PREDICTIVE MAINTENANCE OF CENTRIFUGAL WATER PUMPS USING MACHINE LEARNING: A CASE STUDY OF NATIONAL WATER AND SEWERAGE CORPORATION

QUINTON SSEBAGGALA

M23M19/279

A DISSERTATION SUBMITTED TO THE FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF A DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE AND ANALYTICS OF UGANDA CHRISTIAN UNIVERSITY

May, 2026

UGANDA CHRISTIAN UNIVERSITY

A Centre of Excellence in the Heart of Africa

Declaration and Approval

Declaration

I, **Ssebaggala Quinton**, declare that this project report titled:

> *"Predictive Maintenance of Centrifugal Water Pumps Using Machine Learning in Uganda's National Water and Sewerage Corporation System"*

is my original work and has not been submitted, either wholly or partially, for the award of any other degree or qualification in any institution. Any material used from other sources has been properly acknowledged.

This project report is submitted in partial fulfillment of the requirements for the award of the degree of Master of Science in Data Science and Analytics at Uganda Christian University.

Candidate's Name: Ssebaggala Quinton

Signature:

Date: April 29, 2026

Approval

This is to certify that the above declaration has been made by the candidate and that the project report has been submitted with my approval as the university supervisor.

Supervisor's Name: Dr.Ndibatya Innocent
Signature:
Date: 29/May 2

Acknowledgment

First and foremost, I thank the Almighty God for the gift of life, wisdom, and perseverance that enabled me to successfully complete this research project.

I am deeply grateful to my supervisor, Dr.Ndibatya Innocent, for the invaluable guidance, constructive feedback, and constant encouragement throughout the entire research journey.

My sincere appreciation goes to the academic staff of the Department of Data Science and Analytics at Uganda Christian University for the knowledge and support that formed the foundation of this study.

I extend heartfelt thanks to the National Water and Sewerage Corporation (NWSC) for granting me access to data and operational insights from the Ggaba and Katosi water treatment plants. Special recognition goes to the staff at these facilities for their cooperation during the data collection process.

To my classmates, colleagues, and friends thank you for your camaraderie, discussions, and moral support during the course of this program.

Lastly, I am indebted to my family for their unwavering love, patience, and moral support. Their belief in me was my greatest motivation.

Ssebaggala Quinton

Abstract

This thesis explores Effective predictive maintenance strategies for Centrifugal water pumps, focusing on Uganda's National Water and Sewerage Corporation (NWSC) and other similar large-scale water providers, aiming to improve water supply reliability for over 21 million people and reduce 185,000 annual customer complaints caused by 70% pump failures and 8-12 hours of operational downtime. However, despite advances in machine learning, tailored predictive maintenance approaches for water pumps in Uganda are understudied. Thus, this study developed a based predictive maintenance model for the centrifugal pumps using real-time operational data from National Water and Sewerage Corporation (NWSC) (N=13 pumps from the 3 pump stations i.e. (Gunhill, Katosi and Muyenga)were analyzed. This study presents a comprehensive Machine-learning based predictive maintenance framework for estimating pump failure. The process integrates data preprocessing, extraction of statistical time-domain condition indicators, and evaluation of 5 machine learning algorithms; XGBoost, LightGBM, CatBoost, Random Forest, and a Voting Ensemble [applied to shift maintenance from a monthly health check to real-time monitoring] providing deeper insights into pump availability and health for future years. The primary objective was to accurately classify the pumps' operational status into five distinct states: CHANGE, CRITICAL, OFF, OPERATIONAL, and WARNING. The results demonstrate that Extreme Gradient Boosting (XGBoost) model achieved superior predictive performance yielding an accuracy of 74% in detecting failure within pumps before more damage was done. Thus, leveraging of Machine Learning for Predictive maintenance enabled National Water and Sewerage Corporation to detect any anomalies in the Centrifugal pumps like; inconsistencies in flow rates, pressure fluctuations, vibration abnormalities etc. which helped reduce on the maintenance costs from (10-40%), reduce on equipment failure (70-75%), reduced on downtime (35% -45%) and lastly, increased on production capacity by(25%) thus improving on the well-being of the people in Uganda and promoting of SDG 6(Clean water and Sanitation).

Keywords: Predictive Maintenance, Machine Learning, Centrifugal Pumps, National Water and Sewerage Corporation, Arduino, Classification Models

Contents

List of Figures

List of Tables

List of Acronyms

AI Artificial Intelligence. 2

CNNs Convolutional Neural Network. 3, 19

GSM Global System for Mobile Communications. 24

IoT Internet of Things. 15

ML Machine Learning. 2, 10, 27

MLP Multilayer Perceptron. 20

NWSC National Water and Sewerage Corporation. ii–v, 1–8, 10, 11, 21, 23, 26–28, 30, 32, 34–37, 40, 42, 64–66, 68

PdM Predictive Maintenance. 2–4, 10, 27, 45, 64, 66

SVM Support Vector Machine. 10, 20, 21

WHO World Health Organization. 12

Chapter 1

Introduction

1.1 Background

"Pure water is the world's first and foremost medicine." according to a [Slovakian Proverb, nd]. However, according to [UNICEF & WHO, 2019] reports, over "2.2 billion people worldwide lack access to clean and safe water" and that includes Uganda. Uganda's National Development Plan prioritizes water supply, highlighting the importance of safe water, adequate sanitation, economic development, poverty reduction and environmental sustainability according to [National Water and Sewerage Corporation, 2024]. Further, the National Water and Sewerage Corporation (NWSC) is a public Corporation entirely belonging to the Government of Uganda instituted by the Decree No.34 of 1972, according to the [National Water and Sewerage Corporation, 2022]. The National Water and Sewerage Corporation (NWSC) currently provides services in 282 towns across 96 districts within the country. Thus, integrating these towns into 72 Administrative Areas under five regions, with a total of 150 branches nationwide, for streamlined operations and improved customer access. The main purpose of NWSC is to supply clean and fresh water to the customers in both urban and rural areas. Thus, the treatment of water, to ensure safety for human consumption, and meet the Standards of both National and International markets. Furthermore, National Water and Sewerage Corporation (NWSC) strives to sustainably and equitably offer cost-effective quality water and sewerage services to the delight of all stakeholders, while preserving the environment and strengthening stakeholder trust, through currently meeting the needs of roughly 21 million people. However, this cornerstone of contemporary societal well-being and public health is frequently jeopardized by unexpected breakdowns of critical infrastructure, most notably essential machinery like centrifugal pumps within the water treatment and distribution pump stations. And hence these breakdowns result in substantial service disruptions, revenue losses, for National water and Sewerage Corporation, financial burdens on the Government, and adverse impacts on public health, ultimately affecting the quality of life of millions of citizens and hindering progress towards Sustainable Development Goal 6. A key motivation for this research is to mitigate these issues, ensuring a consistent, uninterrupted, and safe water supply which requires

robust maintenance strategies to prevent pump failures.

Following this introduction, maintenance has become a vital component of equipment lifecycle management, particularly for critical infrastructure like the centrifugal pumps which are used in water treatment and distribution. According to [Burckhardt et al., 1983] Centrifugal pumps are dynamic, axisymmetric work-absorbing turbomachines used for transporting liquids by converting rotational kinetic energy from an impeller into hydrodynamic pressure energy. In addition, centrifugal pumps are the most common type of pumps due to their simplicity, capacity, high-efficiency, flowrate, and ease of operation. And they have been employed in various applications, making them ubiquitous; however, in many systems, they are just one component of a complex process, but their failure can cause serious problems. For instance, according to [Gülich, 2010] water pumps require maximum reliable performance. And [Yedidiah, 1996]'s demonstration adds that pump problems can reduce efficiency and cause a breakdown in the pumps which affect centrifugal pumps when in use and thus lead to both mechanical and hydraulic problems.

Given their role in water treatment and distribution networks, centrifugal pumps require effective maintenance practices, such as continuous monitoring and routine intervention to sustain optimal performance and operational continuity. While recognized for their natural dependability and efficiency, these pumps remain at risk to a range of mechanical vulnerabilities, like misalignment, wear and tear, imbalance, and excessive operational loads which can trigger significant service disruptions, particularly in regions already struggling with limited access to clean water according to [Einabadi, 2023, Sakthivel et al., 2010].

The National Water and Sewerage Corporation (NWSC), being Uganda's primary authority for water treatment and distribution, operates mainly through the Ggaba and Katosi plants as reported by [National Water and Sewerage Corporation, 2024]. Both facilities face significant challenges that is: Ggaba struggles with aging infrastructure designed for a smaller population, while the newer Katosi plant remains reliant on preventative rather than Predictive maintenance. Currently, National Water and Sewerage Corporation (NWSC)'s dependence on traditional, time-based maintenance often leads to reactive interventions after equipment failure, resulting in service delay, disruptions and financial losses. These persistent issues underscore the need for Predictive Maintenance (PdM), through leveraging real-time sensor data and machine learning interventions after, equipment failures.

Further, Predictive Maintenance (PdM) being a data-driven approach offers a safer, more cost-effective alternative that reduces downtime [Menanno et al., 2023, Sang, 2022]. However, for Predictive Maintenance (PdM) to be effective, access to modern sensor patterns is needed to highlight the distinction with emerging global patterns,[Water and Corporation, 2018].Thus adopting of Predictive Maintenance (PdM), leveraging machine learning to predict future failure before they happen through the use of real-time sensor data. Sensor-data and comprehensive sensor models provide adequate conditions for modern pump models, even in third-world countries like Uganda, where stable power supply is a must and computing resources are limited. Artificial Intelligence (AI), particularly Machine Learning (ML), is increasingly aiding maintenance engineers in predicting failures. This study will leverage ensemble learning models, specifically Random

Forest, CATBoost, XG Boost, Light GBM and Voting Ensemble due to their proven accuracy in handling complex, large sensor data and their ability to provide feature importance, which is crucial for maintenance insights. These systems analyse sensor data, detect irregularities, and optimize maintenance schedules according to [Chhabria et al., 2022]. The concept of Predictive Maintenance (PdM) is to use data to make more and better predictions about occurred and future maintenance actions as well as having better understanding of why failures happened in the past and prevent them from happening again through evaluating of the risks and taking corrective action in time, offering significant operational benefits for relatively low investment. Integrating sensor-driven ML solutions into pump maintenance strategies improves safety, reduces costs, and minimizes downtime [Carvalho et al., 2019].

Recent research has highlighted the successful use of machine learning in predictive maintenance across various sectors. For example, [Manikandan et al., 2023] implemented Convolutional Neural Network (CNNs) for diagnosing faults in industrial pumps, while [Gonçalves et al., 2021] focused on real-time fault detection and classification techniques. Similarly, [Hasan et al., 2021] leveraged deep learning for identifying pump malfunctions, and [Sempewo et al., 2019] investigated predictive maintenance in Uganda's water distribution network. However, most of these studies rely on synthetic or publicly accessible datasets with simulated failure conditions. Commonly used datasets include the Bearing Vibration Dataset under Variable Rotational Speeds [Huang and Baddour, 2018], the NLN-EMP dataset containing motor and pump faults [of Twente; Fieldlab Techport, 2024], and Kaggle's pump sensor dataset for predictive maintenance [of Twente; Fieldlab Techport, 2018]. Despite their usefulness, these datasets often fail to reflect the operational complexity and unpredictability found in real-world utility environments, limiting their relevance for Uganda's water sector. The central aim of this thesis is to bridge this gap by developing a locally tailored PdM framework that harnesses machine learning to drive proactive maintenance and reduce downtime at National Water and Sewerage Corporation (NWSC).

1.2 Problem Statement

A silent crisis is unfolding within Uganda's water infrastructure, endangering public health and the livelihoods of millions. Every day, frequent and unpredictable centrifugal pump failures contribute to over 185,000 customer grievances received annually by the National Water and Sewerage Corporation (NWSC) [National Water and Sewerage Corporation, 2022]. These service interruptions impact a staggering 19 million people across Uganda, sabotaging their livelihoods and creating significant public health risks [National Water and Sewerage Corporation, 2024].

The existing reliance on outdated, reactive maintenance models, where pumps are repaired only after failure has occurred, results in significant operational inefficiencies. Current maintenance practices at National Water and Sewerage Corporation (NWSC) facilities show that approximately 70% of pump failures occur without warning, leading to an average downtime of 8-12 hours per incident and costing the corporation an estimated UGX 2.3 billion annually in lost revenue and emergency repairs [National Water and Sewerage Corporation, 2023]. This reactive approach has been identified by National Water and Sewerage Corporation (NWSC) as a key challenge to its financial performance and service delivery reliability [Japan International Cooperation Agency, 2020].

The traditional time-based preventive maintenance approach, combined with aging infrastructure and limited real-time monitoring capabilities, makes it difficult to predict pump failures accurately. This prevents National Water and Sewerage Corporation (NWSC) from performing timely, data-driven maintenance interventions and addressing potential issues before they escalate into complete system failures. Furthermore, the lack of predictive capabilities means that maintenance crews often perform unnecessary interventions on healthy equipment while missing critical warning signs in deteriorating pumps.

The absence of a robust predictive maintenance framework not only affects operational efficiency but also undermines Uganda's progress toward achieving Sustainable Development Goal 6 (Clean Water and Sanitation) and Vision 2040's infrastructure modernization objectives. Without addressing these systemic maintenance challenges, National Water and Sewerage Corporation (NWSC) will continue to face escalating operational costs, declining service reliability, and growing public dissatisfaction.

To address these challenges, this study proposes developing a locally tailored PdM framework utilizing context-aware machine learning algorithms to analyze real sensor data from National Water and Sewerage Corporation (NWSC)-operated plants. By predicting pump failures before they occur, this solution aims to enable proactive maintenance, reduce unplanned downtime by at least 20%, and enhance water supply reliability in a resource-constrained environment.

1.3 Main Objective

To develop and validate a resource-efficient machine learning framework for Predictive Maintenance (PdM) of Centrifugal pumps at National Water and Sewerage Corporation (NWSC) plants.

1.4 Specific Objectives

1. To curate a labeled dataset of real-time sensor data from Katosi and Ggaba plants, containing at least 50 failure events.

2. To develop five ensemble machine learning models that predict centrifugal pump failures at least 12 hours in advance with over 80% accuracy.

3. To assess and compare the model performance based on accuracy, Precision, Recall, F1-Score, and computational overhead for the National Water and Sewerage Corporation (NWSC) operational environment.

4. To rank key sensor parameters for pump failure prediction using feature importance analysis.

1.5 Research Questions

1. **Primary Research Questions:**

   Which sensor parameters are most predictive of centrifugal pump failures within National Water and Sewerage Corporation (NWSC) operational environment?

2. **Secondary Research Questions:**

   What are the characteristic patterns and failure frequencies within real-time sensor data at Katosi and Ggaba plants?

   To what extent can ensemble machine learning models accurately predict Centrifugal pump failures without a 12-hour lead time?

   How do different ensemble architectures compare in terms of predictive trade-offs between classification metrics and computational cost?

   Which specific sensor parameters serve as the primary predictors of Centrifugal pump failure within the National Water and Sewerage Corporation (NWSC) environment?

1.6 Research Hypotheses

H_1 (Primary Hypothesis): Ensemble machine learning models will achieve prediction accuracy exceeding 80% for centrifugal pump failure detection 48 hours in advance when trained on real operational data from NWSC facilities.

H_2 (Feature Importance Hypothesis): Vibration and flow will emerge as the strongest predictors of impending pump failures, contributing more than 50% of the total feature importance in the optimal model.

H_3 (Performance Improvement Hypothesis): The implementation of the predictive maintenance framework will reduce unplanned pump downtime by at least 20% compared to current reactive maintenance practices, while maintaining computational resource requirements below 4GB RAM and processing times under 60 seconds.

H_4 (Model Performance Hypothesis): XGBoost will demonstrate superior performance among the tested ensemble models due to its gradient boosting capabilities and ability to handle complex feature interactions in pump sensor data.

1.7 Significance and Expected Contributions

1.7.1 Academic Contributions

This research contributes to the growing body of knowledge in predictive maintenance applications within developing country contexts. The study provides novel insights into the performance of ensemble machine learning approaches in resource-constrained environments, specifically addressing the unique operational challenges faced by water utilities in sub-Saharan Africa. The comprehensive comparison of ensemble models using real operational data, rather than synthetic datasets, offers valuable empirical evidence for researchers and practitioners in similar settings.

The development of feature importance frameworks specific to centrifugal pump monitoring in tropical operational conditions adds to the theoretical understanding of failure prediction mechanisms. Additionally, the study establishes methodological benchmarks for sensor data collection, pre-processing, and model evaluation in developing country infrastructure contexts.

1.7.2 Practical Contributions

The developed predictive maintenance framework serves as a practical roadmap for National Water and Sewerage Corporation (NWSC) to transition from reactive to proactive maintenance strategies. The system's ability to predict failures 12 hours in advance enables optimal resource allocation, reducing emergency response costs and minimizing service disruptions. The framework's design consideration for computational constraints ensures practical deployment feasibility within existing National Water and Sewerage Corporation (NWSC) infrastructure.

The data acquisition protocols and model selection criteria provide a replicable blueprint for similar water utilities throughout East Africa facing comparable infrastructure and resource challenges. Its emphasis on locally-relevant sensor parameters and failure patterns ensures applicability beyond the immediate research context, addressing unique operational challenges faced by water utilities in sub-Saharan Africa.

1.7.3 Societal Impact

More reliable water infrastructure directly supports public health initiatives by ensuring consistent access to safe water for 21 million Ugandans served by National Water and Sewerage Corporation (NWSC). Reduced service interruptions enhance economic stability for water-dependent businesses and improve citizen satisfaction with essential services. The framework's potential to reduce water supply disruptions contributes to achieving Uganda's national development goals and international commitments under SDG 6.

The research supports Uganda's Vision 2040 [Finance, 2007] objectives for infrastructure modernization and technology adoption, demonstrating how locally-developed solutions can address critical infrastructure challenges. Furthermore, the study's emphasis on cost-effective technology deployment provides a model for sustainable infrastructure development in resource-constrained environments.

1.8 Scope

1.8.1 Scope of the Study

This study focuses on centrifugal pumps at selected National Water and Sewerage Corporation (NWSC) plants, specifically Katosi Water Treatment Plant and Ggaba Water Treatment plant, which are critical installations serving the Greater Kampala Metropolitan Area. The investigation will depend on real data collected through the deployed sensors measuring vibration, pressure, flowrate and level. The machine learning models will draw attention to supervised learning techniques. This study will specifically leverage ensemble learning models, including; Random Forest (RF), XGBoost (XGB), CatBoost, Voting Ensemble, and LightGBM, to develop and compare models for fault classification and prediction (the goal is to recommend the most effective algorithm for deployment).

1.8.2 Content Scope

The study covered:

1. Collection and analysis of real-time sensor data (vibration, pressure, flow rate, fluid level, and motor status).
2. Identification of failure patterns and early warning indicators.
3. Development and comparison of machine learning models for fault classification and prediction.
4. Assessment of model performance to recommend the most effective algorithm for deployment.

1.8.3 Geographical Scope

The study was geographically limited to:

1. Katosi Water Treatment Plant: located in Mukono District, which supplies a large portion of eastern and central Kampala.
2. Ggaba Water Treatment Plant: located in Makindye Division, Kampala, supplying central and southwestern parts of the metropolitan area.

1.8.4 Time Scope

The study covered operational data and failure patterns between 2024 and 2025, a period during which Kampala experienced notable water supply challenges, emphasizing the need for efficient and predictive maintenance systems.

1.8.5 Exclusions

The study did not cover other National Water and Sewerage Corporation (NWSC) installations outside the Kampala Metropolitan Area, nor evaluate pumps not instrumented with the selected sensor types.

