IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-5, May 2018 E-ISSN: 2347-2693

A Two-Stage Learning Method For Fault Detection of Machines Using Mechanical Big Data

M.S. Jamadagni[1*], K.A. Nagarkar[2], V.S. Badame[3] , C.R. Chindwin[4] , V.M. Lomte[5]

[1,2,3,4,5] Computer Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune University, Pune, India

**Corresponding Author: mugdha1.jamadagni@gmail.com, Tel.: +91-8087999768*

Available online at: www.ijcseonline.org

Accepted: 16/May/2018, Published: 31/May/2018

Abstract— Intelligent fault diagnosis is a promising instrument to deal with mechanical big data because of its capacity in quickly and proficiently handling gathered signals and giving exact diagnosis outcomes. Feature extraction is done manually in most of the traditional techniques which required previous knowledge along with diagnostic expertise. Such procedures take favourable position of human inventiveness is tedious and work escalated. The main possibility of unsupervised component discovering that utilization of intelligence systems to learn raw data, a two-stage learning technique is proposed for intelligent analysis of machines. In the first stage vibration signal is utilized to get a grasp on features from mechanical vibration signals. In the next stage, softmax regression is used to classify the health conditions depends on the studied features. The approach is verified by a motor bearing dataset and a locomotive bearing dataset. It can be seen that using this method high diagnosis accuracy can be obtained. Also, the proposed method reduces the need of human labour making it preferable than the existing methods.

Keywords— Mechanical big data, unsupervised feature learning, sparse filtering, softmax regression, intelligent fault diagnosis.

I. INTRODUCTION

In current industries , machines have turned out to be more programmed, exact and effective than ever, which influences their health condition monitoring. As the information for the most part is gathered quicker than it is investigated by diagnosticians, how to successfully extract qualities from mechanical big data and precisely distinguish the comparing health conditions turns into an earnest research subject as of now. Intelligent fault analysis is fit for preparing massive collected signals effectively and giving precise fault diagnosis results. Ordinarily, the procedure incorporates three major steps as signal acquisition, feature extraction and selection, and fault categorization. In the first step the signals can be acquired since they contain important information regarding mechanical vibration signals. This feature extraction aims at extracting the ideal features from the collected signals. Further pre-processing algorithms are applied followed by classification algorithm. The histogram, which is our contribution, facilitates in better understanding of learned features. It also helps to compare the features.

Rest of the paper is organized as follows, Section I contains the introduction, Section II explains a few of the existing techniques in brief , Section III contains the system design of the experimented system, Section IV gives the mathematical model for the system, section V describes the results and discussion, Section VI concludes research work with future directions.

II. RELATED WORK

The paper [1] presents a HACE theorem that differentiates the features, and presents a Big Data processing model. This approach includes on request collection of information resources, mining and examination, user interest modeling, and security and protection contemplations. Weaknesses are: High performance computing stages are required. These platforms requires deliberate outlines to utilize the Big Data productively. In paper [2], the use of big data in industrial applications is to acquire a process which is cheap and error free, while maintaining its performance. The drawback of the system is that the basic research concentrating on big data solution is basic. The paper [3] gives an audit on the condition checking and fault diagnostic advances for wind turbines. The capacities, abilities, and constraints of the signs and flag preparing methods for WT CMFD. Disadvantages are: Requiring additional space to mount sensors, need additional capital and Quality & Maintenance costs and increasing WT system complexity.

It has a high computational complexity. The paper [4] represents, the structures of a planetary gearbox and a fixed-axis one are briefly introduced. Also, the singular behaviours and fault characteristics of planetary gearboxes are identified and analyzed. The paper [5] put forth an efficient way to predict the geometrical parameters and detect welded defects by using optical features. The feature vector affects the estimation and classification accuracy which can be gained by using wavelet packet decomposition principal component analysis. It gives an innovative information driven-based approach for laser welding process monitoring and defects analysis.

Disadvantages of Existing System:

1. Conventional artificial intelligent methodologies fail to extract and arrange the discriminative information from raw data directly.
2. The features are conventionally extracted and chosen by a particular issue and presumably unsatisfactory for different issues.
3. Lack of a comprehensive understanding of mechanical big data, it is often difficult to ensure the extracted features carrying optimal information to classify the mechanical faults.

III. SYSTEM DESIGN

Unsupervised learning holds the ability to overcome the said shortcomings in the conventional intelligent detection techniques. Unsupervised feature learning is an arrangement of algorithms considering how to well train the AI systems with the unlabeled raw information in order to automatically study the discriminative features required for classification. Unsupervised feature learning has been broadly utilized as a part of various fields, for example, picture grouping, question discovery, discourse acknowledgment, and so forth, and best in class performances are acquired in these fields. The fig.1 shows the framework, the features are specifically gained from mechanical raw signs and a classifier is used to arrange the mechanical deficiencies in light of these learned features. Following the new intelligent fault finding system, we propose a two-stage learning technique. In the first learning stage, sparse filtering, which is seen as a two-layer network, is utilized to learn the features from the mechanical vibration signals. At that point in the second learning stage, softmax regression, which is additionally a two-layer network, is prepared to consequently classify the mechanical health conditions.

