

## International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

*IJCSMC, Vol. 3, Issue. 5, May 2014, pg.659 – 668*

### **RESEARCH ARTICLE**

# **Classification of Multiple Fault in an Automobile Engine Using Statistical and ANN Based Classifiers**

Shankar N. Dandare  
Associate Professor in Electronics & Tele. Engg.  
Electronics Department,  
College of Engineering, Pusad, Maharashtra-State, India  
Email: [sndandare@rediffmail.com](mailto:sndandare@rediffmail.com)

S. V. Dudul  
Professor & Head, Dept of Applied Electronics,  
Sant Gadge Baba Amravati University, Amravati,  
Maharashtra-State, India  
Email: [svdudul@gmail.com](mailto:svdudul@gmail.com)

## **Abstract**

The proposed system follows a model-based approach based on Digital Signal Processing and Artificial Neural Network. Fault Detection and Isolation (FDI) of an Automobile Engine have been carried out using acoustic signals which is captured from the engine. This method is based on parameter estimation, where a set of parameters is used to check the status of an engine and a model based approach is employed to generate several symptoms indicating the difference between faulty and non-faulty status.

In this work, experimentation is carried out on Maruti Suzuki Alto Four Stroke (MSAFS) Engine. There are many more types of faults which may be developed because of wear and tear or lack of maintenance but, the database is generated only for three different types of faults and classification of the same is carried out. The signal normalization, signal conditioning, signal decompositions, analog to digital conversion and feature extraction were carried out by using the algorithm written in MATLAB R2010B. The paper describes the performance of statistical and Artificial Neural Network (ANN) based classifiers for individual and multiple faults and finally the optimal classifiers are proposed based on classification accuracy. It is observed from the experimental results that ANN based classifiers are more appropriate than statistical classifiers. It is also observed that the magnitude of MSE is under permissible limits and percentage Average Classification Accuracy (% ACA) is also reasonable.

**Keywords:** Signal Processing, Statistical Classifiers, Four Stroke Engine and ANN Based Classifiers

## **1. Introduction**

Today, transportation technology, especially car, grows fast, but many drivers do not know how to work for their car. Fault Detection is not an easy for inexperienced mechanic or driver because it is needed a lot of knowledge for finding the fault. Therefore, they extremely depend on expert mechanic. Looking into this the FDI system is proposed to detect the fault in an incipient stage to avoid the inconvenience. The work carried out in this area is discussed below.

The classical approaches are limited for checking of some measurable output variables because they do not give a deeper insight and usually do not allow a fault diagnosis. Model-based methods of fault detection were developed by using input and output signals and applying dynamic process models. These methods are based, on parameter estimation, parity equations or state observers, Rolf Isermann, (2005), suggested the model based approach. The goal is to generate several symptoms indicating the difference between nominal and faulty status. Based on different symptoms fault diagnosis procedures follow, determining the fault by applying classification or inference methods [1]. The fault can be isolated if the residual associated with the matched isolation estimator remains below its corresponding adaptive threshold, whereas at least one of the components of the residuals associated with all the other estimators exceeds its threshold at some finite time [2]. R. J. Howlett, (1996 & 1999), a neural network technique was described for determination of air-fuel ratio in the engine. The voltage waveforms across the spark plug were used for monitoring the engine and for fault diagnosis or control [3, 4]. The soft computing (SC) methods were surveyed by R. J. Patton *et al.* (2001), in this study, the use of SC methods was considered an important extension to the quantitative model-based approach for residual generation in FDI [5]. Wang Weijie, *et al.* (2004), proposed the engine vibration signals for fault diagnosis. A model of wavelet neural networks was constructed based on wavelet frame theory and neural networks technology [6]. The FDI system in dynamic data from an automotive engine air path using artificial neural networks was investigated by M. S. Sangha *et al.* (2005). A generic SI mean value engine model was used for experimentation. Several faults were considered, including leakage, EGR valve and sensor faults, with different fault intensities. RBF neural networks were trained to detect and diagnose the faults, and also to indicate fault size [7]. Jian-Hua Zhang, *et al.* (2010), Proposed a fault diagnosis using Adaptive Neuro-Fuzzy inference system (ANFIS). ANFIS was applied to build a fault diagnosis model of automobile engine and induce cloud model of fan-out, outputting results were used to detect the performance parameter failure for the automobile engine [8]. Zhe Wang, *et al.* (2011), proposed a Fault Diagnosis Model for Automobile Engine using gradient descent genetic algorithm and optimization of system parameters have been carried out using neural network learning algorithm [9]. Hamad A., *et al.* (2012), proposed a RBF network to classify the faults. The performance of the developed scheme was assessed using an engine benchmark, the Mean Value Engine Model (MVEM) with Matlab/Simulink. Six faults have been simulated on the MVEM, including four sensor faults, one component fault and one actuator fault [10]. Sensor fault detection, isolation (FDI) and accommodation has been investigated by M. S. Sangha, *et al.* (2012), along with detection of unknown faults for an automotive engine. Radial basis function (RBF) neural networks were used for fault diagnosis [11]. Madain M., *et al.* (2010), proposed the fault diagnosis using the sound samples. The fault under test was compared with the faults in the database according to their correlation, normalized mean square error, and formant frequencies values and the best match was considered fault detection. The developed system can be useful for the inexperienced technicians as a training module for them [12].

