Modeling a Fault Detection Predictor in Compressor using Machine Learning Approach based on Acoustic Sensor Data

Divya M.N[1]

Research Scholar, VTU Research Centre, MSRIT, VTU, Belagavi, Karnataka, India
School of ECE, REVA University, Bengaluru, India

Narayanappa C.K[2]

Department of Medical Electronics, M.S. Ramaiah Institute
of Technology, Bangalore, Karnataka, India

Gangadharaiah S.L[3]

Department of Electronics and Communication,
M. S. Ramaiah Institute of Technology, Bengaluru, India

Abstract—**Proper functioning of the air compressor ensures stability for many critical systems. The ill-effect of the breakdown caused by the wear and tear in the system can be mitigated if there exists an effective automated fault classification system. Traditionally, the simulation-based methods help to extend to identify the faults; however, those systems are not so effective enough to build real-time adaptive methods for the fault detection and its type. This paper proposes an effective model for the fault classification in the air compressor based on the real-time empirical acoustic sensor time-series data were taken on a sampling frequency of 50Khz. In the proposed work, the time-series datais transformedinto the frequency domain using fast Fourier transforms,where half of the signals are considered due to its symmetric representation. Afterward, a masking operation is carried out to extract significant feature vectors fed to the multilayer perception neural network. The uniqueness of the proposed system is that it requires less trainable parameters, thus reduces the training time and imposes lower memory overhead. The model is benchmarked with performance metric accuracy, and it is found that the proposed masked feature set-based MLP-ANN exhibits an accuracy of 91.32%. In contrast, the LSTM based fault classification model gives only 83.12% accuracy, takes more training time, and consumes more memory. Thus, the proposed model is realistic enough to be considered a real-time monitoring system of the fault and control. However, other performance metrics like precision, recall, and F1-Score are also promising with the LSTM based fault classifier.**

Keywords—*Air-compressor; fault detection; LSTM; multi-layer perception; ANN; acoustic sensor data*

I. INTRODUCTION

The air compressors (AC) play a significant role in essential functions like fuel injection and metal finishing in the aircraft's design [1]. The ACs are used widely in thermal plants [2], power generation systems [3], vehicle propulsion [4] and pipeline systems [5], etc. In building an effective quality control system, the compressor simulation plays an essentialrole in evaluating the tolerable pressure by the different components of the aircraft while in transit [6]. Specifically, the aircraft manufacturer depends on the very high quality of the compressors for every phase of the

production for the operative success of the functions. Another important aspect is that the aircraft components thwart the contamination due to mixing the air with the lubricants in ACs. Fig. 1 demonstrates the common applications in aircraft ACs.In order to design the complete line of the product of anair compressor, a compress air system is used. Two main air compressors are widely used: i) Rotary screw AC, and ii) Reciprocating ACs used depending upon the application requirements.

The automated fault classification (AFC) problem in compressors has attracted researchers to address the issues and build solution paradigms in the recent past so that early detection can minimize the damage caused to the overall system. The early warning systems are a step towards preventive maintenance [7], which is broadly classified into two categories,namely: i) maintenance on breakdown [8] and ii) condition-based maintenance (CBM) [9-10]. However, the CBMhas taken an edge over the maintenance on breakdown because the CBMperforms both the detection and seclusion of the faults that occurred in the early stage of the breakdown itself. The early and intelligent fault detection system [11] synchronizes with the conditions [12] of the various machine aspects dynamic changes in pressure, temperature, vibration, or acoustics [13]. Dhosi et al. [14] reviewed the correlation of vibrations and fault for various machines like pumps, turbines, and compressors. However, the most predominant measurement for fault detection is an acoustic signal.Some of the recent studies include the fault diagnosis in the planetary gearbox [15], Polak et al. [16], have highlighted the fault detection in the compressor, combustion engines using acoustic signals, last but not least, Ahmed et al. [17] reviews the use of acoustic measurement for the power unit fault detection in aircraft.

This paper proposes fault classification in the air compressor based on the real-time empirical acoustic sensor data. However, the challenging point in designing such an intelligent fault detection system (IFDS) is to identify the sensitive points from where the signals are acquired, and that is performed by the mechanism of the sensitive point or position analysis (SPA) [18-19]. The raw data collected from all the

sensitive positions are exposed to various noises requiring appropriate denoising treatments [20]. The learning model does not take these de-noised data directly into its computing model. Therefore, the significant information representation of the acoustic signals using mathematical models plays a vital role in accuracy for both dimensionality reduction and predictor performance. Various representations of the signals, including time, frequency, and time & frequency domains, are found in the designing of the fault detection systems for pipeline leakage [21], mechanical compound fault [22], centrifugal compressor [23], and reciprocating compressor [24]. Theoretically, the detection probability and accuracy are assumed to be higher if more features are considered while designing the detection models [25]. However, in contrast, the machine learning-based model performances degrade with the higher number of features [26] due to ill-effect posed by high polynomiality complexities and approximation.

Furthermore, this is handled by using an effective feature selection mechanism towards dimension reduction techniques like Partial least squares (PLS) used in the gas turbine's compressor blade [27]. Many other popular ones used in machines related faults detection systems are PCA [28], the variance of the ICA [29], etc. Moreover, there exist various types of faults which is trained to the model using these feature sets and many of the learning models as a function approximator such that if Yn represents all the feature set, then a function Fn emerge out as Fn(Yn) → C, where C is the class or the type of the faults. Popular learning models used in the fault classification in the machine are linear-SVM used for vehicle power systems [30], CNN for bearing fault classification [31], and LSTM for compressor valves [32]. Towards achieving better classification performance, this paper has presented an effective model for the fault-classification in the air compressor using a multilayer perceptron neural network that overcomes memory overhead, unlike Long Short Term Memory (LSTM) based approaches. The paper utilizes the autistic signal data prepared by Verma et al. [33] from 24 sensor positions based on SPA from an air compressor. The entire process flow of this method is described in the respective sections in this paper as a snap-view given in Fig. 2.

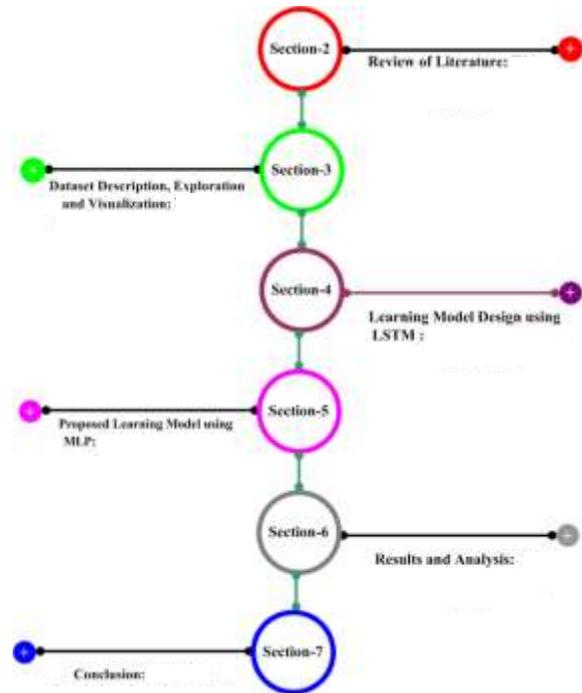Fig. 1. Air Compressor Application in Aircraft.

Fig. 2. Snap View of the Paper Organization.

As mentioned in the above Fig. 2 the remaining part of the paper are organized as follows: Section II discusses the related work in the context of AC fault prediction, Section III presents dataset visualization and analysis; Section IV discusses the implementation of LSTM model; Section V discusses implementationof the proposed model based on MLP; Section VI presents the result and performance analysis of the proposed system and finally entire work of this paper is concluded in Section VII.

II. REVIEW OF LITERATURE

This section discusses the related work carried out towards air compressor fault prediction. In the existing literature, there are two kinds of approaches used to predict fault AC classification. The first one is the model-based approach, and the second is the data-driven approach [34]. The model-based approaches utilize mathematical modeling for machine life estimation and fault prediction [35-36]. In contrast, data-driven approaches are based on statistical analysis and soft computing approaches like machine learning, deep learning, and evolutionary. However, model-based approaches involve complicated procedures to describe attributes of the mechanical system [37]. Conditions of a mechanical system like air compressors can be analyzed by processing sensory data using data-driven and soft-computing approaches [38]. The work carried out by Ouadine et al. [39] has applied a neural network optimized using a genetic algorithm for predicting Aircraft AC-bearing fault. The dataset used in this study consists of vibratory signals captured from different bearings defects. The features are determined based on spectral density estimation, and prediction outcomes were compared with discriminant analysis classifier.Li et al. [40] introduced an intelligent fault detection system for mechanical rotating systems. The study uses a recurrent neural network (RNN) with fuzzy logic. The
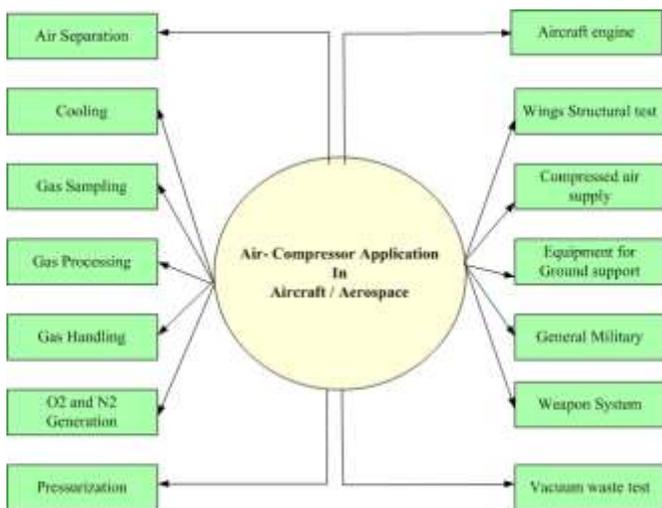
RNN filters the input signal, and then the filtered signal is fed to fuzzy logic to detect faults. Ghorbanian and Gholamrezaei [41] investigated the application of different machine learning mechanisms in the context of analyzing performance compressors. The authors have utilized general regression neural (GRN) and MLP to simulate the performance of these models. The result indicated GRN is less associated with mean error and performed well with the experimental data but limited to interpolationfactor. On the other hand, MLP was evaluated, and the result indicated the most favorable outcome to analyze compressor performance.The work of Aravinth and Sugumaran [42] adopted a statistical feature extraction approach and random forest (RF) classifier to monitor and predict the fault in the AC to avoid regular failure in industrial and domestic applications. In this study,the accelerometer sensory signal is processed via a statistical approach, and RF is applied to detect the type of fault in AC.Fan et al. [43] have considered the case study of vehicle communication and presented their work on predicting AC breakdown using data streamed by the vehicles. The authors have used histogram analysis to model the signal. However, the histogram is a more straightforward approach that can determine the deviation in the signal to some extent. The study of Cui et al. [44] suggested an intelligent model for the early detection of faults in AC. The approach used in this study is based on the construction of an adaptive matrix based on the PCA and backpropagation techniques. This matrix is constructed to store the signals and determine a function of deviation in the signal pattern.