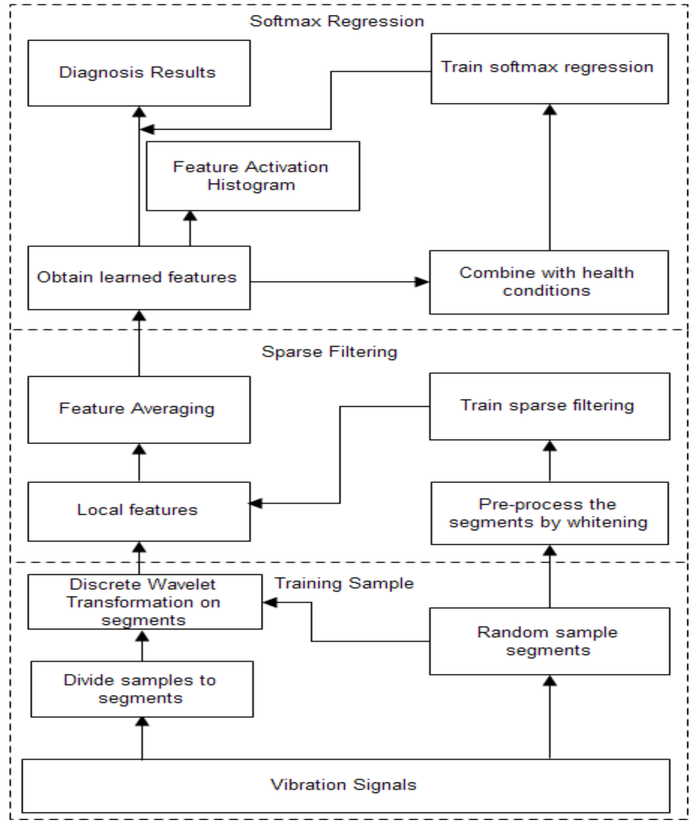

Figure 1. Block Diagram of Proposed System

The learning stage has the following three steps. We first train sparse filtering and get its weight matrix. At that point, the prepared sparse filtering is used to catch the local features from each example. At last, these features are averaged in order to acquire the features of each example. In this paper, we use an average path rather than the aggregate way. Averaging is important since it enhances discriminative features that the segments share with each other and suppresses random features caused by noise. Finally, the health conditions of the test samples are chosen by the softmax regression with the help of learned features.

Advantages of Proposed System:

1. Learns features that catch discriminative data from vibrations in an unsupervised way.
2. It obtains high diagnosis accuracy.
3. It reduces the need of human labor.
4. It makes intelligent fault diagnosis manage big data more easily.

IV. MATHEMATICAL MODEL

1. Sparse Filtering

Given a training set $\{x^i\}_{i-1}^{M}$, where $x^i \in R^{N\times 1}$ is a sample and M is the sample number, sparse filtering maps the samples onto their features $f^i \in R^{L\times 1}$ using a weight matrix $W \in R^{N\times L}$.

$$f_l^i = W_l^T x^i \qquad (1)$$

Where f_l^i compares to the lth feature of the ith test. sparse filtering advances a function by using l2– normalized features[10]

2. Softmax Regression

As in [10], we have a training set $\{x^i\}_{i-1}^M$, with its label set $\{y^i\}_{i-1}^M$, where $x^i \in R^{N\times 1}$ and $y^i \in \{1,2,\dots,K\}$. For each input sample x^i, the model attempts to estimate the probability $P(y^i = k|X^i)$ for each label of k = 1,2,…,K. Thus, the hypothesis of softmax regression will output a vector that gives K estimated probabilities of the input sample x^i belonging to each label. Concretely, the hypothesis $h_\theta(x^i)$takes the form:

$$\boldsymbol{h}_\theta(\boldsymbol{x}^i) = \begin{bmatrix} p(y^i = 1|\boldsymbol{x}^i;\boldsymbol{\theta}) \\ p(y^i = 2|\boldsymbol{x}^i;\boldsymbol{\theta}) \\ \vdots \\ p(y^i = K|\boldsymbol{x}^i;\boldsymbol{\theta}) \end{bmatrix} = \frac{1}{\sum_{k=1}^{K} e^{\boldsymbol{\theta}_k^T x^i}} \begin{bmatrix} e^{\boldsymbol{\theta}_1^T x^i} \\ e^{\boldsymbol{\theta}_2^T x^i} \\ \vdots \\ e^{\boldsymbol{\theta}_K^T x^i} \end{bmatrix}$$

