

# Effect of noise on support vector machine based fault diagnosis of IM using vibration and current signatures

Purushottam Gangsar<sup>a</sup>, and Rajiv Tiwari<sup>b</sup>

<sup>a</sup>Department of Mechanical Engineering, SGSITS Indore, MP, 452003, India

<sup>b</sup>Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam, 781039, India

**Abstract.** This paper analyzes the effect of noise on support vector machine (SVM) based fault diagnosis of IM (IM). For this, a number of mechanical (bearing fault, unbalanced rotor, bowed rotor and misaligned rotor) and electrical faults (broken rotor bar, stator winding fault with two severity levels and phase unbalance with two severity levels) of IM are considered here. The vibration and current signals are used here for the diagnosis. Different experiments were performed in order to generate these signals at various operating condition of IM (Speed and Load). Time domain feature are then extracted from the raw vibration and current signals obtained from the experiments. Then, the noise are added in the raw signals and the same features are extracted from this corrupted signals. The features from the original and corrupted signals are used to feed the classifier. The one-versus-one multiclass SVM are used here to perform multi-fault diagnosis. The comparative analysis of performance of the SVM classifier using data with and without noise is presented.

## 1 Introduction

IMs (IMs) are one of the most widely used electric machines in all type of industrial applications like machine tools, petrochemical, textile, agriculture and power plants. They are most popular among other motors due to their mechanical simplicity, robustness, low cost and adaptability to wide variety of applications. Although IMs are reliable, they usually fall out of service due to severe operating conditions and dynamic loading conditions. A small motor fault can degrade the performance of the system and sometime permanent damage the system. Faults of IM can cause massive financial loss and serious human injury. Therefore, condition monitoring and fault diagnosis of IM is a significant work to early detect the fault to avoid such serious problems in industry [1]. The majority of research related to IM fault detection has concentrated on condition monitoring and detection of electrical fault i.e., stator winding faults and broken rotor bar faults [2]. Few research works have been reported on detection of mechanical faults in IM. However, chances of failure of IM is around 50 % due to mechanical faults like bearing faults, unbalanced rotor, bowed and misalignment rotor [2]. For fault detection of motor several noninvasive condition monitoring techniques have been used based on the current, voltage, acoustics, angular speed and vibrations [3]. Many papers have been reported on identification of IM faults using the current and the vibration; however,

they mostly deal with a single fault at a time [4]. However, in real machines, it is very common to have multi faults. In addition, various methods have been reported for fault diagnosis in IM such as, spectral analysis, discrete wavelet transform, Hilbert-hung transform, and Park's vector approach [5–7].

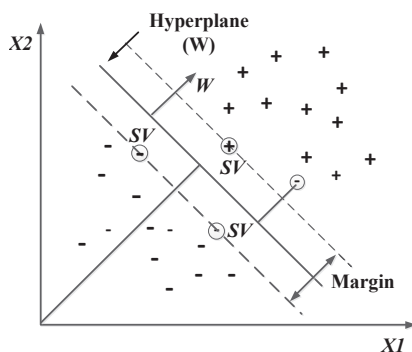
In last decades, various artificial intelligence methods, such as neural network, fuzzy logic, and combination of these two, have been applied to enhance the effectiveness of electrical machines fault diagnosis [8–9]. AI methods are data based methods and do not require any comprehension about the motor model and parameters. They only require a database of both faulty and healthy conditions for extraction of feature and fault diagnosis. Also these methods is used to take care of a number of faults at a time. Nowadays, the SVM is very popular AI method among others in fault diagnosis due to its attractive feature like higher generalization even with few data samples and lowest training time [10]. From literature, it is noted that almost all the researchers have been considered the IM fault diagnosis based on the data acquired in the laboratory and diagnosis have been found to be effective [11, 12, 13]. However, if the same fault diagnosis problem is considered for an industrial setting, there is almost always some difference between the training and testing distributions, even in the large-sample limit. These differences can be due to variations in operating conditions, changes in environment conditions, variations of fault types of same class, sensor noise and etc. This can degrade the performance of the classifier. Therefore, from industry point of view it is very important to consider the effect of noisy data on classifier performance. From literature review, it is evident that AI based faults diagnosis of IMs under noisy condition is still uncommon and has a lot of potential.

