

# Air Compressor Fault Diagnosis Through Vibration Signals using Statistical Features and J48 Algorithms

S. Aravinth<sup>1\*</sup>, K. Rajesh Kanna<sup>2</sup> and V. Sugumar<sup>2</sup>

<sup>1</sup>Bannari Amman Institute of Technology, Sathyamangalam - 638401, Tamil Nadu, India; s.aravinth2015@vit.ac.in

<sup>2</sup>School of Mechanical and Building Sciences (SMBS), VIT University, Chennai Campus, Chennai - 600127, Tamil Nadu, India; rajesh.kannak2014@vit.ac.in, v\_sugu@yahoo.com

## Abstract

**Objectives:** The fault diagnosis in reciprocating air compressor system was done through this article using vibration signals from accelerometer for both healthy and faulty conditions. **Methods/Analysis:** This article presents a condition monitoring strategy for compressor through vibration signals using accelerometer data in identifying five common faults of air compressor these were simulated manually. These vibration signals were processed through machine learning technique, where statistical features were extracted and the features contributing to the maximum classification accuracy were selected. The J48 decision tree algorithm is used in predicting the compressor faults in early stages. **Findings:** High classification accuracy of 98.33% was obtained for fault detection in compressor system. **Application/Improvements:** The proposed model can be used for regular monitoring of air compressor.

**Keywords:** Fault Diagnosis, J48 Algorithms, Reciprocating Air Compressor, Vibration Signals

## 1. Introduction

Air compressors are machineries used widely to support industries in pressurized air requirement, handle chemicals in the required pressure, also found in various air conditioning systems and used to power pneumatic tools and systems; the machinery is often subjected to mechanical wear, the prime need for an intelligent Fault Detection and Diagnosis (FDD) system is that an air compressor is not located in places where frequently accessible from the work place. The recent practices are that is that both primitive and advanced algorithms are used in building a classification model for diagnosis. Hence, a continuous

monitoring of compressors is one of the important phenomena to achieve good productivity.

Cylinder pressure using pressure transducer and Instantaneous angular speed using optical encoder system are used together in combination to achieve fault diagnosis in earlier stages of reciprocating air compressor<sup>1</sup>. It is difficult to diagnose faults if numbers of fault types are more. Hence, RBF Neural Network was used to train sample data based on which overall sound signals that were achieved through condition monitoring was processed in order to detect faults of compressor<sup>2</sup>. Bearing fault diagnosis<sup>3</sup> is done through sound signals extraction of condition monitoring through machine learning approach and

\*Author for correspondence

features were further selected through decision tree based on important instances from top priority basis and further signals were classified in C4.5 decision tree algorithms at an accuracy level 95% in fault detection<sup>4</sup>. In order to avoid irritation during compressor running in near field as well as far field, sound radiation measurement is done to find major sound source of compressor unit with low spatial resolution. This data is considered as one of major strategy in this study to incorporate changes in compressor design to achieve reduction noise level of compressor<sup>5</sup>. Performance monitoring and fault detection of compressor in petro chemical industries is done through signal processing techniques such as time domain analysis and frequency domain analysis in order to diagnose faults of compressor<sup>6</sup>. An article was presented which utilizes pressure sensor and condenser temperature sensor in air compressor motor diagnostic for the application of air conditioning and refrigeration system which possess motor protector that stops while it undergoes an out specification condition<sup>7</sup>. One of the most common failures in compressor relates to valve mounting or valve fluttering and piston ring damage. A paper proposes a simulated model of common valve failure and piston ring failure using pressure sensor for monitoring of pressure pulsation through the compressor in both healthy as well as failure state to achieve the difference in compressor operation<sup>8</sup>. Overall monitoring of reciprocating air compressor was done to achieve optimization in fault detection using 4 different methods of SVM such as one against one, one against all, fuzzy decision function, DDAG and which gives the diagnostic accuracy of 98.03% in detecting faults of air compressor<sup>9</sup>. Valve fault diagnosis in compressor is done through Information entropy SVM method to analyse the characterization of vibration signal to achieve low nonlinear patten in compressor operation and further the analysed signal were classified through SVM<sup>10</sup>. Condition monitoring for the diagnosis of air compressor fault detection in valve leakage and valve spring deterioration is done through numerical simulation and experimental study in five different physical processes such as speed-torque characteristic of induction motor, cylinder pressure variation, crankshaft rotational motion, flow characteristic through valves and vibration of valve plates<sup>11</sup>. Air compressor faults such as valve leakage, inter-cooler leakage and loose belt drive are commonly occurred which were monitored through vibration signal extraction. These extracted features were selected based on important instances and processed through artificial neural

