**DIAGNOSING THE HEALTHY OR DEFECTIVE STATE OF A CENTRIFUGAL PUMP BASED ON VIBRATION, PRESSURE AND CURRENT DATA USING A DEEP LEARNING ALGORITHM**

*Shivam Gautam, Rajiv Tiwari [[1]](#footnote-1)*

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

**Abstract**

*Throughout operating condition, centrifugal pumps may suffer due to operationally developed faults, likely to result in a disruption in long - term operation. As a result, tracking the health of the centrifugal pumps is crucial for avoiding unwelcome stoppages, which might result in the failure of the overall system. This study uses data from several types of sensors collected over a frequency range based on the cavitation hydraulic phenomenon to assess the health or fault of the pump. Data from the healthy pump and the impeller fault pump is considered. The pump ran at various levels of obstruction and frequency to collect data from vibration signals, motor current signals, and pressure signals based on time domain. At various obstruction levels and pump rotation frequencies, detection based on binary classification using the deep learning algorithm is provided as binary classification outperformed in any other algorithm such as kernel SVM, random forest, etc. Multiple statistical attributes are extracted from this data and fed into an artificial neural network (ANN) model and predicting the condition of pump. Moreover, for data classification, hyperparameters are tuned to optimized the model result. The methodology's robustness, as well as that of the developed and tuned model, is also tested and presented to demonstrate their dependability. The proposed failure detection methodology was found to be very efficient during observation. Cavitation is the hydraulic phenomenon used.*

Keywords: Centrifugal pump; Healthy or Faulty state; Machine Learning; Artificial Neural Network

**1. Introduction**

The most common type of hydraulic rotating equipment for home, industrial, medical, and agricultural usage is the centrifugal pump (CP). CP main purpose is to move liquid and deliver it to a high head. The liquid enters through the suction pipe, travels to the pe impeller eye and goes through a number of impeller blades. Due to the centrifugal force generated by the impeller rotating blades, liquid is propelled radially outward and is then forced into a casing before exiting into the piping system below. In several industrial areas, hard particles of various types of pollutants may be present in the liquid being pumped. Impurities may clog the pipe, which results in unstable pump flow. 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. Therefore, it is essential to identify the different types of problems by the examination of their signatures. Most CP defects can be divided into three categories: (i) Mechanical causes, such as worn bearings, bent shafts, unbalanced rotors, and loose parts; (ii) system causes, such as partial or plugged strainers, clogged impellers or suction lines, and installation errors; and (iii) operational causes, such as cavitation, speed and flow issues, and insufficient immersion of the suction pipe. Manufacturing or operational flaws could be to blame for bearing issues. [1]

Leaks and obstructions, as well as flow interruptions or mechanical flaws (such as burst impellers, bearing issues, and bent rotors), are the most frequent causes of CP defect. Suction obstruction issues might result from utilising contaminated working fluids or from damaging the pipe's surface. Such flow obstructions cause the flow rate to decrease and secondary flow, or recirculation, to appear. As flow separation widens, vortices form, which causes a drop in local pressure and the appearance of vapour bubbles. Also, the creation of bubbles in centrifugal pumps is not desirable because it reduces the head that is being created and leads to holes at the surfaces of the pumps due of the creation of micro-jets. Treating CP failures as separate problems is therefore not viable. One defect can make another one appears worse if it already exists. As a result of the ensuing flow instabilities, this causes unpleasant vibrations. Maintenance strategies are readily available to maintain the assets in the sectors. [2]

Ranawat [3] evaluated the accuracy and efficiency of Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) techniques and compared their performance to identify the most suitable technique for fault diagnosis in centrifugal pumps. Furthermore, this paper also proposed the use of various signal processing techniques, such as wavelet transform and empirical mode decomposition, for feature extraction from the pump vibration signals. The use of these techniques allows for better classification of faults, which enhances the accuracy of fault diagnosis. Hasan [4] Proposed a framework that combines scalogram-based imaging and deep learning for more accurate fault diagnosis in centrifugal pumps, which can reduce downtime and maintenance costs. The framework uses deep learning algorithms and scalogram-based imaging to diagnose various centrifugal pump defects, such as misalignment, unbalance, and bearing problems. Wang et. al [5] postulated the use of a combination of Computational Fluid Dynamics (CFD) and optimized the design of a centrifugal pump and evaluated the effects of design parameters on pump performance. The proposed approach offers a more efficient and accurate way of optimizing design parameters for improved pump efficiency.