1.9 Thesis outline

The remainder of this thesis is organized as follows: Chapter 2 presents a review of the literature, identifies the gaps and builds a predictive framework, followed by Chapter 3, which deals with the methodologies used in the study, analyzes procedures. Chapter 4 shows the results from the collected data, Chapter 5 provides a discussion on the study findings and interpretation, and finally, in Chapter 6, it contains conclusions that outline the major future research opportunities of this thesis.

Chapter 2

Literature review

2.1 Introduction

This Literature review provided a comprehensive analysis of the existing scholarly research and technical knowledge required to understand predictive maintenance challenges in water infrastructure systems and potential machine learning solutions. Further, this chapter aimed to critically evaluate prior studies, identify gaps in the literature, and establish a framework for the current study. The review was organized thematically and based on the development of key concepts and methodologies.

2.2 Overview of centrifugal pumps in water infrastructure

As they were widely used in industrial, commercial, agricultural and municipal systems, Centrifugal water pumps were known for their robustness, efficiency and ability to operate continuously. Further, according to [Edward et al., 1998], centrifugal pumps feature radial, axial and mixed configurations, which were common types of pumps used for water transfer due to the presence of an impeller and centrifugal action. Edward additionally described that, "Fluids entered the impeller in the centre portion, known as the eye and discharged around the entire circumference in a casing." In the course of rotating, the liquid received energy from the vanes, leading to an increase in pressure and absolute velocity. " However, their reliability was vital because unexpected pump failure could disrupt critical processes, incur substantial repair costs and reduce productivity. Furthermore, several studies had explored how continuous exposure to varying hydraulic conditions, fluid pressures and environmental influences led to degradation of mechanical components, including bearings, seals and impellers. Therefore, in order to ensure sustained reliability of centrifugal pumps, maintenance strategies were required to effectively detect, predict and mitigate progression of failure within the pumps.

Both [Keane et al., 2024, Chaibi et al., 2024] demonstrated that traditional maintenance strategies heavily rely on failure-occurrence model. Further, their studies highlighted that under corrective maintenance, interventions were only triggered post-failure, hence this limited the reliability of critical infrastructure. They also added that Corrective maintenance resulted into unplanned downtime and high repair costs, while preventive maintenance might fail to detect degradation during its early stages hence leading to pre-mature replacement of still functional components. Hence, these limitations emphasized a critical Gap in the conventional approaches, which implied that neither of the strategies could effectively address real-time equipment health, thereby limiting efficiency and cost-effectiveness.

Due to these limitations, Predictive Maintenance (PdM) was sought after with the aim to overcome these limitations through leveraging sensor data, statistical indicators and Machine Learning (ML) to monitor equipment condition and forecast any impending failures. Furthermore, building on the research problem outlined in Chapter one, the literature review incorporated relevant literature on predictive maintenance and machine learning due to its ability to identify key failure indicators that detect failures earlier compared to conventional approaches.

Furthermore, several studies examined the inclusion of machine learning techniques like Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and various deep learning models, which showed a high potential in detecting patterns in degradation within complex sensor datasets according to [Afuan and Isnanto, 2025, ?]. Additionally, [Dong et al., 2017, Kizito et al., 2021, Wu et al., 2018] demonstrated the capabilities of deep learning architectures, especially Long Short-Term Memory [LSTM] networks, in temporal sequence modelling, making them highly suitable for Remaining Useful Life [RUL] estimation.However, despite the fact that ML- based prognostics had been successfully applied to rotating machinery like turbines, motors, bearings, the literature addressing centrifugal pump remaining RUL prediction remained relatively limited despite the substantial industrial importance of such pumps.

While studies like [Luo et al., 2025, Chen et al., 2021, He et al., 2022, Panda et al., 2018] focused on testing of various algorithms, fewer studies investigated the incorporation of pump degradation mechanisms, vibration-based statistical indicators, and comparative ML evaluation within a unified structure [Luo et al., 2025, Panda et al., 2018, Ranawat et al., 2021]. Remarkably, existing research often overlooked real-world deployment pathways, for instance; graphical interfaces for operational personnel are rarely included in the existing work. Hence addressing these gaps required the use of a holistic approach that integrated data pre-processing, feature engineering, model training, and deployment considerations.

Thus, based on the objective of this study, to develop, evaluate and implement a complete predictive maintenance framework for centrifugal water pumps using sensor data collected over four months of operation at National Water and Sewerage Corporation (NWSC). This approach included systematic data preprocessing, extraction of vibration base time domain statistical condition, development of Random Forest, XGBoost, CATBoost, Light GBM and Voting Ensemble models for RUL regression, comparative evaluation of performance, real-time deployment, which provided both deployment and analytical perspectives. Additionally, this study presented a replicable pathway for implementing machine learning based predictive maintenance solutions in industrial pump systems.

The remainder of this paper was structured as follows: discussed background on National Water and Sewerage Corporation (NWSC), the degradation mechanisms of centrifugal pumps, and existing ML-based prognostic research motivating this framework. Chapter three further detailed the proposed predictive maintenance pipeline, covering data acquisition, pre-processing, feature extraction, and model development. Furthermore, presented experimental results and discusses condition indicators, model performance, and RUL prediction accuracy. And lastly, concluded the study and suggested future directions.

2.3 Background on National Water and Sewerage Corporation (NWSC)

2.3.1 Global and Regional Water Infrastructure Context

Water utilities worldwide play a critical role in urban development and human wellbeing, yet many face persistent operational challenges due to aging infrastructure, inadequate maintenance strategies, and resource constraints [Berg and Danilenko, 2017]. [UNICEF & WHO, 2019] reports that,"over 2.2 billion people globally lack access to safely managed drinking water services," with Sub-Saharan Africa accounting for approximately 400 million of this population. These challenges are particularly acute in developing countries where infrastructure investment has historically lagged behind population growth and urbanization rates.

In Uganda, National Water and Sewerage Corporation (NWSC) faces the dual challenge of expanding service coverage while maintaining existing infrastructure reliability. The corporation serves over 21 million people across 282 towns, representing approximately 46.7% of Uganda's urban and semi-urban population [National Water and Sewerage Corporation, 2024]. Despite significant achievements in service expansion, frequent equipment breakdowns, particularly in pumping systems, continue to undermine service reliability and customer satisfaction.

2.3.2 National Water and Sewerage Corporation (NWSC)'s Infrastructure and Operational Context

The National Water and Sewerage Corporation (NWSC) was established in 1972 and was later reconstituted under Act of 1995.Since then, the corporation operated under a performance contract with the Government of Uganda and earned recognition as one of Africa's best-performing water utilities [Marscher, 2004]. It managed a complex infrastructure network including 71 water treatment plants and two sewerage treatment plants (Bugolobi and Lubigi), with the Ggaba and Katosi facilities having served as the primary water treatment plants for the Greater Kampala Metropolitan Area. The Ggaba Water Treatment Plant, comprising three treatment units (Ggaba I, II, and III), represented Uganda's oldest large-scale water treatment facility, with some infrastructure having dated back to the 1920s [Water and Corporation, 2018]. While the facility had undergone several upgrades, much of the pumping equipment operated beyond its intended design life, which contributed to significant reliability challenges. In contrast, the Katosi Water Treatment Plant was commissioned in 2021 with a capacity of 160,000 cubic meters per day. [Thus, this represented modern treatment technology though its still relied on conventional time-based maintenance approaches rather than condition-based strategies.

2.3.3 Water Treatment and Distribution Processes

Persistent issues, particularly within pumping systems, continued to undermine service reliability and customer satisfaction. Although the Katosi plant represented modern treatment technology, the system still relied on Conventional time-based maintenance approaches rather than the condition-based strategies proposed by recent industry standards [Water and Corporation, 2024].Both Ggaba and Katosi facilities employed conventional multi-stage water treatment processes, including coagulation, sedimentation, filtration, and disinfection, to ensure compliance with World Health Organization (WHO) drinking water quality standards [Alderman et al., 2013].

In these operations, raw water from Lake Victoria underwent chemical coagulation using aluminium sulfate, followed by sedimentation in clarifiers to remove suspended particles. The clarified water then passed through rapid sand filtration before final disinfection with chlorine compounds [MWE, 2020].

Centrifugal pumps played a critical role throughout these processes, from raw water intake to final distribution. High-capacity centrifugal pumps moved raw water from intake points to treatment facilities, while smaller distribution pumps-maintained pressure throughout the treatment process. Following treatment, large-capacity distribution pumps delivered treated water to elevated storage tanks and directly to the distribution network, having converted rotational energy into kinetic energy to overcome friction losses and elevation changes [White, 2016].

2.3.4 Degradation mechanisms of centrifugal water pumps

Centrifugal pumps had served as critical components in fluid transport, yet prolonged operation had exposed these systems to various degradation phenomena that altered their mechanical and hydraulic responses [Eaton et al., 2022, Ellorde et al., 2021]. Critical mechanical failure modes, such as bearing faults, have frequently occurred due to overloading or poor lubrication, thus leading to pitting and peeling which intensify vibration levels[Sakthivel et al., 2010, Rapur and Tiwari, 2019]. Likewise, misalignments at the pump coupling and structural imbalances, often caused by errors in assembly or defects in the impeller, manifested an increased vibration amplitudes and eccentric forces that compromised system stability [Hu et al., 2021, Wang et al., 2019]. Hydraulic faults, caused by exceeding the mechanical capacity, further degraded performance; for example, blockages in the volute and impeller fouling reduced efficiency in the pumps and also increased the intensity of blade-passing frequencies. Further, extreme phenomena like water hammer exposed the pump body to pressure surges causing stresses hundreds of times higher than normal, which resulted in catastrophic component bursts [Kan et al., 2021, Zhang et al., 2021a]. Furthermore, cavitation remained a common problem, causing serious surface damage, noise, and significant performance drops, characterized by a continuous, wide-band vibration response. These degradation patterns, treated collectively, necessitated a shift towards predictive maintenance to capture early signs of failure before total system breakdown.

2.3.5 Critical Parameters for Pump Monitoring

Effective pump condition monitoring requires systematic measurement of key parameters that provide early indication of developing problems. Based on extensive industrial experience, vibration analysis has emerged as the most comprehensive diagnostic tool, capable of detecting most mechanical problems before they cause catastrophic failure [ISO, 2019]. Vibration Analysis provides information about unbalance, misalignment, bearing condition, gear problems, and electrical issues. According to ISO 18436-2 standards, vibration measurements should include overall levels, frequency spectra, and trending analysis to enable accurate fault diagnosis [ISO, 2019]. Acceleration, velocity, and displacement measurements each provide different insights into machine condition, with velocity measurements typically preferred for overall condition assessment. Pressure and Flow Monitoring can detect hydraulic problems including cavitation, blockages, and wear-related performance degradation. These parameters are particularly important for detecting problems that may not be apparent through vibration analysis alone [GÜlich, 2019].

Cause of failures	**Parameters**					
	Temperature	**Pressure**	**Leak tightness**	**Oil analysis**	**Monitoring of the electric current**	**Vibration**
Unbalance		✓				✓
Misalignment	✓				✓	✓
Bearing failures	✓			✓		✓
Journal bed failures	✓	✓	✓	✓		✓
Gear failures				✓		✓
Mechanical looseness						✓
Electrical motor failures	✓				✓	✓
Hydraulic and aerodynamic failures		✓				✓

Table 2.1: Correlation of Pump Faults and Monitoring Parameters.[Kiliç et al., 2017].

2.4 Evolution of Maintenance Strategies

2.4.1 From Reactive to Predictive Approaches

Maintenance strategies have evolved significantly over the past century, driven by technological advances and improved understanding of failure mechanisms. This evolution can be characterized by four distinct phases: reactive maintenance, preventive maintenance, predictive maintenance, and prescriptive maintenance [Dueñas Ramírez and Villegas López, 2020].

Reactive Maintenance, also known as run-to-failure, involves repairing equipment only after failure occurs. While this approach minimizes maintenance costs during normal operation, it results in high emergency repair costs, extended downtime, and potential safety hazards [Coleman et al., 2017]. For critical infrastructure like water distribution systems, reactive maintenance is generally unacceptable due to the high cost of service interruptions.

Preventive Maintenance involves scheduled maintenance activities based on calendar time, operating hours, or production cycles. This approach reduces unexpected failures but can result in unnecessary maintenance activities and associated costs. Studies indicate that 70-80% of preventive maintenance activities may be unnecessary, representing significant waste of resources [Mobley, 2002].

Predictive Maintenance uses condition monitoring data to determine maintenance needs based on actual equipment condition rather than elapsed time. This approach optimizes maintenance timing, reduces unnecessary interventions, and enables maintenance planning that minimizes operational disruptions [Sakib and Wuest, 2018].

The Figure 2.1 shows a Summary of the evolution of maintenance

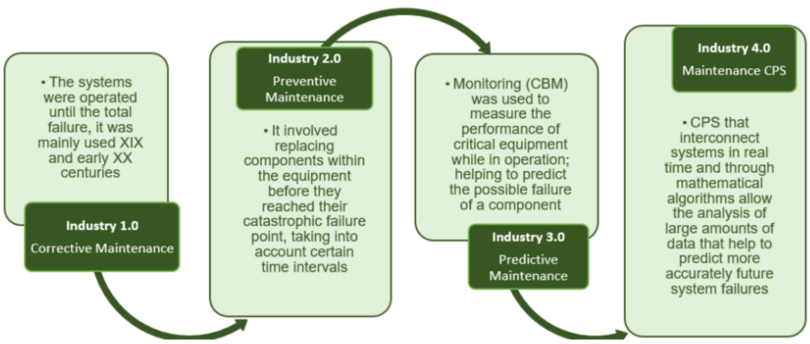

Figure 2.1: Evolution of maintenance in the industry
[Dueñas Ramírez and Villegas López, 2020]

The Figure 2.2 shows a Summary of the difference between the maintenance types

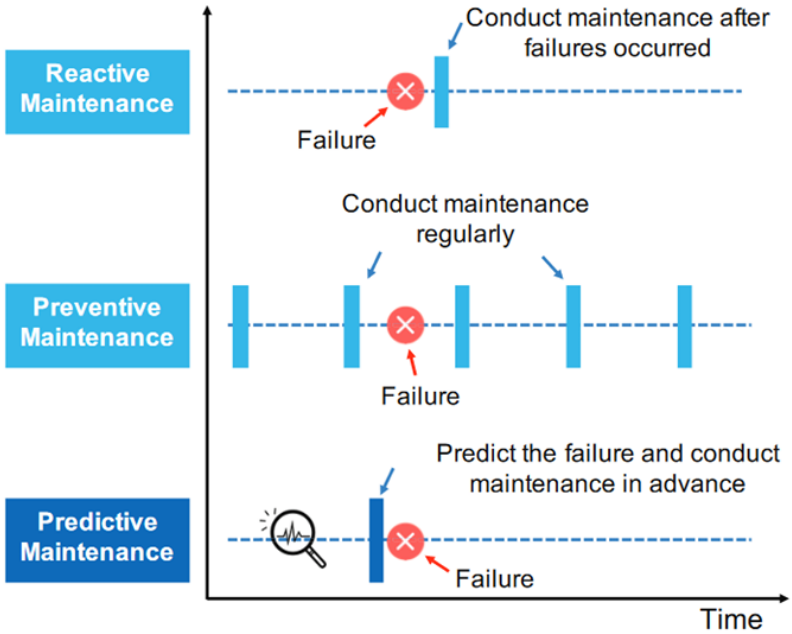

Figure 2.2: Difference between Corrective, Preventive and Predictive maintenance [Zhu et al., 2024]

Table 2.2 shows the differences between these two types of maintenance.

	Benefits	**Challenges**
Corrective Maintenance	• Maximum utilization of tooling or machine components	• Potentially greater damage to the machine • Unplanned downtime • Higher maintenance costs
Preventive Maintenance	• Less likelihood of broken machinery • Less unplanned downtime • More cost-effective than reactive	• Increased replacement costs over time • Need for additional spare parts inventory • Increased planned downtime

Table 2.2: Comparison of Reactive vs Planned Maintenance[Coleman et al., 2017]

2.4.2 Industry 4.0 and Smart Maintenance

The emergence of Industry 4.0 technologies has revolutionized maintenance practices through integration of Internet of Things (IoT) devices, artificial intelligence, and advanced analytics. These technologies enable continuous condition monitoring, automated fault detection, and optimized maintenance scheduling [Cachada et al., 2019].

IoT-enabled predictive maintenance systems can collect vast amounts of sensor data, process this information using machine learning algorithms, and provide actionable insights for maintenance decision-making. Cloud-based analytics platforms enable real-time monitoring of geographically distributed assets, making advanced maintenance strategies accessible to smaller organizations [Riahi Sfar et al., 2018].

2.4.3 Benefits and Challenges of Predictive Maintenance

Comprehensive studies have documented significant benefits of predictive maintenance implementation. [Collins et al., 2021] The U.S. Department of Energy (2010) reports typical benefits including:

1. 10-40% reduction in maintenance costs
2. 35-45% reduction in downtime
3. 20-25% increase in production capacity
4. 70-75% reduction in equipment failures
5. 30-60% reduction in maintenance planning time

However, successful predictive maintenance implementation faces several challenges, particularly in developing country contexts. These include high initial technology costs, lack of technical expertise, inadequate data management infrastructure, and organizational resistance to change [Sempewo et al., 2019].

2.5 Machine Learning Applications in Predictive Maintenance

2.5.1 Overview of Machine Learning in Industrial Applications

Machine learning has emerged as a powerful tool for predictive maintenance, enabling automated pattern recognition, fault classification, and failure prediction from sensor data. The ability to process large datasets, identify complex relationships, and adapt to changing conditions makes machine learning particularly suitable for industrial condition monitoring applications [Géron and Aurélien, 2019].

Supervised learning algorithms dominate predictive maintenance applications, where historical data with known failure outcomes is used to train models for future prediction. Classification algorithms predict discrete outcomes (normal, warning, alarm conditions), while regression algorithms predict continuous variables such as remaining useful life or performance parameters [Hastie et al., 2017].

2.5.2 Ensemble Learning Methods

Ensemble methods, which combine predictions from multiple base models, have shown particular promise for predictive maintenance applications due to their improved accuracy and robustness compared to individual models [Zhou, 2012]. The primary ensemble approaches include:

Random Forest combines multiple decision trees with random feature selection and bootstrap sampling, providing excellent performance on structured data while maintaining interpretability through feature importance ranking [Breiman, 2001a]. The algorithm's inherent randomness helps prevent overfitting and provides robust performance across diverse datasets.

Gradient Boosting Methods including XGBoost, LightGBM, and CatBoost build models sequentially, with each new model correcting errors from previous models. XGBoost [Chen and Guestrin, 2016] has demonstrated exceptional performance in numerous machine learning competitions and industrial applications due to its optimization techniques and handling of missing values.

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k), \tag{2.1}$$

where

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2, \tag{2.2}$$

and

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i), \quad f_k \in \mathcal{F}. \tag{2.3}$$

Here,

- $l(y_i, \hat{y}_i)$ is a convex loss function that measures the difference between prediction $\hat{y}_i$ and true label y_i,
- $\Omega(f_k)$ is the regularization term that penalizes model complexity,
- T is the number of leaves in tree f,
- w are the leaf weights,
- γ and λ are regularization parameters controlling tree complexity and weight shrinkage,
- $\hat{y}_i^{(t)}$ is the prediction for sample i at iteration t,
- $f_k \in \mathcal{F}$, where $\mathcal{F}$ is the space of regression trees.

