

B.Tech. Project Report
On
**Development of a Digital Twin for a
Submersible Pump for Preventive Maintenance**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR
THE AWARD OF THE DEGREE OF
BACHELOR OF TECHNOLOGY

(Mechanical Engineering)

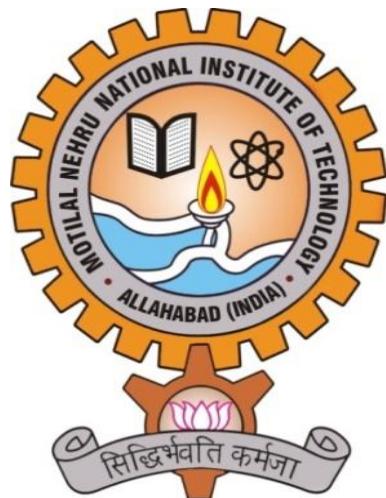
SUBMITTED BY

Vikas Gaur – 20213038

Uttam Kumar Gupta – 20213050

Shushanshu Kumar Bharti – 20213133

Umesh Kumar – 20213140



Mechanical Engineering Department
MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY
ALLAHABAD
PRAYAGRAJ - 211004, INDIA

December 2024

CANDIDATE'S DECLARATION

This is to certify that the work presented in the project report titled "**DEVELOPMENT OF A DIGITAL TWIN FOR A SUBMERSIBLE PUMP FOR PREVENTIVE MAINTENANCE**" is submitted in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology in Mechanical Engineering at **MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY, ALLAHABAD**. It is a genuine record of the work carried out by us under the guidance and supervision of Dr. Skylab Paulas Bhore during the period from July 2024 to December 2024. The content of this report has not been submitted to any other University or Institute for the award of any degree.

Vikas Gaur

Uttam Kumar Gupta

Shushanshu Kumar Bharti

Umesh Kumar

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Signature of Supervisor

Abstract

The submersible pump industry plays a crucial role in meeting the growing demand for efficient water management solutions. However, maintaining submersible pumps involves significant effort, particularly in monitoring their performance and identifying potential faults that may lead to failures. The aim of this project is to address the challenges of pump maintenance by developing a system based on a *Digital Twin* of the submersible pump, integrating advanced data acquisition and machine learning techniques to predict the remaining useful life of the pump.

This system is designed to tackle maintenance issues by employing an Arduino-based data acquisition system equipped with various sensors, including pressure sensors, current sensors, and flow sensors. These components enable real-time monitoring of pump performance, allowing for the identification of faults such as impeller and diffuser wear or blockages in the flow system.

The pressure, current, and flow sensors provide critical data inputs to assess the pump's operating conditions, while the Arduino-based control system acts as the central hub for data collection and processing. The collected data is used to train a machine learning model capable of detecting anomalies and predicting the pump's remaining useful life.

Faults such as impeller and diffuser damage or flow blockages are simulated and analyzed to understand their impact on the pump's performance. By combining these fault conditions with sensor data, the system can accurately predict degradation trends and support proactive maintenance decisions. This approach not only minimizes downtime but also ensures the longevity and efficiency of the submersible pump system.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the Director of Motilal Nehru National Institute of Technology (MNNIT), Allahabad, for giving us the opportunity to undertake this project. Our deepest gratitude goes to our advisor, Dr. Skylab Paulas Bhore, for his invaluable guidance, unwavering dedication, and meticulous attention throughout the entire project. We would also like to extend our utmost respect to the members of our B. Tech. Project Monitoring Committee, namely Prof. K.N. Pandey (Chairman), Dr. D.K. Shukla and Dr. Skylab Paulas Bhore, for their patience, continuous motivation, and diligent supervision, which greatly contributed to the successful completion of our project. Finally, we are thankful to Prof. K.N. Pandey, the Head of the Department, for providing us with the necessary resources and facilities to carry out our work.

Vikas Gaur – 20213038

Uttam Kumar Gupta – 20213050

Shushanshu Kumar Bharti – 20213133

Umesh Kumar – 20213140

TABLE OF CONTENT

Section	Title	Page No.
Candidate's Declaration	Candidate's Declaration	02
Abstract	Abstract	03
Acknowledgement	Acknowledgement	04
Chapter 1:	Introduction	09
1.1	Introduction: Submersible Pump Monitoring and Predictive Maintenance	09
1.2	How Predictive Maintenance Improves Submersible Pump	10
1.3	Our Project: Advancing Smarter Management for Submersible Pumps	10
1.4	Recent Updates and Consumer Base	10
1.4.1	Ongoing Research and Development	11
1.5	Our Project: A Step Towards Predictive Maintenance	11
Chapter 2:	Literature Review	13
2.1	Previous Studies	13
2.2	Objective	14
Chapter 3:	Project Work	16
3.1	Overview of the Submersible Pump's Hardware Components	16
	Data Collection in Healthy Condition	
3.2	Current Sensor and Installation on Test Setup	17
3.2.2	Flow Sensor and Installation on Test Setup	18
3.2.3	Pressure Sensor and Installation on Test Setup	19
	Calibration	
3.3.1	Calibration of Current Sensor	21
3.3.2	Calibration of Flow Sensor	22
3.3.3	Calibration of Pressure Sensor	23
3.4	Collection of Data in Normal/Healthy Condition	25
	Fault Creation	
3.5.1	Impeller Fault	25
3.5.2	Blockage Valve Fault	27
3.5.3	Diffuser Fault	27
3.5.4	Bush Bearing Fault	28
3.6	Data Collection in Faulty Condition	28
3.7	Fault Combination of Blockage, Impeller, Diffuser, and Bearing Fault	29
	Data Processing and Feature Engineering	
3.8	Data Preprocessing	30
3.9	Feature Engineering	32

3.9.1	Central Tendency Features	32
	Mean (Arithmetic Average)	32
	Median	32
	Mode	32
	Maximum	32
	Minimum Value	33
3.9.2	Dispersion Features	34
	Variance	34
	Standard Deviation	34
3.9.3	Signal Processing Features	35
	Root Mean Square (RMS)	35
	Crest Factor	36
	Skewness	36
3.10	Model Development	38
Chapter 4:	Results and Discussion	39
4.1	Data Analysis	39
4.2	Correlation Analysis	41
4.3	Significant Effect of Pressure, Current, and Flow	43
4.3.1	Effect of Pressure Value on Introducing Fault	43
4.3.2	Effect of Flow Rate on Introducing Fault	44
4.3.3	Effect of Current Value on Introducing Fault	44
4.4	Fault Creation and Related Table for Impeller, Diffuser, Bearing, and Blockage Valve	45
4.5	Machine Learning Model Preparation	48
4.5.1	Support Vector Classifier (SVC)	48
4.5.2	K-Nearest Neighbors (KNN)	48
4.5.3	Random Forest (RF)	48
4.5.4	XGBoost	49
4.6	Machine Learning Model Performance	49
4.6.1	Training, Testing, and Validation	52
4.6.2	Eight Features Plot Vs Sample Rate	53
4.7	Feature Plot of Accuracy in Random Forest ML Model	63
4.8	Performance and Confusion matrix	64
	Conclusion	70
References	References	73

LIST OF FIGURES

Fig No.	Name of Figure	Page No.
Fig 3.1	Block Diagram Submersible	18
Fig 3.2	Circuit Diagram of Current sensor (ACS712, 30A) with Arduino (Micro-Controller) For Data Acquisition	19
Fig 3.3	Circuit Diagram of Flow sensor (0-100L/min) with Arduino (Micro-Controller) For Data Acquisition	21
Fig 3.4	Circuit Diagram of Pressure sensor (0-100L/min) with Arduino (Micro-Controller) For Data Acquisition	22
Fig 3.5	Current Probe	24
Fig 3.6	Digital Flow Meter	25
Fig 3.7	Experimental Setup and Data Acquisition Using Arduino and Sensors	26
Fig 3.8	Represent the Healthy and Faulty condition of Impeller of Submersible Pump	27
Fig 3.9	Blockage Value with Degree of Rotation	28
Fig 3.10	Represent the Healthy and Faulty condition of Diffuser of Submersible Pump	29
Fig 3.11	Represent the Healthy, Carbon Brush and Fault condition	30
Fig 3.12	Representation of Outliner	33
Fig 3.13	Represent Maximum Value	35
Fig 3.14	Represent the minimum value	36
Fig 3.15	Represent Standard Deviation	37
Fig 3.16	Representation of Crest Factor	39
Fig 3.17	Represent the Skewness Nature	40
Fig 3.18	Breakdown of the main ML model architectures and tasks	41
Fig 4.3	The supervised ML workflow	52
Fig 4.4	Plot of hyperparameter for different Accuracy	63

LIST OF TABLES

Table of Contents

Table No.	Title	Page No.
Table 2.1	Reported investigations on the use of ML techniques for the prediction of Remaining life of Submersible Pump and related variables	15
Table 3.1	Specification of Current Sensor	19
Table 3.2	Specifications of Flow Sensors	21
Table 3.3	Specification of Pressure Sensor	23
Table 3.4	Value for the calibration of Flow sensor by Analytical Method	25
Table 3.5	Angle of Rotation of Blockage Valve	26
Table 3.6	Combinations of faults used in test cases	30
Table 4.1	Classification Report of Data	41
Table 4.2	Impeller Fault Parameter	45
Table 4.3	Impeller Fault Parameter	46
Table 4.4	Diffuser Fault Parameter	46
Table 4.5	Combination of Blockage, Impeller, Diffuser and Bearing Fault	47
Table 4.6	Blockage Fault Parameter	47
Table 4.7	Comparison of Performance Matrix for Flow Rate	64
Table 4.8	Hyperparameter for Flow Rate	64
Table 4.9	Performance Matrix of Different Algorithms for Current	67
Table 4.10	ML Model with Their Hyperparameters	67
Table 4.11	Performance Matrix of Different Algorithms for Combination	68
Table 4.12	ML Model with Their Hyperparameters	68

Chapter – I

INTRODUCTION

1.1 Introduction: Submersible Pump Monitoring and Predictive Maintenance

Submersible Pumps: A Versatile Powerhouse

Submersible pumps are essential tools widely used in modern industries and homes for handling fluids efficiently. They are versatile, playing a critical role in tasks like providing water for households, managing wastewater, and supporting industrial operations. These pumps are designed to work while being fully submerged in the fluid they pump, making them highly effective and reliable.

One major advantage of submersible pumps is that they don't require a separate system to prime them, which simplifies their operation and reduces maintenance needs. By being submerged, these pumps also avoid many issues like air leaks that can disrupt traditional pumps. This makes them especially useful in scenarios where consistent and dependable performance is necessary.

Overall, submersible pumps save time and effort while providing reliable fluid handling solutions, making them a popular choice in a wide range of applications, from domestic water systems to complex industrial processes.

Why Submersible Pumps Are Becoming More Important

Submersible pumps are becoming more essential as cities grow and industries expand. With more people needing water and climate challenges increasing, having reliable water systems is more important than ever. These pumps help by making it easier to draw water from the ground, supply clean water to homes, and handle wastewater effectively. Their ability to work underwater makes them a key solution for modern water management needs.

Enhancing Submersible Pump Reliability with Predictive Maintenance

Submersible pumps are vital in various industries, but like any mechanical equipment, they face challenges such as wear and tear. Ensuring their reliability is critical for uninterrupted operations and cost savings. Here's a detailed look at how predictive maintenance is transforming the way submersible pumps are maintained:

Challenges in Traditional Maintenance Approaches

1. Wear and Tear Over Time:

- Submersible pumps operate in challenging environments, making them prone to issues like impeller erosion, seal failure, and motor damage.

2. Costly Failures:

- When pumps fail unexpectedly, it can lead to downtime, production losses, and high repair costs.

3. Reactive or Time-Based Maintenance:

- Traditional methods often follow a fixed schedule or wait until a failure occurs, which may result in over-maintenance or delayed repairs, leading to inefficiency.

Predictive Maintenance: Predictive maintenance uses advanced technologies like sensors, data analytics, and machine learning to monitor equipment performance and predict potential issues before they occur.

1.2 How Predictive Maintenance Improves Submersible Pumps

1. Real-Time Monitoring:

- Sensors collect data on critical parameters such as pressure, flow rate, temperature, and current usage.
 - This data is analysed to detect anomalies that indicate early signs of wear or potential faults.
- 2. Early Fault Detection:** Issues like blockages, impeller wear, or motor inefficiency can be identified early, allowing for timely intervention.
 - 3. Reduced Downtime:** By addressing problems before they escalate, predictive maintenance minimizes unexpected breakdowns, keeping operations running smoothly.
 - 4. Optimized Maintenance Schedules:** Instead of following a fixed schedule, maintenance is performed only when needed, saving time and resources.
 - 5. Cost Savings:** Preventing major failures reduces repair costs and extends the lifespan of pumps, leading to long-term savings.
 - 6. Enhanced Pump Lifespan:** Regular monitoring and timely interventions ensure that the pump operates efficiently for a longer period.
 - 7. Data-Driven Insights:** Over time, data collected from sensors can provide valuable insights into performance trends, helping improve pump design and maintenance strategies.

1.3 Our Project: Advancing Smarter Management for Submersible Pumps

Our project focuses on improving the efficiency and reliability of submersible pumps by creating a smart monitoring and maintenance system. This system is designed to bring together advanced sensor technology and intelligent machine learning algorithms, offering real-time insights into the pump's health and performance.

The primary goal is to help users detect issues early and take preventive measures before they turn into costly failures. By analysing key parameters like pressure, flow, and electrical usage, the system will identify irregularities that may indicate wear, blockages, or other faults. This allows for timely interventions, reducing unexpected breakdowns and saving both time and money.

Additionally, the system is designed to optimize maintenance schedules. Instead of following a fixed maintenance routine, users can address specific needs based on the pump's condition, minimizing unnecessary downtime and extending its lifespan.

Our project not only aims to make pump maintenance more efficient but also contributes to reducing operational costs and improving overall performance. With this technology, we hope to set a new standard for managing submersible pumps in a smarter, more proactive way.

1.4 Recent Updates and Consumer Base

In recent years, technological advancements have significantly enhanced the capabilities and reliability of submersible pumps. Manufacturers have introduced innovative features like energy-efficient motors, advanced materials, and intelligent control systems to optimize performance and minimize environmental impact. As a result, the demand for submersible

pumps continues to grow, driven by expanding urbanization, industrialization, and agricultural activities.

The consumer base for submersible pumps is diverse, encompassing:

- **Domestic Households:** For water supply in rural and urban areas, drainage of basements and cellars, and irrigation of gardens.
- **Agriculture:** For irrigation of crops, livestock watering, and drainage of fields.
- **Industry:** For water supply, wastewater treatment, mining operations, and oil and gas extraction.
- **Construction:** For dewatering construction sites and tunnel projects.

1.4.1 Ongoing Research and Development

Researchers and engineers are actively exploring innovative solutions to improve the performance, reliability, and energy efficiency of submersible pumps. Key areas of research include:

- **Advanced Materials:** Developing materials that are resistant to corrosion, abrasion, and extreme temperatures to enhance pump durability.
- **Intelligent Control Systems:** Implementing advanced control algorithms to optimize pump operation, reduce energy consumption, and detect early signs of failure.
- **Real-time Monitoring and Diagnostics:** Utilizing sensor technologies to monitor critical parameters like pressure, flow rate, vibration, and temperature, enabling predictive maintenance and fault diagnosis.
- **Energy Efficiency:** Designing pumps with higher efficiency motors, variable speed drives, and hydraulic optimization techniques to minimize energy consumption.
- **Environmental Impact:** Developing eco-friendly pumps with reduced noise levels and minimal environmental footprint.

1.5 Our Project: A Step Towards Predictive Maintenance

In this project, we aim to contribute to the advancement of submersible pump technology by developing a predictive maintenance system. By leveraging the power of data science and machine learning, we seek to improve the reliability and lifespan of these vital machines.

Our system incorporates the following key components:

1. Data Acquisition:

- **Pressure Sensor:** Measures the pressure generated by the pump, providing insights into its performance and potential issues.
- **Current Sensor:** Monitors the electrical current drawn by the motor, indicating its health and efficiency.
- **Flow Sensor:** Tracks the flow rate of the fluid being pumped, helping to identify any blockages or leaks.

2. Fault Creation and Data Collection:

- We simulate various fault conditions, such as increased vibration, reduced flow rate, and power fluctuations, to generate a comprehensive dataset.
- The sensors continuously collect data on the pump's performance under normal and faulty conditions.

3. Machine Learning Model Development:

- We employ advanced machine learning algorithms to analyze the collected data and identify patterns that correlate with specific faults or impending failures.
- The model is trained on the labeled dataset to learn the relationships between sensor readings and potential issues.

4. Predictive Maintenance Strategy:

- The trained model is used to predict the remaining useful life of the pump based on real-time sensor data.
- Early warning signals are generated to alert operators of potential failures, allowing for timely maintenance and preventing unexpected downtime.

By combining data-driven insights with advanced machine learning techniques, our project aims to provide a valuable tool for both household and industrial users of submersible pumps. This system can help optimize maintenance schedules, reduce operational costs, and minimize environmental impact, ultimately contributing to a more sustainable and efficient future.

Chapter - 2

LITERATURE REVIEW

2.1 Previous Studies

This study introduces an innovative approach to developing a preventive maintenance system for submersible pumps using advanced sensor technology and data acquisition techniques. The authors propose the integration of pressure, flow, and current sensors to monitor the operational health of submersible pumps, enabling early detection of anomalies such as blockages, wear in impellers and diffusers, and motor inefficiencies. The paper elaborates on the hardware configuration, software implementation, and fault-detection algorithms designed to predict potential failures. Experimental results validate the system's accuracy and effectiveness in identifying early warning signs of pump degradation. The study highlights the advantages of using these sensors, including their precision, cost-effectiveness, and adaptability to diverse operating conditions. The authors conclude that the system offers practical and economical benefits, making it ideal for a wide range of applications.

Another area explored is the role of machine learning in enhancing predictive maintenance strategies. The review examines how data from installed sensors can be leveraged to train machine learning models for fault prediction and remaining useful life (RUL) estimation. Two prominent approaches are discussed: supervised learning, where historical sensor data with labeled faults is used to train models, and unsupervised learning, which detects anomalies by analyzing deviations from normal operational patterns. The paper highlights the strengths of these methods, such as improving predictive accuracy and reducing reliance on fixed maintenance schedules. Limitations, such as the need for high-quality data and computational resources, are also addressed.

By analyzing existing research, the review underscores the transformative potential of integrating sensors, data acquisition, and machine learning for submersible pump maintenance. This combined approach enables efficient monitoring, proactive interventions, and optimized maintenance schedules, reducing downtime and operational costs. Future research directions include improving sensor robustness, refining algorithms for better fault classification, and exploring the application of real-time monitoring systems in diverse industrial settings.

Ref. No.	Author and year	Title
[1]	Fernandes et al. (2020)	ML algorithms to predict boiler failures Seven days in advance.
[2]	Mousmoulis et al. (2017) investigates the process of cavitation in centrifugal pumps, analyzing flow images, vibrations, and acoustic emissions from three different impellers.	ML applications for damage detection in centrifugal pumps.
[3]	Dr. Sharda Chhabria, Rahul Ghata, Varun Mehta, Ayushi Ghosekar (2022)	Predictive maintenance using machine learning on water pump.
[4]	Jessica Brasil, Carla Maitelli, Joao Nascimento, Osvaldo Maitelli-Filho and Edney Galvao	Diagnosis of Operating Conditions of the Electrical Submersible Pump via Machine Learning

Table 2.1. Reported investigations on the use of ML techniques for the prediction of Remaining life of Submersible Pump and related variables.

Importance of Machine Learning in Submersible Pump Maintenance

1. **Real-Time Fault Detection:** Machine learning analyses sensor data, like pressure and flow, to identify anomalies early, preventing unexpected breakdowns.
2. **Predicting Remaining Useful Life (RUL):** ML models estimate how long pump components will last, allowing timely maintenance and reducing downtime.
3. **Optimized Maintenance Schedules:** ML ensures maintenance is performed based on actual pump conditions, saving time and resources compared to fixed schedules.
4. **Cost Efficiency:** Early detection of issues and optimized maintenance reduce repair costs and extend the pump's lifespan, improving overall operational efficiency.

2.2 OBJECTIVE

The primary goal of this research is to make a digital twin of submersible pump. The specific objectives include

- To Develop an ESP test setup to examine and improve its preventive maintenance
- Collection of data from submersible pump test setup in lab
- Development of data driven ML model
- Prediction of faults and remaining life of System using ML.
- Develop an accurate digital twin model using machine learning to simulate the behavior of the physical asset.

- Integrate the physical twin and digital twin through an IoT network
- Indication of Fault in virtual Model.

Chapter – 3

Project Work

3. Methodology

The methodology adopted in this study focuses on developing a predictive maintenance model for submersible pumps using data acquired from **pressure, flow, and current sensors**. The approach involves the following stages:

3.1 Overview of the Submersible Pump's Hardware Components:

Overview of the Submersible Pump's Hardware Components:

1. 1 HP Motor:

The submersible pump is powered by a 1 HP electric motor, which drives the pump to circulate water. The motor is controlled through an Arduino-based system for monitoring and automation.

2. Arduino:

An Arduino microcontroller is used as the central control unit for the submersible pump. It interfaces with sensors (pressure, current, and flow) and processes the data to manage the pump's performance, ensuring optimal operation.

3. Pressure Sensor:

A pressure sensor is employed to monitor the pressure within the pump system. It provides real-time data to the Arduino, enabling it to detect any variations in pressure, which could indicate blockages or irregular functioning.

4. Current Sensor (ACS712):

The ACS712 current sensor monitors the current drawn by the motor. This data helps in detecting overcurrent conditions, indicating motor malfunction, or ensuring efficient operation by measuring the load.

5. Flow Sensor:

A flow sensor measures the water flow rate being pumped. The Arduino uses this data to ensure the desired flow rate is maintained, adjusting the motor's speed or shutting it off if necessary.

6. Blockage Valve (Rotatable):

A rotatable blockage valve is integrated into the system to control water flow and detect faults. The valve can rotate from **0° (fully open)** to **90° (fully closed)**. The Arduino adjusts the valve position to address any detected blockage, ensuring the system can respond appropriately by regulating the flow or shutting down the pump to prevent damage.

Together, these components form an integrated system that enables real-time monitoring, control, and fault detection for the submersible pump, ensuring smooth operation and early detection of issues like blockages or malfunctions.

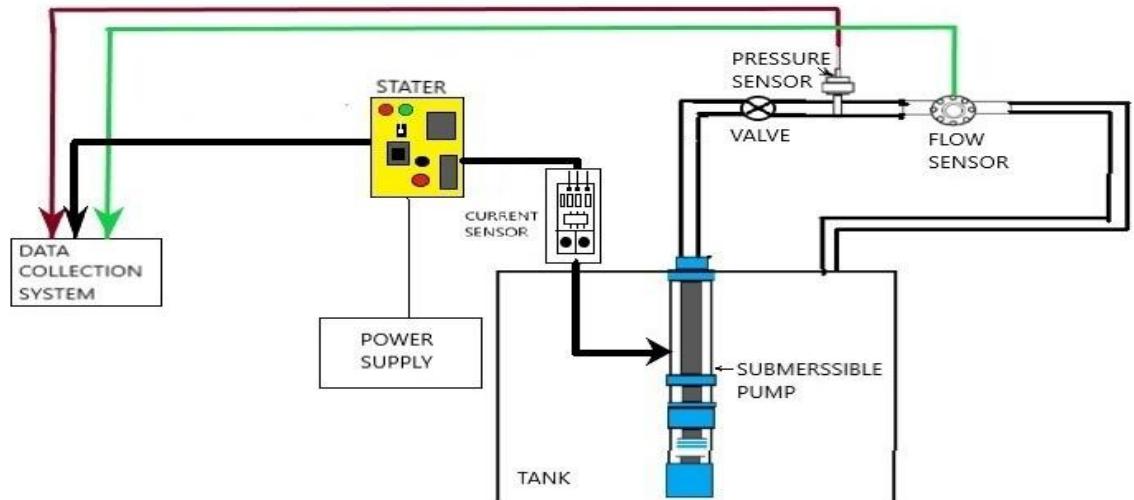


Fig 3.1: Block Diagram Submersible Pump Hardware Component with Experimental Setup

3.2 Data Collection in Healthy Condition:

Sensors: Pressure, flow rate, and motor current sensors were installed to monitor the pump's operational parameters in real-time.

Data Logging: Sensor readings were captured continuously using a data acquisition system (DAQ), ensuring synchronized collection of all three parameters.

