eng

MDPI

Article

Vibration Analysis of a Centrifugal Pump with Healthy and Defective Impellers and Fault Detection Using Multi-Layer Perceptron

Masoud Hatami Garousi [1], Mahdi Karimi [2], Paolo Casoli [1], Massimo Rundo [3] and Rasoul Fallahzadeh [1,*]

[1] Department of Engineering for Industrial Systems and Technologies, University of Parma, 43124 Parma, Italy; masoud.hatamigarousi@unipr.it (M.H.G.); paolo.casoli@unipr.it (P.C.)
[2] Department of Engineering, Bu-Ali Sina University, Hamedan 65175-38695, Iran; m_karimi@basu.ac.ir
[3] Department of Energy, Politecnico di Torino, 10129 Turin, Italy; massimo.rundo@polito.it
* Correspondence: rasoul.fallahzadeh@unipr.it

Abstract: Centrifugal pumps (CPs) are widely utilized in many different industries, and their operations are maintained by their reliable performance. CPs' most common faults can be categorized as mechanical or flow-related faults: the first ones are often associated with damage at the impeller, while the second ones are associated with cavitation. It is possible to use computational algorithms to monitor both failures in CPs. In this study, two different problems in pumps, the defective impeller and cavitation, have been considered. When a CP is working in a faulty condition, it generates vibrations that can be measured using piezoelectric sensors. Collected data can be analyzed to extract time- and frequency-domain data. Interpreting the time-domain data showed that distinguishing the type of defect is not possible. However, indicators like kurtosis, skewness, mean, and variance can be used as input for the multi-layer perceptron (MLP) algorithm to classify pump faults. This study presents a detailed discussion of the vibration-based method outcomes, emphasizing the benefits and drawbacks of the multi-layer perceptron method. The results show that the suggested algorithm can identify the occurrence of different faults and quantify their severity during pump operation in real time.

Keywords: vibration analysis; multi-layer perceptron; centrifugal pump; cavitation; fault detection

Citation: Garousi, M.H.; Karimi, M.; Casoli, P.; Rundo, M.; Fallahzadeh, R. Vibration Analysis of a Centrifugal Pump with Healthy and Defective Impellers and Fault Detection Using Multi-Layer Perceptron. *Eng* **2024**, *5*, 2511–2530. https://doi.org/10.3390/eng5040131

Academic Editor: Antonio Gil Bravo

Received: 21 August 2024
Revised: 3 October 2024
Accepted: 4 October 2024
Published: 8 October 2024

1. Introduction

Centrifugal pumps are important in oil, gas, and other industries. Therefore, knowledge and understanding of the behavior of these types of pumps are essential. Experimental tests are a suitable method for troubleshooting devices. For example, measuring and identifying different vibration frequencies related to the pump impeller, rotating shafts, and bearings is a crucial test in diagnosing potential issues and ensuring the efficient and reliable operation of the pump. By identifying specific vibration patterns, it is possible to detect early signs of wear or damage, allowing for timely maintenance and reducing the risk of unexpected failures. Vibration signals in the time, frequency, or time-frequency domains provide information about the status of equipment and devices [1–3]. In CPs, various problems, such as impeller failure and cavitation, can occur. Detecting and fixing these defects requires a detailed understanding of how the device works and the reasons for the defect. Predictive maintenance involves monitoring the vibration of rotating machines to discover defects in the early stages and prevent the progression of failure. Predictive maintenance helps to determine the conditions of the working equipment to define the appropriate maintenance time [4–7]. In other words, predictive maintenance uses the actual operating conditions of factory equipment and devices to optimize the overall operation of the equipment [8–10].

Mousmoulis et al. [11] investigated cavitation in centrifugal pumps, concentrating on its emergence and growth in three distinct impellers. They also demonstrated that impeller geometry significantly influences cavitation behavior, affecting inception points and behavior. Using flow visualization, acoustic emissions, and vibration monitoring, they found that accelerometers and acoustic emission sensors were effective in detecting cavitation onset well before the total head drop. Cavitation inception occurs at higher flow rates, consistently appearing at the suction side of leading blade edges. Al-Obaidi [12] investigated the effect of pump rotational speed on the performance and detection of cavitation using vibration signals. The results showed that the amount of cavitation has a direct relationship with the rotational speed and the flow rate. Mousmoulis et al. [13] investigated the vibration analysis, flow observation technique, and wave propagation technique on the plexiglass impeller pump to detect cavitation. They showed that the cavitation values increased with the increase in flow rate. Sun et al. [13] used the Hilbert spectrum to detect cavitation in a CP. They concluded that the Hilbert transform is suitable for processing transient and unsteady signals, and time-frequency domain characteristics can be extracted. Murovec et al. [14] used the acoustic diffusion method to detect cavitation in a radial CP. Using acoustic criteria and mathematical parameters, they provided an effective method for predicting and classifying cavitation in a CP. Kumar et al. [15] used time-frequency domain analysis and support vector machines to detect faults in a CP. At first, they trained the machine using the obtained data, and then they used it to identify defects such as cavitation and bearing defects. Azizi et al. [16] employed the hybrid feature selection technique to identify the degree of cavitation. Using this method, it is possible to differentiate between three states: no bubbling, limited bubbling, and cavitation that occurred at the outflow. Birajdar et al. [17] studied the sources of vibration and noise in CPs and the methods of troubleshooting this type of pump. They listed the sources of vibration in CPs as unbalanced rotors, defective impellers, broken shafts, damaged bearings, and cavitation at critical speeds. They concluded that cavitation is created randomly and in a high-frequency energy range, but it has sometimes been mistaken for blade passing frequency. They presented a diagram for the range of the frequency spectrum. Sakthivel et al. [18] investigated fault diagnosis and fault prediction in CPs with the help of a soft computing approach. They used different classification patterns, such as support vector machines, genetic algorithms, and wavelet analysis, to classify the defects. Six parameters, including the bearing defect, impeller defect, leakage defect, impeller and bearing defect together, and cavitation, were considered, and an algorithmic decision tree was used to select the defect. They analyzed the results using a genetic algorithm (GA), support vector machine (SVM), and wavelet analysis (WA) and finally observed that both the GA and SVM have better performance compared to other classifications for defects. Askari et al. [19] investigated the troubleshooting and diagnosis of the cavitation phenomenon by using the vibration analysis method and provided solutions to solve it. They investigated the abnormal performance of the CPs of a refining unit with the modal analysis method and finally determined the vibration sources of the pump according to the frequency range determined for cavitation, predicting the existence of this phenomenon in the pump. Al-Braik et al. [20] investigated the troubleshooting of CPs to detect impeller defects. They conducted a test on a healthy impeller and five impellers with varying degrees of defects on the vane tips. Vibration data were collected at different flow rates, and they obtained the frequency spectrum for these conditions. They showed that discrete spectral components at vane-passing frequencies and higher-order harmonics of the shaft frequency are effective for diagnosing impeller faults and that primary spectral components of the turbulent flow occur above 1 kHz. Spadafor et al. [21] conducted dynamic system simulations and optimizations of a CP to investigate impeller failure. They aimed to determine whether the failure was due to metallurgical issues, such as improper heat treatment, corrosion, or mechanical factors. After replacing the pump with another impeller with the same geometry but made from a stainless steel alloy, they observed the same failure results. Vibration analysis revealed that torsional vibrations introduced by motor oscillations were

affecting the pump. Their experimental research, using a hydraulic response system, demonstrated that pressure fluctuations were closely related to shaft speed variations and changes in engine torque, ultimately causing damage to the pump impeller. Wang and Cheng [22] used wavelet transmission (WT) and a partially linearized neural network (PNN) to extract the characteristics of cavitation from the vibration signals and then used non-dimensional symptom parameters like the mean and standard deviations as input for the artificial neural network (ANN) to determine whether the pump was healthy or defective. The results obtained showed that WT successfully processes cavitation vibration signals, and other processing methods are also effective. Barrio et al. [23] investigated radial loads on centrifugal pump impellers and found significant unsteady components that were challenging to estimate. They showed that the unsteady component (the fluctuating part of the radial load on the impeller) could represent 40% to 70% of the average load when operating at off-design conditions (deviations from the optimal flow rate and pressure for which the centrifugal pump was designed). The CFD simulations demonstrated strong agreement with the experimental data in terms of both global performance and unsteady pressure distribution. Their findings highlight the importance of accounting for unsteady loads in pump design to prevent fatigue failure. Casoli et al. [24,25] studied a vibration signal-based method for fault identification and classification in hydraulic axial piston pumps. Based on the vibration signals and the use of SVM and ANN, they proposed an algorithm to detect the health state of a variable displacement axial piston pump. By using the time-sampling raw signals, they achieved a satisfying accuracy. They showed that the proposed algorithm can identify the faults in the axial piston pump for each working condition.

Following the studies in the literature, vibration analysis and fault detection in CPs are presented in this paper. Most of the researchers explored the utilization of statistical parameters in the frequency domain as input for an ANN. Moreover, there has been a noticeable absence of emphasis on investigating the time-domain, frequency-domain, and classification methods, such as MLP, within the scope of centrifugal pumps. The main objective of this work is to present the practical vibration analysis method for a CP and use statistical parameters of the time domain, including mean, kurtosis, variance, RMS, and skewness in the MLP algorithm to identify the faults in CPs. This study offers superior advantages of MLP over fast Fourier transform (FFT). Unlike FFT, MLP automates the feature extraction process from time-domain signals, eliminating the need for manual interpretation and reducing reliance on expert knowledge. Furthermore, MLP's ability to handle non-stationary and transient signals enhances its effectiveness for real-time fault detection.