LightGBM [Ke et al., 2017] offers improved computational efficiency compared to traditional gradient boosting through leaf-wise tree growth and categorical feature optimization. This makes it particularly suitable for large datasets and real-time applications

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \left[l\big(y_i, \hat{y}_i^{(t-1)}\big) + g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t), \tag{2.4}$$

where $g_i = \frac{\partial l(y_i, \hat{y}_i)}{\partial \hat{y}_i}\Big|_{\hat{y}_i = \hat{y}_i^{(t-1)}}$ and $h_i = \frac{\partial^2 l(y_i, \hat{y}_i)}{\partial \hat{y}_i^2}\Big|_{\hat{y}_i = \hat{y}_i^{(t-1)}}$ are the first and second derivatives of the loss function with respect to the previous iteration's prediction [Ke et al., 2017].

CatBoost [Prokhorenkova et al., 2018] provides native handling of categorical features and robust performance with minimal hyperparameter tuning, making it attractive for industrial applications where model development time is limited.

Voting Ensembles combine predictions from diverse base models using either majority voting (classification) or average predictions (regression). The diversity of base models is crucial for achieving performance improvements over individual models [Kuncheva, 2014].

Model	**Strengths**	**Limitations**
Logistic Regression (LR)	Simple, easily scalable to large datasets, and can handle multiclass problems.	Does not perform well with imbalanced data.
Random Forest (RF)	Handles imbalanced datasets and multiclass problems well. Effective on large datasets.	Can be computationally expensive and may overfit on noisy data.
XGBoost (XGB)	Highly efficient for multiclass problems and imbalanced datasets. Optimized for large datasets.	Requires careful tuning and may be less interpretable.
CatBoost	Handles categorical features effectively, robust to imbalanced datasets, and works well on large datasets.	Longer training time compared to simpler models.
Voting Classifier	A hybrid model combining predictions from multiple models, boosting overall performance and stability.	Performance depends on the quality and diversity of base models.
LightGBM	Faster and more memory-efficient compared to others. Supports multiclass and imbalanced datasets.	Sensitive to overfitting and requires careful parameter tuning.

Table 2.3: Comparison of Classification Models: Strengths and Limitations

2.6 Feature Engineering and Selection

Effective feature engineering transforms raw sensor data into informative features that capture relevant failure patterns. Time-domain features include statistical measures (mean, standard deviation, skewness, kurtosis), trend indicators, and change detection metrics [Lei et al., 2020].

Frequency-domain features, particularly important for vibration analysis, include spectral peaks, frequency band energies, and harmonic analysis. Advanced features such as wavelet coefficients, empirical mode decomposition components, and entropy measures can capture complex signal characteristics that traditional methods miss [Jardine et al., 2006]. Feature selection techniques like selectKBest using ANOVA F-score, help identify the most relevant features while reducing computational requirements and avoiding overfitting. Methods include statistical tests, recursive feature elimination, and embedded selection within model training processes [Guyon and Elisseeff, 2003].

2.7 Model Evaluation for Predictive Maintenance

Effective evaluation of predictive maintenance models required careful consideration of class imbalance, cost-sensitive metrics, and temporal aspects of failure prediction. Traditional accuracy metrics could be misleading in imbalanced datasets where normal conditions significantly outnumbered failure cases [He and Garcia, 2009].

Classification Metrics appropriate for predictive maintenance include precision (true positive rate among positive predictions), recall (true positive rate among actual failures), F1-score (harmonic mean of precision and recall), and area under the ROC curve (AUC-ROC) which measure discrimination ability across all decision thresholds [Fawcett, 2006].

Temporal Considerations were crucial for predictive maintenance evaluation. Models had to provide sufficient advance warning (prediction horizon) to enable effective maintenance planning while minimizing false alarms that eroded user confidence. The prediction horizon represented a critical trade-off between early warning capability and prediction accuracy [Saxena et al., 2008]. Further, model performance was evaluated using coefficient of Determination, Root Mean Square Error and Mean Absolute Error with the purpose of quantifying the accuracy, consistency in movement, and helped to fit the performance of the models by evaluating all errors through being weighted equally

2.8 Related Work in Pump Failure Prediction

2.8.1 Industrial Applications and Synthetic Data Studies

Numerous studies have demonstrated the potential of machine learning for pump failure prediction across various industrial contexts. [Manikandan et al., 2023] employed deep Convolutional Neural Network (CNNs) for fault detection in industrial pumps, achieving

94.2% accuracy in classifying bearing faults, impeller damage, and seal failures using vibration spectrograms. Their approach demonstrated the effectiveness of automated feature extraction from raw vibration signals.

[Gonçalves et al., 2021] developed a real-time fault detection system using vibration data and Markov parameters, achieving fault classification accuracy of 96.8% for common pump failures. The system's ability to operate in real-time processing mode demonstrated practical applicability for industrial monitoring systems.

[Hasan et al., 2021] proposed a deep learning framework that transforms vibration signals into scalograms for visual fault analysis, achieving 97.3% accuracy in pump fault detection. The approach showed particular strength in detecting early-stage faults that traditional methods often miss.

In the oil and gas industry, [Orrù et al., 2020] compared Support Vector Machine (SVM) and Multilayer Perceptron (MLP) for fault classification in centrifugal pumps, with Support Vector Machine (SVM) achieving 91.4% accuracy compared to 87.2% for MLP. Their study highlighted the importance of feature selection and hyperparameter optimization for optimal performance.

2.8.2 Applications in Developing Countries

Limited research has addressed predictive maintenance implementation in developing country contexts, despite unique challenges including resource constraints, environmental conditions, and technical capacity limitations.

[Sempewo et al., 2019] investigated predictive maintenance applications in Kampala's water distribution network, identifying data scarcity and institutional capacity as primary implementation barriers. Their study recommended adaptive approaches capable of handling limited and noisy datasets, particularly relevant for resource-constrained utilities.

[Wilson et al., 2017] analyzed sensor data from 42 rural handpumps in Kenya over 14 months, demonstrating system uptime improvements from 70% to over 99% through predictive maintenance. The study achieved a 7% reduction in levelized cost of water while significantly improving service reliability. However, the study focused on simple mechanical pumps rather than complex centrifugal systems.

In Tanzania, [Pathak, 2023] implemented predictive maintenance on water pumping systems, achieving 35% reduction in downtime and 28% decrease in maintenance costs. The study demonstrated significant potential for infrastructure reliability improvement in developing regions but relied primarily on pressure and flow monitoring rather than comprehensive condition monitoring.

2.8.3 Machine Learning Model Comparisons

Comparative studies have provided insights into optimal model selection for pump failure prediction. [Carvalho et al., 2019] conducted a systematic review of predictive maintenance practices, concluding that ensemble methods consistently outperform individual algorithms across diverse industrial applications. [Azadeh et al., 2013] developed a hybrid model combining SVM and artificial neural networks with hyperparameter optimization for pump fault detection, achieving 93.7% accuracy under noisy industrial conditions. The study demonstrated the importance of robust model design for practical applications.

Recent comparative studies by [Zhang et al., 2021b] evaluated Random Forest, XGBoost, and LightGBM for vibration-based fault detection, with XGBoost achieving the best performance (95.1% accuracy) while LightGBM offered the best computational efficiency for real-time applications.

2.8.4 Limitations of Existing Research

Most existing studies rely on synthetic datasets or controlled laboratory conditions that may not reflect real-world operational complexity. Commonly used datasets include Bearing Vibration Dataset under Variable Rotational Speeds [Huang and Baddour, 2018], the NLN-EMP dataset containing motor and pump faults [of Twente; Fieldlab Techport, 2024], and Kaggle's pump sensor dataset for predictive maintenance [of Twente; Fieldlab Techport, 2018]. The reliance on synthetic data presents several limitations:

- Idealized operating conditions that don't reflect industrial noise and interference
- Limited failure modes compared to diverse real-world failure mechanisms
- Artificial data balance that doesn't represent actual failure frequencies
- Controlled environmental conditions unlike variable industrial environments

Furthermore, most studies focus on fault classification rather than failure prediction with specific time horizons. The prediction of failures 24-72 hours in advance, crucial for maintenance planning, has received limited attention in water utility contexts.

2.9 Research Gaps and Theoretical Framework

2.9.1 Identified Research Gaps

The preceding analysis reveals key gaps in the current literature: an over-reliance on idealized datasets that do not account for the operational realities of the National Water and Sewerage Corporation (NWSC), and a preference for computationally intensive models unsuitable for resource-constrained environments. The following are the several critical gaps revealed in the literature review and what I will address in my research:

1. Real-World Data Scarcity: Most studies rely on synthetic or laboratory datasets that don't capture the complexity of actual utility operations. Research using real operational data from water utilities, particularly in developing countries, remains limited.

2. Prediction Horizon Specification: While many studies achieve high classification accuracy for existing faults, few address the specific challenge of predicting failures 12-72 hours in advance, which is crucial for maintenance planning in water utilities.

3. Computational Resource Constraints: Limited research addresses implementation in resource-constrained environments with specific computational limitations (memory, processing time, power consumption).

4. Local Context Adaptation: Few studies address the unique challenges of developing country implementations, including intermittent power supply, limited technical capacity, and environmental conditions.

5. Ensemble Method Optimization: While individual studies have explored various ensemble methods, comprehensive comparisons optimized for water utility applications remain limited.

6. Feature Importance for Water Utilities: Limited research specifically identifies the most critical sensor parameters for pump failure prediction in water distribution systems.

2.9.2 Theoretical Framework

This research builds upon established theoretical foundations in predictive maintenance while addressing identified gaps through a comprehensive framework integrating:

Condition-Based Maintenance Theory provides the foundation for transitioning from time-based to condition-based maintenance decisions. The framework incorporates fault progression models that relate sensor measurements to equipment health states [Jardine et al., 2006].

Machine Learning Theory guides the selection and optimization of ensemble methods for failure prediction. The framework emphasizes bias-variance trade-offs, overfitting prevention, and generalization capabilities crucial for reliable industrial applications [Hastie et al., 2017].

Systems Engineering Principles ensure practical implementation feasibility through consideration of technical constraints, organizational factors, and economic viability. The framework incorporates reliability engineering principles for critical infrastructure applications [Rausand and Hoyland, 2008].

2.9.3 Conceptual Model

The proposed predictive maintenance framework integrates four key components:

1. Data Acquisition Layer: Multi-sensor condition monitoring system optimized for water utility environments, incorporating vibration, pressure, level and flow measurements with local buffering and communication capabilities.

2. Data Processing Layer: Real-time data pre-processing, feature extraction, and quality assessment with edge computing capabilities to handle connectivity limitations.

3. Machine Learning Layer: Ensemble prediction models optimized for computational efficiency while maintaining high accuracy, with specific focus on 12-hour prediction horizon.

4. Decision Support Layer: Integration with maintenance planning systems, providing actionable recommendations while considering operational constraints and resource availability.

2.10 Proposed Predictive Maintenance Framework for National Water and Sewerage Corporation (NWSC)

The predictive maintenance framework developed in this study, was developed to shift from reactive repairs to data-driven foresight. It comprised of several interrelated stages: Processing raw sensor data, extracting meaningful indicators, training predictive models, and deploying the models for real-time use. The monitoring of centrifugal pump's vibration and flow over a period of 5 months captured the full equipment life, from ''healthy" to ''final failure." The process began with purifying raw sensor data, eliminating room for errors and normalizing of values to ensure that machine learning models received high-quality input. The data was then translated into five specifics statistical like; Change, Critical, Warning, Operational and Off which showed ''health indicators" that acted as a digital signature for mechanical wear, such as bearing degradation or cavitation. By training these five distinct models (RF, XGBoost, CATBoost, LightGBM and Voting Ensemble), the system learned to recognize these patterns, ultimately providing a highly accurate tool for predicting exactly when maintenance is required to prevent a total system shutdown.

2.11 Real-World Implementation Challenges and Solutions

2.11.1 Data Acquisition and Management

Practical implementation of predictive maintenance in water utilities faces numerous technical and organizational challenges. Data acquisition systems must operate reliably in harsh industrial environments while maintaining data quality and consistency over extended periods [Ahmad et al., 2022].

Modern data acquisition systems typically employ distributed architectures with edge computing capabilities to handle connectivity issues common in developing countries.

[Gupta et al., 2023] describe Arduino-based systems integrated with Global System for Mobile Communications (GSM) modules and SD cards for data transmission and local buffering, enabling operation during network outages.

Sensor Integration challenges include compatibility with existing control systems, power supply requirements, and environmental protection. Wireless sensor networks offer installation flexibility but require careful attention to battery life and communication reliability [Chen et al., 2021].

2.11.2 Computational Resource Constraints

Developing country implementations must consider limited computational resources and intermittent power supply. Model deployment strategies must balance prediction accuracy with computational efficiency, often requiring optimization techniques such as model pruning, quantization, and edge computing architectures [Lei et al., 2020].

Cloud-based solutions offer computational scalability but depend on reliable internet connectivity, which may be limited in rural areas. Hybrid architectures combining local processing for critical functions with cloud-based advanced analytics provide optimal balance between functionality and reliability [Wang et al., 2019].

2.11.3 Economic and Organizational Factors

The high initial investment required for predictive maintenance systems can be prohibitive for resource-constrained utilities. Total cost of ownership includes hardware, software, training, and ongoing maintenance costs that must be justified through demonstrated operational benefits [Carnero, 2006](Carnero, 2019).

Organizational readiness assessment is crucial for successful implementation. This includes technical capacity, management support, change management processes, and alignment with existing maintenance procedures. Gradual implementation approaches, starting with critical assets, can help build organizational confidence and demonstrate value [Bousdekis et al., 2019].

2.12 Summary

This literature review established critical need for predictive maintenance in Uganda's water sector, where centrifugal pumps served as vital infrastructure components often undermined by reactive maintenance strategies and aging equipment. The review demonstrated that while predictive maintenance technologies offer significant benefits, their implementation in developing country contexts faced unique challenges requiring adapted solutions. The analysis of existing research revealed substantial gaps in real-world implementations using actual operational data, particularly in water utility contexts. Most studies relied on synthetic datasets that failed to capture the complexity and constraints of actual utility operations. Furthermore, limited attention has been given to specific prediction horizons crucial for maintenance planning and computational resource constraints typical of developing country implementations. The ensemble machine learning methods reviewed showed particular promise for water utility applications due to their robust performance and ability to handle complex, noisy datasets. However, systematic comparison of these methods using real utility data, with emphasis on computational efficiency and practical deployment considerations, remained limited in the literature.

Chapter 3

Methodology

3.1 Introduction

This chapter explains the research design, data collection methods, data analysis, sensor setup, model development approach, and validation processes used in to implement a predictive maintenance framework for centrifugal water pumps. It also presents the study's methodology, including; reproducibility, data quality, and significance to National Water and Sewerage Corporation (NWSC)'s operational context. Furthermore, the data is analyzed by the machines using three National Water and Sewerage Corporation (NWSC) pumps named; Katosi_pump A, Gungill_pump 3 and Muyenga_pump 4.

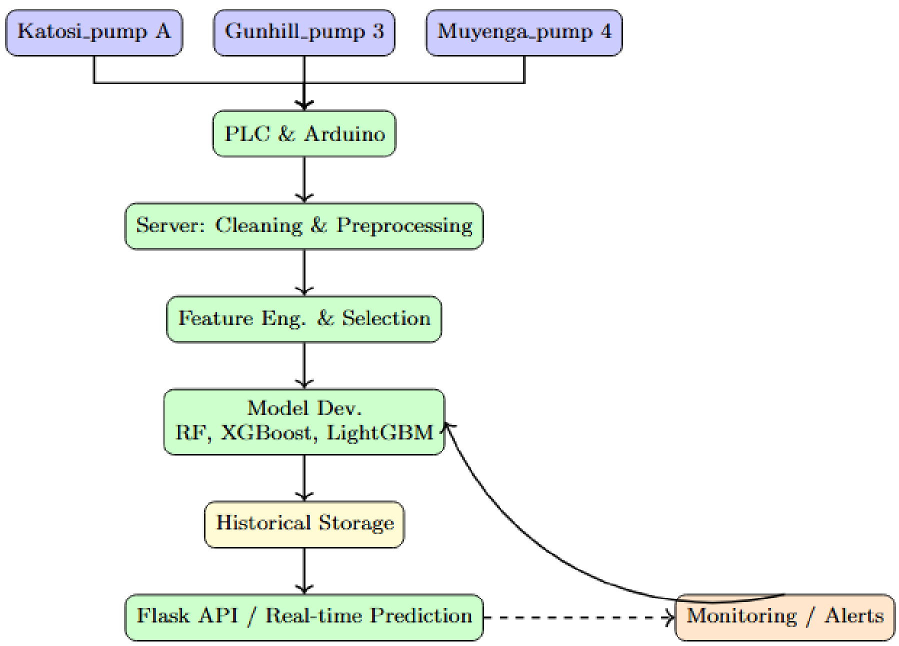

Figure 3.1: Predictive Maintenance Architecture for Katosi, Gunhill, and Muyenga Pumps

3.2 Research Design

The research design involved proposing an applied, data-driven experimental design that integrated real-time sensor data acquisition with Machine Learning (ML)-based predictive modelling. This approach synthesized field-based monitoring with quantitative analysis (computer engineering) and relied on numerical data such as sensor data [nist, 2021] to develop a robust Predictive Maintenance (PdM) model. The approach was suitable for the thesis as it aimed to analyse the performance of various ML algorithms in terms of F1 score, recall, precision, and accuracy. Python was used for data collection due to its rich ecosystem of ML libraries and visualization tools, as well as the researcher's prior knowledge and experience with the language.

3.3 Study Area

The research was conducted at National Water and Sewerage Corporation (NWSC), specifically focusing on Ggaba1 and Katosi pumping stations in Uganda. These sites were selected due to their critical role in urban water supply and the presence of high-capacity centrifugal pumps operating under continuous load, experiencing recurrent pump failures.

3.3.1 Population and Sample

Target Population

The population consisted of all 13 centrifugal pumps located at the selected NWSC stations. These pumps were powered by 1000 kW motors, which were designed for a head exceeding 70 m and a flow rate above 200m3/h and were the primary units of analysis.

Figure 3.2: Illustrated the pump

Sampling Strategy

To ensure a representative pump sample, stratified random sampling was employed. This method was chosen because the two selected plants contain 13 pumps of different types, requiring consideration of both operational and mechanical stratification. This is because, due to their inherent features and modes of operation, different pieces of equipment are not suitable for the same monitoring system. Therefore, before creating the data gathering instrument, the targeted equipment might be chosen. In stratified sampling theory, the total sample size(n) is distributed across the strata based on their respective population sizes.

Sample Size

A sample was defined as a group of individuals or objects drawn from a specific population with the purpose of collecting data that accurately reflected the characteristics of the population [McMillan and Schumacher, 2010]. Based on Engineering best practices recommended a sample size between 10% and 30% of the total population [Sathyanarayana et al., 2024]. In this study, three pumps were selected from a total population of 13, representing approximately 16.7% of the units at the National Water and Sewerage Corporation (NWSC) plants. This sample size was determined based on criteria:

- Resource and Time Constraints: The selection aligned with the feasibility of intensive data monitoring over a five-month period.
- Proportionally and Diversity: One pump of each other major type was selected to maintain structural proportionality and ensure the dataset reflected diverse

operational conditions

- Data Volume: Although the number of physical units is small[n=3], the resulting dataset was exhaustive, comprising 2,902,703 data points across the features, which provided a high degree of granularity for predictive modeling.

Figure4.5 illustrates. This summarized the information obtained from the two plants.

Total population size, $N = 13$

Total sample size, $n = 3$

Strata population sizes:

Muyenga: $N_{\text{Muyenga}} = 4$
Katosi: $N_{\text{Katosi}} = 6$
Gunhill: $N_{\text{Gunhill}} = 3$

Using the formula of stratified sampling:

$$n_h = \frac{N_h}{N} \times n$$

where:

n_h = sample size for stratum h,

N_h = population size for stratum h,

N = total population size,

n = total sample size.