Operational Context: Data was collected under varying load conditions, runtime durations, and environmental influences to capture a comprehensive range of operational scenarios and fault conditions.

3.2.1 Current Sensors and Installation on Test Setup

The ACS712 is a fully integrated, linear current sensor that uses a Hall Effect to measure AC or DC current. It's a popular choice for use with Arduino because it's easy to use and provides accurate results.

Working: The ACS712 uses a copper conduction path to measure current. The current flowing through the copper path generates a magnetic field that is sensed by the Hall IC. This magnetic field is converted into a proportional voltage.

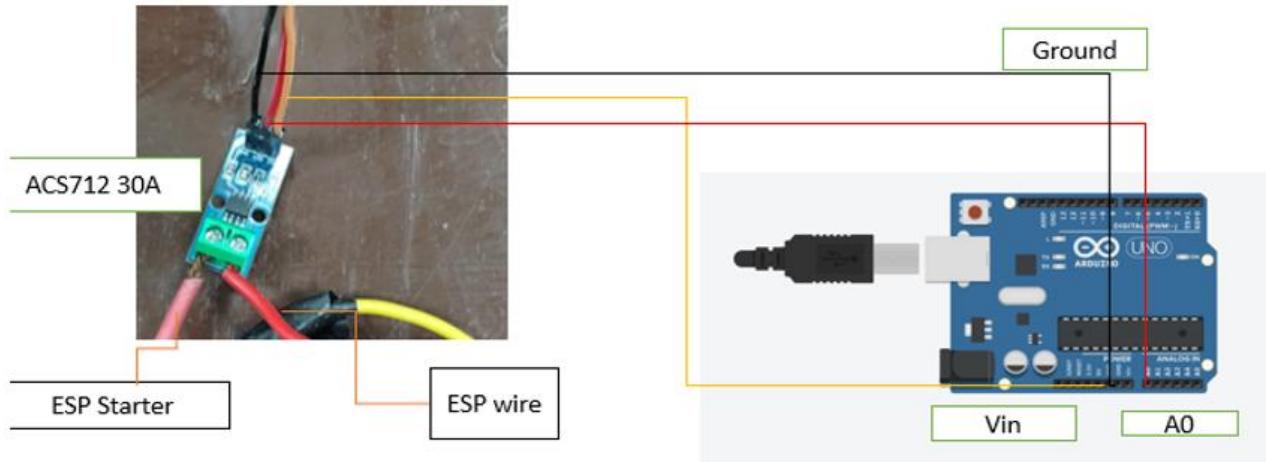


Fig 3.2 : Circuit Diagram of Current sensor (ACS712, 30A) with Arduino (Micro-Controller) For Data Acquisition

Parameter	Value
Current Range	0 – 30 A
Supply Voltage	4.5 – 5.5 V
Sensitivity	66-185mV/A

Table 3.1: Specification of Current Sensor

Connection with Arduino

Motor Starter to ACS712:

- Connect the live (phase) wire from the motor starter to IP+ of the ACS712.
- Connect IP- of the ACS712 to the live wire leading to the motor.

ACS712 to Arduino:

- VCC → Arduino 5V
- GND → Arduino GND
- OUT → Arduino A0

3.2.2 Flow Sensor and Installation on Test Setup

A flow sensor, specifically a turbine-type, measures the flow rate of liquids. It consists of:

- Rotor (Turbine Blades): Positioned in the flow path, these blades rotate as the fluid flows through.
- Housing: Typically made of durable materials like plastic, stainless steel, or brass, it encases the turbine and directs the fluid.
- Magnetic or Optical Sensor: Detects the rotation of the turbine blades. Each rotation corresponds to a specific volume of fluid, allowing calculation of the flow rate.

The sensor outputs electrical pulses proportional to the flow rate, used for monitoring or control applications.

Working Principle:

- The turbine inside the flow sensor spins as fluid flows through it.
- A magnetic or optical sensor detects the turbine's rotation, generating electrical pulses.
- The pulse frequency is proportional to the flow rate.

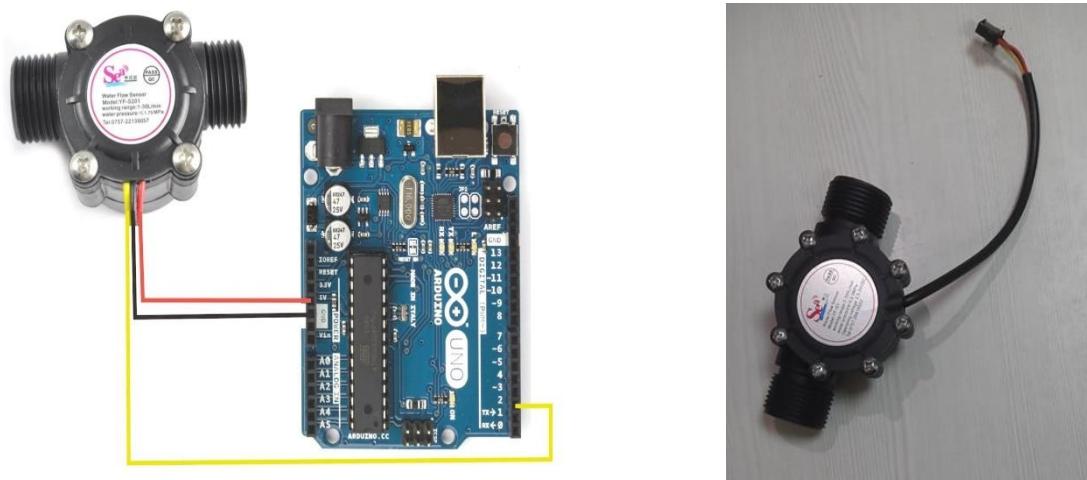


Fig 3.3: Circuit Diagram of Flow sensor (0-100L/min) with Arduino (Micro-Controller) For Data Acquisition

Parameter	Value
<u>Range</u>	<u>0 – 100 L/min</u>
<u>Water Pressure</u>	<u>< 2.4MPa</u>
<u>Operating Voltage</u>	<u>3.5 – 24 V</u>

Table 3.2: Specifications of Flow Sensors

Connection with Arduino:

- Connect the **VCC** and **GND** of the flow sensor to the Arduino's **5V** and **GND** pins.
- Connect the **signal pin** of the flow sensor to a digital pin (e.g., **A2**) on the Arduino.
- Use an external pull-up resistor if required and ensure proper power supply for accurate readings.

3.2.3 Pressure Sensor and Installation on Test Setup

A pressure sensor measures the force exerted by a fluid (liquid or gas) on a surface, typically used for monitoring pressure levels in systems.

Water pressure sensors are typically **piezoelectric**, **strain gauge**, or **capacitive** type, depending on the application and accuracy requirements.

Piezoelectric Pressure Sensor:

- Uses piezoelectric materials that generate an electrical charge when subjected to pressure.
- Often used for dynamic pressure measurement, such as in water flow or shock-wave applications.

Working: A piezoelectric pressure sensor generates an electrical charge when pressure is applied to a piezoelectric material (e.g., quartz or ceramic). The pressure causes deformation in the material, resulting in a voltage or charge proportional to the applied pressure. This charge is then amplified and converted into a measurable output signal. These sensors are ideal for dynamic pressure measurements and offer high sensitivity.

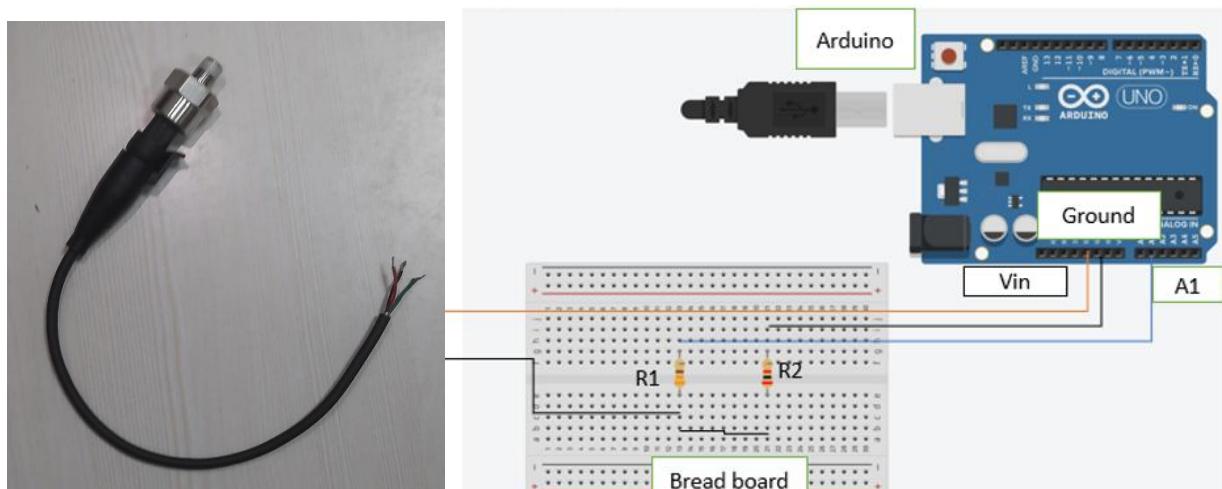


Fig 3.4: Circuit Diagram of Pressure sensor (0-100L/min) with Arduino (Micro-Controller) For Data Acquisition

Parameter	Value
Operating Voltage	5 ± 0.5 V DC
Operating Current	< 10mA
Working Pressure Range	0 – 1.6 MPa
Maximum pressure	3Mpa
Response time	< 2ms

Table 3.3: Specification of Pressure Sensor

Connection with Arduino:

- Connect **Vin** and **Ground** pins of the Arduino to the breadboard's power rails.
- Connect the **A1** pin on Arduino to the middle row of the breadboard where **R2** is connected.
- Use **R1** and **R2** as resistors to form the required circuit for your sensor, ensuring proper connections to the breadboard also ensure proper power supply for accurate readings.

3.3 Calibration

Calibration is the process of adjusting and verifying the accuracy of a device or instrument by comparing its measurements to a known standard or reference. This ensures the device provides reliable and precise results within the specified range.

3.3.1 Calibration of Current Sensor

We have used two methods to calibrate the current sensor **standard calibration**, performed under controlled conditions using precise reference equipment, and **field calibration**, conducted on-site to ensure accurate performance in real-world applications.

- a) **Field Calibration:** Conducted directly. Using Current Probe



Fig 3.5: Current Probe

- b) **Analytical Method:** The calibration of the current sensor was performed using the analytical method, which involves applying the basic power formula: **Power = Voltage × Current**. This approach ensures precise measurement by correlating the sensor's output with calculated values.

Example to Understand Calibration:

Let us understand this with an example:

- The voltage supplied is **220 V**, and the current drawn by the motor is **6.5 A**.
- The **input power** (power supplied) is calculated using the formula:

$$\text{Input Power} = \text{Voltage} \times \text{Current}$$

$$\text{Input Power} = 220 \times 8.5 = 1430 \text{ W}$$

Since the motor is an electrical component, it has a **power factor**, just like all other electrical appliances, and the pump's efficiency also comes into play.

- The **output power** (useful power delivered) is calculated using the formula:
Output Power = Voltage × Current × Power Factor × Efficiency
- **Power Factor:** Accounts for the phase difference between voltage and current in an AC system.

Substituting the values:

$$\text{Output Power} = 220 \times 6.5 \times 0.87 \times 0.6 = 750.805 \text{ W}$$

This calculation highlights the importance of considering both power factor and efficiency when analysing the performance of electrical devices like pumps.

Factors Affecting the value of power factor and Efficiency:

1. The power factor of a submersible motor changes over time due to several factors, primarily related to wear, environmental conditions, and operational changes.
2. Without proper maintenance and care, the power factor of a submersible motor noticeably deteriorates after **2–5 years**, leading to reduced efficiency and higher energy costs.
3. The **efficiency of a submersible motor** also changes over time, primarily due to aging, wear, and environmental factors.
4. The efficiency of a submersible motor typically deteriorates gradually, becoming significant over **3–7 years** if proper maintenance isn't performed. This leads to higher energy consumption and reduced performance.

5. The efficiency and power factor of submersible pumps can degrade over time due to factors such as insulation wear, overheating, increased resistance in windings, or improper load conditions.
6. The typical power factor of submersible pumps is usually between 0.6 and 0.9 [1], while their efficiency ranges from 50% to 70% [2], depending on the operating conditions and design. Factors such as motor load, cable length, and water pressure impact these values over time.

3.3.2 Calibration of Flow Sensor

- a) **Field Calibration:** Conducted directly. By using Flow meter
- b) **Analytical Method:** which involves considering head loss under varying operational conditions. This approach ensures precise measurement by correlating the sensor's output with theoretical flow rate calculations derived from fluid dynamics principles.



Fig 3.6: Digital Flow Meter

Head loss was analysed under different pressure levels and pipe conditions to achieve accurate calibration for real-world applications.

Tank Volume: 1000 liters (1 m³)

Pipe Length: 1 meter

Pipe Diameter: 1 inch (0.0254 meters)

Velocity Calculation: $v=Q/A$, Where A = is the cross-sectional area of the pipe.

Fill Time Calculation: $t=V/Q$, Where V = volume and Q = Discharge

Head Loss Calculation: $hf=(fLv^2)/2gD$, Where f =Darcy's Friction Factor

Flow Rate (L/min)	Velocity (m/s)	Head Loss (m)	Fill Time (min)	Darcy Friction Factor	Cross-Sectional Area (m ²)	Volume (L)
65.8	2.164	0.352	15.2	0.0375	0.000507	1000
67.3	2.214	0.369	14.86	0.0375	0.000507	1000
68.8	2.263	0.385	14.53	0.0375	0.000507	1000
70.3	2.312	0.402	14.22	0.0375	0.000507	1000
71.8	2.362	0.42	13.93	0.0375	0.000507	1000

73.3	2.411	0.437	13.64	0.0375	0.000507	1000
------	-------	-------	-------	--------	----------	------

Table 3.4: Value for the calibration of Flow sensor by Analytical Method

3.3.3 Calibration of Pressure Sensor

During the calibration process, we used the following equipment:

Pressure Source: A submersible pump providing adjustable water pressure.

Reference Standard: A calibrated pressure gauge with traceable accuracy.

Control Mechanism: Pressure regulating valves for precise control of pressure levels.

Sensor Under Test (pressure sensor): The pressure sensor attached to the submersible pump output.

Procedure

1. We first ensured that the submersible pump and pressure system were free of leaks and air bubbles. After verifying the integrity of the system, we powered on the pressure sensor and allowed it to stabilize before proceeding.

2. We opened all valves to release any residual pressure and brought the system to atmospheric pressure. After confirming the reference gauge read zero, we adjusted the pressure sensor's zero point to match the reference gauge reading.

3. We gradually increased the pressure from the zero-pressure point using the control valve. At each calibration interval (10%, 25%, 50%, 75%, and 100% of the sensor's range), we recorded the readings from both the reference gauge and the pressure sensor. We allowed sufficient time for pressure stabilization at each point before recording the values.

4. After reaching the maximum pressure, we gradually decreased the pressure back to the minimum level. We recorded readings at the same intervals used during the increasing pressure calibration to check for hysteresis effects.

5. we repeated the entire process of increasing and decreasing the pressure. Consistency between readings in successive cycles confirmed the reliability of the sensor's calibration.

3.4 Collection of Data in Normal/Healthy Condition

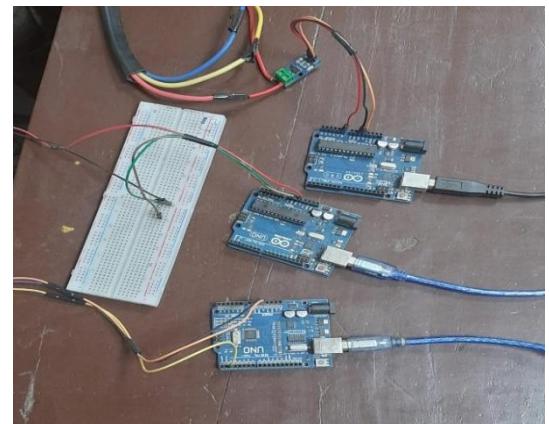


Fig 3.7: Experimental Setup and Data Acquisition Using Arduino and Sensors

3.5 Fault Creation:

3.5.1 Impeller:

Definition: A rotating component of a pump or compressor that moves fluid by converting rotational energy into flow energy.

Working Principle: It uses centrifugal force. As the impeller spins, it accelerates the fluid outward, increasing its velocity and pressure. This motion directs the fluid into the desired direction, typically through a pump or pipeline.

- The pump consists of 8 impellers
- Images contain the Back side, Front side and Faulty condition in Impeller



Fig 3.8.1: Back Side



Fig 3.8.2: Front Side



Fig 3.8.3: Stacks of Impeller and Diffuser



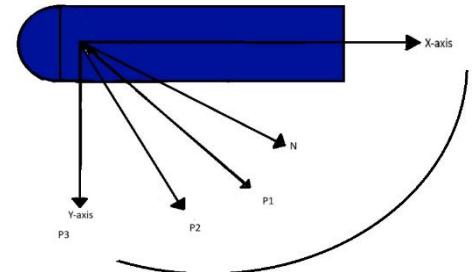
Fig 3.8.4: Fault Created

Fig 3.8: Represent the Healthy and Faulty condition of Impeller of Submersible Pump

3.5.2 Blockage Valve



blockage valve



Code	Condition
N	Close Upto 0 deg
B1	Close Upto 40 deg
B2	Close Upto 80 deg

- Purpose:** Used to control the water flow and prevent backflow or overpressure in the pipeline.
- Operation:** Can be manually or automatically controlled based on system requirements.
- Importance:** Helps in maintaining desired water pressure and protecting the pump from damage due to overloading.

3.5.3 Diffuser

Definition: A stationary component that slows down the fluid velocity and converts kinetic energy into pressure energy.

Working Principle: As water exits the impeller at high velocity, the diffuser guides and decelerates it, reducing turbulence and increasing pressure efficiently before directing it to the pump outlet.

- The pump consists of 8 Diffusers
- Images contain the Back side, Front side and Faulty condition in Diffusers



Fig 3.10.1: Back Side



3.10.2: Front side



Fig 3.10.3:Fault Creation

Fig 3.10: Represent the Healthy and Faulty condition of Diffuser of Submersible Pump

3.5.4 Bush Bearing Fault

- i. **Carbon Brush Fault:** Faults in the carbon brush of a submersible pump occur due to excessive wear, improper alignment, or debris buildup. Prolonged use, poor maintenance, and electrical arcing can cause the brushes to degrade, leading to reduced conductivity and motor performance. Regular inspection and replacement are essential to prevent such issues.
- ii. **Bush Bearing Fault:** Bush bearing faults in submersible pumps arise from inadequate lubrication, excessive wear, misalignment, or contamination with dirt or debris. These issues can lead to increased friction, overheating, and vibration, ultimately reducing pump efficiency and lifespan. Timely maintenance, proper lubrication, and ensuring alignment can help prevent such faults.



Fig 3.11.1: Healthy



Fig 3.11.2: Carbon Brush



Fig 3.11.3: Fault

Fig 3.11: Represent the Healthy, Carbon Brush and Fault condition

3.6 Data Collection in Faulty condition

The data was collected to analyse the behaviour of the submersible pump system under different fault conditions. The faults analysed include Blockage Fault (B), Impeller Fault (I), Diffuser Fault (D), and Bearing Fault (BF). The Bearing Fault at the faulty condition, while the other faults were varied as specified.

Data Collection Considerations:

1. Parameter Observations:

For each combination, note how the system behaves with respect to parameters such as:

- Flow rate
- Pressure
- Energy consumption
- Pump efficiency or performance degradation

2. Testing Procedure:

- Record data under controlled conditions.
- Start with the normal condition (B0I0D0BF) and then introduce faults one by one.
- Ensure to measure and record the system's response to each fault combination.

3. Outcome Analysis:

- Compare the performance of the pump under different fault conditions to the normal condition (B0I0D0BF).
- Identify patterns in system behaviour, such as increased wear and tear, energy loss, or significant drops in efficiency.

3.7 Fault Combination of Blockage, Impeller, Diffuser and Bearing Fault

Specifications:

- B (Blockage Fault): 0°, 40°, 80°
- I (Impeller Fault): 0, 4, 8
- D (Diffuser Fault): 0, 4, 8

These are the 30 possible combinations that include the normal condition and the Bearing Fault (BF).

S. No.	Blockage	Number of faulty diffusers	Number of faulty impellers	Bush bearing wear	Discharge (litre/min)	Pressure (kPa)	Current (A)
1	1	0	4	0	60.9591	71	6.1629
2	1	4	8	1	46.1188	64	6.0932
3	1	8	4	1	57.5889	65.2	6.2246
4	2	4	4	0	19.6601	67.25	6.2221
5	0	8	4	1	17.1554	61.5	6.13
6	0	4	4	1	48.9231	68.5	6.4434
7	2	4	8	1	20.8919	64.1	6.2803
8	1	8	8	1	12.4744	62	6.2102
9	1	4	4	1	47.2861	68	6.3876
10	0	0	4	1	51.8618	71.54	6.5905
11	1	8	4	0	40.7295	64.83	6.1648
12	1	4	4	1	34.0523	68.56	6.67
13	2	4	0	1	28.3774	70.05	6.75
14	1	4	0	0	43.1153	70.79	6.2910
15	0	4	8	1	47.4400	60.75	6.17
16	0	4	0	1	41.9846	70.79	6.1599
17	1	8	0	1	40.0293	67.81	6.1853
18	1	0	4	1	56.0010	68.56	6.4466
19	1	4	0	1	48.0585	67.81	6.4959
20	1	4	4	1	46.6776	69.2	6.1124
21	1	0	0	1	57.5488	67.1	6.2221
22	1	4	4	1	42.0467	67.82	6.2250
23	1	0	8	1	53.4715	67.82	6.3469
24	2	4	4	1	14.9514	67.1	6.4315
25	0	4	4	0	49.9946	68.56	6.0797
26	2	0	4	1	15.5479	71.54	6.3828
27	2	8	4	1	13.1038	64.08	6.1617
28	1	4	8	0	46.2015	64.83	6.1912
29	1	4	4	1	44.1906	66.81	6.2842

Table 3.6: combinations of faults used in test cases

3.8 Data Preprocessing:

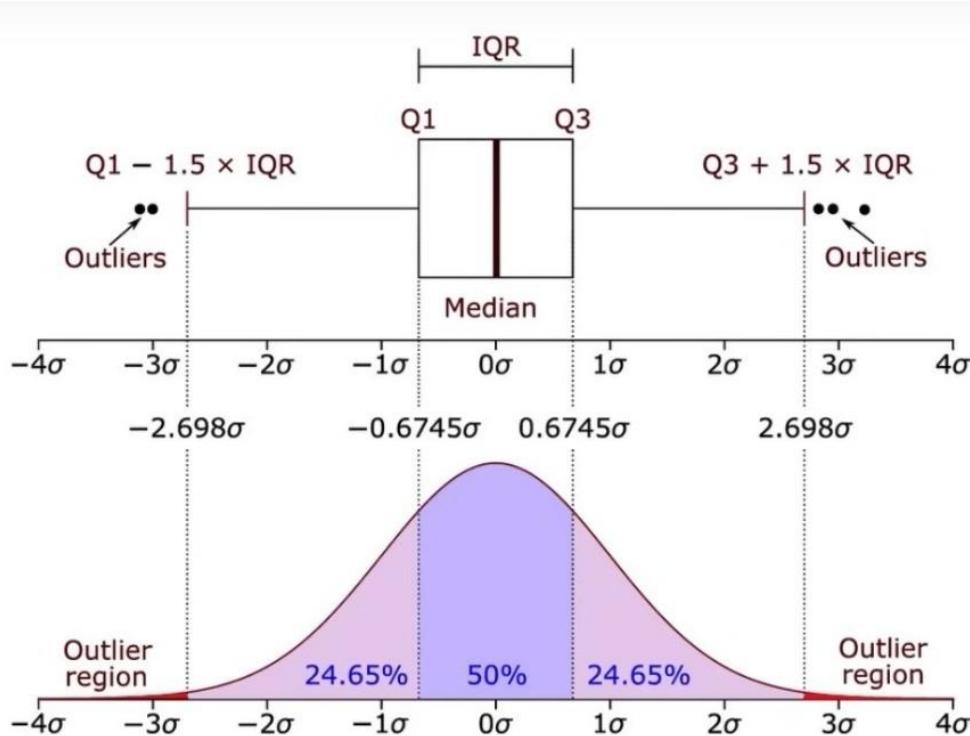


Fig 3.12: Representation of Outliner

- **Cleaning:** Missing or inconsistent data entries caused by sensor noise or transmission errors were handled using interpolation techniques.
- **Normalization:** Sensor readings were normalized to ensure uniformity and prevent bias in machine learning models.
- **Outlier Removal:** Statistical methods such as z-scores or **interquartile range (IQR)** were used to identify and remove anomalies that could skew model training.

