Diagnosis of different degree of blockage levels in the Centrifugal Pump Employing Vibration, Pressure and Current Data through Artificial Neural Network

Shivam Gautam, Rajiv Tiwari [1], D.J. Bordoloi

Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

Abstract

The appearance of flow instabilities like the blockage severity, impeller cut flaws, pitted cover plate flaws can cause to diminish the efficiency of centrifugal pump (CP), and may result in excessive vibration and noise, and their failure may lead to the system imploding. To bridge the gap of downfall in the efficiency of CP, it is crucial that a system can be created to monitor the condition of the CP and must be maintained. The present work proposes at identifying and determining the severity of various blockage levels in the inlet pipe with three different kinds of pumps using three distinct sensors. One pump works faultlessly (healthy pump), another has cuts artificially made on the impeller blade, and the third has pits artificially created on the cover plate. The inlet pipe blockage mimics pump blockage which is made more severe step by step. As the blockage gets worse and the flow slows down, recirculation starts, causing vapor bubbles to form. Utilizing a mechanical modulating valve, the inlet flow area of the pipe is partitioned into six intervals (0%, 16.7%, 33.3%, 50%, 66.6%, and 83.33%) to replicate pump blockage. This obstruction directly influences vibrations, current line signals, and fluid dynamic pressure. To gather data across a spectrum of blockage levels and operational frequencies (30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz, and 60 Hz), a combination of a pressure transducer, accelerometer, and current probes were strategically employed in this investigation. Multiple sets of statistical features were extracted from the data, and through various algorithms, the most effective combined statistical feature set was determined. In this domain, the combination of standard deviation, mean, and entropy demonstrates superior performance compared to other features. This feature set was input into an ANN model, which is developed by optimizing parameters like hidden layer count, neurons, epochs and then the results of this investigation are then compared with existing literature. It has been noted that employing combinations of multiple sets of statistical features significantly improves the accuracy in identifying obstruction levels, often achieving near-perfect accuracy for various feature sets (nearly 100% across various combinations). In comparison to other SOTA methods, this approach achieves higher accuracy, ranging from 2.41% to 15.69% across different metrics. This study presents a method to classify inlet pipe blockages into various levels, enhancing maintenance prioritization and reducing downtime and repair costs, ensuring long-term equipment health and operational efficiency. The fault prediction methodology proves highly robust across various CP operating conditions.

Keywords: Centrifugal Pumps; Blockage; Sensors; Machine Learning; Artificial Neural Network.

1. Introduction

A pump is essential component in the industrial and manufacturing purposes due to their functionality as well as their lifespan. CPs typically operate based on forced vortex motion within the impeller; as

[1] Corresponding author.
E-mail address: rtiwari@iitg.ac.in (R. Tiwari).

the water exits radially, it transitions to free vortex motion. The liquid enters in the suction pipe and then moves to the central part of the pump, called as ‘eye’ and passes through a number of blades, called ‘impeller’. The impeller accelerates it in the radially outward direction because of the centrifugal force imparted by the rotating blades, which leads liquid into a casing, from where it exits into the downstream piping system. The normal operating speed range of CP is between 1000 and 3000 rpm, and they operate in harsh circumstances. CPs assure the smooth operation of the process and are essential components of many industrial plants. Multiple types of contaminants in the form of hard particles may be found in the liquid being pump in industry sectors. Impurities may obstruct pipe, which causes instability in flow in the pump. Pump malfunctions could halt the plant's process flow or reduce its efficiency, failing to produce the desired results. Furthermore, if the defects are not addressed at the appropriate stage, the life of the pumps is drastically reduced [1].

CP flaw is most typically caused by flow disruptions or mechanical faults (ruptured impellers, bearing problems, and bent rotors), as well as leaks and obstructions. On using polluted working fluids or inducing damage to the surface to the pipe, suction obstruction problems can happen. Such flow obstacles lead the flow rate to decline and recirculation of flow, a secondary flow, to emerge. Vortices arise as flow separation expands, which leads to a decrease in local pressure and the emergence of vapour bubbles. Additionally, any bubble generation in CP is not preferred because it minimizes the head, which is being developed; and ends in holes at CP surfaces owing for the generation of micro-jets. It is therefore impractical to treat CP failures as independent faults. The presence of one flaw can amplify the appearance of another. To maintain the assets in the industries, maintenance strategies are available [2].

Perovic et al. [3] created the fault signatures through linking spectral features to specific faults and analysing their behaviour in the presence of faults by using fuzzy logic, and the cavitation, obstruction, and impeller damage faults were examined. Chudina [4] used the generated noises to identified the beginning of cavitation. The initiation and progression of cavitation can be specifically monitored using a noise spectrum structure. A specific frequency tone, at 147 Hz, has been found through experiments, and it has been found to be highly reliant on the cavitation process and its progression and also, to estimate net positive suction head (NPSH). Dister [5] used the vibration and line current data for the predictive maintenance, and also techniques were suggested for online health monitoring of pumps using currents, vibrations, acoustics, and pressure variations. Zouari [6] facilitated the learning of data in an industrial setting and designed a system made up of multiple networks in order to simplify the learning data in a commercial setting. Samanta et al. [7] presented his work by conducting comparisons between neural networks (NNs) and support vector machines (SVMs) based on bearing fault data, and genetic algorithm was employed to improve performance. Also, it was revealed that the selection of features significantly influences the performance of classifier.

Singh M., Shaik A.G. [8] presented a method for detecting faulty bearings in a three-phase induction motor and Stockwell transform is used to extract features from motor vibration signals. These features are then utilized by a support vector machine (SVM) for fault classification and localization. Barakat M., Badaoui M.El. and Gulliet F. [9] proposed a hard competitive growing neural network (HCGNN) for diagnosing small bearing faults. The HCGNN model is trained using vibration signals from bearings to learn patterns and classify faults. This approach offers an effective and accurate method for diagnosing small bearing faults. Zhao R., Yan R. et al. [10] utilized in machine health monitoring. These methods leverage the power of neural networks to automatically learn and extract valuable insights from large amounts of sensor data. Deep learning enables effective fault detection, classification, and predictive maintenance, enhancing machine reliability and reducing downtime. Kane P.V. and Andhare A.B. [11] applied to gear fault diagnosis using an artificial neural network (ANN). By analyzing the sound characteristics produced by faulty gears, psychoacoustic

principles are employed to extract relevant features. These features are then input to an ANN for accurate gear fault diagnosis, enabling effective maintenance strategies.

Wong [12] presented work on detection that is applied to vibration signal monitoring and mono-block CP with defective bearings, seals, impellers, and cavitation. It was used as a test to predict the performance of pump and found that measuring high frequency noise was a convenient way of detecting cavitation. Widodo and Yang [13] presented a survey of publications from 1996 to 2006 in depth, which was based on a literature review on the use of SVM in machine condition monitoring and diagnosis. It may be deduced that he SVM in machine condition monitoring and diagnosis is trending towards expertise orientation and problem-oriented domain since year 2006. Rajakarunakaran et al. [14] proposed model took into account a total of 7 categories of defects, including 20 failures from the CP system. Using testing technique, the constructed neural network model was simulated for pattern classification of defect data, and almost 100% accuracy was attained.

Sakthivel et al. [15] differentiated the cavitation noise using noise detection techniques and employing vibration signal to diagnose mono block CP defects and monitoring the health of the pump including taking into consideration of the bearing, seal, and impeller flaws. Nasiri et al. [16] analyzed and presented vibration signature to automate the cavitation defect in CPs. Three healthy/faulty conditions—normal, mild, and fully developed cavitation were identified using the NN. If there is only one sensor, it should be used in the radial position; if there are two, they should be used in the radial and rear positions; and if there are three, the sensors should be used in the radial, back, and front positions; this will ensure that cavitation is detected accurately.

Azadeh et al. [17] analyzed with the SVM and ANN models, and developed an algorithm to classify two distinct faults in CPs to handle noisy data. Vibration signature analysis was used in a system that was described to automate the cavitation defect identification in CP. Three healthy/faulty conditions- normal, mild, and fully developed cavitation were identified using an NN. Abdulkarem et al. [18] used time and frequency domain vibration analysis and investigated the impeller fault of a CP. In this domain, the power spectrum and the vibration index at a given frequency serve as defect indicators, and provided the NN and fuzzy-neural network (FNN) for the defect identification of pumps. Rapur and Tiwari [19] considered impeller cracking and suction obstruction categorization of faults in pumps by analyzing the vibration signal. Blockage levels in middle range (40% to 70 %) were shown to have consistently excellent classification accuracy. To comprehend the onset of severe blockage, it is advised that a threshold restriction of 65% be placed on classification accuracy.

Lu et al. [20] analyzed the entry and exit pressure parameters used to develop the cavitation in the pump through experimenting and simulating. The size of the dominant frequency rose initially with the incidence and development of cavitation before eventually declining. The pump head reduced by 0.77% from non-cavitation circumstances, as shown by the experimental and numerical simulation data, and this might be interpreted as a sign of impending cavitation. Zhao et al. [21] used the datasets from rolling element bearings and planetary gearboxes, and introduced disturbances into the input data of each autoencoder in order to ensure reliable feature reconstruction and efficient high-level feature extraction. The practicality and effectiveness of employing the deep learning method in conjunction with SoftMax regression were showcased when applied to feature extraction and fault diagnosis using non-stationary and non-linear vibration data. Wang et al. [22] introduced and diagnosed a fault utilizing complementary ensemble empirical mode decomposition (CEEMD), sample entropy (SampEn), and random forest (RF), alongside an applied practical fault detection for the CP. It is observed that the CEEMD outperforms the EMD in dealing with CP vibration signals by adding white noise to the original signals. With the CEEMD, SampEn, and RF together, these provided outstanding fault diagnosis performance for CPs.

Azizi et al. [23] performed number of tests on a model pump, and the pump case vibrations were recorded. The EMD algorithm was used to deconstruct vibration signals. A total of 15 statistical features were taken. A hybrid feature selection technique was proposed to choose the most efficient

feature subset in order to improve the accuracy of the computerized diagnosis system. As a result, the overall size of the feature set was lowered while also increasing the accuracy rate to its highest attainable level of 100%. Rapur and Tiwari [24] accounted obstructions and impeller cracks. The hyper parameter of the classifier and the input features produced from the time-domain vibration data were set optimally. The study also demonstrates how this issue may be used in various industries to assess the severity of a CP blockage fault. This study examined the features of five different obstruction circumstances and cavitation. Bordoloi and Tiwari [25] used the features retrieved from the time-domain vibration signal coming from the pump casing and bearing block. Optimization techniques were used to choose the kernel and SVM parameters. Also, with greater pump speeds and better fault prediction accuracy was shown.

Panda et al. [26] performed training and testing at greater rotational speeds using the SVM, the accuracy of prediction in multi-class fault classification (various levels of obstruction) was moderately higher. At higher rotational speeds, the oncoming bubble creation was also shown to be quite correctly anticipated. Rapur and Tiwari [27] introduced a statistical feature called the inverse of standard deviation, σ^{-1}. This trait is proven to have enormous promise for defect diagnostics because it lowers variance, σ^2, while maintaining the trends of standard deviation. Also, it was noted that the combination of motor current and vibration data characteristic was particularly promising for investigation of CP flaws classification. All of the multiclass classifications could be completed by the SVM classifier created utilizing μ (mean) and σ^{-1} features with nearly flawless accuracy. Chakravarthy et al. [28] examined various faults in centrifugal water pumps powered by induction motors, which was used in agricultural fields and proposed a algorithm to identify and classify the faults. Alabied et al. [29] proposed an approach on intrinsic time scale decomposition (ITD) for feature extraction and SVM for classification of health conditions. ITD is a powerful tool for extracting the most useful information from motor current signals. Lei et al. [30] detected the health conditions of machines automatically using conventional machine learning ideas and also transfer learning theories have the potential to create diagnostic models that can be used to fill the gap by allowing the knowledge of diagnosis to be applied to various diagnosis tasks. Dewangan et. al. [31] developed a methodology for identifying and assessing the severity of blockages and cavitation in centrifugal pumps using fluid pressure data and deep learning-based binary classification with accuracy of 95.8%. This approach aims to prevent operational failures and system breakdowns by providing accurate monitoring and fault detection.

Sha Y. et. al. [32] proposes a novel cavitation detection framework for valves using XGBoost with adaptive feature selection. It addresses small-sample issues with a non-overlapping sliding window and enhances feature extraction using FFT. The method significantly improves cavitation detection accuracy, achieving a 4.67% increase for binary classification and an 11.11% increase for four-class classification over traditional XGBoost. Orrù et al. [33] used the SVM and the multilayer perceptron (MLP) algorithms, and compared both. The MLP exhibited superior classification performance by accurately identifying two out of the four failures that occurred within the specified timeframe. In contrast, the SVM approach demonstrated higher precision but poorer recall for the positive class. Chen [34] altered the traditional Mahalanobis distance fault for CPs with the introduction of the *k*-nearest neighbour algorithm (KNN). The fault prediction accuracy of the off-balance state was up to 82% using this technology, which used vibration signals to detect distinct CP fault types. They proposed a deep transfer convolutional neural network (DTCNN). ResNet-50 was chosen as the pre-trained deep convolutional neural network model and was transferred to solve the problem of bearing fault classification using the transfer learning concept. The proposed method was validated using two datasets from motor bearings and self-priming CPs. The proposed model was tested on two well-known datasets, and it achieved a near-perfect prediction accuracy, surpassing alternative conventional machine-learning and deep learning techniques. Sunal et al. [35] examined

the most recent advancements in machine learning-oriented pump condition monitoring and fault diagnosis research and development.

Dewangan and Tiwari [36] utilized ANNs to detect and classify cavitation and blockage faults using time-domain analysis. They obtained the classification performance and confusion matrix for different conditions with speed and without consideration of speed as an input. Rapur and Tiwari [37] suggested a flexible approach, which was based on a multi-class SVM classifier with hyperparameter optimization for characterizes the CP fault condition and identifying the traits that will classify the 33 CP errors. The selected combined characteristics with $1/\sigma$ were performing well at classifying faults. The methodology was tested at 8 distinct operational frequencies to determine how robust the established technique was. The developed algorithm demonstrated remarkable precision across all speeds in categorizing various forms of flow instabilities, mechanical faults, and their combined occurrences. Rapur et al. [38] approached for meeting the needs for machine maintenance is condition-based maintenance (CBM). Modern CBM techniques strive to minimize human involvement in the real-time defect identification. This review outlines the unexplored areas related to CP fault diagnostics and gives future researchers in this subject a good start. Zaman W., et. at. [39] introduced the SGST method, combining STFT, GAN, and Swin Transformer, to diagnose cavitation faults in hydraulic machinery, achieving 98.6% accuracy. It enhances feature extraction and data augmentation, significantly improving fault diagnosis accuracy over existing methods. Sunal C.E. et. al. [40] detected centrifugal pump faults using real operational data, employing binary classification of visual features from DQ/Concordia patterns with ResNet-34, achieving up to 85.51% classification accuracy. Indriawati K. et. al. [41] approached leverages transfer learning from image detection to solve real-life engineering problems. This study developed a fault detection system for centrifugal pumps, specifically the Medium Pressure Oil Pump (MPOP) and the Water Injection Pump (WIP), using residual feature extraction in the time domain and a statistical approach. The system achieved 91.67% accuracy for MPOP, 94.8% for WIP, and over 99% accuracy during online monitoring simulations.

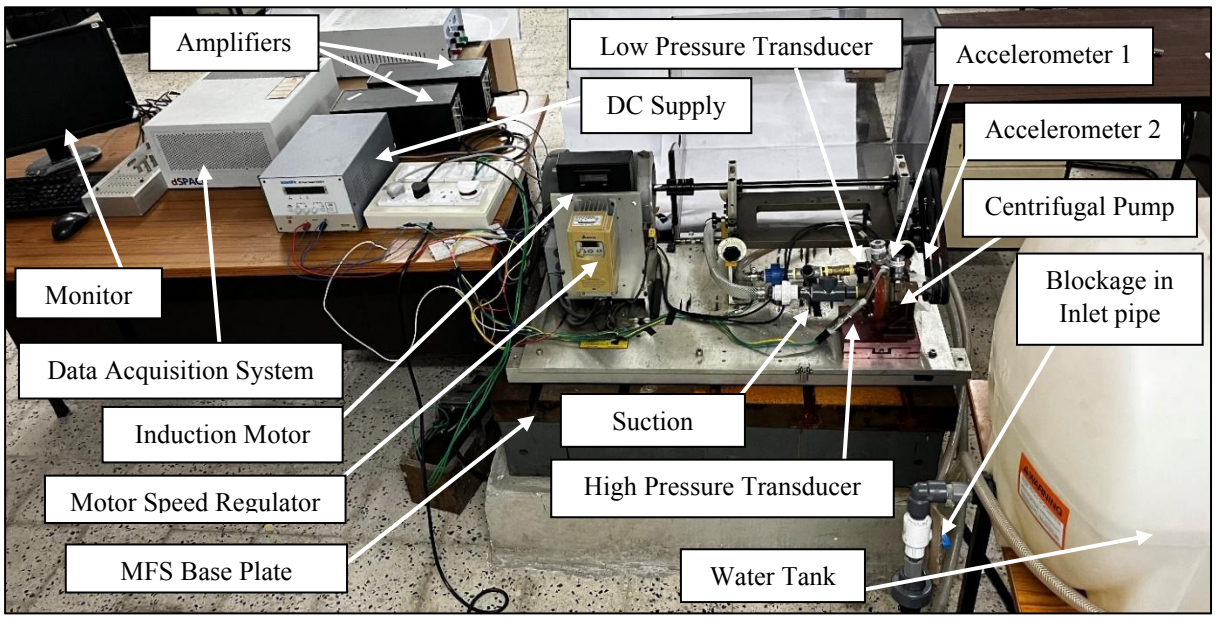
Our review of the literature underscores the importance of selecting the right signals for data collection in fault identification. Research highlights that pressure, acceleration, and motor line current signals are the most informative regarding the condition of CPs. However, no existing studies have combined all three signals—pressure, acceleration, and motor line current—to detect various degree of blockages in healthy and two different faulty CPs. Newer algorithms and data classification techniques are surpassing older methods in performance. A significant gap in current research is the limited focus on detecting various blockage degrees in the inlet pipe of centrifugal pumps based on different features extraction. Our study aims to address this gap by focusing on degree of blockage detection in the inlet pipe of CPs. What makes our work unique is the innovative combination of acceleration, pressure, and motor line current data to detect different degrees of blockage using an artificial neural network (ANN) algorithm using different features. Furthermore, we are inspired to explore various statistical features derived from basic data, using them either together or separately, to enhance the accuracy of detection of various blockage degree.