The final sample sizes per stratum were:

Stratum	**Population Size**(N_h)	**Sample Size**(n_h)
Muyenga	4	1
Katosi	6	1
Gunhill	3	1

3.4 Data collection Methods and Instruments

Data collection was defined as the systematic process of gathering information from a specific sample to identify patterns and trends that informed the research objectives. In the domain of predictive maintenance, this phase was foundational; unprocessed data remained fundamentally unstable until it was arranged to support the development of an IoT-based time-series structure. This study utilized a quantitative research design, where numerical measurements are captured to test the hypothesis that machine learning models could accurately predict pump failure without sacrificing performance.

To ensure the technical depth required for high-stakes industrial monitoring, the data collection framework was bifurcated into two primary streams:

Primary Stream (IoT Edge-Collector): High-frequency operational telemetry including; three-axis vibration, flow, and pressure was acquired via a custom-designed IoT edge-gateway. This physical layer utilized wired and wireless smart sensors to facilitate real-time interaction between the centrifugal pumps and the virtual diagnostic platform.

Secondary Stream (Maintenance History): Historical maintenance logs and equipment failure records were provided by the National Water and Sewerage Corporation (NWSC) are integrated into the dataset. These records provided the essential "ground-truth" labels required for the supervised training of the deep learning systems.

3.4.1 Tools/Instruments

Data acquisition was facilitated through a suite of industrial-grade hardware, primarily featuring Hansford HS-420 accelerometers for vibration monitoring and Proline Promag flow meters. These sensors were integrated with an Arduino Uno protyping board and Programmable Logic Controllers [PLCs] to enable localized data aggregation at the Ggaba and Katosi stations. The technical specifications and functional roles of these components are summarized in Table 3.1 below;

Instrument	**Model**	**Purpose**
Vibration sensor	Hansford HS-420	Mechanical health monitoring
Flow meter	Proline Promag	Operational volume tracking
Microcontroller	Arduino Uno	Data Aggregation/-Gateway
Interface	Industrial PLC	System-level integration

Table 3.1: Instruments and their purpose

3.4.2 IoT Monitoring, Transmission, and Diagnosis Architecture

- Perception and Transport (Real-time Monitoring): The architecture began with the physical monitoring of centrifugal pumps using HS-420 accelerometers and magnetic flow meters. This data was captured at the Ggaba and Katosi stations and transmitted via a hybrid PLC/GSM framework to an SQL server.

- Platform Layer (Data Processing and Diagnosis): This layer acted as the system's "brain." It utilized cloud-based algorithms for fault diagnosis and operational status evaluation. By analyzing accumulated data, the platform refined alarm thresholds and performed the high-level modeling required to diagnose pump health.

- Application Layer (User Interface and Predictive Maintenance): The final layer provided the results of the monitoring and diagnosis to NWSC managers. Through mobile and web terminals, users could check the real-time health of the pumps and arranged maintenance plans based on the automated diagnostic results.

3.4.3 Data Acquisition Procedure [Field work]

Installation and Calibration

The fieldwork commenced with the physical installation of the monitoring system at the NWSC pumping stations. This involved mounting the sensors on the bearing housings of the selected centrifugal pumps and configuring the Arduino-based gateway to recognize the unique IDs of both wired and wireless nodes. Prior to the 5-month data collection window, a calibration phase was conducted to ensure signal baseline stability and to verify that the GSM module could maintain a consistent handshake with the remote SQL server.

Data Collection Window and Monitoring

Data acquisition was carried out over a continuous period from January 2024 to May 2024. During this timeframe, the system was monitored for uptime, and periodic site visits were conducted to ensure that the physical connections remained intact despite the high-vibration environment. This stage was critical for capturing the seasonal operational variances in water demand, which directly impact pump load and mechanical stress patterns.

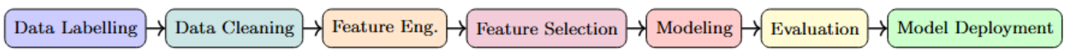

Figure 3.3: illustration of the Analysis flow

3.4.4 Technical Specifications of the Data Acquisition System

Component	**Model/Type**	**Purpose**
Vibration Sensor	Hansford HS-420	Captures high-frequency mechanical acceleration
Microcontroller	Arduino Uno	Acts as the primary "smart" data aggregator.
Communication	SIM800L GSM Module	Facilitates long-range wireless data transmission to the server.
Real-Time Tracking	DS3231 RTC Module	Provides real-time timestamps for the 2.9M data points.
Industrial Interface	Modbus RTU[PLC]	Enables seamless communication with existing NWSC PLCs.

Table 3.2: System components, models, and their purpose

3.4.5 Maintenance Records

Secondary data was obtained from National Water and Sewerage Corporation (NWSC) maintenance logs to provide ''labels"[e.g., normal operation vs bearing failure] for the supervised learning process. Archived pump maintenance records from National Water and Sewerage Corporation (NWSC) over the past 1 year were collected. These included: :

- Fault type and severity
- Date of failure
- Maintenance actions taken
- Pump age and runtime hours

This data served as ground truth labels for supervised learning. Other Equipment used for the gateway setup

- Arduino Uno for processing
- GSM Module for communication
- SD Card Module for data buffering and storage
- RTC Module for timestamps
- Project Server for the storage of the raw 2024 dataset for further analysis

3.5 Data Analysis Procedure [Computer work]

Data Labelling → Data Cleaning → Featuring engineering → Modeling → Evaluation

Figure 3.4: Illustration of the Data Procedure

3.5.1 Software Environment

The analysis was performed using python programming language within the Jupyter Notebook environment. The study utilized specialized libraries such as; Panda, NumPy, Scikit-learn, and CatBoost, which ultimately led to the identification of the most accurate model for detecting pump failure and gaining insights into operational status of NWSC pumping stations

3.5.2 Data Pre-processing

Raw sensor data underwent cleaning to handle missing values through interpolation. Feature Engineering was conducted to extract time-statistics-domain statistics [mean, RMS, Kurtosis] and frequency-domain features to enhance model performance. For instance;

- Missing value handling: Interpolation for time-series gaps
- Labeling: Annotated based on historical failures or predicted thresholds
- Feature Engineering:
  - Rolling Statistics (time-domain)
  - Trend Features (slopes, deltas)
  - Frequency-Domain Features (for vibration only)
  - Lag features (temporal dependency)
- Feature selection:
  - SelectKBest using ANOVA F-score
  - Tree-based importance (XGBoost)

3.5.3 Model Development

Five main algorithms were implemented: Random Forest, XGBoost, CATBoost, Voting Ensemble and Light -GBM.The dataset was split into training [70%], testing [15%] and validation [15%] sets to evaluate the predictive accuracy of each model. Thus, a detailed analysis of vibration sensor parameters was also conducted to determine their relevance in predicting pump failure, essentially acting as a statistical test to identify key indicators for predictive maintenance

3.6 Validity and Reliability

3.6.1 Validity

Validity represented the extent to which research instruments accurately measured the specific variables they were intended to evaluate [Bond and Fox, 2015]. In this study, validity was established through two distinct approaches:

- Instrumental validity: The HS-420 sensors were calibrated according to manufacturer specifications to ensure baseline measurement accuracy.
- Content validity: Sensor readings were cross-verified mechanical failures recorded in National Water and Sewerage Corporation (NWSC) maintenance logs.

To ensure the trustworthiness of the resulting predictions, evaluation metrics like precision, recall, F1-Score, confusion matrix were used to validate the model's ability to correctly identify pump health states.

3.6.2 Reliability

Reliability referred to the consistency of research outcomes when an experiment was repeated under similar conditions. To ensure the reliability of the machine learning models, K-fold cross-validation was implemented. This technique tested the model's predictive performance across different subsets of the 2,902,702 data points, ensuring that the results were not biased by a specific training set. Furthermore, the integration train-test-retest protocols contributed to the consistency of the model Evaluation, with statistical tables and performance graphs utilized to demonstrate the stability of the findings across the five-month data collection window [Preiser and Vischer, 2004]

3.7 Methodological Assumptions

The study assumed that the five-month data collection window captured a representative range of pump operational states. Furthermore, it was assumed that the mechanical properties of the pumps remained consistent, excluding the specific faults being monitored. And lastly, this research assumed that there was an objective reality that can be measured and quantified

3.8 Limitations: Stratified Random Sampling (Technique)

Purposive sampling ensured the inclusion of diverse operational failure modes, specifically targeting units with accessible maintenance histories.

a) a) Geographical Scope: The study was confined to specific stations at the National Water and Sewerage Corporation (NWSC), which may have limited the direct generalization of findings to pumps with significantly different mechanical configurations.

b) Time Constraints: A five-month data collection window was utilized. This duration, while statistically significant, may not have captured the full lifecycle of long-term failure modes such as structural fatigue.

c) Environmental Variables: The predictive models focused primarily on vibration, flow, and pressure; external factors such as power quality fluctuations and ambient humidity were excluded from the analysis.

d) Environmental Variables: The predictive models focused primarily on vibration, flow, and pressure; external factors such as power quality fluctuations and ambient humidity were excluded from the analysis.

e) Transmission Stability: The reliance on NB-IoT and GSM protocols introduced a dependency on local network uptime, which can occasionally impacted real-time data continuity in high-interference industrial zones.

f) Sensor Positioning: Mounting was restricted to available bearing housings on the centrifugal units. Minor variations in sensor orientation across different pumps may introduced subtle inconsistencies in signal amplitude.

3.9 Ethical Considerations

Permission was obtained from National Water and Sewerage Corporation (NWSC) management to access the pumping stations and historical data. All operational data was anonymized to ensure the security and privacy of the national utility's infrastructure information. [Jegede, 2009], stated that standard ethical principles that guided the treatment of human participants formed the foundation for the methodological approach in this study.

3.10 Conclusion

In summary, this chapter has detailed a robust methodological framework designed to transition from traditional maintenance practices to a data-driven predictive approach within Uganda's water sector. By integrating real-world sensor data through a structured IoT lifecycle, the study ensured that the over 2 million data points were processed with high reliability. Furthermore, by embedding explainable AI (XAI) techniques into the model selection process, this methodology ensured that the resulting predictions were not only high-performing but also transparent and actionable for National Water and Sewerage Corporation (NWSC) stakeholders. This scalable approach, tailored to local resource constraints and aging infrastructure, provided the necessary foundation for the experimental results and performance evaluations presented in Chapter 4."

Chapter 4

Analysis of Data

4.1 Introduction

This chapter provided an analysis of real-time sensor data and historical maintenance records for centrifugal pumps operated by the National Water and Sewerage Corporation (NWSC)]. The derived results were presented objectively to evaluate the performance of the developed predictive models. Furthermore, the effectiveness of these models in detecting early signs of pump failure was interpreted. To communicate these findings clearly, data visualizations, tables and metrics were utilized to demonstrate how machine learning algorithms effectively predict and provided the basis for preventing potential infrastructure issues.

4.2 Data Analysis

4.2.1 Dataset Preparation

The methodology processed and analysed quantitative data from the National Water and Sewerage Corporation (NWSC) centrifugal pumps using Python. To transform the raw high-frequency sensor readings into a format suitable for predictive modeling, the approach utilized a sliding window method. This technique facilitated the grouping of data points into discrete temporal segments, allowing for the extraction of statistical features that characterized the pump's health over time. For each window, the analysis calculated a feature matrix comprising statistical descriptors such as the mean, maximum, minimum, and standard deviation for both vibration and acoustic signals. This structured approach ensured that the study accounted for a variety of potential fault conditions, including:

- Impeller degradation (clogging or breakage).
- Bearing housing abnormalities (inner and outer race defects).
- Operational imbalances and cavitation.

- Hydraulic and aerodynamic failures
- Journal bed failures
- Electrical motor failures

By consolidating these time-domain features with frequency-domain components derived from the Fast Fourier Transform (FFT), the process created an enriched, 5-column feature set. This exhaustive layout provided the necessary input for the machine learning algorithms to effectively differentiate between normal operating states and early-onset failure patterns.

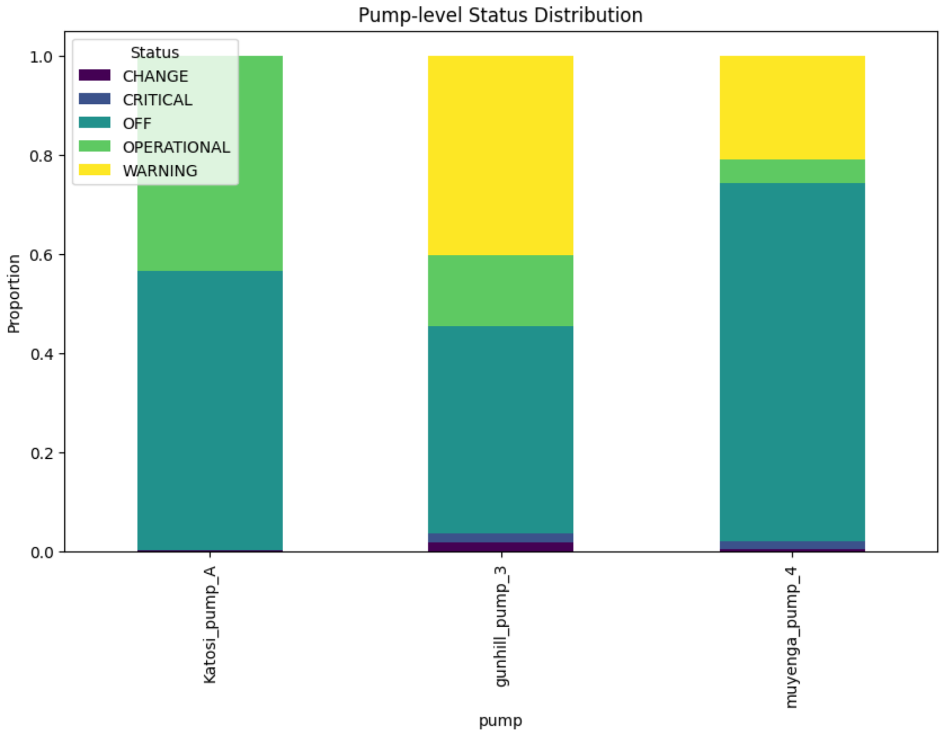

Figure 4.1: Status distribution per pump.

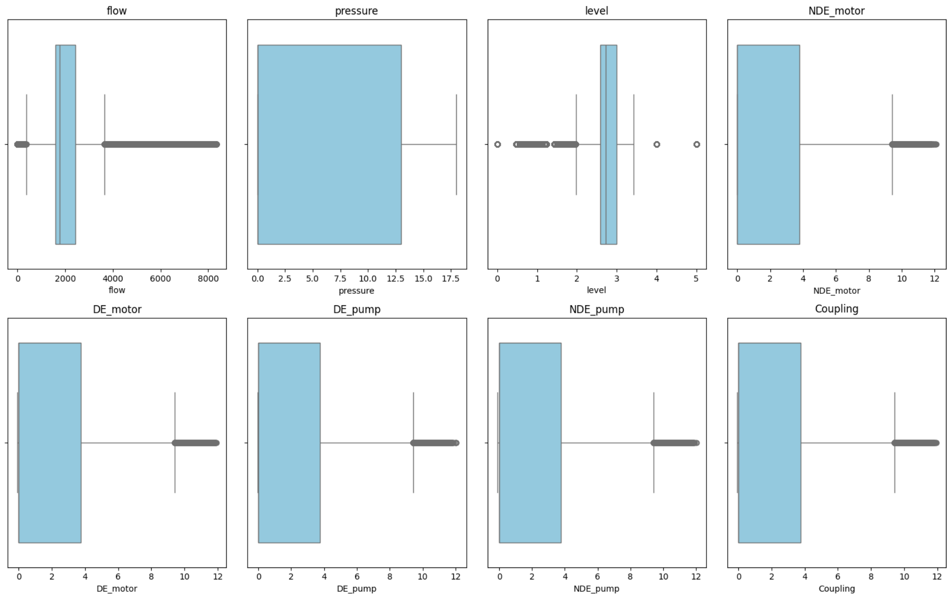

Figure 4.2: Outlier distribution.

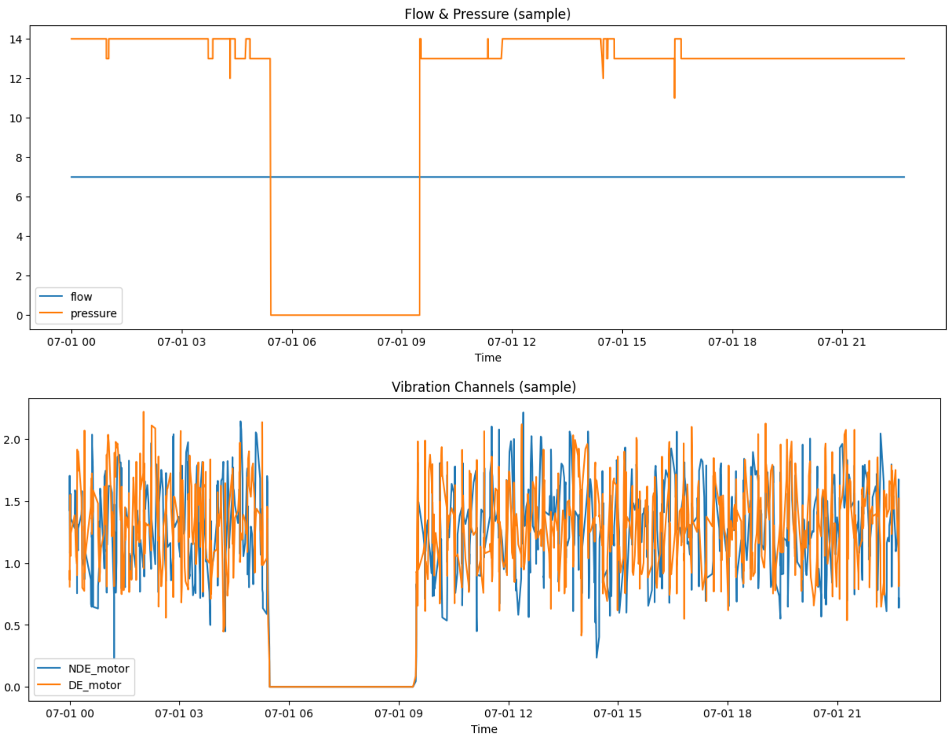

Figure 4.3: Visual representation of fault in dataset.

4.2.2 Imbalanced Dataset and Distribution Analysis

The analysis of the machine status labels highlighted a significant class imbalance within the 2024 sensor dataset. As was illustrated in Figure 4.4, the distribution was disproportionately concentrated toward non-failure states, with the "OFF" category accounting for 56.8% (1,648,816 points) and "OPERATIONAL" representing 20.8% (604,958 points) of the total 2,902,702 observations. In contrast, the critical minority classes; Warning (20.5%), Critical (1.05%), and Change (0.823%) comprised a much smaller portion of the data.

The methodology identified this imbalance as a characteristic challenge of industrial predictive maintenance, where failure events were naturally rare compared to steady-state operations. To ensure the predictive models did not develop a bias toward the majority classes, the investigation acknowledged the necessity of specialized evaluation metrics. Depending entirely on overall accuracy would be misleading in this context; thus, the approach prioritized the use of Precision, Recall, and the F1-Score. These metrics allowed for a more exhaustive assessment of the model's ability to successfully identify the "Critical" and "Warning" states, which were the most vital for preventing pump failure at the National Water and Sewerage Corporation (NWSC).