Based on the review of related literature the observation are made as under

- The fault diagnosis is not standardized among vehicle manufacturers.
- There is an uncertainty in scheduling vehicle.
- A complete diagnosis may need special equipments and trained technician's help, which could make the diagnosis very expensive.
- Fault diagnosis is difficult because of hybrid control systems.
- Multiple sensors are required to detect the multiple faults and hence, the system may be complex.
- Sensitive Sensors with high accuracy and precision are required.
- Maintenance of sensors is difficult.

By considering the different approaches for FDI system, the MSFAS Engine model is selected for experimentation. The specification of engine model is given in the Table 1. The proposed system follows a model-based approach based on Digital Signal Processing and Artificial Neural Network with single sensor system. The knowledge database is generated by recording the sound variation signals at different speeds and with each gear position in healthy and faulty conditions. The normalization, signal conditioning and analog to digital conversion were carried out by using the algorithm written in MATLAB R2010B. Statistical and ANN based classifiers are employed to classify the faults correctly.

The experimentation is carried out at “Automobile Engineering Laboratory, Department of Electronic and Mechanical Engineering, Babasaheb Naik College of Engineering, Pusad. District Yavatmal. (M.S.)” and Research laboratory of Department of Applied Electronics, Faculty of Engineering and Technology, Sant Gadge Baba Amravati University, Amravati”. The data acquisition system consists of an automobile engine along with the microphone as a sensor to capture the acoustic signal, signal recording, signal conditioning and signal processing system. There are many more types of faults which may be developed in the automobile engine but, the database is

generated for only three different types of faults that are Air Filter Fault (FF), Insufficient Lubricants Fault (ISL) and Insufficient Fuel Supply (ISF).

It is worthwhile to notice that the proposed system may be designed and attached to every newly produced engine, so that the fault can be detected at an incipient level. It is also suggested that the proposed FDI system can be extended to detect any number of faults. The proposed FDI system can be used as one type of tool to know the status of the engine and will act as a guide for maintaining the vehicle in good condition that will save our time and inconvenience. By considering the necessity of FDI system the broad objectives of proposed FDI system are listed as under.

- It is possible to detect the faults at an incipient stage.
- To improve productivity & reliability of an automobile.
- To facilitate unskilled or less skilled automobile staff to work more efficiently.
- To reduce the maintenance cost and down-time of an automobile.
- To avoid vehicular accidents because of inadequate maintenance.
- To prevent the monetary loss of customer (in the event of a wrong diagnosis).
- As a tool for training inexperienced people.
- To improve knowledge of driver in diagnosing the fault.
- The proposed FDI system is simple, reliable and flexible.
- It is single sensor system based on acoustic signal.

Table1: Specifications of Automobile Engines, Microphone and Sound Recorder

MSAFS Engine	Microphone	Sound Recorder
<i>Engine Displacement (cc) : 796</i>	Unidirectional Cardioid	AD/DA conversion : 24 bits, 44.1 kHz
<i>Maximum Power: 46bhp @6200rpm</i>	<i>Frequency : 50Hz -18KHz</i>	<i>Format : WAV</i>
<i>Engine Type : In-Line Engine, 3-Cylenders</i>	<i>Impedance : 32 Ohm</i>	<i>Bit Rates : 64/96/128/ 160/192 /256/320 kbps</i>
<i>Gear Box : 5 Speed</i>	<i>Sensitivity : 62 dB</i>	<i>Frequency Response : 20 Hz to 20 kHz</i>
<i>Compression Ratio : 9:1</i>	<i>Connector : 3.5 mm</i>	<i>USB Interface</i>
<i>Maximum Torque: 62Nm @3000rpm</i>	<i>Impedance : 1K ohm</i>	
<i>Valves Per Cylinder : 4</i>		