This study is separated into the upcoming sections as follows: Section 2 discusses the experimental setup and data capture. Explanation of data classification methods is covered in Section 3. Section 4 highlights the performance and the accuracy of fault detection. Section 5 presents the final conclusions.

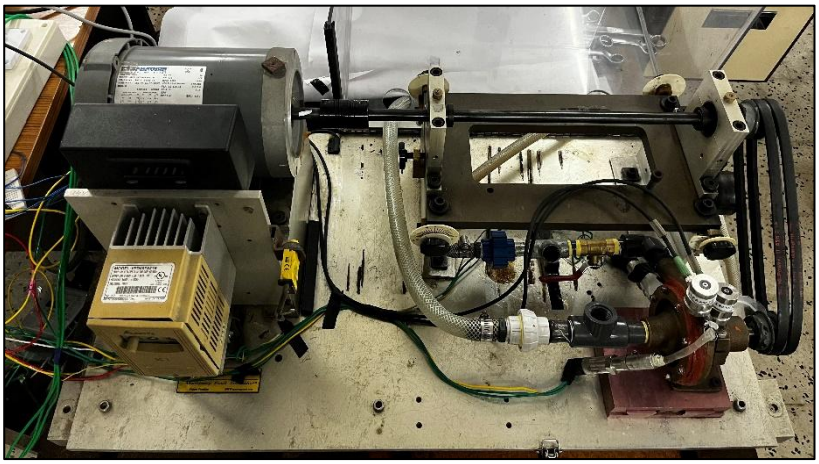
2. Experimental setup and Experimentation description

Machine Fault Simulator (MFS) provided by Spectra-Quest^{TM} was used for the experimentation. Figure 1 shows the experimental setup and the close look of CP. A pump was installed on the fixed base of the MFS and is driven by a double-belt pulley system. Within the MFS setup, a 3-phase

induction motor is linked to the rotor via a flexible coupling. The rotor is supported by two bearings at its extremities, and these bearings are securely positioned between two stationary plates. The rotor shaft is attached to the CP using belt drives to power it. The speed of the pump could be varied by varying motor frequency using a variable frequency drive (VFD).



(a)

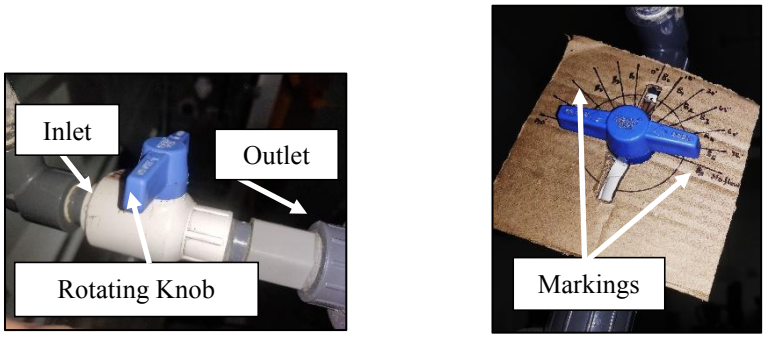


(b)

Figure 1: (a) Experimental set-up (b) A close look

Leak-proof fittings were made sure to be on the pump. Figure 2 illustrates the presence of manual modulating valves at the inlet of the water tank, facilitating flow regulation.

To prevent cavitation-related issues, the water tank is strategically positioned to ensure an ample head at the pump inlet. By using a mechanical modulation valve at the suction and discharge ends of the line connected to the pump outlet, the pressure can be varied. The data is acquired with 5000 samples per set. For every pump working situation, 150 sets of data aggregating 150 s are recorded. Each set of data has a 0.1-second timeframe.



(a) (b)

Figure 2: Mechanical modulating valves (a) Rotating knob (b) Markings on the knob

The pump is operated at a speed of 30 Hz to 60 Hz with an interval of 5 Hz. The motor was heated up at a speed higher than 60 Hz. Six equal intervals are marked on the mechanical modulation valve, so that different degrees of clogging could achieved by governing valve in different intervals. B0 employs 0% clogging (unrestricted flow, no disturbance), B1 indicates 16.7% clogging, B2 reveals 33.3% clogging, B3 displays 50% clogging, B4 shows 66.6% clogging, and B5 indicates 83.33% clogging.

Accelerometers: A pump with two triaxial accelerometers affixed with sensitivities of 100.3, 100.7, and 101.4 mV/g (accelerometer-1) and 101, 101.1, and 101.4 mV/g (accelerometer-2) in the *x*, *y*, and *z* directions, respectively, were used to analyses acceleration. Figure 3 depicts the orientations and accelerometers.

Pressure Transducers: For measuring liquid pressure, two pressure transducers from "Nictech," a sensitive silicon chip is used. Current fluctuation can be used to interpret the response from the pressure sensors when they are operating in the operational range (0-60 psi) (4–20 mA). High- and low-pressure transducers are shown in Figure 4.

Current sensing probes: The current line reading was recorded using "The Keysight 1146B" current probes, as shown in Figure 5. Current of 100 mA to 10 A rms can be measured with current probes.

Data Acquisition System: The NI PXI - 4472 is an 8-channel and NI PXI -6251 is a 16-channel dynamic signal acquisition module for high accuracy frequency domain measurements. The eight channels of the NI PXI -4472 and the 16 channels of the NI PXI -6251 simultaneously digitize input signals over a bandwidth of DC to 45 kHz. Data Acquisition system is shown in Figure 6.

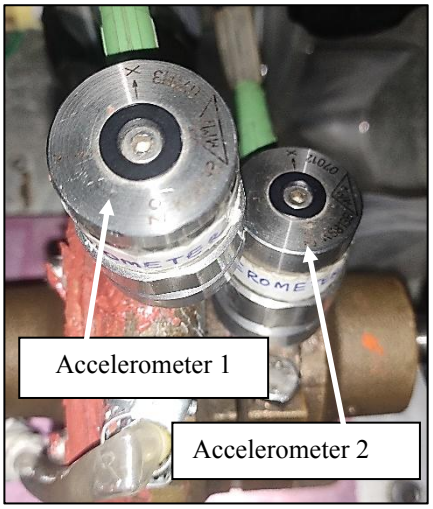


Figure 3: Sensor Placement with Tri-axial Accelerometers on Pump Casing and Bearing Housing.

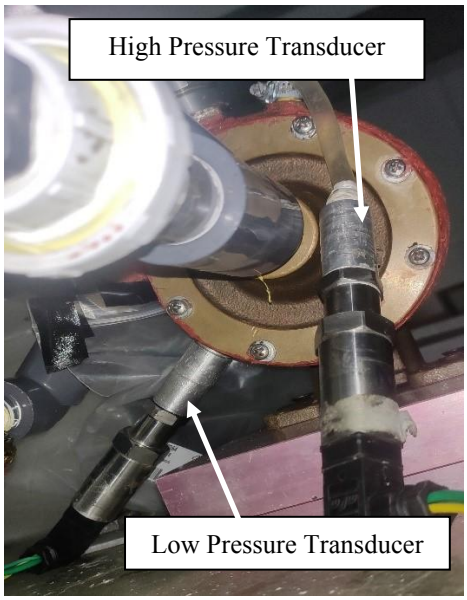


Figure 4: High- and low-pressure transducers mounted on the pump.

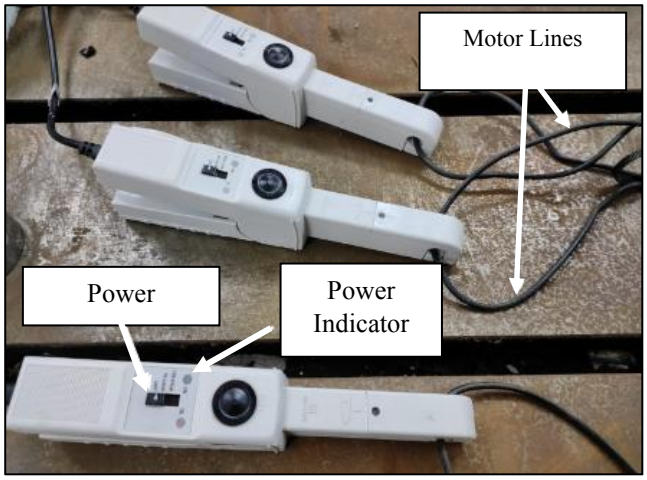


Figure 5: Current wires passing through Current probes

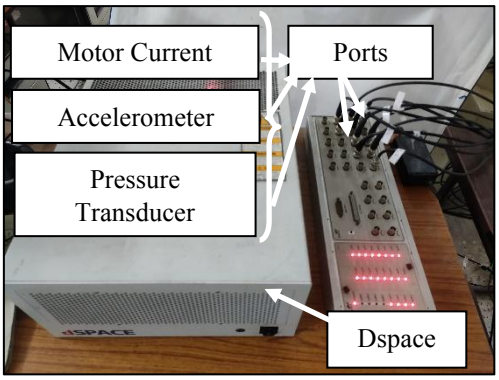


Figure 6: Dspace for collecting data

2.1. Description of experimentation

On the MFS, an experiment was performed to collect data at various levels of flow obstruction (B0, B1, B2, B3, B4, B5) at different frequencies (30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz and 60 Hz). For each combination of pump running speed and block level, data collection was done. In order to obtain comprehensive insights into the fault, three distinct types of sensor measurements were taken which are accelerometers, pressure transducers, and current probes. Accelerometers measure the system's vibration, low and high-pressure transducers measure the system's pressure, and current probes measure the current in the lines. To capture vibration-related data, two accelerometers were strategically placed. The first accelerometer (accelerometer-1) was affixed to the pump casing, while the second accelerometer (accelerometer-2) was positioned on the bearing housing. On the pump casing, two pressure transducers measure the low and high pressure of the system. Three current probes were deployed to measure the motor line current. To collect the data, all of the sensors were linked to the DAQ system's various channels. The control desk software utilized to record data that was gathered by the sensors. Abbreviations used in this study and the details of DAQ are given in Table 1 and Table 2 respectively.

Table 1: Abbreviations	
σ	*Standard Deviation*
μ	*Mean*
κ	*Kurtosis*
χ	*Skewness*
x	*Input to the neural network*
w	Weight vector
b	Bias vector
$\hat{y}$	Predicted class
z	Loss function
T	Transpose
n	Total number of data points in a collection
x_i	Amplitude of each data point
CP	Centrifugal Pump
HP	Healthy Pump
IF	Impeller fault
PF	Cover plate fault
RF	Random Forest
SVM	Support Vector Machine
DT	Decision Tree
KNN	K-Nearest Neighbours
XGB	XGBoost
ANN	Artificial Neural Network
SDG	Stochastic Gradient Descent

Table 2: Data Acquisition System outlines	
Blockage levels	B0 (Full Flow, no obstruction), B1 (1/6 obstruction), B2 (1/3 obstruction), B3 (1/2 obstruction), B4 (2/3 obstruction), B5 (5/6 obstruction)
Frequency	30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz and 60 Hz
Quantity of each case in every pump condition	6 (Blockage levels) × 7(Frequency) = 42 cases
No of measurement sets for each combination of blockage level and frequency	150 sets
Data collection time for each blockage level and frequency combination	150 sec
Timeframe for a single set collection	1 sec
Sample captured in one second	5000

Three-wire setup: Pressure sensor have output in the Ampere form, which cannot measure by the DAQ, so to make it possible we have to convert the signals into the voltage and it can be possible by the connection shown in Figure 7 with the help of resistor connected in the bread board. According to the Ohm's law, the voltage across a conductor is directly proportional to the current flowing through it, provided all physical conditions and temperatures remain constant. This setup is called the three-wire setup.

Simulink Model: A Graphical MATLAB-Based Environment for Modelling, Simulation, and Analysis of Multi-Domain Dynamical Systems as shown in Figure 8. The description of the model is given as the step size used is 0.0001 with infinity time and fixed step of time. Ode (Runga-Kutta method) solver is used. This model facilitates the ingestion of data from 11 sensors, visible on the monitoring screen within the control desk app. The acquired data is subsequently processed by a Data Acquisition (DAQ) system, resulting in raw output data.

2.2. Measurement Procedure

Vibration, pressure, and current lines signatures were used for diagnosis of blockage level. Control Desk software was used to configure data acquisition. To conduct time domain measurements, a sampling rate of 5000 samples per second was employed, resulting in the acquisition of 5000 samples. Data collection was conducted over a period of 150 seconds. For each of the sensors, 5000 x 150 non-overlapping data points were captured.

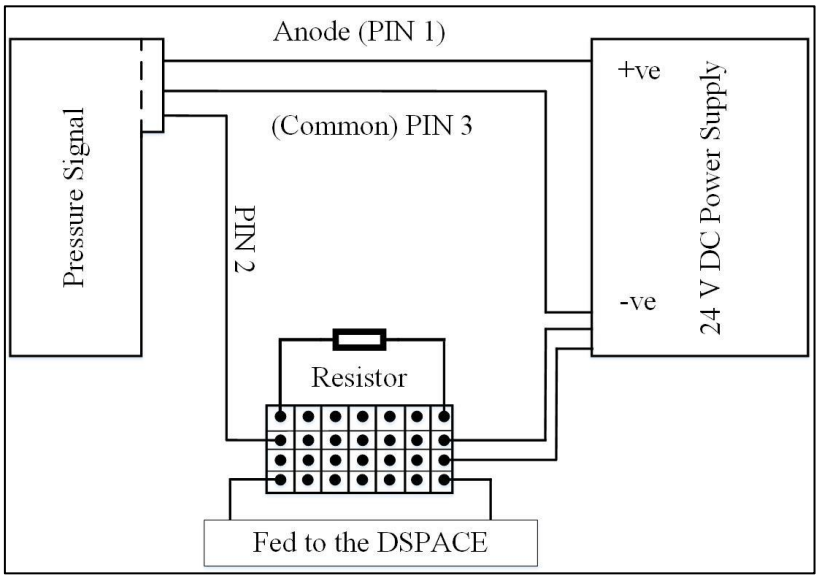


Figure 7: Three wiring setup to convert pressure transducer signals (Ampere form) to voltage form (measurable form by DAQ)

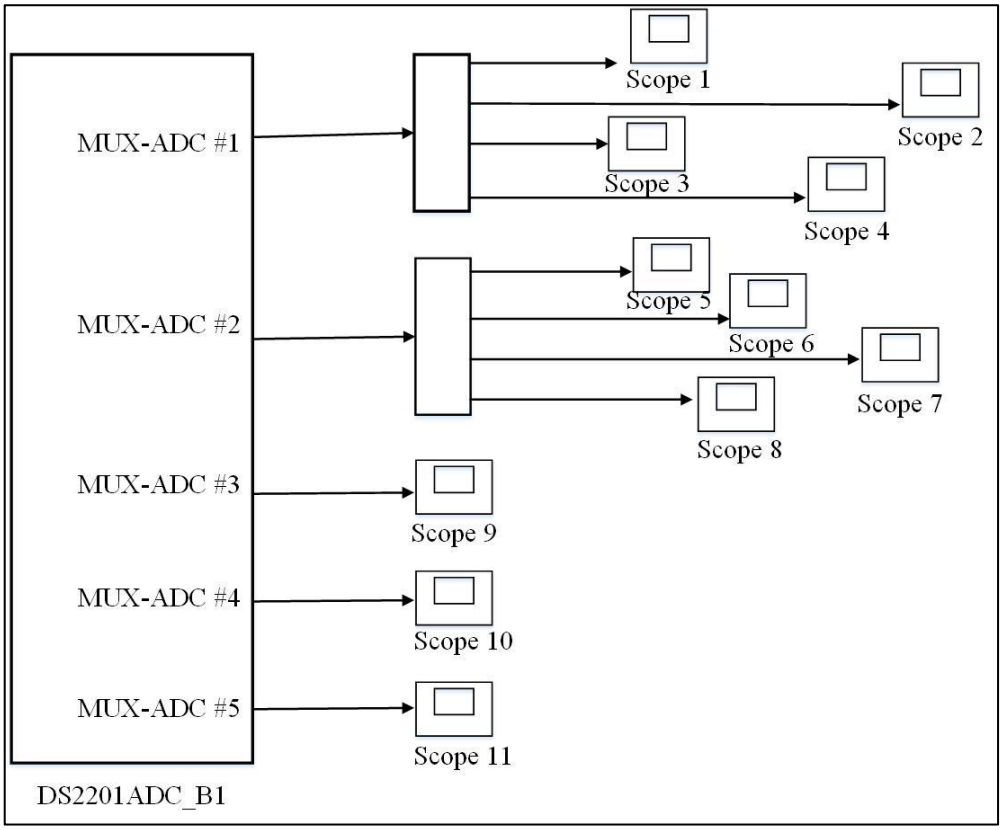


Figure 8: Simulink Model for display the sensors data signal in the monitor screen

The data were recorded on system hard drive, saved in separate DAQ measurement files in .csv files for each disturbance at individual CP speeds. The time domain data can later be converted to any domain using various transformations. A healthy pump, a pump with an impeller defect, and a pump with a cover plate defect were sequentially mounted on the MFS. Impeller fault and cover plate fault is shown in Figure 9. The impeller plate has cuts on the blades and the cover plate fault has number of pits on it and these are supposed to formed due to the high-pressure bubbles burst and hits the metal wall in the form of micro jets and the cuts on the impeller blade and pits on the cover plate, both are the actual fault type of the CP, which is created artificially so that signals created to the artificial fault can developed actual signals. The cavitation will cause such an obvious and regular failure but here it is created artificially by third party and available in our lab. Damage can arise from the dynamic water flow's swirling motion within the turbine, triggering a range of effects. Dynamic water flow over turbine blades can lead to damage in various ways in which some are as follows: (1) Cavitation: Cavitation occurs when water flowing over turbine blades experiences pressure drops, causing vapor bubbles to form and collapse, leading to shockwaves that erode the blade's surface, potentially causing cuts. (2) Vortex Fluctuations: Swirling water causes pressure changes and stress on blade surfaces, leading to fatigue and cuts over time. (3) Abrasive Erosion: High-speed, sand-laden water erodes blades, creating cuts and grooves in sediment-rich areas. The design and construction of the pump-turbine system could be one reason. The striking of water may be effects more on the cover plate than turbine blade. Hence cover plate has a serious issue for cavitation than turbine blade. For each of these 3 pump conditions, a suction blockage defect with six severity levels was introduced.