2. Methodology

In this section, the test method is explained, and then the multi-layer perceptron neural network method is introduced.

2.1. Test Method

The test bench used, as seen in Figure 1, is made of several parts, such as a DC motor, a voltage converter, suction and discharge tanks, a centrifugal pump (radial flow), a pressure sensor, and a flow and pressure adjustment valve. Vibration signals in our experiment were measured using a piezoelectric accelerometer under the trade name 2224C (Endevco, Halifax, NC, USA), which was mounted horizontally and parallel to the pump axis. Specifically, the accelerometer was installed on the body of the centrifugal pump, near the impeller housing, and on the horizontal ventilation bolt, as this area is highly sensitive to vibrations induced by impeller defects or cavitation. Equipment such as a piezoelectric accelerometer, an amplifier, and a DS20080A two-channel oscilloscope card (OC) -(TNM Electronics, Tehran, Iran) were used to collect the signals. The specifications of the Endevco 2224C accelerometer are presented in Table 1.

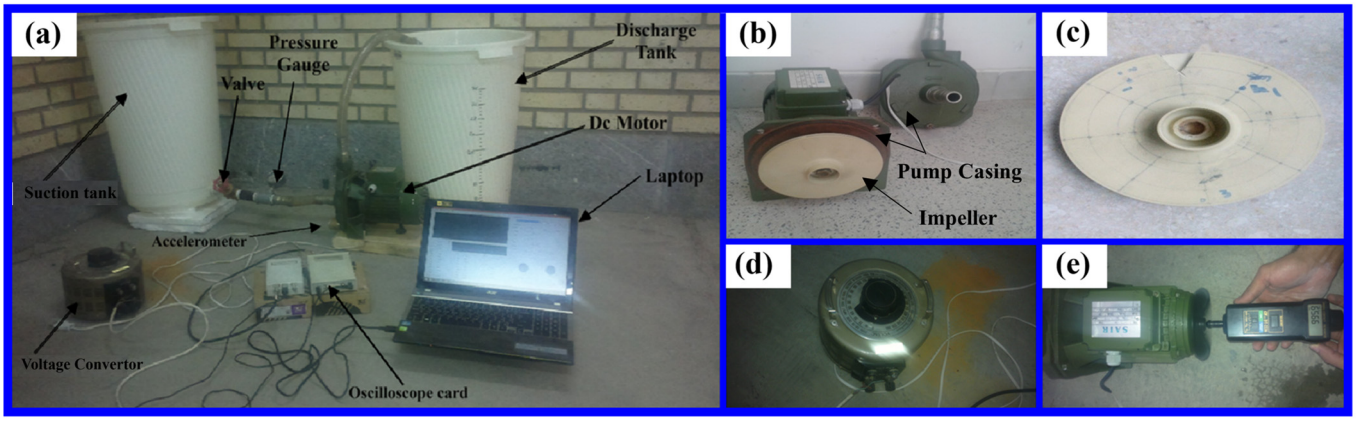


Figure 1. (**a**) Laboratory test rig, (**b**) pump with healthy impeller, (**c**) defective impeller, (**d**) voltage converter, and (**e**) tachometer.

Table 1. Key parameters of the Endevco 2224C accelerometer.

Attribute	Testing Conditions/Remarks	Measurement Unit
Sensitivity	±12	pC/g
Sinusoidal limit	1000	g
Shock limit	2000	g
Operating temperature	−55–+177	°C (°F)
Frequency response	0.1–10,000	±1 dB Hz

The pump used in this experiment is the SAER CMP76, which is used to pump the water with an impeller made of polyethylene. The pump's and impeller's specifications are shown in Table 2.

Table 2. Main parameters of SAER CMP76 centrifugal pump.

Name	Value
Flow rate	0–6 m^3/h
Head	21–29 m
Efficiency	75%
Impeller inlet diameter	36 mm
Impeller outlet diameter	148 mm
Impeller outlet width	2 mm
Power	0.75 kW
Blade number	6
Specific speed	58.45
Flow coefficient	0.00165
Head coefficient	0.1233

To test a case representing faulty mechanical conditions, the impeller was damaged by removing a triangular piece from it, as shown in Figure 1c. This type of defect in pump impellers, known as a notch defect, usually occurs at the outlet diameter because of the higher speed. The size of the notch defect can indeed vary depending on several factors, such as the operational conditions of the pump (e.g., speed and flow rate), the material of the impeller, and the duration of exposure to mechanical stress or cavitation. In our study, we removed a triangular piece of the impeller to simulate a typical notch defect often observed at the outlet diameter due to high-speed and tension operations. However, in real-world scenarios, the size and shape of the defect can differ. Notch defects may grow over time as the pump continues to operate under faulty conditions, especially if cavitation or abrasive particles are present in the fluid, further aggravating the damage. Thus, our experiment used a controlled defect size for repeatability, but the same methodology can

be applied to varying sizes of impeller damage. Indeed, changing the size of the notch will change the amplitude of the peaks, but the frequency range will remain constant. The defect notch was created specifically on the impeller shroud, not the diffuser or volute. Regarding the dimensions, the notch was a triangular cut with approximately 10 mm sides and a 15 mm height, located at the impeller outlet where the tension is highest. The notch was positioned facing the shroud. Figure 1 depicts a test bench with both healthy and defective impellers, while Figure 2 shows a schematic representation of the test bench. "AS" in Figure 2 refers to the accelerometer sensor used to measure vibration signals.

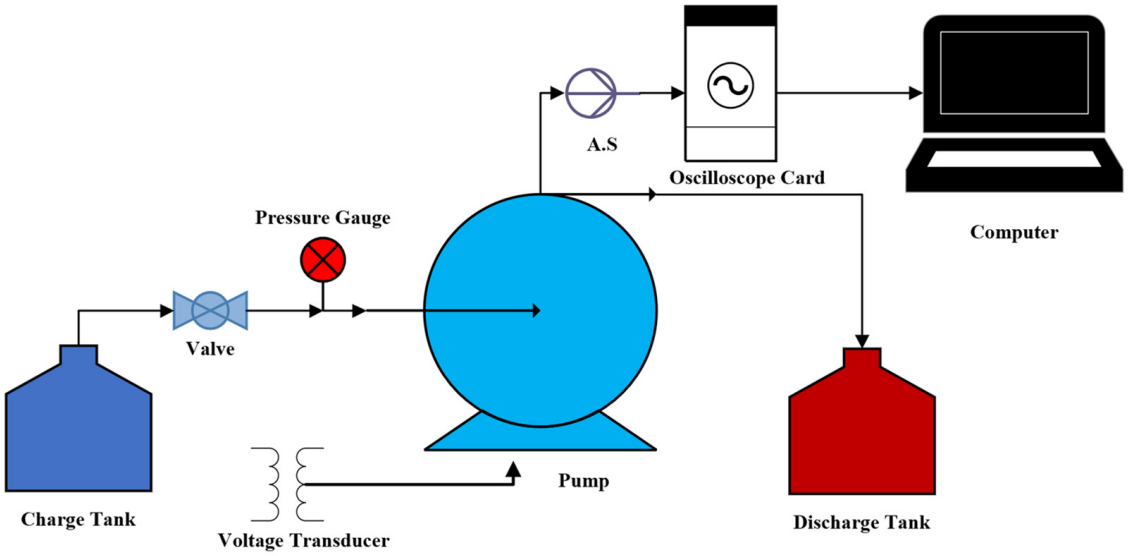


Figure 2. Schematic diagram of the experimental setup.

A single-phase induction motor is directly coupled to the pump and drives it. For changing the rotational speed of the motor, a voltage converter was used, and the speed was measured with a portable tachometer. The pump speed ranges from 500 to 2880 r/min. The tank that supplies water for the pump was positioned to ensure sufficient head at the inlet and to prevent any inherent cavitation. Vibration signals were acquired by the accelerometer sensor and transferred to the OC and computer to save the data. The sampling frequency of the OC was 100,000 samples per second. When the pump is operating, the procedure consists of gradually closing the valve on the inlet pipe to create cavitation until the bubbles are visible through the transparent suction pipe. The cavitation bubbles were visually detected through the transparent suction pipe, and no additional optical devices were used for this purpose. The transparent pipe provided sufficient visibility for the manual observation of cavitation onset and progression during the experiments. The procedure was first carried out with a healthy impeller and then with the damaged one.

2.2. Multi-Layer Perceptrons

An artificial neural network (ANN) is characterized by its internal connectivity and associative connections used to process information. In most cases, an ANN is a flexible system that adjusts its structure in response to input from either the outside or inside domain [26,27]. One of the main types of ANN is the feedforward neural network (FNN), characterized by the unidirectional flow of information between its layers. A significant sub-branch of the FNN is the multi-layer perceptron (MLP). MLP neurons typically use nonlinear activation functions, allowing the network to identify complex patterns in the data. An output signal is generated by every unit (neuron), and this signal is a function of the sum of its inputs. The output as a function of the input and weights is expressed as:

$$y_i = f\left(\sum x_i w_i\right) \quad (1)$$

The activation function can be any function; however, sigmoid functions (also known as the hyperbolic tangent) are most frequently employed. An MLP consists of layers stacked one after another, with varying numbers of processing units in each layer. The units in the first layer are fully connected to the units in the hidden layer and receive input from the external environment [27]. Meanwhile, the units in the hidden layer are fully connected to the units in the output layer, and the units in the output layer produce the MLP's output, as shown in Figure 3.