## 2. Methodology

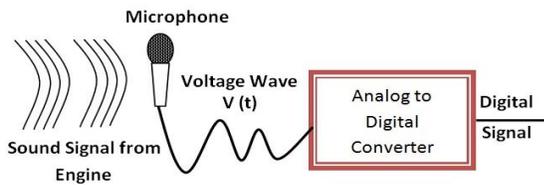


Fig 1: Signal Capturing System

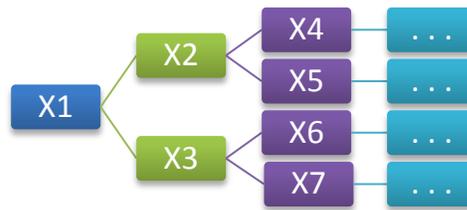


Fig 2: Signal Decomposition Technique

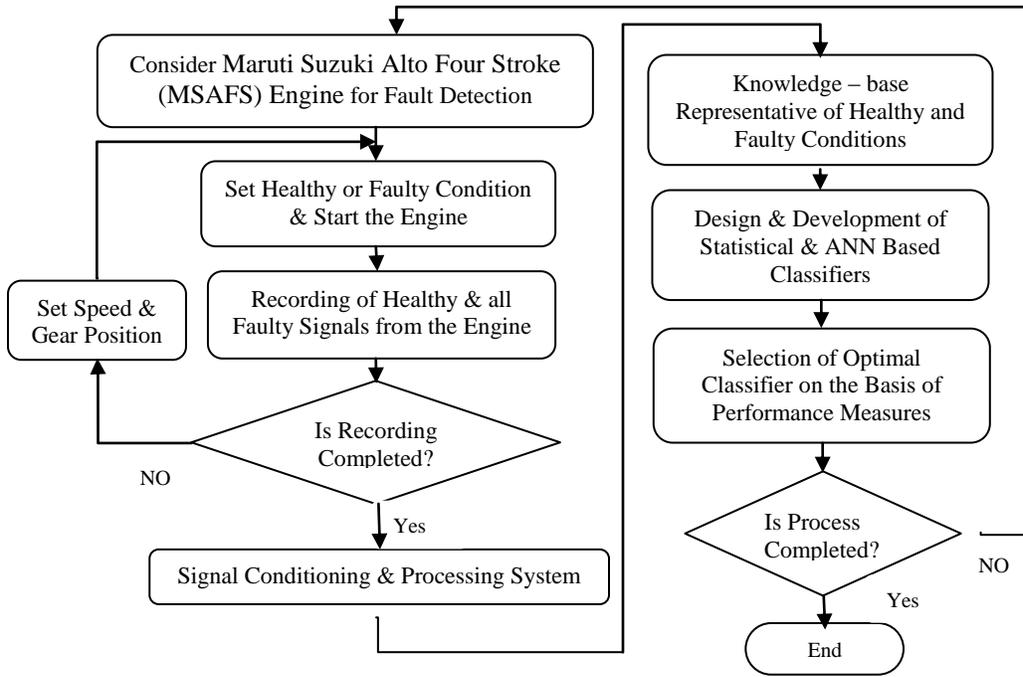


Fig 3: Working of Proposed FDI System

The unidirectional cardioid Microphone has been used as a sensor to acquire the sound variations from an automobile engine in normal and faulty conditions as shown in Fig 1. A unidirectional cardioid microphone is sensitive to sounds from only one direction, which it is facing only. These microphones are more commonly used for live vocal, speech & instrument performances as they are good at rejecting background sounds from other unwanted directions. The MP3 sound recorder is used to record the sound variations of an Automobile Engine in a wave format. The specifications of Microphone and MP3 recorder are given in Table 1. Signal is divided into different frames by using the signal decomposition technique is shown in Fig 2. Initially, the engine is started in healthy (normal) condition and signals are recorded at different speed and each gear positions. The MSAFS engine consists of five different gears including one neutral gear. The four signals were recorded in each gear position at 1200, 1500, 1800 and 2100 RPM. Therefore, there will be a collection of 80 recorded signals for three different faults and one neutral condition. The feature extraction is also carried out of each frame of the signal. These extracted features will be considered as a knowledge database for classification of three different faults. The details of working of Fault Detection system are shown in Fig 3 and experimental setup of engine is shown in Fig 4.



Fig 4: Experimental Setup for MSAFS Engine

### 3. Observations of Recorded Signals and Pattern of Features

The nature of recorded signals is observed in time domain and features are extracted of each signal before and after decompositions of the signals. The recorded signals are plotted as shown in Fig 5A, Fig 5B and Fig 5C. It is noticed from the signal plot that the most of the signals are overlapped and their nature is found to be highly complex and also the amplitude of the faulty signal is greater than that of the healthy condition signals.