Further, identify early fault signature, a threshold is computed based on the mechanism of the sliding statistical window method. Work towards evaluating trustworthiness and prediction reliability on the AC in the Ammonia Plantis considered by Musyafa et al. [45]. Chen et al. [46] presented an LSTM-oriented approach for classifying compressor breakdown using aggregated sensory data.The performance of the presented model is evaluated using information captured from large heavy-duty vehicles. The authors have formulated a classification task to identify whether a compressor fault will occur within the specified horizon. An LSTM learning model is used to predict, and its performance is evaluated against the RF classifier. The experimental outcome exhibited that RF slightly outperforms LSTM regarding AUC. However, the prediction outcome from LSTM shows stability over time, maintain stability in the trend of healthy faulty classification. Another work carried out in a similar direction by Yang et al. [47] suggested an AC fault classification mechanism using on lifting wavelet approach. Initially, this study has decomposed the vibration signal of the AC wavelet; and further statistical features of decomposition are computed as the AC faults. In the classification process,the probabilistic model-based supervised classifier is employed to predict the fault class. The study outcome suggested that faulty features determined using a wavelet-based approach provide comprehensive fault features that lead to higher accuracy by the supervised classifier.

III. DATASET DESCRIPTION, EXPLORATION, AND VISUALIZATION

This section describes the process of data creation, its description, and an exploratory analysis and Visualization that decides the line of research for the data presentation for the learning model to overcome the memory constraints of the traditionally applied LSTM based fault classification system for higher accuracy.

A. Data Description

The Department of Electrical engineering of the Indian Institute of Technology (IIT), Kanpur, India having an air compressor of the single-stage reciprocating type. An effort by Dr. N.K Verma and his team to provide an open-source dataset [33][34] on the following specification as in Table I.

In the work of Verma et al. [33], all sensors, as in Fig. 3, records raw data every 5 seconds at the sampling rate of each sensor at 50 kHz and gets stored into the respective eight files as a structure shown in Fig. 3.

A closer analysis of each record shows that it consists of precisely 225 'dat' files in all the folders. Since there are 24 sensors at the sensitive points and an additional one sensor kept at a far distance to record the acoustic data at every 5 seconds, there are nine different timeslots; thus, as per Fig. 4, there are 9 x 25 = 225 files for each fault.

TABLE I.　　SPECIFICATIONS

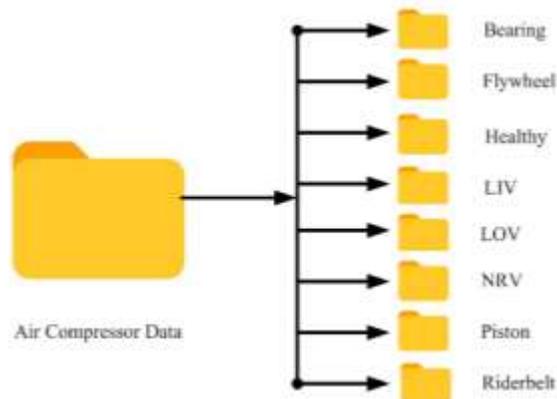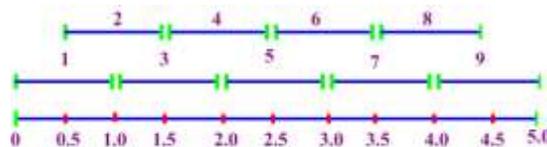| SI. No	Specification	Values
1	Air pressure range	0-500Ib/m^2, 0-35 Kg/cm^2
2	Induction Motor	5Hp, 415V, 5Am,50Hz,1440rpm
3	Presure Switch	Type Pr-15, Range 100-213PSI

Fig. 3.　A File Structure for the Data.

Fig. 4.　Time Slot for the Data Records from the Sensors.

TABLE II.   CLASSES OF THE FAULT

Sl. No	Fault class	Abbreviation
1	Bearing fault	BF
2	Piston Ring fault	PRF
3	Flywheel fault	FWF
4	Rider Belt fault	RBF
5	Leakage Inlet Value fault	LIVF
6	Leakage Outlet Valve fault	LOVF
7	Non-Return Valve fault	NRVF

The raw data undergoes various pre-processing stages, including filtering to eliminate the undesirable frequency component using FIR filter at 400 Hz cut-off frequency (COF) threshold and low pass filter with COF of 12 kHz to obtain the valuable information.Further, clipping, smoothing, and normalization operations are performed to obtain the pre-processed data. The operation for extracting the name of the fault classes is applied and as in Table II. Moreover, the fault class is categorized as seven faults classes and consists of the normal class.

1) Bearing fault: Bearing fault in the compressor arises when there is malfunctioning in the bearings, which are meant to make the compressor wheels running smoothly. Either bearing may break or may get imbalanced due to wear and tear. Due to this, friction in the machine will increase, and noise will arise.

2) Piston fault: Piston is the major part of the mechanism which converts rotatory motion into linear or vice versa. If there is a fault in the piston RPM of the entire machine may reduce. Moreover, due to this, the full sound of the machine will get less loud.

3) Flywheel: Flywheel is the main storage of kinetic energy in any machine. The main source of rotary motionis its motor or IC engine, which may not provide continuous energy to the machine. Hence if there is a fault in the flywheel due to wear and tear, the wheel spins faster; however, it can store less kinetic energy. Since it spins faster, the frequency of the sound may increase.

4) Leakage in inlet valve: This fault occurs when the inlet valve of the compressor leaks, the pressure in the cylinder also reduces significantly. The noise becomes lessloud since the compressor is no longer working at optimal efficiency. The speed of the piston will reduce, and the frequency of the noise also reduces.

5) Leakage in outlet valve: Contrary to the previous problem, high frequency and loud noises will appear when the leakage is in the outlet valve. This is because the pressure in the cylinder and speed of the piston remains the same, but still, the air escapes from the outlet valve with high pressure. This causes extra noises of various frequencies to appear.

6) NRV fault: NRV refers to a Non-return valve, which means that the valve will close when air tries to flow in the opposite direction. The fault arises when the air starts hitting the NRV valve in the opposite direction, which might be caused due to blockage or damage. In either case, there will be an impulsive load on the valve. This will induce noises of low frequency to appear along with the rest of the noises.

All these faults including normal is encoded as {0,1,2,3,4,5,6,7}. In order to take the recording, typically, faults were inducted into the AC. Fig. 5 illustrates the placement of the 24 acoustic sensors (AS) setup at four typical locations as {top of the piston: 6, NRVF: 6, Opp. of NRV Side:6, Opp. Of FW Side: 6}, where 6 → number of AS.

B. Data Exploration and Visualization

Initially, all the 225 data files stored into the respective directories are read and converted into the 2D vector of size 50,000 x 1, as each file contains reading at every '1' sec at a 50KHz sampling rate. The explicit procedure for this operation is as in algorithm 1. That can be understood using the flow chart in Fig. 6.

Algorithm-1: 2D Vector representation of the processed data

Input: C
Output:P, R
Start
Initialize P, R
for ∀ fault_Dir ∈ C
f_{read}(fault_Dir)
for ∀ file ∈ fault_Dir
Fn←f_{ioin}(fault_Dir, fault, file)
Text ←f_{read}(Fn)
CSV ←f_{split}(Text : ,)
V ←f_{float}(CSV)
Append V → *P* &fault→*R*
End
End
End

| Top Of Piston | NRV Side | Opp. of NRV Side | Opp.of Flywheel Side |

Fig. 5.   Position for the Sensitive Position Analysis on the different Sides of the Air Compressor.
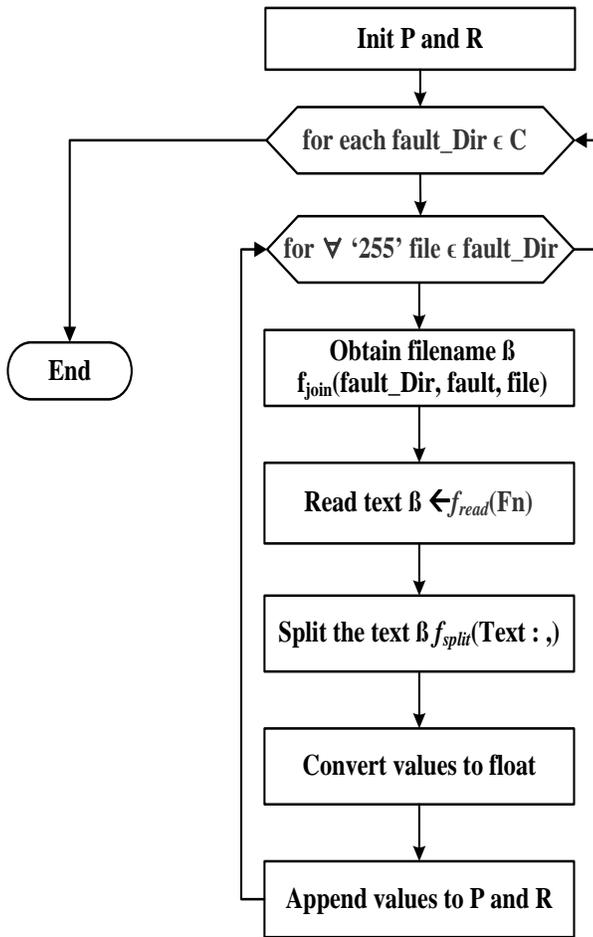
Fig. 6. Process Flow of the 2D Data Representation of the Pre-process Sensor Data of Faults.

In the process of this stage of the data representation, all the data stored into respective folders(fault_Dir) ∈, the main folder (C) are read.Then ∀'255' file ∈ fault_Dir are read to obtain the file names (Fn) of the data by joining the strings: {fault_Dir, fault, and the file name.wav}, however when the 'Fn' is read, it is in the string format that gets converted into the comma-separated string that gets converted into the number types as values(V). Further,the value and the fault are updated to the initialized prediction vector (P) and response (R).

C. Signal Transformation

This section presents the transformation of the time-domain signal into the frequency domain, as demonstrated in Fig. 7.

The time-domain audio signal is transformed into the frequency domain using the numerical expression given in equation 1.

$$x[k] = \sum_{n=0}^{N} x[n] \cdot e^{\frac{-2\pi j}{N} \cdot Kn} \tag{1}$$

Where $x[k]$ refers to the frequency domain using fast Fourier transformation, N denotes samples per second (sampling frequency), $n \in [0, 50000)$, $x[n]$ indicates samples in the time domain, $e^{\frac{-2\pi j}{N} \cdot Kn}$ is the Euler's formula, and coefficient of e, $\frac{-2\pi j}{N} \cdot n$ denotes rotation, and finally, K indicates the amplitude of AS signal at a particular frequency.

1) Time domain analysis: The time-domain signal represents variation in quantity concerning time as a waveform. The advantage of considering AS signals in the time domain are highlighted as follows:

Advantages

- Minor changes in the AS signaling pattern can be represented in the time domain.

- If there is time-sensitive data, a particular noise occurring only in the first few seconds of the machine's starting can be represented in the time domain.

- Phase shifts can be easily recognized in the time domain representation.

However, even considering these advantages time domain may not be a suitable representation of the data since the readings are taken after sometime of the machine being started. In a mechanical system like AC, there are absolutely no changes of occurrences of phase shifts. Phase shifts occur only in electronic systems. In a complex machine-like AC, many types of audio signals are mixed. There may be sound coming from the main cylinder, sound from the flywheel, and minor sounds due to friction between moving parts. The disadvantages of AC signal analysis in the time domain are described as follows:

Disadvantages

- AC signals are captured at different frequencies and often change depending on the sampling frequency of the sensors.