network and Genetic Algorithms. 93.05% and 95.50% are the acquired accuracy level for time-domain analysis and frequency domain analysis<sup>12</sup>. Vibration characteristics were easily achieved in air compressor where its performance were inspected using virtual instrumentation by analysing the interrelationship between temperature, pressure, flow, speed etc...and which are processed using Lab View for overall monitoring<sup>12</sup>. Fault diagnosis system can be carried out for gear box, brake fault, wind turbine<sup>13</sup>, pump, bearings and TPMS.

**Table 1.** Common compressor faults

| S NO. | FAULT                                    | COMMON CAUSE   | EFFECT  |
|-------|--|--|---|
| 1     | Inlet valve fault (IVF)                  | Valve fluttering associated with inlet reed damage         | Loss in discharge pressure. Build up crank case temperature.                  |
| 2     | Outlet valve fault (OVF)                 | Valve fluttering associated with outlet reed damage        | Loss in discharge pressure. Build up crank case temperature.                  |
| 3     | Valve plate leakage (VPL)                | Damage in valve seating                                    | Loss in discharge pressure. Damage crank case with micro fracture.            |
| 4     | Check valve fault (NVHL)                 | Leakage through the valve housing and the connecting valve | Loss in outlet pressure leading to pressure not built up in storage cylinder. |
| 5     | Inlet and outlet valve fluttering (IOVF) | Both inlet and outlet valve reed damage                    | Catastrophic damage in pressure developed within the whole system.            |

## 2. Experimental Setup

The experimental setup shown in Figure 1 consists of a single stage reciprocating air compressor, in which the common faults were simulated. The data acquisition process were carried out using a piezoelectric accelerometer connected to NI Data acquisition system, where analog data was acquired by fast Fourier transform method as digital data and the processed data were used to classify the different faults acquiring the statistical features from the sensor data<sup>13</sup>. From experimental setup, procedural

development of experimental methods or process in step by step arrangement is projected in methodology shown in Figure 2. The fault detection and diagnosis method was incorporated, where training data were classified using J48 decision tree algorithm with the statistical features as the contributing attributes.



Figure 1. Experimental setup.

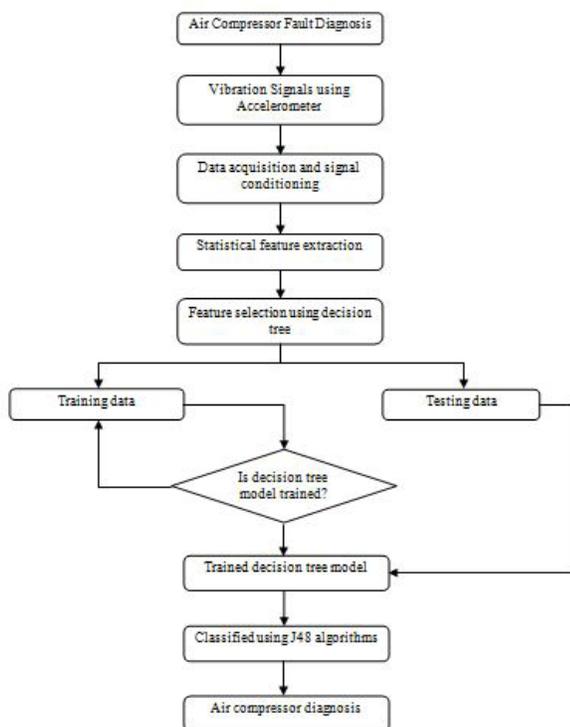


Figure 2. Flow chart of Air Compressor fault diagnosis.