Gangsar and Tiwari [6] evaluated the use of wavelet-based features, current, and SVM algorithms to diagnose faults in induction motors under various operating conditions. The effectiveness of SVM algorithms for defect diagnosis was assessed under different load levels and speeds. Panda and Tiwari [7] The study suggests using SVM methods based on vibration measurements and signal processing techniques such as wavelet decomposition and principal component analysis to anticipate problems and evaluate the effectiveness of SVM algorithms in predicting flow obstructions and cavitation with varying conditions and severity. Kumar et. al [8] utilized multi-source data and a deep learning algorithm to determine intake pipe blockage levels in centrifugal pumps at a range of speeds results in a more precise and effective approach of recognising blockages than existing techniques.

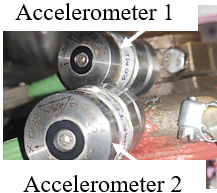
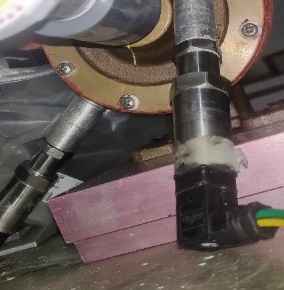
According to the literature, to identify obstructions, vibration, pressure, and motor line signals have not yet been merged. There are still few studies dealing with locating faulty or healthy state of CP. Hence, all the elements motivate researchers to investigate state of the CP.

This study is separated into the upcoming sections as follows: Section 2 discusses the experimental setup and data capture. Section 3 explained of data classification methods is covered. Section 4 highlights steps for NN model and the state of CP. Section 5 presents the final conclusions.

**2. Description of the experimental design and procedure**

The Machine Failure Simulator was used to conduct the experiment on a centrifugal pump. The pump was mounted on a fixed base and powered by a double-belt pulley system with a 3-phase induction motor. The pump had leak-proof fittings and a manual modulating valve to control water flow. Each batch of data contained 5000 samples recorded 150 datasets for 150 seconds with a timeframe of 0.1 seconds. The pump was run at speeds from 30 to 60 Hz with varying levels of blockage at a gap of 5 Hz is achieved through a mechanical modulation valve with six equal intervals marked on it so that varied levels of blockage can be achieved by regulating the valve at various intervals. No clogging is used in B0, B1 shows 16.7% clogging, B2 shows 33.3% clogging, B3 shows 50% clogging, B4 shows 66.6% clogging, and B5 shows 83.33% clogging.

*Accelerometers*: To analyse acceleration, a pump was equipped with two triaxial accelerometers (refer to Figure 1(a)) with sensitivities of 101, 101.1, and 101.4 mV/g (accelerometer-2) and 100.3, 100.7, and 101.4 mV/g (accelerometer-1) in the x, y, and z axes, respectively. *Pressure Transducers*: Two pressure transducers (refer to Figure 1(b)) from "Nictech," a sensitive silicon chip, are used to measure the pressure of liquids. When the pressure sensors are functioning within the operational range (0-60 psi), current fluctuation can be used to understand the response from the sensors (4–20 mA). *Current sensing probes*: "The Keysight 1146B" current probes were used to record the current line reading. Current probes can measure currents between 100 mA and 10 A rms. *Data Acquisition System (DAQ)*: For high accuracy frequency domain measurements, the NI PXI -4472 dynamic signal acquisition module has 8 channels while the NI PXI -6251 dynamic signal acquisition module has 16 channels. The 16 channels of the NI PXI -6251 and the eight channels of the NI PXI -4472 both digitise input signals simultaneously across a DC to 45 kHz bandwidth.