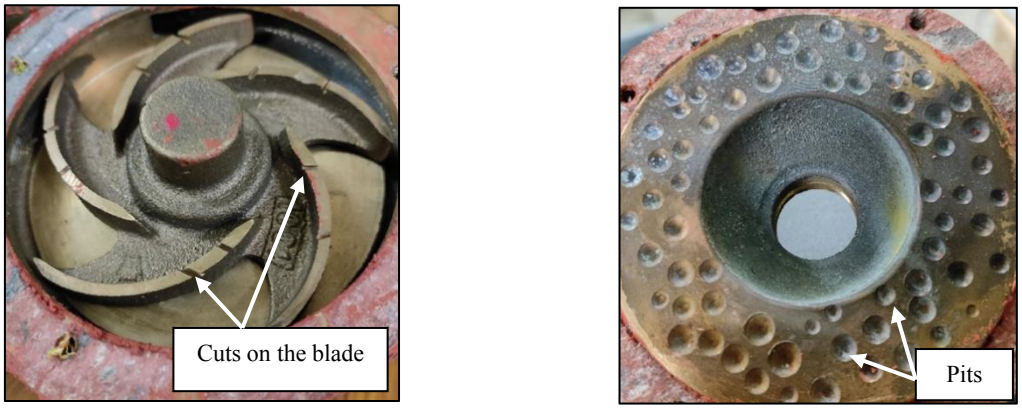
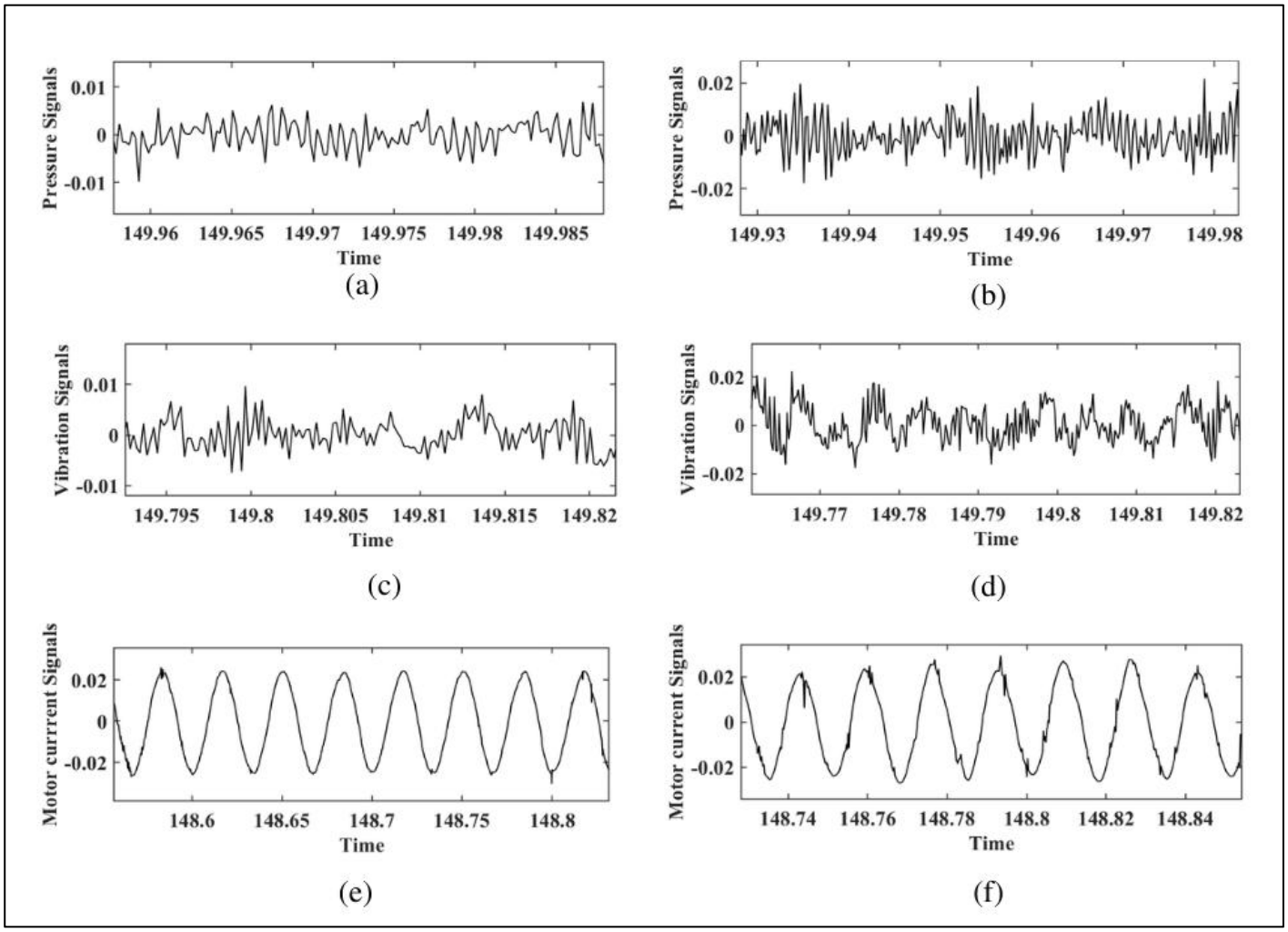


Figure 9: Impeller fault (left) and Cover plate fault (right)

In each pump state, the pump was operated at a series of different speeds. A variable frequency drive was used to operate the induction motor from 1800 rpm (30 Hz) to 3600 rpm (60 Hz) in steps of 300 rpm (5 Hz). Vibration, pressure, and current signatures were used for blockage level diagnosis. For each faulty condition, data were collected at 5000 samples/s. Data for each condition was collected for 150 seconds, resulting in a total of 150 records collected for each fault condition. When the pump is faulty and the suction clogging level is full flow (without any obstruction level) and with an obstruction level of B5 by using the signatures of vibration, pressure, and line current at 30 Hz of impeller fault pump configurations are shown in Figure 10. It is observed that as flow restriction increases to the maximum level (B5), the pressure, vibration, and current signal pulsations also increase, as shown in Figure 10 (b, d, f), compared to the conditions with no flow restriction as shown in Figure 10 (a, c, e). As the liquid flow area diminishes, the pressure of the liquid decreases correspondingly. When this pressure falls below the vapor pressure of the liquid, bubbles begin to form. Increased flow restriction leads to more bubble formation, which in turn elevates vibration levels and pressure pulsations. Also, bubble formation occurs even at a low blockage level with a high running speed. The description of each pump condition corresponding to each in Figure 10 can be found in Table 3.

Table 3: Description of Figure 10 Pressure, vibration, and motor current signals of impeller fault pump without obstruction level and final obstruction level (B5)

Figure No.	**Signal**	**Blockage Levels**
10(a)	Pressure	B0
10(b)	Pressure	B5
10(c)	Vibration	B0
10(d)	Vibration	B5
10(e)	Current	B0
10(f)	Current	B5



Units: Time (in seconds) and Pressure, Vibration and Motor current signals (in volts)

Figure 10: Pressure, vibration, and motor current signals of impeller fault pump without obstruction level and final obstruction level (B5) at 30 Hz (a) Impeller fault pump pressure signals without an obstruction (b) Impeller fault pump pressure signals with an obstruction of level B5 (c) *Impeller fault pump vibration signal without an obstruction* (d) *Impeller fault pump vibration signal with an obstruction level of B5* (e) *Impeller fault pump current signals without an obstruction* (f) *Impeller fault pump current signals with an obstruction level of B5*

2.3. Description of Fault Set

A total of 18 different fault conditions of the CP have been considered in this study. Healthy pump with no blockage (HP0), healthy pump with suction blockages (HPb), Impeller fault with no blockage (IF0), impeller fault with suction blockages (IFb), cover plate fault with no blockage (CP0) and cover plate fault with suction blockages (CPb), where $b = 1,2,3,4,5$. The fault set has six classes. We have identified a total of 18 faults, as outlined in Table 4. These faults encompass six distinct degrees of blockage levels across three pumps. Our classification involves organizing the fault set based on the blockage level. Each class comprises data from all three pumps, and within each class, we vary the blockage level. By incorporating all six blockage levels, we create a comprehensive fault set conditions that make up these sets have been explained in Table 4. Class 1 (B0): HP0, PC0, IF0 represents a condition with zero blockage, allowing full flow. This class includes all datasets with no blockage across all three pumps: one healthy pump and two faulty pumps.

Table 4: Fault classification sets description	
Fault Set	**Classes (labels)**
1	Class 1(B0): HP0, PC0, IF0 Class 2(B1): HP1, PC1, IF1 Class 3(B2): HP2, PC2, IF2 Class 4(B3): HP3, PC3, IF3 Class 5(B4): HP4, PC4, IF4 Class 6(B5): HP5, PC5, IF5

3. Data classification methodology

The experiments yielded pressure, acceleration, and current signatures that were used for fault categorization at various frequency. Deep learning-based neural networks are employed for fault classification. A machine learning method called deep learning enables computer systems to get better with practise and data. When solving challenging environmental issues in the real world, this method works incredibly well. Supervised learning and unsupervised learning are two subcategories of machine learning algorithms. In this study, supervised learning has been used to achieve its goals. The programme is instructed to divide/classify the data in accordance with various characteristics. The method creates the function called $y = f(x)$, which bridges the input and output. Deep learning uses optimization algorithms to enhance the performance of fault classification. A machine learning algorithm is created by combining various optimization algorithm elements, such as an optimization method, a model, and a dataset. In real-world scenarios, the algorithm seeks to reduce training error rather than finding the optimum function. Various hyperparameters can also be tuned to enhance the performance of a learning algorithm. A subfield of machine learning known as deep learning. It is a discipline where independent learning and growth depend on the study of computer algorithms. The machine learning relies on more basic principles, whereas deep learning makes use of artificial neural networks designed to replicate the ways of people think and perceive information. A typical neural network consists of several layers of interconnected neurons, known as neural nodes. These layers include a single input layer, one or more hidden layers, and a single/multiple output layer. When a node receives an input, it performs a transformation on that input and then sends an output based on the result of that transformation. This study utilizes an architecture comprising one input layer, six hidden layers, and an output layer with five classes. Here is the fundamental structure of a Neural Network.

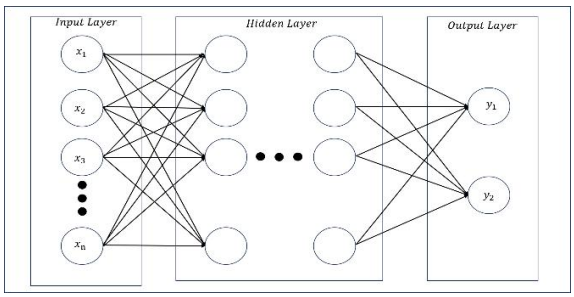


Figure 11: Structure and Function of Neural Networks

The network receives input **x** from the input layer. Following the input layer, several hidden layers extract data that becomes increasingly important. Finally, the output layer predicts the kind of class of the input. During training, the network aims to establish a mapping between the input data vector x and its corresponding labelled category y. In this process, the neurons within the initial hidden layer

compute a fundamental vector function, denoted as z, which serves as the starting point for mapping. This function involves the weighted sum of the input x along with a bias. The expression for **z** is as follows:

$$\mathbf{z} = \mathbf{w}^T \mathbf{x} + \mathbf{b} \tag{1}$$

Here, w and b represent the weight and bias vectors, respectively, while the superscript T signifies the transpose operation. The variable x denotes the input supplied to the neural network.

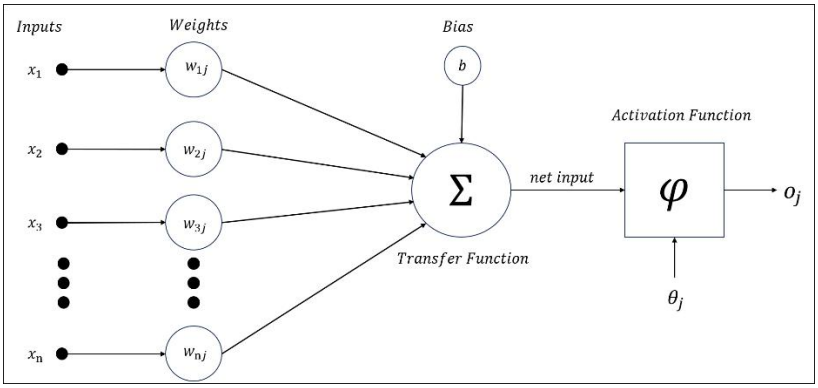


Figure 12: Neuron in a neural network

An activation function is applied to compute the vector function **z**. It is discovered that the "Rectified linear unit" (ReLU) activation function performs better than all other activation functions. The subsequent hidden layer receives the output, which includes the bias of the weighted sum of the activation function. Ultimately, data categorization is achieved through the output layer. Notably, the 'SoftMax' activation function proves effective within the output layer for multiclass classification tasks. This is due to its capacity to yield values within the 0 to 1 range, facilitating the creation of a probability distribution across various events. The chance of every target class out of all possible target classes are returned by this function. Just using simple mapping, the discussed process assists in completing a complex routing.

$$s\left(z_i\right) = \frac{\exp\left(z_i\right)}{\sum_{i=0}^{k} \exp\left(z_i\right)} \tag{2}$$

For error computation, the sparse categorical cross entropy loss function is used in this work. The weights and biases of nodes of the neural network will be adjusted in accordance with the law of gradient descent.

$$\begin{aligned} w_{new} &= w - \alpha\, dw \\ b_{new} &= w - \alpha\, db \end{aligned} \tag{3}$$

where, w_{new} is the updated weight, α is the learning rate, w is the previous weight and dw is the partial derivative of the cost function with respect to the weight. Also, b_{new} is the updated bias, b is the previous bias and db is the partial derivative of the cost function with respect to the bias. The local sparse categorical cross entropy loss function is used to prevent situations from developing. As our data classes are mutually exclusive, we use sparse categorical loss function as its computation is faster. Expression for this loss function is:

$$L = -\log\left\{p\left(s \in c\right)\right\} \tag{4}$$

where L refers to loss function, s represents samples, c refers to classes, and $s \in c$ addresses to sample s belongs to class c. Enhancing the classifier's performance can be achieved by adjusting different hyperparameters. Improperly adjusting the hyper-parameters results in either "under fit" or "over fit"

circumstances. In this paper, open-source deep learning libraries 'Keras' and 'Scikit learn' are utilized in Python (Jupyter) software.

4. Results and Discussions

As stated previously, the purpose of this study is to recognise severity of the block level utilized multiclass classification with the help of multiple sensor data. Only six statistical features are considered and retrieved from raw sensor data in order to avoid excessive duplication in the input data. The picking of the features for a given set of data and defect is not predetermined. As a result, features such as mean (μ), standard deviation (σ), mode (M), Entropy (S), kurtosis (κ), and skewness (χ) are considered, which can refer to various data aspects. As a result, only six features are examined. We have used the multi sensor signals to capture the more features, which can help us to predict the best and accurate result. With the help of raw data, we have found six features of each sensor (total 11) separately, and we have used the six feature indicators.

4.1. Feature Introduction

Mean (μ): Mean represents the statistical average of values of the data points of the signal.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5}$$

where n is the total number of data points in a collection and x_i is amplitude of each data point.