Algorithms:

1) Sparse Filtering

It does not directly model the data (i.e. it has no reconstruction error term in the cost function). The aim of the algorithm is to form a dictionary D that provides a sparse representation by reducing the normalized entries in a feature value matrix. For each iteration of the algorithm:

1. ℓ2 normalization across rows
2. ℓ2 normalization across columns
3. Objective function = ℓ1 norm of normalized entries

2) PCA Algorithm:

Step 1: Normalize the data
Step 2: Calculate the covariance matrix
Step 3: Calculate the eigen values and eigen vectors
Step 4: Choosing components and forming a feature vector
Step 5: Forming Principal Components

3) Softmax Regression

Step 1: Initialize constants and parameters
Step 2: Load data
Step 3: Implement softmaxCost
Step 4: Gradient checking
Step 5: Learning parameters
Step 6: Testing

V. EXPERIMENTAL RESULTS

The training motor bearing signals collected by Case Western Reserve University website. And testing dataset is thecollection of vibration signals using Intelligent Maintenance System (IMS) bearing dataset. The experimental evaluation of the vibration signals were collected from the drive end of a motor in the test rig under four different conditions: normal condition, outer race fault (OF), inner race fault (IF), and roller fault (RF). For OF, IF, and RF cases, vibration signals for three different severity levels (0.18, 0.36, and 0.53 mm) were collected separately. The signals were all collected under four load conditions (0, 1, 2 and 3 hp) and the sampling frequency was 12 kHz.The testing vibration signals samples are shown below in fig. 2

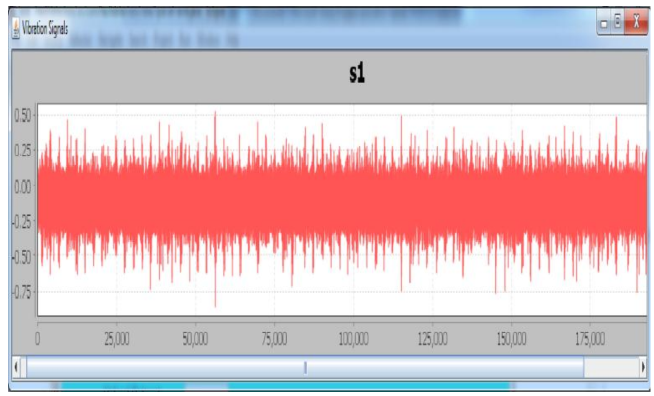

Figure 2. Vibration Signals in DWT format

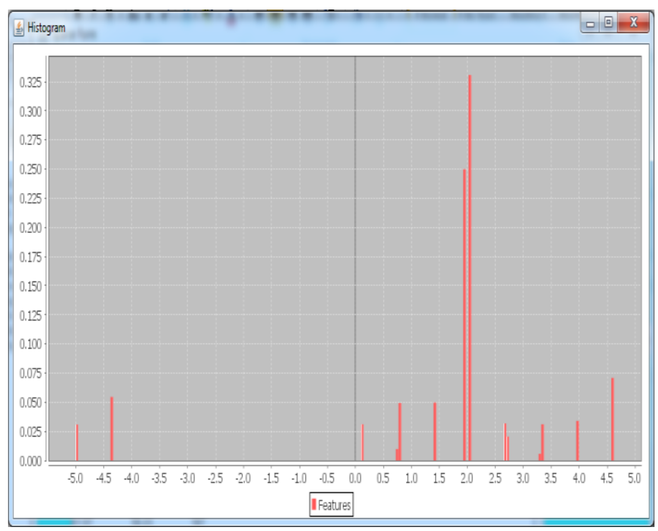

Figure 3. Histogram of Sparse Filtering features

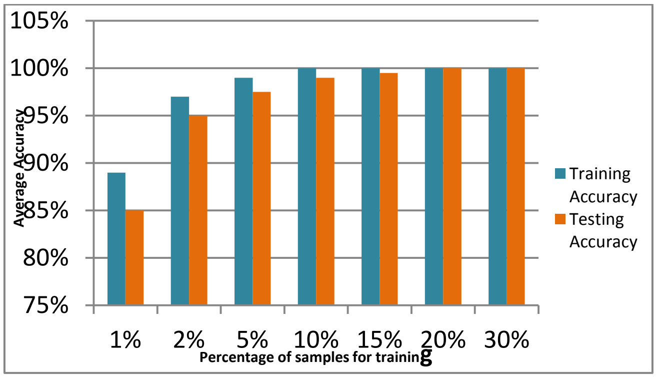

Figure 4. Accuracy analysis between training and testing on different samples

Table 1. The used CWRU bearing data set number (SN) of every classification group