3.9 Feature Engineering:

Introduction

Feature engineering is a critical process in machine learning that involves creating, transforming, and selecting the most informative features to improve model performance. This guide explores various statistical and signal processing features, their formulas, and physical interpretations.

3.9.1. Central Tendency Features

3.9.1.1 Mean (Arithmetic Average)

The mean is the average value of a set of numbers, calculated as the sum of all values divided by the total number of values. It represents the central tendency of a dataset.

Formula: $\mu = (x_1 + x_2 + \dots + x_n) / n$

Physical Significance:

- Represents the central point of a dataset
- Sensitive to extreme values
- Useful for normally distributed data
- Indicates the average or typical value in a signal or measurement

3.9.1.2 Median

Formula:

- i. For odd number of observations: Middle value
- ii. For even number of observations: Average of two middle values

Physical Significance:

- Less sensitive to outliers compared to mean
- Represents the middle point of a dataset
- Provides a robust central tendency measure
- Useful for skewed distributions or data with extreme values

3.9.1.3 Mode

Formula: The value that appears most frequently in a dataset

Physical Significance:

- Represents the most common value
- Useful for categorical data
- Can indicate dominant characteristics in a signal or measurement
- Helpful in understanding recurring patterns

3.9.1.4 Maximum

The maximum value in a dataset is the largest value among all the observations. It represents the upper bound of the data and provides insight into the range or extent of the dataset.

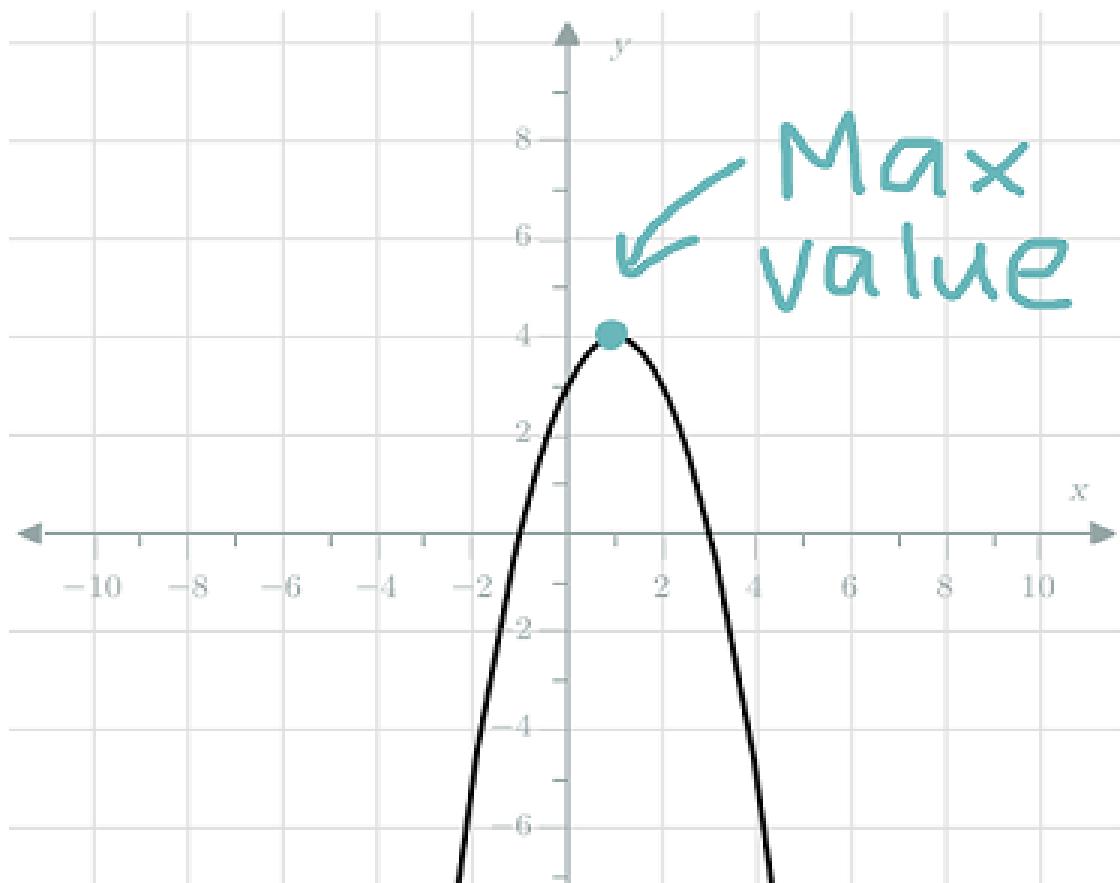


Fig 3.13: Represent Maximum Value

Formula:

If x_1, x_2, \dots, x_n are the values in the dataset:

$$\text{Maximum Value} = \max(x_1, x_2, \dots, x_n)$$

Significance

- **Range Calculation:** The maximum value is essential in determining the range of the data, which is $\text{Range} = \text{Max} - \text{Min}$.
- **Data Extremes:** It helps identify outliers and understand data spread.
- **Decision-Making:** In fields like finance, sports, or engineering, the maximum value indicates the peak performance or stress limits.

3.9.1.5 Minimum Value

The **minimum value** in a dataset is the smallest value among all the observations. It represents the lower bound of the data and provides insight into the dataset's extent.

Formula:

If x_1, x_2, \dots, x_n are the values in the dataset:

Minimum Value= $\min(x_1, x_2, \dots, x_n)$

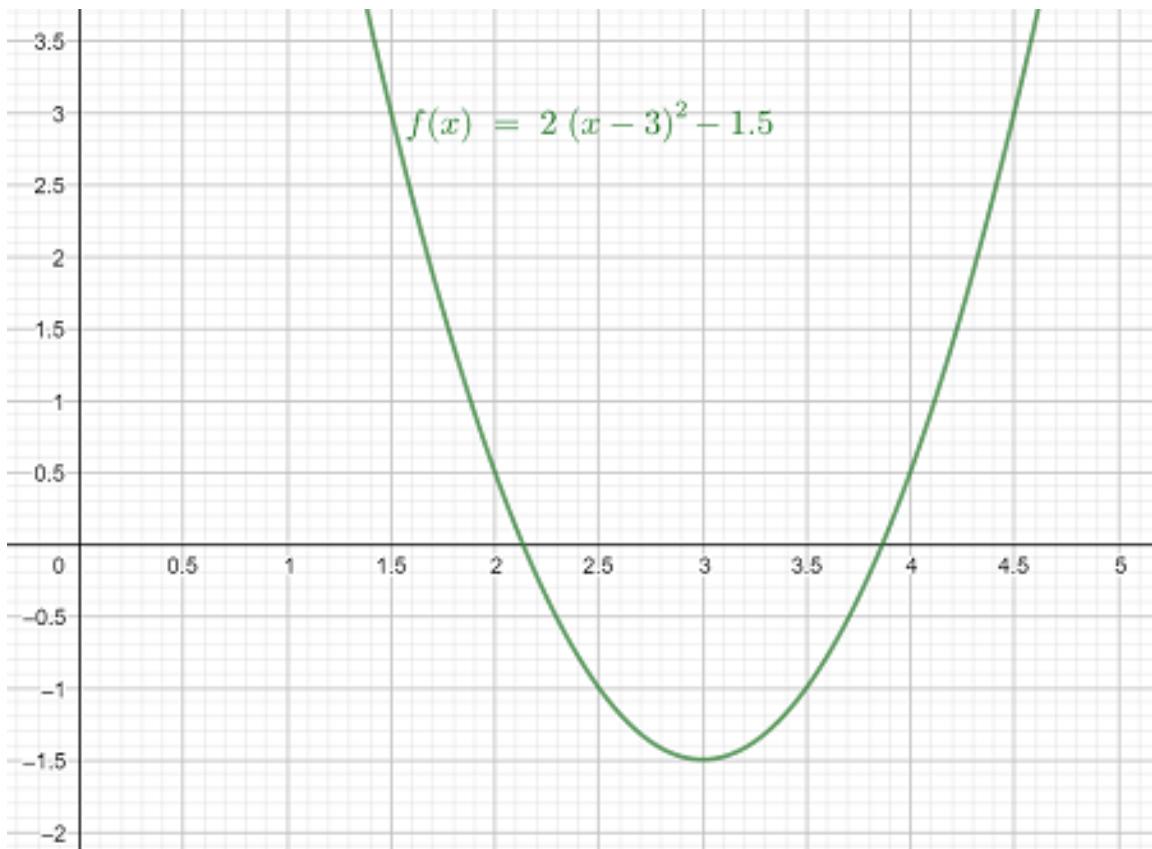


Fig 3.14: Represent the minimum value

Significance

- **Range Calculation:** The minimum value helps determine the range, $\text{Range}=\text{Max}-\text{Min}$, which measures data spread.
- **Data Extremes:** It identifies the lowest point in the dataset, often useful for detecting anomalies or outliers.
- **Benchmarking:** The minimum value acts as a baseline in various fields, such as performance analysis or quality control.

3.9.2 Dispersion Features

3.9.2.1 Variance: Variance is a statistical measurement of the spread between numbers in a data set.

Formula: $\sigma^2 = \sum (x_i - \mu)^2 / n$

Physical Significance:

- Measures the spread or dispersion of data points
- Indicates how far data points are from the mean
- Higher variance suggests more spread-out data

3.9.2.2 Standard Deviation

The standard deviation (SD) is a measure of the amount of variation or dispersion in a dataset. A lower standard deviation indicates that data points are closer to the mean, while a higher standard deviation shows more spread-out data.

Formula: $\sigma = \sqrt{(\sum(x_i - \mu)^2 / n)}$

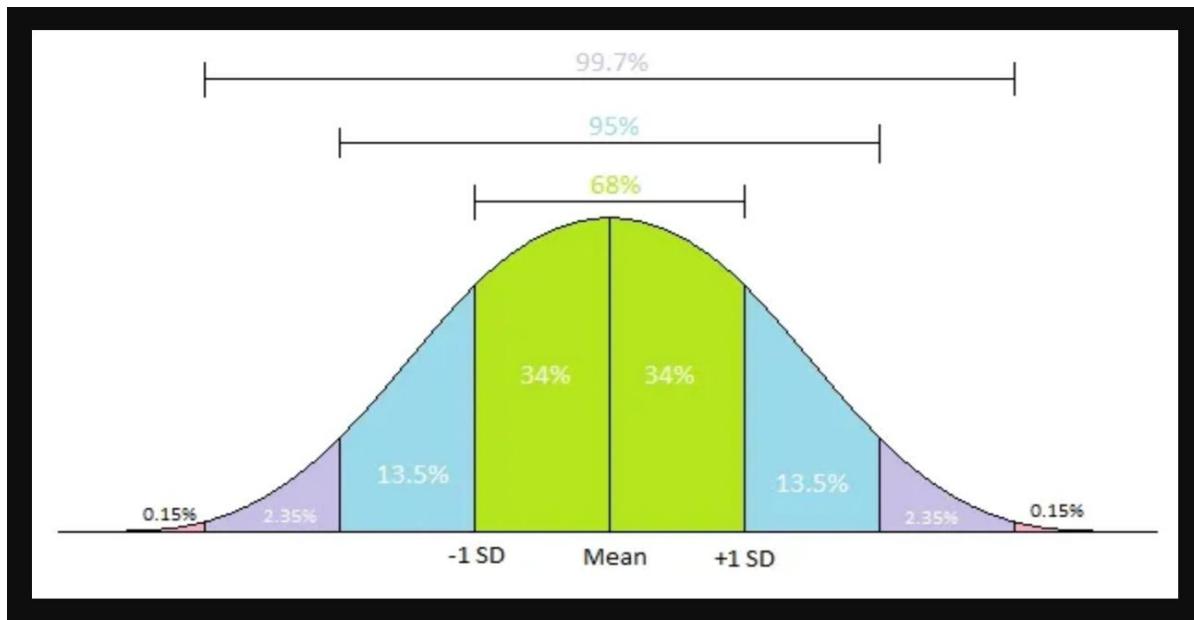


Fig3.15: Represent Standard Deviation

Physical Significance:

- **Measure of Spread:** SD quantifies the dispersion of data relative to the mean.
- **Comparison of Variability:** It allows comparison of consistency or variability between datasets.
- **Probabilistic Analysis:** In a normal distribution, SD helps identify probabilities (e.g., 68% of data lies within 1 SD, 95% within 2 SDs).
- **Error and Risk Assessment:** SD is crucial in fields like finance and quality control for assessing risks and uncertainties.

3.9.3. Signal Processing Features

3.9.3.1 Root Mean Square (RMS): the arithmetic means of the squares of a given set of numbers.

Formula: $RMS = \sqrt{(\sum x_i^2 / n)}$

Physical Significance:

- **Effective Value:** RMS is used to determine the equivalent DC value of an AC signal in terms of power.
- **Energy Quantification:** It provides a measure of the energy or power delivered by the waveform.
- **Signal Comparison:** RMS allows for comparing signals of different shapes and magnitudes by representing their effective values.

- **Applications:** Widely used in electrical systems, audio engineering, and vibration analysis to measure effective performance.

3.9.3.2 Crest Factor

Formula: Crest Factor = Peak Amplitude / RMS Value

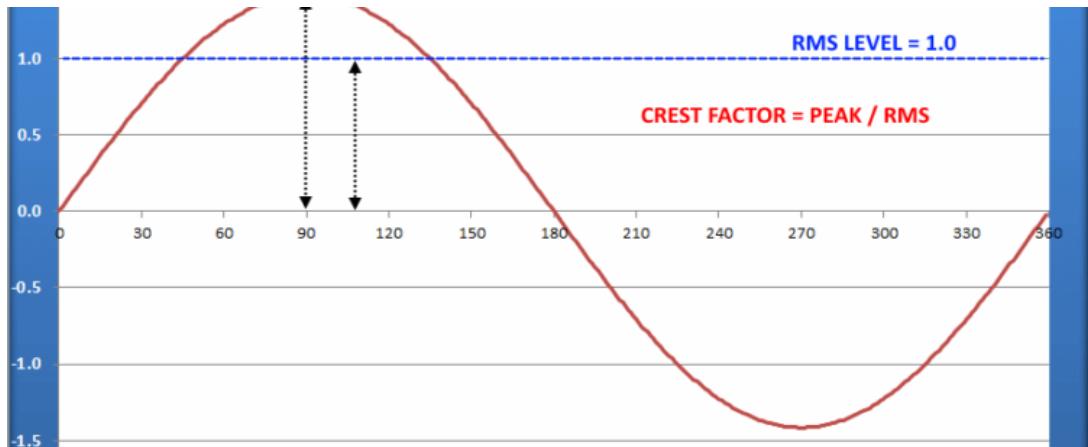


Fig3.16: Representation of Crest Factor

Physical Significance:

- **Waveform Shape Analysis:** The crest factor provides insight into the shape and sharpness of the waveform. Higher values indicate more pronounced peaks.
- **Signal Quality Assessment:** In audio signals, a high crest factor may indicate less distortion or clipping.
- **Equipment Stress Testing:** High crest factor signals can indicate potential stress on electrical equipment, as they experience larger peak values relative to their RMS levels.
- **Fault Detection:** In vibration analysis, abnormal crest factors can signal potential machinery faults.

3.9.3.3 Skewness:

Skewness is a statistical measure that describes the asymmetry of a dataset's distribution around its mean.

- **Positive Skewness:** The tail is longer on the right; most data values are concentrated on the left.
- **Negative Skewness:** The tail is longer on the left; most data values are concentrated on the right.
- **Zero Skewness:** The data is symmetrically distributed.

Formula: Skewness = $\Sigma(x_i - \mu)^3 / (n * \sigma^3)$

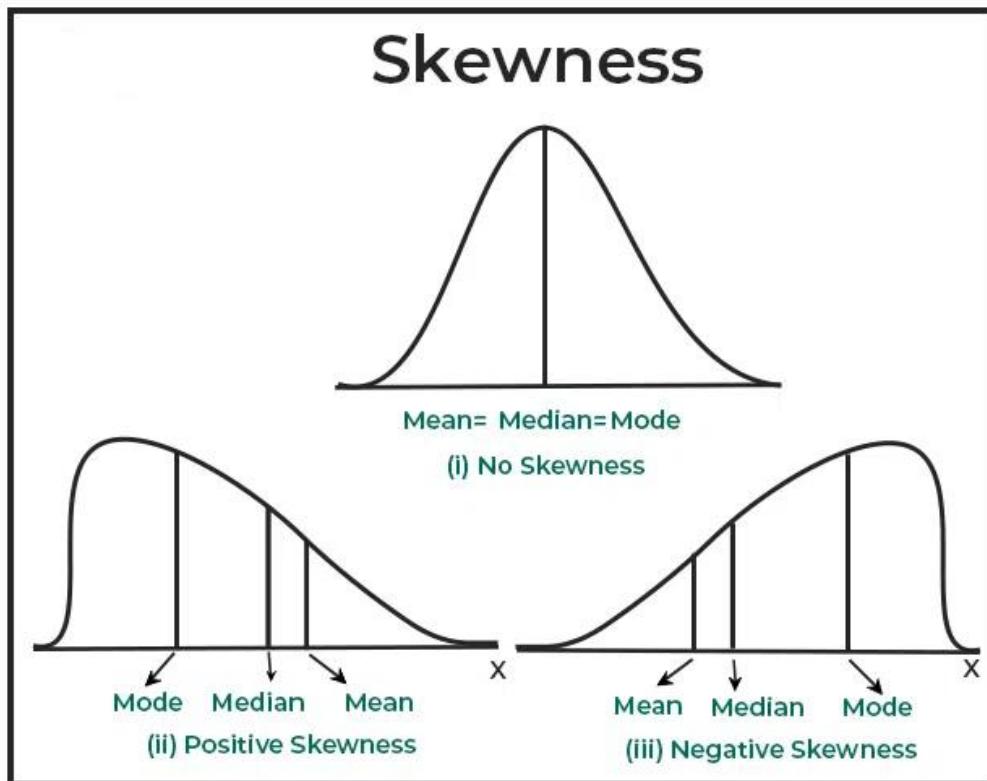


Fig3.17: Represent the Skewness Nature

Physical Significance:

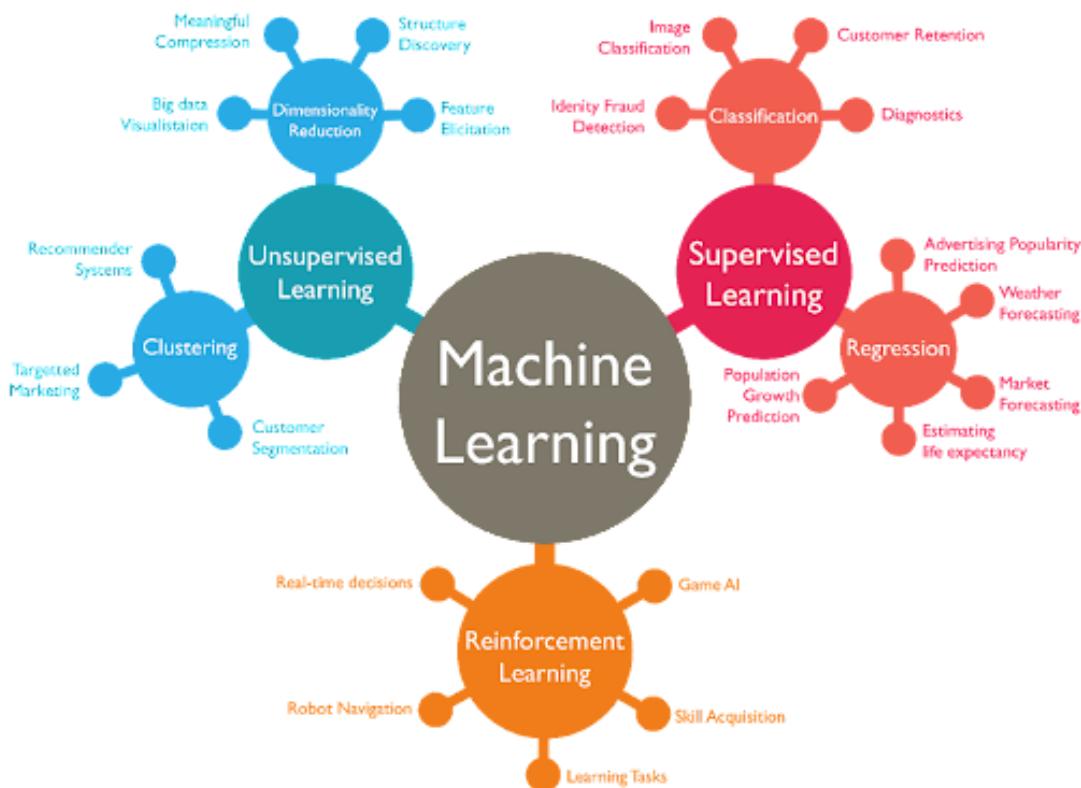
- **Shape Analysis:** Skewness helps understand the shape and balance of the dataset's distribution.
- **Outlier Detection:** It identifies if a dataset has extreme values skewing the distribution.
- **Model Selection:** In statistical modeling, skewness indicates whether transformations are needed for normality.
- **Practical Insights:** Skewness is essential in fields like finance to understand risk (e.g., returns skewed negatively indicate more frequent losses).

Features such as **pressure fluctuations**, **flow rate variations**, and **current spikes** were analysed for their relevance to pump health.

Temporal features (e.g., rate of change over time) were extracted to identify early signs of wear or malfunction.

3.10 Model Development:

Fig



3.18: Breakdown of the main ML model architectures and tasks. Source: [AWS](#)

- **Algorithm Selection:** Various machine learning algorithms, including Random Forest, Gradient Boosting, K-nearest neighbour and Support Vector Machines (SVM), were evaluated to predict maintenance requirements.
- **Training:** The dataset was split into training (20%), validation (20%), and testing (60%) sets.

Hyperparameter Tuning:

A **hyperparameter** is a setting or configuration external to the model that controls its training process and performance. Unlike model parameters (e.g., weights in a neural network), hyperparameters are set before training and do not change during training.

- **Learning rate:** Controls the step size in weight updates.
- **Number of layers:** Defines the architecture of a neural network.
- **Max depth:** Limits the depth of a decision tree.
- **Batch size:** Determines how many samples are processed at once during training.
- Hyperparameters influence how well a model learns and generalizes but must be chosen manually or tuned using methods like grid search or random search.

Chapter 4

Result and Discussion

The primary objective of this study was to evaluate and compare multiple machine learning techniques for predicting the performance of submersible pumps based on pressure, flow rate, and current data. Models such as Random Forest (RF), Gradient Boosting (GB), xGBoost (XGB), and support vector classifier developed to predict pump performance using key input parameters. These included essential features like pressure (P), flow rate (F), current (I), and operational conditions.

The dataset, consisting of 29 tests with 1000 observations of each test case, underwent comprehensive preprocessing and feature engineering to enhance model accuracy. Grid search and randomized cross-validation techniques were applied to optimize hyperparameters, with model evaluation based on 8 features of which include mean, maximum, minimum, skewness, crest factor, form factor, and RMS value.

Following the identification of the optimal model, a feature importance analysis was conducted using XG Boost algorithm to determine the most influential variables affecting pump performance. Additionally, feature dependence plots were generated to provide interpretive insights into how these key features influenced performance matrix.

This method allowed for a detailed understanding of each variable's contribution to the pump's efficiency and reliability, offering valuable insights for model-driven optimization and operational improvements in submersible pump systems.

4.1. Data Preprocessing and Feature Engineering

Our dataset encompassed a wide range of operational and statistical parameters critical for predicting submersible pump performance. The features included mean, maximum value (max), minimum value (min), crest factor, form factor, root mean square (RMS), skewness, and kurtosis. These parameters were carefully selected to represent the operational and dynamic conditions influencing pump efficiency and reliability. Ensuring the dataset's consistency and suitability for robust machine learning model training posed a significant challenge, necessitating extensive preprocessing and feature engineering.

To handle categorical variables, such as fault types or operational modes, techniques like one-hot encoding, label encoding, and ordinal encoding were applied. These methods transformed the categorical data into numerical representations, resulting in 8 distinct columns corresponding to each category. This approach effectively captured categorical differences while avoiding the introduction of undesired ordinal relationships.