Standard Deviation (σ): The standard deviation is a measurement of how far data point values deviate from the data set's mean value.

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{\left(x_i - \mu\right)^2}{\left(n-1\right)}} \tag{6}$$

Entropy (S): Entropy serves as a universal indicator of system disorder and is mathematically represented as-

$$S = \sum_{i=1}^{n} P(\mathrm{x}_i)\log_{10} P(\mathrm{x}_i) \tag{7}$$

Skewness (χ): Skewness is an indicator of how asymmetrically the probability distribution is distributed around the mean.

$$\chi = \sum_{i=1}^{n} \frac{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^3\right)}{\sigma^3} \tag{8}$$

Kurtosis (κ): Kurtosis is the measure of the extent of spikiness or flatness of the data points.

$$K = \sum_{i=1}^{n} \frac{\left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^4 \right)}{\sigma^4} \quad (9)$$

4.2. Test of feature performance

Three approaches were used to examine the effectiveness of the retrieved statistical features as follows: first individual sensor data are considered separately, second is the combination of the few bests performing features, and lastly on all the best features set.

4.2.1. Individual performance of features

It can be computationally expensive to use raw data directly. Additionally, it can contain redundant data, which would have an impact on how the network is trained. The size of the data and the temporal complexity can be reduced while still capturing meaningful information by extracting features. The different statistical features considered in this study are standard deviation, variance, mean, median, mode, kurtosis and skewness. As there is no specific rule to choose the algorithm therefore these extracted features fed as input to a Kernel SVM, Random Forest, Logistic Regression algorithm. At each frequency (say, 30 Hz), datasets from blockage levels B1 to B5 are collected and analysed using various algorithms to determine the classification accuracy of each individual feature (say, σ) as shown in Figure 11. This process is repeated at 45 Hz and 60 Hz and take the average of them also. Each individual feature is tested for their performance with different algorithms. It is found that RF performs better than Kernel SVM, Logistic Regression algorithm. Also, in the RF model with mean gives accuracies of 98.2% at 30 Hz, 98.87% at 45 Hz and 99.81% at 60 Hz, with overall average accuracy of 98.96 %, which is the highest amongst all the statistical features considered. So, the RF model is considered for further combination of feature calculations and the mean is considered as the best statistical feature among all others but the standard deviation and the entropy are also close to the mean. So, the mean, standard deviation and entropy will take into consideration while finding the blockage severity accuracy and other features are ignored as of their low performances as shown in the Figure 11. The description of the RF is given as the estimator used is ten with ten cross validation fold and the criteria used is entropy. It is found that features work less accurately at lower frequencies (30 Hz) compared to higher ones (60 Hz). This shows that the individual feature matters less for accuracy at low frequency but does vary at higher frequencies (50 Hz - 60 Hz). It seems that signals change a bit at lower frequencies, while more bubbles form at higher frequencies during our experiment. Using just one feature at low frequencies does not give us useful information. So, we tried combining different features, as shown in Figure 12, and saw that this improved how well the classifier works.

4.2.2. Combination performance of features

A single feature is insufficient at the slow speed to gather any insightful data. Various combinations of these qualities are tested to overcome this issue, one can observe that by using the combination of the features there is an enhancement in the classifier's performance (refer to Figure 12). Now this pattern can be observed that by increasing the frequency, the performance of the classifier is increasing. At higher blockage levels and frequencies, the complex interplay of fluid dynamics, system response, and sensor signals provides the model with more informative data for accurate blockage detection and classification. Higher frequencies result in the flow of fluid through the system becoming more turbulent, leading to clearer signals due to increased effects such as vibration,

pressure pulsations, and current lines, along with increasing obstruction levels, which result in more pronounced changes in pressure, velocity, and other flow parameters. This indicates an improvement in signal capture, as shown in Figure 12, where we have combined (σ,μ), (σ,s), and (μ,s) to find that the classification accuracy of the combined features is better able to predict the signals compared to individual features. The combination feature, which performs best is the standard deviation and mean, with accuracies of 98.80% at 30 Hz, 99.05% at 45 Hz and 99.35% at 60 Hz, and with an average accuracy of 99.06%, which is the highest amongst all the statistical features set considered. Now several two best feature combination (σ and μ) are chosen, which is shown in Figure 12. As speed increases, performance of classifier also increases because combining of all the statistical features gives better result as shown in Figure 12.

4.2.3. Performance with combination of all features

All these features set (σ,μ,s) can be used together for further improvement in the classification accuracy. The standard deviation, mean, and entropy achieved accuracies of 98.86% at 30 Hz, 99.44% at 45 Hz, and 99.67% at 60 Hz, averaging at 99.26%. This highest average accuracy among all considered statistical features indicates the precision of this feature set in capturing signals. After considering all three features, the difference in accuracy with speeds is depicted as indicated in Figure 12. It is evident that the classification accuracy rises with frequencies, which is a sign that obstruction level severity is enhancing by the classification accuracy. The trend indicates that classification accuracy rises with both higher frequencies and more severe blockages, suggesting a link between accuracy and cavitation severity. Low accuracy points to typical signal variation, whereas high accuracy hints at pronounced signal changes due to increased bubble formation and more numerous signals. The depicted performance of all features is shown in Figure 12. As a result, all the features are chosen at once for the remaining of the study.

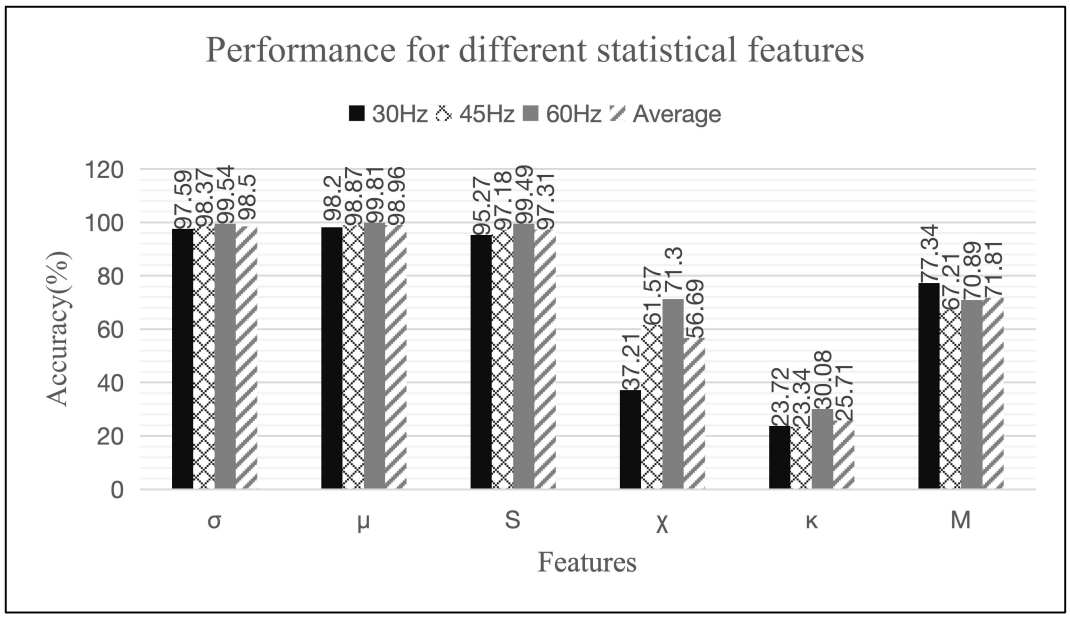


Figure13: Random Forest algorithm used to identify the performance for different statistical features at three different frequencies and their average

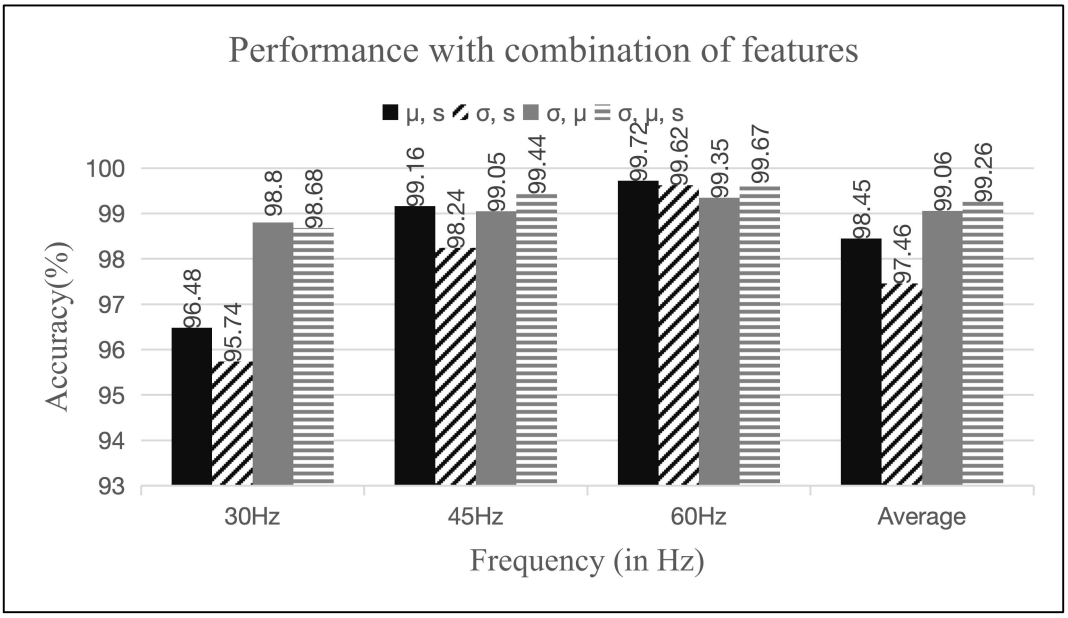


Figure 14: Classifier performance versus Frequency with combination of the features at three different frequencies

4.3. Classification algorithms for blockage severity

A variety of methods can be applied for the classification purposes for a given input. Determining the best algorithm for the application in issue is essential. In this investigation, a spectrum of classification techniques was explored, encompassing XGBoost (XGB), Decision Tree (DT), K-Nearest Neighbours (KNN), Random Forest (RF), and the intricate realm of Artificial Neural Network (ANN). Data from all operating frequencies are used. Pressure signals, vibration signals and motor current signals are also taken into consideration. The accuracy of classification of these algorithms have been presented in Figure 13(a). All classification methods are tuned and the neural network demonstrates to be the best classifier in a tuned state. The neural network exceeds all other classification techniques with a fault set accuracy of 99.38%.

In Section 4.4, the steps of tuning the Neural Networks are described and then this model is used to predict the blockage severity, which is of five classes as class 1 (B0) is ignored because it is without any disturbance (fully flow). So, classes from class 2 (B1) to class 6 (B5) is taken into consideration.

4.4. Steps for tunning the Artificial Neural Network

Before going to start tuning process, few things we have taken into consideration by working iteratively again and again over the model. We used a systematic approach for hyperparameter tuning, focusing on manual exploration and iterative adjustments based on real-world performance data. We targeted critical factors like the number of hidden layers, batch size, epochs, activation function and weight initializations at different train test ratio to enhance model performance and optimized our model with up to six hidden layers and five output layers. We also fine-tuned batch sizes around 40 to strike a balance between computational efficiency and model accuracy. We have considered six hidden layers because it is found that less than the six hidden layers are not performed that much better as six hidden layers do. The number of neurons set in each hidden layer is 55 and the output layer has 5 neurons as five classes (B1 to B5). In the hidden layer Relu activation function is used and for the output layer, Softmax activation function

is used as it is found that both of them have the best combination. The number of epochs found to be 130 as at this number the model performance is at its peak. It is found that if we increase the epochs greater than the 130 then the model performance started decreasing, and 40 minibatch sizes were found by performing iteratively. Table 5 provides the tuned neural network model configuration. To avoid the overfitting, dropout regularization is also used, which randomly drops a number of neurons in a neural network during model training on the basis of probability. In this model, we found 0.5 works best as the dropout regularization probability.

4.4.1. Final pre-processing step – Feature scaling

When numerical input variables are scaled to a standard range, many machine learning algorithms perform better. In this investigation, distinct signatures had quite diverse numerical values. This might result in various gradient descent step sizes for certain features. Feature scaling is used to make the gradient descent to the minima more flexible. The feature scaling method used in this study is standard scaling, often known as standardisation. Mathematically standardization represents as

$$X' = \frac{X - \mu}{\sigma} \tag{10}$$

4.4.2. Selection of the training testing ratio

In neural networks, various training testing ratio like as 25:75, 30:70, 50:50, 80:20, 90:10 is tested to get the best classification accuracy. Over all of the consideration, 80:20 training testing ratio worked best. So, 80:20 training testing ratio is taken into consideration in further neural network. The training accuracy of 80:20 ratio is 97.41% and the testing accuracy trigged with the 92.16%. The better accuracy with an 80:20 train-test split compared to a 90:10 split is likely because the larger test set (20%) provides a more comprehensive and stable evaluation of the model's performance, reduces the risk of the model being too tailored to the training data, and results in more reliable and realistic performance metrics. It can be shown in Figure 13(b).

4.4.3. Selection of the optimization algorithm

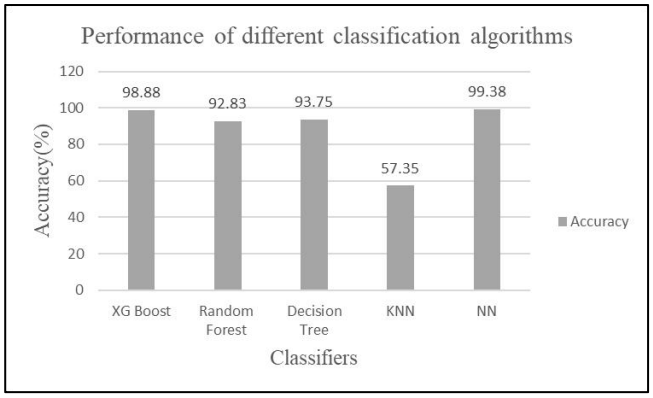
In neural networks, optimization algorithms train the weights and biases associated with each network neuron in an effort to lower the cost function. For this study, many optimization methods, including stochastic gradient descent (SDG), RMSprop, Adadelta, Adam, and Nadam. However, Nadam have been taken into consideration as it performs best in all others. The classification accuracies of the models using these optimization algorithms have been presented in Figure 13(c).

4.4.4. Selection of Network weight initialization

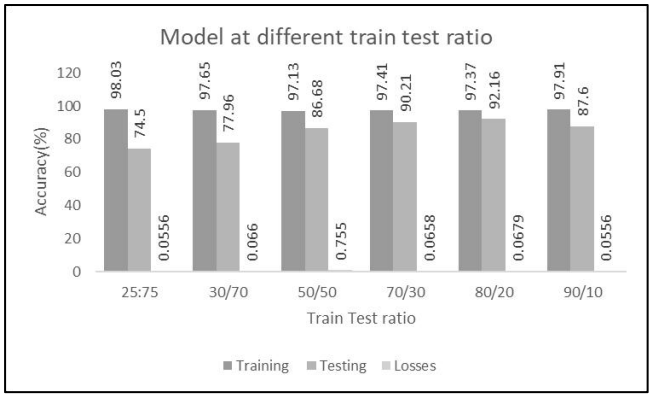
It is crucial to initialise network weights correctly to prevent gradients from inflating and disappearing. It may also further effect on how the network is trained. This study considers several weight initialization methods, including Uniform, Lecun Uniform, Normal, Zero, Glorot Normal, Glorot Uniform, He Normal, and He Uniform. The classification accuracies of the models using these weight initializations have been presented in Figure 13(d). Lecun uniform with a classification training accuracy of 97.43% and classification testing accuracy of 91.45% performs the best amongst all the weight initializers considered and hence chosen for further analysis.

4.4.5. Selection of the activation function

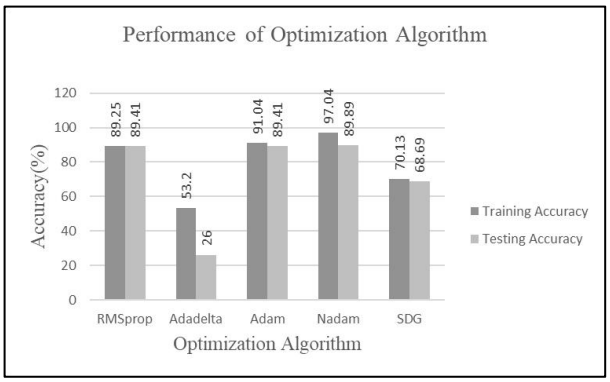
When creating any neural network, the activation function must be carefully chosen. How well the network model learns the training dataset is determined by the activation function of the hidden layer. The kind of output we can get depends on the activation function in the output layer. The output layer of the algorithm uses the Softmax activation function because this is a multiclass classification problem. It involves testing each of the available relevant functions to find the one that is most suited to the application before deciding on the activation function for hidden layers. For this aim, many activation functions like Softmax, Softplus, Softsign, RELU, Tanh, Sigmoid, Hard-sigmoid, and Linear activation function are taken into consideration. The classification accuracies of the models using these weight initializations have been presented in Figure 13(e). Relu activation function is most effective for this application, with a classification accuracy of 97.41%. Additionally, it should be mentioned that the Lecun uniform weight initialization method complements the Relu activation function the best.