- Time-domain analysis may not be suitable to determine the fault accurately because of a high number of captured signal overlapping.

- The ambient noises are removed in the adopted dataset [BP]. However, there are common sounds recorded by all the sensors, for example, the sound of the motor being captured by ASs since it transmits efficiently through the metal shell of the AC.
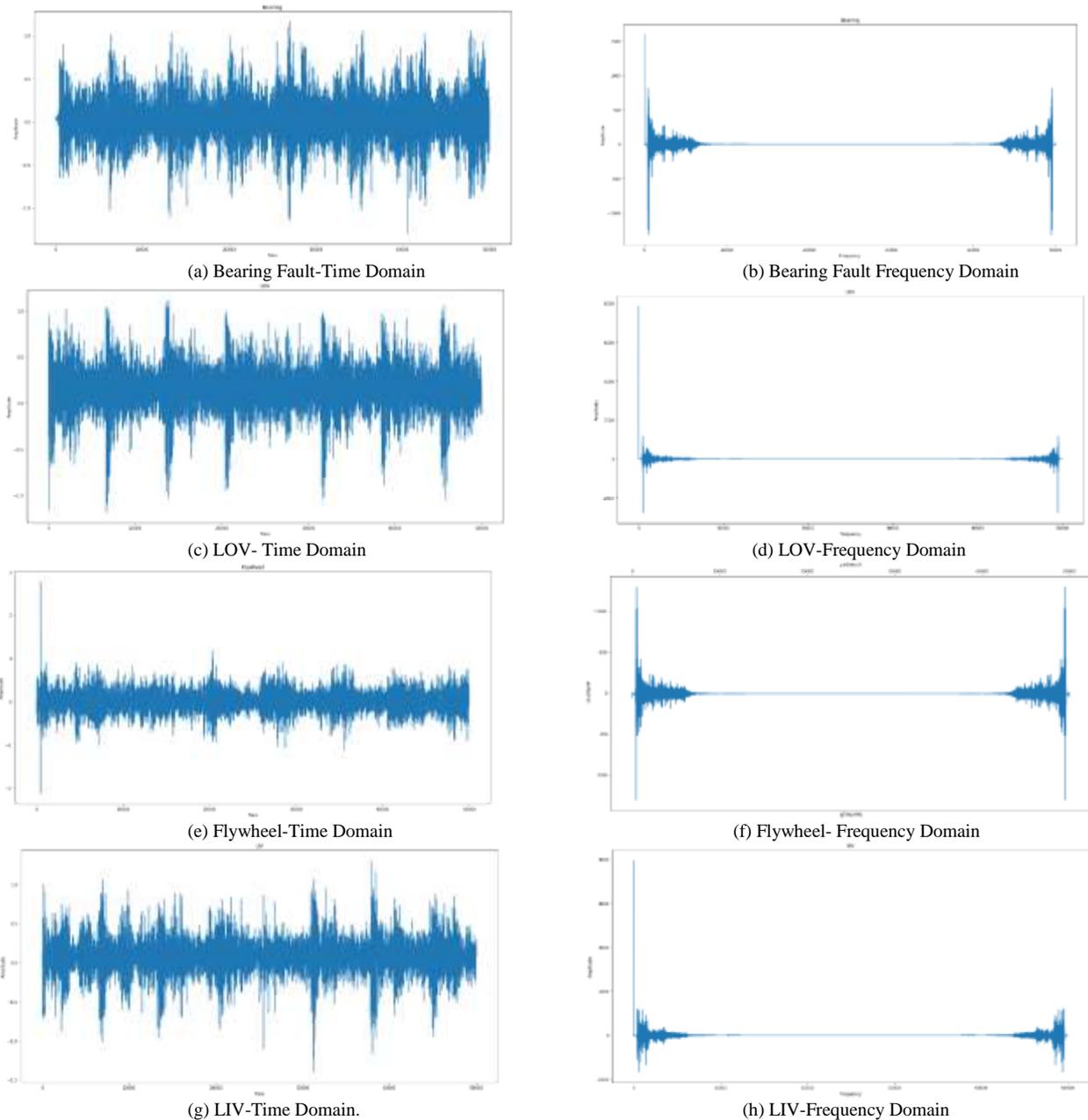
(a) Bearing Fault-Time Domain

(b) Bearing Fault Frequency Domain

(c) LOV- Time Domain

(d) LOV-Frequency Domain

(e) Flywheel-Time Domain

(f) Flywheel- Frequency Domain

(g) LIV-Time Domain.

(h) LIV-Frequency Domain

Fig. 7. Represents the Acoustic Signals∀ Fault Classes in a Time Domain and Frequency Domain Such that AS ∀ Fault Classes ∈ Time-Domain = {7a, 7c, 7e, 7g, 7i, 7k,7m,7o}, Whereas AS ∀ Fault Classes ∈ Frequency Domain = {7b, 7d, 7f, 7h, 7j, 7l,7n,7p}.

2) Frequency domain analysis: The AS signals in the frequency domain represent the amplitude of the quantity over various frequencies. The signal in the frequency domain is called a spectrum. There are many advantages of representing AS signal in frequency domain described as follows:

Advantages

- Any frequency domain transformation works as a frequency un-mixer.

- Easier to find out which instrument faults by looking at variations in the natural frequencies in the spectrum. For example, when there is a fault in the bearing, the friction increases and produces high-frequency noise from rubbing the metal pieces. So, the higher frequency noise becomes more dominant when there is a bearing fault present. Hence, such fault can be easily recognized.

Disadvantages

- In the frequency domain representing phase, shifts are quite challenging tasks. However, phase shits are not important in a mechanical system.

Further, a descriptive statistical analysis is performed better to understand the overall data through the summarization

process and generate actionable information from the signal representation data. It provides '225 x 8 = 1800' samples.

Each has nine statistically significant computations belonging to a set {min, Q1, Q2, Q3, max, count, standard deviation, mean, fault-type}. Table III provides some random samples subset from the complete descriptions.

The count of all the samples is 50,000, indicating that there is no need to work on the cleaning process as there are no missing values in the value point in the sample or data point.

However, a better correlation is analyzed through the histogram of a sample for each fault type as shown in Fig. 8 provides a better visual perception of the data pattern.

TABLE III. A SUBSET OF A SAMPLE OF DESCRIPTIVE STATISTICAL ANALYSIS OF THE DATA

Sample No	count	Min	Q1	Q2	Q3	Max	Std Dev.	Mean	Fault type
0	50,000	-1.5920	-0.049589	0.049434	0.147045	1.3448	0.186775	0.048568	1
1	50,000	-1.2099	-0.082617	0.034085	0.149682	1.2805	0.206034	0.032862	1
2	50,000	-1.1444	-0.070414	0.059485	0.188635	1.1400	0.223090	0.058771	1
3	50,000	-1.1913	-0.095192	0.045985	0.188565	1.1978	0.242442	0.044639	1
4	50,000	-1.1168	-0.025945	0.082214	0.190672	1.2462	0.194634	0.080915	1
...	...	...	...	...	...	...	...	...	...
1795	50,000	-1.3833	-0.105585	0.000489	0.107895	1.3087	0.192076	0.000854	8
1796	50,000	-1.2674	-0.111110	0.002802	0.116820	1.3919	0.206447	0.003726	8
1797	50,000	-1.5291	-0.052908	0.049118	0.155342	1.2516	0.194609	0.049839	8
1798	50,000	-1.2897	-0.166673	-0.074163	0.019368	1.6754	0.167009	-0.074875	8
1799	50,000	-1.4939	-0.075400	0.035780	0.148352	1.5342	0.204002	0.035833	8

8 (a) Healthy   8(b) Flywheel   8 (c) LIV

8 (d) LOV 8 (e) NRV 8 (f) PISTON
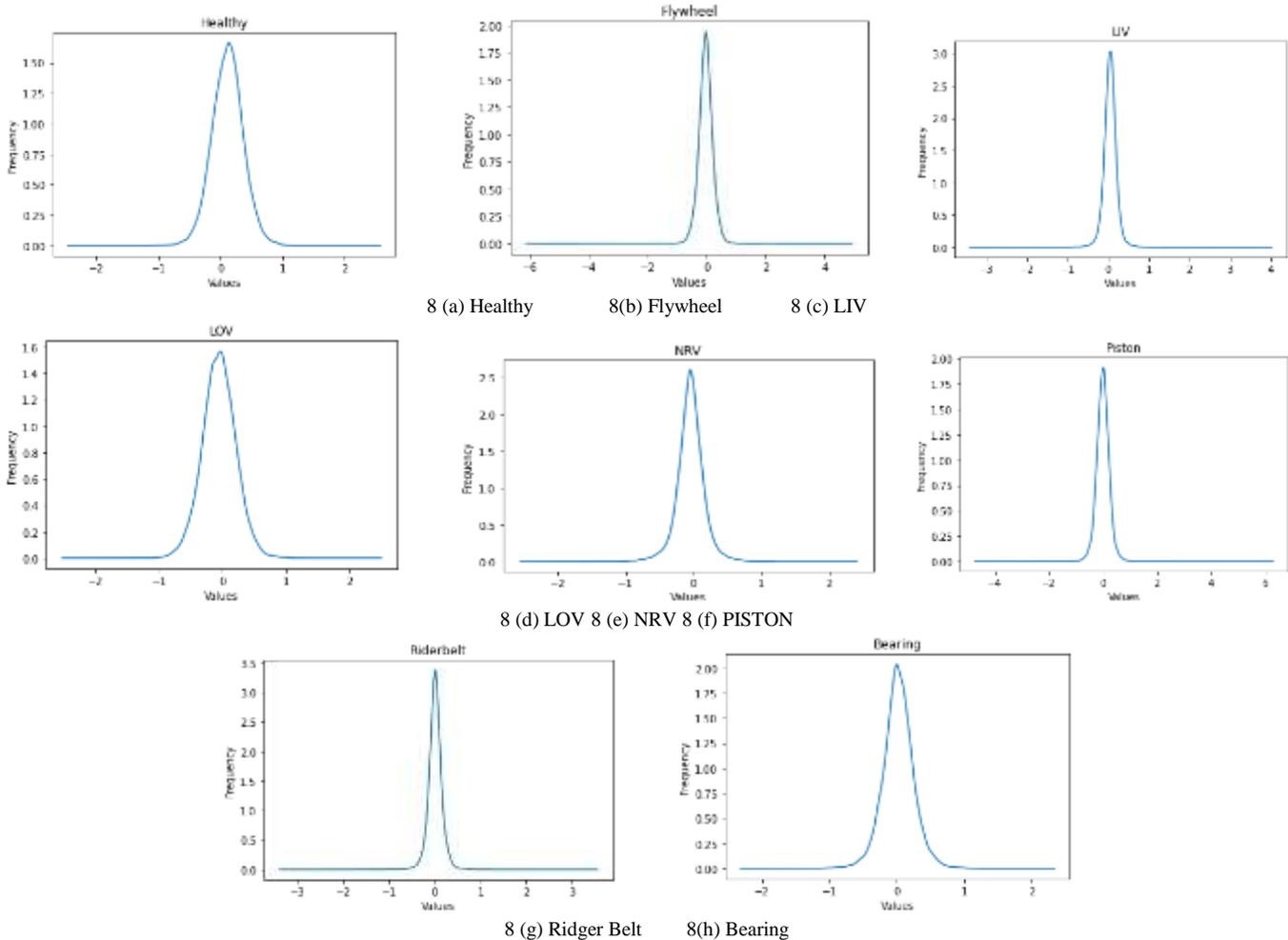
8 (g) Ridger Belt   8(h) Bearing

Fig. 8. Histogram Plot of a Sample Set of each Category of Fault.