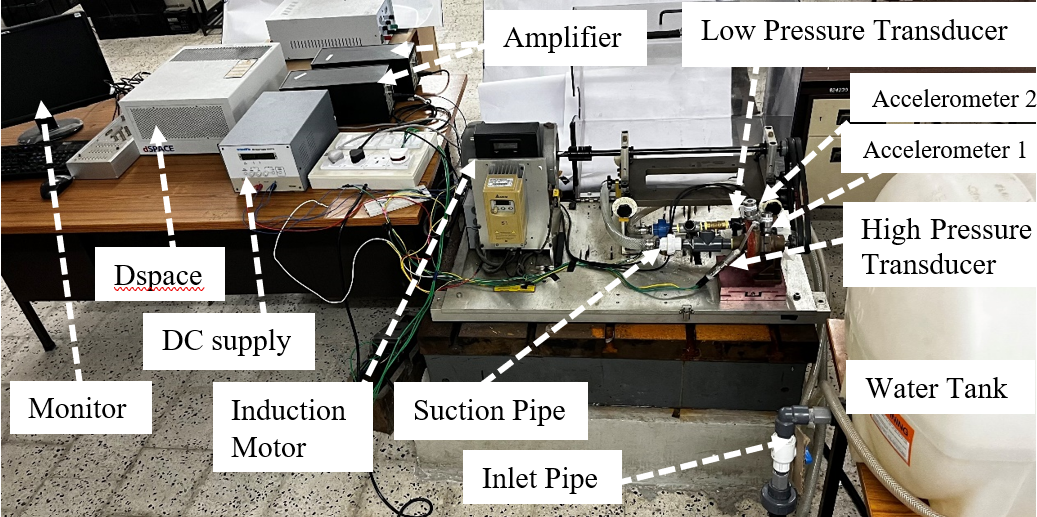
  

(a) (b) (c)

*Figure 1: (a) Tri-axial accelerometers glued to the pump casing and bearing housing (b)* *High- and low-pressure transducers mounted on the pump (c) Impeller fault*

**2.1 An overview of the experiment**

An experiment was run on the MFS to gather data at different levels of flow blockage. Data gathering was carried out for each combination of pump running speed and block level. In order to get comprehensive information about the defect, three distinct measurements acceleration, fluid pressure, and motor line current were taken using several sensors, including accelerometers, pressure transducers, and current probes. Figure 2 depicted the experimental setup.



*Figure 2: Experimental set-up*

**2.2 Evaluating Procedure**

For fault diagnostics, vibration, pressure, and current signatures were used. Data acquisition system (DAQ) was set up using Control Desk software. The 5000 samples were collected at a sampling rate of 5000 samples per second for a time domain measurement. There were 150 seconds of data collection. 5000 x 150 non-overlapping data points were collected for each sensor.

The information was saved in distinct DAQ measurement files in .csv files for each disturbance at various centrifugal pump speeds and was recorded on the system hard drive. The details of DAQ are given in Table 1. The time domain data can then be transformed using a variety of methods into any domain. Sequentially mounted on the MFS were a healthy pump, a pump with a defective impeller, and a pump with a defective cover plate. Figure 1(c) displays an impeller defect fault. The cover plate fault has several pits on it and the impeller plate has cuts on the blades; these are thought to have occurred when high-pressure bubbles burst and impacted the metal wall as tiny jets. A suction blockage defect with six severity levels was included for each of these three pump problems. The pump was run at a variety of speeds during each pump state. The induction motor was run on a variable frequency drive in stages of 300 rpm from 1800 rpm (30 Hz) to 3600 rpm (60 Hz) (5 Hz). For fault diagnostics, vibration, pressure, and current signatures were used. Data were gathered at 5000 samples/s and 5000 samples/record for each defective situation. A total of 150 recordings were gathered for each fault situation after 150 seconds of data collection for each condition.