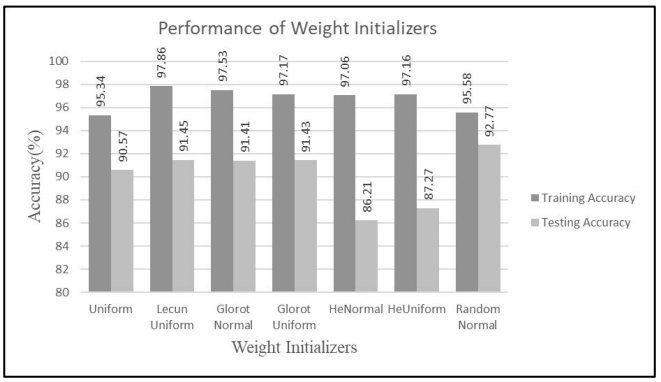
(a)



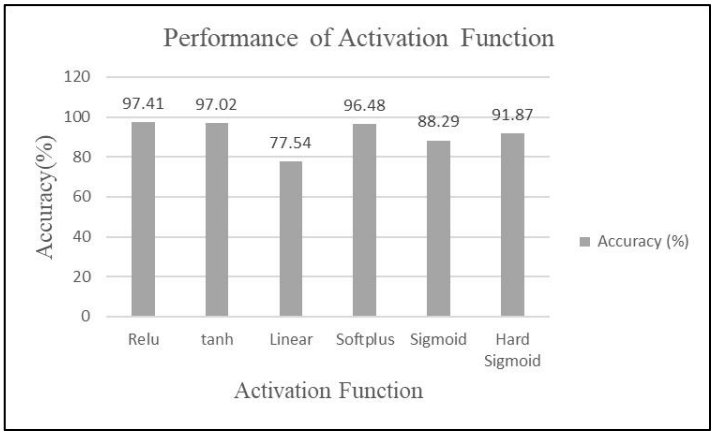
(b)



(c)



(d)



(e)

Figure 15: Performance while making the Neural Network (a) Different classification algorithm at their tuned states in which Neural Network performs best (b) Selection of train-test ratio in which 80:20 ration found best (c) Selection of optimization algorithm in which Nadam executes best (d) Selection of weight initializers in which Lecun Uniform shows best performance (e) Selection of activation function in which Relu found best.

Table 5*: *Configuration of the neural network used	
Hidden layers	6
No. of neurons in each hidden layer	55
Activation function used in each hidden layer	Relu
Activation function used in output layer	Softmax
Optimizer	Nadam
No. of epochs performed	130
Mini batch size	40
Weight initialisation type	Lecun Uniform
Signals used	All
Frequency considered	30 Hz – 60 Hz

4.5. Test of blockage severity level

At different levels of obstruction, the potency of features is also assessed (B1, B2, B3, B4 and B5). Testing for blockage occurs at a stepping size of 5 Hz and an operatinh

g frequency of 30 Hz to 60 Hz.

4.5.1. Blockage severity level with the single feature

The blockage severity levels are predicted under the single feature, which perform best in all other that is the standard deviation, mean, and entropy. We predicted the five-blockage level from B1 (Low severity level) to B5 (High severity level). Figure 14 displays the categorization accuracy provided by each feature at various levels of obstruction across the whole frequency range (30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz and 60 Hz) and with all the sensor data together. One can see that the classification accuracy does not follow a set pattern when only considering one characteristic. For enhanced reliability in results, employing combinations of these features is advisable.

4.5.2. Blockage severity level with the combination of features

Combinations of these traits can be employed to provide results that are more dependable. Figure 14 displays the fluctuation in classification accuracy with extent of obstruction for several feature combinations. There is a pattern formed and can be seen in the Figure 14. The pattern of increasing the accuracy of the severity level as the blockage level is elevated. Now like as in individual feature, there is no peaks and valleys found. This can be seen in classifier performance for all of the classes is higher than 90%, suggesting that the level of obstruction may be identified with better precision. We can also observe that the classifier performance for the same type of accelerometer, pressure, and motor current sensor data are very close, indicating that the test results can repeatable. The relationship between the categorization accuracy and blockage level can be seen to follow a pattern. We can see that most blockage levels can now be detected more accurately across with the combo of features. Also, only a small area around the actual blockage level is affected by inaccurate blockage level detection, which is a sign of qualitative growth. The degree of blockage influences how accurately the classification is made. When all the features are considered, the classifier will perform better.

4.5.3. Blockage severity level with the all of features

In this subsection, classification accuracy is elevated as the number of blockage severity level is increased, and is demonstrated in Figure 14. The investigation of the performance of features reveals that when used collectively, these features improve data comprehension and defect prediction.

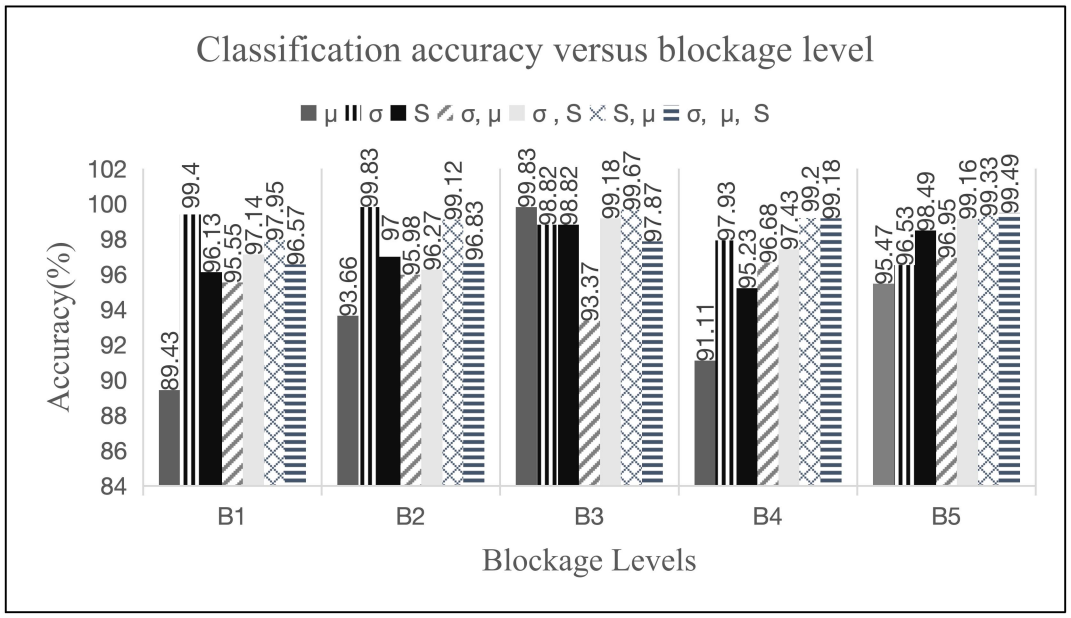


Figure 16: Performance of the classifier vs obstruction level using at five different obstruction levels

The classifier performance is high for greater blockage levels, which implies that one may forecast the severity of the blockage issue based on the classification accuracy. All of the above features would be combined in the remaining study. The confusion matrix of the model for all feature blockage is presented in Table 6. An C x C matrix called a confusion matrix is used to assess the effectiveness of a classification model, where C is the total number of target classes. The projected values of the machine learning model are compared to the actual target values in the matrix.

Table 6: Confusion matrix for all combined features at five different blockage classes

		Predicted Levels			
Actual Levels	**2369**	0	32	0	52
	0	**2365**	26	43	7
	0	25	**2246**	0	25
	9	2	9	**2286**	0
	0	5	0	7	**2285**

Table 6 shows the classification confusion matrix when all statistical features are considered by the classifier. We can see that the classifier's performance in estimating the degree of obstruction has increased.

As we observe the increasing severity of blockage, it becomes evident that combining all statistical features and all signals of dynamic pressure, vibration, and current yields the most accurate results. This study enables the prediction of the level of obstruction in the inlet pipe, facilitating preventive maintenance to avoid breakdowns. Different blockage levels (B1, B2, B3, B4, B5) exhibit prediction accuracies of 96.57%, 96.83%, 97.87%, 99.18%, and 99.49%, respectively. The near-perfect prediction of blockage levels in the inlet pipe, reaching close to 100%, is significant for preventing shutdowns. This research delineates the degree of blockage in the inlet pipe of the CP, with Class B1 exhibiting 96.57% classification accuracy. Notably, as blockage severity increases, classification accuracy also rises, culminating in 99.49% accuracy for Class B5, nearing 100%. Ultimately, this predictive capability serves as a valuable tool for proactive maintenance, ensuring uninterrupted operation of the system and minimizing downtime.

In this study, a deep learning model has been developed for efficient pump health diagnosis. The decision-making steps that went into selecting the input signals, the feature used for extraction, the optimization algorithm, network weight initialization, activation function, and the number of hidden layers, number of neurons used in each layer, and dropout regularization technique to be used for avoid overfitting in the model. These all contributed to the creation of the final tuned model. Pressure, vibration, and current motor signal are utilized to illustrate the interdependence between the obstruction and blockage severity. Figure 15 displays the flow chart for the entire process which refers to the entire procedure of data processing. Table 7 gives the comparison between this study and other literature papers which shows that comparing to other SOTA, this method is more accurate by varying 2.41% to 15.69%. In comparison to Perovic et al., whose minimum accuracy is 83.8%, this present work shows an improvement of 15.69%. Similarly, compared

to Wang et al., who achieved a maximum accuracy of 97.08%, this present work demonstrates an increase of 2.41%.

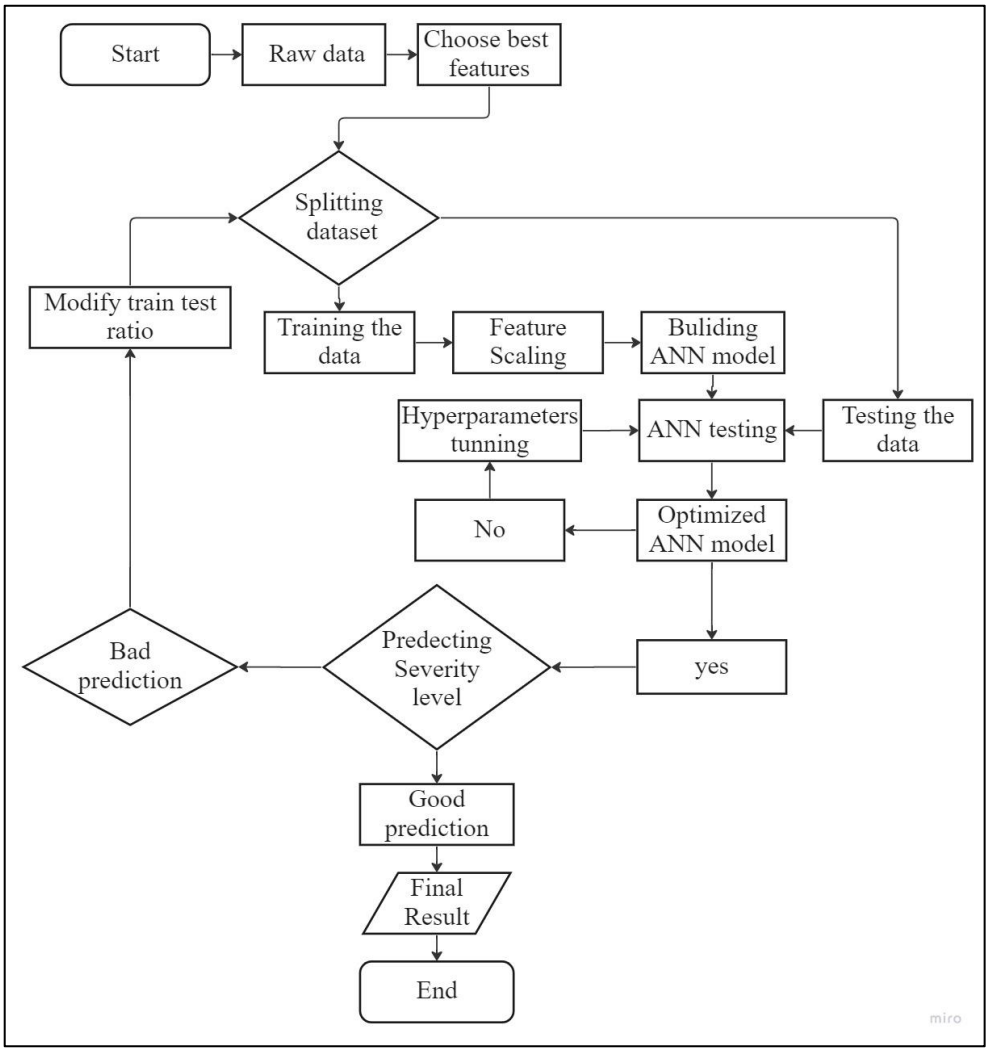


Figure 17: An overview of the data processing flowchart for the present study

Table 7: Comparison between this work and research studies

S.no.	Authors	Experimental Setup	Type of Flaws	Signals used	Domain	Features used	AI approach used	Results	Robustness & Generalization
1.	Panda et al. [26]	Lab experiments; accelerometer-based	Obstruction and cavitation	Vibration	Time	σ, μ, κ, χ, crest factor, S	SVM	93.4%	Effective for tested conditions; limited generalization
2.	Rapur and Tiwari [37]	Controlled tests; power spectrum data	Impeller break and sucked obstruction (Multi-fault conditions)	Vibration	Time	σ, μ, S	SVM	91.25%	Robust across speeds; generalizable to various faults
3.	Bordoloi and Tiwari [25]	Multiple fault conditions; varying speeds	Obstruction and cavitation	Vibration	Time	σ, κ, χ	SVM	94.8%	Effective for specific speeds; limited generalization
4.	Perovic et al. [3]	Two pumps; motor current spectral analysis	Impeller defect, cavitation, and obstruction	Motor line current	Frequency	Noise accumulation and current amplitude	Fuzzy system	83.8%	Good performance with tested pumps; needs adaptation
5.	Wang et al. [22]	CEEMD and vibration signals for fault diagnosis	Issues with the bearing and impeller	Vibration	Time	Sample Entropy	Empirical Mode Decomposition and random forest	97.08%	Effective for specific faults; computationally intensive
6.	Zouari et al. [6]	Industrial bench; real-time accelerometer data	Air injection, partial flow, and cavitation	Vibration	Time	13 set of features	Neural network and neuro fuzzy technique	96%	Adaptable; performance may vary with conditions
7.	Dewangan A. [31]	Pressure transducer;	Obstruction and cavitation	Pressure; Vibration;	Pressure	Time	κ, χ, σ, μ	Deep Learning	95.80%	High accuracy

		blockage simulation		Current				Algorithm		for tested conditions; future work needed
8.	Present work	Three pumps; multi-sensor data collection	Obstruction levels; Impeller cuts; Pitted cover plate	Vibration; Pressure; Current	Pressure, acceleration and current line motor	Time	μ, S, σ	Deep Learning Algorithm (ANN)	99.49%	High accuracy for blockage levels; effective across pumps

5. Conclusions

Successful deep learning-based fault diagnosis in the CPs using time domain multiple sensor data are presented. The progression of a pump blockage is believed to be gradual, closely associated with the potential occurrence of cavitation. Thus, the timely identification of blockages holds paramount significance. The impeller and pits on cover plate faults are induced artificially. Different levels of flaw classes (B0, B1, B2, B3, B4 and B5) are considered as class B0 is ignored because it is without any disturbance (fully flow). The obstruction severity of the CP can be identified by the successful application of pressure signatures from pressure transducers, vibration signatures from accelerometers, and current signatures from motor currents. Accuracy of the fault prediction is enhanced when data is collected from several sensors. The combination of various sensors and features yielded a better improved accuracy. The mean, standard deviation, entropy, mode, kurtosis, and skewness are taken from the time-domain signals representing to all the signals. It has been discovered that the combination of these features provides accurate fault severity identification. The mean was the best individually feature used to extract information from these data in time domain and a combination of standard deviation, mean, and entropy is found best for the all-combined statistical features set. Development of the artificial neural network model has been done successfully. Standardization was then done to scale the different input features as a part of feature scaling. The hyperparameter tuning was done to improve the classification accuracy of the model. The number of hidden layers found are six and number of neurons found in each layer is 55 and iteration performed is 130. ReLU is used in the hidden layer while the Softmax function is applied to the output layer. The ultimate optimized ratio for training and testing is 80:20. The dropout regularisation method is used to check for the presence of overfitting and rule it out with a probability of 0.5. The tuned model predicted the severity of suction blockages with an accuracy of 99.49%. The prediction accuracies for blockage levels B1, B2, B3, B4, and B5 are 96.57%, 96.83%, 97.87%, 99.18%, and 99.49%, respectively, nearing near-perfect accuracy close to 100%. This study effectively identifies blockage severity in the CP's inlet pipe, showing increasing accuracy with higher severity levels.

A method presented to classify inlet pipe blockages into five levels, helping us prioritize maintenance efforts effectively. By tackling blockages early, we reduce downtime, cut emergency repair expenses, and boost the overall reliability of our systems. Using data-driven insights lets us make informed decisions, optimizing our maintenance schedules and ensuring our equipment stays healthy in the long run. This approach is key to improving efficiency and keeping our systems running smoothly day in and day out. It has also been observed that the more severe the blockage, the more classifiable it becomes. The obtained accuracy was adequate to foretell the degree of obstruction, and the results also show that collecting data from multiple sources is always advantageous. The classification accuracy and confusion matrix inferred to the accelerometer sensor, motor current sensor, and pressure sensor data indicated that all these signals can be used to determine severity level of the obstruction level.