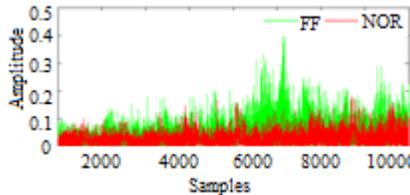


Fig 5A: Signals for NOR & FF

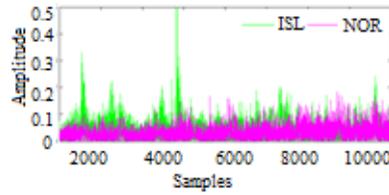


Fig 5B: Signals for NOR & ISL

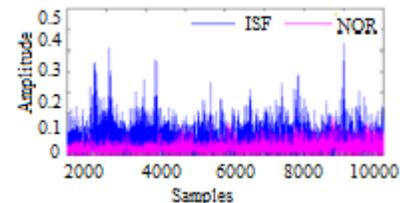


Fig 5C: Signals for NOR & ISF

The knowledge data base is generated by extracting the features of healthy and faulty conditions signals. The extracted features are Mean, Energy, Maximum Value, Minimum Value, Standard Deviation, Variance and Mode for three different types of faults and out of which Variance vs. Energy and Standard Deviation vs. Energy features are plotted as shown in Fig 6A and Fig 6B. After observing the overlapping nature of features and non separable decision boundaries, the decision is taken to employ the soft computing approach to classify the faults. At the beginning the statistical classifiers are employed to classify the faults which are explained in the subsequent section.