*Table 1: Data Acquisition System outlines*

|  |  |
| --- | --- |
| Blockage levels | B0 (Full Flow), B1 (1/6 obstruction), B2 (1/3 obstruction), B3 (1/2 obstruction), B4 (2/3 obstruction), B5 (5/6 obstruction) |
| Frequency | 30, 35, 40, 45, 50, 55, 60 Hz |
| Quantity of all fault conditions | 6 × 7 = 42 |
| Sets of measurements for each combination of obstacle and frequency | 150 |
| sample timings | 5000 |
| Timeframe for a single data collection | 1 sec |
| Data collection time for each failure combination | 150 sec |

**2.3. Declaration of Fault Set**

In this study, a total of 18 different CP fault circumstances have been taken into account. Healthy pumps without obstructions (HP0), healthy pumps with blockages at the suction end (HPb), impeller faults without blockages (IF0), and impeller faults with blockages at the suction end (IFb), where b = 1, 2, 3, 4, 5. By bracketing the fault circumstances into the appropriate classes with these conditions, fault sets can be produced. There are two classes, one of which is in a healthy state and the other of which has the faulty state. Table 2 provides an explanation of the both fault sets.

*Table 2: Fault classification sets description*

|  |  |
| --- | --- |
| Fault Set | Classes (labels) |
| 1 | Class 1: HP0, HP1, HP2, HP3, HP4, HP5  Class 2: IF0, IF1, IF2, IF3, IF4, IF5 |

**3. Methodology for data classification**

This study uses supervised learning to classify faults at various frequencies using pressure, acceleration, and current lines. Supervised learning classifies data based on different traits, creating a function connecting input and output. Deep learning uses neural networks to process data and extract sophisticated features. In deep learning, input x is passed through hidden layers to extract important data, and the output layer predicts the class of the input. The expression for **z** is as follows:

 (1)

Where x is the input, w and b are the weight and bias vectors, respectively, with T representing the transpose. Use "Rectified linear unit" (ReLU) activation function to compute vector function z in hidden layers, then apply bias and weight to obtain output in the next hidden layer, and use Sigmoid activation function in the output layer for binary class classification. The sigmoid function is a mathematical function commonly used in artificial neural networks as an activation function. It maps any input value to a value between 0 and 1, and is defined as

 (2)

The sigmoid function is useful in modelling decision making or probability-based scenarios, where the output represents the likelihood of a certain event occurring. The binary loss function which is used in binary classification and it compares the predicted probability of the positive class to the target label of 0 or 1. Binary loss functions include binary cross-entropy and measures the difference between the predicted probability and actual label. binary cross-entropy loss function is expressed as

 (3)

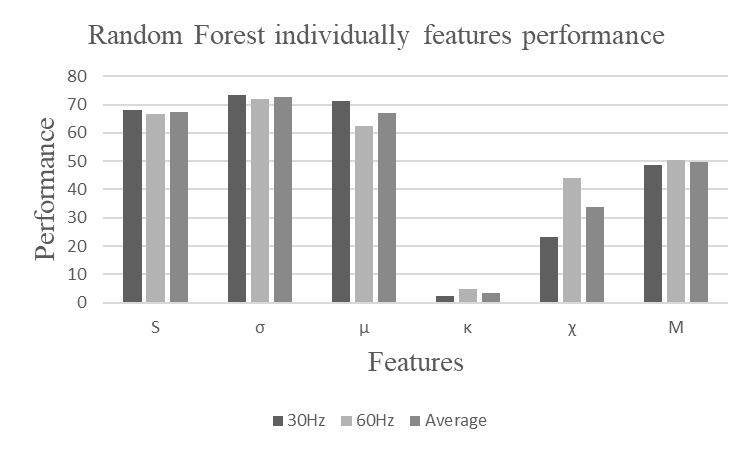
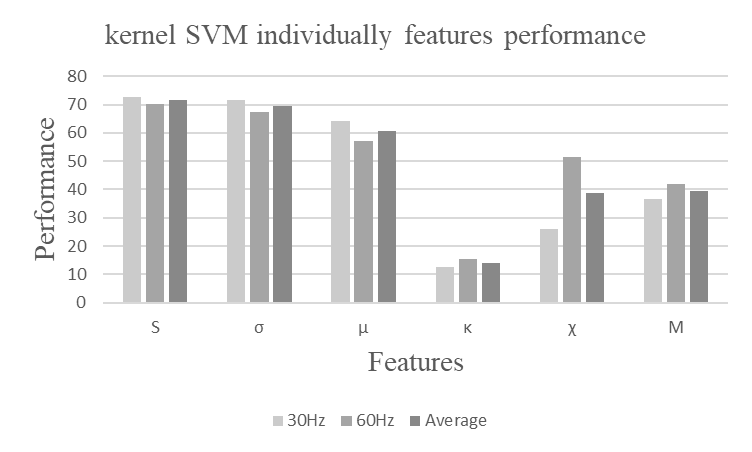
where p(y) is the expected probability for each of the N locations, and y is the label. The open-source Python (Jupyter) deep learning libraries "Keras" and "Scikit learn" are used in this study.