The present study leveraged time domain data. In forthcoming research, the integration of frequency domain data for blockage level detection could be explored similarly to time domain data, thus allowing for a comparative analysis of outcomes. The same data classification technique may be utilized to examine various other traits derived from same sensor data for monitoring of the obstruction. Also, the same study can be conducted using a CNN model.

References

[1] Tiwari, R., 2017, Rotor Systems: Analysis and Identification, CRC Press: Boca Raton, USA.
[2] Tabar M., Majidi S., Poursharifi Z., 2011 Investigation of recirculation effects on the formation of vapor bubbles in centrifugal pump blades, World Academy of Science, Engineering and Technology, vol. 73, pp. 521–526.
[3] Perovic, S., Unsworth, P.J., Higham E.H., 2001 Fuzzy Logic System to Detect Pump Faults from Motor Current Spectra, Conference Record - 36th IAS Annual Meeting, Chicago, USA, September 30- October 4 2001, vol. 1, pp. 274–280, https://doi.org/10.1109/IAS.2001.955423.
[4] Chudina M., 2003, Noise as an indicator of cavitation in a centrifugal pump, Askerceva 6, 1000 Ljubljana, Slovenia, vol. 49(4), pp. 463–474, https://doi.org/10.1134/1.1591303.
[5] Dister, C.J., On-line health assessment of integrated pumps IEEE Aerospace Conference Proceedings, Big Sky, Montana, USA, March 8-15,2003, vol. 7, pp. 3289–3294, https://doi.org /10.1109/AERO.2003.1234172.
[6] Zouari R., Sieg-Zieba S., Sidahmed M., Fault Detection system for centrifugal pumps using neural networks and neuro-fuzzy techniques,13 october 2004.
[7] Samanta B., Al-Balushi K.R., Al-Araimi S.A., Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection, Engineering Applications of Artificial Intelligence, Muscat, Sultanate of Oman, September 18 2003, vol. 16: pp. 657–665, https://doi.org/10.1016/j.engappai.2003.09.006.
[8] Singh M. and Shaik A.G., Faulty bearing detection, classification and location in a three-phase induction motor based on Stockwell transform and support vector machine, Measurement: Journal of the International Measurement Confederation, vol. 131, pp. 524-533, Jodhpur, Rajasthan, India, 7 September 2018, https://doi.org/10.1016/j.measurement.2018.09.013.
[9] Barakat M., Badaoui M. El and Guillet F., Hard competitive growing neural network for the diagnosis of small bearing faults, Mechanical Systems and Signal Processing, vol. 37, pp. 276-292, Roanne, France, 14 February 2013, http://dx.doi.org/10.1016/j.ymssp.2012.11.002.
[10] Zhao R., Yan R., Chen Z., Mao K., Wang P. and Gao R.X., Deep learning and its applications to machine health monitoring, Mechanical Systems and Signal Processing, *vol.* 115, pp. 213-237, Case Western Reserve University, United States, 14 June 2018, https://doi.org/10.1016/j.ymssp.2018.05.050.
[11] Kane P. V., and Andhare A. B., Application of psychoacoustics for gear fault diagnosis using artificial neural network, *Journal of Low Frequency Noise, Vibration and Active Control,* vol. 35(3), pp. 207–220, Nagpur, India, 2016, http://dx.doi.org/10.1177/0263092316660915.
[12] Wong M.L.D., Jack L.B., Nandi A.K., Modified self-organising map for automated novelty detection applied to vibration signal monitoring, Mechanical Systems and Signal Process, Brownlow Hill, Liverpool, United Kingdom, March 10 2005, vol.20, pp. 593–610, https://doi.org/10.1016/j.ymssp.2005.01.008.
[13] Widodo A., Yang B., 2007 Support vector machine in machine condition monitoring and fault diagnosis, Mechanical Systems and Signal Process, Nam-gu, Busan, Republic of Korea,December 19 2006, vol. 21, pp.2560–2574, https://doi.org/10.1016/j.ymssp.2006.12.007.
[14] Rajakarunakaran, S., Venkumar, P., Devaraj, D., Rao, K. S. P., Artificial neural network approach for fault detection in rotary system, Applied Soft Computing, Tamilnadu, India, June 10 2007 , vol. 8(1), pp. 740–748, https://doi.org/10.1016/j.asoc.2007.06.002.
[15] Sakthivel N. R., Sugumaran, Babudevasenapati S., 2010 Vibration based fault diagnosis of monoblock centrifugal pump using decision tree, Expert Systems with Applications, Chennai, India, vol. 37(6), pp. 4040–4049, https://doi.org/10.1016/j.eswa.2009.10.002.
[16] Nasiri M.R., Mahjoob M.J., Alizadeh H., 2011 Vibration signature analysis for detecting cavitation in centrifugal pumps using neural networks, IEEE International Conference on Mechatronics, ICM 2011 - Proceedings, Tehran, Iran, April 13-15, 2011, pp. 632–635, https://doi.org/10.1109/ICMECH.2011.5971192.

[17] Azadeh A., Saberi M., Kazem A., Ebrahimipour V., Nourmohammadzadeh, A., Saberi Z, 2013 A flexible algorithm for fault diagnosis in a centrifugal pump with corrupted data and noise based on ANN and support vector machine with hyper-parameters optimization, Applied Soft Computing, Tehran, Iran, 27 June 2012, vol.13(3), pp. 1478–1485, https://doi.org/10.1016/j.asoc.2012.06.020.

[18] Abdulkarem W., Amuthakkannan R., Al-Raheem K., Centrifugal Pump Impeller Crack Detection Using Vibration Analysis International Institute of Engineers, 2nd International Conference on Research in Science, Engineering and Technology, March 21-22 2014, Dubai ,UAE, https://doi.org/10.15242/iie.e0314606.

[19] Rapur S., Tiwari R., Severity Assessment and Classification of Blockage of Centrifugal Pumps in Frequency Domain of Vibration Data Using Support Vector Machine Algorithms Severity assessment and classification of blockage of centrifugal pumps in frequency domain of vibration, Guwahati, India, September 2016.

[20] Lu J., Yuan S., Luo Y., Yuan J., Zhou B., Sun H.,2014 Numerical and experimental investigation on the development of cavitation in a centrifugal pump, Proceedings of the Institution of Mechanical Engineers, Journal of Process Mechanical Engineering, China, November 12 2014, vol. 230(3), pp. 171–182. https://doi.org/10.1177/0954408914557877.

[21] Zhao W., Wang Z., Lu C., Ma J., Li L., Fault diagnosis for centrifugal pumps using deep learning and softmax regression, Proceedings of the 12th World Congress on Intelligent Control and Automation (WCICA), Guilin, China, June 12-15, 2016, pp.165–169, https://doi.org/10.1109/WCICA.2016.7578673.

[22] Wang, Y., Lu, C., Liu, H., and Wang Y., Fault diagnosis for centrifugal pumps based on complementary ensemble empirical mode decomposition, sample entropy and random forest, Proceedings of the World Congress on Intelligent Control and Automation (WCICA), June 12-15 2016, Guilin, China. doi:10.1109/wcica.2016.7578401, https://doi.org/10.1109/WCICA.2016.7578401.

[23] Azizi, R., Attaran, B., Hajnayeb, A., Ghanbarzadeh, A., & Changizian, M., Improving accuracy of cavitation severity detection in centrifugal pumps using a hybrid feature selection technique, Measurement: Journal of the International Measurement Confederation, vol. 108, pp. 9–17, Ahvaz, Iran, May 9 2017, https://doi.org/10.1016/j.measurement.2017.05.020.

[24] Rapur J.S., Tiwari R., Experimental time-domain vibration- based fault diagnosis of centrifugal pumps using support vector machine, Measurement: Journal of the International Measurement Confederation, Guwahati, India, July 8 2017, https://doi.org/10.1115/1.4035440.

[25] Bordoloi D.J., Tiwari R., Identification of suction flow blockages and casing cavitations in centrifugal pumps by optimal support vector machine techniques, Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 39, pp. 2957–2968, Guwahati, India, January 9 2017, https://doi.org/10.1007/s40430-017-0714-z.

[26] Panda A.K., Rapur J.S., Tiwari R., Prediction of flow blockages and impending cavitation in centrifugal pumps using Support Vector Machine (SVM) algorithms based on vibration measurements, Measurement: Journal of the International Measurement Confederation vol. 130, pp. 44–56, Guwahati, India, July 31 2018, https://doi.org/10.1016/j.measurement.2018.07.092.

[27] Rapur J.S., Tiwari R., On-line Time Domain Vibration and Current Signals Based Multi-fault Diagnosis of Centrifugal Pumps Using Support Vector Machines, Journal of Nondestructive Evaluation 38, Guwahati, India, November 13 2018, https://doi.org/10.1007/s10921-018-0544-7.

[28] Dutta N., Umashanka S., Shankar V.K.A., Padmanaban S., Leonowicz, Z., Wheeler P., Centrifugal Pump Cavitation Detection Using Machine Learning Algorithm Technique, IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial

and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Vellore, India, Esbjerg, Denmark, Poland, United Kingdom, 2018. https://doi.org/10.1109/eeeic.2018.8494594.

[29] Alabied S., Daraz A., Rabeyee K., Alqatawneh I., Gu F., Ball A.D., Motor Current Signal Analysis Based on Machine Learning for Centrifugal Pump Fault Diagnosis, 25th International Conference on Automation and Computing (ICAC), Proceedings of the 25th International Conference on Automation & Computing, Lancaster United Kingdom, 5-7 September 2019 https://doi.org/10.23919/iconac.2019.8895057.

[30] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K., Applications of machine learning to machine fault diagnosis: A review and roadmap, Mechanical Systems and Signal Processing, vol. 138, United Kingdom, June 9 2020, https://doi.org/10.1016/j.ymssp.2019.106587.

[31] Dewangan A., Bordoloi D.J., Tiwari R., Blockage and cavitation detection in centrifugal pumps from dynamic pressure signal using deep learning algorithm, Journal of the International Measurement Confederation, vol. 173, Guwahati, India, October 26 2020, https://doi.org/ 10.1016/j.measurement.2020.108676.

[32] Sha Y., Faber J., Gou S., Liu B., Li W., Schramm S., Stoecker H., Steckenreiter T., Vnucec D., Wetzstein D., Widl A., An acoustic signal cavitation detection framework based on XGBoost with adaptive selection feature engineering, Journal of the International Measurement Confederation, vol. 192, Frankfurt Institute for Advanced Studies, Germany, February 25 2022. https://doi.org/10.1016/j.measurement.2022.110897

[33] Orrù P. F., Zoccheddu A., Sassu L., Mattia C., Cozza R., Arena S., Machine Learning Approach Using MLP and SVM Algorithms for the Fault Prediction of a Centrifugal Pump in the Oil and Gas Industry, Sustainability (Switzerland), vol. 12(11), pp. 4776, Cagliari, Italy, June 23 2022. https://doi.org/10.3390/su12114776.

[34] Chen Y, Yuan J, Luo Y, Zhang W., Fault Prediction of Centrifugal Pump Based on Improved KNN, Shock and Vibration 2021, Zhenjiang, China, October 4 2021, https://doi.org/10.1155/2021/7306131.

[35] Sunal C.E., Dyo V., Velisavljevic V. Review of Machine Learning Based Fault Detection for Centrifugal Pump Induction Motors IEEE Access, vol. 10, pp. 71344–71355, Luton, United Kingdom, June 23 2022, https://doi.org/10.1109/ACCESS.2022.3187718.

[36] Kumar D., Dewangan A., Tiwari R., Bordoloi D.J., Identification of inlet pipe blockage level in centrifugal pump over a range of speeds by deep learning algorithm using multi-source data, Measurement: Journal of the International Measurement Confederation. vol.186, Guwahati, Assam, India, December 2021, https://doi.org/ 10.1016/j.measurement.2021.110146.

[37] Rapur, J. S., & Tiwari, R., Automation of multi-fault diagnosing of centrifugal pumps using multi-class support vector machine with vibration and motor current signals in frequency domain, Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 40(6), Guwahati, Assam, India, April 19 2018. https://doi.org/10.1007/s40430-018-1202-9.

[38] Rapur, J. S., Dewangan A., Tiwari, R., Bordoloi D.J., Review: Measurement-Based Monitoring and Fault Identification in Centrifugal Pumps, Journal of Harbin Institute of Technology (New series), CNKI, ISSN 1005-9113, CN 23-1378/T.

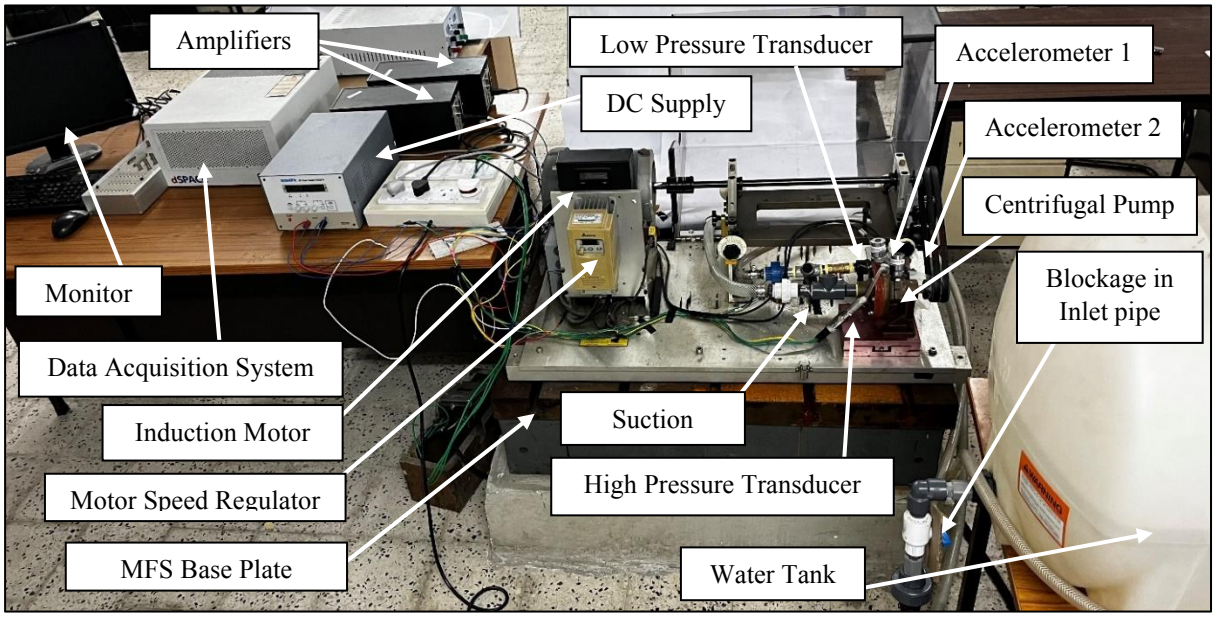
[39] Zaman W., Ahmad Z., Siddique M.F., Ullah N., Kim J.M., Centrifugal Pump Fault Diagnosis Based on a Novel SobelEdge Scalogram and CNN, Sensors, University of Ulsan, Republic of Korea, June 01 2023, https://doi.org/10.3390/ s23115255.

[40] Sunal C.E., Velisavljevic V., Dyo V., Newton B., Newton J., Centrifugal Pump Fault Detection with Convolutional Neural Network Transfer Learning, Sensors, University of Bedfordshire, Luton LU, 11 April 2024, https://doi.org/10.3390/ s24082442.

[41] Indriawati K., Yugoputra G.F., Habibah N.N., Yudhanto R., Artificial Neural Network-Based Fault Detection System with Residual Analysis Approach on Centrifugal Pump: A Case Study, Sepuluh Nopember Institute of Technology, Indonesia, Vol. 20, April 7 2023, https://doi.org/10.15282/ijame.20.1.2023.10.0795.

List of Figures

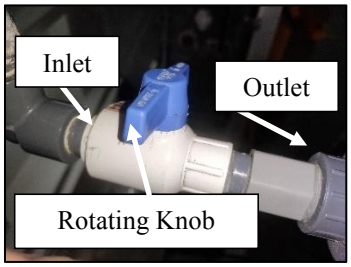
List of Table



(a)

(b)

Figure 1: (a) Experimental set-up (b) A close look



(a) (b)

Figure 2: Mechanical modulating valves (a) Rotating knob (b) Markings on the knob

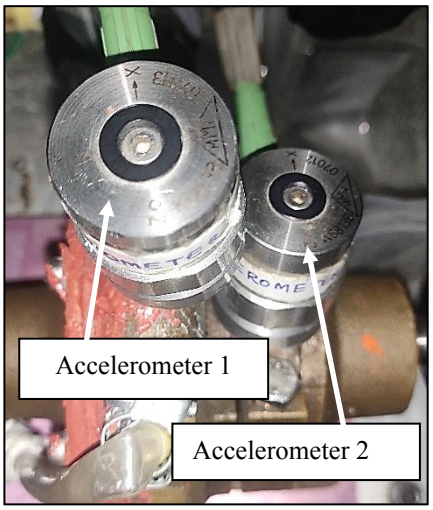


Figure 3: Sensor Placement with Tri-axial Accelerometers on Pump Casing and Bearing Housing.

