined faults on induction motors," Mech. Sys. and Sig. Process, vol. 30, pp. 123-130, 2012. doi:10.1016/j.ymssp.2012.01.021

[7] P. Konar, J. Sil, P. Chattopadhyay, "Knowledge extraction using data mining for multi-class fault diagnosis of induction motor," Neurocomputing, vol. 166, pp. 14-25, 2015. doi:10.1016/j.neucom.2015.04.040

[8] P. Konar, P. Chattopadhyay, "Multi-class fault diagnosis of induction motor using Hilbert and Wavelet Transform," Appl. Soft Computing, vol. 30, pp. 341-352, 2015. doi:10.1016/j.asoc.2014.11.062

[9] J. Seshadrinath, B. Singh, B. K. Panigrahi, "Vibration analysis based interturn fault diagnosis in induction machines," IEEE Trans. Ind. Inf, vol. 10, pp. 340-350, 2014. doi:10.1109/TII.2013.2271979

[10] R. J. Romero-Troncoso, R. Saucedo-Gallaga, E. Cabal-Yepez, A. Garcia-Perez, R. H. Osornio-Rios, R. Alvarez-Salas, H. Miranda-Vidales, N. Huber, "FPGA-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference," IEEE Trans. Ind. Elect, vol. 58, pp. 5263-5270, 2011. doi:10.1109/TIE.2011.2123858

[11] J. Seshadrinath, B. Singh, B.K. Panigrahi, "Incipient interturn fault diagnosis in induction machines using an analytic wavelet-based optimized bayesian inference," IEEE trans. neural networks and learning sys, vol. 25, pp. 990-1001, 2014. doi:10.1109/TNNLS.2013.2285552

[12] J. Seshadrinath, B. Singh, B. K. Panigrahi, "Single-turn fault detection in induction machine using complex-wavelet-based method," IEEE Trans. Ind. Appl, vol. 48, pp. 1846-1854, 2012. doi:10.1109/TIA.2012.2222012

[13] H. Garcia-Perez, R. d. J. Romero-Troncoso, E. Cabal-Yepez , R. A. Osornio-Rios, "The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors IEEE Trans. Ind. Electronics, vol. 58, pp. 2002-2010, 2011. doi:10.1109/TIE.2010.2051398

[14] H. Ordaz-Moreno, R. d. J. Romero-Troncoso, J. A. Vite-Frias, J. R. Rivera-Gillen, A. Garcia-Perez, "Automatic Online Diagnosis Algorithm for Broken-Bar Detection on Induction Motors Based on Discrete Wavelet Transform for FPGA Implementation," IEEE Trans. Ind. Elec, vol. 55, pp. 2193-2202, 2008. doi:10.1109/TIE.2008.918613

[15] E. Cabal-Yepez, A. G. Garcia-Ramirez, R. J. Romero-Troncoso, A. Garcia-Perez, Roque A. Osornio-Rios, "Reconfigurable monitoring system for time-frequency analysis on industrial equipment through STFT and DWT," IEEE Trans. Ind. Inf, vol. 9, pp. 760-771, Jan. 2013. doi:10.1109/TII.2012.2221131

[16] H. Garcia-Perez, R. J. Romero-Troncoso, E. Cabal-Yepez, R. A. Osornio-Rios, J. d. J. Rangel-Magdaleno, H. Miranda, "Startup current analysis of incipient broken rotor bar in induction motors using high-resolution spectral analysis," IEEE Symp. Diagnostics for Electrical Machines, pp. 657-663, 2011. doi:10.1109/DEMPED.2011.6063694

[17] H. M. Knight, S. P. Bertani, "Mechanical fault detection in a medium-sized induction motor using stator current monitoring," IEEE Trans. Energy Conv, vol. 20, pp. 753-760, 2005. doi:10.1109/TEC.2005.853731

[18] P. Zhang, Y. Du, T. G. Habetler, B. Lu, "A survey of condition monitoring and protection methods for medium-voltage induction motors," IEEE Trans. Ind. Appl, vol. 47, pp. 34-46, 2011. doi:10.1109/TIA.2010.2090839

[19] P. S. Panigrahy, P. Konar, P. Chattopadhyay, "Application of data mining in fault diagnosis of induction motor," IEEE Int. Conf. Control, Measurement and Instrumentation, 2016, pp. 274-278. doi:10.1109/CMI.2016.7413754

[20] P. Konar, P. S. Panigrahy, P. Chattopadhyay, "Tri-Axial Vibration Analysis Using Data Mining for Multi Class Fault Diagnosis in Induction Motor," Int. Conf. Mining Intelligence and Knowledge Exploration Springer International Publishing, 2015, pp. 553–562. doi:10.1007/978-3-319-26832-3_52

[21] M. Kantardzic, Data mining: concepts, models, methods, and algorithms. Second Ed., John Wiley & Sons, 2011. doi:10.1002/9781118029145

[22] H. Jurek, Y. Bi, S.Wu, C. D. Nugent, "Clustering-based ensembles as an alternative to stacking," IEEE Trans. Knowledge and Data Eng, vol. 26, pp. 2120-2137, 2014. doi:10.1109/TKDE.2013.49

[23] J. Xia, L. Bombrun, T. Adalı, Y. Berthoumieu, C. Germain, "Spectral–Spatial Classification of Hyperspectral Images Using ICA and Edge-Preserving Filter via an Ensemble Strategy," IEEE Trans. Geoscience and Remote Sensing, vol. 54, pp. 4971-4982, 2016. doi:10.1109/TGRS.2016.2553842

[24] S. Džeroski, B. Ženko. Is combining classifiers with stacking better than selecting the best one?. Machine learning, pp. 255-273, 2004. doi:10.1023/B:MACH.0000015881.36452.6e

[25] S. Arlot, A. Celisse. A survey of cross-validation procedures for model selection. Statistics surveys, pp. 40-79, 2010. doi:10.1214/09-SS054

[26] Z. Deng, F. L. Chung, S. Wang, "Robust Relief-Feature Weighting, Margin Maximization, and Fuzzy Optimization," IEEE Trans. on Fuzzy Systems, vol. 18, pp. 726-744, 2010. doi:10.1109/TFUZZ.2010.2047947

[27] A. Ambarwari, Y. Herdiyeni and T. Djatna, "Combination of Relief Feature Selection and Fuzzy K-Nearest Neighbor for Plant Species Identification," Int. Conf. on Advanced Computer Science and Information Systems, 2016, pp. 315–320. doi:10.1109/ICACSIS.2016.7872767