For continuous features, standardization was performed to normalize their scales, ensuring that each variable contributed proportionally to the machine learning model. This preprocessing step reduced biases caused by varying ranges in the dataset and improved model stability and accuracy.

After preprocessing, the dataset, consisting of 8 key features, was used to train various machine learning models, including Support Vector Classifier (SVC), XGBoost, Random Forest, and k-Nearest Neighbors (KNN). These models were evaluated to determine their ability to classify faults accurately and identify the most effective sensor for fault detection.

The performance of each model was assessed using metrics such as precision, recall, and F1-score, ensuring a fair comparison. Among the models, the one that delivered the highest accuracy and generalization capabilities was identified as the most suitable for sensor-based fault classification. This approach not only optimized fault detection but also provided valuable insights into selecting the best sensors for monitoring submersible pump operations, contributing to enhanced reliability and predictive maintenance strategies.

A detailed statistical analysis of the dataset, summarizing key metrics such as range, mean, standard deviation, and variance for each feature and the target variable, highlighted significant trends. For instance, variability in RMS and crest factor demonstrated a strong correlation with pump performance, underscoring their importance as predictors in the modelling process.

The combination of thorough preprocessing, feature standardization, one-hot encoding, and data augmentation ensured that the dataset was clean, comprehensive, and well-suited to the demands of advanced machine learning models, including Support Vector Classifier (SVC), XGBoost, Random Forest, and k-Nearest Neighbor (KNN). These techniques collectively enhanced the models' ability to predict submersible pump performance with high accuracy and reliability.

	Part	F_mean	F_max	F_min	F_rms	F_std_dev	F_skewness	F_crest_factor	F_form_factor
count	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000
mean	500.447800	40.318339	49.469639	33.284011	40.592287	4.097777	-0.026762	1.229211	1.009909
std	288.596145	15.230923	28.240491	14.206236	15.201806	2.161530	1.618833	0.409490	0.015861
min	1.000000	11.597938	17.000000	8.000000	11.925456	0.000000	-10.198039	1.000000	1.000000
25%	250.750000	28.030612	34.650000	18.900000	28.174204	3.139683	-0.832775	1.089544	1.003100
50%	500.500000	45.960000	51.000000	42.000000	46.176617	4.090554	-0.218236	1.183282	1.004514
75%	750.250000	50.307692	60.000000	42.000000	50.402318	4.514651	0.273898	1.264438	1.006328
max	1000.000000	73.093220	531.000000	60.000000	85.605075	45.364866	11.656794	6.855768	1.197287

(a)

	I_mean	I_max	I_min	I_rms	I_std_dev	I_skewness	I_crest_factor	I_form_factor
count	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000
mean	6.352343	6.401701	6.307495	6.352470	0.027187	0.144203	1.007702	1.000020
std	0.341513	0.357521	0.341249	0.341579	0.032104	0.582888	0.009549	0.000936
min	5.403435	5.979000	0.000000	5.794581	0.007603	-3.139651	1.000588	1.000001
25%	6.167145	6.204000	6.117000	6.167195	0.021820	-0.231429	1.005629	1.000006
50%	6.256211	6.307000	6.221000	6.256265	0.025247	0.134498	1.007124	1.000008
75%	6.410151	6.463000	6.359000	6.410271	0.029415	0.503802	1.008931	1.000010
max	8.098292	14.343000	8.018020	8.098551	3.477615	2.936954	1.928221	1.122050

(b)

	P_mean	P_max	P_min	P_rms	P_std_dev	P_skewness	P_crest_factor	P_form_factor
count	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000	30000.000000
mean	67.692253	85.709821	35.204734	68.655628	11.487158	-0.808151	1.248831	1.014239
std	2.925096	3.471938	3.137445	2.956299	0.660528	0.088718	0.027320	0.001242
min	60.369498	77.605000	29.133000	61.301823	9.221580	-1.592140	1.186542	1.009948
25%	65.587626	82.998000	32.994000	66.515772	11.017776	-0.871595	1.228981	1.013293
50%	67.827329	85.860000	35.280000	68.819341	11.478365	-0.799738	1.237038	1.014355
75%	69.342756	87.768000	37.240000	70.296060	11.932148	-0.751241	1.265399	1.015209
max	77.089855	95.400000	47.000000	78.034080	13.952505	-0.442202	1.336508	1.019488

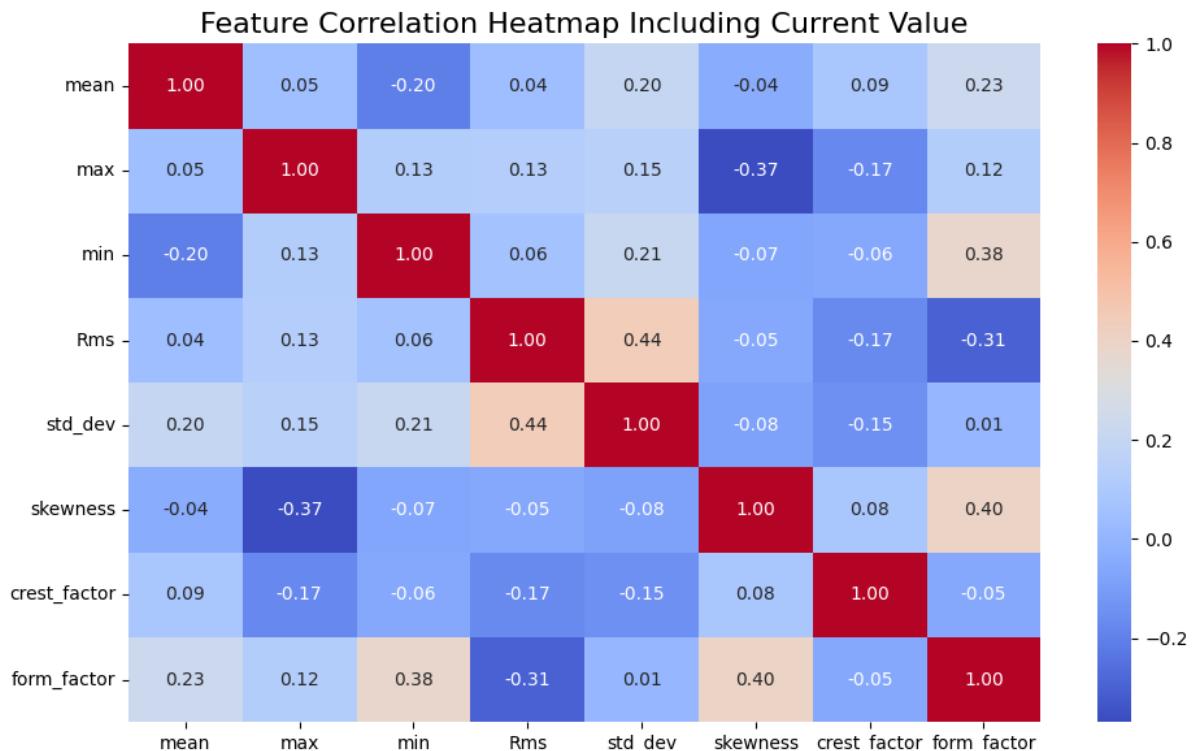
(C)

Table 4.1 Classification Report of Data

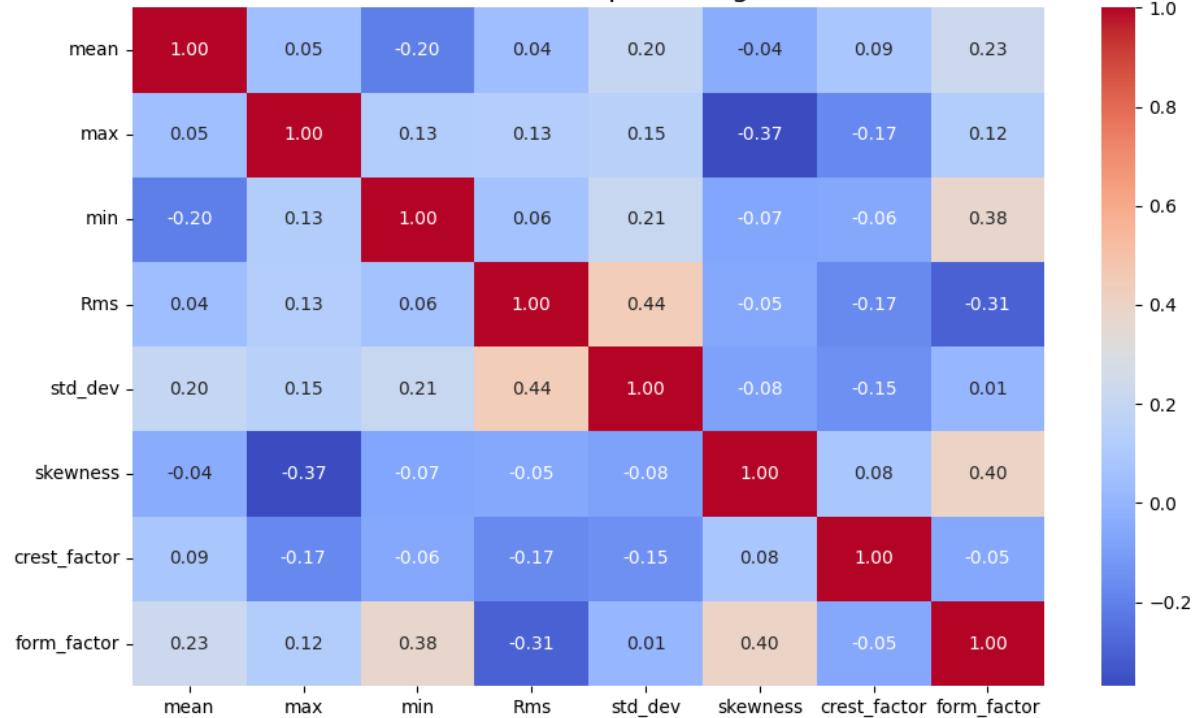
4.2. Correlation Analysis

A detailed correlation analysis was performed to examine the relationships between input features and the target variable, submersible pump performance. This analysis provided valuable insights into how each parameter influences pump operation and efficiency. A Pearson correlation matrix was utilized to quantify the strength and direction of linear relationships among the features and the performance metric. The results were visualized using a heatmap, enabling a clear interpretation of the interactions.

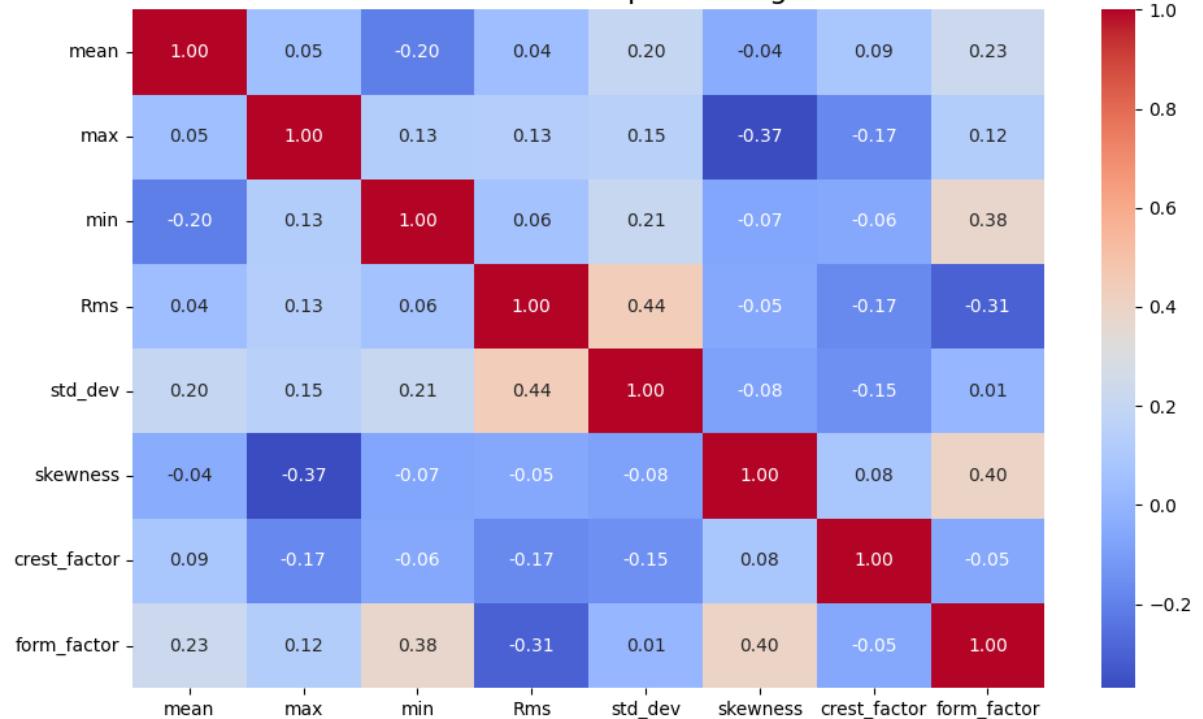
This analysis was crucial for identifying the most significant predictors, detecting potential redundancies among features, and understanding the overall structure of relationships within the dataset. These insights guided the feature selection process, ensuring the inclusion of parameters with the highest impact on pump performance for further modelling and evaluation.



Feature Correlation Heatmap Including Pressure Value



Feature Correlation Heatmap Including Flow Value



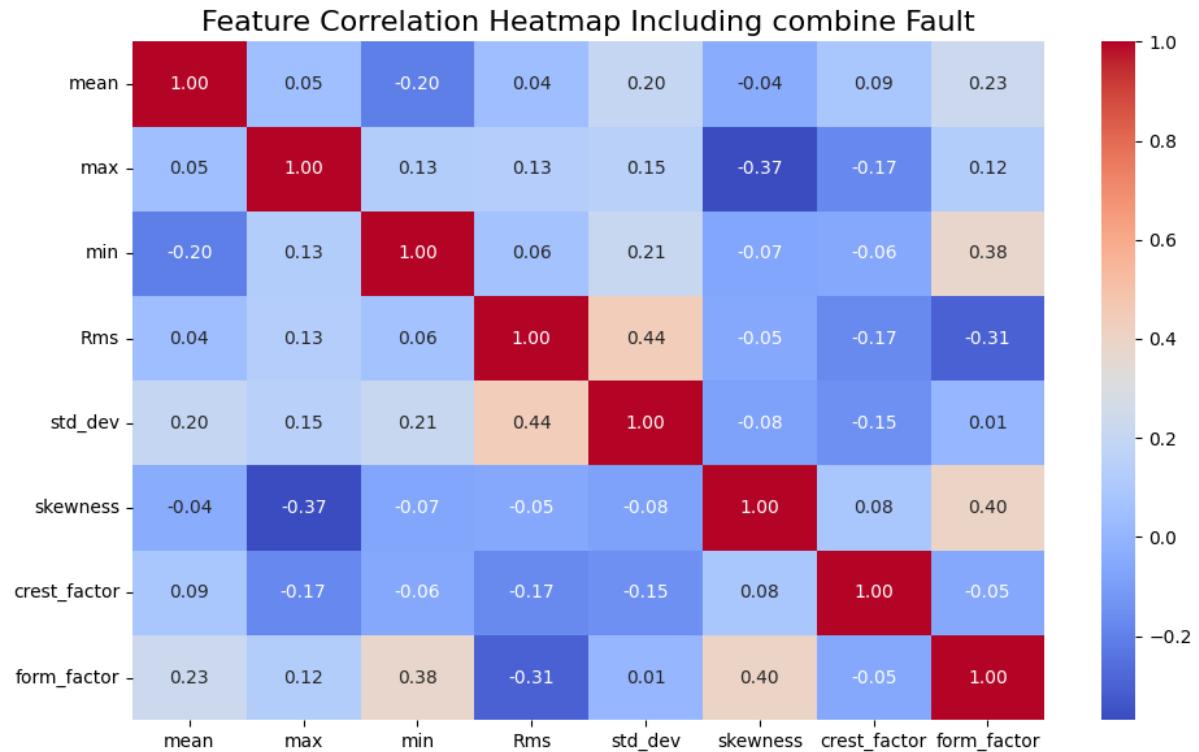


Fig 4.1: Feature Correlation Heatmap

4.3 Significant Effect of Pressure, Current and Flow

The pressure, flow, and current are essential parameters that significantly impact the performance and efficiency of a submersible pump. Pressure determines the pump's capability to push water to a specific height or through a complex piping system, making it crucial for maintaining consistent water delivery. Flow rate, on the other hand, measures the volume of water transported over a given period, directly affecting the pump's capacity to meet system demands. Current reflects the electrical power consumption of the pump and serves as an indicator of its operational efficiency. Variations in these parameters can signal potential issues, such as blockages, mechanical wear, leaks, or inefficiencies in energy usage. Regular monitoring and analysis of these factors not only ensure optimal performance but also help in identifying and addressing faults early, thus extending the pump's lifespan and reducing operational costs.

4.3.1 Effect in Pressure value on Introducing Fault

Pressure plays a crucial role in the operation of a submersible pump and can contribute to various faults if not maintained within optimal limits. Here's how pressure affects different components and faults:

- Impeller Faults:** Excessive pressure can cause stress on the impeller, leading to deformation, cracking, or imbalance. This can reduce the pump's efficiency and cause vibrations that accelerate wear.

2. **Diffuser Faults:** High pressure can create uneven flow distribution, leading to erosion or damage to the diffuser. Over time, this may reduce its ability to guide water efficiently, affecting overall performance.
3. **Bearing Faults:** Increased pressure can overload the pump system, transmitting additional forces to the bearings. This can cause overheating, wear, and even bearing failure, resulting in noisy operation and reduced lifespan.
4. **Blockage Faults:** If the pressure rises due to blockages in the pump or pipeline, it can strain the pump components, causing damage to the impeller, diffuser, or seals. This can lead to decreased flow rate and potential overheating of the motor.

Proper pressure management and regular maintenance are essential to prevent these issues and ensure the smooth and efficient operation of the pump.

4.3.2 Effect in Flow Rate on Introducing Fault

Flow rate is a key operational parameter in submersible pumps, and deviations from optimal levels can contribute to several faults in the system:

1. **Impeller Faults:** A low flow rate can lead to cavitation, where vapor bubbles form and collapse on the impeller's surface, causing pitting and erosion. High flow rates, on the other hand, can exert excessive hydraulic forces on the impeller, leading to wear or mechanical stress.
2. **Diffuser Faults:** An improper flow rate can disrupt the smooth guidance of water through the diffuser. Low flow may result in uneven water distribution, causing turbulence and wear, while high flow can erode the diffuser surfaces, reducing its efficiency.
3. **Bearing Faults:** Irregular flow rates can lead to imbalanced hydraulic forces, which may transfer additional load to the bearings. This can result in uneven wear, overheating, or even failure of the bearings over time.
4. **Blockage Faults:** Low flow rates often indicate partial blockages in the pump or pipeline. This can cause pressure to build up, straining the system and potentially damaging the impeller, diffuser, or seals. High flow rates through a partially blocked system can worsen the blockage or lead to leaks.

Maintaining an appropriate flow rate and addressing any deviations promptly is critical to preventing these faults and ensuring the efficient and reliable operation of the pump.

4.3.3 Effect in Current value on Introducing Fault

Current is a vital parameter in the operation of a submersible pump, and deviations from normal levels can indicate or contribute to faults in various components:

- Impeller Faults:** An increase in current may suggest that the impeller is overloading the motor, possibly due to debris, wear, or improper alignment. This can strain the motor and cause overheating, leading to reduced efficiency or damage.
- Diffuser Faults:** If the diffuser is damaged or misaligned, it can create hydraulic resistance, causing the motor to draw more current to maintain performance. This additional load can accelerate motor wear and reduce energy efficiency.
- Bearing Faults:** High current often points to increased friction in the system, which can occur if the bearings are worn, improperly lubricated, or misaligned. Over time, this can lead to bearing failure and potential motor damage.
- Blockage Faults:** A blockage in the pump or pipeline can restrict water flow, forcing the motor to work harder and draw excessive current. This not only increases energy consumption but also raises the risk of overheating and damage to the pump components.

Monitoring current levels is essential for detecting faults early and preventing damage to the pump. Regular maintenance and timely intervention can help ensure smooth and energy-efficient operation.

4.4 Fault Creation and Related Table for Impeller Diffuser Bearing and Blockage Valve

Impeller Fault Parameter	Value
No. of Faulty Impeller	0,4,8
No. of Faulty Vanes	2,4,6
Wear Depth	2,4,6mm
Wear Length	3,6,9mm

Table 4.2: Impeller Fault Parameter

Impeller Fault Code	Fault Condition
N	Healthy impeller
I0	No impeller has faults
I2	4 impellers have faults
I3	8 impellers have faults

Table 4.3: Impeller Fault Parameter

Diffuser Fault Parameter	Value
Wear Length	3,6,9mm
Wear Depth	2,4,6mm
No. of Faulty Vanes	1,3
No. of Faulty Diffuser	0,4,8

Table 4.4: Diffuser Fault Parameter

S. No.	Blockage	Number of faulty diffusers	Number of faulty impellers	Bush bearing wear	Discharge (litre/min)	Pressure (kPa)	Current (A)
1	1	0	4	0	60.9591	71	6.1629
2	1	4	8	1	46.1188	64	6.0932
3	1	8	4	1	57.5889	65.2	6.2246
4	2	4	4	0	19.6601	67.25	6.2221
5	0	8	4	1	17.1554	61.5	6.13
6	0	4	4	1	48.9231	68.5	6.4434
7	2	4	8	1	20.8919	64.1	6.2803
8	1	8	8	1	12.4744	62	6.2102
9	1	4	4	1	47.2861	68	6.3876
10	0	0	4	1	51.8618	71.54	6.5905
11	1	8	4	0	40.7295	64.83	6.1648
12	1	4	4	1	34.0523	68.56	6.67
13	2	4	0	1	28.3774	70.05	6.75
14	1	4	0	0	43.1153	70.79	6.2910
15	0	4	8	1	47.4400	60.75	6.17
16	0	4	0	1	41.9846	70.79	6.1599
17	1	8	0	1	40.0293	67.81	6.1853
18	1	0	4	1	56.0010	68.56	6.4466
19	1	4	0	1	48.0585	67.81	6.4959
20	1	4	4	1	46.6776	69.2	6.1124
21	1	0	0	1	57.5488	67.1	6.2221
22	1	4	4	1	42.0467	67.82	6.2250
23	1	0	8	1	53.4715	67.82	6.3469
24	2	4	4	1	14.9514	67.1	6.4315
25	0	4	4	0	49.9946	68.56	6.0797
26	2	0	4	1	15.5479	71.54	6.3828
27	2	8	4	1	13.1038	64.08	6.1617
28	1	4	8	0	46.2015	64.83	6.1912
29	1	4	4	1	44.1906	66.81	6.2842

Table 4.5: Combination of Blockage, Impeller, Diffuser and Bearing Fault

Table 4.6: Blockage Fault Parameter

Code	Blockage Condition
N	Close Upto 0 deg
P1	Close Upto 40 deg
P2	Close Upto 80 deg

4.5 Machine Learning Model Preparation with SVC, KNN, Random Forest, and XGBoost

4.5.1 Support Vector Classifier (SVC):

SVC is a classification algorithm that works by finding a hyperplane that best separates data points into different classes. It is particularly effective for small datasets and cases where the decision boundary is not linear. Hyperparameters like the kernel type (linear, rbf, or poly), regularization parameter C, and gamma (for non-linear kernels) significantly impact its performance.

Significance

SVC: Best for small and medium-sized datasets where clear margins exist between classes. It excels in high-dimensional spaces.

Hyperparameters and Their Impact

- i. C: Controls the trade-off between a smooth decision boundary and correctly classifying training points.
- ii. kernel: Defines the type of boundary (linear or non-linear).
- iii. gamma: Determines the influence of each training example.

4.5.2 K-Nearest Neighbors (KNN):

KNN is a simple, instance-based learning algorithm that classifies data points based on the majority class of their nearest neighbours. The key hyperparameter for KNN is the number of neighbors (k) and the distance metric (e.g., Euclidean or Manhattan). It is easy to implement but can be computationally intensive for large datasets.

Significance

KNN: Useful for problems where interpretability and simplicity are prioritized. Works well for smaller datasets.

Hyperparameters and Their Impact

- i. k: Number of neighbors to consider; too high can smooth boundaries, too low may overfit.
- ii. distance metric: Determines how neighbors are identified (Euclidean, Manhattan).