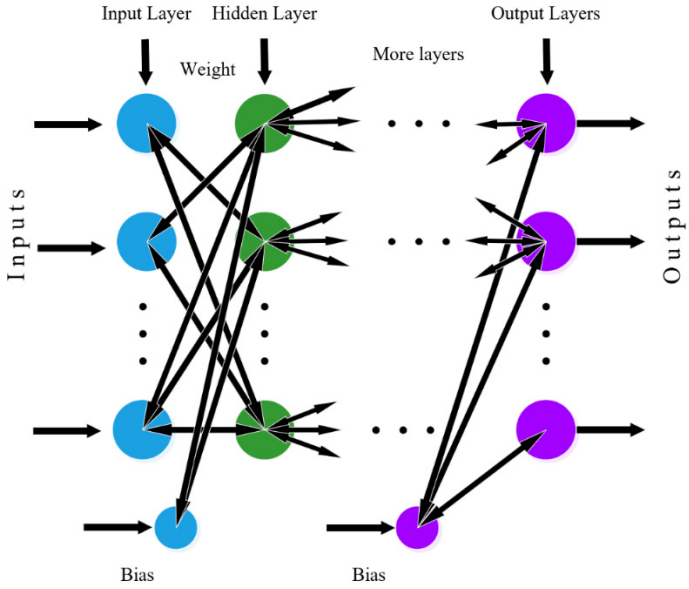


Figure 3. Architecture of the MLP network.

2.2.1. Training Method

To achieve a desired task, an MLP must first be trained. This means that the network's connection weight values need to be defined so that it can produce the correct output for each input pattern. A training algorithm calculates these appropriate weights. There are numerous training methods and their variations. For instance, the backpropagation algorithm is a commonly used method where the error is propagated backward through the network to iteratively adjust the weights. This process involves passing an input forward through the network to obtain an output, calculating the error by comparing this output to the expected result, and then updating the weights to minimize this error over many iterations. It should be noted that another goal of training a neural network is to achieve a good generalization ability rather than merely memorizing the training set. In other words, the network should generate accurate output values for inputs that were not seen during training [28]. During training, the early stopping technique [29–31] is employed to enhance the network's generalization performance and prevent overtraining. This method involves selecting a validation set that is distinct from the training set. The validation error serves as the stopping condition during the training process. In this study, early stopping criteria are implemented to achieve optimal performance. The training algorithm adopted in this study is the Levenberg–Marquardt algorithm.

2.2.2. Levenberg–Marquart Algorithm

The Levenberg–Marquardt (LM) algorithm is an optimization method that effectively combines the gradient descent and Gauss–Newton algorithms to solve nonlinear least squares problems. It aims to minimize the sum of squared errors between predicted and actual outputs. The algorithm updates parameters iteratively using the following formula:

$$P_{k+1} = P_k - \left(J_k^T J_k + \lambda I\right)^{-1} J_k^T e_k \quad (2)$$

where J_k is the Jacobian matrix of partial derivatives, e_k is the error vector, λ is a damping factor, and I is the identity matrix. The damping factor λ adjusts the blend between gradient descent and the Gauss–Newton method. For large λ, the algorithm behaves like gradient descent, ensuring stability, while for small λ, it behaves like the Gauss–Newton method, ensuring fast convergence. This adaptive approach makes the LM algorithm particularly suitable for training neural networks and curve fitting, providing a balance between efficiency and robustness, although it can be memory-intensive due to the computation of the Jacobian matrix [32].

2.3. Statical Parameters as Input Data

Statistical parameters like mean and variance are primarily used for data normalization and understanding the spread of data. Skewness and kurtosis help in understanding the asymmetry and tailenders of data, guiding data transformation processes, while root mean square (RMS) provides an accurate measure of the magnitude of varying quantities, making it essential for assessing prediction errors and fluctuations in data. The accuracy and reliability of an MLP can be significantly improved by ensuring that the data input into the model is well-prepared by utilizing these statistical properties [33].

In the following, statistical parameters like mean, kurtosis, variance, skewness, and RMS are described.

Calculating the average of measured data is done by:

$$\overline{Z} = \frac{Z_1 + Z_2 + \dots Z_n}{n} \tag{3}$$

where $\overline{Z}$ is the mean value, $Z_1, Z_2, \dots Z_n$ are measured data at each time, and n is the time of data recorded.

Variance is the parameter used to measure the distribution and dispersion of data around their average value:

$$\mathrm{Var}_x = \frac{1}{N}\left[\sum_{i=1}^{N}(x_i - \overline{x})^2\right] \tag{4}$$

where N is the number of samples, x is the time signal sample, and $\overline{x}$ is the mean value.

Standard deviation is the amount of variation or dispersion in a set of data values. It indicates how much individual data points differ from the mean (average) of the dataset.

$$\sigma = \sqrt{\mathrm{Var}_x} = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}(x_i - \overline{x})^2\right]} \tag{5}$$

Skewness measures the distortion or asymmetry of a signal, which can be visually interpreted by examining the signal's distributional shape, and is given as:

$$\chi = \frac{\sum_{i=1}^{N}(x_i - \overline{x})^3}{N\sigma^3} \tag{6}$$

where $\overline{x}$ is the mean value, and σ is the standard deviation of the data [33].

Kurtosis is defined as the fourth-order moment of data and shows the degree of a peak in a set of data [33]. In other words, kurtosis is an index to detect larger peaks among the data, or it can be said that it determines the shape of a data distribution and is given as:

$$K = \frac{N\sum_{i=1}^{N}(x_i - \overline{x})^4}{\left[\sum_{i=1}^{N}(x_i - \overline{x})^2\right]^2} \tag{7}$$

Our interpretation of the time-domain data also included the use of the root mean square (RMS) parameter. The RMS value is a statistical measure of the magnitude of a

varying quantity. It is especially useful in contexts where the values can be both positive and negative [33]. RMS is given as:

$$RMS_x = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N} x_i{}^2\right]} \quad (8)$$

where x is the time signal sample, N is the number of samples, and i is the sample index [33].

The statistical convergence of skewness and kurtosis was ensured by employing a sufficiently large sample size during data collection and analysis. Additionally, to measure the stability of the statistical assessment, a test was conducted across multiple trials to confirm their reliability.

2.4. Training Procedure

The training procedure for an MLP network involves feeding the network with input data and corresponding target outputs and then adjusting the network's weights and biases through backpropagation to minimize the error between predicted and actual values. The neural network employed here consists of three layers: the input, middle, and output layers. The input layer has neurons that represent the normalized features of the extracted vibration signals. To minimize the mean square error or functional function between the network's output and the target, networks are alternately trained. The application also automatically generates the network's initial weights and biases. This is the reason why the network has undergone several iterations of training. Only 1 and 0, which denote the healthy and faulty states of the pump, respectively, can be the target values for the two output neurons. However, since the combination of the intermediate and output transfer functions is considered a mix of hyperbolic and linear transfer functions, the neural network's output can range from zero to one and may even exceed one. Extracted statistical parameters from the time domain were collected in the form of a 5 × 20 matrix. The number 5 represents the statistical features used, and 20 is equivalent to the data in four different functional states and tested at five rotational speeds (500, 1000, 2000, 2500, and 2880 rpm). From the 5 × 20 matrix, a 5 × 16 matrix was used for training the algorithm, while a 5 × 4 matrix was reserved for validation to determine if the algorithm can correctly anticipate faults. The target matrix of the neural network was chosen as 2 × 1. The data used for training and testing the ANN is shown in Table 3.

Table 3. Input data of neural network.

M, h500		M, h2880	M, hc500		M, hc2880	M, d500		M, d2880	M, dc, 500		M, dc, 2880
K, h500		K, h2880	K, hc500		S, hc2880	K, d500		S, d2880	K, dc, 500		K, dc, 2880
S, h500	···	S, h2880	S, hc500	···	S, hc2880	S, d500	···	S, d2880	S, dc, 500	···	S, dc, 2880
V, h500		V, h2880	V, hc500		V, hc2880	V, d500		V, d2880	V, dc, 500		V, dc, 2880
RMS, h500		RMS, h2880	RMS, hc500		RMS, hc2880	RMS, d500		RMS, d2880	RMS, dc, 500		RMS, dc, 2880

In this matrix, M is the mean, K is the kurtosis, S is the skewness, V stands for the variance, and RMS shows the root mean square value. Also, h represents healthy status, hc is for health with a cavitation status, d is for the defect status, and dc is for a defect with cavitation status. The numbers after the letters show the rotational speed. For example, M, hc2880 means the mean value of the health with a cavitation status at 2880 rpm.

There are ones and zeros in the target matrix. In this selected target matrix, if the output of the neural network is closer to the value of one, it is a sign of the health of the pump, and if it is closer to the value of zero, it indicates the presence of a defect.

The flowchart of the process for training ANNs is shown in Figure 4.