## 2.1 Sampling Rate Calculation

The sample rate determination and the sensor mounting position were the most important parameters to be taken into consideration in the pre-data acquisition stage. Here, the sensor position to acquire the data for the fault diagnosis was positioned as per the API standards and from the study it was clear that the faults occurs normally at 2 or 3 times the operating speed of the compressor. Hence the sample rate was calculated as below:

- Rated speed (experimental setup - compressor): 550 rpm
- Fault frequency:  $(550 \times 3) / 60 = 27.5$  Hz
- By rule minimum sample rate:  $2 \times 27.5$  Hz = 55 Hz

For better data acquisition the sample rate was fixed to be 12 kHz. The common faults occurring with the single stage reciprocating air compressor are tabulated in Table 1. From the acquired data the statistical features were extracted to obtain the hidden data defining the signature signal of each fault.

## 3. Machine Learning

Machine learning consist of three phases; Feature extraction, feature selection, feature classification. The vibration signal will be used to extract the statistical features to yield the required parameters.

### 3.1 Feature Extraction

In Feature extraction the statistical parameters such as 'mean', 'mode' etc. are extracted from the vibration signal. Spread sheet software was used to complete the task. Using accelerometer vibration signals are extracted from air compressor and these signals were extracted using statistical analysis by considering several parameters. 11 parameters of statistical analysis are taken into consideration for the vibration signals to be extracted from compressor fault detection such as mean, standard deviation, standard error, range, median, minimum, maximum, sum, kurtosis, skewness and sample variance. From these parameters, important parameter or instances pertaining to accuracy level alone selected under each and every condition of compressor<sup>14</sup>.

### 3.2 Feature Selection

This work utilizes C4.5 based binary tree creation of J48 algorithms, where it uses a divide and conquers approach in the growth of tree with nodes and instances. Basic steps in construction of tree are<sup>15</sup>:

- To check whether all cases belong to same class.
- For each and every attribute, calculate all the information and information gain.
- Best set of the attribute is selected based on correctly classified instances.

The decision tree consists of the nodes and leaves in which the end node represents the class of which is classified based on the features. The top most node of the decision tree represents the most significant feature classifying the classes' i.e. the compressor faults. The features in the hierarchy of the decision tree represent the features in the order of their significance contributing to the classification accuracy. These extracted features were arranged in each and every node of a tree based on prioritization.

Instances such as standard deviation, standard error, range, sum, minimum, maximum, skewness, kurtosis, medium, mean, sample variance are categorized into nodes before selection based on its importance .In which, standard deviation, standard error, kurtosis, sum, mean and range were the instances placed in the nodes of the decision tree show in Figure 3.

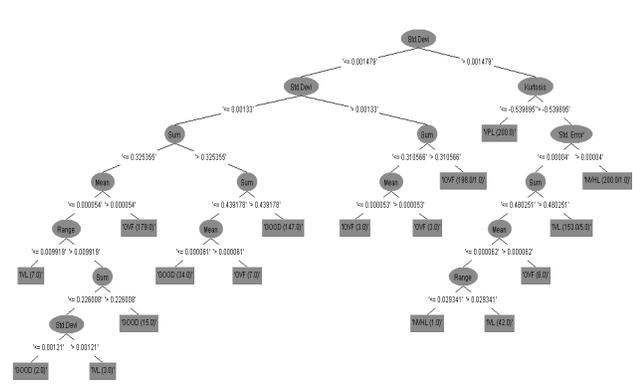


Figure 3. Decision tree.

### 3.3 Feature Classification

J48 is a tree based knowledge representation methodology which consists of branches, root, nodes and leaves to define classification rules<sup>14</sup>. J48 decision tree algorithm (C4.5) has two phases, the building phase and the

pruning phase. In the building phase, J48 builds decision tree by using the concept of information theory. The tree has a single root node for the entire training set. For every partition, a new node is added to the decision tree. For a set of samples in a partition S, a test attribute X is selected for further partitioning the set into  $S_1, S_2, \dots, S_L$ . New nodes for S are created which are added to the decision tree as children. The construction of decision tree depends on the test attribute X.

J48 uses entropy based information gain as the selection criteria. As per information theory, entropy is a measure of the uncertainty in a random variable. The expected reduction in entropy due to the partitioning of the examples according to the given feature gives the information gain. It is a measure of the capability of a given attribute to separate its training examples according to the target function. When the variation in data becomes large, the decision tree becomes large leading to more inaccuracy due to under fitting or overtraining. Thus, for better classification accuracy, the trees must be pruned to remove less reliable branches.