As shown in Fig. 8, the amplitude ranges from a higher value to a lower value depending on the fault type. The histograms for the respective fault types, as in Fig. 8(a) to Fig. 8(h), indicates the repetitions of amplitudes with thecentral tendency of each curve to zero. However, Fig. 8(d) and 8(h) show multiple central tendencies with zero and a little higher peak in Bearing and a lower peak in the case of LOV fault. The detailed observatory description for the rest ofthe distribution is as below:

- *Flywheel:* As it can be observed that the flywheel curve is wider compared to both above histograms. (When we call it wider, observe the x-axis. The curve is landing at -2,+2) due to this, it can be concluded that when there is a flywheel fault, the noise of a particular frequency from the machine gets louder. This is an important indicator.

- LIV: It can be observed here that the noise will become less loud compared to normal operation. This is quite understandable since LIV stands for leakage in the inlet valve. Moreover, due to this, pressure will reduce, and the loudness of the machine will also reduce.

- LOV: Contrary to the previous example, when there is a leakage in the outlet valve (LOV), it will induce another high-frequency noise. Upon closer inspection of the peak, there are two peak points present. The lower one is for the formal operation, and the higher one is for the noise. The air will escape with a much higher velocity from the outlet valve. Moreover, due to this, high-frequency noise is induced.

- NRV: NRV or non-return valve occurs only when the air hits the NRV with the impulsive load. The purpose of NRV is to ensure the unidirectional flow of air. Except for this, the machine is in normal working condition. Hence, this is very similar to a normal

operation. However, due to lack of pressure in output, some noises are not present.

- PISTON: In this case, however, the outside piston is malfunctioning. Since both flywheel and piston are external components to the main turbine, this histogram looks very similar to flywheel fault.

- Ridge Belt: Ridge belt is the belt connecting the flywheel to some machine tool or energy converter. If this is at fault similar effect of flywheel fault is produced in the case of acoustics.

- Healthy: In a healthy air compressor, it can be observed that the central tendency of the KDE plot is little towards the positive side of the plot. This means that there is a very low-frequency noise is present when the compressor is working normally. This could be due to the rotation of the wheels and bearings.

IV. Learning Model Design using LSTM

LSTM is a specific Recurrent Neural Network (RNN) class, which is most suitable for predicting time-domain sequences and their long-term dependencies more accurately than ordinary machine learning models.RNN considers that the association amongcellsisformulated as a directed graph. The previous state of the cell may be recurrent, which gives the network the ability to "remember" the information. With this exclusive structure, RNN can make decisions based on previous output value and current Input. However, RNN encountered the issue of exploding andgradient vanishing during the training phase.LSTM has been conceptually designed to address the issue of vanishing and exploding gradients. The LSTM network has a unique structure called a cell (neuron), allowing it to control the flow of information in the network. The elementary unit structure of the LSTM cell is shown in Fig. 9.
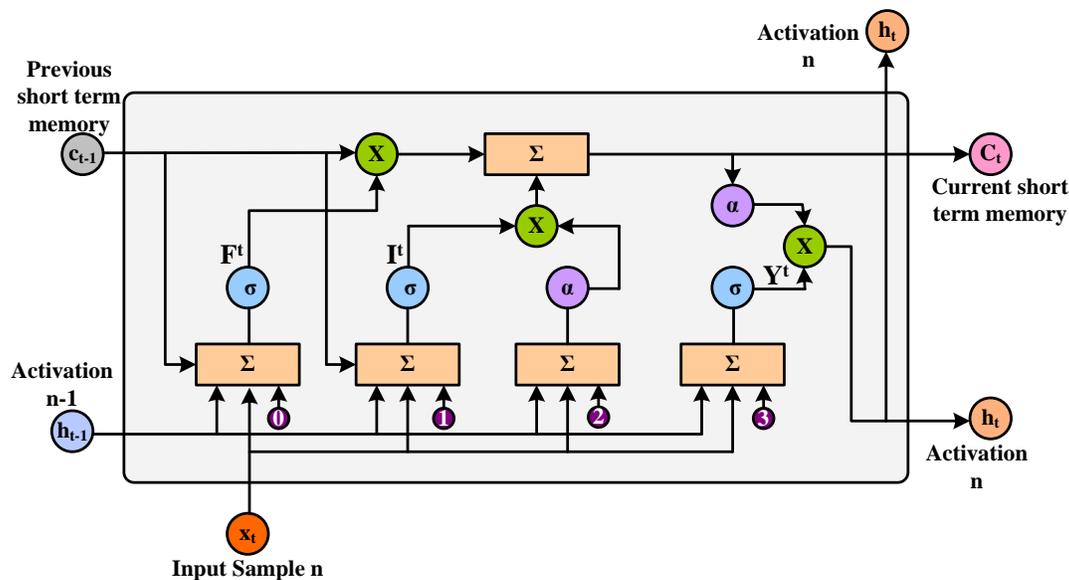

Fig. 9. Basic Structure of LSTM Cell.

In Fig. 9, the basic structure of the LSTM cell is demonstrated that utilizes vector connection by different functions such as sigmoid 'σ'and hyperbolic tangent 'α' with point-by-point addition '∑' and multiplication '⊗' operations. The cell has knodes such as input nodes (x_t) that takes input samples in the form of vector to the LSTM, activation-n (h_t) shows the output of a node, the current short-term memory, or current state of cell (C_t) where both $\{h_t, C_t\} \epsilon \mathbb{R}^{k \times 1}$, previous short-term memory (C_{t-1}) indicates the previous state of cell and activation n-1 (h_{t-1}) Shows the output of the previous node. Moreover, to have better control and memorize the flow of information, the LSTM cell utilizes gating mechanisms such as input gate I^t, forget gateF^tAnd the output gateY^t, where each cell gate such that $\{I^t, F^t, Y^t\} \epsilon \mathbb{R}^{k \times 1}$. The I^t utilizing x_t and h_{t-1} determines what value to use to decide the value of C_t. The operation of updatingC_t by I^tgate numerically expressed as follows:

$$I^t = \sigma(w_i \otimes [h_{t-1}, x_t] + b^I) \tag{1}$$

$$C_t^\alpha = \alpha(w_C \otimes [h_{t-1}, x_t] + b^C) \tag{2}$$

Where, in equation (1) I^t denotes input gate of the cell at timestep 't' (occurrence of LSTM cell), the variable w_i and b^I are the weights and bias of 'σ' sigmoid operation between I^t and Y^t. In equation (2) C_t^α denotes values of cell state generated by α, the variable w_C and b^C denotes weight and bias of α operation between C_t and h_t.The next F^t gate decides what information from C_{t-1} to be considered to update C_t. The operation of information flow by $F^t$the gate is numerically expressed as follows:

$$F^t = \sigma(w_F \otimes [h_{t-1}, x_t] + b^F) \tag{3}$$

Where, in equation (3) F^t denotes forget gate of the cell at 't', the variable w_F and b^F are the weights and bias of 'σ' sigmoid operation between F^t and I^t. Using equations (1), (2), and (3), the operation of C_t can be numerically expressed as follows:

$$C^t = F^t \otimes (C_{t-1}) + I^t \otimes C_t^\alpha \tag{4}$$

Further, the Y^t gate determines what information in C_t become value of h_t. The operation of information flow by $Y^t$the gate is numerically expressed as follows:

$$Y^t = \sigma(w_Y \otimes [h_{t-1}, x_t] + b^Y) \tag{5}$$

$$h_t = Y^t \otimes \alpha(C_t) \tag{6}$$

Where, in equation (5) Y^t denotes output gate of the cell at 't', the variable w_Y and b^Y are the weights and bias of 'σ' Y^t. In equation (6) h_t denotes the output of LSTM network computed point-by-point multiplication of previous equation (5) and function α with the input argumentC_t. In the proposed work, the task of compressor fault prediction using time-domain AS signals is regarded as a sequence classification problem.Therefore, the proposed study explores implementinga deep learning mechanism, particularly LSTM, for large-scale time-domain AS signals modeling for fault prediction in AC.The proposed architecture learning model for AC fault classification is demonstrated in Fig. 10.
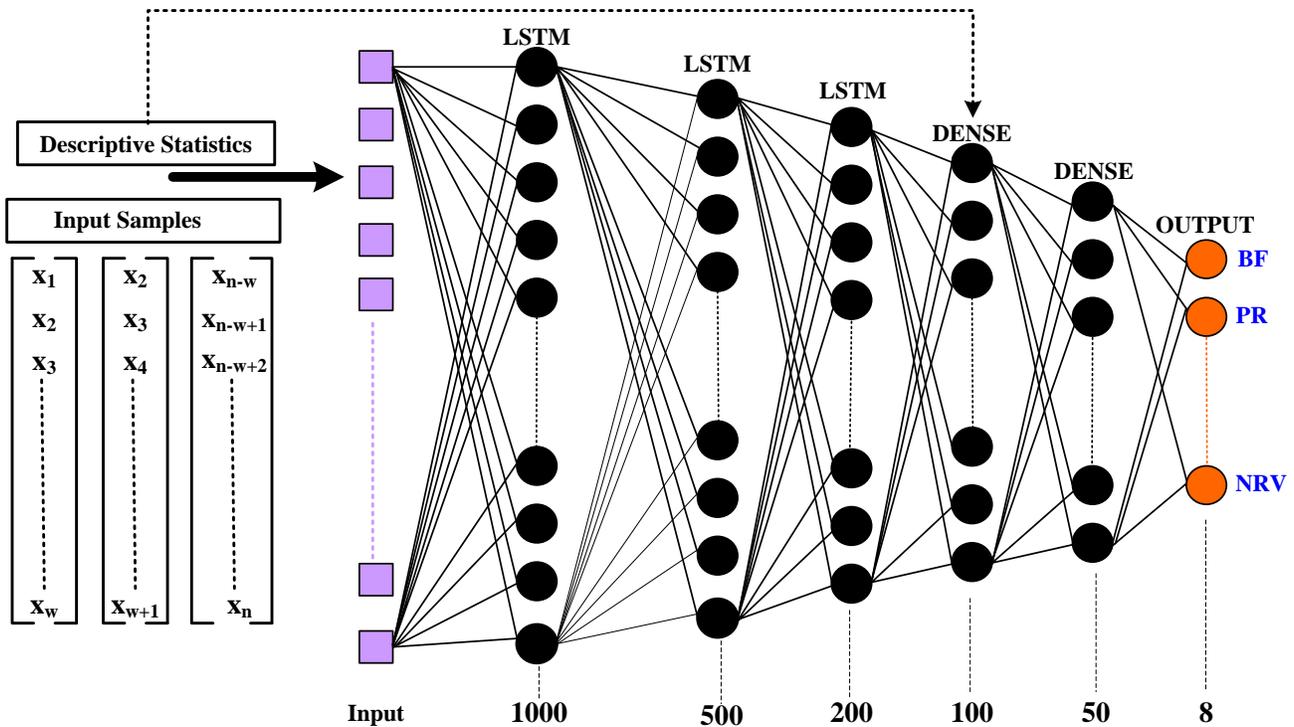

Fig. 10. LSTM Model for AC Fault Classification.