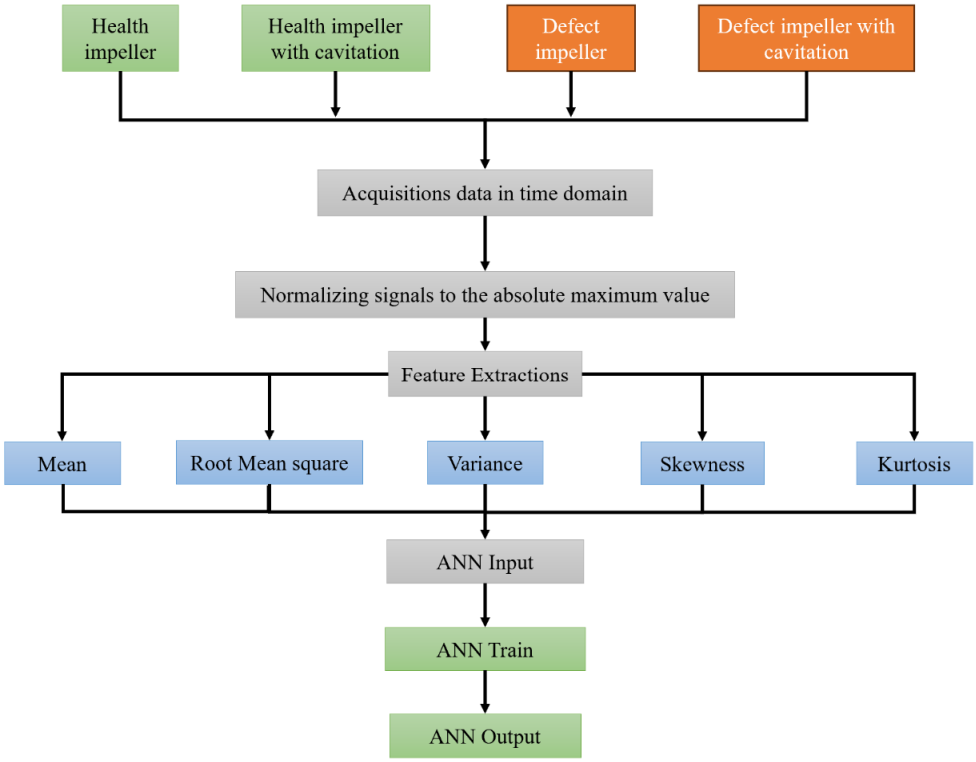


Figure 4. Flowchart of the neural network training process.

3. Results and Discussion

Time-domain analysis was conducted, and graphs were generated using the data obtained from the accelerometer. In the next step, statistical factors such as the mean, kurtosis, skewness, variance, and RMS were extracted from the normalized time signals. Subsequently, frequency-domain analysis was performed using Fourier transform to diagnose and monitor the faults. Finally, an MLP neural network was employed to predict the defects based on the statistical data.

3.1. Time-Domain Analysis

The time domain displays the range value of signals and vibrations measured by the accelerometer at moments during sampling. Basic troubleshooting techniques were performed to analyze the signals in the time domain. Since the maximum input frequency is 48 Hz (the maximum rotational speed of the pump is 2880 rpm), the periodicity of the signals is 0.0208 s. The interval of the horizontal axis can be selected from 0 to 208 samples, and since the sampling rate is equal to 10,000 samples per second, this interval shows the periodicity of the signal with the input frequency. However, due to the nature of the randomness of the signal, this interval was chosen from 0 to 416. It should be noted that the value of the data collection frequency has been chosen to prevent the occurrence of an aliasing phenomenon. It is important to make sure that the sampling rate for transient monitoring is high enough to capture the parameters of the system dynamics. Data were normalized to fix the data range between −1 and 1 based on this formula:

$$\widetilde{x} = 2 \times \left(\frac{x - \min(x)}{\max(x) - \min(x)} \right) - 1 \quad (9)$$

In this formula, x is the original value, $\widetilde{x}$ is normalized value, and min(x) and max(x) are the minimum and maximum values in the dataset, respectively. The normalized graphs in the time domain for different states of the pump at 2880 rpm are shown in Figure 5.

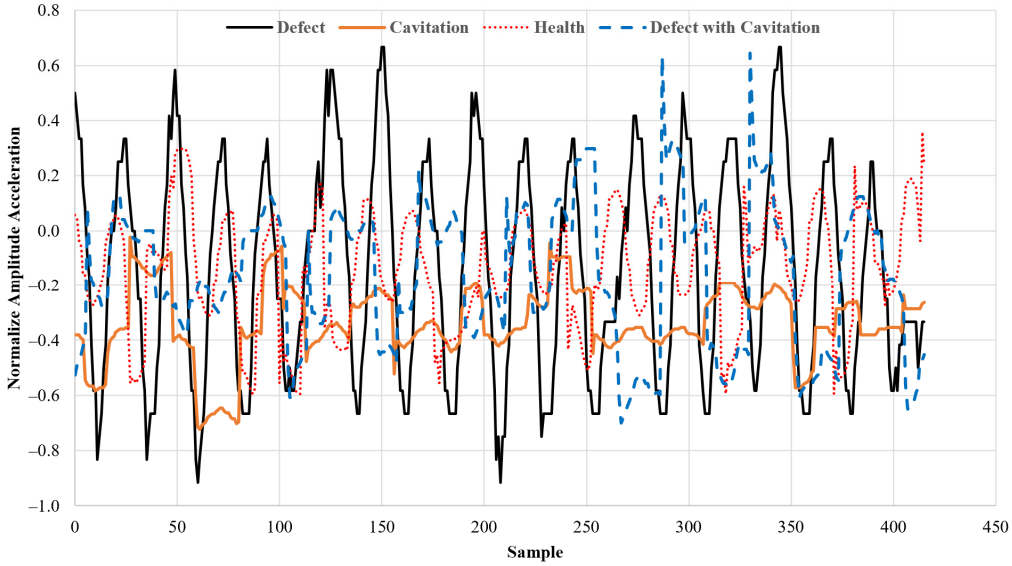


Figure 5. Time domain diagrams for 4 different statuses of the pump (2880 rpm).

The negative values in Figure 5 indicate that the measured vibration signals exhibit a shift in their amplitude relative to the baseline level. This could suggest that the vibrations are oscillating around a mean value, which may occur due to factors like noise in the system or the inherent characteristics of the signal processing method used. As it is well known, detecting pump defects using this diagram is challenging. Consequently, variations in the instantaneous time signals do not provide a comprehensive understanding of identifying the pump defects. Therefore, further analysis is needed to analyze the signals in the time domain and detect the defects using various statistical features. For better diagnosis and analysis of the time-domain graphs, statistical parameters such as the mean, variance, and kurtosis could be extracted. The mean value is presented in Figure 6.

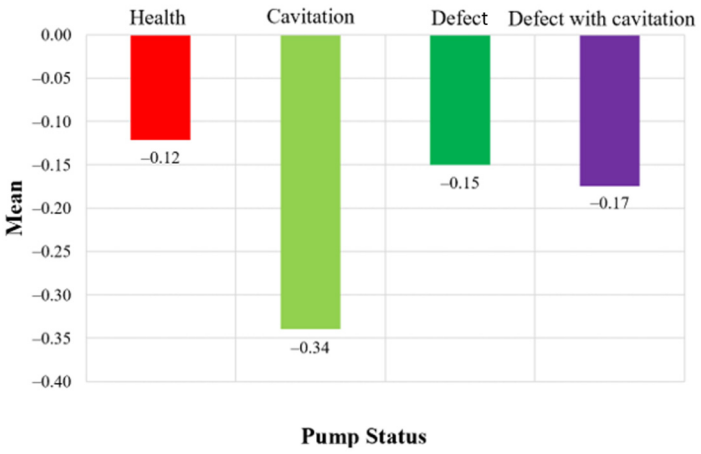


Figure 6. Mean values for different pump statuses.

This analysis shows that the average signal level varies across different pump states, with the cavitation state showing the most significant decrease, followed by the defect with cavitation state, the defect state, and finally, the healthy state with the smallest decrease.

The kurtosis diagram for the extracted data for the 2880 rotational speed is shown in Figure 7.

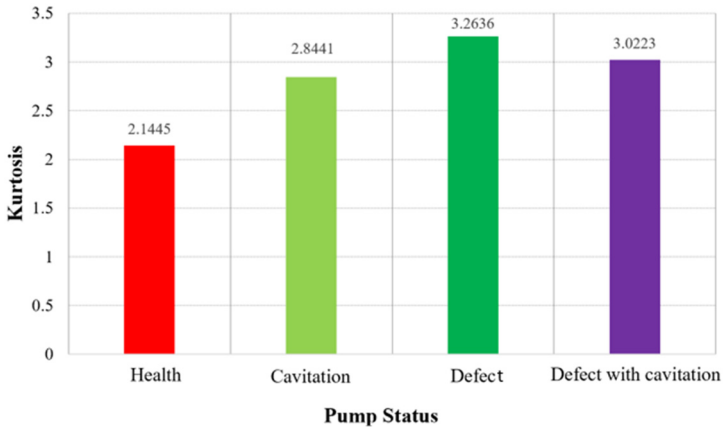


Figure 7. Kurtosis values for different pump statuses.