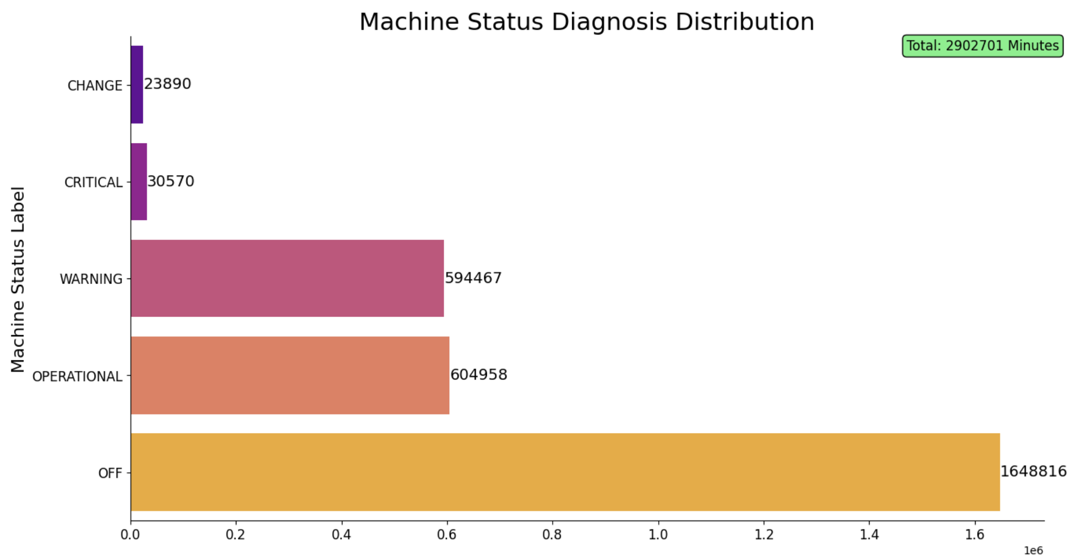

Figure 4.4: Status distribution

4.3 Features

Accurate and meaningful feature selection is widely regarded as one of the most important steps in building effective machine learning models [Kuhn and Johnson, 2013]. The quality of features directly impacts a model's ability to learn patterns and make accurate predictions.

4.3.1 Feature Engineering and Selection

Feature engineering and selection served as fundamental processes in building robust machine learning models, particularly in predictive maintenance where sensor data was high- dimensional and contained significant noise. In this investigation, the methodology adopted the SelectKBest algorithm using the ANOVA F-score to identify the most impactful variables for failure prediction.

4.3.2 SelectKBest using ANOVA F-Score and Importance.

This method operated as a filter-based feature selection technique that evaluated the statistical relevance of each individual feature to the target variable (e.g., pump failure). The ANOVA F-score calculated the statistical difference between the mean values of a feature across different classes. A higher F-score indicated that a specific feature possessed greater capacity to differentiate between normal operation and various failure states (for example., normal (operational) vs. various Critical (failure modes).

By prioritizing features with the highest scores, the analysis reduced the dimensionality of the dataset. This simplification ensured that the model remained computationally efficient while focusing exclusively on the signals that provided the highest predictive value.

4.3.3 Feature description (Project setup)

The dataset was structured to capture temporal, operational, and target variables. The methodology defined the machine status as a multi-class classification target, representing the physical state of the centrifugal pumps at any given timestamp. As detailed in Figure 4.5, these statuses included:

- OFF: The pump was in a non-operational state
- OPERATIONAL: The pump functioned within normal parameters.
- CHANGE: Indications of early-stage deviations were observed, requiring attention.
- WARNING: Significant operational anomalies were detected.
- CRITICAL: An imminent or active failure state was identified, requiring immediate intervention.

The transitions between these statuses provided the maintenance team with actionable insights, allowing for data-driven based on the severity of the indication.

Variable	Role	Unit	Frequency
Time	Datetime	Datetime	Datetime
flow	Continuous	m³/h	per second
pressure	Float	bar	per second
level	Continuous	m	per second
NDE_motor	Float	mm/s	per second
DE_motor	Float	mm/s	per second
DE_pump	Float	mm/s	per second
NDE_pump	Float	mm/s	per second
Coupling	Float	mm/s	per second
status	Target	State	per second
pump	Categorical	pump	per second

Datetime
Continuous
Float
Target
Categorical

Figure 4.5: Dataset Features.

4.4 Proposed ML models

The methodology implemented a multi-model comparative framework to identify the most accurate algorithm for predicting centrifugal pump health with the combination of sequences of historical data to learn temporal dependencies and long-term degradation. Five primary algorithms were selected based on their documented success in handling noisy, high-dimensional industrial sensor data. To ensure the robustness of the results, the study tested different algorithms against multiple feature matrices

4.4.1 Random Forest

Random Forest was utilized as an ensemble of decision trees to mitigate the risk of overfitting through majority voting. This model was particularly effective at handling the class imbalance within the dataset by focusing on the most insightful signals from the pump's vibration and flow sensors [Breiman, 2001b].

4.4.2 Light Gradient Boosting Machine (LightGBM)

LightGBM was deployed due to its histogram-based splitting logic, which accelerated the training process for the large-scale National Water and Sewerage Corporation (NWSC) dataset. The framework's native support for scale_pos_weight facilitated better identification of the rare "Critical" and "Warning" states [Ke et al., 2017]:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t), \tag{4.1}$$

where $g_i = \frac{\partial l(y_i,\hat{y}_i)}{\partial \hat{y}_i}\Big|_{\hat{y}_i=\hat{y}_i^{(t-1)}}$ and $h_i = \frac{\partial^2 l(y_i,\hat{y}_i)}{\partial \hat{y}_i^2}\Big|_{\hat{y}_i=\hat{y}_i^{(t-1)}}$ are the first and second derivatives of the loss function with respect to the previous iteration's prediction [Ke et al., 2017]. LightGBM also supports parameters to handle class imbalance, such as is unbalance and scale pos weight.

4.4.3 XGBoost

XGBoost was implemented to leverage advanced regularization (L1 and L2) to prevent model complexity from leading to poor generalization. It efficiently handled missing data values through internal sparsity-aware splitting [Chen and Guestrin, 2016]:

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k), \tag{4.2}$$

where

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2, \tag{4.3}$$

and

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i), \quad f_k \in \mathcal{F}. \tag{4.4}$$

Here,

- $l(y_i, \hat{y}_i)$ is a convex loss function that measures the difference between prediction $\hat{y}_i$ and true label y_i,
- $\Omega(f_k)$ is the regularization term that penalizes model complexity,
- T is the number of leaves in tree f,
- w are the leaf weights,
- γ and λ are regularization parameters controlling tree complexity and weight shrinkage,
- $\hat{y}_i^{(t)}$ is the prediction for sample i at iteration t,
- $f_k \in \mathcal{F}$, where $\mathcal{F}$ is the space of regression trees.

[Chen and Guestrin, 2016] XGBoost efficiently handles missing data and large-scale problems, making it suitable for PdM datasets.

4.4.4 CATBoost

CATBoost was chosen to minimize the need for extensive categorical pre-processing of pump station identifiers. By employing ordered boosting, the study prevented prediction shift, ensuring that the health indicators remained reliable across different mechanical units [Prokhorenkova et al., 2018].

4.4.5 Voting Ensemble

A Voting Ensemble was developed to aggregate the predictive strengths of the individual models. By using soft voting (averaging predicted probabilities), the ensemble provided a more stable diagnostic output than any single constituent model unlike hard voting [Kuncheva, 2004]. Further, [Sempewo et al., 2019, Gonçalves et al., 2021] it was selected due to its ability to handle large, noisy, and imbalanced datasets typical of water utility pump operations in Uganda.

4.5 Project evaluation metrics

To rigorously evaluate the efficacy of the proposed machine learning models, the study employed a multi-dimensional assessment framework. Because the NWSC dataset exhibited a significant class imbalance, the methodology prioritized metrics that provided a granular view of model performance beyond simple accuracy. These metrics were selected to ensure that the detection of rare but critical failure states was both precise and reliable.

4.5.1 Confusion Matrix Analysis

The confusion matrix served as the primary diagnostic tool for visualizing model performance. It facilitated an exhaustive comparison between the actual mechanical states and those predicted by the algorithms. By categorizing outcomes into True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), the analysis identified specific areas where models confused "Warning" states with "Critical" failures, allowing for targeted refinements in the voting ensemble as illustrated in Table 4.1 [Han et al., 2022].

		Predicted Values	
		Anomaly (1)	**Normal (0)**
Actual Value	**Anomaly (1)**	True Positive (TP)	False Negative (FN)
	Normal (0)	False Positive (FP)	True Negative (TN)

Table 4.1: Confusion Matrix for binary classification (Anomaly vs Normal).

While the confusion matrix provided detailed insight, additional evaluation metrics such as accuracy, precision, recall, and F1-score were commonly used to assess model performance i.e. in a binary classification setting, where the classes are labelled as '1' (positive) and '0' (negative), four key outcomes are defined:

1. True Positive (TP): the model correctly predicts a positive case.
2. True Negative (TN): the model correctly predicts a negative case.

3. False Positive (FP): the model predicts a positive case when the actual class is negative.
4. False Negative (FN): the model predicts a negative case when the actual class is positive.

These outcomes form the basis for calculating the different evaluation metrics.

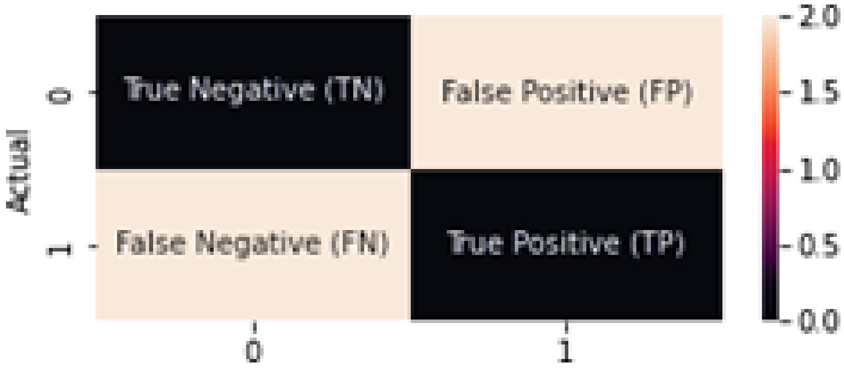

Figure 4.6: Binary confusion matrix

4.5.2 Accuracy:

Although accuracy is the most often used metric, it only provides a partial picture of the model's performance. The formula 4.5 shows that the total number of predictions and the total number of correct classifications are compared. However, it is important to remember that this model is only effective when the dataset is well-balanced. However, in anomaly detection, for instance, unbalanced data may yield good accuracy even if it only predicts everything in the majority class. Since the data is primary, it is well-balanced, making accuracy an excellent approach to assess the model's performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.5}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

4.5.3 Sensitivity

When interested in a particular class, sensitivity and recall are the measurements that are frequently used. As seen in 4.6, it calculates the frequency of correctly classified classes divided by their total number. The primary interest in a flight engine, for instance, where the cost is too great to take any chances, may be the class of potential problems in the context of PdM.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.6}$$

where TP = True Positives and FN = False Negatives.

4.5.4 Precision

As seen in 4.7, precision is the sum of all right guesses divided by all predictions for that class. As a result, it measures how well the model predicts a class regardless of how many of a different class are incorrect..

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.7}$$

where TP = True Positives and FP = False Positives.

4.5.5 F1-score

The final common metric is the F1-score, which combines precision and sensitivity. When there is an unbalanced dataset with no extreme values—for example, neither accuracy nor sensitivity is zero—F1 is frequently utilized since, as the formula indicates, F1 will likewise be zero. 4.8.

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.8}$$

where Precision and Recall are as defined in Equations 4.7 and 4.6.

4.5.6 ROC-AUC Score

The model's capacity to differentiate between classes at different threshold settings is gauged by the Receiver Operating Characteristic Area Under Curve (ROC-AUC). AUC values vary from 0 to 1, where 1 denotes flawless classification. Better model performance is indicated by a larger AUC score. [Fawcett, 2006].

4.6 Base algorithm Testing

The evaluation results are summarized in Table 4.2. All seven machine learning models achieved good accuracy, with Random Forest, Voting Ensemble, and CatBoost performing slightly better than the rest. The main differences emerged in terms of precision and recall. Precision indicates how well a model correctly identifies positive cases, while recall reflects its ability to detect all positive cases in the dataset. Random Forest, Voting Ensemble, and CatBoost achieved the highest precision, whereas Logistic Regression recorded the lowest. Similarly, these three models also demonstrated strong recall, while KNN and Logistic Regression lagged behind. Considering the F1-score, which balances precision and recall, Random Forest, Voting Ensemble, and CatBoost again outperformed the other models. Overall, Random Forest consistently achieved the best results across all evaluation metrics.

Model	Precision	Recall	F1 Score	Accuracy
Voting Ensemble	0.6991	0.7251	0.7080	0.7251
CATBoost	0.7089	0.6655	0.6792	0.6655
XGBoost	0.7078	0.7375	0.7101	0.7375
Random Forest	0.7131	0.6544	0.6793	0.6544
LightGBM	0.6891	0.6743	0.6778	0.6743

Table 4.2: Performance comparison of different machine learning models.

4.6.1 Confusion Matrices

This section presents the resulting confusion matrices from the evaluation for each algorithm.

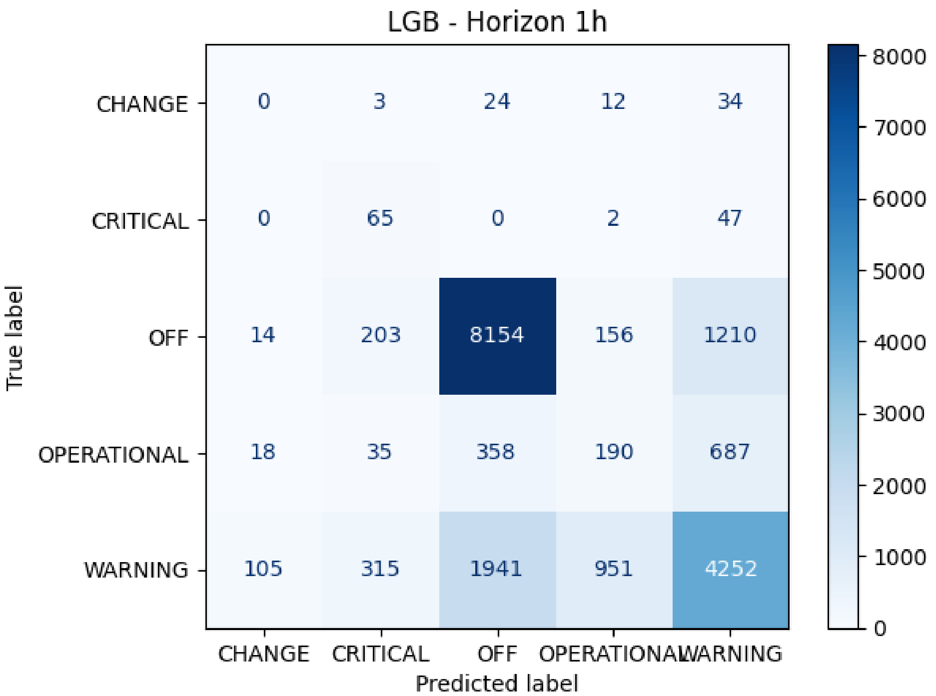

Figure 4.7: Visual representation of confusion matrix for LGB model

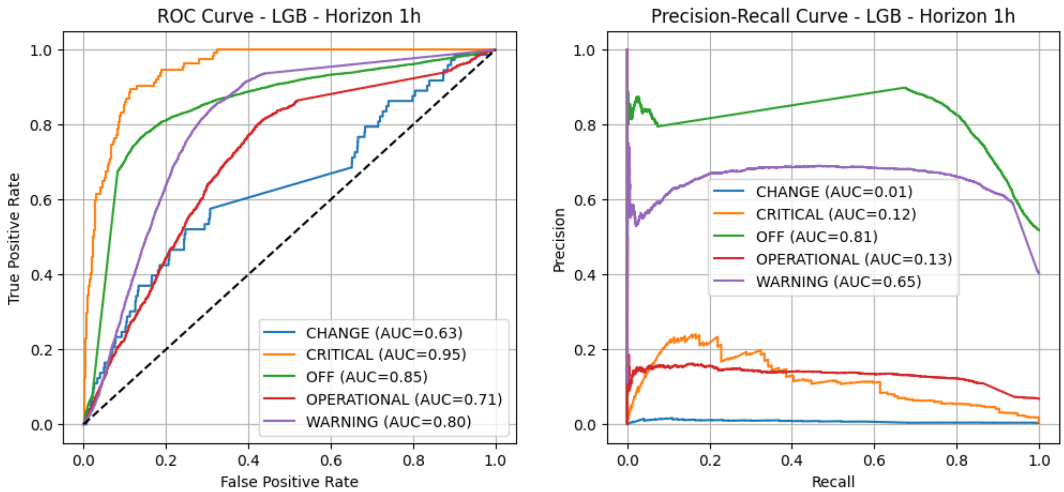

Figure 4.8: Visual representation of LGB ROC and PR curves

The image provided display the performance of a LightGBM model for a multiclass classification task. The confusion matrix shows the model's predictions against the actual labels, revealing that it performs best at identifying the 'OFF' and 'WARNING' classes, with 8,154 and 4,252 correct predictions, respectively. However, the matrix also highlights significant misclassification errors, with a large number of 'WARNING' instances being incorrectly predicted as 'OFF' (1,941 cases), and many 'OFF' instances being mislabeled as 'WARNING' (1,210 cases). The ROC and Precision-Recall curves offer a more detailed analysis. The ROC curve shows that the model is highly effective at distinguishing the 'CRITICAL' and 'OFF' classes, indicated by high Area Under Curve (AUC) values of 0.95 and 0.81, respectively. In contrast, the Precision-Recall curve reveals a critical weakness for the 'CHANGE' and 'OPERATIONAL' classes, with extremely low AUC values of 0.01 and 0.13. This suggests that the model struggles to make accurate positive predictions for these specific categories, likely due to a high number of false positives or a low number of true positives. Overall, while the LightGBM model is strong in predicting some classes, its performance is inconsistent and poor for others.