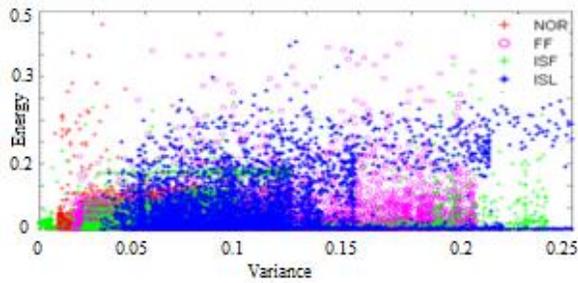


Fig 6A: Scatter plot for Variance vs. Energy

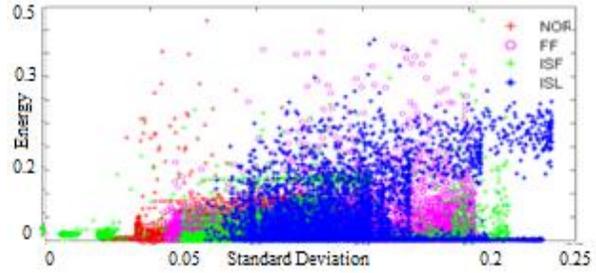


Fig 6B: Scatter plot for Standard Deviation vs. Energy

#### 4. Classification of Faults Using Statistical Classifiers

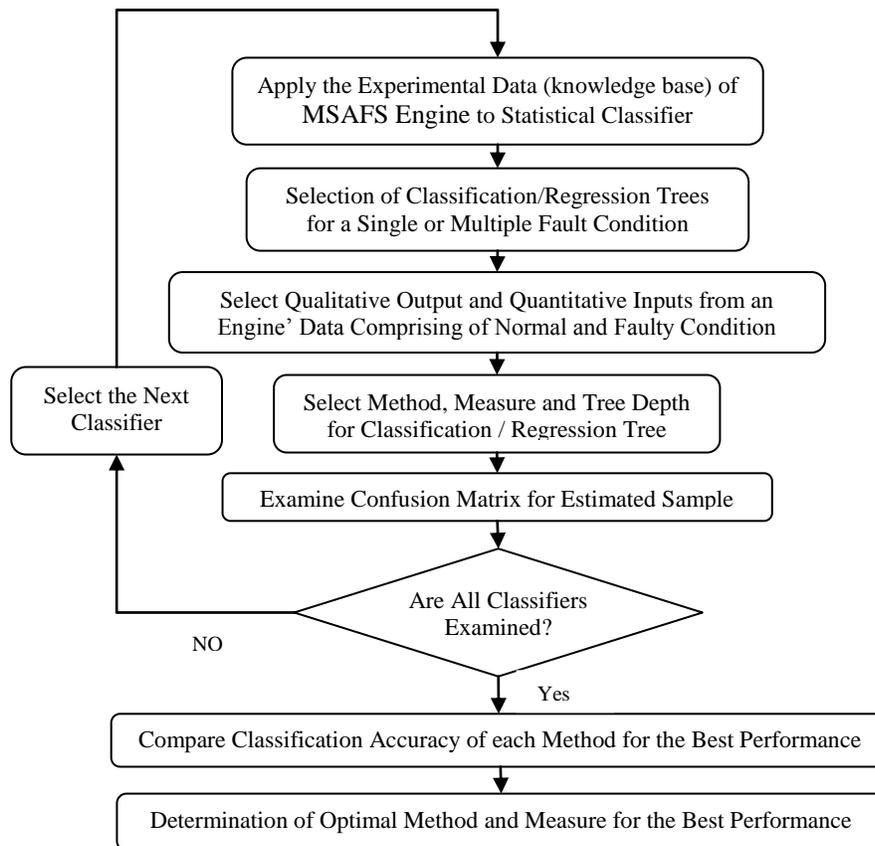


Fig 7: Process Flow Diagram for Statistical Classifiers

Table 2: Classification of Single Fault Condition

Method	Measure	% ACA for FF	% ACA for ISF	% ACA for ISL
<i>CHAID &amp; Exh. CHAID</i>	<i>Pearson</i>	<i>87.50%</i>	<i>85.00%</i>	<i>77.50%</i>
<i>CHAID &amp; Exh. CHAID</i>	<i>Likelihood</i>	<i>87.50%</i>	<i>77.50%</i>	<i>77.50%</i>
<i>C&amp;RT</i>	<i>Gini</i>	<i>87.50%</i>	<i>85.00%</i>	<i>85.00%</i>
<i>C&amp;RT</i>	<i>Towing</i>	<i>70.00%</i>	<i>82.50%</i>	<i>67.50%</i>
<i>Quest</i>		<i>75.00%</i>	<i>82.50%</i>	<i>77.50%</i>

The statistical classification is carried out using *XLSTAT* and Process Flow Diagram for Statistical Classifiers is shown in Fig 7. In this the classification and regression trees have been employed to classify the

faults. The feature matrix comprising of 7 inputs and one categorical output has been applied to statistical classifiers. The size of each feature matrix is 20×8 including seven inputs and one categorical output. Therefore, knowledge database consists of 40 rows and 8 columns. The performance has been observed with respect to *CHAID-Pearson*, *CHAID-Likelihood*, *Exhaustive-CHAID-Pearson*, *Exhaustive-CHAID-Likelihood*; *C&RT (Gini)*, *C&RT-Towing and Quest*.

#### 4.1 Classification of Combined Three Faults in MSAFS Engine

The knowledge database of combined three different faults is applied to the Statistical Classifier for classification. The size of each feature matrix is 20×8 including seven inputs and one categorical output. Therefore, knowledge database consists of 80 rows and 8 columns. The performances of statistical classifiers were observed and results are depicted in Table 3. It is observed that the percentage classification accuracy for *CHAID & Exh. CHAID (Pearson)* is found to be maximum 65.00% amongst all statistical classifier and it is not reasonable. Therefore, the analysis is pursued by using ANN based Classifiers for possibility of improvement in performance and which is explain in subsequent section.

Table 3: Performance of Statistical Classifier for Combined Three Faults

Faults	<i>CHAID &amp; Exh. CHAID (Pearson)</i>	<i>CHAID &amp; Exh. CHAID(Likelihood)</i>	<i>C&amp;RT (Gini)</i>	<i>C&amp;RT (Towing)</i>	<i>Quest</i>
<i>FF</i>	30.00%	40.00%	50.00%	60.00%	45.00%
<i>ISL</i>	85.00%	25.00%	30.00%	30.00%	35.00%
<i>ISF</i>	70.00%	35.00%	75.00%	45.00%	60.00%
<i>NOR</i>	75.00%	70.00%	60.00%	35.00%	70.00%
<i>Total % ACA</i>	65.00%	42.50%	53.75%	42.50%	52.50%

## 5. Classification of Faults using ANN Based Classifiers

Subsequently, the analysis is continued using different configuration of ANN based classifiers such as *Multilayer Perceptron (MLP)*, *Generalized Feedforward (GFF)*, *Modular Neural Network (MNN)*, *Jorden & Elman Network (JEN)*, *Radial Basis Function (RBF)*, *Self Organizing Feature Map (SOFM)*, *Principal Component Analysis (PCA)*, *Time Lagged Recurrent Network (TLRN)*, *Recurrent Network (RN)* and *Support Vector Machine (SVM)*. The working of ANN based classifiers is shown in Fig 8. The percentage Classification Accuracy has been observed for all ten types of ANN based classifiers. The input layer of the ANN contains seven neurons corresponding to seven inputs. One categorical output denotes a type of fault or healthy condition of an engine. As there are three different types of faults and one state indicating healthy condition, the number of neurons in the output layer must be four (three neurons corresponding to three different faults and one neuron corresponding to healthy condition). Three data partitions, namely, Training, Cross Validation (CV) and Testing were used with different tagging order. Every time, ANN is retrained three times with different random initialization of connection weights and biases with a view to ensure true learning and generalization.

### 5.1 Classification of Single Fault in a MSAFS Engine using ANN based Classifiers

The three types of faults are considered for analysis is Air Filter Fault (FF), Insufficient Lubricant Faults (ISL) and Insufficient Fuel Supply Fault (ISF). All ten types of ANN based classifiers are employed to classify the faults. Initially, the classification of single faults has been carried out followed by multiple faults. The size of the feature matrix for a single fault with single frame is 40×8, i.e. 20×8 matrix for normal signal and 20×8 matrix for single fault Signal. The feature matrix was fragmented into three parts in the ratio of 2:1:1. This feature matrix is then applied to each ANN based classifier to classify the faults.