The kurtosis value indicates a distribution with higher peaks and heavier tails than a normal distribution. The kurtosis value for the cavitation state is 2.84; this kurtosis value is more than that of the healthy state, indicating that the cavitation state has a higher peak and heavier tails compared to the healthy state. The kurtosis value for the defect state is 3.26; this value shows that the defect state has the highest kurtosis value, suggesting a spike distribution with heavier tails compared to all other states. This analysis shows that the defect with and without cavitation states exhibit higher peaks and heavier tails in their distributions, indicating more extreme values. In contrast, the healthy state has a flatter distribution, suggesting fewer extreme variations.

The unequal distribution of a signal about its mean value is measured by its skewness. Therefore, in the next step, the value of the signal's skewness was calculated, as shown in Figure 8.

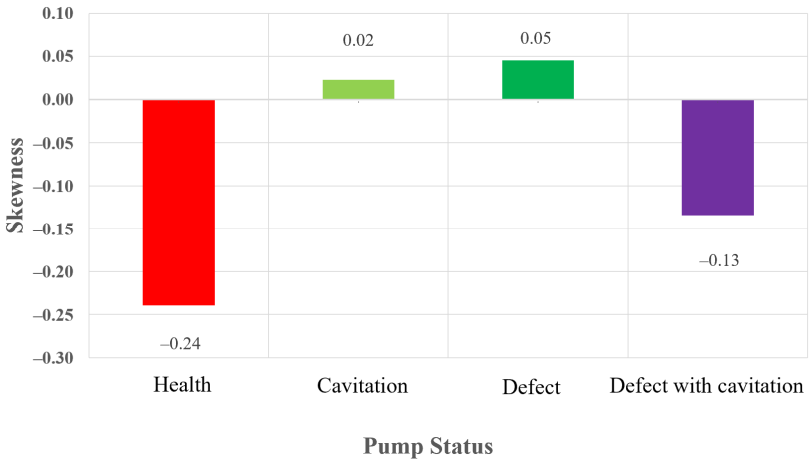


Figure 8. Skewness values for different pump statuses.

The negative skewness value indicates that the signal for the healthy pump status shows a significant left skew distribution, suggesting that the data points are spread out more on the left side of the mean. A skewness value close to zero indicates that the signal distribution for the cavitation state is almost symmetrical. A slightly positive skewness value indicates that the signal for the defect state has a longer tail on the right side of the distribution, suggesting that the data points are spread out more on the right side of the mean. The negative skewness value indicates that the signal for the defect with cavitation state has a longer tail on the left side of the distribution, like the healthy state, but less than that. This analysis suggests that the skewness of the signal varies depending on the pump status, with the healthy and defect with cavitation states showing left-skewed distributions

and the cavitation and defect states showing near-symmetrical or slightly right-skewed distributions. Figure 9 shows the RMS values for 2880 rpm.

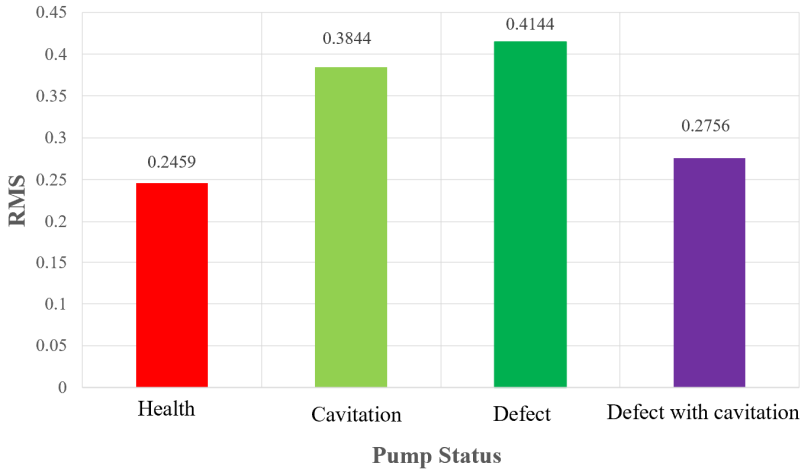


Figure 9. RMS values for different pump statuses.

The RMS value for the healthy pump status indicates a relatively lower value of the signal compared to the other states. The defect with a cavitation state has a higher RMS value. The RMS values in the chart provide a clear indication of the pump's condition. Lower RMS values suggest stable and healthy operation, while higher RMS values indicate increased signal power or intensity due to cavitation or defects.

Figure 10 shows the variance values for the obtained data.

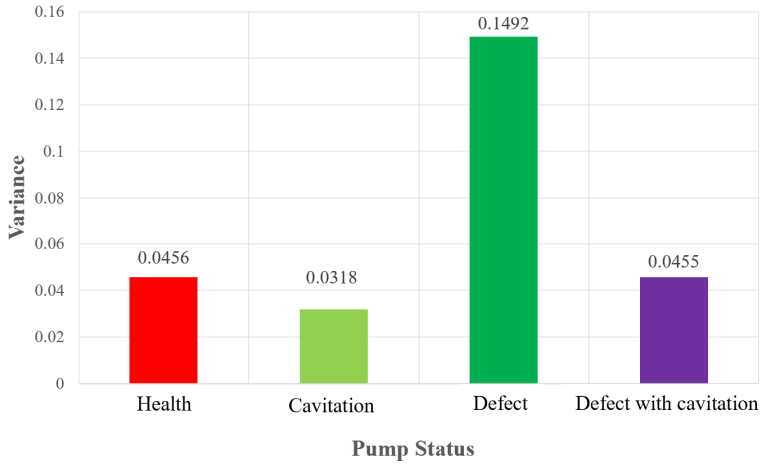


Figure 10. Variance diagram.

The variance value for the healthy pump status indicates a moderate level of spread in the signal values. The cavitation state shows a lower variance value, indicating less spread in the signal values compared to the healthy state. The defect state has the highest variance value, indicating a significant level of spread in the signal values. The defect with the cavitation state has a variance value like the healthy state, indicating a moderate level of spread in the signal values. This analysis suggests that the diffusion of signal values varies across different pump states, with the defect state showing the most significant variation, followed by the healthy and defect with cavitation states, and finally, the cavitation state with the least variation.

In general, the results of time-domain analysis are challenging, revealing that different pump states exhibit distinct statistical features. Specifically, the defect state shows the highest kurtosis and variance, indicating extreme and widely spread values, while the healthy

state maintains the lowest RMS and moderate variance, suggesting stable operation. These variations underscore the necessity of using multiple statistical parameters to effectively diagnose pump conditions. While time-domain analysis alone may not provide comprehensive insights into complex signal patterns, the statistical parameters calculated in this domain can be used to train artificial intelligence networks, such as ANNs, to establish suitable criteria for fault diagnosis in devices. It should be noted that one of the effective methods for fault diagnosis is through frequency-domain analysis. The next section will explore the analysis of data in the frequency domain.

3.2. Frequency Domain Analysis

In the frequency domain, the range of signals is shown in terms of amplitude and frequency. Therefore, with the help of fast Fourier transform, time-domain signals can be converted into a frequency spectrum. Fourier transform is given as:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \tag{10}$$

where f(t) is the time-domain data and F(ω) is Fourier transforms f(t) in the time domain.

Figure 11 shows the frequency spectrum of the installed sensor in a healthy state for a rotational speed of 2880 rpm. In this figure, the horizontal axis shows the frequency range in Hz, and the vertical axis shows the amplitude of the signal in decibel voltage (dbv).

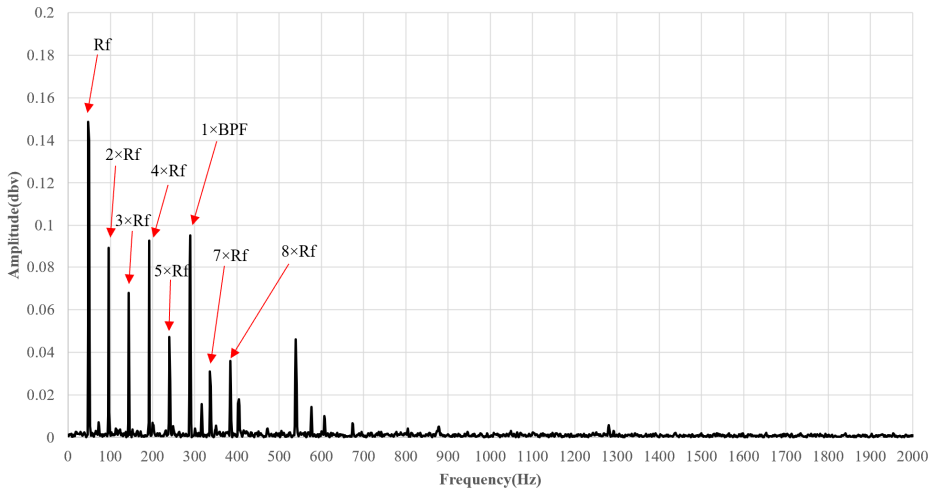


Figure 11. Frequency spectrum of the pump in a healthy state.

Pressure fluctuations can be detected at discrete frequencies that are multiples of the rotation frequency and the number of blades. These frequencies are also called blade passage frequencies (BPF). They are generated by the interaction of the rotating blades with a stationary component. Each blade passing a fixed point produces a distinct frequency, known as the blade pass frequency:

$$BPF = N \times Rf \tag{11}$$

In this equation, N is the number of blades, and Rf is the rotational speed in revolutions per second.