## 4. Results and Discussion

Decision tree algorithm provides advantage to one with easy understanding and classification of an observed data-set. The J48 or C4.5 algorithm works on a simple 'if-else' based classification, depending upon the given attributes or features of the dataset. The outcome of a decision tree algorithm consists of an inverted tree representation starting with the root or the attribute that contributes the most to classification of the given classes<sup>15</sup>. The root is followed by branches providing further detail about the classification and leaves are the end nodes that give the conclusion to the classification. In a decision tree algorithm suitable pruning factors and confidence factor are governed to get an optimal outcome of the classification show in Table 2.

Table 2. Effect of confidence factor and minimum number of objects

| S.No | Confidence factor | Minimum number of object | Accuracy |
|------|-------------------|--------------------------|----------|
| 1    | 0.25              | 2                        | 97.75%   |
| 2    | 0.25              | 1                        | 98%      |
| 3    | 1                 | 2                        | 97.42%   |
| 4    | 3                 | 2                        | 97.42%   |

## 4.1 Effect of Confidence Factor (C)

The value of confidence factor lies between 0 and 1. For better classification accuracy from the selected attributes the value of 'c' was varied from 0 to 1. From the outcome it is clear that the classification accuracy reduces gradually<sup>16</sup>, and constant at c values after 0.25. It is found that for the 'c' value of '0.25' yields the maximum classification accuracy of 98%.

## 4.2 Effect of Minimum Number of Objects (M)

The most significant minimum number of objects (m) is chosen by which maximum classification accuracy is achieved. However, minimum number of objects being the most important pruning factor the value of m was varied from '1' to '100'. From Table 2, it is clear that the classification accuracy is constant for lower values of 'm' and varies gradually with increase in the value of 'm'. It was found that for the m value of 1 gives maximum classification accuracy of 98%. Hence, the optimum value of 'm' was found to be 1 with a decision tree of nominal size shown in Figure 4.

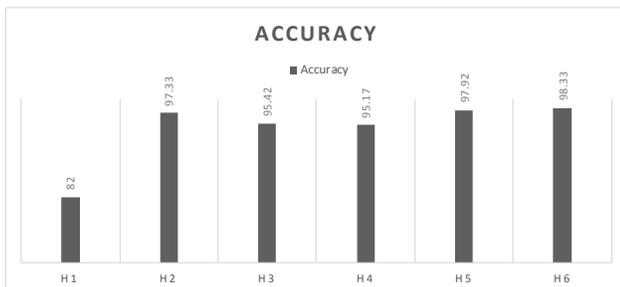


Figure 4. Feature selection accuracy.

Where,

- H1: Standard deviation and condition
- H2: Standard deviation, kurtosis and condition
- H3: Standard deviation, kurtosis, sum and condition
- H4: Standard deviation, kurtosis, sum, standard error and condition
- H5: Standard deviation, kurtosis, sum, standard error, mean and condition
- H6: Standard deviation, kurtosis, sum, standard error, mean, range and condition

From the decision tree the following inferences were obtained:

- The algorithm required only 5 significant attributes i.e. the extracted statistical features namely standard deviation, standard error, mean, range, sum and kurtosis to form an optimum decision tree with 98.33% of classification accuracy<sup>17</sup>.
- The statistical feature standard deviation forms the root of the decision tree, defining it to be the most significant in the fault diagnosis of the reciprocating air compressor.

The confusion matrix was constructed using 10 fold cross validation method, the diagonal elements gives the correctly classified instances out of 200 signals that were taken and out of which diagonal element shown in Table 3 gives the correctly classified instances and apart from diagonal elements gives the misclassified instances during signal classification<sup>18</sup>.