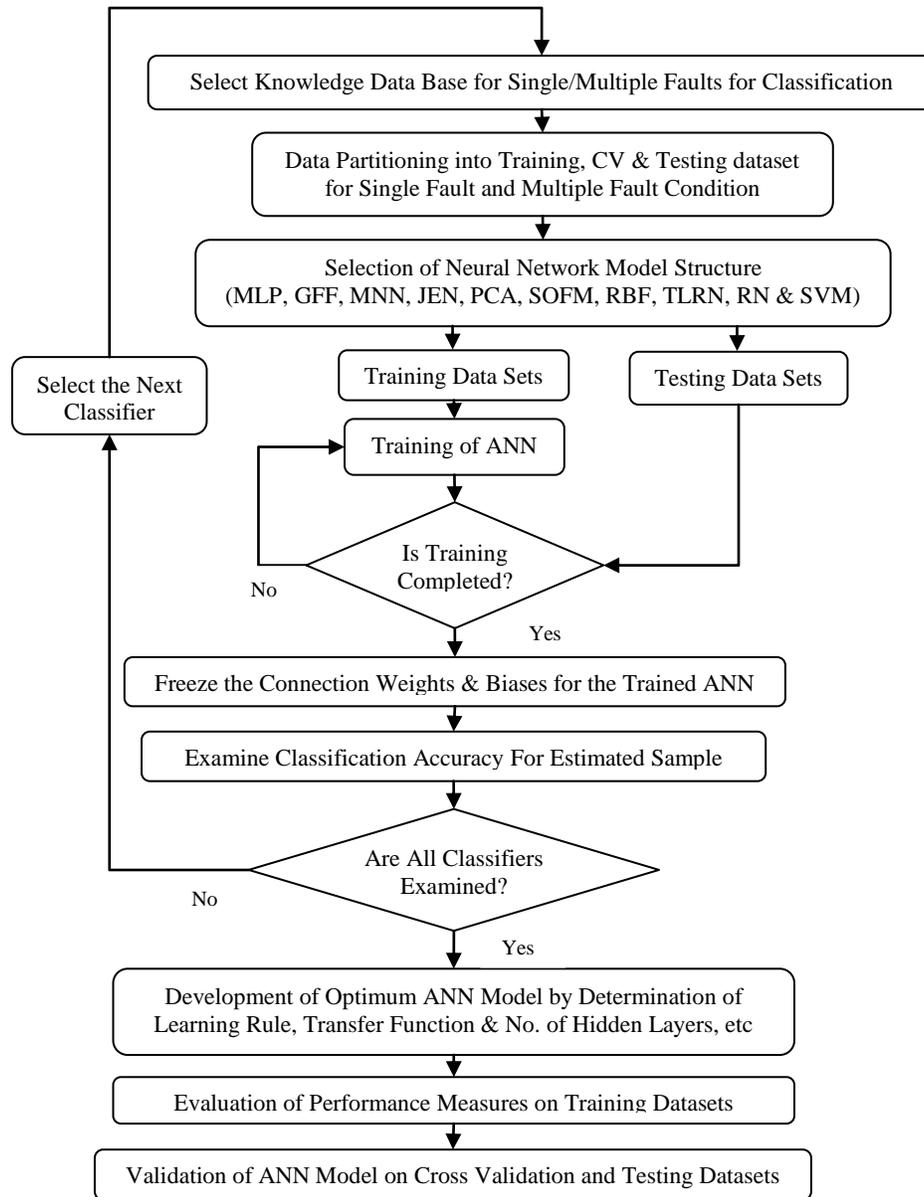


Fig 8: Working of ANN Based Classifiers for Classification of Faults in a MSAFS Engine

Table 4: Performance of ANN Based Classifiers for Single Fault in %ACA

ANN	Air Filter Fault			Insufficient Fuel Supply			Insufficient Lubricants Fault		
	Testing	CV	Training	Testing	CV	Training	Testing	CV	Training
MLP	85.0	95.5	98.7	84.8	82.1	92.5	100.0	100.0	100.0
GFF	85.0	95.5	98.7	82.6	82.1	91.3	100.0	100.0	100.0
MNN	85.0	95.5	98.7	82.6	82.1	91.3	100.0	100.0	100.0
JEN	85.0	95.5	98.7	84.8	79.8	92.5	100.0	100.0	100.0
PCA	85.0	95.5	98.7	89.9	89.9	95.0	89.0	88.2	96.3
RBF	87.5	94.9	97.5	89.9	85.4	97.5	86.1	97.1	98.8
SOFM	87.5	92.7	98.7	83.1	87.1	95.0	100.0	97.8	100.0
TLRN	72.5	89.9	96.2	75.8	77.0	93.8	92.7	89.0	97.5
RN	85.0	89.9	97.5	84.8	79.3	91.3	93.5	97.1	100.0
SVM	92.5	97.7	100.0	87.1	80.3	100.0	88.2	93.5	100.0

The performance of all ten ANN classifiers is depicted in Table 4 with percentage Average Classification Accuracy (%ACA). It is observed that the classification accuracy of each ANN based classifier is reasonably good. Hence the classification is continued for combining all three faults.