Therefore, in this work, the effect of noise on SVM based fault diagnosis of IM is explored. In this work, ten possible mechanical as well as electrical fault condition of IM have been considered to take care of multi-faults situation. The current as well as vibration signals have been used for fault diagnosis, to take care of individual effectiveness in detection of electrical and mechanical faults. The raw data is acquired in controlled condition in the laboratory using an experimental test rig. Then the noisy data is obtained by adding white Gaussian noise in the raw data. Then suitable features have been extracted from the raw and noisy data. Finally the diagnosis is considered when the raw data is used for training and testing of SVM then the same is considered for noisy data and a comparative analysis is done. The fault diagnosis has been performed and investigated for high load condition and a range of motor speed in order to check the robustness of it.

## 2 Introduction to SVM

Support vector machine (SVM) is first developed by Boser, Guyon and Vapnik in 1992 [14, 15]. The basic SVM is based on binary classifier and its principles can be illustrated in two-dimensional domain as shown in Figure 1. It shows the classification of data of two classes, class A (square) and class B (circle). Now, the SVM construct a hyperplane or set of hyperplane, which separates two classes and finally select an optimal separating hyperplane by maximizing the margin, namely, the distance between the closest data point of two classes. Data points which define the margin are known as support vectors. The function that corresponds to the hyperplane is linear for a linearly separable data. The basic SVM was developed for the binary classification problem; however by combining them, the multiclass classification problem can be solved. Hsu and Lin in 2002 [16] compared these three techniques and concluded that the one-against-one and direct-acyclic graph techniques are

competitive techniques. In this work, LIBSVM software [17] has been used for the multiclass fault diagnosis of IM using the one-verses-one technique.



**Figure 1.** Illustration of basic principal of SVM

### 3 Experimentation and data processing

The experiment was conducted on a machinery fault simulator (MFS) as shown in Figure 2. The MFS consist of a 0.5 hp, 50 Hz, 3-phase IM, coupled with a shaft using a flexible coupling, the shaft is coupled with a gearbox by using a pulley-belt drive. A variable frequency drive (VFD) and a magnetic brake with a gear box are attached to change the speed and external load of IM, respectively. Ten motors with different fault condition (i.e., no defect- ND, five electrical related, i.e. broken rotor bar (BRB), stator winding fault with severity level 1 and level 2 (SWF1 and SWF2, respectively), phase unbalance with severity level 1 and level 2 (PUF1 and PUF2, respectively), and four mechanical related, i.e. bearing fault (BF), unbalanced rotor (UR), bowed rotor (BR) and misaligned rotor (MR), are installed one by one on the MFS for generating the data. A tri-axial accelerometer and three AC current probes are used to acquire vibration signals in three orthogonal directions and three-phase motor current signals, respectively. A national instruments data acquisition system (DAQ) is used to acquire the data. A signal monitor with NI-Lab VIEW data acquisition software is used to analyze the data. The raw data were acquired in time domain with 1000 Hz sampling rate and 10000 sample points. In total 25 raw data-sets (25×10000 sample points) were collected for each faulty conditions. Measurements were taken for various speeds from 10 Hz to 40 Hz at 5 Hz interval for high external load (0.565 N-m, i.e. 0.55 % of rated torque).



**Figure 2.** Experimental test rig for data generation

For the fault diagnosis based on SVM, input features (or vector) must be selected appropriately. In order to overcome the problem with raw data, three effective statistical features, i.e. the standard deviation ( $\sigma$ ), skewness ( $\gamma$ ) and kurtosis ( $\kappa$ ) are now extracted from the raw time domain vibration and current signals. These features are very sensitive to any changes in IM faults as per literatures [18]. Now, the 2% additive white Gaussian noise is added to the raw time domain signal in order to produce corrupted signal. Three effective features are now extracted from the corrupted signals.

#### 4. Fault diagnosis based on SVM

In order to perform the fault diagnosis, first total feature data sets are divided into the training and testing data in a ratio of 80 % and 20 %, respectively. In this study, the Gaussian RBF kernel, which is one of the very well-liked and claimed to be very effectual kernel are used. Its parameters (i.e., regularization parameter,  $C$  and kernel parameter,  $\gamma$ ) are selected by training the classifier using training data based on the grid-search technique along with cross-validation method. The best pair of ( $C$ ,  $\gamma$ ) is selected corresponding to best training accuracy and by using this pair the final training is performed to build an optimal SVM model. This model would be used to perform final fault prediction using the test data sets. In order to classify different IM faults, different label is assigned to each fault such as: ND-1, BRB-2, PUF1-3, PUF2-4, SWF1-5, SWF2-6, BF-7, UR-8, BR-9, and MR-10. The testing is performed using test data, to determine the prediction accuracies of different faults.