The amplitude of such pressure fluctuations depends on the number of blades, diffuser design parameters, and operational parameters. Figure 12 shows the periodicity in the centrifugal pump. It can be observed that two main dominant frequencies are present in the frequency range; the dominant frequency in this range is associated with the shaft rotating frequency (Rf), the blade passing frequency (BPF), and their harmonics. These effects result from the interaction between the impeller blades and the pumped flow, as well as the significant interaction between the impeller blades and the stationary components. Blade

passage frequencies appear at the lower end of the frequency spectrum. However, they are not evident for a specific number of blades in most spectra, which can be attributed to the unique vibration characteristics and design of the pump. Blade passage frequencies typically appear at the lower end of the spectrum, as they correspond to the fundamental mechanical vibrations of the rotating blades. Figure 12 shows the frequency spectrum of the pump in the state of cavitation.

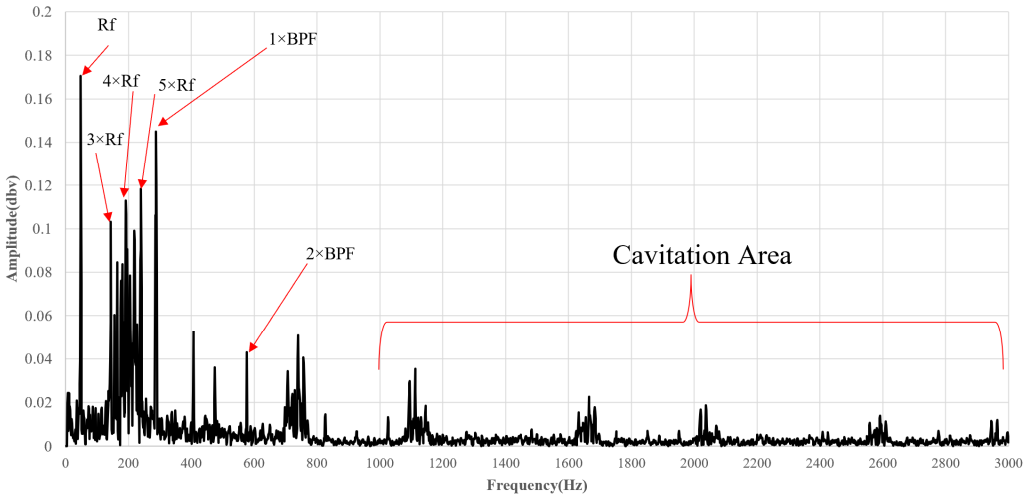


Figure 12. Frequency spectrum of the pump in a cavitation state.

From Figure 12, the first rotational frequency occurs at 48 Hz with an amplitude of 0.17 dbv. The first harmonic amplitude in healthy status was 0.148, so the amplitude of harmonics increased in this status. The amplitude of the first and second BPF in this status increased, and the amplitude of the other harmonics increased. In fact, more bubbles form and burst in the pump as cavitation progresses, which causes the pump to vibrate more intensely, leading to increased fluctuations in amplitude. Cavitation occurs in the high-frequency range and can be seen; the frequency peaks in the high-frequency range increase with the start of cavitation. In Figure 13, the frequency spectrum for the defective impeller is shown.

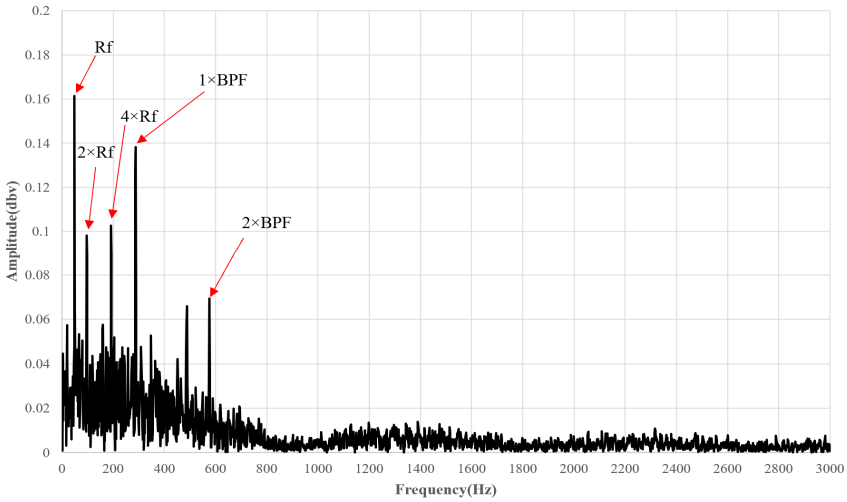


Figure 13. Frequency spectrum of the pump under a faulty impeller condition.

In Figure 13, the amplitude of 1×Rf and 2×Rf, compared to the healthy status, has increased. The BPF also increased in this status from 0.092 to 0.138 dbv. Additionally, the frequency range of the other harmonics has increased. A defective impeller introduces

additional mechanical vibrations and irregularities in the flow, and it is clear that many other frequencies are produced by various sources, including the driving motor and the bearing misalignment, as observed. It can be concluded that in a pump with a defective impeller, the maximum peak corresponding to the Rf and blade passing frequencies increases.

Figure 14 shows the frequency spectrum for the state of the defective impeller with cavitation. In this spectrum, the frequency harmonic is shown, along with the frequency range of cavitation.

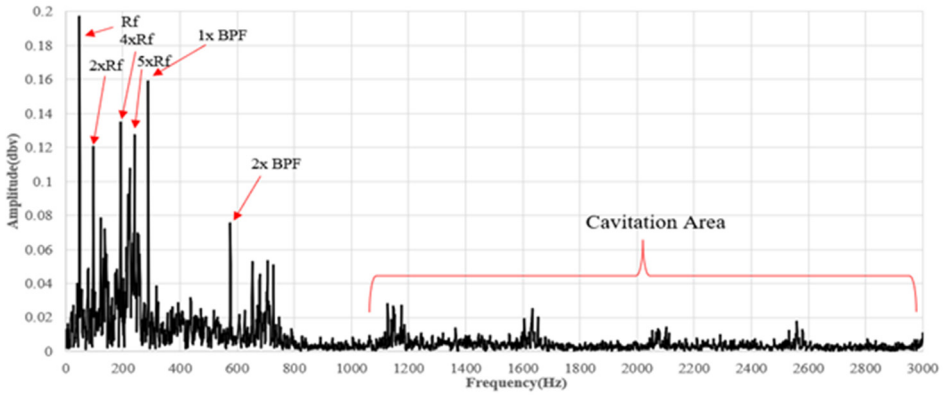


Figure 14. Frequency spectrum of the pump with cavitation-induced impeller damage.

In Figure 14, the amplitudes of the harmonics of 1× Rf, 2× Rf, 4× Rf, and 5× Rf have increased, and these harmonics have higher amplitudes compared to the cavitation and defect statuses. The BPF also increases compared to the defective and cavitation statuses. The combined effects of cavitation and a defective impeller significantly amplify across all of the vibration levels. The interaction between impeller defects and cavitation bubbles generates complex vibration patterns, leading to increased amplitude in both the low-frequency and high-frequency ranges. In fact, defects in the impeller cause turbulence, which alters the pressure contours on both the impeller and the pump body. This results in an increase in amplitude in the frequency domain.

3.3. Fault Diagnosis Using ANN

While the analysis of vibration signals in the frequency domain is effective for detecting faults in pumps, interpreting these signals accurately requires a comprehensive understanding of signal processing techniques and the operational dynamics of pumps. This interpretation becomes even more complex when considering signals across multiple domains, such as time, frequency, and time-frequency, which demand experience and expertise. To address the challenge of accurately diagnosing pump faults, an automatic, fast, and reliable troubleshooting method has been developed. ANNs can analyze vast amounts of vibration data, learn from patterns, and provide predictive maintenance insights, thereby reducing downtime and maintenance costs. By automating the signal interpretation process, these AI-based methods enhance the efficiency and reliability of pump condition monitoring, making it accessible, even to those with limited expertise in signal processing.

After trying different layers and making many try-and-error processes, the number of hidden layers was finally chosen to be ten layers. A linear transfer function was used for the output layer, and a hyperbolic tangent transfer function was used for the intermediate layer to spread the data regarding zeros and ones. To predict the output value from the primary matrix, which is a 5 × 20 matrix, a 5 × 16 matrix was assigned for training the neural network, and a 5 × 4 matrix that included the five statistical parameters for four states of the pump at specific speeds was assigned as a test matrix. The first column of this matrix includes a healthy impeller with cavitation at a rotational speed of 1000 rpm; the second column includes a defective impeller at a rotational speed of 2000 rpm; the third column includes a defective impeller with cavitation at a rotational speed of 2500 rpm; and the fourth column includes a healthy impeller at a rotational speed of 2880 rpm.

Figure 15 shows the regression percentages for eleven iterations of the neural network.