4.5.3 Random Forest (RF):

Random Forest is an ensemble learning algorithm that uses multiple decision trees to improve accuracy and reduce overfitting. It randomly selects features and data subsets to build each tree, making it robust to noise. Key hyperparameters include the number of trees (n_estimators), maximum depth of trees, and the minimum number of samples required to split a node.

Significance

Random Forest: Provides high accuracy and works well with larger datasets. It handles missing data and categorical features effectively.

Hyperparameters and Their Impact

- i. n_estimators: Number of trees; too few can underfit, too many increase computation.
- ii. max_depth: Limits tree depth to prevent overfitting.
- iii. min_samples_split: Controls tree growth by requiring a minimum number of samples to split nodes.

4.5.4. XGBoost:

XGBoost (Extreme Gradient Boosting) is an advanced ensemble method based on boosting. It builds trees sequentially, where each new tree corrects the errors of the previous one. It is highly efficient and handles missing values well. Key hyperparameters include the learning rate (eta), the number of boosting rounds (n_estimators), and maximum tree depth.

Significance

XGBoost: Ideal for large and complex datasets with non-linear relationships. It is computationally efficient and often achieves the best performance.

Hyperparameters and Their Impact

- i. learning_rate: Step size for weight updates, smaller values lead to slower but more accurate learning.
- ii. n_estimators: Number of boosting rounds, too many may overfit.
- iii. max_depth: Limits tree depth to prevent overfitting.

4.6. Machine Learning Model Performance

A diverse set of machine learning algorithms was employed to classify faults and predict outcomes, including K-Nearest Neighbors (KNN), Support Vector Classifier (SVC), Random Forest, and XGBoost. These models were selected for their varying complexities and their ability to capture both simple and intricate relationships within the dataset.

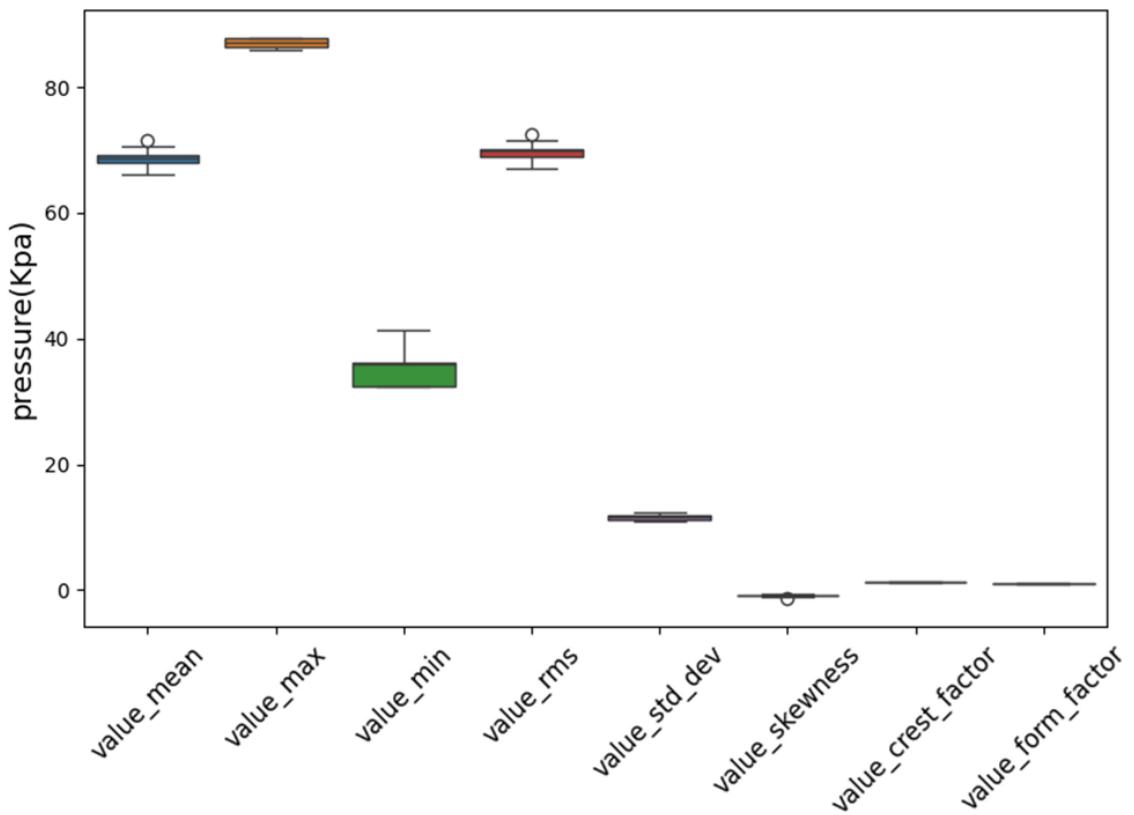
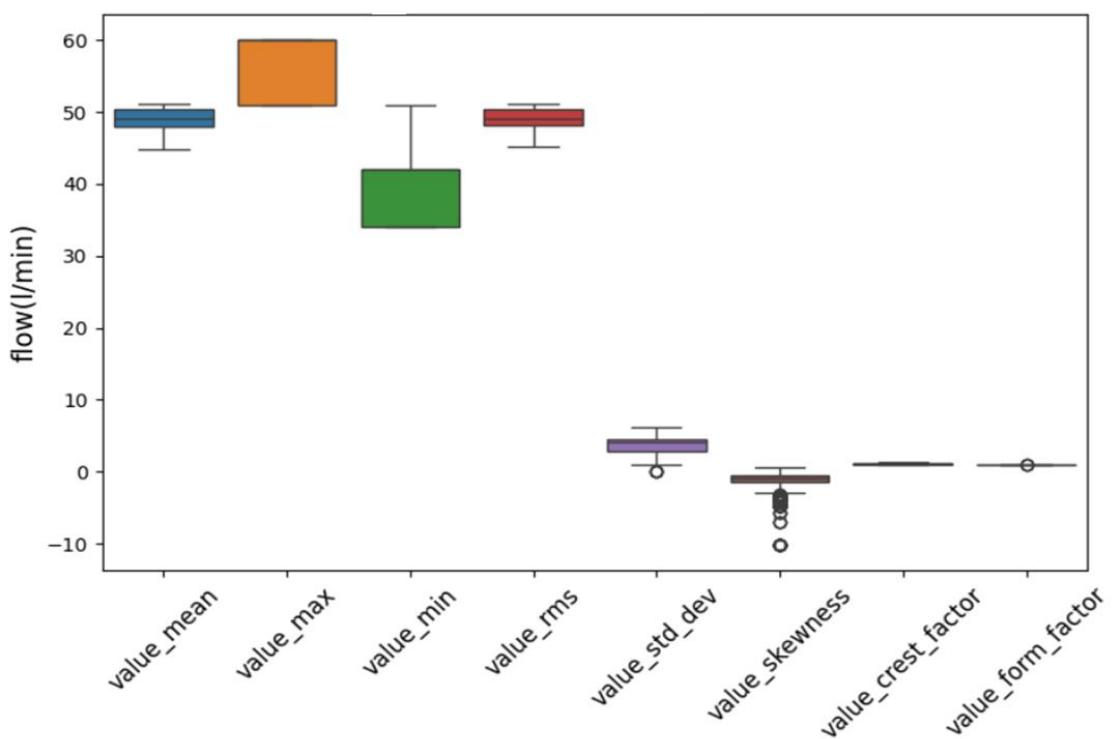
To assess their effectiveness, performance metrics such as accuracy, precision, recall, and F1-score were utilized. These metrics provided a holistic view of each model's predictive power and overall reliability.

- **KNN** demonstrated solid performance in cases where feature scaling and distance metrics were crucial, relying on proximity-based decision-making.
- **SVC** performed well in datasets with clear class separations, particularly excelling in high-dimensional spaces due to its flexible kernel options.
- **Random Forest** stood out as the most robust model, effectively managing complex feature interactions and resisting overfitting through its ensemble approach.
- **XGBoost** exhibited exceptional efficiency and accuracy, leveraging boosting techniques to iteratively refine predictions, making it particularly suited for handling non-linear relationships.

Among these models, Random Forest emerged as the most accurate, showcasing its ability to handle intricate feature interactions and deliver reliable predictions. However, XGBoost was a close competitor, offering faster computations and strong performance in complex scenarios.

Box Plot: A box plot is a tool of visualization to see how data is distributed.

- The horizontal line within the box is the median
- Upper and lower edges of the box are Q3 and Q1
- Interquartile range specifies range of middle 50% of the data
- Whiskers: These are the smallest and largest value of data that are not considered outliers



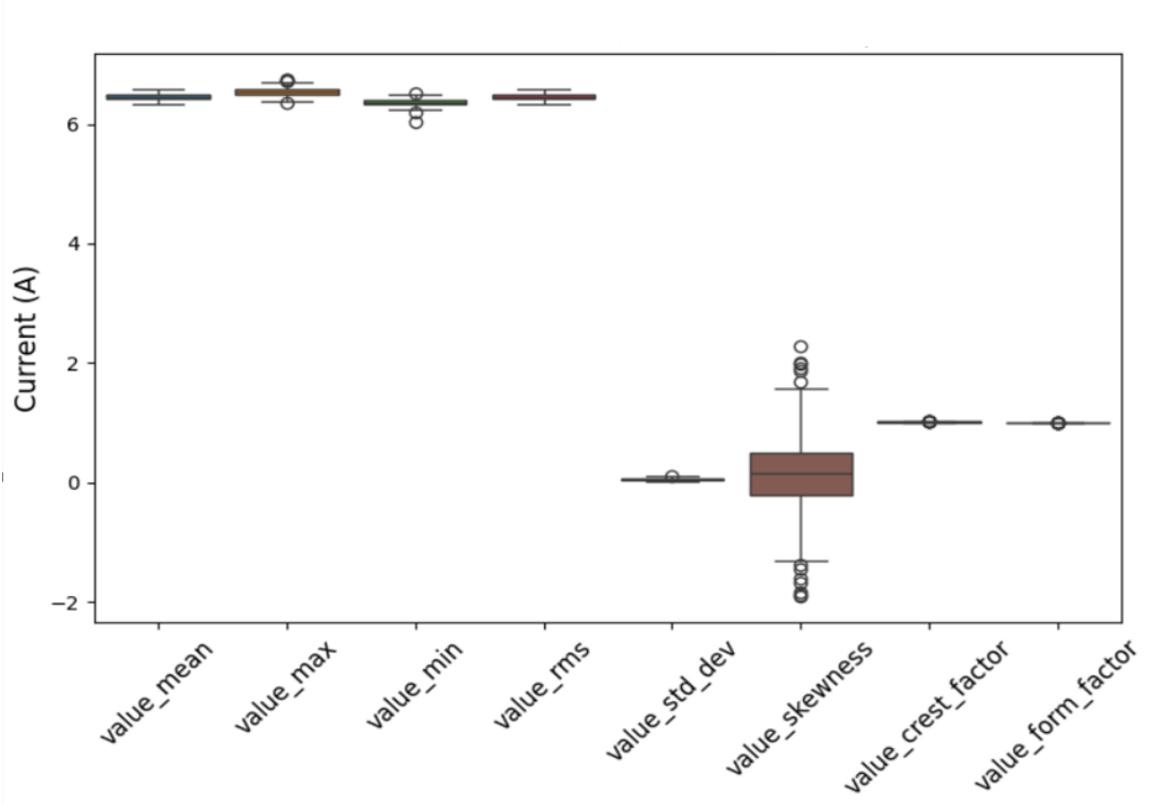


Fig 4.2: Box Plot for the Flow Rate, Pressure and Current

4.6.1 Training, Testing, and Validation

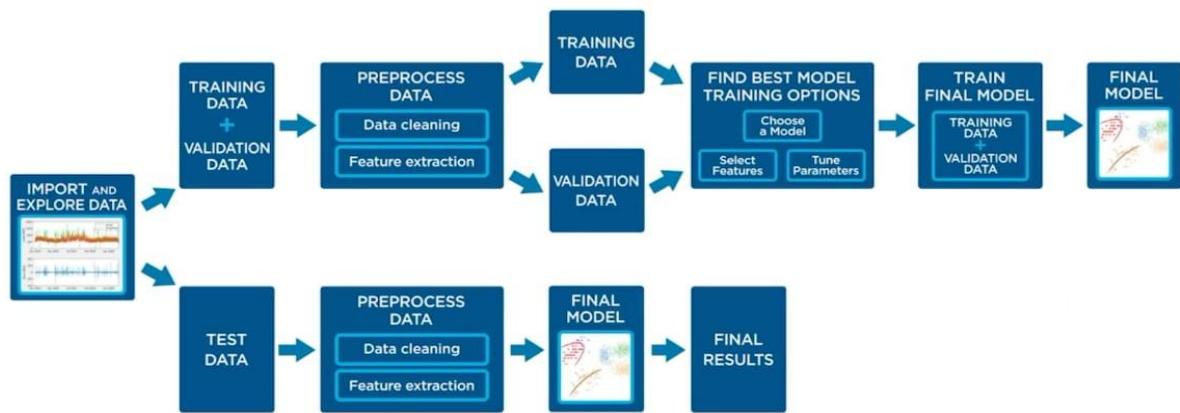


Fig 4.3: The supervised ML workflow [12]

Dataset: Prepared a dataset of 30,000 records of each Sensor pressure, flow and current with 8 features and labels. This dataset is split into:

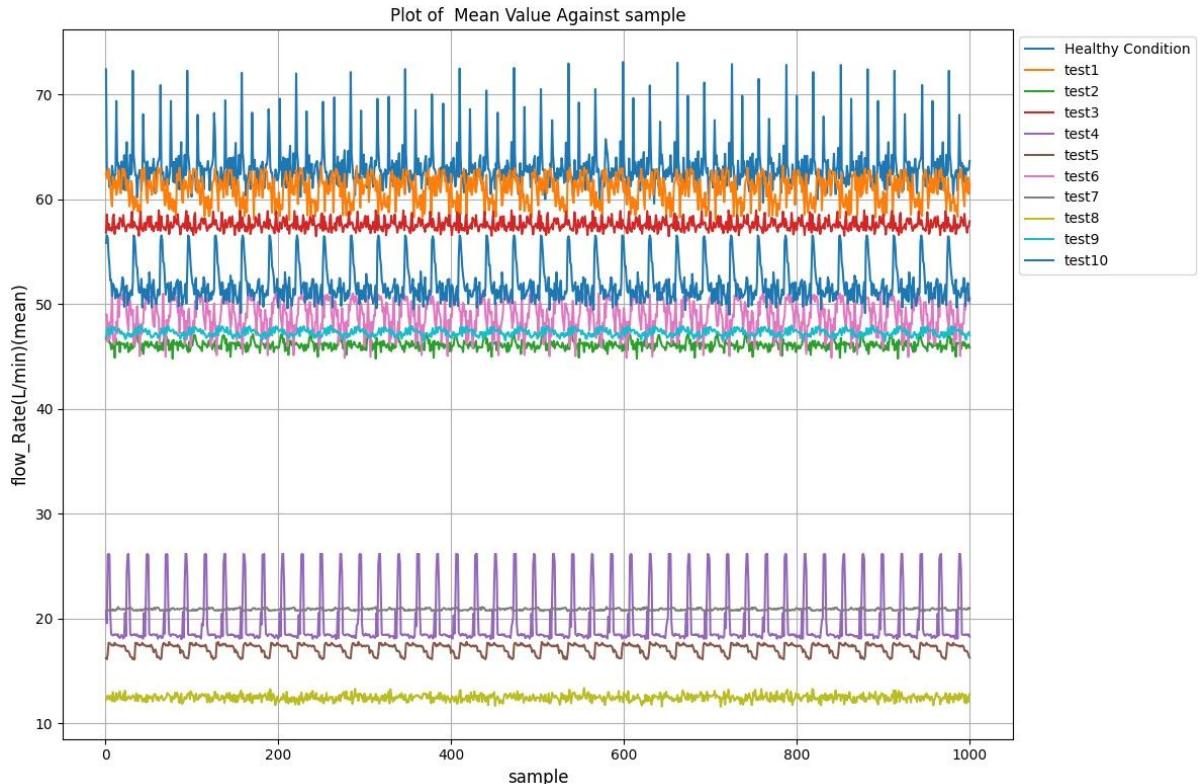
- Training Data (20%): Used to train the model.
- Testing Data (80%): Used to evaluate model performance on unseen data.

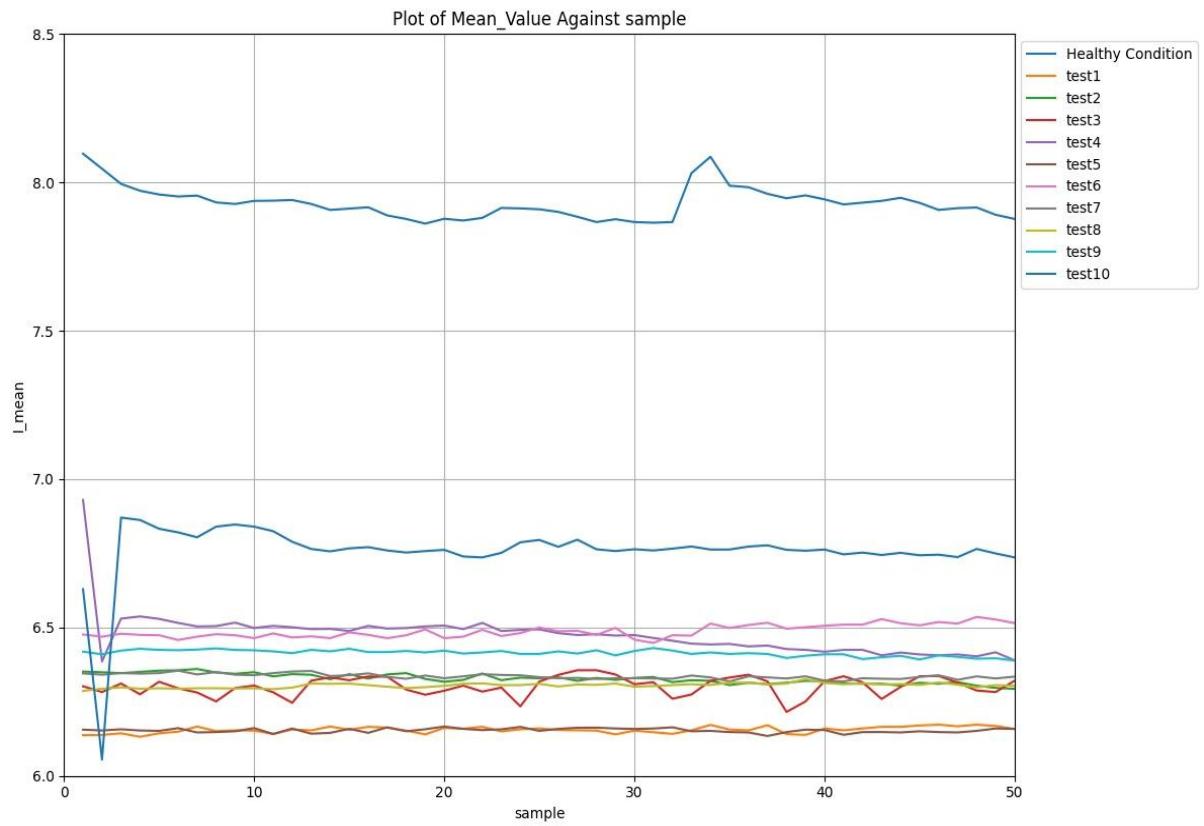
Validation: During training, validation is done using techniques like k-fold cross-validation to ensure the model generalizes well and to tune hyperparameters effectively.

By training and validating these algorithms on the dataset, you can compare their performance based on metrics like accuracy, precision, recall, or F1-score. Hyperparameter tuning through grid search or random search ensures that each algorithm achieves optimal performance for the given problem.

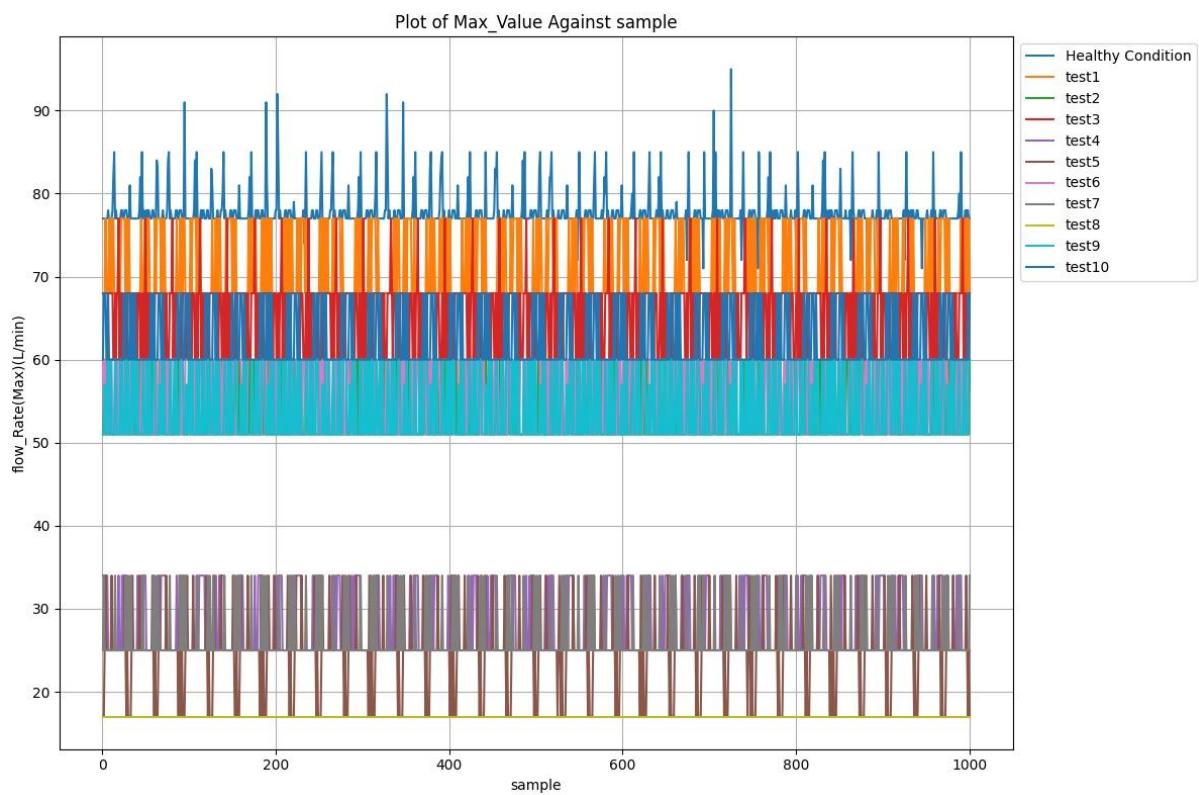
4.6.2 Eight Features Plot Vs Sample Rate