**4. Observations and Opinions**

This study aims to use multiple sensor data to perform multiclass classification and recognize the severity of block utilization. To avoid excessive duplication in input data, only six statistical features are selected from raw sensor data. The selection of these features is not predetermined and can vary based on the defect and data being analysed. The six features considered are mean, standard deviation, mode, entropy, kurtosis, and skewness, which cover different aspects of the data.

**4.1 Performance of separately and combined features**

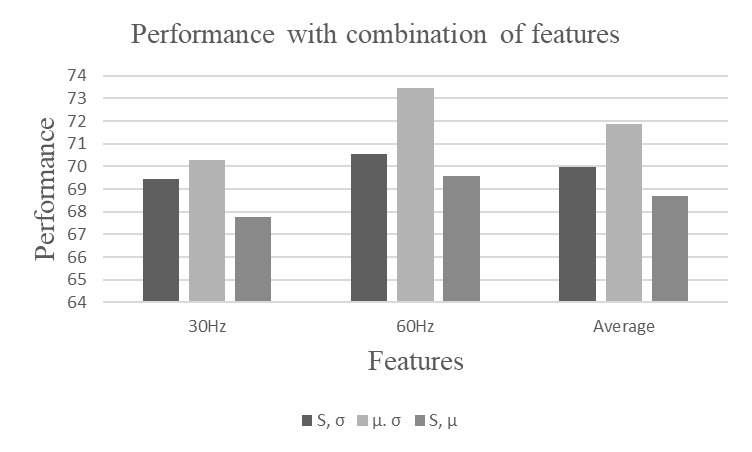
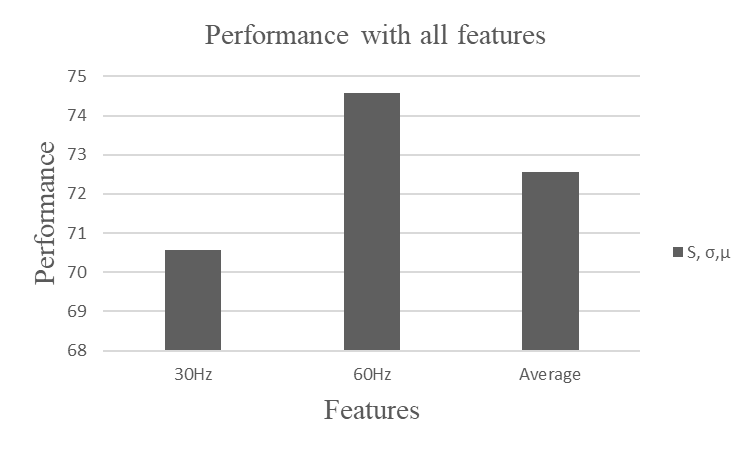
To make processing faster and more efficient, meaningful information can be extracted from raw data by considering statistical features such as standard deviation, variance, mean, median, mode, kurtosis and skewness. In this study, different algorithms including Kernel SVM and Random Forest were used to test the performance of these individual features at three different speeds (30 Hz, 60 Hz). It is depicted in figure 3 (a,b). Results showed that Random Forest performed better than other algorithms, particularly. Therefore, the Random Forest model was selected for further feature calculations with 10 estimators and entropy criterion.