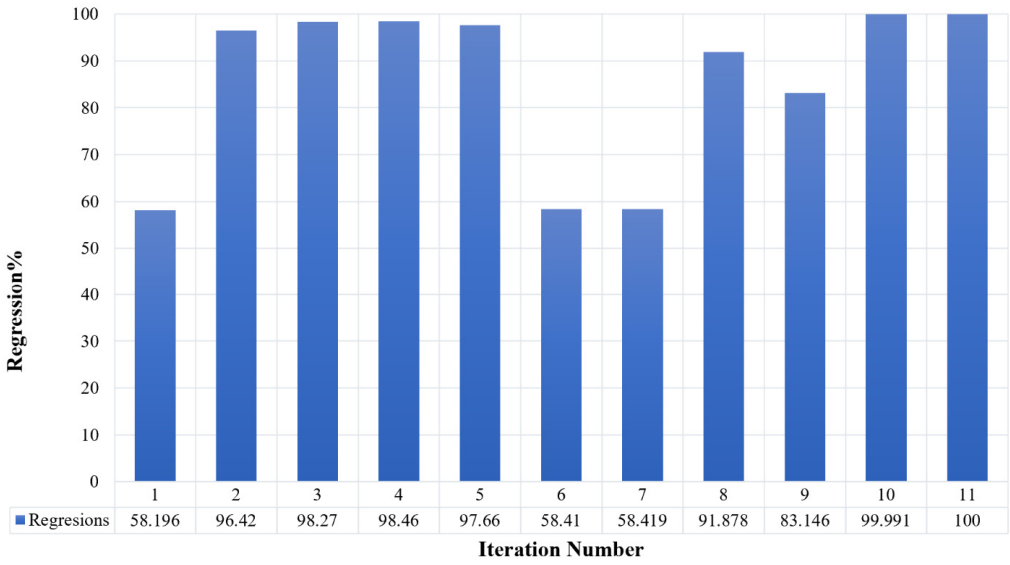
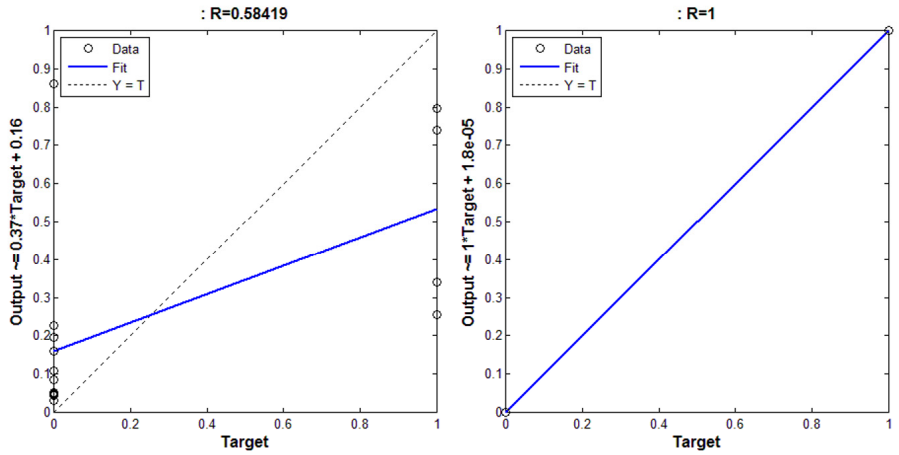


Figure 15. Percent accuracy of the trained network.

In Figure 15, the best regression results are observed for iterations 10 and 11. To ensure the accuracy of network prediction, the columns have been randomly extracted at different speeds. In Figure 16, the regression percentage is shown for seven and eleven repetitions.



(a) (b)

Figure 16. Regression graph: (**a**) seventh iteration and (**b**) eleventh iteration.

The charts represent the performance of an ANN in predicting target values. The horizontal axis shows the target values (T), while the vertical axis shows the output values produced by the ANN (Y). The open circular symbols in Figure 16 represent the target values during the neural network's training process. These symbols indicate the expected outputs for the various states of the pump being analyzed. This value represents the correlation coefficient, indicating the strength and direction of the linear relationship between the target and output values. An R-value of 0.58419 suggests a moderate positive correlation. The fit line equation is given as Output $\approx 0.37 \times$ Target + 0.16. This indicates that the ANN's output is only about 37% of the target value plus a small constant offset (0.16). The slope of 0.37 suggests the ANN is under-predicting the target values. In this figure, the dashed line represents the ideal scenario where the output perfectly matches the target (Output = Target). As is clear from Figure 15, the regression percentage for the

seventh iteration is 58.419%, and for the eleventh iteration, it is 100%. The eleventh iteration has the best performance, so it has been used to train data and predict faults in the network simulation part. The performance of the ANN based on the mean squared error is shown in Figure 17.

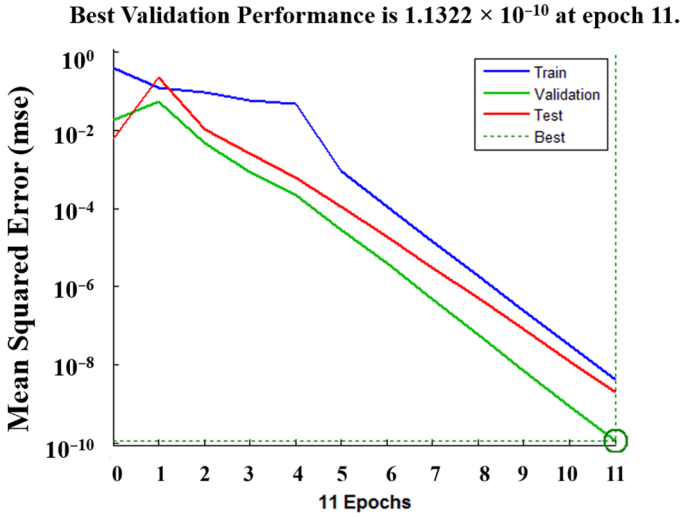


Figure 17. Mean square error for the training artificial neural networks.

In Figure 17, the horizontal axis (epochs) represents the number of training epochs, which is the number of times the learning algorithm has processed the entire training dataset. The mean squared error (MSE) in the vertical axis represents a measure of the difference between the predicted and actual values. A lower MSE indicates better model performance. In this chart, the blue line represents the MSE for the training dataset over the epochs, the green line represents the MSE for the validation dataset over the epochs, the red line (test) represents the MSE for the test dataset over the epochs, and the dashed green line represents the best validation performance achieved during the training process. The chart demonstrates effective training of the ANN, with the MSE decreasing consistently across training, validation, and test datasets. The model shows initial overfitting but quickly improves, achieving optimal performance by epoch eleven. As it is clear, the best validation performance is 1.1322×10^{-10} in epoch eleven. This interpretation indicates that the ANN training process is highly effective, with the model achieving excellent performance and minimal error by the final epoch.

To accurately predict faults in the pump, a neural network model was employed, consisting of three layers, ten neurons per layer, and trained over eleven iterations. The results of this neural network's prediction were then calculated. Subsequently, the final prediction matrix, sized 5×4, as mentioned earlier, was input into the algorithm to analyze and validate the outcomes. This approach not only ensured precise fault detection but also demonstrated the robustness and efficiency of using neural networks for predictive maintenance in pump systems. The results are shown in Figure 18.

In this diagram, the first column is a number close to zero, which indicates that the pump is defective (a healthy impeller with cavitation). The second column is a number close to zero, which indicates that the pump is faulty (a defective impeller). The third column is a number close to zero, which indicates that the pump is defective (a defective impeller with cavitation), but the fourth column is a number close to one, which indicates that the pump is healthy. As is evident, the neural network used has worked well. Therefore, it can be said that the designed neural network is successful in detecting the fault.

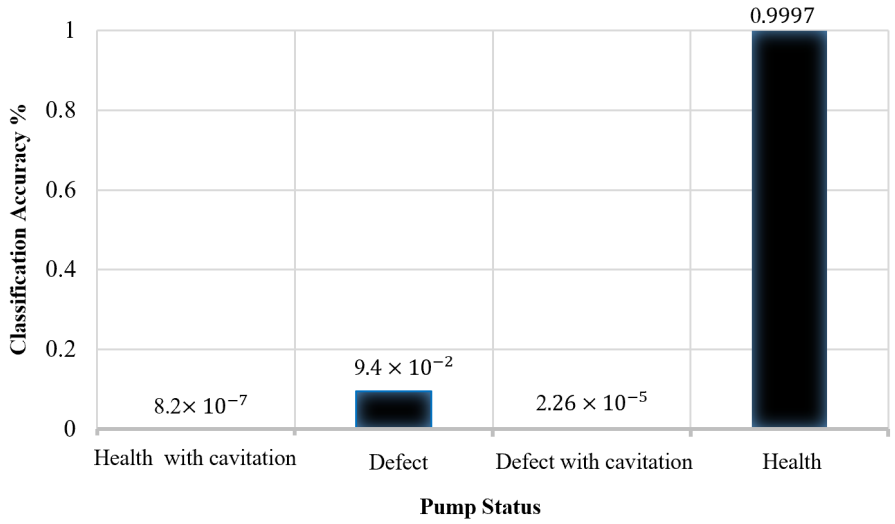


Figure 18. The result prediction of ANN.