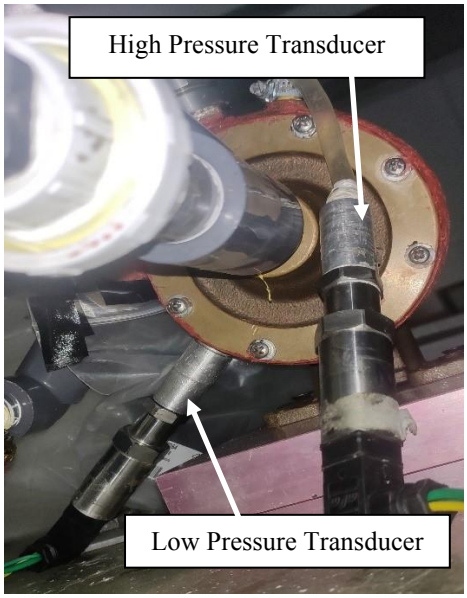


Figure 4: High- and low-pressure transducers mounted on the pump.

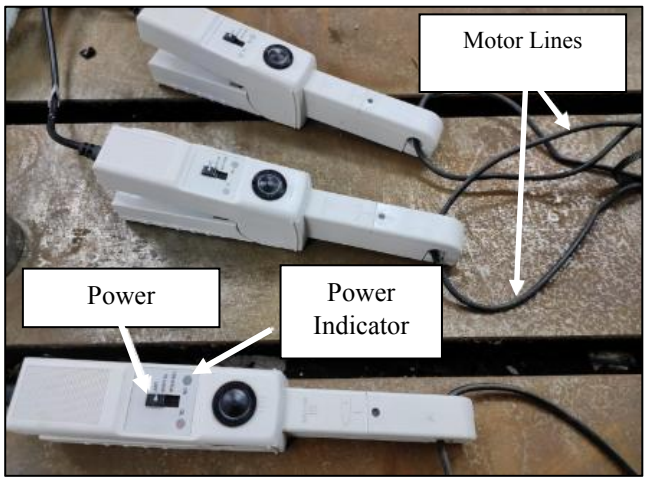


Figure 5: Current wires passing through Current probes

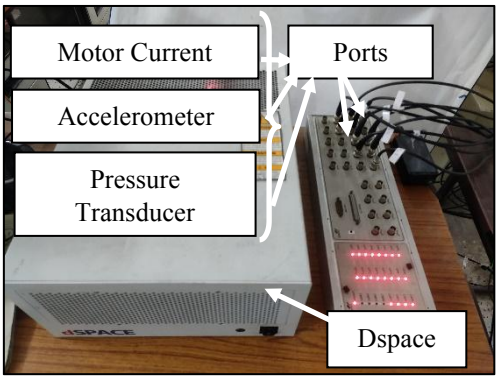


Figure 6: Dspace for collecting data

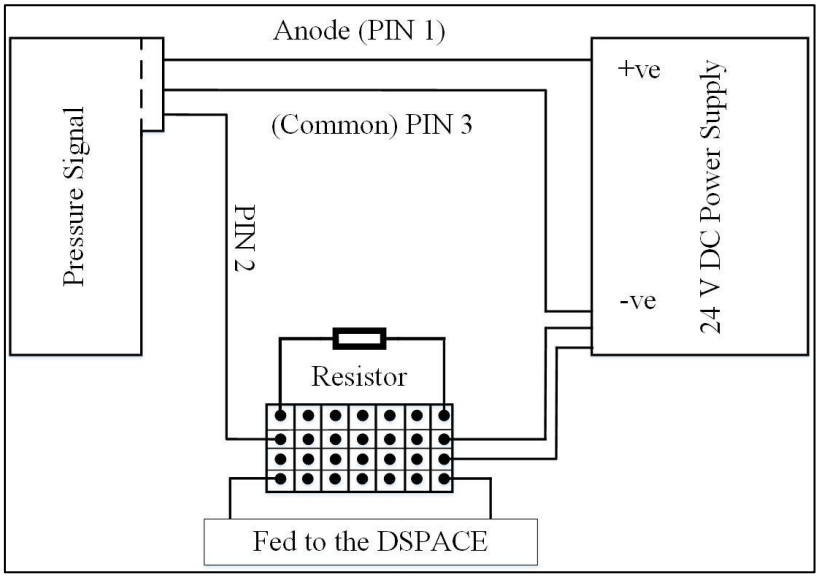


Figure 7: Three wiring setup to convert pressure transducer signals (Ampere form) to voltage form (measurable form by DAQ)

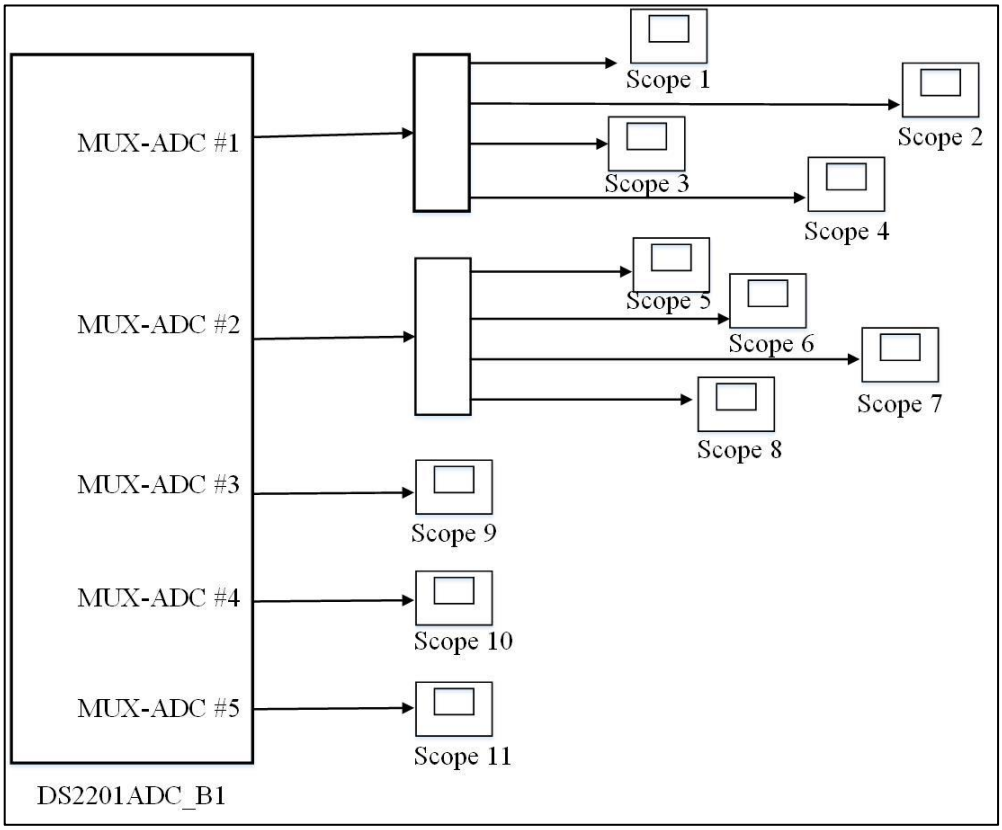


Figure 8: Simulink Model for display the sensors data signal in the monitor screen

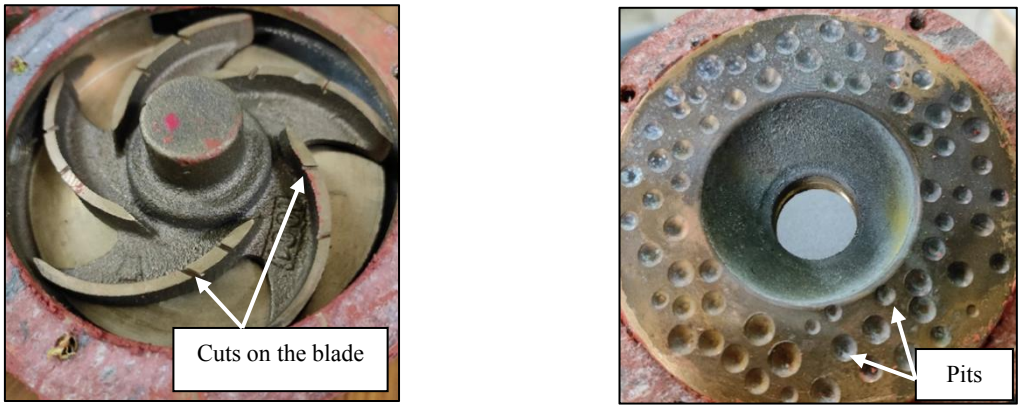
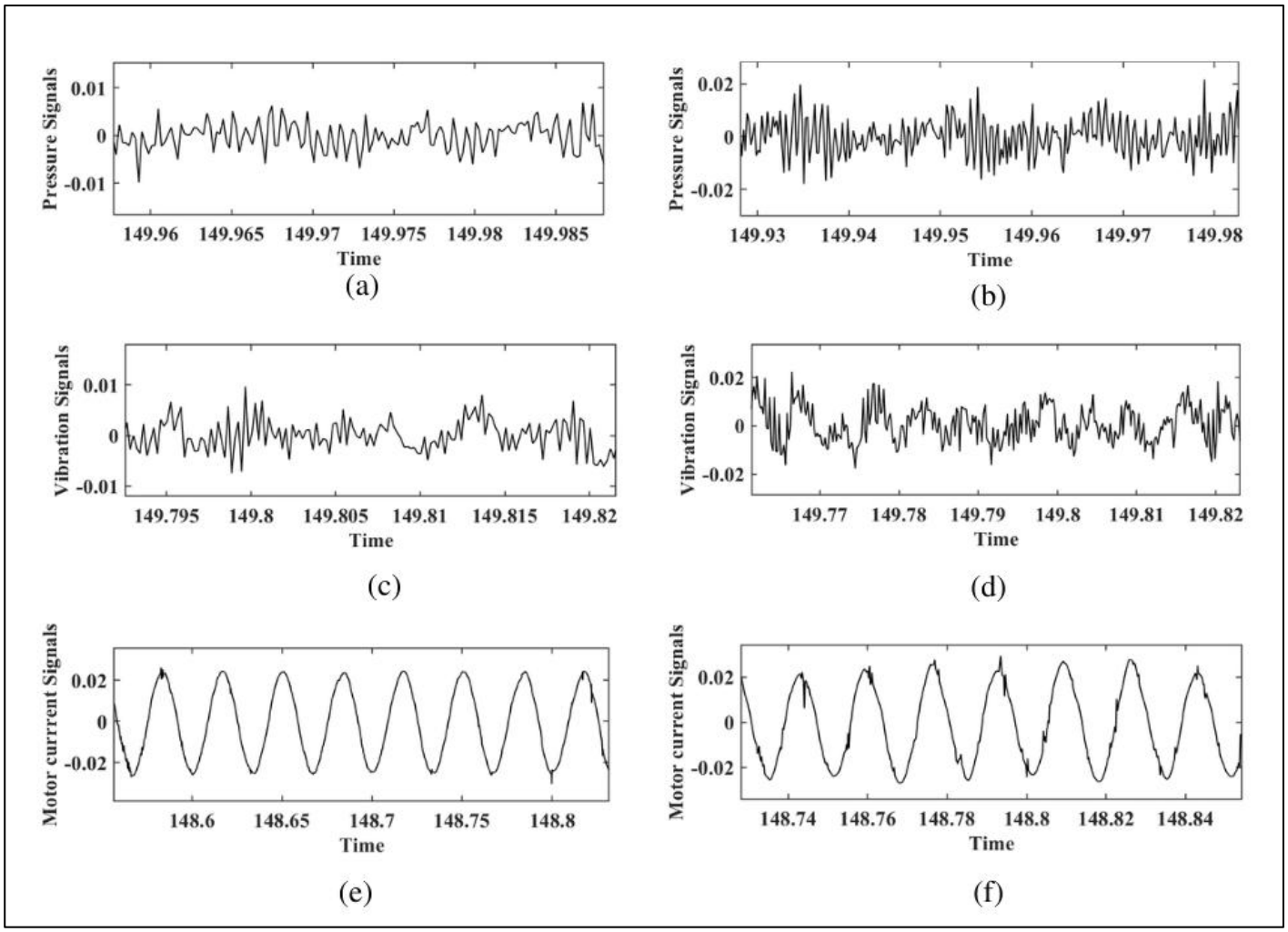


Figure 9: Impeller fault (left) and Cover plate fault (right)



Units: Time (in seconds) and Pressure, Vibration and Motor current signals (in volts)

Figure 10: Pressure, vibration, and motor current signals of impeller fault pump without obstruction level and final obstruction level (B5) (a) Impeller fault pump pressure signals without an obstruction (b) Impeller fault pump pressure signals with an obstruction of level B5 (c) *Impeller fault pump vibration signal without an obstruction* (d) *Impeller fault pump vibration*

signal with an obstruction level of B5 (e) *Impeller fault pump current signals without an obstruction* (f) *Impeller fault pump current signals with an obstruction level of B5*

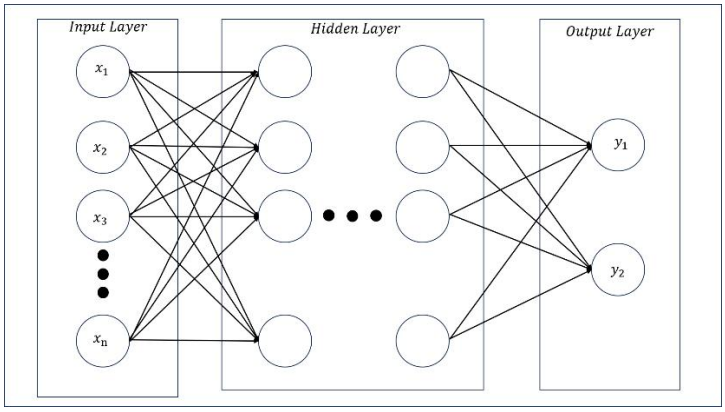


Figure 11: Structure and Function of Neural Networks

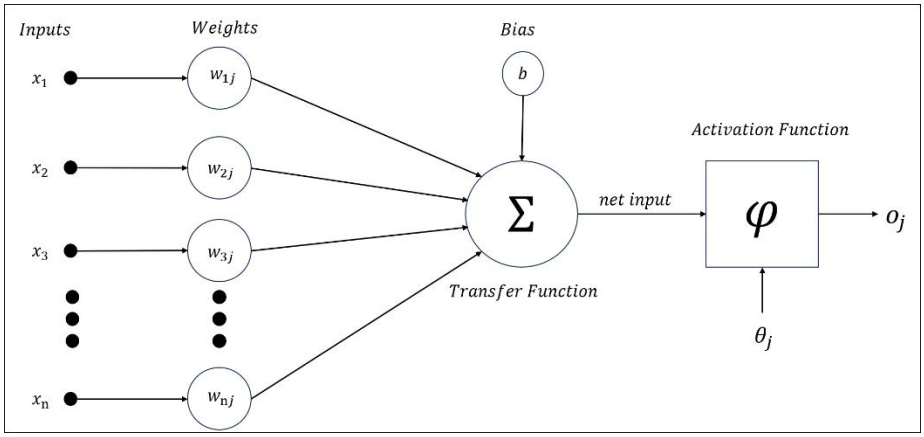


Figure 12: Neuron in a neural network

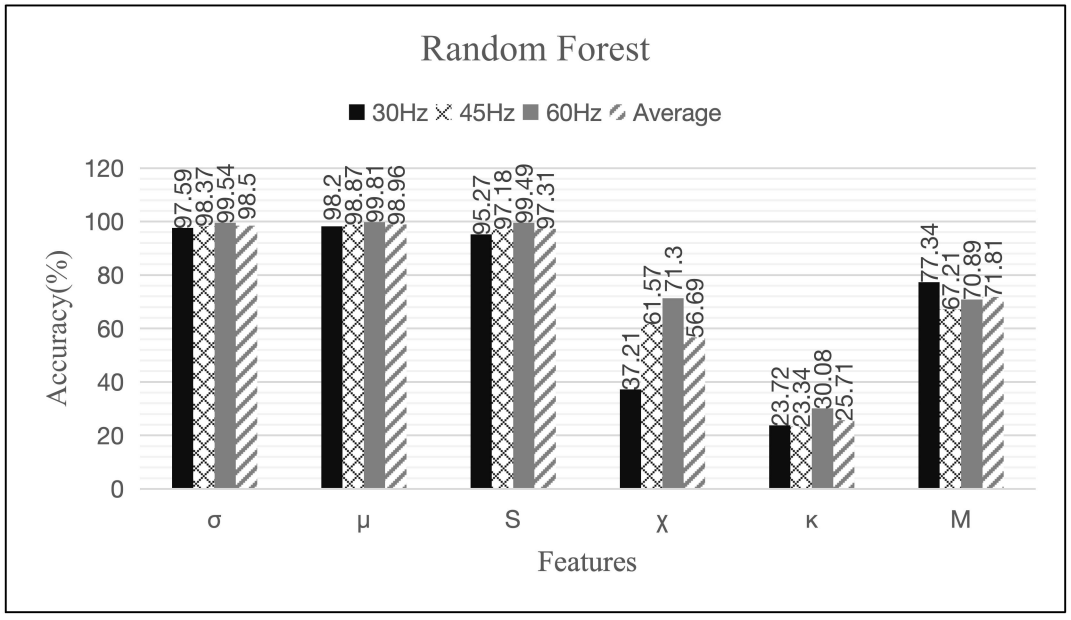


Figure13: Random Forest algorithm used to identify the performance for different statistical features at three different frequencies and their average