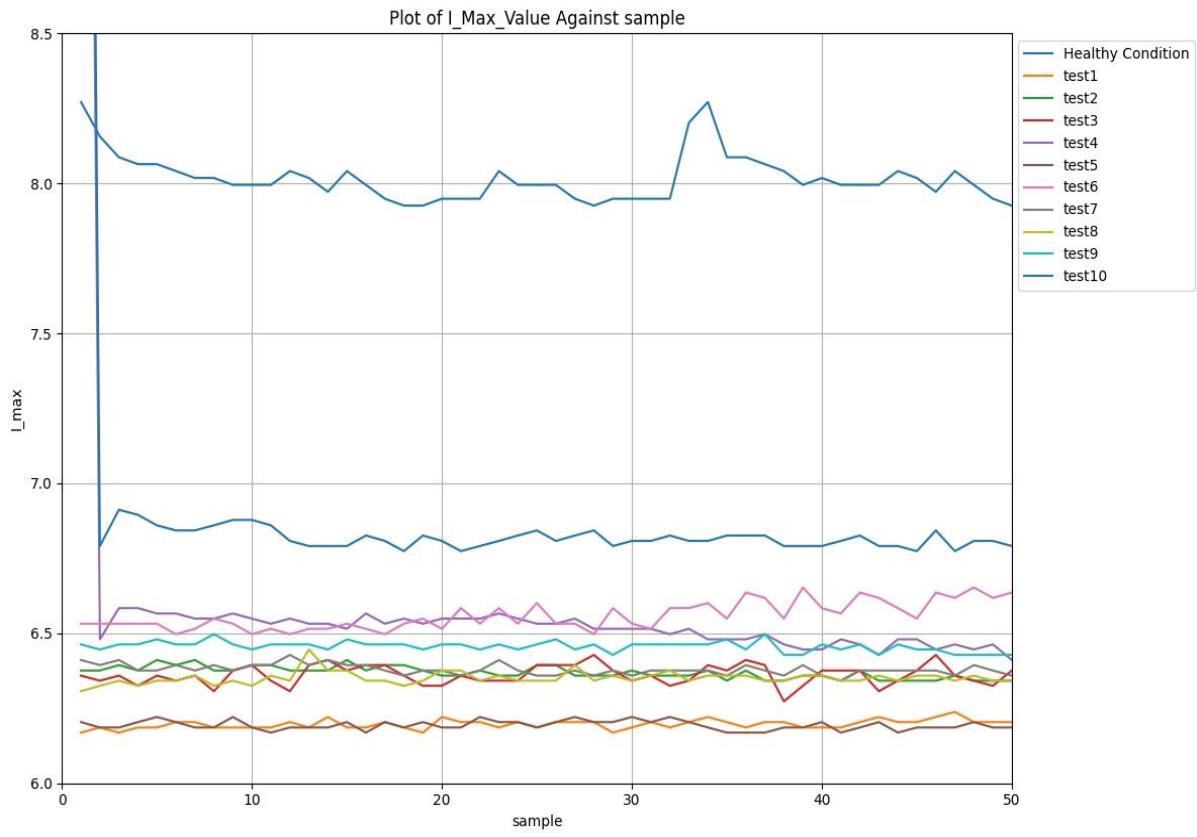
1. Feature Plot of Mean



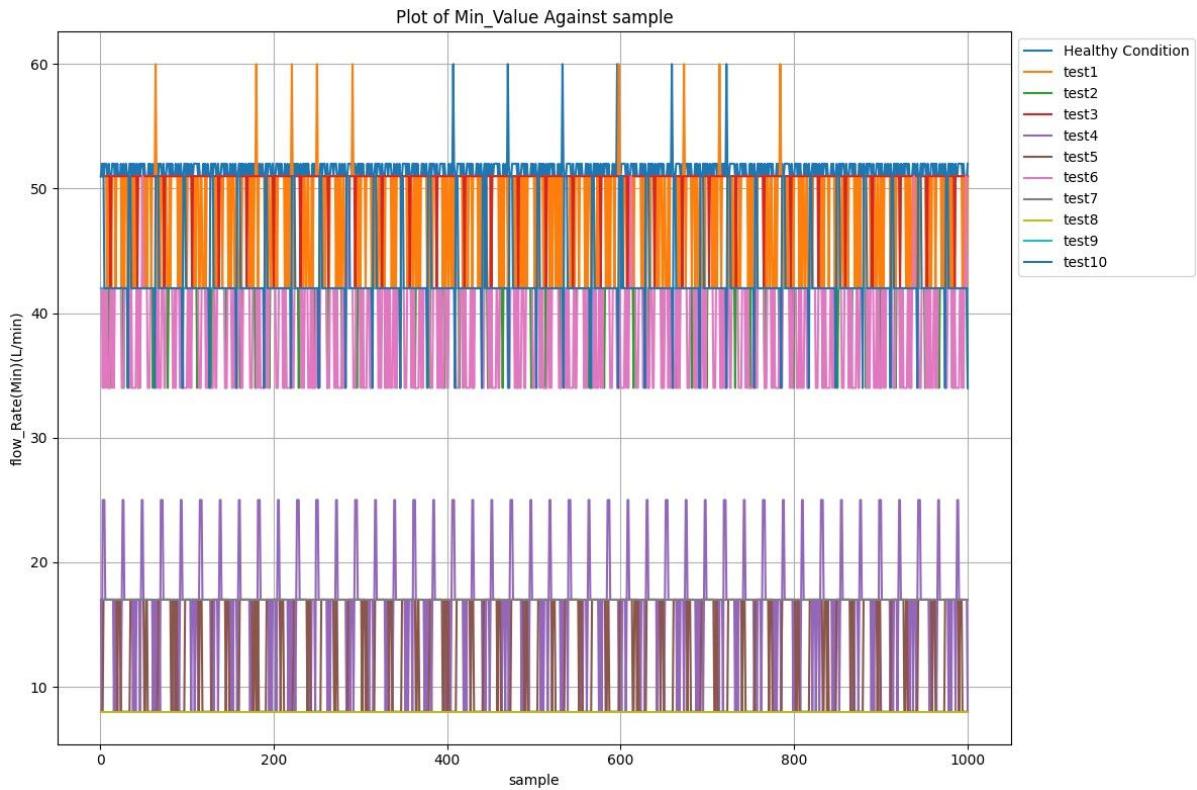


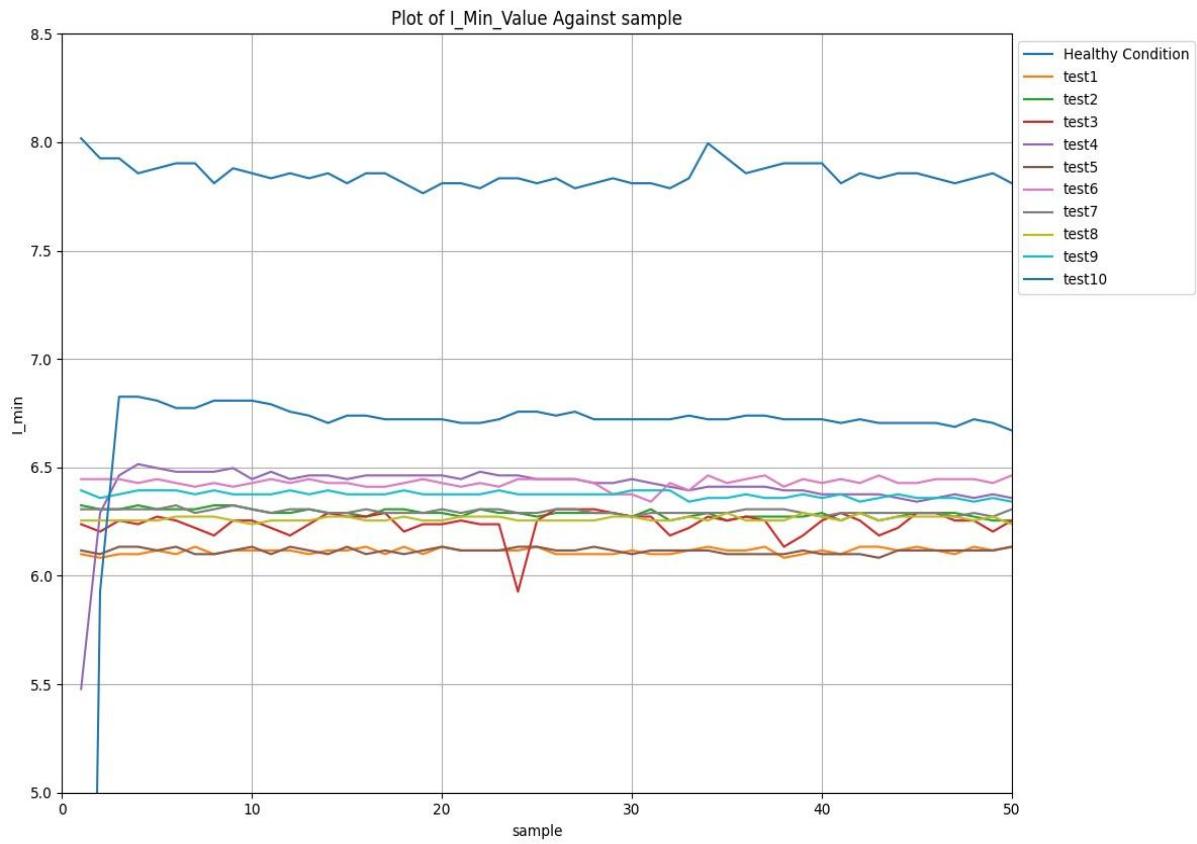
2. Feature Plot Maximum Value (Max_value)



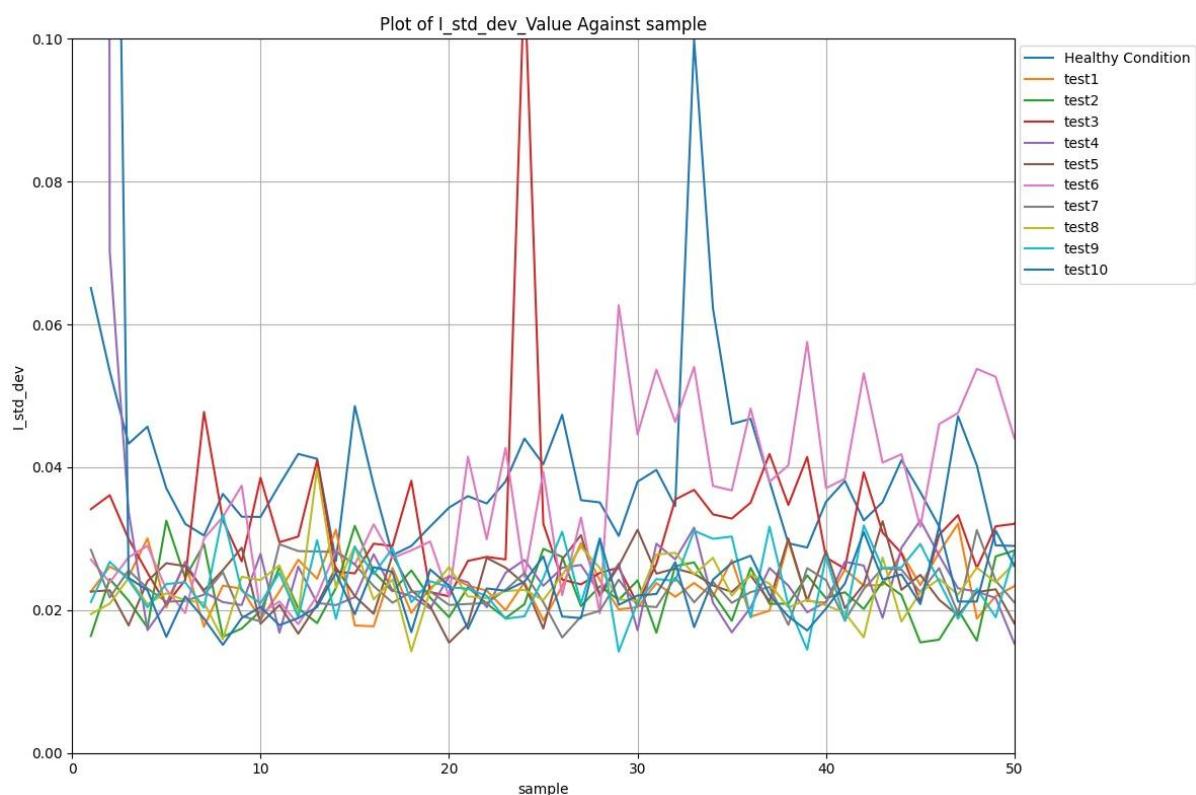
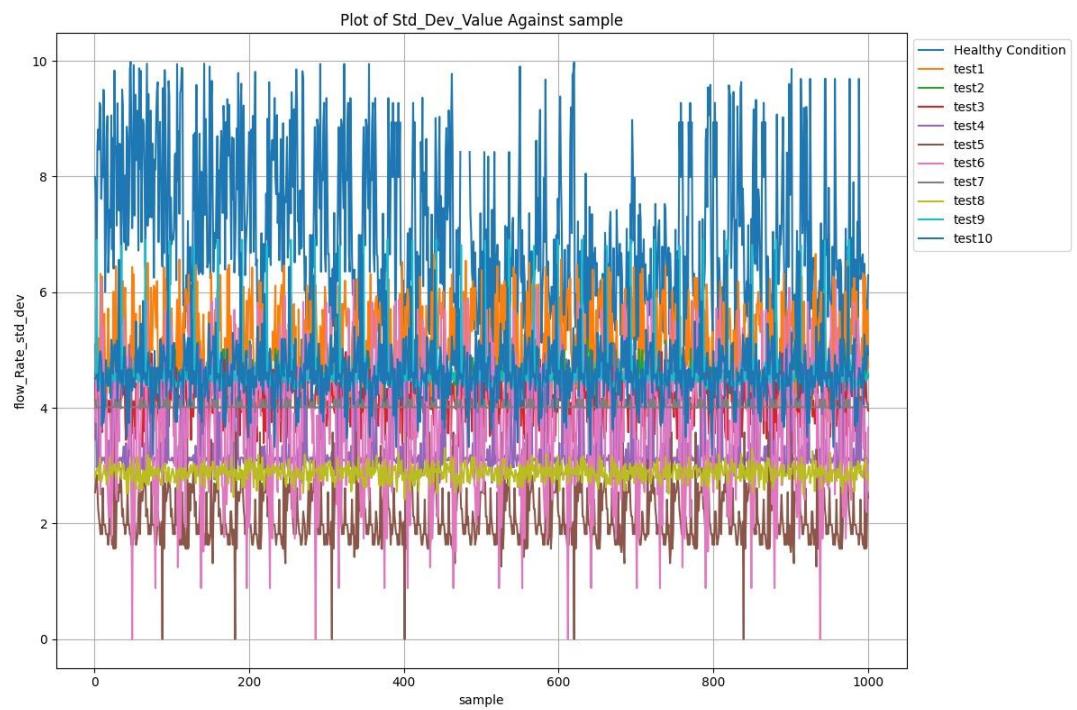


3. Feature Plot of Minimum value (Min_value)

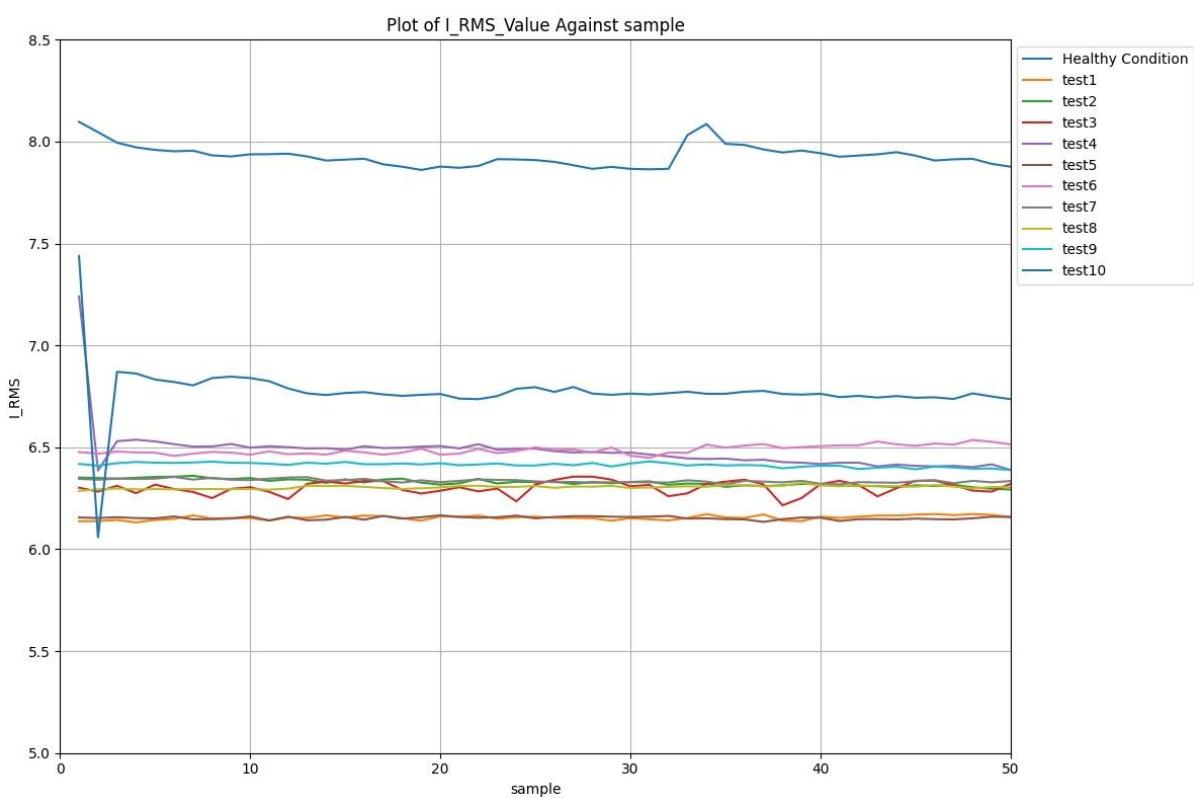
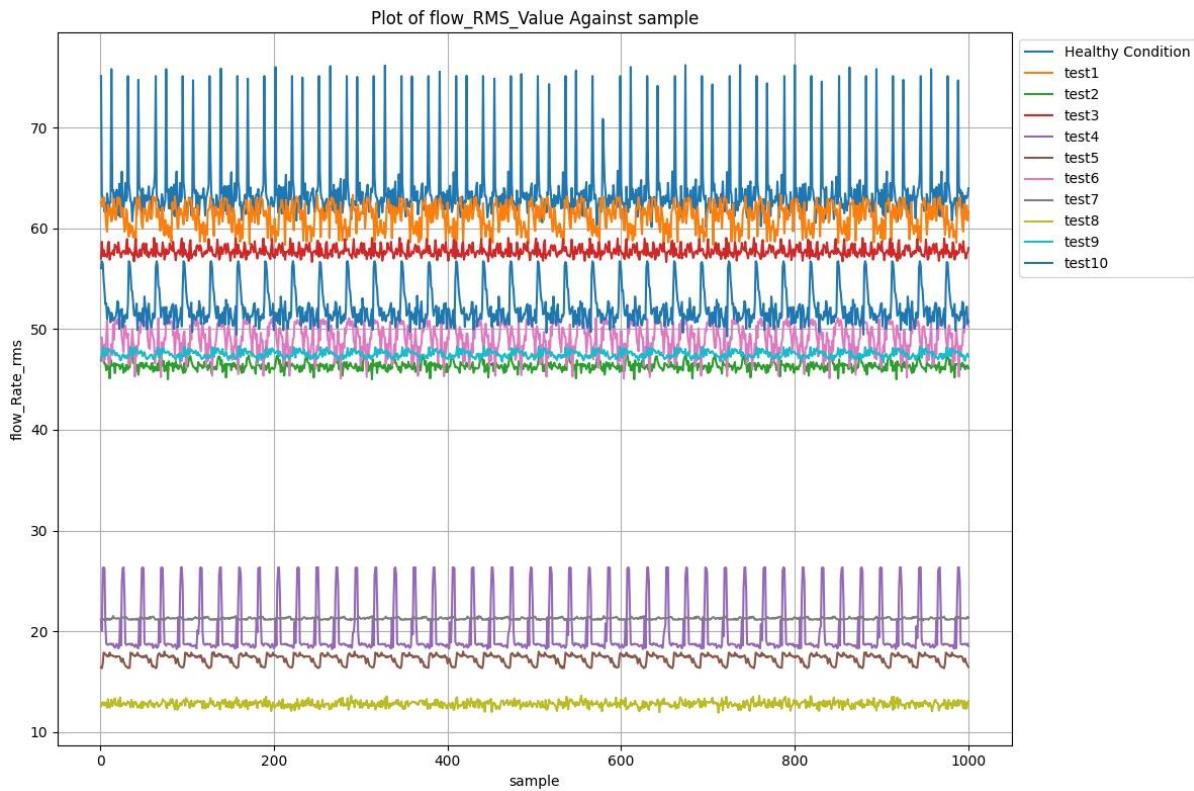




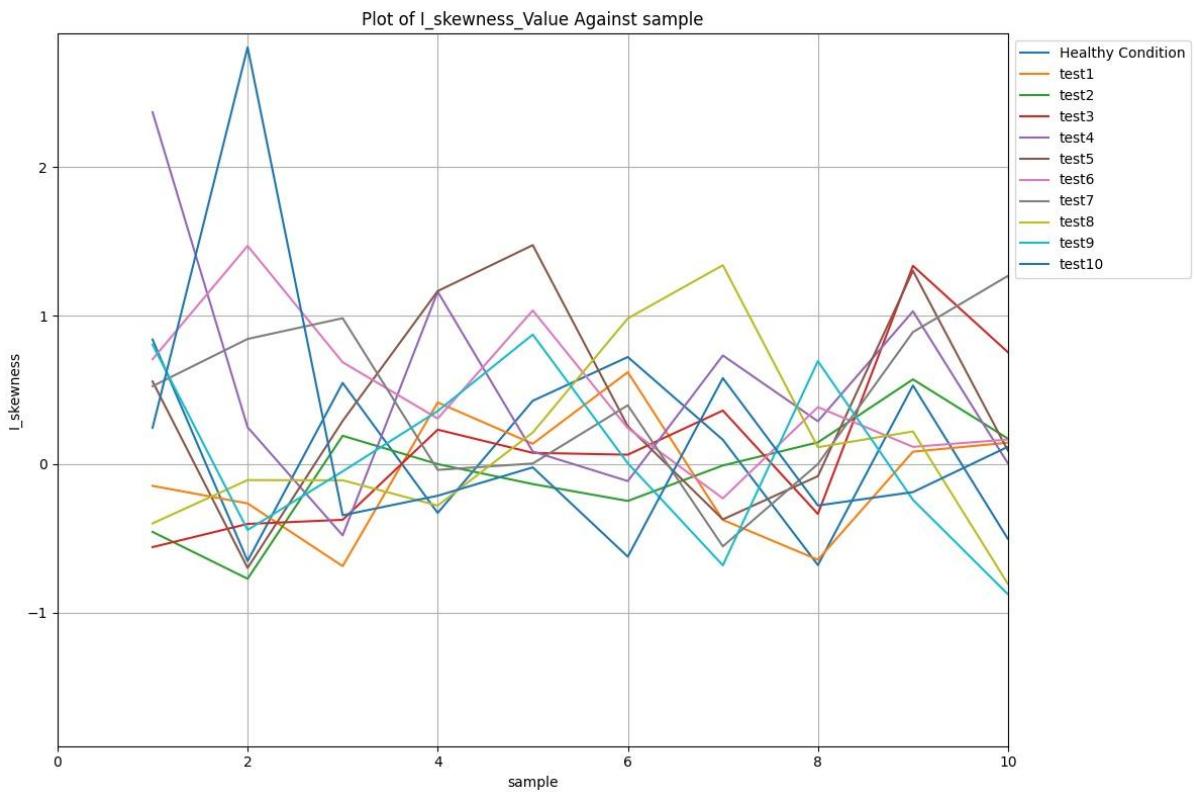
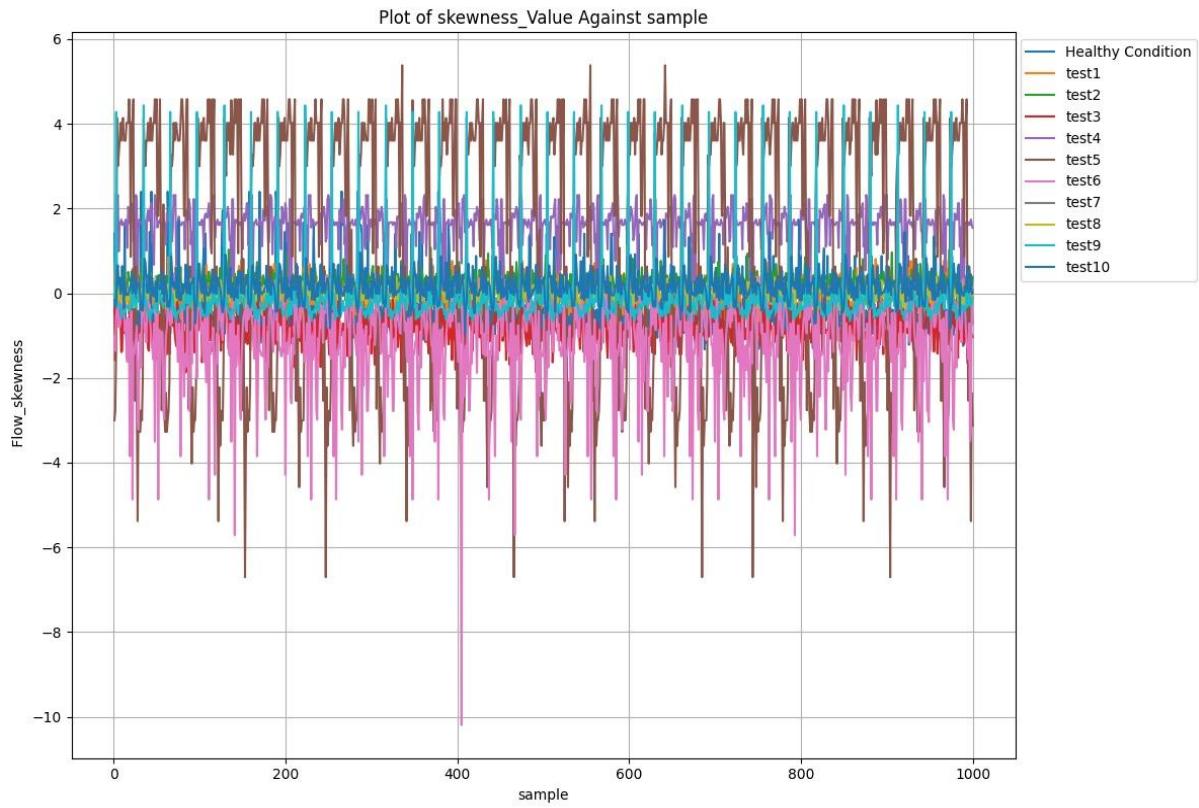
4. Feature Plot Standard Deviation (Std_dev)