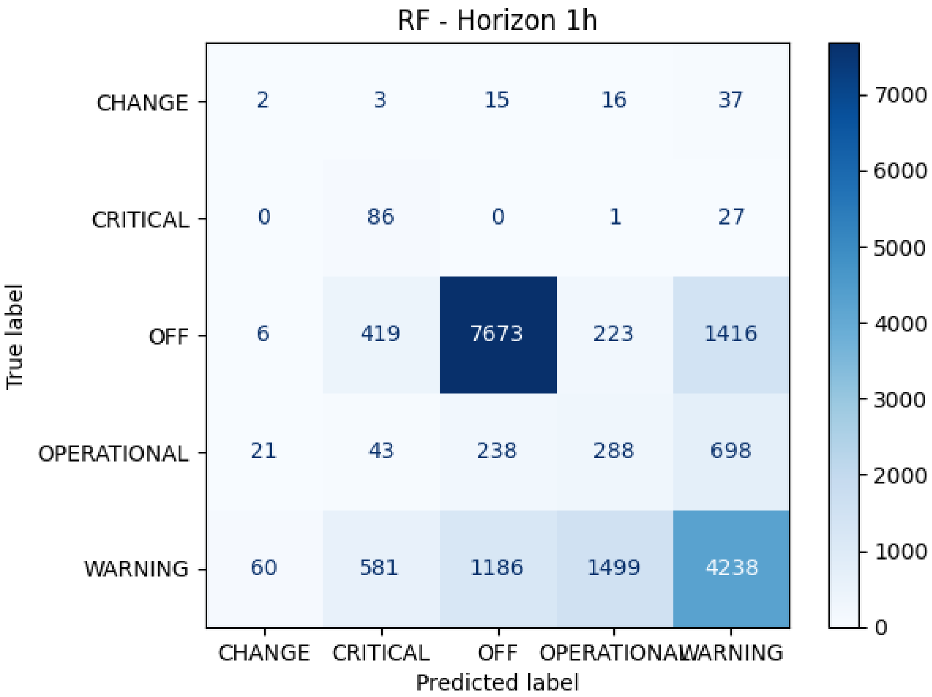

Figure 4.9: Visual representation of confusion matrix for RF model

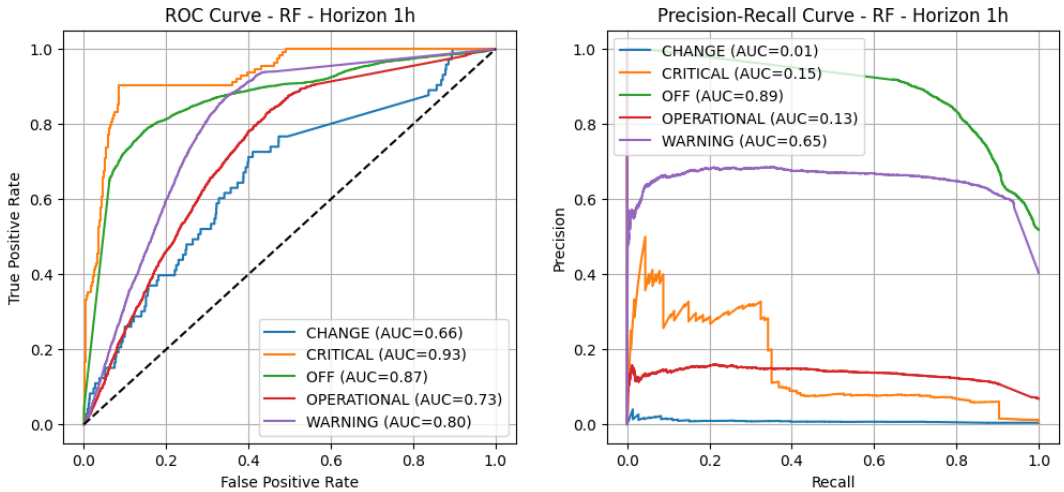

Figure 4.10: Visual representation of RF ROC and PR curves

4.6.2 Overall Model Performance:

The model achieves an accuracy of 0.6655, meaning it correctly classifies approximately 66.55% of the instances. It demonstrates a macro average precision of 0.3440, recall of 0.4769, and f1-score of 0.3333 across all classes. The weighted average precision stands at 0.7089, recall at 0.6655, and f1-score at 0.6792, considering the support for each class. The model achieves an AUC score of 0.8030, indicating its ability to distinguish between classes.

4.6.3 Class-Specific Performance:

The report details the model's performance for individual classes: 'CHANGE', 'CRITICAL', 'OFF', 'OPERATIONAL', and 'WARNING'. For the 'CHANGE' class, the model achieves a precision of 0.0126, a recall of 0.0411, and an f1-score of 0.0193, with 73 instances supporting this class. For the 'CRITICAL' class, it shows a precision of 0.0588, a high recall of 0.8772, and an f1-score of 0.1101, based on 114 instances. For the 'OFF' class, the model demonstrates strong performance with a precision of 0.7896, a recall of 0.8397, and an f1-score of 0.8139, supported by 9737 instances. For the 'OPERATIONAL' class, it achieves a precision of 0.1410, a recall of 0.0831, and an f1-score of 0.1045, with 1288 instances. For the 'WARNING' class, the model exhibits a precision of 0.7182, a recall of 0.5434, and an f1-score of 0.6186, supported by 7564 instances.

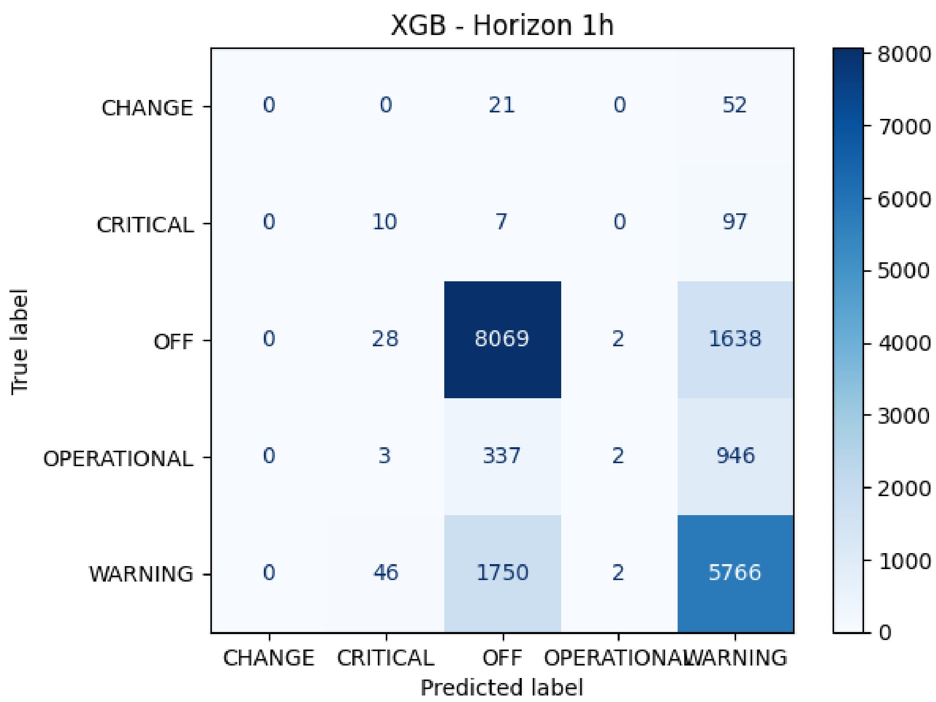

Figure 4.11: Visual representation of confusion matrix for XGB model

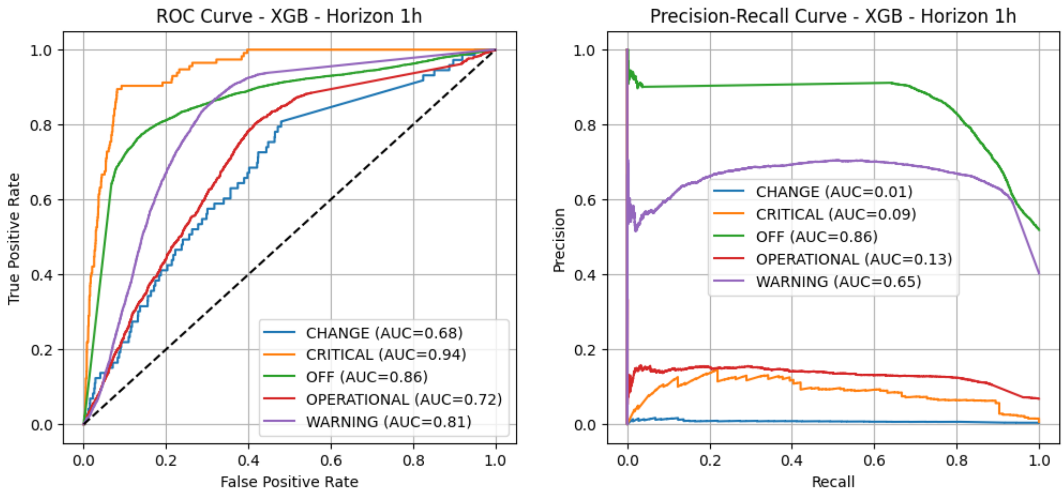

Figure 4.12: Visual representation of XGB ROC and PR curves

This classification report summarizes the performance of a machine learning model. The model achieves an overall accuracy of 0.6544, a precision of 0.7131, a recall of 0.6544, and an F1-score of 0.6793. It also shows an AUC score of 0.7976794972646977. The report details performance metrics for individual classes: CRITICAL achieves a precision of 0.0760, recall of 0.7544, and F1-score of 0.1380 with 114 samples, OFF demonstrates a precision of 0.8421, recall of 0.7880, and F1-score of 0.8142 with 9737 samples, OPERATIONAL records a precision of 0.1421, recall of 0.2236, and F1-score of 0.1738 with 1288 samples, WARNING presents a precision of 0.6605, recall of 0.5603, and F1-score of 0.6063 with 7564 samples. Additionally, the report provides aggregated metrics: The macro average for precision is 0.3486, for recall is 0.4707, and for F1-score is 0.3514, the weighted average for precision is 0.7131, for recall is 0.6544, and for F1-score is 0.6793. The total number of samples processed is 18776. CHANGE shows a precision of 0.0225, recall of 0.0274, and F1-score of 0.0247 with 73 samples.

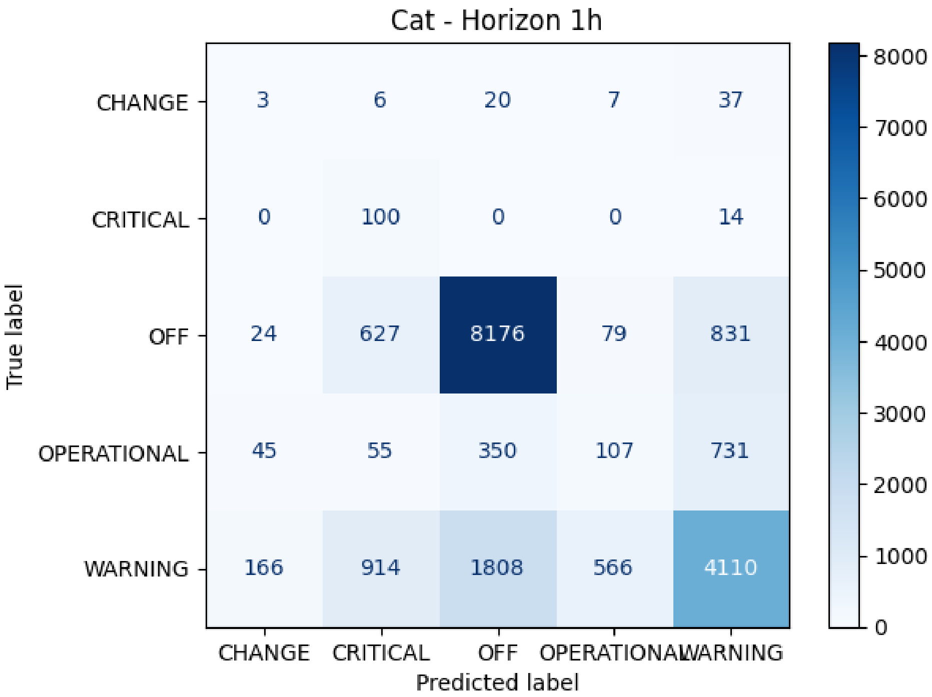

Figure 4.13: Visual representation of confusion matrix for CAT model

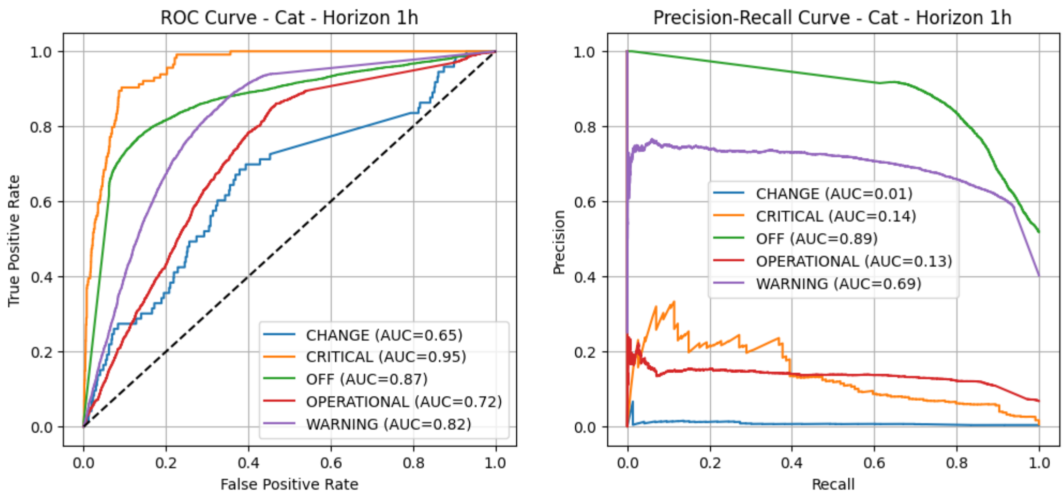

Figure 4.14: Visual representation of CAT ROC and PR curves

The confusion matrix shows the classification results for a multiclass problem with five distinct labels: CHANGE, CRITICAL, OFF, OPERATIONAL, and WARNING. The diagonal values represent the number of correct predictions, while the off-diagonal values indicate misclassifications. For example, the model correctly predicted 8,176 instances of the 'OFF' label but misclassified a significant number of 'WARNING' cases as 'OFF' (1,808 instances) and 'OPERATIONAL' (566 instances). The ROC curve and Precision-Recall curve further evaluate the model's performance for each class. The ROC curve plots the true positive rate against the false positive rate at various thresholds, with the Area Under the Curve (AUC) indicating the model's overall discriminative ability for each class. The AUC values suggest the model performs best on the 'CRITICAL' and 'OFF' classes (AUCs of 0.95 and 0.87, respectively) and worst on the 'CHANGE' class (AUC of 0.65). The Precision-Recall curve provides a different perspective, showing the trade-off between precision and recall for each class. The low AUC for the 'CHANGE' class (0.01) on the Precision-Recall curve suggests that the model struggles to identify positive instances of this class without a high rate of false positives. Overall, the provided charts collectively offer a comprehensive view of the CATBoost model's strengths and weaknesses across the different classes.

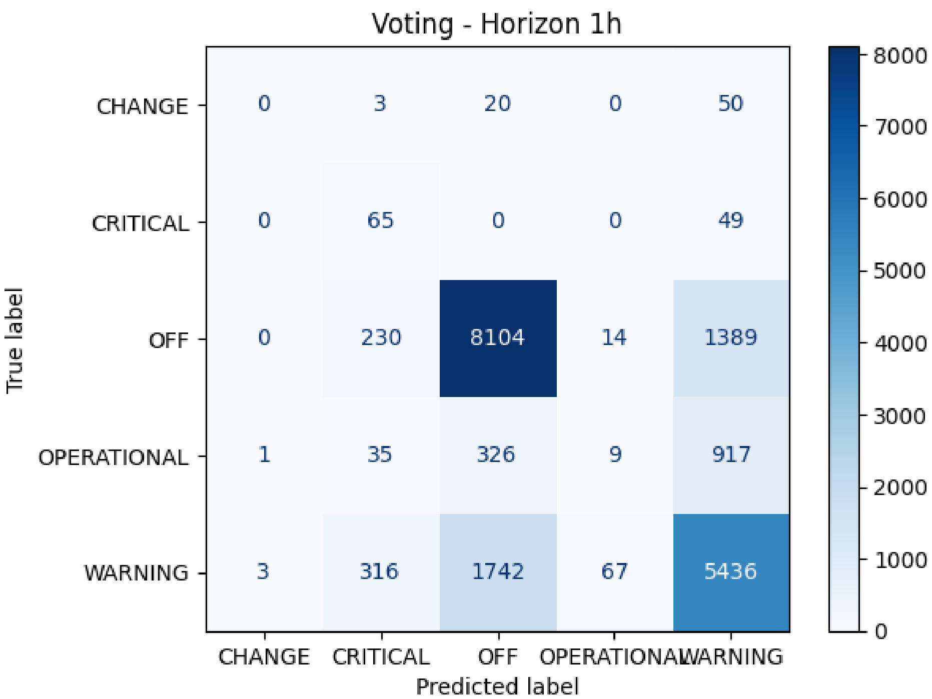

Figure 4.15: Visual representation of confusion matrix for Voting model

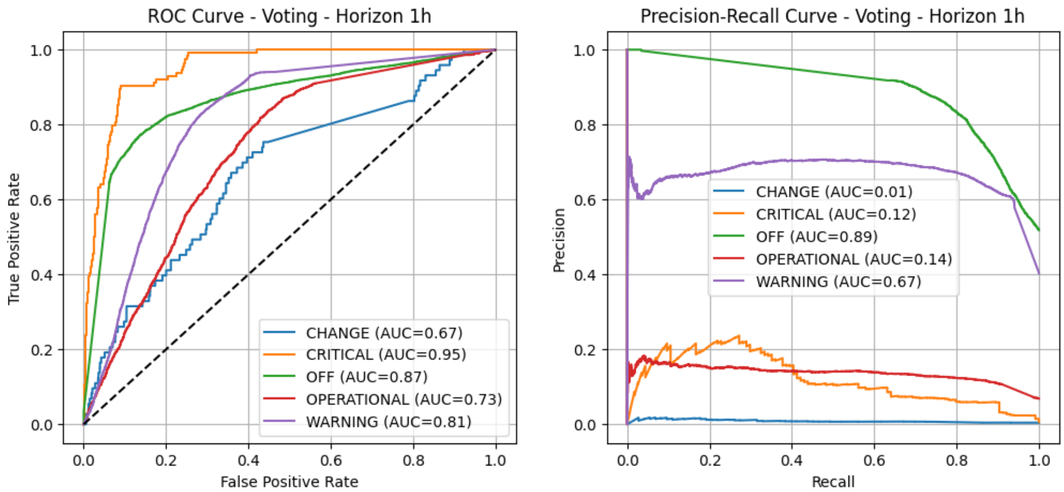

Figure 4.16: Visual representation of Voting ROC and PR curves

4.7 Definition of training, Validation and test sets

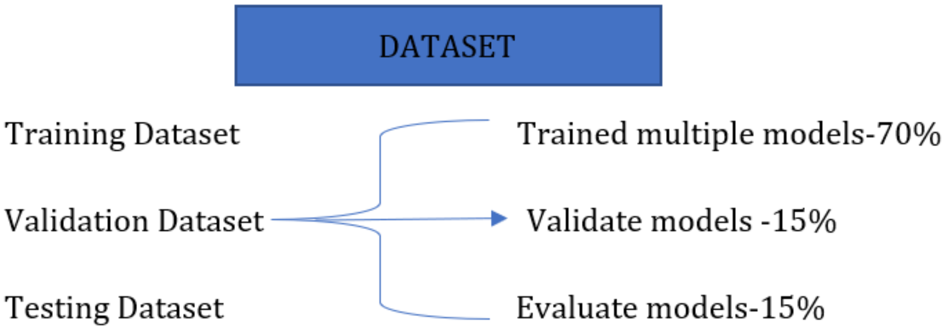

Figure 4.17: Visual representation of dataset splitting.

As per the Experiment section, the development set consists of entirely data, excluding the periods of time when faults occurred (that are part of the test set). The development set is split chronologically using a time-based approach, with 70% of the data allocated to training, 15% to validation, and the remaining 15% to testing. The test set comprises of three available faults. Per classification, a two days interval of normal behaviour data before and after the incident is also obtained. Another time interval lets us analyse the behaviour of the model when the pump shifts from a normal to an abnormal phase. Prediction models are trained using supervised machine learning techniques. Each time step in the dataset is manually labeled with one of the actual pump states CHANGE, CRITICAL, OFF, OPERATIONAL, or WARNING based on maintenance records and expert technician analysis. These categorical labels are then encoded numerically for model training while preserving the original status names for evaluation. The resulting dataset contains two columns: the timestamp and the labeled pump state, enabling accurate evaluation through classification reports, confusion matrices, and ROC/PR curves.

Chapter 5

Results

5.1 Introduction

This chapter presented the findings from the model optimization phase and evaluated the performance of various machine learning frameworks. The study established a robust technical baseline for predictive maintenance at the NWSC by analyzing multi-sensor data and vibration signatures. Furthermore, the research demonstrated the best parameters and strategies discovered during the process and derived specific conclusions regarding the deployment of the final model.