Here two cases of fault diagnosis have been considered, i.e. first the diagnosis with the raw signal and second diagnosis with the corrupted signal. The diagnosis is performed for seven different IM speed. Table 1 shows the fault diagnosis results with the raw signal. The average of overall accuracies is found to be 99.71 %. The overall accuracies are 100 % at all speeds except 30 Hz speed (98 %). That means all the faults as well as considered severities are diagnosed perfectly. From here it can say that, the combination of  $\sigma$ ,  $\gamma$  and  $\kappa$  feature is capable of diagnosing IM faults for all speeds. In addition, the result indicates that proposed diagnosis does not depend on the motor speed that means diagnosis can be performed at any speed with good prediction performance. Table 2 shows the fault diagnosis with the corrupted signal. In this case, the average of overall accuracies is found to be 89.52 %. The overall accuracy is varying from 80 % (15 Hz) to 100 % (30 Hz). It is noted that the overall accuracy is comparatively lower below 20 Hz motor speed. This is because the fault conditions like SWF1, SWF2, UR and BR could not be diagnosed at lower speeds. The minimum overall accuracy is 91.67 % for speeds above 25 Hz. For this speed range, 1-9 % misclassification of features is occurred and this is mainly due to the misclassification of features of SWF2, However at lower speeds misclassification is occurred for SWF1, SWF2, UR and BR. In order to show the percentage data misclassification, a confusion matrix is added for 15 Hz in Table 3. From this table, it is found that the SWF2 data are misclassified with the data of SWF1. The reason may be, SWF1 and SWF2 are the two severity levels of the same fault i.e., stator winding fault. So if there is not large difference between the two severity levels of the fault, then there is less chances of producing distinct features in current signals. In addition, UR features are misclassified with the BR, this may be due to that they produces nearly similar vibration signature. Table 1 and Table 2 show the comparative analysis of prediction performance of data with no and 2 % external noise. From these tables it can be observed that, the overall prediction performance is reduced by 10 % in case of noisy data as compared to original raw data. Still the prediction performance is quite good at higher speeds in case of noise data. That means the proposed methodology can perform the diagnosis successfully (at higher speeds) even in noisy conditions.

**Table 1.** Prediction performance with data of no additive external noise

Train speed (Hz)	Test speed (Hz)	Prediction accuracy, %										
		ND	BR B	PUF 1	PUF 2	SWF 1	SWF 2	BF	UR	BR	RM	Overall
10	10	100	100	100	100	100	100	100	100	100	100	100
15	15	100	100	100	100	100	100	100	100	100	100	100
20	20	100	100	100	100	100	100	100	100	100	100	100
25	25	100	100	100	100	100	100	100	100	100	100	100
30	30	100	100	100	100	100	80	100	100	100	100	98
35	35	100	100	100	100	100	100	100	100	100	100	100
40	40	100	100	100	100	100	100	100	100	100	100	100
	Avg.	100	100	100	100	100	97.1	100	100	100	100	<b>99.71</b>

**Table 2.** Prediction performance with data of 2 % additive external noise

Train Speed (Hz)	Test Speed (Hz)	Prediction accuracy, %										
		ND	BR B	PUF 1	PUF 2	SWF 1	SWF 2	BF	UR	BR	MR	Overall
10	10	100	100	100	100	83.33	100	100	50	0	100	83.33
15	15	100	100	100	100	66.67	16.67	100	33.3	83.3	100	80
20	20	100	100	100	100	83.33	0	100	100	100	100	88.33
25	25	100	100	100	100	100	16.67	100	100	100	100	91.67
30	30	100	100	100	100	100	100	100	100	100	100	100
35	35	100	100	100	100	100	33.33	100	83.3	100	100	91.67
40	40	100	100	100	100	100	16.67	100	100	100	100	91.67
	Avg.	100	100	100	100	91.67	40.48	100	80.9	83.3	100	<b>89.52</b>

**Table 3.** Confusion matrix for feature misclassification (data of 2 % additive external noise) at 15 Hz

Faults	ND	BRB	PUF1	PUF2	SWF1	SWF2	BF	UR	BR	MR	Individual accuracy
SWF1	0	0	0	0	<b>66.67</b>	33.33	0	0	0	0	66.67
SWF2	0	0	0	0	83.33	<b>16.67</b>	0	0	0	0	16.67
UR	0	0	0	0	0	0	0	<b>33.33</b>	66.67	0	33.33
BR	0	0	0	0	0	0	0	16.67	<b>83.33</b>	0	83.33