4. Conclusions

In this study, the vibration analysis of a CP with a healthy and defective impeller, with and without cavitation, was performed at different rotational speeds in two domains of time and frequency. The data obtained from the time domain were used as input for the neural network to predict faults in the pump. The results prove that the data obtained from the normalized time domain do not provide interpretable information. Therefore, statistical factors such as the kurtosis, mean, and variance were calculated using these data to interpret the results. The calculated data reveal that the results are difficult to interpret and that, in certain situations, different pump statuses exhibit diverse statistical patterns. To proceed, the FFT diagrams of the frequency domain were used to identify the defect, and it was shown that all defects in the CP can be detected and diagnosed using these diagrams. Finally, the results of this research demonstrate that the MLP properties of vibration and current signals could indicate pump function under faulty conditions. The application of the MLP algorithm in this study has demonstrated its superiority. The key advantage of using MLP is its ability to process and learn from large volumes of time-domain data and automatically extract meaningful patterns. This reduces dependency on expert knowledge for signal interpretation and allows for real-time monitoring and diagnosis. The MLP's robustness in handling non-stationary and transient signals further emphasizes its potential as a powerful tool for predictive maintenance, enabling more accurate and timely interventions to prevent pump failures. To make the model more sensitive to the type of issue, we will attempt to investigate this in subsequent research. The current work's next phase is concentrated on creating an MLP that can classify all the defects into different groups. To enhance the algorithm's capability to differentiate between various faulty states, we plan to expand our training dataset to include a broader range of defect types. This can be achieved by modifying the ANN architecture to support multi-class output that allows the model to recognize and differentiate among multiple faulty conditions. The proposed method can also be used for frequency-domain analysis to investigate pump fault detection. This study focused on a specific notch defect in the impeller. Next, research could also focus on a wider range of defects, such as blade cracks, impeller pitting, and other mechanical failures, or changing the dimensions of defect on the impeller to assess how well the ANN generalizes to different fault conditions.

Author Contributions: Conceptualization, M.H.G. and M.K.; Writing—original draft preparation, M.R. and P.C.; Writing—review and editing, M.K. and R.F.; Investigation, M.H.G.; Visualization, M.R. and P.C.; Supervision, M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

ANN	Artificial neural network
BPF	Blade passage frequency
CP	Centrifugal pump
FNN	Feedforward neural network
GA	Genetic algorithm
LM	Levenberg–Marquardt
MLP	Multi-layer perceptron
OC	Oscilloscope card
PNN	Partially linearized neural network
RMS	Root mean square
SVM	Support vector machine
RF	shaft rotating frequency
WA	Wavelet analysis

References

1. Gülich, J.F. Pump Types and Performance Data. In *Centrifugal Pumps*; Springer: Cham, Switzerland, 2014; pp. 43–78.
2. Muralidharan, V.; Sugumaran, V.; Indira, V. Fault diagnosis of monoblock centrifugal pump using svm. *Eng. Sci. Technol. Int. J.* **2014**, *17*, 152–157. [CrossRef]
3. Muralidharan, V.; Sugumaran, V. A comparative study of naïve bayes classifier and bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. *Appl. Soft Comput.* **2012**, *12*, 2023–2029. [CrossRef]
4. Flett, J.; Bone, G.M. Fault detection and diagnosis of diesel engine valve trains. *Mech. Syst. Signal Process.* **2016**, *72–73*, 316–327. [CrossRef]
5. Moosavian, A.; Najafi, G.; Ghobadian, B.; Mirsalim, M.; Jafari, S.M.; Sharghi, P. Piston scuffing fault and its identification in an ic engine by vibration analysis. *Appl. Acoust.* **2016**, *102*, 40–48. [CrossRef]
6. Dolenc, B.; Boškoski, P.; Juričić, Đ. Distributed bearing fault diagnosis based on vibration analysis. *Mech. Syst. Signal Process.* **2016**, *66–67*, 521–532. [CrossRef]
7. Orhan, S.; Aktürk, N.; Çelik, V. Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: Comprehensive case studies. *NDT E Int.* **2006**, *39*, 293–298. [CrossRef]
8. Rizal, M.; Ghani, J.A.; Nuawi, M.Z.; Haron, C.H.C. Online tool wear prediction system in the turning process using an adaptive neuro-fuzzy inference system. *Appl. Soft Comput.* **2013**, *13*, 1960–1968. [CrossRef]
9. Xiang, J.; Zhong, Y.; Gao, H. Rolling element bearing fault detection using ppca and spectral kurtosis. *Measurement* **2015**, *75*, 180–191. [CrossRef]
10. Gan, M.; Wang, C.; Zhu, C. Multiple-domain manifold for feature extraction in machinery fault diagnosis. *Measurement* **2015**, *75*, 76–91. [CrossRef]
11. Mousmoulis, G.; Karlsen-Davies, N.; Aggidis, G.; Anagnostopoulos, I.; Papantonis, D. Experimental Analysis of Cavitation in a Centrifugal Pump Using Acoustic Emission, Vibration Measurements, and Flow Visualization. *Eur. J. Mech.-B/Fluids* **2019**, *75*, 300–311. [CrossRef]
12. Al-Obaidi, A.R. Investigation of effect of pump rotational speed on performance and detection of cavitation within a centrifugal pump using vibration analysis. *Heliyon* **2019**, *5*, e01989. [CrossRef] [PubMed]
13. Sun, H.; Luo, Y.; Yuan, S.; Yin, J. Hilbert spectrum analysis of unsteady characteristics in centrifugal pump operation under cavitation status. *Ann. Nucl. Energy* **2018**, *114*, 607–615. [CrossRef]
14. Murovec, J.; Čurović, L.; Novaković, T.; Prezelj, J. Psychoacoustic approach for cavitation detection in centrifugal pumps. *Appl. Acoust.* **2020**, *165*, 107323. [CrossRef]
15. Kumar, A.; Kumar, R. Time-frequency analysis and support vector machine in automatic detection of defect from vibration signal of centrifugal pump. *Measurement* **2017**, *108*, 119–133. [CrossRef]
16. 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* **2017**, *108*, 9–17. [CrossRef]

17. Birajdar, R.; Patil, R.; Khanzode, K. Vibration and Noise in Centrifugal Pumps: Sources and Diagnosis Methods. In Proceedings of the 3rd International Conference on Integrity, Reliability and Failure, Porto, Portugal, 20–24 July 2009; Volume 81, pp. 2631–2644.
18. Sakthivel, N.R.; Nair, B.; Sugumaran, V. Soft computing approach to fault diagnosis of centrifugal pump. *Appl. Soft Comput.* **2012**, *12*, 1574–1581. [CrossRef]
19. Askari, A.; Alekathir, J.; Mator, F.; Rezazade, A. Troubleshooting and diagnosis of cavitation in centrifugal pump using vibration analysis and providing solution to fix it. In Proceedings of the Second Technical Conference of Monitoring and Troubleshooting, Tehran, Iran, 24 February 2008. (In Persian).
20. Al-Braik, A.; Hamomd, O.; Gu, F.; Ball, A. Diagnosis of impeller faults in a centrifugal pump using vibration signals. *J. Sound Vib.* **2014**, *78*, 695–712.
21. Espadafor, F.J.; Villanueva, J.B.; García, M.T.; Trujillo, E.C. Experimental and dynamic system simulation and optimization of a centrifugal pump-coupling-engine system. part 1: Failure identification. *Eng. Fail. Anal.* **2011**, *18*, 1–11. [CrossRef]
22. Wang, H.; Chen, P. Intelligent diagnosis method for a centrifugal pump using features of vibration signals. *Neural Comput. Appl.* **2009**, *18*, 397–405. [CrossRef]
23. Barrio, R.; Fernandez, J.; Blanco, E.; Parrondo, J. Estimation of radial load in centrifugal pumps using computational fluid dynamics. *Eur. J. Mech.-B/Fluids* **2011**, *30*, 316–324. [CrossRef]
24. Casoli, P.; Pastori, M.; Scolari, F.; Rundo, M. A vibration signal-based method for fault identification and classification in hydraulic axial piston pumps. *Energies* **2019**, *12*, 953. [CrossRef]
25. Casoli, P.; Bedotti, A.; Campanini, F.; Pastori, M. A methodology based on cyclostationary analysis for fault detection of hydraulic axial piston pumps. *Energies* **2018**, *11*, 1874. [CrossRef]
26. Paliwal, M.; Kumar, U.A. Neural networks and statistical techniques: A review of applications. *Expert Syst. Appl.* **2009**, *36*, 2–17. [CrossRef]
27. Kutlu, Y.; Kuntalp, M.; Kuntalp, D. Optimizing the performance of an mlp classifier for the automatic detection of epileptic spikes. *Expert Syst. Appl.* **2009**, *36*, 7567–7575. [CrossRef]
28. Haykin, S. *Neural Networks*; Prentice Hall: Hoboken, NJ, USA, 1999.
29. Amari, S.I. Training error, generalization error, and learning curves in neural learning. In Proceedings of the 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, Dunedin, New Zealand, 20–23 November 1995; pp. 4–5.
30. Demuth, H.; Beale, M. *Neural Network Toolbox for Use with MATLAB*; The MathWorks Inc.: Natick, MA, USA, 1998; pp. 10–30.
31. Hagiwara, K.; Kuno, K. Regularization Learning and Early Stopping in Linear Networks. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000: Neural Computing: New Challenges and Perspectives for the New Millennium, Como, Italy, 27 July 2000; pp. 511–516.
32. Ranganathan, A. The Levenberg-Marquardt Algorithm. *Tutor. LM Algorithm* **2004**, *11*, 101–110.
33. Shankar, P.M. *Probability, Random Variables, and Data Analytics with Engineering Applications*; Springer: Berlin/Heidelberg, Germany, 2021; Volume 473.