Table 3. Confusion matrix

|      | GOOD | IOVF | IVL | NVHL | OVF | VPL |
|------|------|------|-----|------|-----|-----|
| GOOD | 195  | 2    | 2   | 1    | 0   | 0   |
| IOVF | 0    | 199  | 0   | 0    | 0   | 1   |
| IVL  | 1    | 0    | 199 | 0    | 0   | 0   |
| NVHL | 0    | 0    | 1   | 199  | 0   | 0   |
| OVF  | 1    | 1    | 8   | 0    | 190 | 0   |
| VPL  | 0    | 0    | 2   | 0    | 0   | 198 |

## 5. Conclusion

The fault diagnosis in reciprocating air compressor system was done through this article using vibration signals from accelerometer for both healthy and faulty conditions. The variation in vibration levels of each condition were analyzed using machine learning approach of extraction, selection and classification. Statistical analysis is used for extraction of vibratory signals and 97.75% is the accuracy level obtained. C and M parameters were changed and rise in accuracy level is obtained by keeping C as 0.25 and M as 1. Applied changes were given as an input to decision tree for selection of best instances. These selected instances were processed through J48 and 98.33% is the highest accuracy level in prediction of air compressor fault detection.

## 6. Reference

1. Elhaj E, Almrabet M. A combined practical approach to condition monitoring of reciprocating compressor using IAS and dynamic pressure. *World Academy of Science, Engineering and Technology*. 2010; 4(3):287–93.
2. Jun XM, yong XY. Fault diagnosis of air compressor based on RBF neural network. *International Conference on Mechatronic Science, Electric Engineering and Computer*; 2011 .
3. Joshua A, Sugumaran V, Amarnath M. Selecting Kernel function of support vector machine for fault diagnosis of roller bearings using sound signals through histogram features. *International Journal of Applied Engineering Research*. 2015; 10(68).
4. Amarnath M, Sugumaran V. Exploiting sound signals for fault diagnosis of bearing using decision tree. *Measurement*. 2013; 46(3):1250–6.
5. Cho YT. Noise source identification of reciprocating air compressor in a normal operating condition using spatially low-resolution sound measurement. *Indian Journal of Science and Technology*. 2015; 8(23):1–9.
6. Zhang Y, Jiang J. Condition monitoring and Fault detection of a compressor using signal processing techniques. *Proceedings of the American Control Conference*; Arlington, Virginia. 2001. p. 4460–5.
7. Jayanth N, Pham H. Compressor diagnostic system. *United States Patent*. 2001 Mar 27.
8. Manepatil SS, Tiwari A. Fault diagnosis of reciprocating compressor using pressure pulsations. *International Compressor Engineering Conference*; Purdue. 2006. p. 1–10.
9. Verma NK, Roy A. An optimized method for reciprocating air compressors based on SVM. *IEEE International Conference on System Engineering and Technology*; 2011. p. 65–9.
10. Cui H, Zhang L. Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method. *Journal of Loss Prevention in the Process Industries*. 2009; 22(6):864–7.
11. Elhaj M, Gu F. Numerical Simulation and experimental study of a two-stage reciprocating compressor for condition monitoring. *Mechanical System and Signal Processing*. 2008; 22(2):374–89.
12. Ahmed M, Gu F. Feature selection and fault classification of reciprocating compressor using a genetic algorithm and a probabilistic neural network. *Journal of Physics*. 2011; 305(1).
13. A. Joshua, V. Sugumaran. Fault diagnostic methods for wind turbine: A review. *ARPN Journal of Engineering and Applied Sciences*. 2016 Apr; 11(7):4654–68.
14. Yin J, Jiang L. Air compressor performances inspection system based on virtual instrument. *International Conference on Computer, Mechatronics, Control and Electronic Engineering*; 2010.
15. Sugumaran V, Sabareesh GR. Fault diagnosis of roller bearing using Kernel based neighborhood score multi-class support vector machine. *Expert System with Application*. 2008; 34(4):3090–8.
16. Bhargava N, Sharma G. Decision tree analysis on J48 Algorithm for Data mining. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2013 Jun; 3(6):1114–9.
17. Patil TR, Sherekar MSS. Performance analysis of naive bayes and j48 classification algorithm for data classification. *International Journal of Computer Science and Applications*. 2013 Apr; 6(2):256–61.
18. Selvakumar J, Prakash AE. Data acquisition system and signal processing technique for bearing fault analysis. *Indian Journal of Science and Technology*. 2016 Aug; 9(31):1–7.