The architecture of the learning model consists of a total three-layer such as i) sequence input layer (I), ii) 5 hidden layers (H), and iii) one output layer (Y), such that learning model $M \in \{I = 1, H = 5, Y = 1\}$. Three hidden layers are configured with 1700 LSTM cells, and the remaining two hidden layers are dense with 158 (regular, deeply connected neurons). The configuration description of the LSTM learning model is presented in Table I. The learning model takes Input as time-domain sensory signals in sequence such that $X_n \in \{x_1, x_2, x_3 \cdots, x_t, \dots x_L\}$ where t denotes timestep, $'L'$ length of time-domain signal samples x and maps to output class $C \in \{C_1, C_2, C_3 \cdots C_8\}$ via hidden layer $H \in \{H_1, H_2, H_3 \cdots H_5\}$ using function $H_k = f\{H_{k-1}, x_k\}$, where function f refers to the LSTM cell method discussed above that generalizes the long-term dependence between the time domain relationship of the Input X_n signal. The LSTM is trained considering the Input X_n in the form of vectors using the sliding windowing (w) approach, where the Input is a sequence of time-domain signals with length L and w+1 window length. The process of window sliding is illustrated in Fig. 11 with window length (w=1000 AS signal samples).

In the above illustrated, the model takes Input as the first window having the first 1000 AS signal. Then, the next window is selected from the second signal sample of the first window, i.e., from the 2nd sample to the 1001st sample. This process is recurrent until all-time-domain input signals are windowed and fed to the LSTM model. The process flow of AC faults prediction using LSTM is shown in Fig. 12.

The system initially imports the dataset, consisting of 1800 AS signals captured at a 50,000 Hz sampling frequency. To execute the sequence classification tasks, splitting the dataset into two sub-datasets, i.e., training and testing sets. The training dataset is used to train the model, and the testing dataset is used to evaluate the model performance. This allows understanding characteristics of the trained model and provides scope for minimizing the effects of overfitting and underfitting of the model. The dataset split is carried out with a ratio of 80%-20% for training and testing, respectively. Therefore, the training dataset consists of 1440 AS signal samples and a testing dataset composed of 360 AS signaling samples. The study further considers feature selection, where descriptive statistics were analyzed in the time domain. In the training phase, a sliding windowing operation is carried to represent AS signals into the fixed-sized frame, which is further processed via the LSTM layer. Its output is then accumulated with the operation of the dense layer that considers descriptive statistics as Input.

The adoption of a dense layer enhances the generalization of the learning model minimizes the issue of overfitting and underfitting during the learning process. In the proposed LSTM architecture, Adam optimizer is used with a categorical cross-entropy loss function to reduce training loss by adjusting learning attributes such as weights, biases, and learning rate. The configuration details of the LSTM model implemented in this study are mentioned in Table IV. Moreover, Softmax activation is used at the output layer of the LSTM model, as it is designed to address multiclass sequence classification problems (i.e., multiple AC faults). A similar procedure is carried out during model testing, and its effectiveness is assessed regarding the accuracy, precision, recall, and F1-score.

Fig. 13 illustrates the learning curve of the LSTM model training performance over 1000 epochs. As the epochs pass by, the reduction in learning also reduces in the LSTM. Generally, learning rate reduction happens when the error is not reducing for more than 5 epochs. The learning rate is reducing rapidly the error is not converging quickly enough in LSTM. It can be observed that a sharp exponential decrease in learning rate in the LSTM model. When evaluated with the testing dataset, the model achieved an accuracy rate of 83.12% in AC fault classification. The following section discussed the proposed learning model based on a multilayer perceptron neural network.
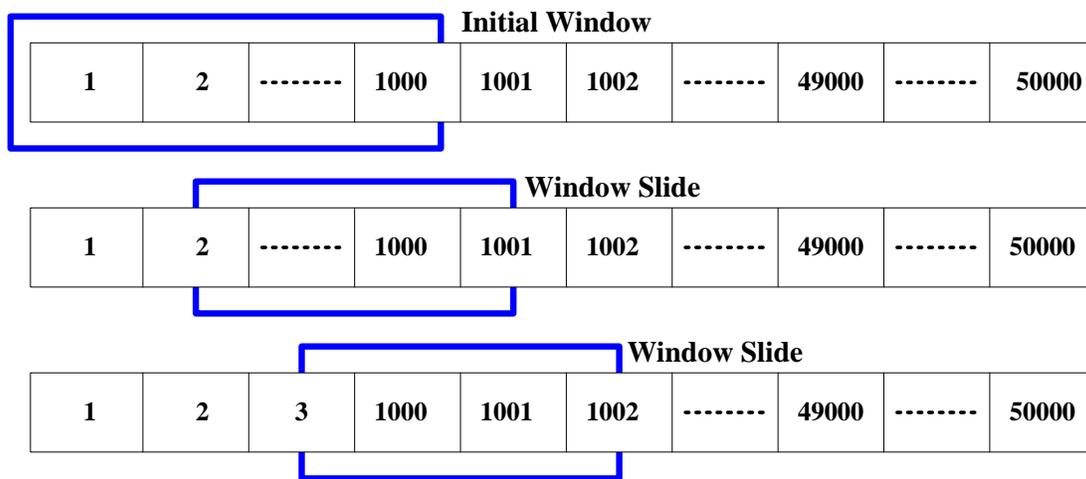

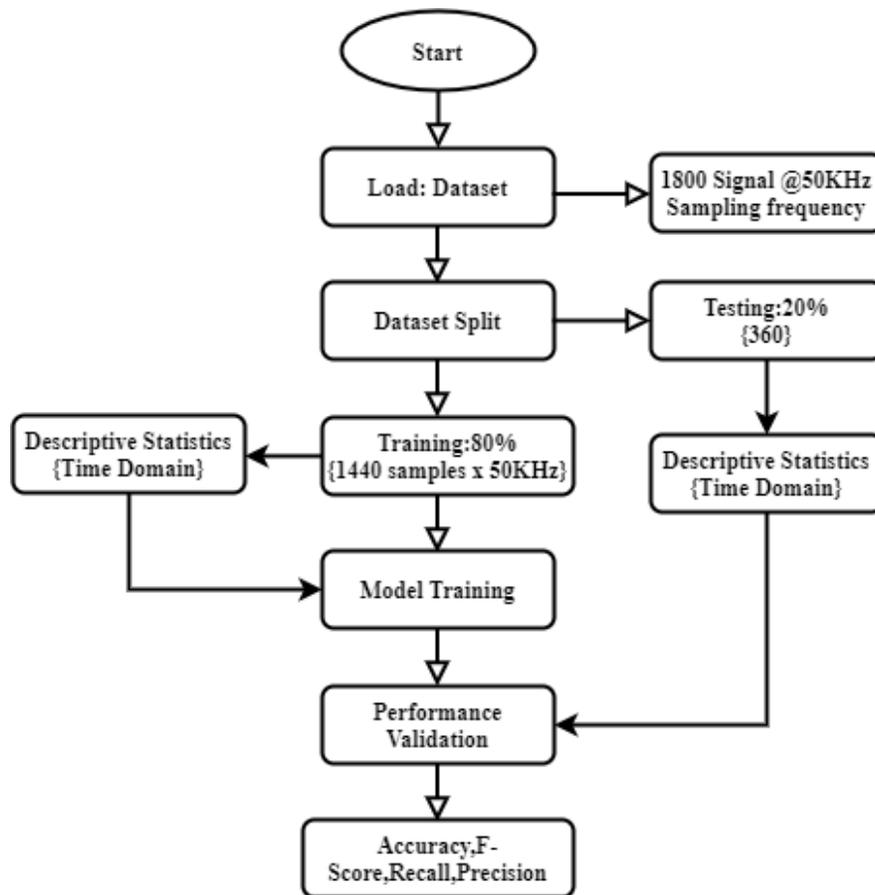Fig. 11. Process of Sliding Window.

Fig. 12. Illustration of the Process Flow of the LSTM for AC Fault Classification.

TABLE IV.   LSTM MODEL CONFIGURATION DETAILS

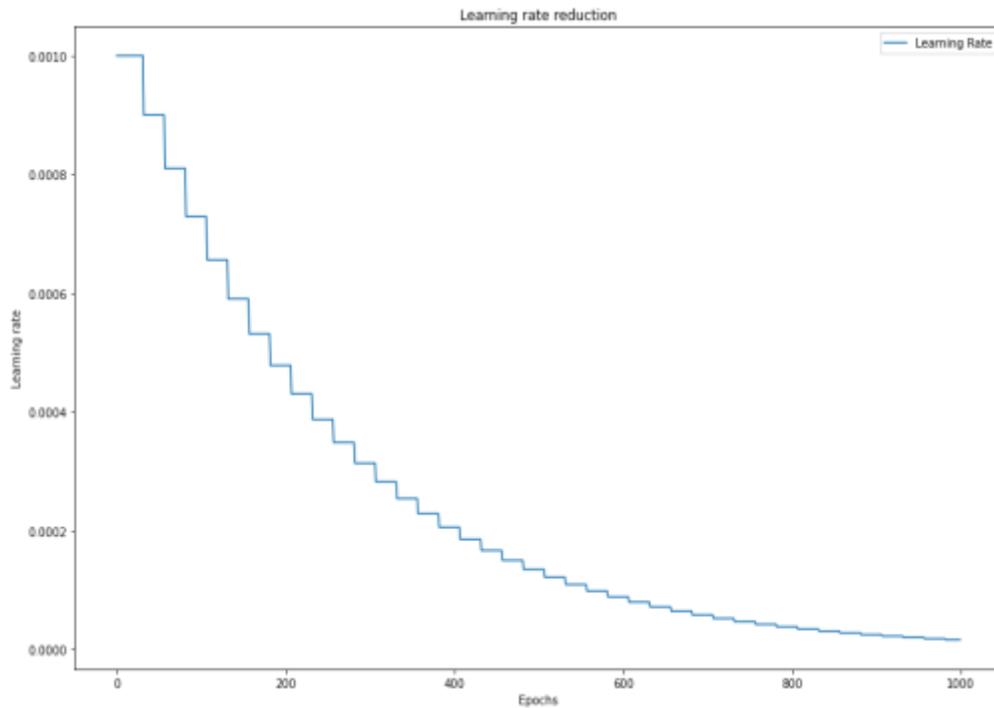|		
Input layer	1	Fixed window length of 1000 AS signal
LSTM layer	3	Specific to time-domain AS Signal
Dense Layer	2	Specific to Descriptive Statistics
Optimizer	Adam	Stochastic optimizer
Activation for output	Softmax	Multiclass classification
Loss Function	Categorical cross-entropy	Computesthe difference between two probability distributions
Min Learning rate	10^{-5}	Minimum permitted learning rate for reduce_lr
Max Learning rate	10^{-3}	Initial learning rate
Epochs	1000	The optimal number of epochs for best generalization

Fig. 13. Learning Curve of LSTM.

V. Proposed Learning Model using Multilayer Layer Perceptron

The neural network consists of an artificial neuron interconnected together by synaptic weights to form a network. Each neuron is modeled by the linear threshold unit, which maps single Input to single output using mathematical operation described as follows:

$$y = f(\sum_{i=1}^{n} W_i x_i + \theta) \tag{7}$$

where y denotes the output of the neuron, W_i indicated synaptic weight, $i \in \{1,2,3 \cdots N\}$, x_i is the Input $i \in \{1,2,3 \cdots N\}$, and θ indicates a threshold function. A non-linear $f(x)$ can be a sigmoid function or a hyperbolic tangent function.