## Conclusions

In this paper, the effect of noise on SVM based fault diagnosis of IM is checked for a number of mechanical and electrical faults. The vibration as well as current signals are used for the effective IM fault diagnosis. In order to check the fault diagnosis performances over different operating speeds of IM, the diagnosis is presented for various speeds when there is a high load on IM. To get the noisy signal, the white Gaussian noise is added to the raw signals

which is acquired in the controlled condition in the laboratory. Three effective features i.e.,  $\sigma$ ,  $\gamma$  and  $\kappa$  are extracted from the original signal and noisy signal. The fault diagnosis is then performed for various operating speeds and a higher load of the IM for both the cases. The results show that when the original raw signal is considered for the diagnosis, perfect prediction is found at all the speeds. However, when the diagnosis is performed with the noisy signals the overall prediction performances reduces by 10 %. Still for noisy data case the diagnosis is better at higher speed i.e., above 25 Hz. It is noted that the 2 % external noise is considered for the present study, however it would be interesting to check the performance of the classifier for higher percentage of noise. The present study has been done at higher load condition, the diagnosis can be checked for light and no load on the IM.

## References

- [1] Y. Zhongming, and W. Bin, Power Electronics and Motion Control Conference, 2000. *Proceedings. IPEMC 2000*. The Third International, vol. **3**, pp. 1353-1358. IEEE (2000).
- [2] S. Nandi, H. A. Toliyat, and X. Li, *IEEE transactions on energy conversion*, **20**(4), pp. 719-729 (2005).
- [3] D. Basak, A. Tiwari, and S. P. Das, *In Industrial Technology, ICIT 2006*. IEEE International Conference on, pp. 3061-3066. IEEE (2006).
- [4] M. E. H. Benbouzid, *IEEE transactions on industrial electronics*, **47**(5), pp. 984-993 (2000).
- [5] S. H. Kia, H. Humberto, and G.A. Capolino, *IEEE Transactions on Industry Applications*, **45**(4), pp. 1395-1404 (2009).
- [6] H. Nejjari, and M. E. H. Benbouzid, *IEEE Transactions on Industry Applications* **36**(3), pp. 730-735 (2000).
- [7] J. Antonino-Daviu, P. J. Rodriguez, M. Riera-Guasp, A. Arkkio, J. Roger-Folch, and R. B. Pérez, *Energy Conversion and Management*, **50**(7), pp. 1810-1820, (2009).
- [8] V. T. Tran, B. S. Yang, M. S. Oh, and A. C. C. Tan, *Expert Systems with Applications*, **36**(2), pp. 1840-1849 (2009).
- [9] L. Li, and C. K. Mechefske, *International Journal of COMADEM*, **9**(3), p. 15, (2006).
- [10] V. A. D. Silva, and R. Pederiva, *Surveillance 7*, International Conference - October 29-30, 2013, Institute of Technology of Chartres, France (2013).
- [11] J. D. Martínez-Morales, E. Palacios, and D. U. Campos-Delgado, *In Electrical Engineering Computing Science and Automatic Control (CCE)*, 2010 7th International Conference on, pp. 176-181. IEEE (2010).
- [12] A. Widodo, B. S. Yang, and T. Han, *Expert Systems with Applications*, **32**(2), pp. 299-312 (2007).
- [13] P. Gangsar, & R. Tiwari, *Mechanical Systems and Signal Processing*, **94**, pp. 464-481 (2018).
- [14] C. J. Burges, *Data mining and knowledge discovery*, **2**(2), pp. 121-167 (1998).
- [15] V.N. Vapnik, *Neural Networks, IEEE Transactions on*, **10**(5), pp. 988-999 (1999).
- [16] C.W. Hsu, and C.J. Lin, *Neural Networks, IEEE Transactions on*, **13**(2), pp. 415-425 (2002).
- [17] C. C. Chang and C. J. Lin, *ACM Transactions on Intelligent Systems and Technology (TIST)*, **2**(3), pp. 27:1-27:27 (2011).
- [18] P. Gangsar, & R. Tiwari, *Journal of Dynamic Systems, Measurement, and Control*, **140**(8), 081014 (2018).