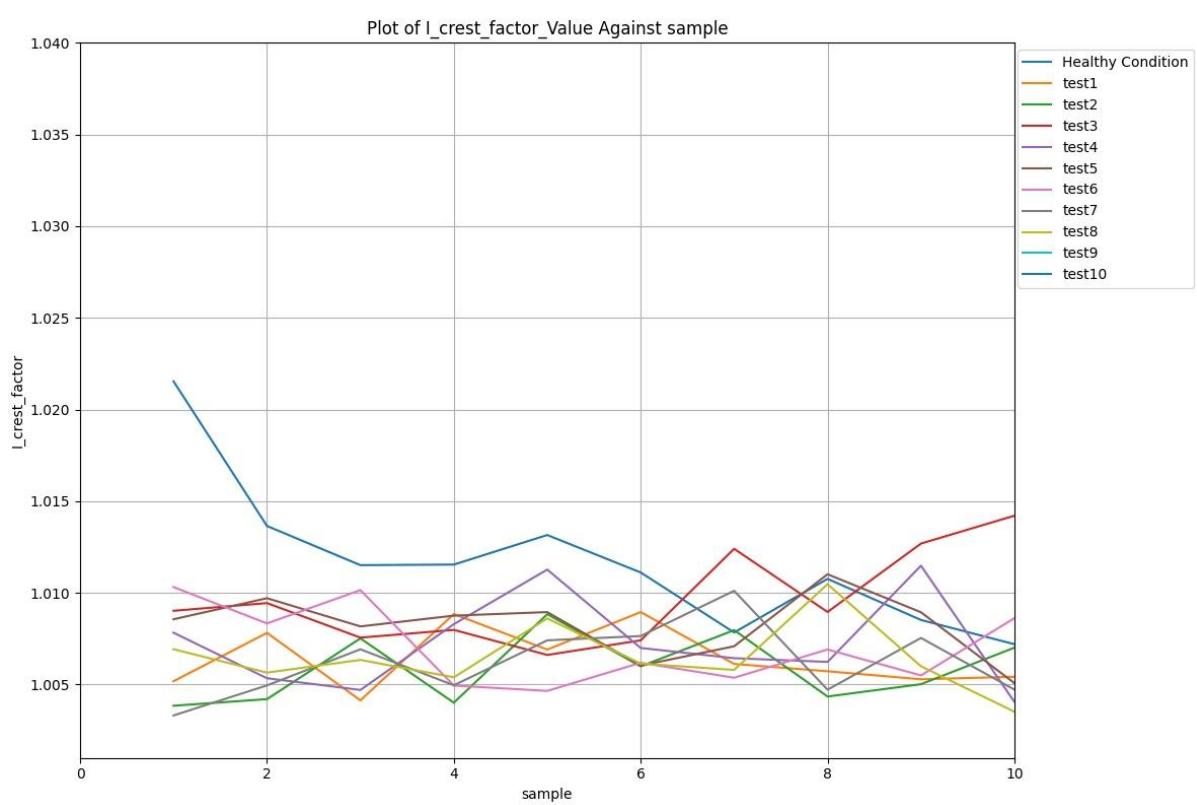
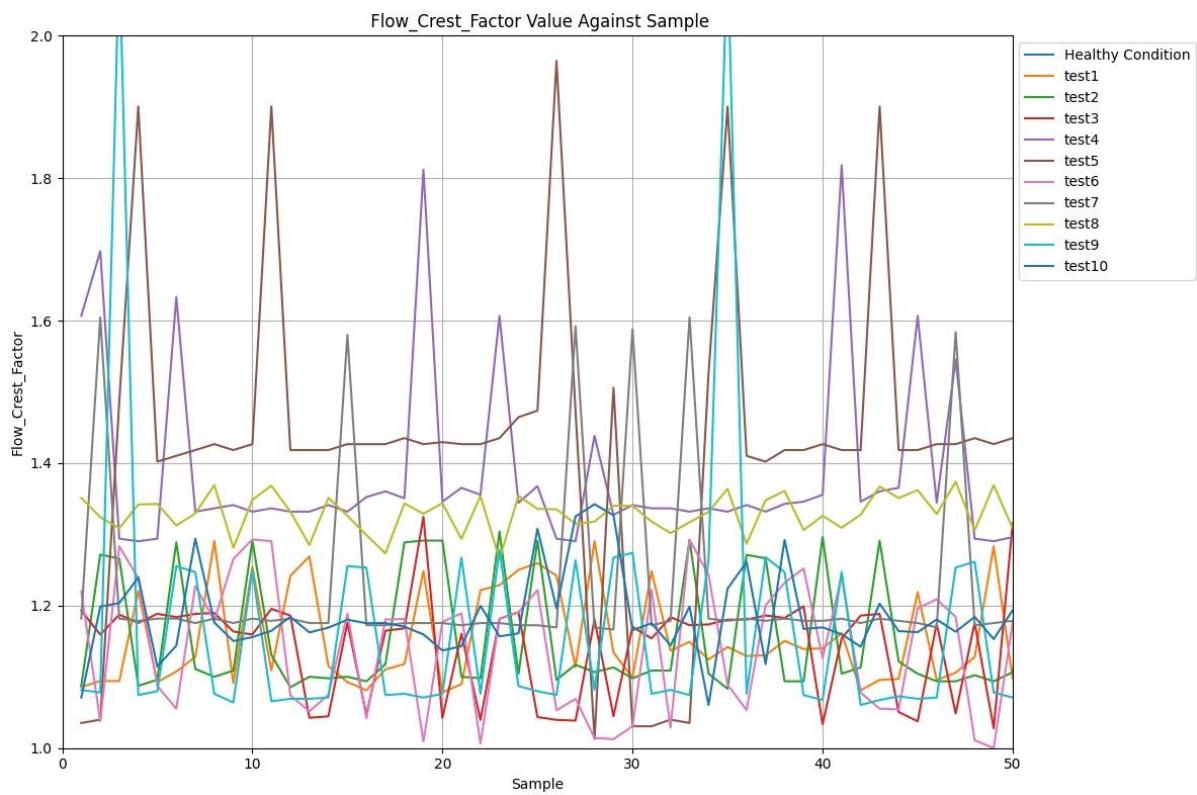
5. Feature Plot of RMS(RMS_Value)



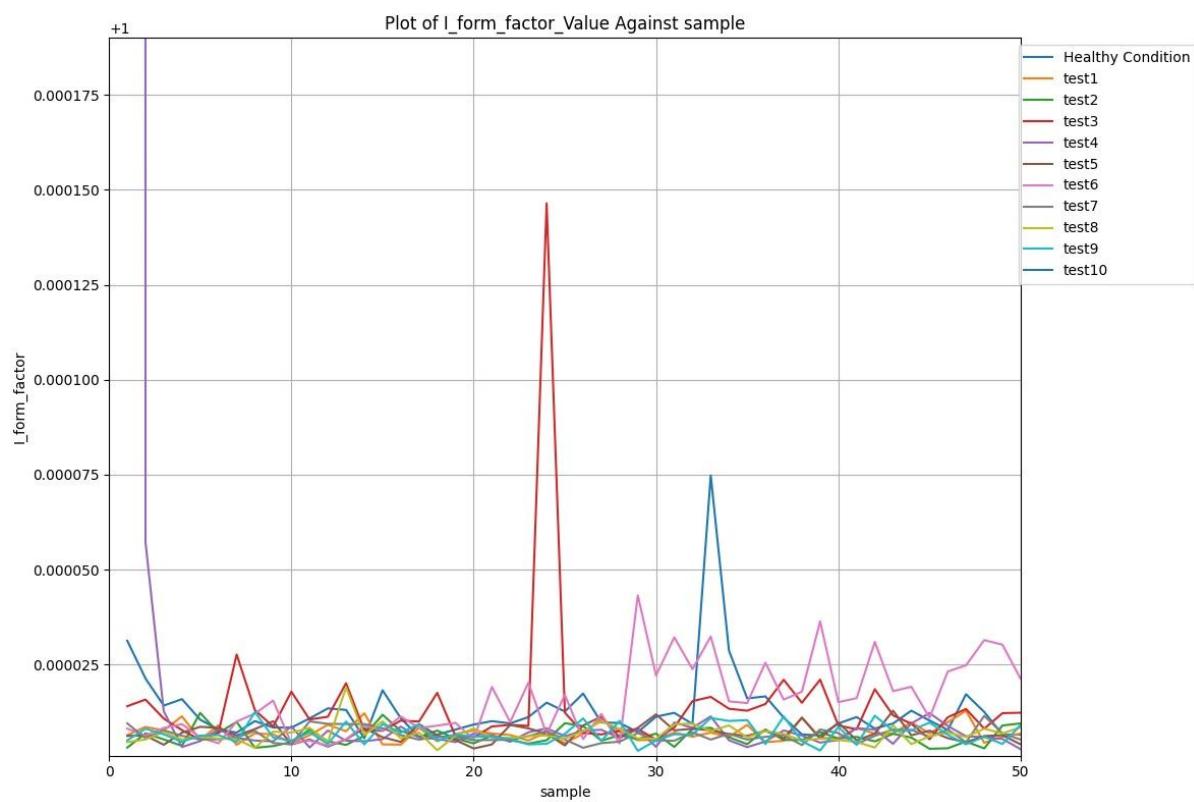
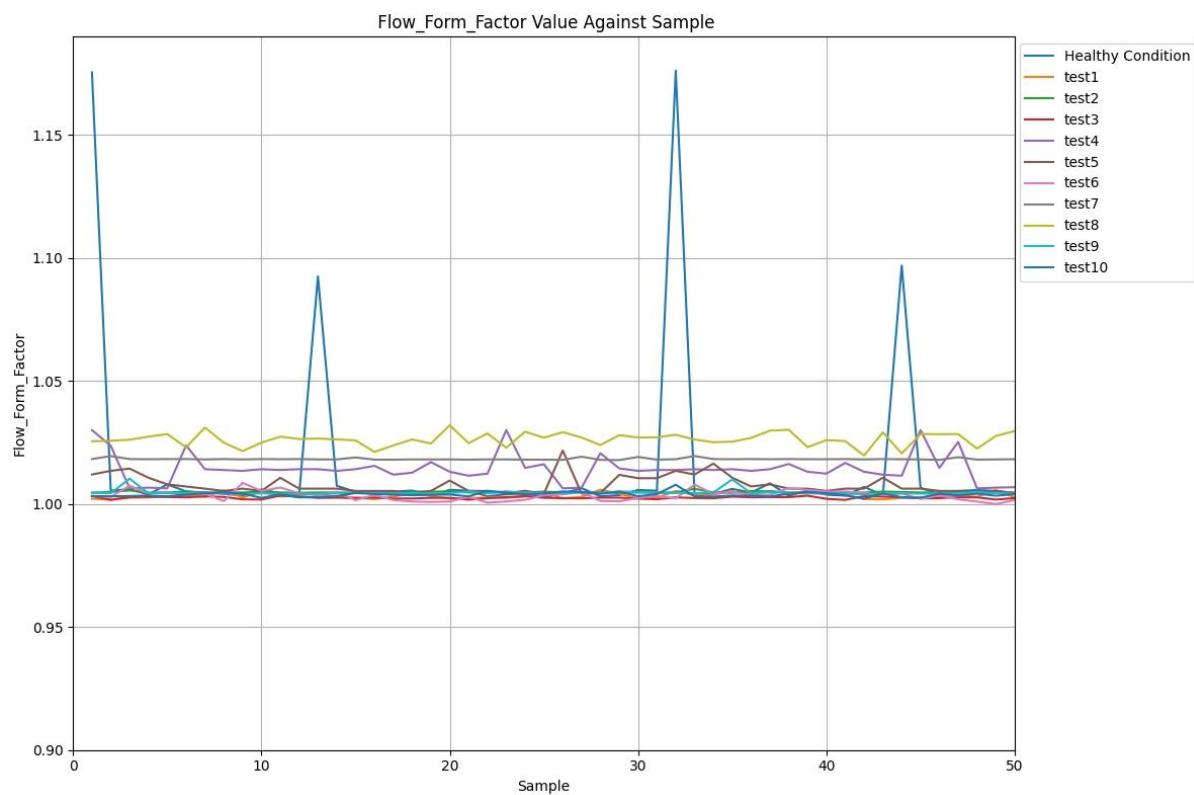
6. Feature Plot of Skewness (skewness_Value)



7. Feature Plot Crest Factor



8. Feature Plot of Form Factor



4.7 Feature Plot of Accuracy in Random Forest ML Model

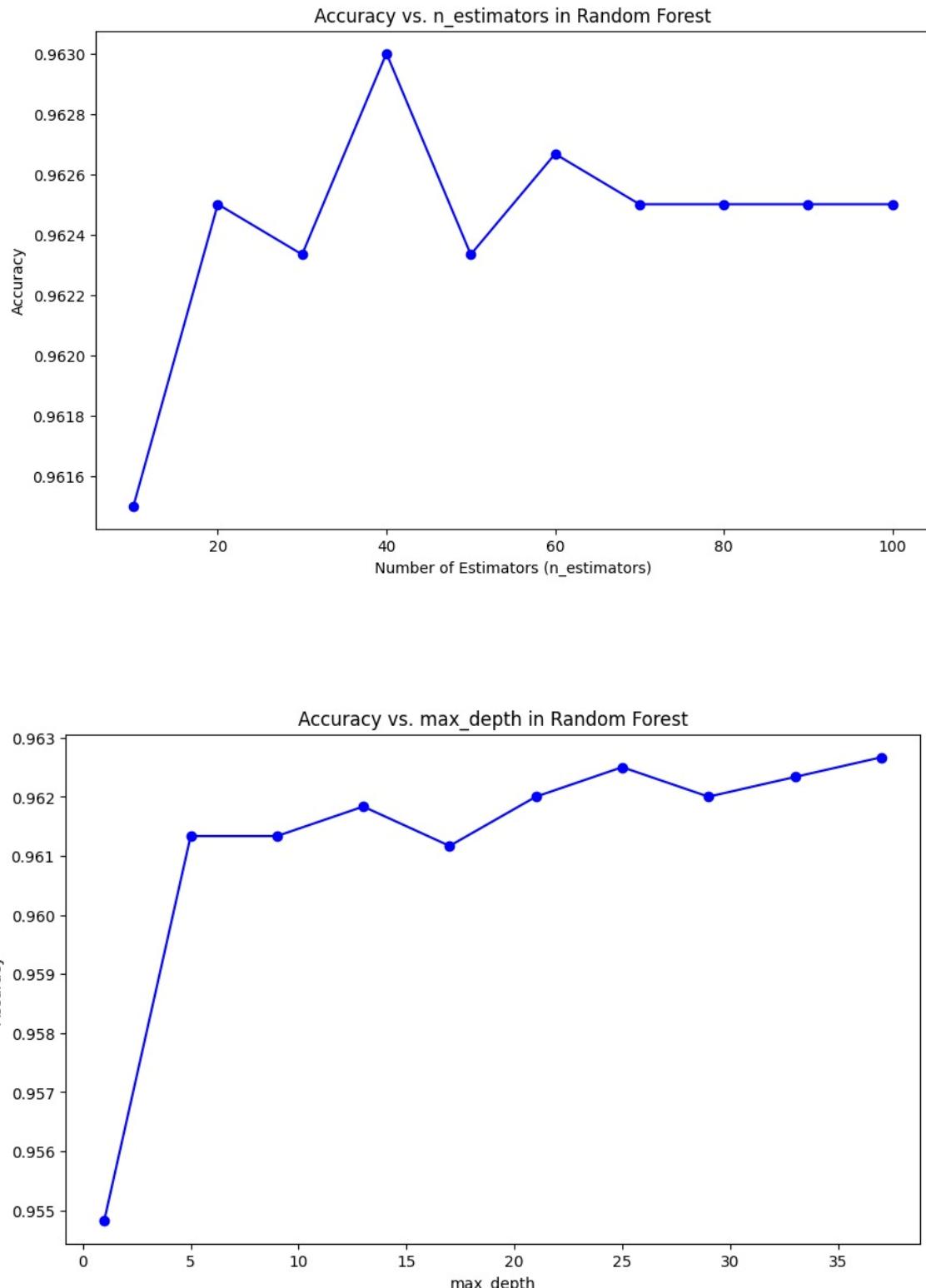


Fig 4.4: Plot of hyperparameter for different Accuracy

4.8 Performance Matrix and Performance Matrix

4.8.1 Performance matrix of Different Algorithm for Flow Rate

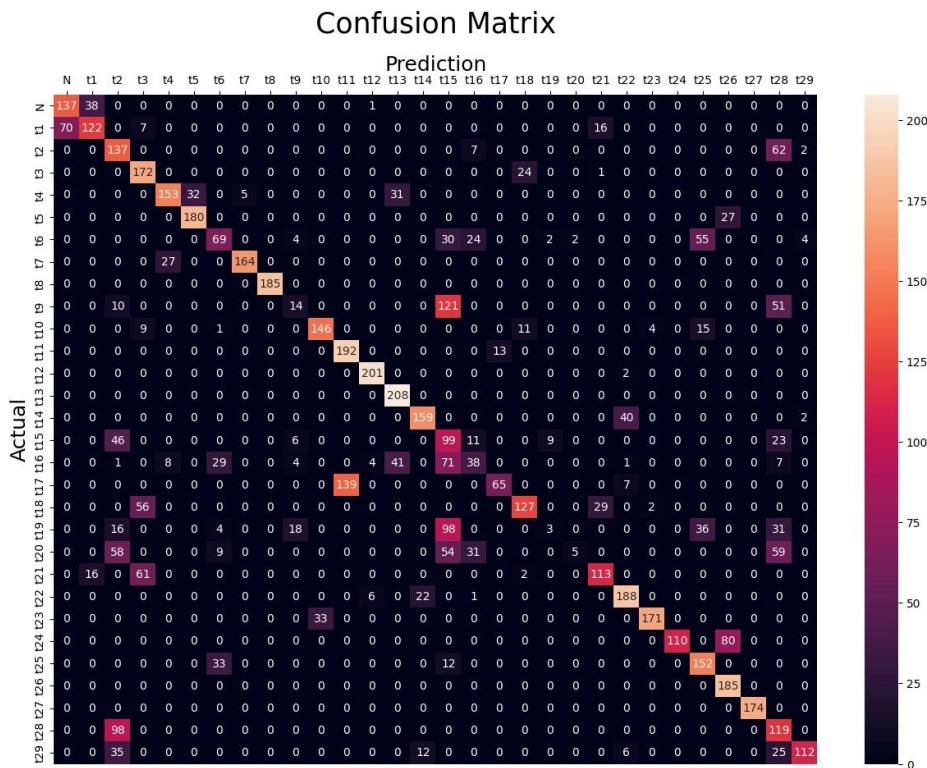
S.No	Algorithm	Accuracy	Precision	F1 score	Recall
1	Support Vector Classifier (SVC)	65%	68.32%	65%	63%
2	K Nearest Neighbours(KNN)	94.06%	94.42%	94.06%	94.07%
3	Random Forest	96.18%	96.61%	96.183%	96.17%
4	Extreme Gradient Boosting(XGBoost)	95.66%	96.42%	95.96%	95.63%

Table 4.7: Comparison of Performance matrix for flow rate

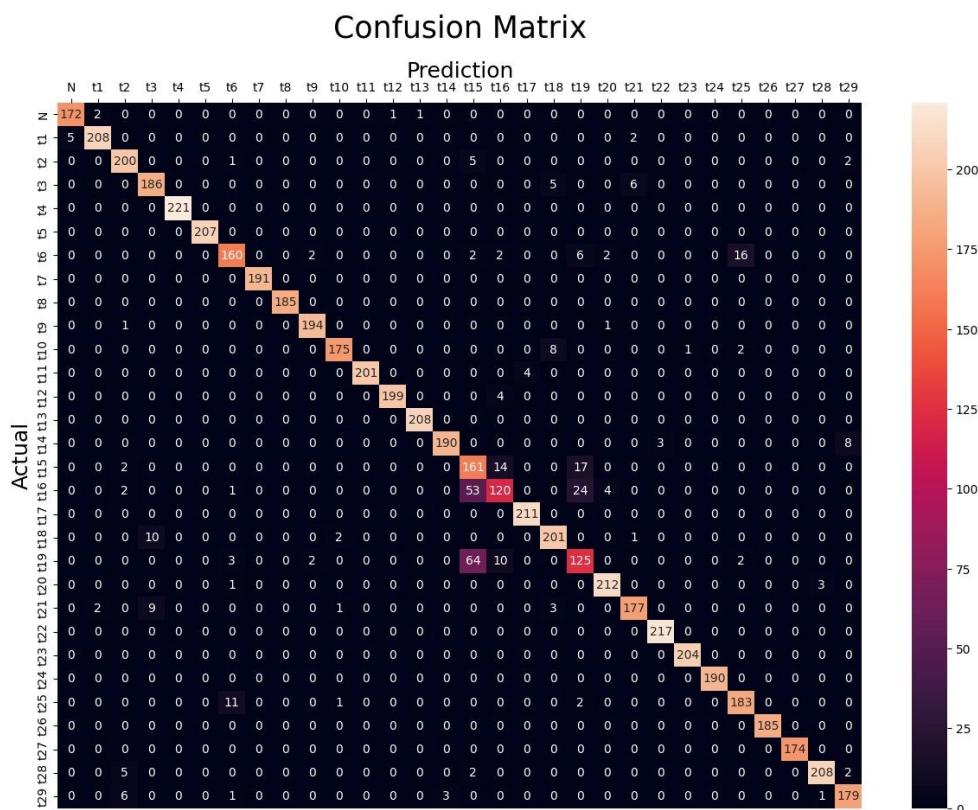
Machine Learning Model	Hyper Parameter
Random Forest	Criteria: entropy max-depth: 30 n_estimator :40
XG Boost	Learning rate:0.1 max_depth:9 n_estimator:1000

Table 4.8: Hyper parameter for flow rate

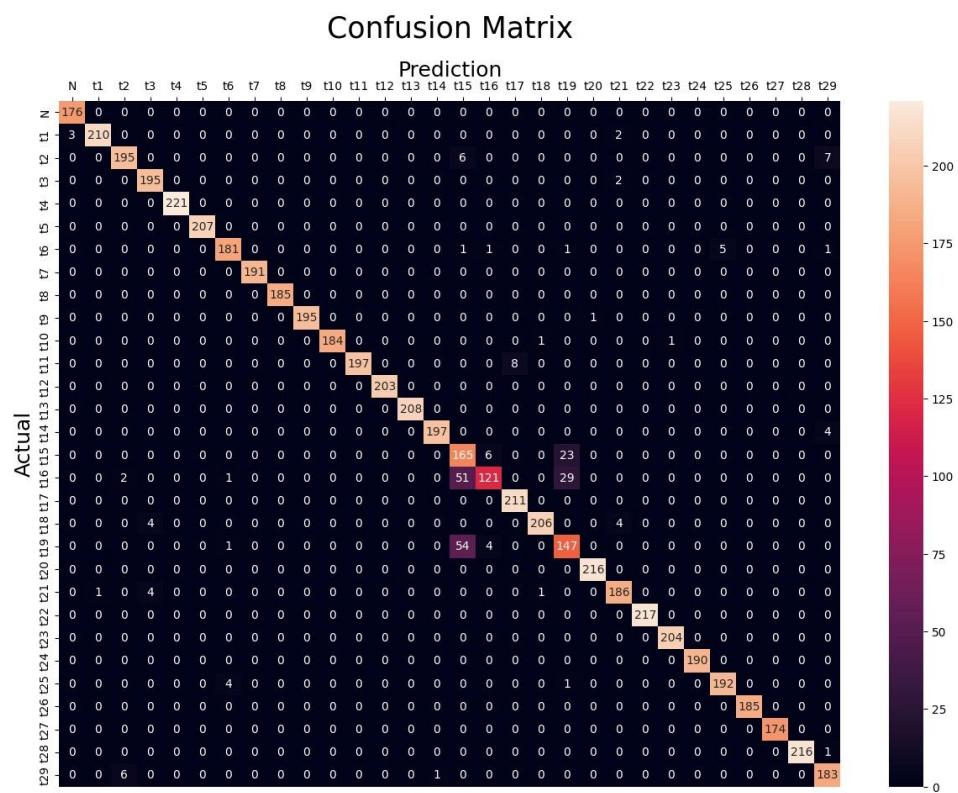
Confusion matrix for SVC Algorithm For Flow Rate