Amultilayer perceptron (MLP)class of NN. In MLP, the signal travels only in a forward direction; numerically, it can be represented as follows:

$$y = f(A_z + \theta_y) \tag{8}$$

$$z = f(B_u + \theta_z) \tag{9}$$

where, y is an $m \times 1$ vector refers to the output of the neurons at the output layer; Z is a $p \times 1$ vector, indicates the outputs of neurons at the hidden layer; u is an $n \times 1$ vector, indicates the feature vector of the input signal; θ_y and θ_z arethe threshold vector for the neurons at the output and hidden respectively; the size of θ_y is $m \times 1$ and θ_z is $P \times 1$, A and B are the matrices of size $m \times p$ and $p \times n$, respectively. Both refer to synaptic weights connecting the hidden layer neuron to the output and the Input and hidden layer neurons. The nonlinearity function to be a sigmoid function, i.e.,

$$f(\alpha) = \frac{1}{1+e^{-\alpha}} \tag{10}$$

The unknown parameters A, B,θ_y and θ_z can be determined viareducing an error criterion such that:

$$J = \sum_{i=1}^{P}(d_i - y_i)^2 \tag{11}$$

Where d_i indicates expected outputs which are required to MLP learn and $i \in \{1,2,3 \cdots P\}$, P indicated a total number of instances.

The proposed system implements MLP to classify frequency domain AS signals to predict AC faults because MLP can address complex non-linear problems. It works with both large and small input data and offers quick prediction after training. All these factors are highly significant to the real-time scenario. Although the LSTM is suitable for time sequence data prediction, it is prone to computational overhead and sometimes overfitting problems. The architecture of the proposed AC fault classification system using MLP is shown in Fig. 14.

The MLP can perform better if the AS signal is better exposed to the MLP; as observed in the data exploration, the frequency of audio samples provides a better insight into AS signal representing the faults. Hence, in the proposed study, the frequency domain AS signal is used to train the MLP model to get better accuracy in classification fault classes. The proposed model is composed of two core modules such as i) Adaptive filter and ii) MLP module. Adaptive filter as a frequency domain bandpassfilter, also known as the digital filter,restricts some frequencies from being givenInput to MLP. The functional process of an implemented digital filter with MLP is shown in Fig. 15.
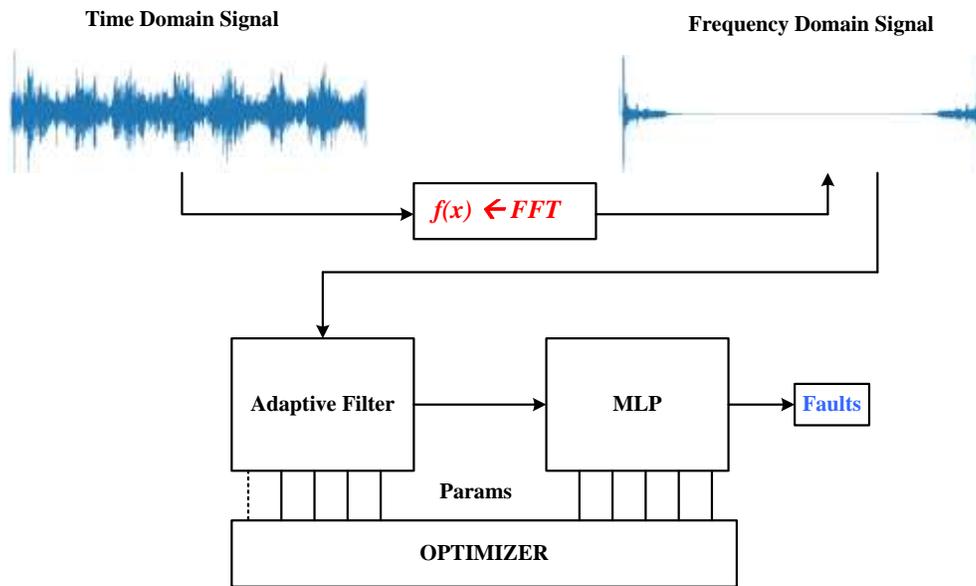
Time Domain Signal **Frequency Domain Signal**

$f(x) \Leftarrow FFT$

Adaptive Filter **MLP** **Faults**

Params

OPTIMIZER

Fig. 14. Schematic Architecture of Proposed Learning Model.

x(k)

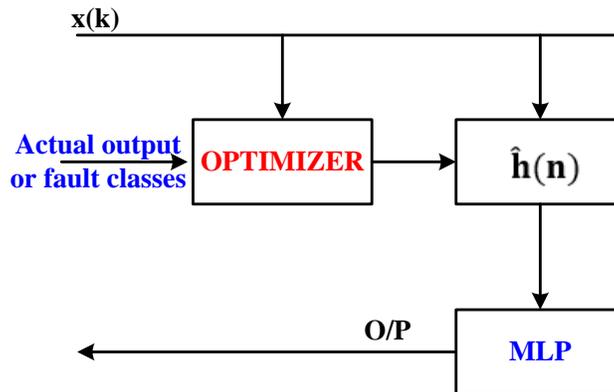**Actual output or fault classes** **OPTIMIZER** $\hat{h}(n)$

O/P **MLP**

Fig. 15. Synchronization of Frequency Domain Bandpass Digital Filter with MLP.

In the real-time scenario, AS-generated signals often may associate with noisy environmental factors, consisting of recursive or redundant frequencies. Since the proposed learning model takes Input in the frequency domain, it is essential to ensure that the Input AS frequency domain signal does not associate with any irrelevant factors to achieve higher accuracy in the output (O/P). The implementation of adaptive digital filter $\hat{h}(n)$ restricts the irrelevant and redundant frequencies before it is being introduced to the MLP. As a result, reduction in the number of input frequency domain signal samples x(k) reduces computational complexity, thereby reducing feature space complexity by removing irrelevant frequency domain AS signaling features. The processed information by $\hat{h}(n)$ representing a precise input, which providesbetter generalization ability to the MLP in the training phase. The architecture of MLP for AC fault prediction is shown in Fig. 16.

The MLP architecture proposed in the current study consists of the single input layer, with input frequency domain

sensory signals such that $x_n \in \{x_1, x_2, x_3 \cdots, x_N\}$ each at 25000 sampling frequency (Nyquist frequency) and mapped to output class $C \in \{C_1, C_2, C_3 \cdots C_8\}$ at output layer via a hidden layer of type dense$D \in \{D_1, D_2, D_3\}$. Since the time domain AS signal is transformed into a frequency domain signal, the theoretical maximum frequency using FFT can be detected always half of the sampling frequency. In the current study, since AS signal sampling frequency is 50 KHz, after transforming to the frequency domain using FFT, Nyquist frequency is 25 KHz. In the proposed MLP architecture, a linear activation function is used at each hidden layer. In the output layer, SoftMax activation is used to deal with the prediction of multiclass AC faults. Therefore, the output layer contains only 8 neurons signifying 8 different outputs. The SoftMax function ensures the sum of all outputs is always 1; hence only the maximum output is selected as the final output with the help of argmax function. A common optimizer is used for both ANN as well as the filter. The optimizer sets the h(n),known as the filter's impulse response. Fig. 17 exhibits the Nyquist frequency sampling process.
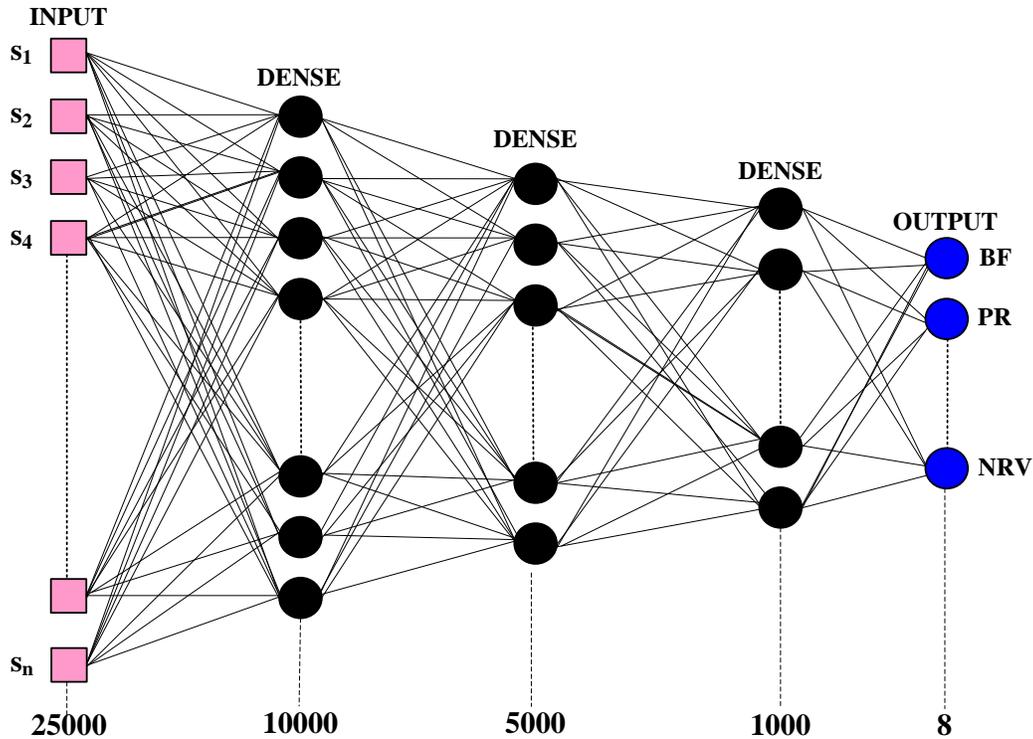
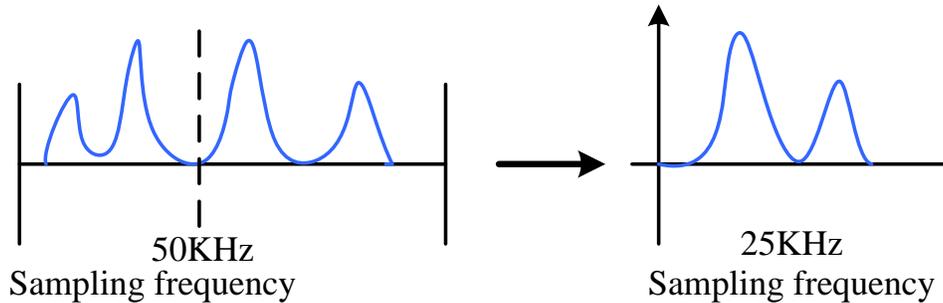Fig. 16. Architecture of Implemented MLP Model.

Fig. 17. Illustration of Nyquist Frequency Sampling.