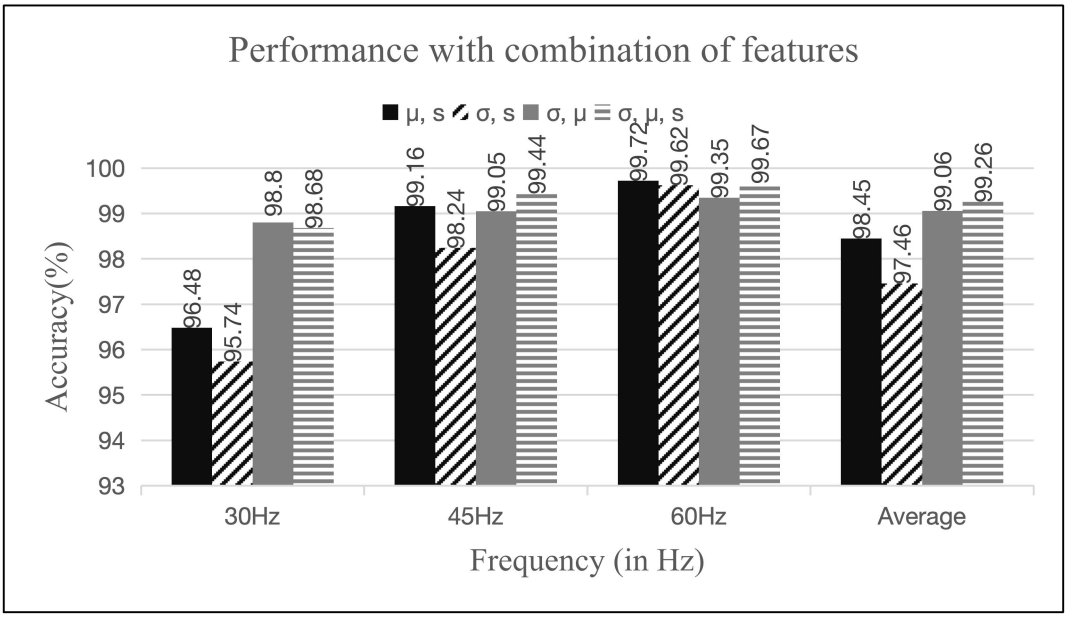
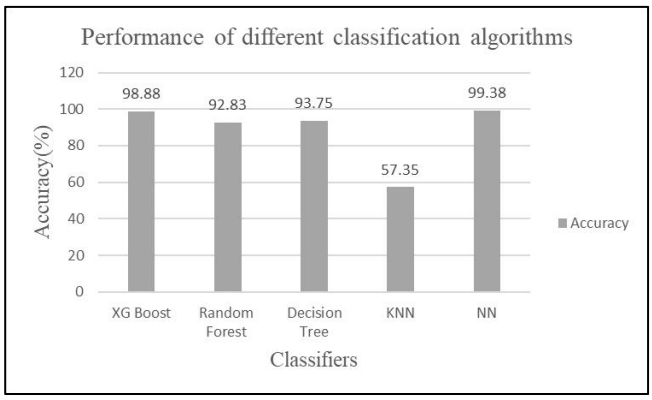
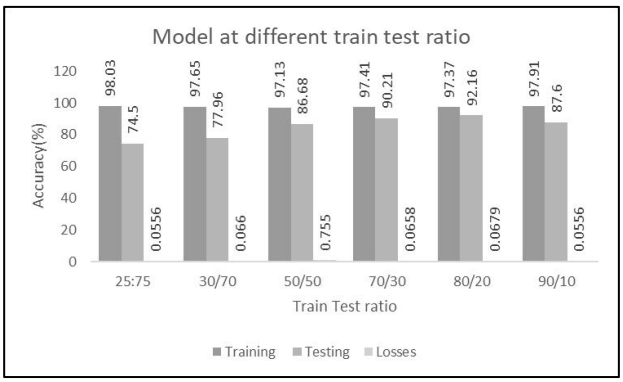


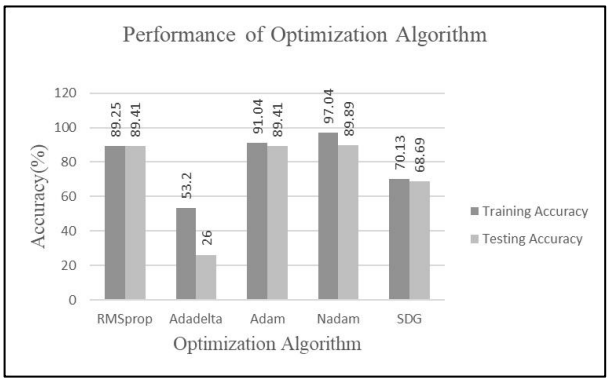
Figure 14: Classifier performance versus Frequency with combination of the features at three different frequencies



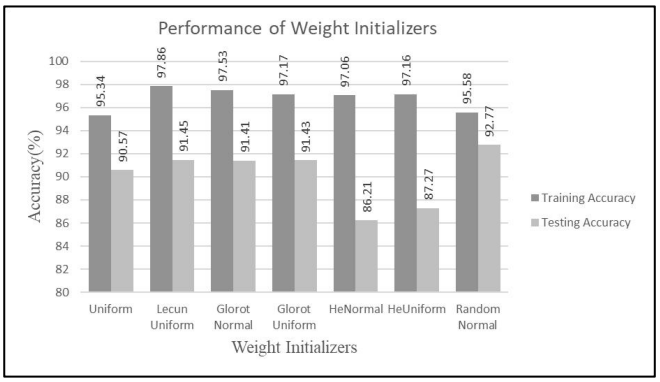
(a)



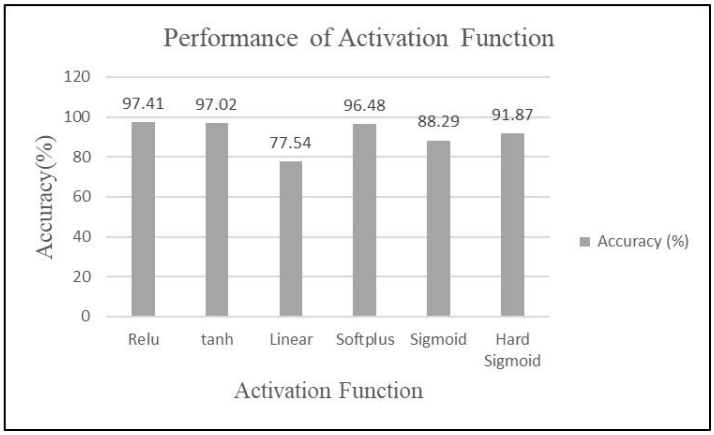
(b)



(c)



(d)



(e)

Figure 15: Performance while making the Neural Network (a) Different classification algorithm at their tuned states in which Neural Network performs best (b) Selection of train-test ratio in which 80:20 ration found best (c) Selection of optimization algorithm in which Nadam executes best (d) Selection of weight initializers in which Lecun Uniform shows best performance (e) election of activation function in which Relu found best

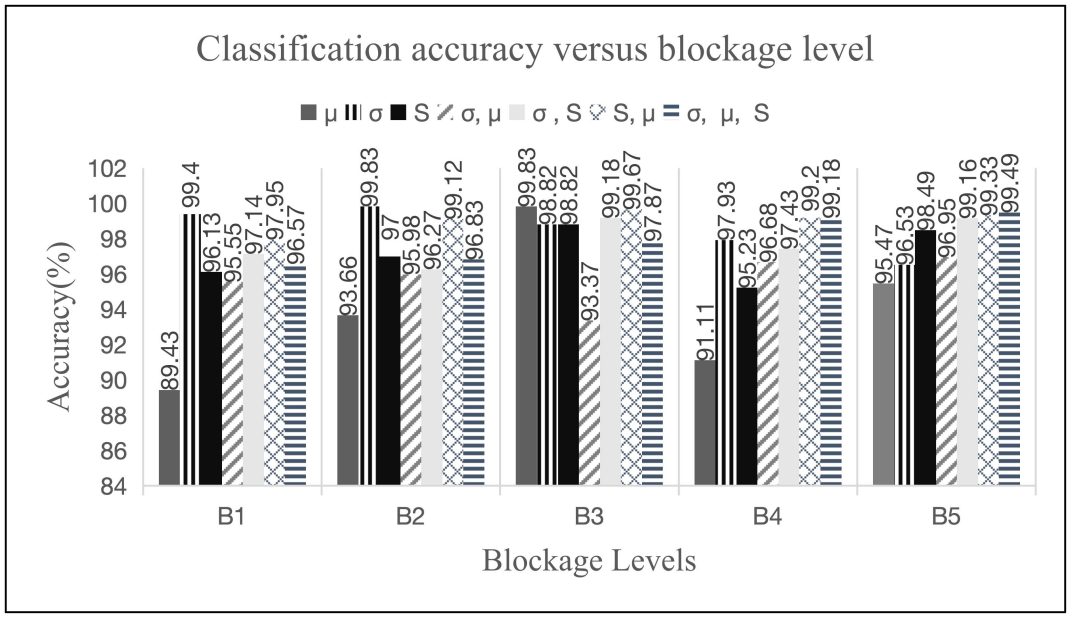


Figure 16: Performance of the classifier vs obstruction level using at five different obstruction levels

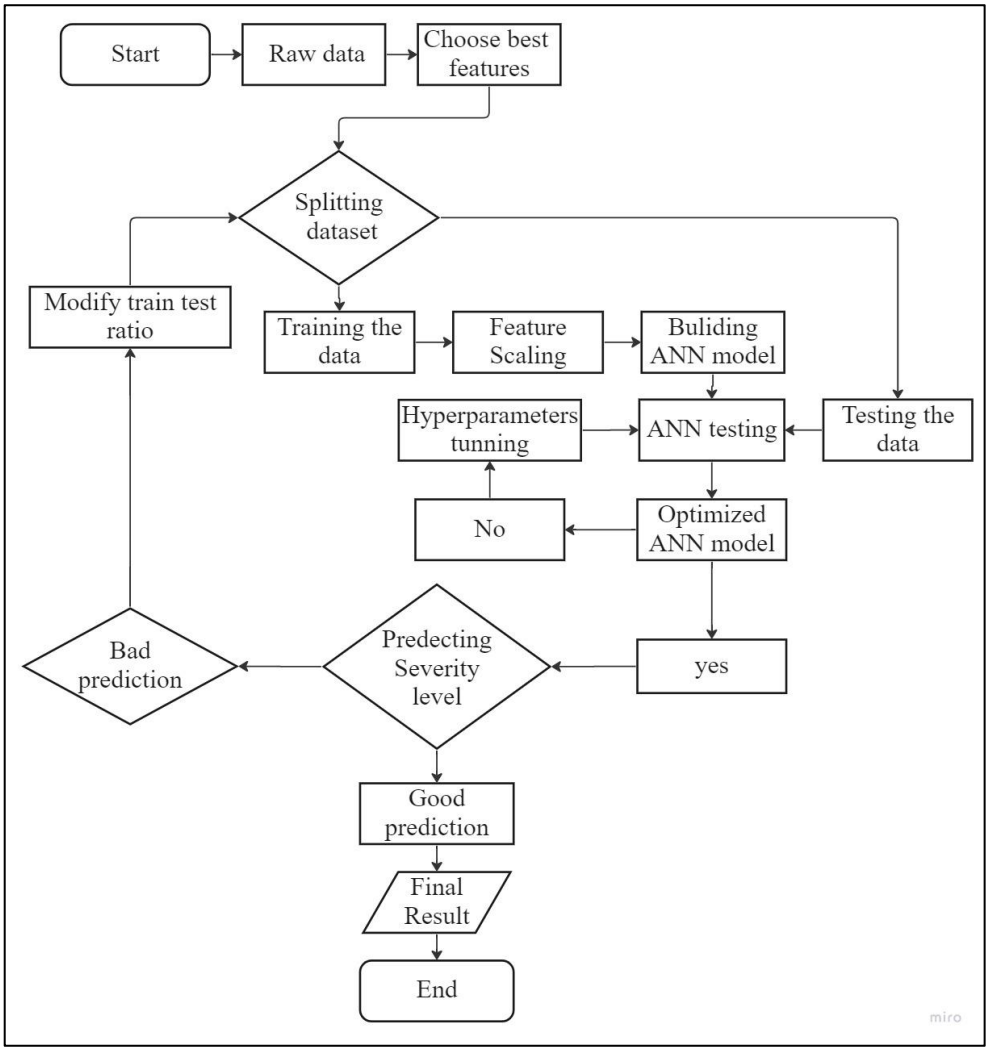


Table 1: Abbreviations	
σ	*Standard Deviation*
μ	*Mean*
κ	*Kurtosis*
χ	*Skewness*
x	*Input to the neural network*
w	Weight vector
b	Bias vector
$\hat{y}$	Predicted class
z	Loss function
T	Transpose
n	Total number of data points in a collection
x_i	Amplitude of each data point
CP	Centrifugal Pump
HP	Healthy Pump
IF	Impeller fault
PF	Cover plate fault

Figure 17: An overview of the data processing flowchart for the present study

RF	Random Forest
SVM	Support Vector Machine
DT	Decision Tree
KNN	K-Nearest Neighbours
XGB	XGBoost
ANN	Artificial Neural Network
SDG	Stochastic Gradient Descent

Table 2: Data Acquisition System outlines	
Blockage levels	B0 (Full Flow, no obstruction), B1 (1/6 obstruction), B2 (1/3 obstruction), B3 (1/2 obstruction), B4 (2/3 obstruction), B5 (5/6 obstruction)
Frequency	30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz and 60 Hz
Quantity of each case in every pump condition	6 (Blockage levels) × 7(Frequency) = 42 cases
No of measurement sets for each combination of blockage level and frequency	150 sets
Data collection time for each blockage level and frequency combination	150 sec
Timeframe for a single set collection	1 sec
Sample captured in one second	5000

Table 3: Description of Figure 10 Pressure, vibration, and motor current signals of impeller fault pump without obstruction level and final obstruction level (B5)

Figure No.	**Signal**	**Blockage Levels**
10(a)	Pressure	B0
10(b)	Pressure	B5
10(c)	Vibration	B0
10(d)	Vibration	B5
10(e)	Current	B0
10(f)	Current	B5

Table 4: Fault classification sets description	
Fault Set	**Classes (labels)**

1	Class 1(B0): HP0, PC0, IF0 Class 2(B1): HP1, PC1, IF1 Class 3(B2): HP2, PC2, IF2 Class 4(B3): HP3, PC3, IF3 Class 5(B4): HP4, PC4, IF4 Class 6(B5): HP5, PC5, IF5

Table 5*: *Configuration of the neural network used	
Hidden layers	6
No. of neurons in each hidden layer	55
Activation function used in each hidden layer	Relu
Activation function used in output layer	Softmax
Optimizer	Nadam
No. of epochs performed	130
Mini batch size	40
Weight initialisation type	Lecun Uniform
Signals used	All
Frequency considered	30 Hz – 60 Hz

Table 6: Confusion matrix for all combined features at five different blockage classes

	Predicted Levels				
Actual Levels	**2369**	0	32	0	52
	0	**2365**	26	43	7
	0	25	**2246**	0	25
	9	2	9	**2286**	0
	0	5	0	7	**2285**

Table 7: Comparison between this work and research studies

S.no.	Authors	Experimental Setup	Type of Flaws	Signals used	Domain	Features used	AI approach used	Results	Robustness & Generalization
1.	Panda et al. [26]	Lab experiments; accelerometer-based	Obstruction and cavitation	Vibration	Time	σ, μ, κ, χ, crest factor, S	SVM	93.4%	Effective for tested conditions; limited generalization
2.	Rapur and Tiwari [37]	Controlled tests; power spectrum data	Impeller break and sucked obstruction (Multi-fault conditions)	Vibration	Time	σ, μ, S	SVM	91.25%	Robust across speeds; generalizable to various faults
3.	Bordoloi and Tiwari [25]	Multiple fault conditions; varying speeds	Obstruction and cavitation	Vibration	Time	σ, κ, χ	SVM	94.8%	Effective for specific speeds
4.	Perovic et al. [3]	Two pumps; motor current spectral analysis	Impeller defect, cavitation, and obstruction	Motor line current	Frequency	Noise accumulation and current amplitude	Fuzzy system	83.8%	Good performance with tested pumps; needs adaptation
5.	Wang et al. [22]	CEEMD and vibration signals for fault diagnosis	Issues with the bearing and impeller	Vibration	Time	Sample Entropy	Empirical Mode Decomposition and random forest	97.08%	Effective for specific faults; computationally intensive
6.	Zouari et al. [6]	Industrial bench; real-time accelerometer data	Air injection, partial flow, and cavitation	Vibration	Time	13 set of features	Neural network and neuro fuzzy technique	96%	Adaptable; performance may vary with conditions

7.	Dewangan A. [31]	Pressure transducer; blockage simulation	Obstruction and cavitation	Pressure; Vibration; Current	Pressure	Time	κ, χ, σ, μ	Deep Learning Algorithm	95.80%	High accuracy for tested conditions; future work needed
8.	Present work	Three pumps; multi-sensor data collection	Obstruction levels; Impeller cuts; Pitted cover plate	Vibration; Pressure; Current	Pressure, acceleration and current line motor	Time	μ, S, σ	Deep Learning Algorithm (ANN)	99.49%	High accuracy for blockage levels; effective across pumps