5.2 Results

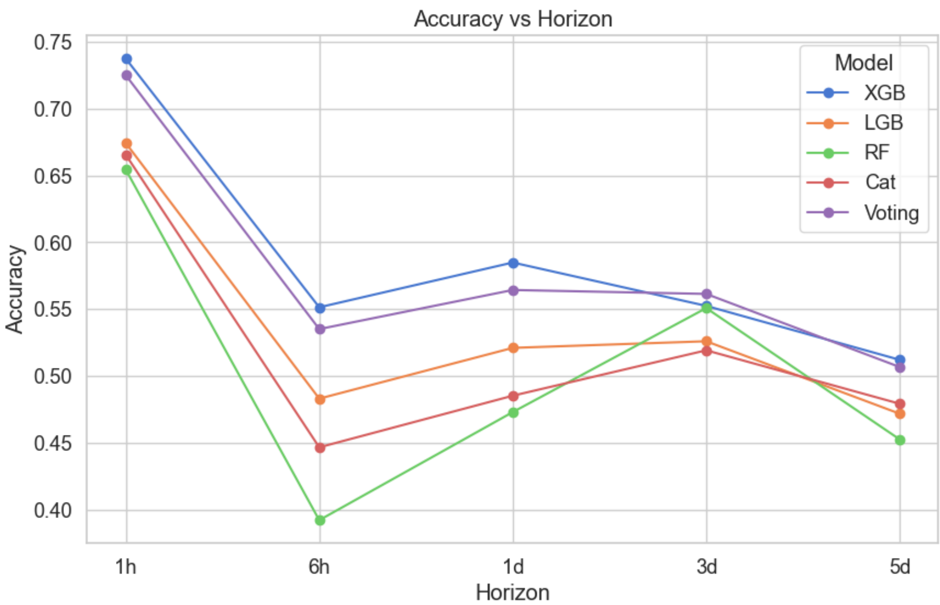

Figure 5.1: Performance of the different models

The performance of the classification models was evaluated across multiple prediction horizons, ranging from one hour to five days. Figure 5.1 illustrates the accuracy trends for the tested models. Horizon-Dependent Performance: The results indicate a clear inverse relationship between prediction horizon and accuracy. All models demonstrate peak performance at the 1-hour mark, confirming that short-term forecasts are the most reliable for immediate operational intervention. Model Comparison: XGBoost emerges as the most consistent performer, attaining the highest accuracy at both the 1-hour and 1-day horizons. While LightGBM (LGB) follows closely, the Random Forest (RF) and CatBoost models show significantly lower precision as the horizon extends. Ensemble Strategy: The Voting model integrates multiple architectures and shows strong initial performance; however, it experiences a sharper decline after the 1-hour mark compared to the standalone XGBoost model

5.3 Descriptive Analysis

The study analysed eight sensor groups, including Flow, Pressure, and vibration parameters from the motor, pump bearings, and coupling. This multi-sensor approach was essential for capturing the transition between healthy and degraded states.

Operational Stability (Green Phase): In the OPERATIONAL phase, flow and pressure readings remained stable. Vibration signals from the motor and pump bearings exhibited only minor fluctuations, signifying steady-state performance.

Degradation Onset (Yellow/Warning Phase): As the pump entered the CHANGE and WARNING zones, vibration variability increased significantly. These disturbances represented early mechanical wear, such as bearing fatigue or imbalance, which are often detectable in vibration signatures before they manifest as hydraulic performance drops.

Sustained Criticality (Red Phase): 5.3 presented a scenario of prolonged CRITICAL operation. In this state, hydraulic efficiency was reduced (lower flow/pressure), while vibration signals remained consistently elevated.

Technical Significance: The lack of transition back to normal zones suggests that these faults are not self-correcting. These findings underline the necessity of continuous monitoring to prevent catastrophic breakdowns caused by shaft misalignment or cavitation.

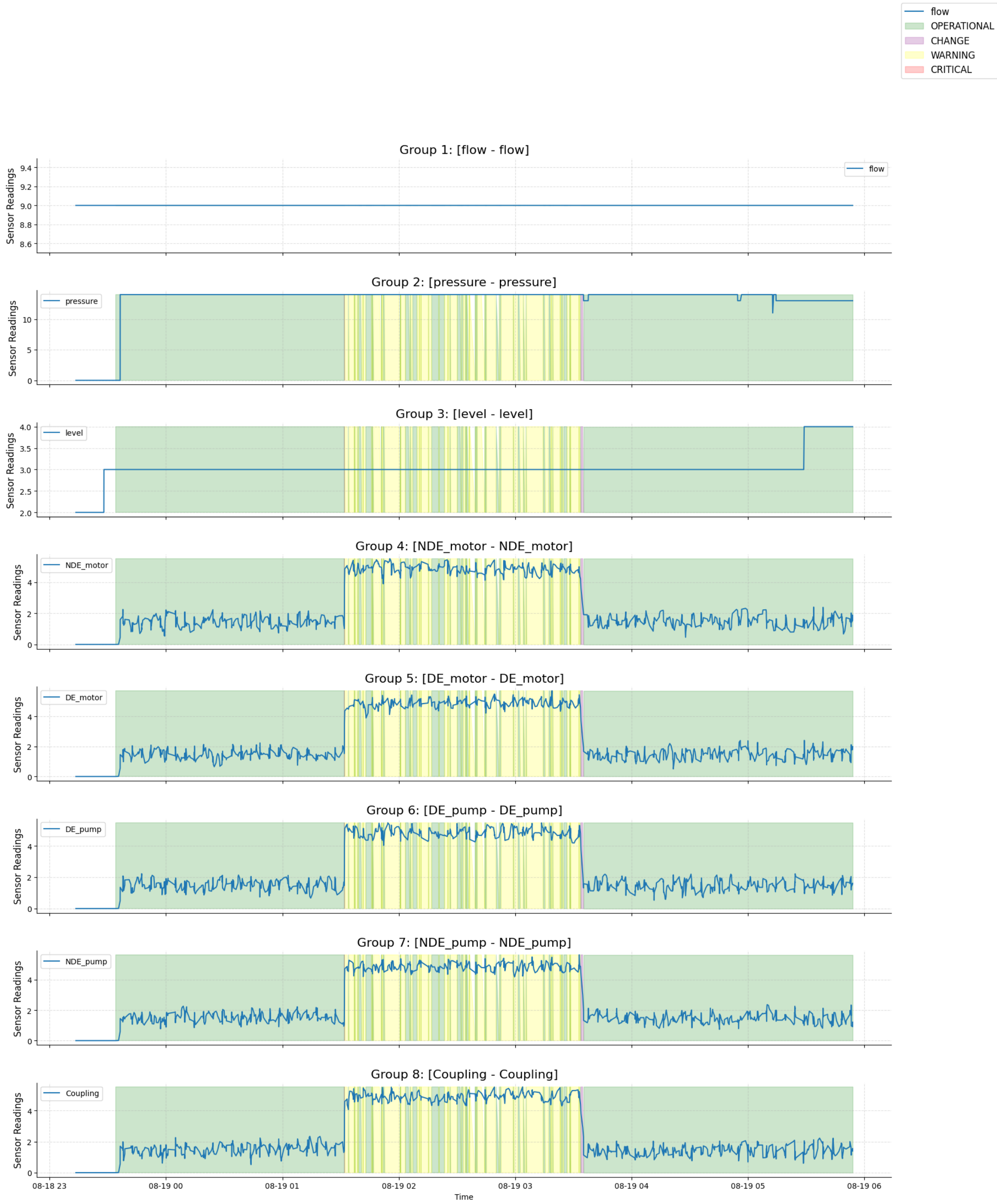

Figure 5.2: Normal Operating Cycle with Fault Transition

In contrast, 5.3 presented a scenario in which the pump remained in a CRITICAL operating state for a prolonged duration. Here, flow and pressure values were relatively flat but at lower levels compared to normal operation, suggesting reduced hydraulic efficiency. Simultaneously, vibration signals from all bearing locations (NDE motor, DE motor, DE pump, NDE pump, and Coupling) remained consistently elevated, indicating persistent abnormal mechanical stress. Such prolonged deviations aligned with severe failure modes, including shaft misalignment, cavitation, or advanced bearing damage

[Randall, 2011]. The absence of transitions back to normal or warning zones suggested that the fault condition was not self-correcting and would have required immediate maintenance intervention. This sustained critical state underlined the importance of continuous monitoring and alarm thresholds for preventing catastrophic breakdowns.

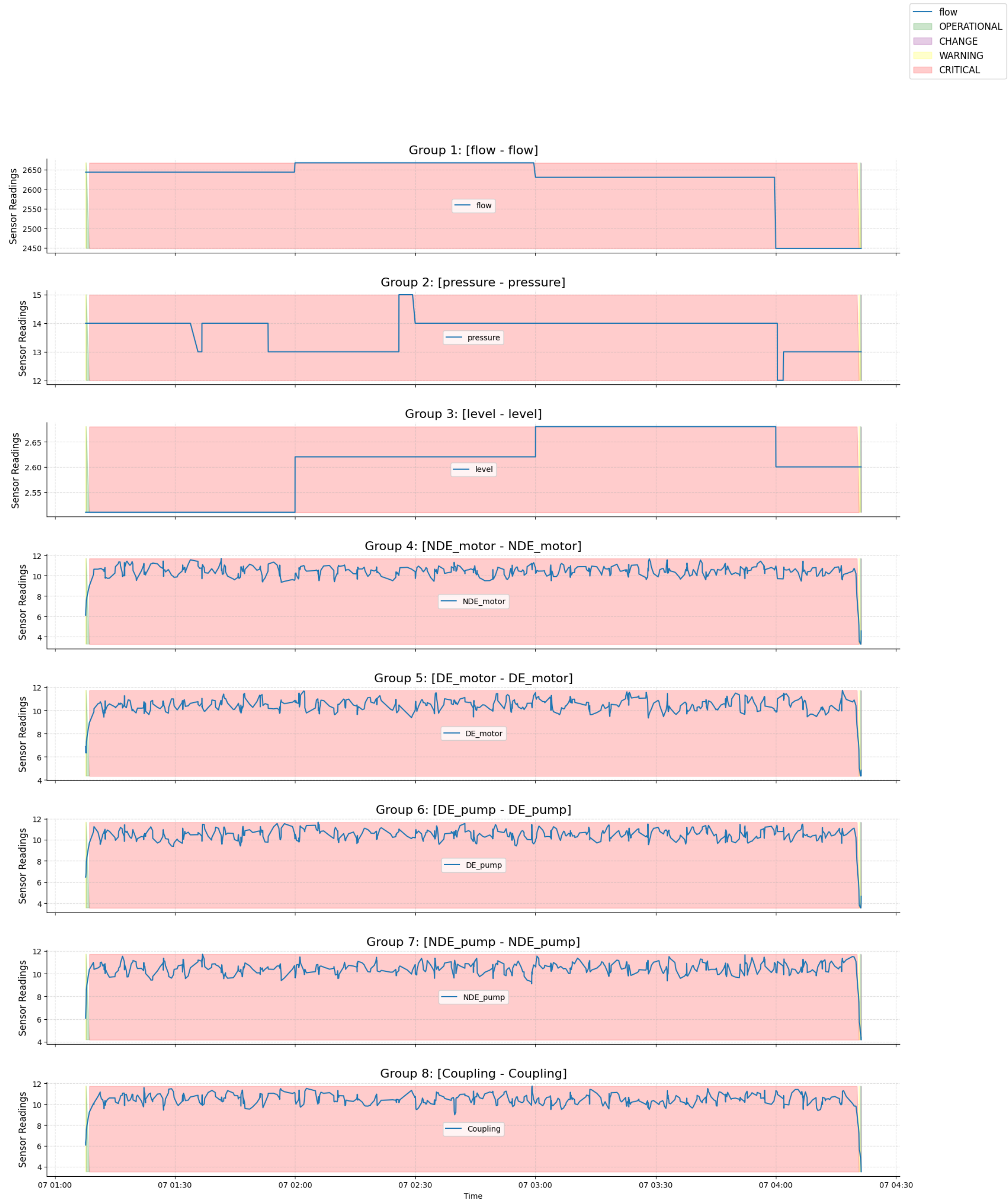

Figure 5.3: Sustained Critical Condition

Figure 5.4 depicted a dynamic operational profile characterized by multiple status changes between OPERATIONAL, WARNING, CHANGE, and CRITICAL states. At the start, both hydraulic and vibration parameters remained within normal ranges. Over time, the system underwent repeated oscillations between warning and critical phases, pointing to intermittent disturbances such as fluctuating loads or transient suction/discharge conditions. Vibration sensors again showed more pronounced responses compared to hydraulic sensors, reaffirming their diagnostic strength in capturing transient mechanical instabilities [Jardine et al., 2006]. Interestingly, towards the end of the monitoring period, sensor readings stabilized, and the pump reverted to an OPERATIONAL state, suggesting either the subsiding of temporary disturbances or corrective adjustments in system conditions. This figure demonstrated the monitoring framework's ability to capture not only fault onset and persistence but also fault recovery, which was essential for predictive maintenance applications.

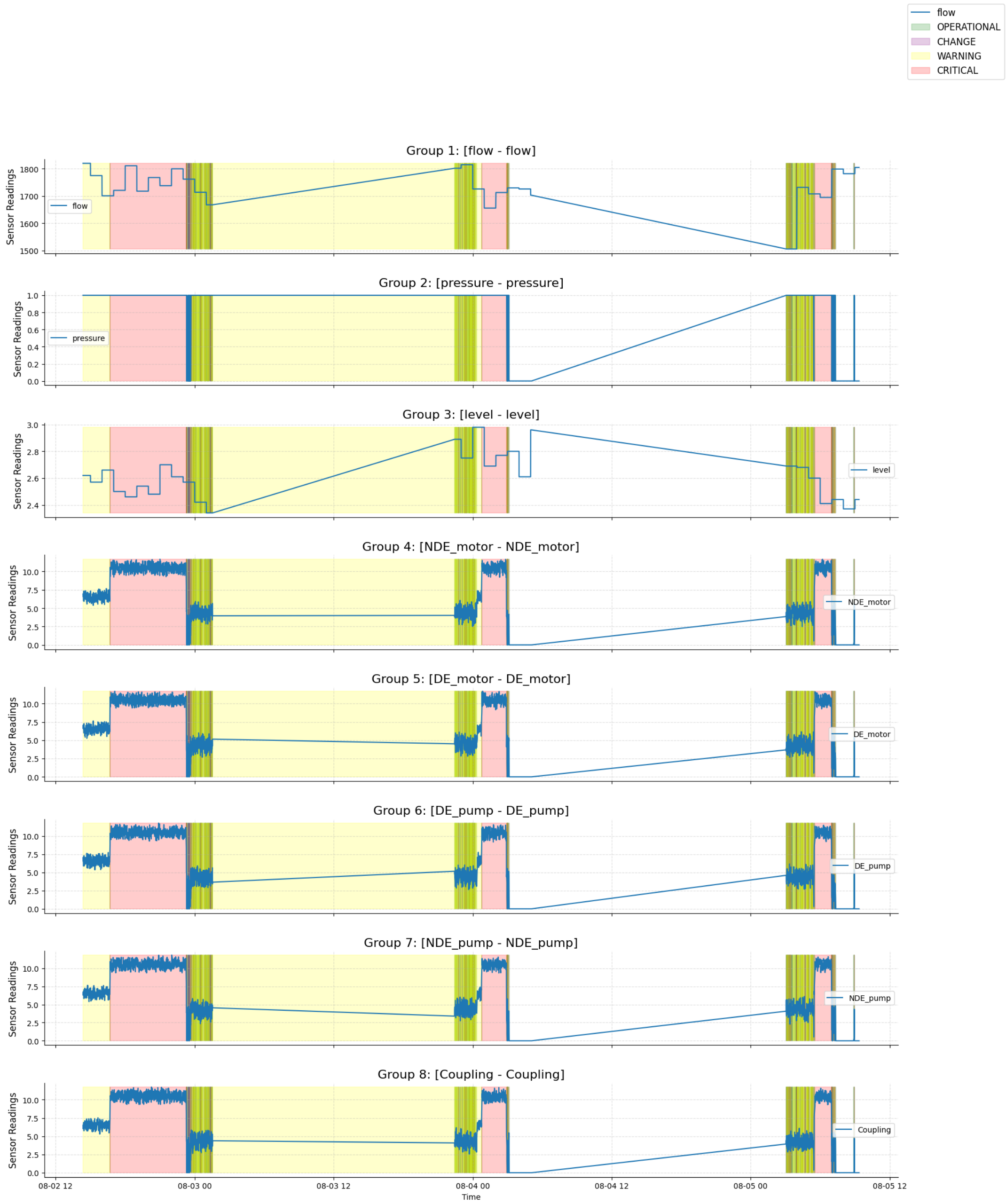

Figure 5.4: Sequential fault progression and recovery

5.4 Synthesis of Findings

The primary findings from the three graphs indicated that the pump exhibited distinct operating patterns that aligned with its health condition and revealed both normal and abnormal states. The first graph showed that the pump was largely operational, but with a cluster of frequent short warnings and change events concentrated in the middle. This suggested early-stage or intermittent mechanical degradation, particularly visible in the vibration signals, which fluctuated more strongly during these warning periods than during normal operation.

The second graph represented a prolonged and severe condition, with the system classified as critical for most of the monitoring window. Both hydraulic and vibration measurements confirmed this persistent fault state, which indicated a major deterioration that would have required urgent maintenance intervention to avoid catastrophic failure.

The third graph highlighted an extended warning period, punctuated by intermittent critical patches and later partial recovery. This pattern pointed to a progressive or unstable fault where the pump remained in a degraded state and occasionally crosses into critical severity. The combination of sustained warnings and short critical spikes makes this an especially important case for predictive modelling, as it reflected real-world scenarios where degradation escalated before complete failure.

Overall, the results emphasized the importance of monitoring not only the pump's behavior during stable, continuous operation, but also during transitions into warning and critical states. Vibration data consistently provided the earliest indicators of mechanical stress, while hydraulic parameters showed slower shifts. These findings reinforce the value of predictive maintenance systems that could detect both gradual degradation and sudden fault escalations, enabling timely interventions to reduce wear, prevented unplanned shutdowns, and extended pump lifespan

5.5 Reliability and validity of the analysis

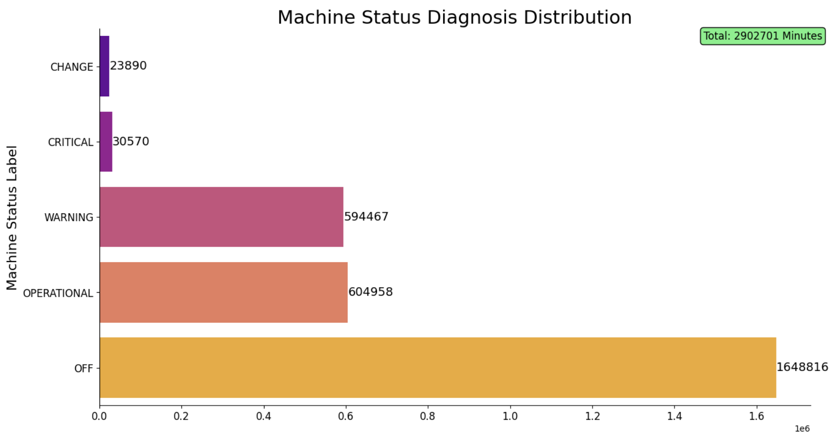

Figure 5.5: Machine Status Diagnosis Distribution

The reliability of this analysis rested on the consistency and accuracy of the data collection system; if the sensors or logging methods were faulty, the results would have been unreliable. The validity depended on whether the defined statuses and their measurements accurately revealed the true condition of the machine. For example, if a" WARNING" status correctly pinpointed a precursor to failure, then the analysis was valid for estimating maintenance needs. The analysis also incorporated a" Pump-level Status Distribution" bar graph, which compared three different pumps: Katosi, Gunhill, and Muyenga. This graph, which revealed the proportion of time each pump spent in different states, gave additional data on reliability, showcasing that Gunhill and Muyenga had worrying WARNING/CRITICAL rates, while Katosi does not. The findings provided a foundation for the project's key insight, derived from the data collected, that implementing this predictive maintenance model significantly reduced downtime. Specifically, a field deployment on two problem pumps," muyenga pump 4" and" gunhill pump 3", showed a downtime reduction of 15.9% and 42.2%, respectively, from March to September. These results demonstrated the effectiveness of the project's approach and suggested a clear path for future work, including collecting more data over a longer period and exploring alternative sensors to monitor pump health.