**5.2 Classification of Combined Three Faults**

In this case, combined three faults are considered for analysis and knowledge base are combined for three different faults with normal signal for classification. The size of the feature matrix for single frame is 80x8 with 80 rows and 8 columns. The 8 columns comprise of 7 inputs and 1 symbolic output (translated into 4 binary outputs). The feature matrix was divided into three parts in the ratio of 2:1:1. The first part of data were used for training the network, the second one used for cross validation and the third one used for testing the network and one by one all ten types of ANN based classifiers were tested. Fig 9 shows the Classification Accuracy of all ten types of classifiers. From the classification, it is observed that the performance of SVM is better amongst all ten types of classifiers. Therefore, subsequent analysis is continued for different types of frames using signal decomposition technique.

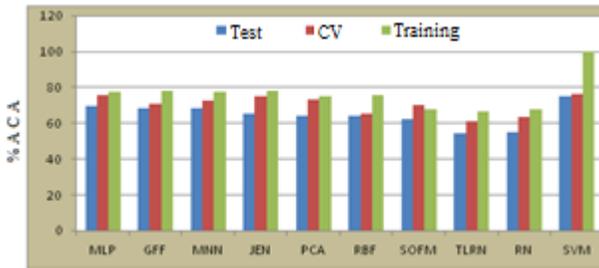


Fig 9: Performance of ANN Based Classifiers

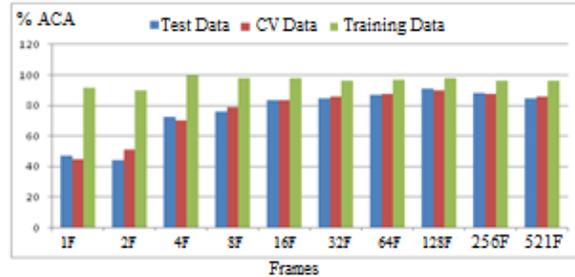


Fig 10: Performance of SVM Based Classifier

**5.3 Performance of SVM for Multiple Faults (FF, ISL and ISF)**

As the performance of SVM based classifier is found to be superior amongst all other ten types of classifiers, the detailed analysis is carried out for Support Vector Machine using the Kernel Adatron algorithm. The Kernel Adatron algorithm transforms inputs into a higher dimensional feature space and then optimally separates data into their respective classes by isolating those inputs which fall close to the data boundaries. Therefore, the Kernel Adatron is especially effective in separating sets of data which share complex boundaries. Therefore, the performance of SVM based classifier is examined for different frames and the performance are plotted in Fig 10. It is observed that %ACA is maximum for 128-frames and further increase in frame size does not seem to improve the classification accuracy. Therefore, subsequent analysis is continued for 128 frames of each signal for combined three faults

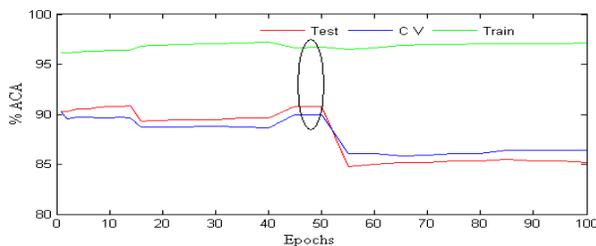


Fig 11: %ACA of SVM Based Classifier

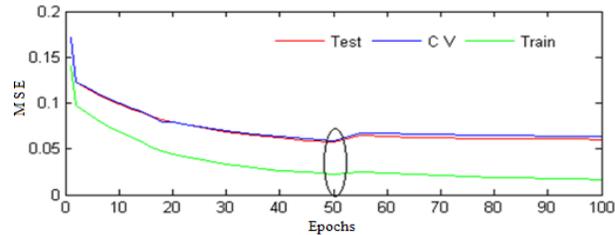


Fig 12: MSE of SVM Based Classifier

The knowledge base for combined three faults with 128 frames comprising of 10240 records was split into three parts in the ratio of 2:1:1. The first part of data was used for training the network, second one was used for cross validation and the third part was used for testing the network. A total dataset has a size 10240 x 8 with 7 inputs and 1 symbolic output (translated into 4 binary outputs). The SVM is retrained three times with healthy & faulty feature matrix derived from an automobile engine. The performance of SVM is exhibited in Fig 11 and Fig 12. It is observed that at the 50<sup>th</sup> epoch, the classification accuracy on the training dataset is found to be above 96% and MSE

is found to be less than 0.05. It is also observed that at the 50<sup>th</sup> epoch, the classification accuracy on the training dataset is found to be reasonable and average MSE is also reasonably less.

