Perspectives in Science (2016) 8, 334—337

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/pisc

Feature selection and classification of mechanical fault of an induction motor using random forest classifier☆

Raj Kumar Patel*, **V.K. Giri**

EED, MMMUT, Gorakhpur 273010, U.P., India

Received 4 February 2016; accepted 10 April 2016
Available online 27 April 2016

KEYWORDS
CART;
Neural Network;
Random Forest;
Sensitivity;
Specificity

Summary Fault detection and diagnosis is the most important technology in condition-based maintenance (CBM) system for rotating machinery. This paper experimentally explores the development of a random forest (RF) classifier, a recently emerged machine learning technique, for multi-class mechanical fault diagnosis in bearing of an induction motor. Firstly, the vibration signals are collected from the bearing using accelerometer sensor. Parameters from the vibration signal are extracted in the form of statistical features and used as input feature for the classification problem. These features are classified through RF classifiers for four class problems. The prime objective of this paper is to evaluate effectiveness of random forest classifier on bearing fault diagnosis. The obtained results compared with the existing artificial intelligence techniques, neural network. The analysis of results shows the better performance and higher accuracy than the well existing techniques.

Introduction

Now a day, industries are more dependent on the rotating machinery. The most popular machine used in the industries is induction motor staring from fractional horse power to large rating. In order to make continuous production, the monitoring of health of these induction motor are most challenging task for the health monitoring engineers (Mortazavizadeh and Mousavi, 2014). Generally, the parameter which is associated with the induction motors are recorded and used for analysis to know the internal condition of the motor. The vibration produced by the motor is one of the vital parameter which is widely used to detect the fault in induction motor. Mostly, the first action on recorded signal is to calculate parameters in time domain but this analysis do not claimed as effective method. Hence, the success of diagnosis depends on the other signal processing

☆ This article belongs to the special issue on Engineering and Material Sciences.

* Corresponding author.
E-mail addresses: rajkp007@gmail.com (R.K. Patel), girivkmmm@gmail.com (V.K. Giri).

http://dx.doi.org/10.1016/j.pisc.2016.04.068

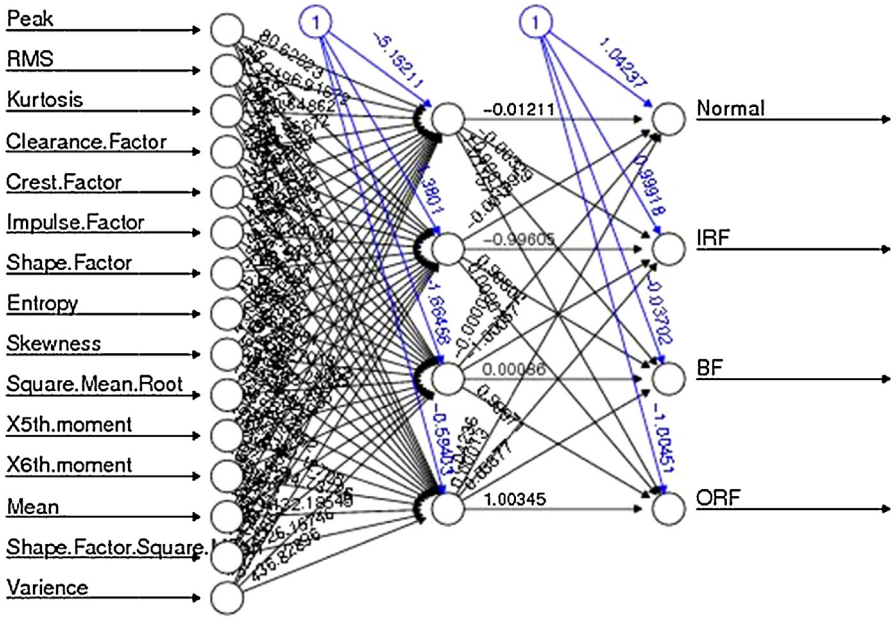

Figure 1 Trained neural network with input, output and training process.

techniques. During last decade, frequency spectrum, envelope detection using Hilbert transform, Empirical Mode Decomposition (EMD), Short Time Fourier Transform (STFT), Wavelet Transform (WT) methods have been used for extracting the features and to identify the fault frequencies (Aiyu et al., 2013; Li et al., 2012; Wang et al., 2013; Xu et al., 2013). Principal Component Analysis (PCA), Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) are very popular for selecting the significant feature for the classification of fault type and its severity (Hu et al., 2014; Unal et al., 2014). Automatic fault detection and online condition monitoring Artificial Neural Network (ANN) and Support Vector Machine (SVM) are enormously used in the literature. The existing classifier used for fault detection is affected when the number of class is more. It has also been reported that the response become sluggish. In order to sort out the mentioned problem in the present paper Random Forest (RF) classifier method has been used. Its performance has been compared with the one of the conventional classifier, i.e. ANN.

Data description

In the present study, bearing vibration records have been selected from the Case Western Reserve Lab data centre for analysis and classification purpose. The data have been recorded from 2 HP three phase induction motor bearing using accelerometer. The accelerometer was mounted on the motor housing at the drive end of the motor. The samples have been taken at 12,000 per second for each of a 16 channels digital audio tape (DAT) recorder. The speed of the shaft is measured as 1797—1752 rpm from no load to full load, respectively. The fault seeded into the races and rolling ball. The size of fault is 0.007 inch, 0.014 inch and 0.021 inch for each cases of fault. The recorded data have been prepared for the classification pupose. The four cases have been taken as Normal, Inner Raceway Fault (IRF),

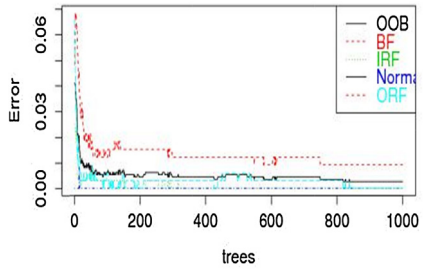

Figure 2 OOB-error of random forest classifier.

Bearing Ball Fault (BF) and Outer Raceway Fault (ORF). The data for classification of Normal case is 160 × 3000 samples and for IRF, BF, ORF the size is chosen of 480 × 3000 for each case. From the prepared data, the 15 statistical feaures have been calculated, and feeded to the neural network and random forest classifier, these features are shown in Fig. 1. The output of the classifier chosen as four related to the condition of the bearing of an induction motor.

Artificial neural network model

Artificial neural networks model is inspired from biological learning process of the human brain and have been developed in form of parallel distributed network. During the training process of the neural networks are fitted to the data by learning algorithms. In this article, supervised learning has been used with input of 15 feature and four output. In the present work, the traditional backpropagation algorithm has been used which is regarded by the usage of a given output that compared to the predicted output and by adjusted of all parameters according to comparision. The parameter of ANNs such as weigts are usully initialized with random values drawn from a standard normal distribution. The first step of the ANN training process is that to compute an output for given inputs and its current weights and output is compared