** **

(a)(b)

*Figure 3: Classifier accuracies for several statistical features on average and at two different frequencies (a) Random Forest algorithm used to identify the performance of different feature at three different frequencies. (b) Kernel SVM algorithm used to identify the performance of different feature* *at three different frequencies.*

The standard deviation was the best statistical feature, with mean and entropy also showing good performance. These features were used to calculate blockage severity accuracy. Accuracy was lower at low frequency (30 Hz) but vary with high frequency (60 Hz). It is due to slight signal pulsation variations, while higher frequency led to more bubbles forming and affecting defect severity. The combination of standard deviation and mean provides the best accuracy among all statistical features. The performance of the classifier also increases with increasing frequency, and two best feature combinations of mean and standard deviation are shown in Figure 4(a).

** **

(a) **(**b)

*Figure 4: Classifier performance versus Frequency (a) using multiple features (b) using all the features*

After analysing all three features, Figure 4(b) shows that accuracy increases with frequency, indicating that cavitation severity improves accuracy. High accuracy indicates significant signal variation due to many bubble formations (i.e., cavitation generation), while low accuracy shows normal signal fluctuation. The combination of standard deviation and mean provides the best accuracy among all statistical features. Therefore, all the features are selected for the remainder of the study.

**4.2. Classification algorithms for blockage severity**

This study compares several classification techniques, including XGBoost, Decision Tree, K-Nearest Neighbours, Random Forest, and Artificial Neural Network to determine the best algorithm for a given input. The study uses data from all operating frequencies, including pressure signals, vibration signals, and motor current signals. The results of the study are presented in Figure 5(a), which shows the accuracy of each algorithm. After tuning, the neural network is found to be the best classifier, with a fault set accuracy of 85.90%. The study then used the tuned neural network to predict the state of CP, which are classified into two classes. Section 4.3 describes the steps for tuning the neural network, which is then used to predict blockage severity.

**4.3. Steps for tunning the Artificial Neural Network**

The tuning process involves iterative experimentation with the model. The study found that using three hidden layers with 100, 50, 15 neurons in each layer respectively and a Softmax activation function at the output layer produced the best results. The study also found that using Relu activation in the hidden layers and 100 epochs with 50 minibatch sizes optimized model performance. To avoid overfitting, dropout regularization method used.

Various training testing ratios were tested in neural networks to determine the best classification accuracy, including 25:75, 30:70, 50:50, 80:20, and 90:10. After evaluating all the ratios, the study found that the 80:20 training testing ratio produced the best results shown in figure 5(b).

To train the weights and biases of each neuron in a neural network, optimization algorithms are used to reduce the cost function. This study considered several optimization methods, such as stochastic gradient descent (SDG), RMSprop, Adadelta, Adam, and Nadam. After testing all of them, Nadam was found to perform the best and was used for the study as shown in figure 5(c).

Proper initialization of network weights is critical to prevent gradients from vanishing or exploding, which can affect how the network is trained. This study considered various weight initialization methods, such as Lecun Uniform, Glorot Normal, Glorot Uniform, He Normal, and He Uniform in which Lecun Uniform performed best in all as shown in figure 5(d).

The choice of activation function is crucial in creating any neural network, as it can impact how well the model learns from the training dataset. For binary class classification problems like this one, the output layer typically uses the Sigmoid activation function. However, when it comes to the activation function for hidden layers, it is important to test different options to find the most suitable one for the application. This study considered various activation functions, including Softplus, Relu, Tanh, Sigmoid, Hard-sigmoid activation function, to determine the most appropriate one for the given problem in which Relu performed best as shown in figure 5(e).

|  |  |
| --- | --- |
| *(a)* | *(b)* |
| *(c)* | *(d)* |
| *(e)* | |

*Figure 5: 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.*

**4.4. Test of healthy or defective state of CP**

Table 3 displays the confusion matrix for the model's performance on all feature blockage. The confusion matrix is a C x C matrix that evaluates the classification model's effectiveness, with C representing the total number of target classes. It compares the predicted values generated by the machine learning model with the actual target values.