Group Number	Fault Diameter (mm)	NC (SN)	IF (SN)	RF (SN)	OF (SN)
1	0.18	100	108	121	133 (6 o'clock)
2	0.18	100	108	121	147 (3 o'clock)
3	0.18	100	108	121	160 (12 o'clock)
4	0.36	100	172	188	200 (6 o'clock)
5	0.53	100	212	225	237 (6 o'clock)
6	0.53	100	212	225	249 (3 o'clock)
7	0.53	100	212	225	261 (12 o'clock)

Table I gives the data set number in seven analyzed data groups with a 12 kHz sampling frequency. The fault depth is 0.11 inches. The fault diameters are 0.18 mm, 0.36 mm and 0.53 mm, respectively. Meanwhile, the OF position is located at the lowest point (6:00 o'clock), the highest point (12:00 o'clock) and horizontal point (3:00 o'clock), respectively.

The fig. 4 demonstrates the execution investigation chart amongst preparing and testing tests. It is sure that the testing precision increments and its standard deviation decreases with the increase in percentage of sample. It can be observed from the figure that the method diagnoses the ten health conditions of the motor bearing dataset with 95.2% accuracy using only 2% of samples for training. At the point when the rate increments to 10%, the testing exactness is 99.66% with a little standard deviation of 0.19%. They indicate that the presented method could perform well even in the situation of lacking the labeled data.

VI. CONCLUSION & FUTURE SCOPE

In this paper, a two-stage learning method is based on unsupervised feature learning. In the method, sparse filtering flexibly learns features. Then the features are fed to softmax regression to classify health conditions in a supervised manner. Through the case studies of the two bearing datasets, it is shown that the proposed method flexibly learns features from raw signals for various diagnosis issues. The proposed method is able to take advantage of unsupervised learning and improve its diagnosis accuracy along with the increase of the number of the unlabeled data. The histogram, which is our contribution, facilitates in better understanding of learned features .It is more effective to use the strategy of local feature averaging in the diagnosis cases of lacking the labeled data for the final classification. The future work is, apply the DWT technique on segments as well as plot histogram results using learned features. So, the system will show the results to users.

VII. REFERENCES

[1] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng*., vol. 26, no. 1, pp. 97-107, Jan. 2014.

[2] S. Yin, and O. Kaynak, "Big data for modern industry: challenges and trends," *Proc. IEEE*, vol. 103, no. 2, pp. 143-146, Feb. 2015.

[3] W. Qiao, and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis, " *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6536-6545, Oct. 2015.

[4] Y. Lei, J. Lin, M. J. Zuo, and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review," *Measurement*, vol. 48, pp. 292-305, Feb. 2014.

[5] D. You, X. Gao, and S. Katayama, "WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 628-638, Jan. 2015.

[6] Y. Lei, Z. He, Y. Zi, and X. Chen, "New clustering algorithm-based fault diagnosis using compensation distance evaluation technique," *Mech. Syst. Signal Process.*, vol. 22, no. 2, pp. 419-435, Feb. 2008.

[7] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418-6428, Nov. 2014.

[8] Y. Lei, Z. He, Y. Zi, and Q. Hu, "Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs," *Mech. Syst. Signal Process.*, vol. 21, no. 5, pp. 2280-2294, Jul. 2007.

[9] M. Kang, J. Kim, J.-M. Kim, A. C. Tan, E. Y. Kim, and B.-K. Choi, "Reliable Fault Diagnosis for Low-Speed Bearings Using Individually Trained Support Vector Machines With Kernel Discriminative Feature Analysis," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2786-2797, May 2015.

[10] Yaguo Lei, *Member, IEEE*, Feng Jia, Jing Lin, Saibo Xing, and Steven X. Ding," An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning Towards Mechanical Big Data.", *IEEE Transctions on Industrial Electronics*, 2016.

Authors Profile

Mugdha Jamadagni is currently pursuing her B.E degree in computer engineering from RMD Sinhgad School of Engineering under the University of Pune. Her areas of interest include Data Mining, Machine Learning and Smart Systems.

Kalyani Nagarkar is currently pursuing her B.E degree in computer engineering from RMD Sinhgad School of Engineering under the University of Pune. Her areas of interest include Data Mining, Artificial Intelligence and Networking.

Vaishnavi Badame is currently pursuing her B.E degree in computer engineering from RMD Sinhgad School of Engineering under the University of Pune. Her areas of interest include Data Mining, Machine Learning and Artificial Intelligence.

Crystal Chnidwin is currently pursuing her B.E degree in computer engineering from RMD Sinhgad School of Engineering under the University of Pune. Her areas of interest include Data Mining, Cloud Computing and Machine Learning.