The process flow of AC faults prediction using MLP is shown in Fig. 18, where the system initially imports the dataset,consistingof 1800 AS signals captured at a 50,000 Hz sampling frequency. To execute the sequence classification task, splitting the dataset into two sub-datasets, i.e., training and testing sets. The training dataset is used to train the model, and the testing dataset is used to evaluate the model performance. The input AC signal is converted to the frequency domain using FFT. In this present study, since the sampling frequency is 50 KHz, using the Nyquistmechanism,the sampling frequency of AS signalis computed at 25 KHz, which is the theoretical maximum frequency of FFT. Next, descriptive statistics of frequency domain AS signalsare computed and processed with a domain bandpass adaptive filter. As a result, redundant frequencies from the given Input AS signals being restricted. The filtered AS signal is further fed to MLP, where training is carried out using linear activation functions at each dense layer. After training the model, the testing dataset is used to evaluate the model. Fig. 19 illustrates the learning curve of the MLP model training performance over 1000 epochs.

In Fig. 19, the learning curve trend exhibits a reduction in learning rate is slower compared to LSTM. This indicates that the error reduces rapidly, and effective generalization of MLP. The next section discusses the performance metrics considered for the proposed learning model performance analysis.
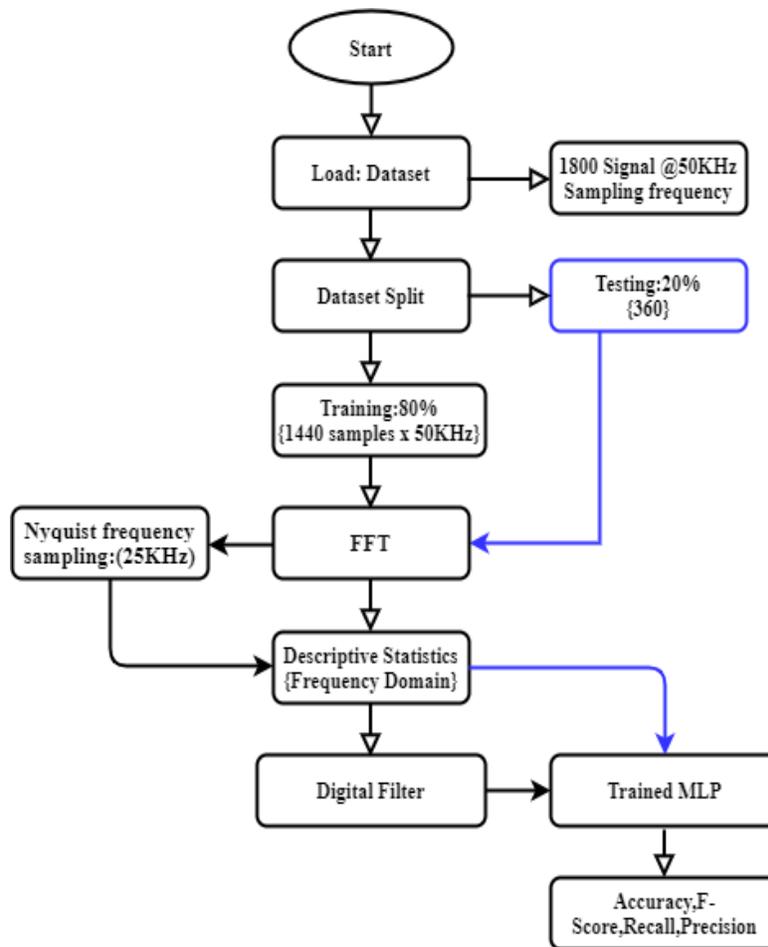
Fig. 18. Illustration of the Process Flow of the MLP for AC Fault Classification.
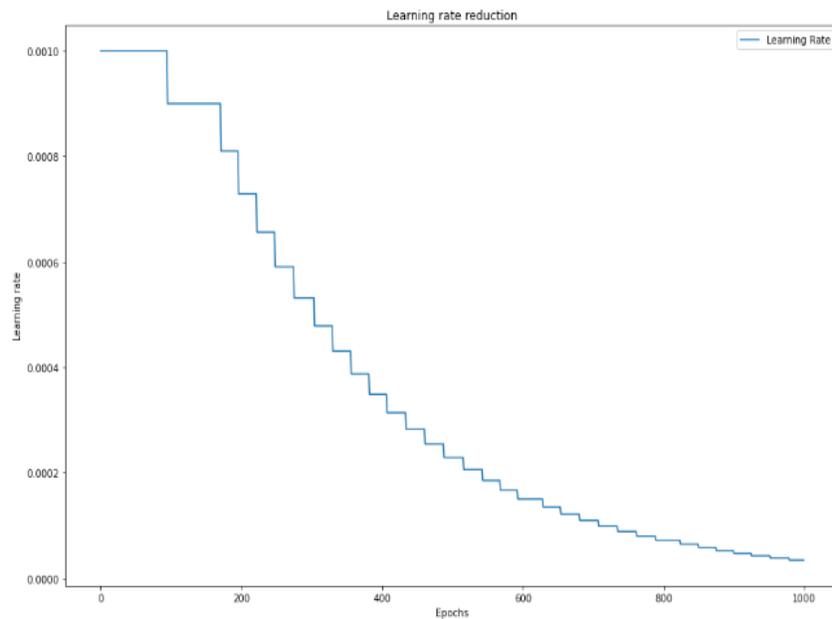

Fig. 19. Learning Curve of MLP.

VI. PERFORMANCE METRICS

Both learning models' outcome and performance evaluation is carried out concerning multiple performance parameters such as accuracy, precision, recall, and F-1 score.

Accuracy (A): Accuracy can be defined as the ratio of correct predictions over a total number of predictions. Therefore, in the current context of the case study, accuracy can be described as follows:

$$A = \frac{\text{Number of correct predictions of AC faults}}{\text{Total number of AC fault classes}}$$

Precision (P): Precision is the ratio of the number of correct predictions over a total number of predictions made to the current fault class.

$$P = \frac{\text{Total number of accurately identified AC fault classes}}{\text{Total number of detected AC faults classes}}$$

Recall (R): Recall is the ratio of correctly predicted values over the number of expected faults, i.e., the total fault classes present in the test dataset. The lower recall represents the inability of the system to detect the particular class. Like precision, even for recall,the weighted average is taken.

$$R = \frac{\text{Total number of accurately identified AC fault classes}}{\text{Total number of expected AC faults classes}}$$

F1 score: This performance metricdescribes the harmonic mean of precision and recall, which truly represents the system's performance.

$$\text{F1\_Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right)$$

VII. RESULTS AND PERFORMANCE ANALYSIS

This section presents the outcome obtained from both implemented learning model and performance analysis and discussesAir Compressor faults classification using acoustic sensor signals. The entire modelling and development of the proposed system are carried out using Python.

A. Analysis of Learning Rate

The comparative analysis concerning the learning rate reduction to access training performance of both LSTM and the proposed learning model.

Fig. 20 presents a comparison of implemented LSTM and Proposed MLP regarding learning rate. It can be analyzed from the learning curve trend that at the beginning, the proposed MLP method takes a little longer time to reduce the learning rate compared to LSTM. However, the proposed MLP model maintains a significant reduction in error during its initial stage of the training process. This indicates that the proposed method has a better optimization in learning and generalization compared to LSTM. It is to be noted that the more the area between the curves better the improvement will be, and if the error in the training phase is not reduced for continuous five epochs, then the learning rate will reduce. The proposed model's learning rate is more, which signifies that the proposed MLP learns faster than the LSTM.
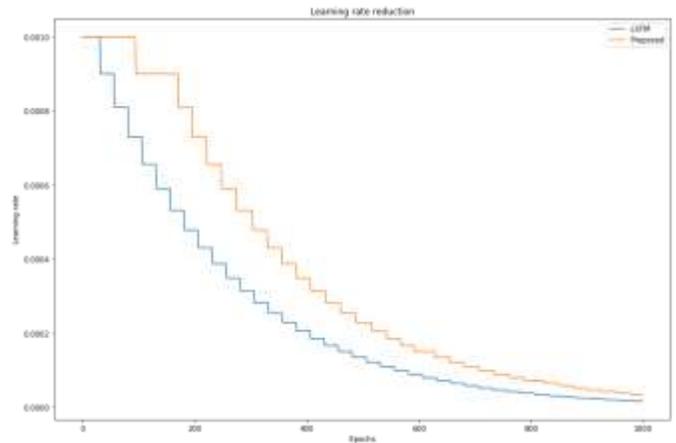

Fig. 20. Comparative Analysis Regarding Learning Rate.

B. Analysis of Classification Performance

The performance analysis is carried out considering multiple evaluation metrics. This is because the system's accuracy is always not a good metric to measure performances, especially when using learning models. The accuracy may not represent the performance of the system completely. If the modelcorrectly predicts fault classes, it indicates a higher accuracy even if it cannot predict negative values. Therefore, performance evaluation of the implemented learning model, i.e., LSTM and the proposed MLP, is carried out considering other two evaluation metrics such as precision and recall rate. However, there is always a trade-off between precision and recall. This is because the precision focuses more on the exactness of the learning models.On the other hand, recall rate focuses more on measuring the completeness of the learning model. For example, suppose the model recognizes the air compressor as faulty. When there are no faults in the air compressor, it is said to have a lower precision rate, which indicates many false positives or biases in the air compression fault predictions. If the model has a higher precision score, then the model is subjected to a low false-positive rate. A low recall rate indicates higher false negatives, and a higher recall rate indicates low false negatives in the prediction result. Since the precision represents the correctness of positive results and recall represents the correctness of negative results, the model should be built to balance both. In order to measure the balance between precision and recall, the F1_score metric is evaluated, which shows the harmonic mean of precision and recall. The harmonic mean is used instead of the regular average since the harmonic mean reduces the effect of extreme values. Table V presents the numerical outcome obtained, followed by the graphical outcome in Fig. 21 to evaluate the implemented learning models.

TABLE V.   NUMERICAL OUTCOME

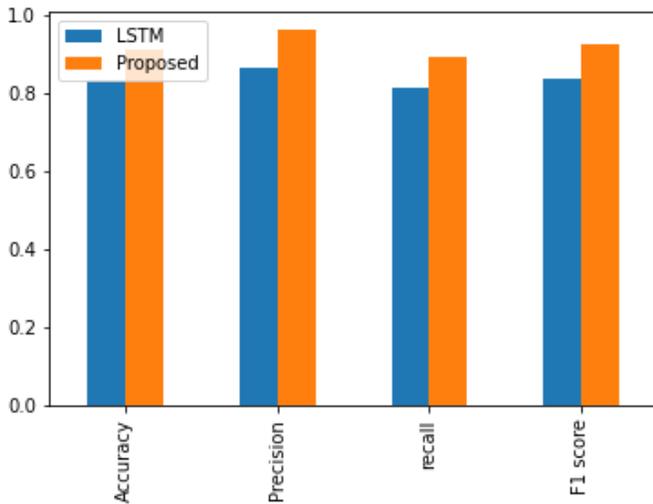| Performance Metrics	LSTM	Proposed
Accuracy	0.83123	0.913234
Precision	0.863425	0.962342
recall	0.815464	0.892342
F1 score	0.838759	0.926021

Fig. 21. Comparative Analysis of Classification Outcome.

In Fig. 21, the comparative analysis exhibits that the proposed learning model outperforms LSTM in all performance metrics. The LSTM achieved an accuracy rate of 83.12%, whereas the proposed model achieved a 91.32% accuracy rate. In the case of precision metric, LSTM has scored 86.34%, and the proposed model has attained 96.23 % of the precision rate. The proposed model achieved an 89.23% recall rate, whereas LSTM achieved an 81.54% recall rate. Also, the proposed model has a higher F1_score than LSTM, i.e., 92.60% and 83.87, respectively. Based on the observation, it can be analyzed that the LSTM is biased to an extent towards the faulty results. This indeed came as no surprise since the data contains very few healthy signal samples than faulty signals. The proposed method also has a difference of approximately 7% between precision and recall. However, these are within acceptable limits, which indicates that the proposed model is better at detecting air compressor faults. However, even then, the system is more reliable when both precision and recall are balanced.