*Table 3: Confusion matrix for all combined features at two different classes*

|  |  |  |
| --- | --- | --- |
|  | Predicted Levels | |
| Actual Levels | **7498** | 14 |
| 2081 | **5268** |

It can be seen from the confusion matrix that all 12766 instances have been correctly classified by the model.

**5. Conclusion**

We provide effective deep learning-based centrifugal pump malfunction diagnostics using time-domain multiple sensor data. It is quite likely that a pump obstruction, which is progressive in nature, is what causes cavitation and causes the shutdown. So, it is crucial to find state of CP as early as possible. Two states, one is healthy and other is impeller faulty state were considered. All the input features are continuous variables while the output is represented as [1 0] for Healthy, [0 1] for Impeller Faulty for cavitation. The condition of a centrifugal pump can be determined using a simple method that uses features including mean, standard deviation, entropy, skewness, kurtosis, and mode. Several features were extracted from the raw data and different algorithms were used to classify them, and bar charts were used to compare the outcomes. Out of all features, individually standard deviation outperformed and, in the combination, standard deviation, mean and entropy outperformed in all. It is found that total 12766 predication were correct with the accuracy of 85.90% using the tuned model. ReLU is used in the hidden layer while the Sigmoid function is applied to the output layer. The ultimate minibatch size is found 50 an epoch is 100. 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 state of the CP. The results also show that collecting data from multiple sources is always advantageous. The same data classification technique may be utilized to examine various other traits derived from same sensor data for monitoring of the obstruction.

**References**

[1] Tiwari, R. (2017). Rotor Systems: Analysis and Identification CRC Press: Boca Raton, USA, https://doi.org/10.1201/9781315230962

[2] Tabar MTS, Majidi SH, Poursharifi Z (2011) "Investigation of recirculation effects on the formation of vapor bubbles in centrifugal pump blades". World Academy of Science, Engineering and Technology, January 2011, 73:521–526

[3] N. S. Ranawat, P. K. Kankar, and A. Miglani, “Fault diagnosis in centrifugal pump using support vector machine and artificial neural network,” *Journal of Engineering Research (Kuwait)*, vol. 9, pp. 99–111, 2021, doi:10.36909/jer.EMSME.13881.

[4] Hasan, M. J., Rai, A., Ahmad, Z., & Kim, J.-M. (2021). "A Fault Diagnosis Framework for Centrifugal Pumps by Scalogram Based Imaging and Deep Learning," IEEE Access, 1–1. doi:10.1109/access.2021.3072854

[5] C. N. Wang, F. C. Yang, V. T. T. Nguyen, and N. T. M. Vo, “CFD Analysis and Optimum Design for a Centrifugal Pump Using an Effectively Artificial Intelligent Algorithm,” *Micromachines*, 2022, Taiwan, doi: 10.3390/mi13081208.

[6] P. Gangsar and R. Tiwari, “Diagnostics of mechanical and electrical faults in induction motors using wavelet-based features of vibration and current through support vector machine algorithms for various operating conditions,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, volume 41, no. 2, 12 January 2019, Guwahati, Assam, India, doi: 10.1007/s40430-019-1574-5.

[7] A. K. Panda, J. S. Rapur, and R. Tiwari, “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*, Guwahati, Assam, India volume 130, 2018, doi: 10.1016/j.measurement.2018.07.092.

[8] D. Kumar, A. Dewangan, R. Tiwari, and D. J. Bordoloi, “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*, volume 186, Guwahati, Assam, India September,2021, doi: 10.1016/j.measurement.2021.110146.

1. Corresponding author.

   E-mail address: rtiwari@iitg.ac.in (R. Tiwari). [↑](#footnote-ref-1)