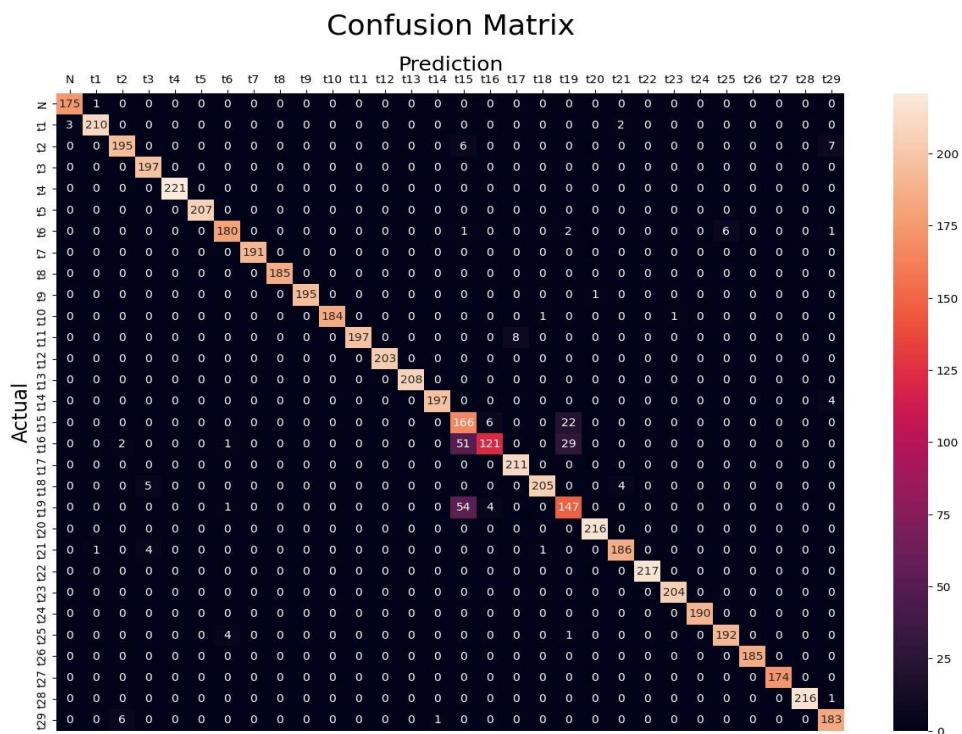
Confusion matrix for KNN Algorithm For Flow Rate



Confusion matrix for Random Forest Algorithm For Flow



Confusion matrix for XGBoost Algorithm For Flow



4.8.2 Performance matrix of Different Algorithm for Current

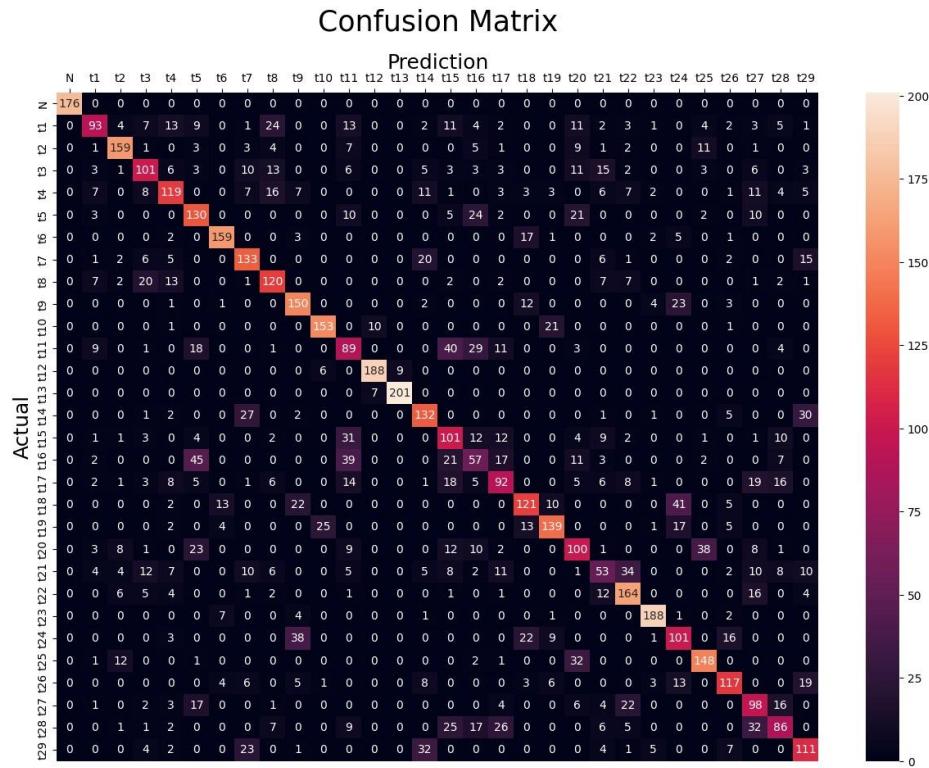
S.No	Algorithm	Accuracy	Precision	F1 score	Recall
1	Support Vector Classifier(SVC)	24.5%	21.9%	25%	23.8%
2	K Nearest Neighbours(KNN)	32%	31.5%	29.8%	28.9%
3	Random Forest	60.2%	60.1%	59.8%	60.2%
4	Extreme Gradient Boosting(XGBoost)	62.9%	63.1%	62.7%	62.8%

Table 4.9 Performance matrix of Different Algorithm for Current

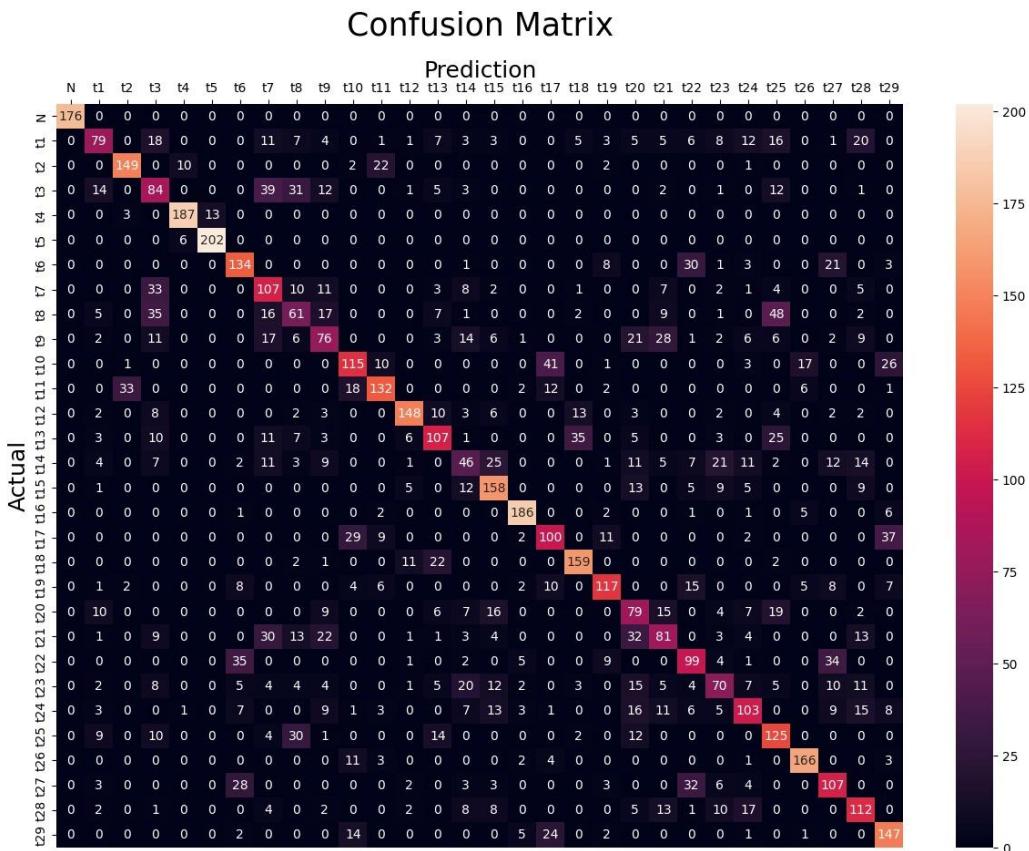
Machine Learning Model	Hyper Parameter
Random Forest	Criteria: entropy max-depth: 30 n_estimator :40
XG Boost	Learning rate:0.1 max_depth:9 n_estimator:1000

Table 4.10: ML model with their Hyperparameter

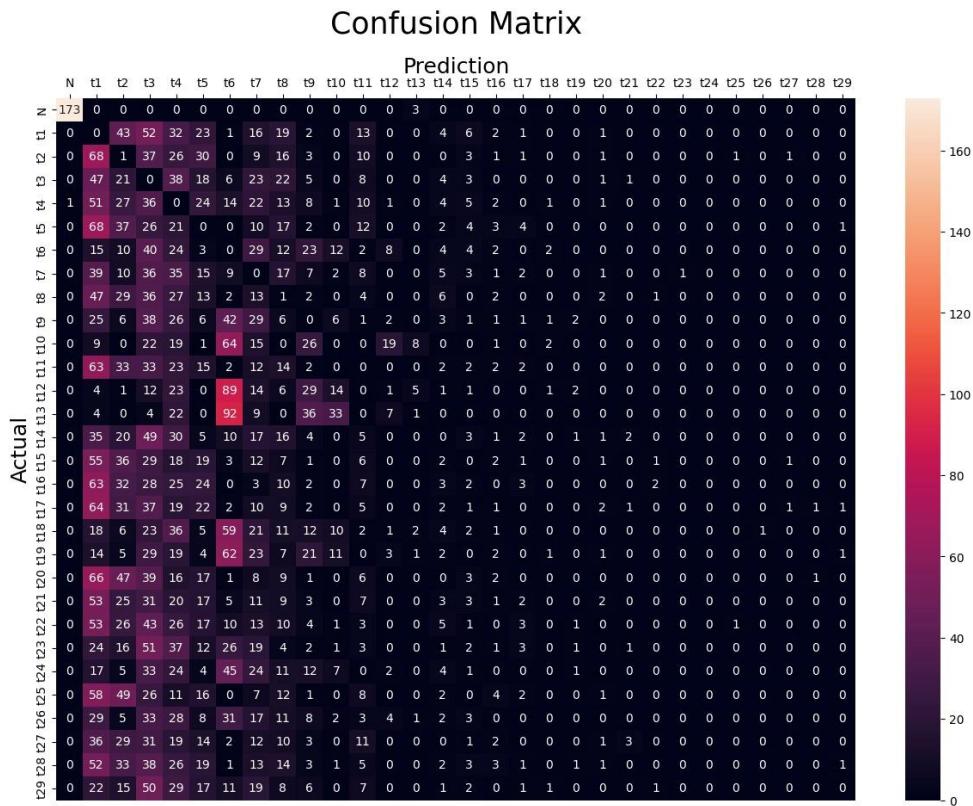
Confusion matrix for XGBoost Algorithm for Current



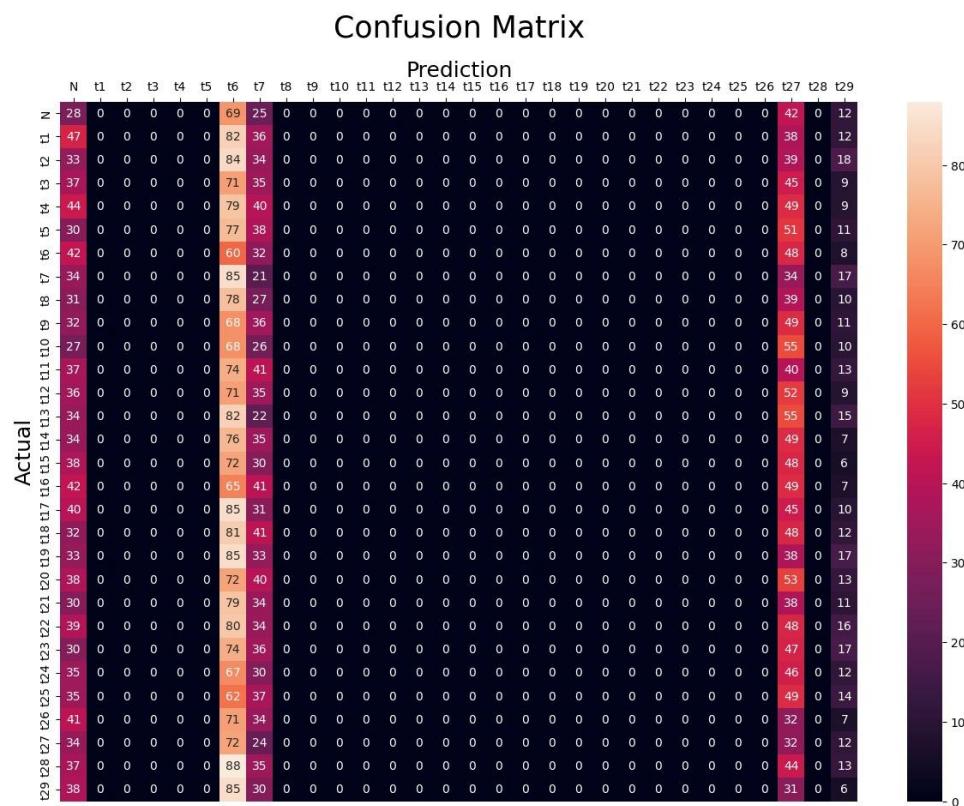
Confusion matrix for Random Forest Algorithm for Current



Confusion matrix for KNN Algorithm for Current



Confusion matrix for SVC Algorithm for Current



4.8.2 Performance matrix of Different Algorithm for Combine Fault

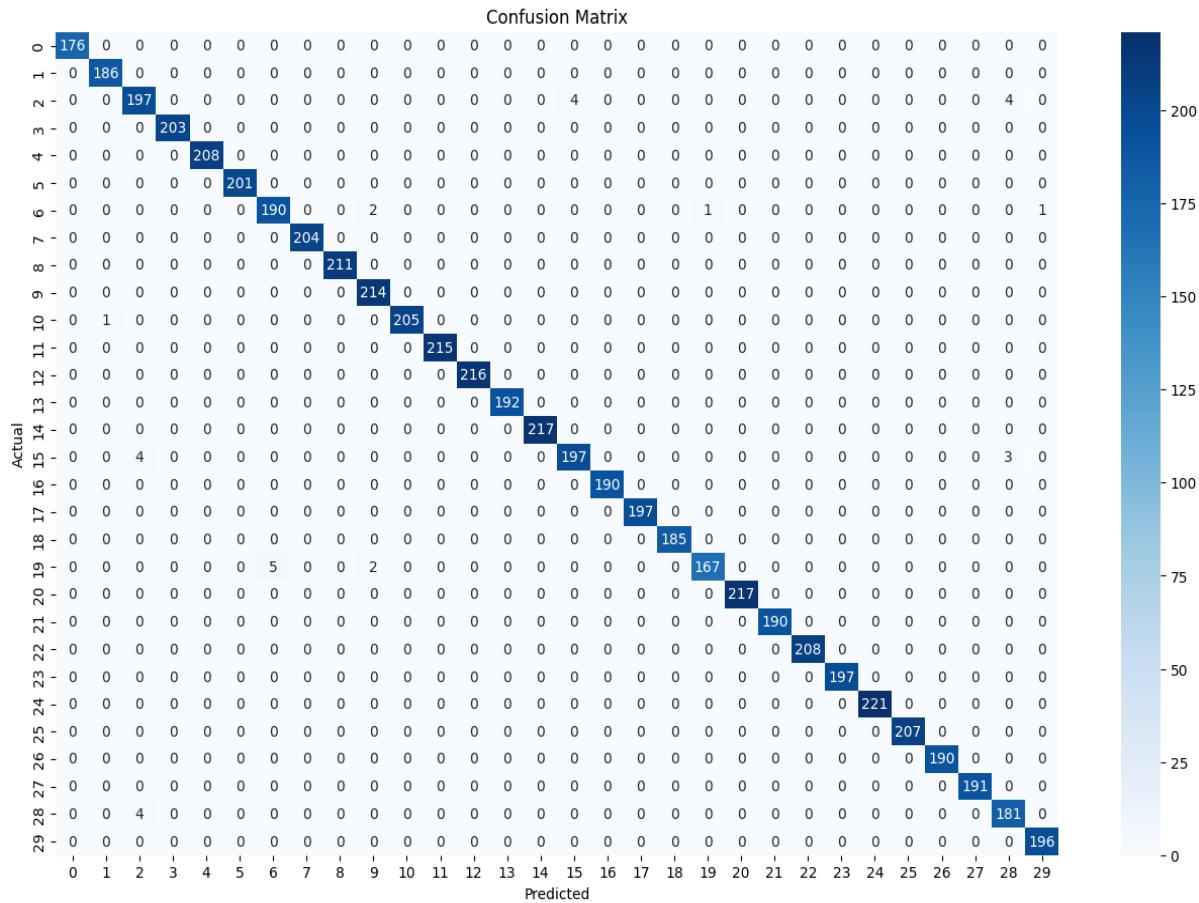
S.No	Algorithm	Accuracy	Precision	F1 score	Recall
1	Support Vector Classifier(SVC)	98.55%	98.66%	98.56%	98.55%
2	K Nearest Neighbours(KNN)	79.75%	81.79%	79.75%	79.75%
3	Random Forest	99.48%	99.49%	99.48%	99.48%
4	Extreme Gradient Boosting(XGBoost)	99.02%	99.02%	99.03%	99.5%

Table 4.11: Performance matrix of Different Algorithm for Combine Fault

S.No	Machine Learning Algorithm	Hyper Parameter
1	Support vector Classifier(SVC)	Regularisation parameter ‘C’: 5 kernel: linear
2	K Nearest Neighbors(KNN)	n_neighbors: 5 Distance metric ‘p’ : 1
3	Random Forest	criterion: entropy max_depth: None, n_estimator: 100
4	XG Boost	learning_rate: 0.2 max_depth: 3 n_estimator:100

Table 4.20: ML model with their Hyperparameter

Confusion matrix for XGBoost Algorithm for Combine Fault



Conclusion

1. Model Performance:

Random Forest and XGBoost consistently outperformed other machine learning models, including SVC and KNN, in detecting faults in submersible pumps. Their superior performance was demonstrated by achieving the highest accuracy, precision, recall, and F1 scores. XGBoost, in particular, stood out due to its ability to handle complex and non-linear relationships in the dataset while offering faster computations compared to other algorithms.

2. Sensor Effectiveness:

- Flow Sensors: Flow sensors emerged as the most effective tool for detecting blockage faults. They provided accurate and timely identification of flow irregularities, which are crucial indicators of potential blockages.
- Pressure Sensors: Pressure sensors were also highly effective but slightly less sensitive than flow sensors. They played a significant role in identifying abnormal pressure variations that signal blockages and other faults.

- Current Sensors: Although current sensors were slightly less effective in blockage detection, they proved valuable in identifying faults like bush bearing wear, impeller damage, and diffuser issues, showcasing their broader utility in fault classification.

3. Key Features:

The analysis revealed that specific statistical features—such as mean, maximum, minimum, and RMS values—had the most significant impact on model performance. These features effectively captured variations in sensor data, allowing models to differentiate between fault types with high accuracy.

4. Dataset Integration:

Combining datasets from different fault scenarios enhanced the models' ability to detect and classify faults, particularly for blockage and impeller-related issues. This integration allowed the algorithms to generalize better across diverse conditions, improving their reliability and robustness in real-world applications.

5. Algorithm Superiority:

Among the four algorithms studied, XGBoost demonstrated exceptional performance. Its boosting mechanism enabled it to correct errors iteratively, leading to highly accurate predictions. Random Forest also performed strongly due to its ensemble approach, which minimized overfitting and managed complex feature interactions effectively.

6. Sensor Data Insights:

Flow sensors emerged as the most appropriate for fault detection, as they provided the earliest and most accurate indicators of blockages. Pressure sensors were ranked second, offering strong support in detecting pressure abnormalities. Current sensors, while less effective for blockage, provided critical data for detecting mechanical faults like impeller, diffuser, and bush bearing failures.

7. Operational Insights:

The study highlights the importance of monitoring key parameters such as flow rate, pressure, and current in submersible pumps. By leveraging these sensor inputs, machine learning models can detect faults early, reducing downtime and preventing costly repairs.

Future Scope:

- i. Prediction of faults and remaining life of System using ML.
- ii. Develop an accurate digital twin model using machine learning to simulate the behavior of the physical asset.
- iii. Integrate the physical twin and digital twin through an IoT network
- iv. Indication of Fault in virtual Model
- v. Incorporating additional sensors such as temperature and vibration sensors could further enhance fault detection accuracy and broaden the range of detectable fault types.

- vi. Exploring more advanced algorithms, including deep learning approaches, could improve predictive capabilities, especially in handling highly complex datasets.
- vii. Developing real-time fault detection systems using the studied models can enhance the operational efficiency of submersible pumps in industrial settings.
- viii. Investigating the long-term effects of parameter fluctuations on pump components may provide deeper insights into preventive maintenance strategies.

This comprehensive study demonstrates that machine learning models, particularly XGBoost, combined with appropriate sensor data (flow, pressure, and current), offer a powerful framework for detecting and classifying faults in submersible pumps.

References:

1. [1] *Sustainability* 2023, 15(15), 11845; <https://doi.org/10.3390/su151511845>
2. [2] *48th IEEE Industrial & Commercial Power Systems Conference*
3. <https://www.sulzer.com/-/media/files/products/pumps/submersible-pumps/technical-articles/lifetime-of-efficiency-a10440.pdf?la=en>
4. Abdalla, Ramez & Samara, Hanin & Perozo, Nelson & Paz, Carlos & Jaeger, Philip. (2022). Machine Learning Approach for Predictive Maintenance of the Electrical Submersible Pumps (ESPs). *ACS Omega*. 7. 10.1021/acsomega.1c05881.
5. Brasil, J.; Maitelli, C.; Nascimento, J.; Chiavone-Filho, O.; Galvão, E. Diagnosis of Operating Conditions of the Electrical Submersible Pump via Machine Learning. *Sensors* 2023, 23, 279. <https://doi.org/10.3390/s23010279>.
6. Machine Learning Approach for Predictive Maintenance of the Electrical Submersible Pumps (ESPs). Ramez Abdalla, Hanin Samara, Nelson Perozo, Carlos Paz Carvaljal, and Philip Jaeger *ACS Omega* 2022 7 (21), 17641-17651 DOI: 10.1021/acsomega.1c05881.
7. Zhu, Y., Zhou, L., Lv, S., Shi, W., Ni, H., Li, X., & Hou, Z. (2023). Research Progress on Identification and Suppression Methods for Monitoring the Cavitation State of Centrifugal Pumps. *Water*, 16(1), 52.
8. Chen L, Wei L, Wang Y, Wang J, Li W. Monitoring and predictive maintenance of centrifugal pumps based on smart sensors. *Sensors*. 2022 Mar 9;22(6):2106.
9. <https://www.youtube.com/playlist?list=PLTDARY42LDV7WGmlzZtY-w9pemvPrKNUZ>
10. https://www.youtube.com/playlist?list=PLKnIA16_Rmvbr7zKYQuBfsVkjolcJgxHH
11. <https://www.scaler.com/topics/course/free-supervised-learning-course/>
12. <https://www.coursera.org/specializations/practical-data-science-matlab>
13. <https://medium.com/@nimrashahzadisa064/supervised-machine-learning-classification-and-regression-c145129225f8>
14. 11. *Sustainability* 2023, 15(15), 11845; <https://doi.org/10.3390/su151511845>