VIII. CONCLUSION

In the proposed work, the study aimed to predict different types of air compressor faults. The analysis was carried out using sensory signals captured from the Acoustic sensors mounted on the Air compressor. The proposed study carried out data visualization and exploratory analysis to characterize the signal features and faults in time and frequency domains. The proposed study is concerned with two aspects of the classification process: the first classification of air compressor faults using the LSTM learning model where the time-domain signal is used as input. On the other hand, the frequency-domain signal is used with a digital filter in the proposed MLP learning model. The result indicated that the proposed learning model outperforms LSTM in accuracy, precision, recall rate, and F1_Score. The outcome shows 83.12% and 91.32% of accuracy achieved by LSTM and MLP, respectively. Also, the learning performance of both models is evaluated. The analysis exhibited that the proposed MLP has less training time compared to LSTM. Therefore, the proposed learning can be claimed to be efficient and suitable for real-time implementation. It has less training time, does not suffer from feature extraction problems, has less memory overhead, and has good generalization ability due to preciseness in the input signal leads to achieving higher accuracy.

REFERENCES

[1] Sastry YS, Kiros BG, Hailu F, Budarapu PR. Impact analysis of compressor rotor blades of an aircraft engine. Frontiers of Structural and Civil Engineering. 2019 Jun;13(3):505-14.

[2] Zhang L, Cui J, Zhang Y, Yang T, Li J, Gao W. Performance analysis of a compressed air energy storage system integrated into a coal-fired power plant. Energy Conversion and Management. 2020 Dec 1;225:113446.

[3] Razmi AR, Afshar HH, Pourahmadiyan A, Torabi M. Investigation of a combined heat and power (CHP) system based on biomass and compressed air energy storage (CAES). Sustainable Energy Technologies and Assessments. 2021 Aug 1;46:101253.

[4] Fang Y, Lu Y, Roskilly AP, Yu X. A review of compressed air energy systems in vehicle transport. Energy Strategy Reviews. 2021 Jan 1;33:100583.

[5] Li J, Yu T. A new adaptive controller based on distributed deep reinforcement learning for PEMFC air supply system. Energy Reports. 2021 Nov 1;7:1267-79.

[6] Sayed E, Abdalmagid M, Pietrini G, Sa'adeh NM, Callegaro AD, Goldstein C, Emadi A. Review of Electric Machines in More/Hybrid/Turbo Electric Aircraft. IEEE Transactions on Transportation Electrification. 2021 Jun 15.

[7] Cui C, Lin W, Yang Y, Kuang X, Xiao Y. A novel fault measure and early warning system for air compressor. Measurement. 2019 Mar 1;135:593-605.

[8] Aikin AR. The process of effective predictive maintenance. Tribology & Lubrication Technology. 2021 Feb 1;77(2):34-40.

[9] Ye Y, Grossmann IE, Pinto JM, Ramaswamy S. Modeling for reliability optimization of system design and maintenance based on Markov chain theory. Computers & Chemical Engineering. 2019 May 8;124:381-404.

[10] Qu Y, Hou Z. Degradation principle of machines influenced by maintenance. Journal of Intelligent Manufacturing. 2021 Feb 11:1-0.

[11] Lv Q, Yu X, Ma H, Ye J, Wu W, Wang X. Applications of Machine Learning to Reciprocating Compressor Fault Diagnosis: A Review. Processes. 2021 Jun;9(6):909.

[12] Alshorman O, Alshorman A. A review of intelligent methods for condition monitoring and fault diagnosis of stator and rotor faults of induction machines. International Journal of Electrical & Computer Engineering (2088-8708). 2021 Aug 1;11(4).

[13] Połok B, Bilski P. Intelligent Diagnostic system for the rachet mechanism faults detection using acoustic analysis. Measurement. 2021 Jun 24:109637.

[14] Doshi, S., Katoch, A., Suresh, A. et al. A Review on Vibrations in Various Turbomachines such as Fans, Compressors, Turbines and Pumps. J. Vib. Eng. Technol. (2021). https://doi.org/10.1007/s42417-021-00313-x.

[15] Yao J, Liu C, Song K, Zhang X, Jiang D. Fault detection of complex planetary gearbox using acoustic signals. Measurement. 2021 Jun 1;178:109428.

[16] Połok B, Bilski P. Intelligent Diagnostic system for the rachet mechanism faults detection using acoustic analysis. Measurement. 2021 Jun 24:109637.

[17] Ahmed U, Ali F, Jennions I. A review of aircraft auxiliary power unit faults, diagnostics and acoustic measurements. Progress in Aerospace Sciences. 2021 Jul 1;124:100721.

[18] Lou F, Key NL. Reconstructing Compressor Non-Uniform Circumferential Flow Field From Spatially Undersampled Data—Part 1: Methodology and Sensitivity Analysis. Journal of Turbomachinery. 2021 Aug 1;143(8):081002.

[19] Verma NK, Salour A. Faults and Data Acquisition. InIntelligent Condition Based Monitoring 2020 (pp. 7-88). Springer, Singapore.

[20] Verma N.K., Salour A. (2020) Pre-processing. In: Intelligent Condition Based Monitoring. Studies in Systems, Decision and Control, vol 256. Springer, Singapore. https://doi.org/10.1007/978-981-15-0512-6_3.

[21] Wang F, Lin W, Liu Z, Wu S, Qiu X. Pipeline leak detection by using time-domain statistical features. IEEE Sensors Journal. 2017 Aug 15;17(19):6431-42.

[22] Yi Z, Pan N, Guo Y. Mechanical compound faults extraction based on improved frequency domain blind deconvolution algorithm. Mechanical Systems and Signal Processing. 2018 Dec 1;113:180-8.

[23] Sun Z, Zou W, Zheng X. Instability detection of centrifugal compressors by means of acoustic measurements. Aerospace Science and Technology. 2018 Nov 1;82:628-35.

[24] Mondal D, Zhen D, Gu F, Ball AD. Fault diagnosis of reciprocating compressor using empirical mode decomposition-based Teager energy spectrum of airborne acoustic signal. InAdvances in Asset Management and Condition Monitoring 2020 (pp. 939-952). Springer, Cham.

[25] Hammond TT, Curtis MJ, Jacobs LE, Tobler MW, Swaisgood RR, Shier DM. Behavior and detection method influence detection probability of a translocated, endangered amphibian. Animal Conservation. 2020 Sep.

[26] McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction.

[27] Hartwig L, Bestle D. Compressor blade design for stationary gas turbines using dimension reduced surrogate modeling. In2017 IEEE Congress on Evolutionary Computation (CEC) 2017 Jun 5 (pp. 1595-1602). IEEE.

[28] Liang X, Duan F, Mba D, Ian B. Centrifugal Compressor Diagnosis Using Kernel PCA and Fuzzy Clustering. InAsset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies 2019 (pp. 373-381). Springer, Cham.

[29] Li X, Duan F, Loukopoulos P, Bennett I, Mba D. Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. Control Engineering Practice. 2018 Mar 1;72:177-91.

[30] Gong CS, Su CH, Tseng KH. Implementation of machine learning for fault classification on vehicle power transmission system. IEEE Sensors Journal. 2020 Jul 20;20(24):15163-76.

[31] Sobie C, Freitas C, Nicolai M. Simulation-driven machine learning: Bearing fault classification. Mechanical Systems and Signal Processing. 2018 Jan 15;99:403-19.

[32] Wang Y, Wang J, Sun J, Liang E, Wang T. Investigation on recognition method of acoustic emission signal of the compressor valve based on CNN and LSTM network. InE3S Web of Conferences 2021 (Vol. 252, p. 02023). EDP Sciences.

[33] N. K. Verma, R. K. Sevakula, S. Dixit and A. Salour, "Intelligent Condition Based Monitoring Using Acoustic Signals for Air Compressors," in IEEE Transactions on Reliability, vol. 65, no. 1, pp. 291-309, March 2016, doi: 10.1109/TR.2015.2459684.

[34] https://www.iitk.ac.in/iil/datasets/ [Accessed on 22-06-2021].

[35] Schwabacher, M., Goebel, K.: A survey of artificial intelligence for prognostics. In: AAAI Fall Symposium, pp. 107–114.

[36] Omer Faruk Eker, Faith Camci, and Ian K Jennions. 2014. A similarity-based ¨ prognostics approach for remaining useful life prediction. (2014).

[37] RachaKhelif, Simon Malinowski, BrigieChebel-Morello, and Noureddine Zerhouni. RUL prediction based on a new similarity-instance based approach. In IEEE International Symposium on Industrial Electronics' 14.

[38] Zhong, G., Yang, G.: Fault detection for discrete-time switched systems in finitefrequency domain. Circuits Syst. Signal Process. 34(4), 1305–1324.

[39] Rigamonti, M., Baraldi, P., Zio, E.: Echo state network for the remaining useful life prediction of a turbofan engine. In: Third European Conference of the Prognostics and Health Management Society.

[40] Ouadine AY, Mjahed M, Ayad H, El Kari A. Aircraft Air Compressor Bearing Diagnosis Using Discriminant Analysis and Cooperative Genetic Algorithm and Neural Network Approaches. Applied Sciences. 2018 Nov;8(11):2243.

[41] Li, X.; Palazzolo, A.; Wang, Z. Rotating machinery monitoring and fault diagnosis with neural network enhanced fuzzy logic expert system. In Proceedings of the ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition, Seoul, South Korea, 13–17 June 2016.

[42] Ghorbanian K, Gholamrezaei M. An artificial neural network approach to compressor performance prediction. Applied Energy. 2009 Jul 1;86(7-8):1210-21.

[43] Aravinth S, Sugumaran V. Air compressor fault diagnosis through statistical feature extraction and random forest classifier. Progress in Industrial Ecology, an International Journal. 2018;12(1-2):192-205.

[44] Fan Y, Nowaczyk S, Antonelo EA. Predicting air compressor failures with echo state networks. InPHM Society European Conference 2016 (Vol. 3, No. 1).

[45] Cui C, Lin W, Yang Y, Kuang X, Xiao Y. A novel fault measure and early warning system for air compressor. Measurement. 2019 Mar 1;135:593-605.

[46] Musyafa A, Kusumawardhani S, Noriyati RD, Justiono H. Evaluation of the Reliability and Prediction Maintenance on the Air Compressor System in Ammonia Plant PT. Petrokimia Gresik. Australian Journal of Basic and Applied Sciences. 2015 May;9(11):853-62.

[47] Chen K, Pashami S, Fan Y, Nowaczyk S. Predicting air compressor failures using long short term memory networks. InEPIA Conference on Artificial Intelligence 2019 Sep 3 (pp. 596-609). Springer, Cham.

[48] Yang WS, Su YX, Chen YP. Air compressor fault diagnosis based on lifting wavelet transform and probabilistic neural network. InIOP Conference Series: Materials Science and Engineering 2019 Oct 1 (Vol. 657, No. 1, p. 012053). IOP Publishing.