## Conclusion

From the relative study and scrupulous comparison of all statistical and ANN based classifiers for automobile engine, it is concluded that the proposed FDI system can provide the best possible solution to early detection of faults in an automobile engine. The main advantage of this system is its simplicity, reliability, cost-effectiveness and compactness requiring a single sensor system. From the meticulous analysis using statistical and ANN based classifiers, it is inferred that ANN based classifiers are more appropriate for fault diagnosis. From the comparative analysis of all ten different types of ANN based classifiers, it is noticed that the Classification Accuracy of SVM based classifier is found to be reasonably acceptable amongst the group of ten ANN classifiers used for the analysis. Therefore, SVM can be used as reasonable classifiers for multiple faults detection in a MSAFS engine. However, MLP and SVM based classifier are seen to be more appropriate classifier for MSAFS Engines as its classification accuracy is higher than all other classifiers. In future, the work can be extended to any number of faults by generating the database for respective faults.

## References

- 1 Rolf Isermann, (2005), "Model-Based Fault Detection and Diagnosis Status and Applications", Elsevier Ltd., Annual Reviews in Control 29 (2005), pp. 71–85.
- 2 X. Zhang, M. M. Polycarpou, and T. Parisini (2002), "A Robust Detection and Isolation Scheme for Abrupt and Incipient Faults in Nonlinear Systems.", IEEE Transactions on Automatic Control, 47 (4), 576 –593.
- 3 Howlett R. J. (1996), "Condition Monitoring and Fault Diagnosis in a Domestic Car Engine Using a Neural Network", IEEE Artificial Intelligence in Consumer and Domestic Products (Digest No. 1996/212), IEE Colloquium Vol 5, pp.1-4.
- 4 R. J. Howlett,(1999), "Monitoring and Control of an Internal Combustion Engine Air-Fuel Ratio using Neural and Fuzzy Techniques" Moulseccomb, Brighton, BN2 4GJ, U.K. Department of Electrical and Electronic Engineering, University of Brighton, International ICSC Symposium on Intelligent Industrial Automation (IIA), Genoa, Italy.
- 5 R. J. Patton, F. J. Uppal & C. J. Lopez-Toribio,(2001) "Soft Computing Approaches To Fault Diagnosis For Dynamic Systems: A Survey" A Proceedings Volume from the 6th IFAC Symposium on *Fault Detection*
- 6 Wang Weijie, Kang Yuanfu, Zhao Xuezheng and Huang Wentao, (2004) "Study of Automobile Engine Fault Diagnosis Based on Wavelet Neural Networks", IEEE- Intelligent Control and Automation, Volume: 2, pp. : 1766 - 1770.
- 7 M. S. Sangha, J. B. Gomm, D. L. Yu, G. F. Page (2005), "Fault Detection and Identification of Automotive Engines Using Neural Networks", Proceedings of the 16th IFAC World Congress, pp. 1166-1166.
- 8 Jian-Hua Zhang, Li-Fang Kong, Zhang Tian and Wei Hao, ( 2010) , "The Performance Parameter Fault Diagnosis for Automobile Engine Based on ANFIS", Web Information Systems and Mining (WISM) -IEEE International Conference on Volume 2, pp. 261 – 264.
- 9 Zhe Wang, Li Fang Kong, and Shi Song Zhu, ( 2011), "Fault Diagnosis Model for Automobile Engine Based on Gradient Genetic Algorithm", IEEE Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2nd International Conference, pp. 1128 – 1131.
- 10 Hamad A., Dingli Yu, Gomm, J.B. and Sangha, M.S. (2012), "Fault Detection and Isolation for Engine Under Closed-Loop Control", IEEE Xplore, Control (CONTROL), UKACC- Proceedings of IEEE International Conference, pp. 431 – 436.
- 11 M. S. Sangha, D. L. Yu and J. B. Gomm (2012), "Sensor Fault Detection, Isolation, Accommodation and Unknown Fault Detection in Automotive Engine Using AI", International Journal of Engineering, Science and Technology, Vol. 4, No. 3, 2012, pp. 53-65.
- 12 Madain M., Al-Mosaiden A. & Al-khassaweneh M. (2010), "Fault Diagnosis in Vehicle Engines Using Sound Recognition Techniques", Electro/Information Technology (EIT), IEEE International Conference, pp. 1 – 4.