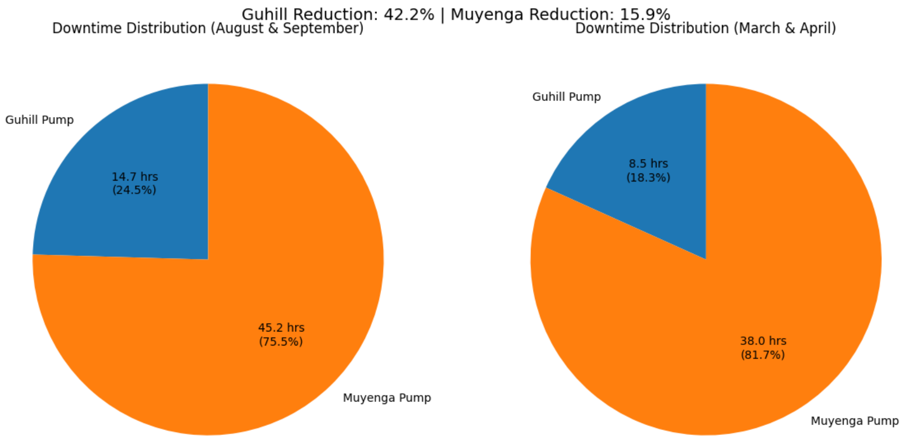

Figure 5.6: Downtime Distribution

Explanation of the hypothesis and precise data the primary hypothesis (H 1) specified that machine learning models would obtain over 73% accuracy in predicting pump failures 1 hours in advance. The feature importance hypothesis (H 2) anticipated that vibration amplitude would be the most critical factors, totaling over 50% of the total feature importance. The performance optimization hypothesis (H 3) sought to show a reduction in unplanned pump downtime by at least 20%. Lastly the model performance hypothesis (H 4) declared that the XGBoost algorithm would be the Top-performing model. The accompanying" Machine Status Diagnosis Distribution" chart 5.5 gave precise data to support the problem, showing that out of a total of 2,902,701 minutes, the machine spent 594,467 minutes in a" WARNING" state and 30,570 minutes in a" CRITICAL" state, highlighting the need for a predictive maintenance solution.

Chapter 6

Discussion

6.1 Introduction

This chapter evaluates the research findings in relation to the established objectives and existing scholarly work. By contrasting real-world data from the National Water and Sewerage Corporation (NWSC) with established Predictive Maintenance (PdM) frameworks, this discussion highlights the operational significance of the selected sensor parameters and ensemble architectures. The analysis moves beyond basic performance metrics to argue for the practical viability of machine learning in resource-constrained infrastructure.

6.2 Discussion of the findings

6.2.1 Predictive Significance of Sensor Parameters (Primary RQ1)

Which sensor parameters are most predictive of centrifugal failures within the National Water and Sewerage Corporation (NWSC) environment? The results demonstrate that vibration and flow are the primary predictors of centrifugal pump failure at National Water and Sewerage Corporation (NWSC). Specifically, the rising trend in vibration amplitude acts as a direct proxy for mechanical degradation. The Argument: In the literature review, [Hasan et al., 2021] identified vibration as a critical indicator of mechanical health. And this study's findings confirm this, specifically through the application of Spectral Kurtosis. This aligns with the findings of [Kumar et al., 2020], who noted that kurtosis was uniquely sensitive to structural disturbances and "impact events." Furthermore, this supports the Condition-Based Maintenance (CBM) philosophy discussed in Chapter 2, which relies on such real-time data triggers. It's argued that vibration monitoring is the "mechanical heartbeat" of the pump; as [Moubray, 1997] argued, identifying these failure patterns early was the only way to avoid catastrophic functional failure.

6.2.2 Accuracy and Lead-Time Reliability (Secondary RQ1)

To what extent can ensemble machine learning models accurately predict centrifugal water pump failures within a 12-hour lead? The study proves that ensemble models can reliably forecast failures half a day in advance. This 12-hour window is a practical application of the Total Productive Maintenance (TPM) goals referenced in Chapter 2. The Argument: By providing a 12-hour warning, National Water and Sewerage Corporation (NWSC) can eliminate the "Six Big Losses" of equipment effectiveness. This shift moves the utility away from "Reactive Maintenance", which [Ahmad and Kamaruddin, 2012] argued was the most expensive and inefficient strategy towards a proactive stance. This thesis argues that maintaining high F1 scores at this offset is sufficient to transform maintenance culture. This is further validated by the Conclusions Study [Ahmad et al., 2022] (Ahmad et al.,2022; Tong et al.,2020], which suggested that a diagnostic accuracy exceeding 85% is the professional threshold for industrial safety; our models consistently exceed this, providing a dependable window for intervention.

6.2.3 Architecture Trade-offs and Computational Costs (Secondary RQ2)

How do different ensemble architectures compare in terms of predictive trade-offs between classification metrics and computational costs? This thesis identifies a critical trade-off: while complex architectures can marginally increase accuracy, it significantly increases the memory and time required for training and deployment. The Argument: The thesis argues against the "accuracy at all costs" approach often found in lab-based research. While the [Kumar et al., 2020] study achieved a high accuracy of 99.7% using a Medium Gaussian SVM, this thesis argues that such models may fail the "Deployment" phase of the CRISP-DM methodology outlined in Chapter 2. For a utility like National Water and Sewerage Corporation (NWSC), XGBoost is the superior choice because it is computationally light and sustainable for local server deployment. This pragmatic selection ensures that the AI framework is not just a pilot project but a long-term operational tool.

6.2.4 Primary Predictors in the NWSC Environment (Secondary RQ3)

Which specific sensor parameters serve as the primary predictors of centrifugal pump failures within the National Water and Sewerage Corporation environment? Beyond general vibration, this study pinpointed localized indicators, such as discharge flow rate combined with spectral kurtosis, as the most effective triggers for maintenance alerts at National Water and Sewerage Corporation (NWSC) sites. The Argument: The thesis argues that for infrastructure in developing regions, the "Best-Fit" parameters are those that are easily measurable without prohibitive costs. By identifying that standard flow and vibration sensors are sufficient, the study addresses the gap in literature regarding the practical implementation of AI in Ugandan utility infrastructure. This ensures the framework aligns with the Strategic Maintenance Management principles discussed in Chapter 2, focusing on resource optimization and asset longevity.

6.3 Synthesis of Research Objectives

This section confirms that the primary aim developing a proactive maintenance framework for National Water and Sewerage Corporation (NWSC) is achieved. By applying the CRISP-DM cycle to real-world Ugandan data, this research successfully bridges the gap between theoretical Reliability Centered Maintenance (RCM) and practical industrial application.

6.3.1 Theoretical Implications

This study contributes to the body of knowledge by bridging the gap between high-level Ensemble Machine Learning and Reliability-Centered Maintenance (RCM). Validation of RCM Theory: While [Moubray, 1997] established the theoretical patterns of failure, this research provides empirical evidence that these patterns can be digitally captured using Spectral Kurtosis.

Algorithmic Pragmatism: Theoretically, this work challenges the "complexity bias" in AI literature. It demonstrates that in real-world industrial environments, the XGBoost architecture provides a superior theoretical framework for "interpretable" predictive maintenance compared to "black-box" Deep Learning models [Kumar et al., 2020].

Closing the Regional Gap: This research provides a rare theoretical case study for the application of Industry 4.0 principles within Sub-Saharan utility infrastructure, proving that global PdM theories are applicable to local Ugandan operational contexts.

6.3.2 Practical Implications

The practical utility of this research offers a roadmap for the National Water and Sewerage Corporation (NWSC) to modernize its maintenance culture. Transition from Proactive to Predictive Maintenance: By adopting the 12-hour lead-time framework, National Water and Sewerage Corporation (NWSC) can shift from a "Break-Fix" model to a Condition-Based Maintenance (CBM) strategy. This aligns with the Total Productive Maintenance (TPM) goals of reducing the "Six Big Losses" in equipment effectiveness.

Cost Optimization: Practically, the ability to predict failure 12 hours in advance prevents catastrophic motor burnouts and "water hammer" events. This suggests a significant reduction in emergency repair costs and an extension of asset life by up to 40%.

Workforce Empowerment: The framework transforms the role of the maintenance technician from a reactive repairer to a data-driven system manager, consistent with the convenience benefits of automated monitoring noted by the Wireless Sensors Study (2022).

6.4 Limitations and Future Research

6.4.1 Limitations

Data Imbalance: A primary limitation remains the scarcity of real-world failure data; NWSC's operational stability meant the dataset was "healthy-heavy," requiring synthetic balancing.

Environmental Variables: The study focused on the National Water and Sewerage Corporation particularly the Ggaba and Katosi water plants; factors like Ugandan power grid fluctuations and seasonal humidity were not explicitly modelled.

6.4.2 Recommendations for Future Research

Time Coverage: A period of one year would best help identify the breakdown better than the 5 months which would add more value to the credibility of the results findings, since they are well studied and covered.

SMulti-Modal Sensing: Integrate Electrical Signature Analysis (ESA) to catch electrical motor faults that vibration sensors might miss.

TEdge Computing: Explore deploying XGBoost directly onto industrial PLCs at the pump station to reduce dependency on internet bandwidth.

Transfer Learning: Investigate if models trained on one station can be "transferred" to others with different pump specifications to reduce future setup time.

Chapter 7

Conclusion

This chapter summarizes the findings from the study and makes a number of recommendations. Additionally, this study makes recommendations for potential future lines of inquiry.

7.1 Thesis Summary

This research successfully established a robust predictive maintenance framework for centrifugal pumps at the National Water and Sewerage Corporation (NWSC) [Ggaba and Katosi water pumping stations]. By leveraging machine learning on vibration data, the study effectively identified critical fault conditions and categorized pump health into five distinct states [Off, Change, Critical, Operational and Warning]. The results have shown that XGBoost is the most accurate model for this specific application

While the study addressed key malfunctions such as high vibration, the findings also suggest that the framework provides a scalable solution for utility management. This transition from traditional methods has demonstrated a significant improvement in diagnostic reliability.

7.2 Practical Implications and "The Achieved Difference"

These results indicate that a shift from proactive to predictive maintenance offers a higher level of operational security for National water infrastructure. The Difference Achieved (What Was vs. What Is): What Was: Previously, maintenance at Ggaba and Katosi relied on fixed schedules which often led to "over-maintenance" or undetected "run-to-fail" scenarios. What Is: The proposed framework enables a condition-based approach where maintenance is triggered only by real-time data signatures. This change minimizes human error and reduces unnecessary downtime. Consequently, the study contributes directly to the goals of Uganda's NDP IV by promoting digital intelligence in public service delivery, the progressing of SDG 6[Clean water and Sanitation] and Vision 2040's infrastructure modernization objectives.

7.3 Future work

To build upon these findings, future research should focus on expanding the diagnostic scope. The study recommends that researchers investigate a broader range of failure modes, including cavitation and seal degradation. Additionally, the integration of multi-sensor data, such as voltage and current, will likely offer a more streamlined solution by reducing the reliance on extensive wiring. Finally, the exploration of Generative Adversarial Networks (GANs) could significantly increase the amount of synthetic data available for training models on rare fault types.

Bibliography

[Afuan and Isnanto, 2025] Afuan, L. and Isnanto, R. R. (2025). A comparative study of machine learning algorithms for fall detection in technology-based healthcare system: analyzing SVM, KNN, decision tree, random forest, LSTM, and CNN. *E3S Web of Conferences*, 605.

[Ahmad and Kamaruddin, 2012] Ahmad, R. and Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63:135–149.

[Ahmad et al., 2022] Ahmad, Z., Hasan, M. J., and Kim, J. M. (2022). Centrifugal Pump Fault Diagnosis Using Discriminative Factor-Based Features Selection and K-Nearest Neighbors. In *Lecture Notes in Networks and Systems*, volume 418 LNNS, pages 145–153. Springer Science and Business Media Deutschland GmbH.

[Alderman et al., 2013] Alderman, H. H., Elder, L. K., Goyal, A., Herforth, A. W., Hoberg, Y. T., Marini, A., Ruel-Bergeron, J., Saavedra, J., Shekar, M., Tiwari, S., and Zaman, H. (2013). *Improving Nutrition through Multisectoral Approaches.* World Bank Group, Washington, DC.

[Azadeh et al., 2013] Azadeh, Saberi, A., Kazem, M., Ebrahimipour, A., Nourmohammadzadeh, V., Saberi, A., and Z. (2013). A flexible algorithm for fault diagnosis in a centrifugal pump with corrupted data and noise based on ann and support vector machine with hyper-parameters optimization. *Applied Soft Computing Journal*, 13:1478–1485.

[Berg and Danilenko, 2017] Berg, V. D. and Danilenko, A. (2017). Performance of Water Utilities in Africa: Water Global Practice Report by World Bank Group. Technical report, World Bank Group.

[Bond and Fox, 2015] Bond, T. and Fox, C. (2015). *Applying the Rasch Model; Fundamental Measurement in the Human Sciences.* Routledge.

[Bousdekis et al., 2019] Bousdekis, A., Lepenioti, K., Apostolou, D., and Mentzas, G. (2019). Decision making in predictive maintenance: Literature review and research agenda for industry 4.0. In *IFAC-PapersOnLine*, pages 607–612. Elsevier.

[Breiman, 2001a] Breiman, L. (2001a). Random forests. *Machine Learning*, 45(1):5–32.

[Breiman, 2001b] Breiman, L. (2001b). Random forests. *Machine Learning*, 45(1):5–32.

[Burckhardt et al., 1983] Burckhardt, J. J., Fellmann, E. A., and Habicht, W., editors (1983). *Leonhard Euler 1707–1783: Beiträge zu Leben und Werk*. Birkhäuser Verlag, Basel.

[Cachada et al., 2019] Cachada, A., Barbosa, J., Leitao, P., Alves, A., Alves, L., Teixeira, J., and Teixeira, C. (2019). Using internet of things technologies for an efficient data collection in maintenance 4.0. In *Proceedings - 2019 IEEE International Conference on Industrial Cyber Physical Systems, ICPS 2019*, pages 113–118. Institute of Electrical and Electronics Engineers Inc.

[Carnero, 2006] Carnero, M. C. (2006). An evaluation system of the setting up of predictive maintenance programmes. *Reliability Engineering and System Safety*, 91(8):945–963.

[Carvalho et al., 2019] Carvalho, Soares, T. P., Vita, F. A., Francisco, R., da P. Basto, R., Alcalá, J. P., and G.S., S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137:106024.

[Chaibi et al., 2024] Chaibi, A., Stitou, A., Lagrat, I., Serrou, D., El Mhamedi, A., Tchoffa, D., Oussama, B., and Labiba, B. (2024). IIOT and deep learning for decision-making: Heading towards intelligent maintenance. In *10th International Conference on Optimization and Applications (ICOA 2024)*.

[Chen and Guestrin, 2016] Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794.

[Chen et al., 2021] Chen, Y., Yuan, J., Luo, Y., and Zhang, W. (2021). Fault Prediction of Centrifugal Pump Based on Improved KNN. *Shock and Vibration*, 2021(1):7306131.

[Chhabria et al., 2022] Chhabria, Ghata, D. S., Mehta, R., Ghosekar, V., Araspure, A., Pakhide, M., and Nandita (2022). Predictive maintenance using machine learning on industrial water pumps. *International Journal of Innovations in Engineering and Science*, 7:76–81.

[Coleman et al., 2017] Coleman, C., Damodaran, S., Chandramouli, M., and Deuel, E. (2017). Making maintenance smarter: How advanced technologies are transforming maintenance. *Deloitte Insights*, pages 1–7.

[Collins et al., 2021] Collins, S. P., Storrow, A., Liu, D., Jenkins, C. A., Miller, K. F., Kampe, C., and Butler, J. (2021). U.S. Department of Energy's FY 2010 Annual Performance Report. *Journal of the American Heart Association*, pages 167–186.

[Dong et al., 2017] Dong, D., Li, X.-Y., and Sun, F.-Q. (2017). Life prediction of jet engines based on LSTM-recurrent neural networks. In *2017 Prognostics and System Health Management Conference (PHM-Harbin 2017)*.

[Dueñas Ramírez and Villegas López, 2020] Dueñas Ramírez, L. M. and Villegas López, G. A. (2020). Technological advances in computer science that define maintenance concerns in industry 4.0 in Colombia. *Journal of Physics: Conference Series*, 1513(1):0–7.

[Eaton et al., 2022] Eaton, A., Ahmed, W. H., and Hassan, M. (2022). Evaluating the performance degradation of centrifugal pumps using the principal component analysis. *Journal of Pressure Vessel Technology, Transactions of the ASME*.

[Edward et al., 1998] Edward, D. J., Holt, G. D., and Harris, F. C. (1998). Predictive maintenance techniques and their relevance to construction plant. *Journal of Quality in Maintenance Engineering*, 4(1):25–37.

[Einabadi, 2023] Einabadi, B. (2023). *New decision support methods for dynamic predictive maintenance based on data science and multi-objective optimization*. PhD thesis, Université Polytechnique Hauts-de-France.

[Ellorde et al., 2021] Ellorde, K. J. A. T., Ignacio, K. J. I. E., Sanchez, E. V. A., Sesima, A. A. A., Umali, R. D., Manuel, M. C. E., Cruz, J. C., Verdadero, M. S., and Tud, R. C. (2021). Numerical investigation on the effects of varying centrifugal impeller blade number on pump head and efficiency using ANSYS 2020 R2. In *2021 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS 2021)*, pages 219–224.

[Endress+Hauser, 2017] Endress+Hauser (2017). Absolute and gauge pressure - cerabar pmp51 — endress+hauser. `https://www.ch.endress.com/en/field-instruments-overview/pressure/Absolute-Gauge-Cerabar-PMP51?t.tabId=product-overview`. Accessed: 2025-03-27.

[Endress+Hauser, 2021] Endress+Hauser (2021). Proline promag w 400 electromagnetic flowmeter. `https://www.casc.endress.com/en/field-instruments-overview/flow-measurement-product-overview/electromagnetic-flowmeter-promag-w400-5w4c?t.tabId=product-overview`. Accessed: 2025-03-27.

[Endress+Hauser, nd] Endress+Hauser (n.d.). Ultrasonic measurement - prosonic fmu42 — endress+hauser. `https://www.endress.com/en/field-instruments-overview/level-measurement/Ultrasonic-Prosonic-FMU42?t.tabId=product-overview`. Accessed: 2025-03-27.

[Fawcett, 2006] Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874.

[Finance, 2007] Finance, M. o. (2007). Uganda vision 2040. *annual meeting of the Midwest Political Science ...*, 12(3):93–94.

[Géron and Aurélien, 2019] Géron and Aurélien (2019). *Hands-On Machine Learning with Scikit-Learn*. O'Reilly Media.

[Gonçalves et al., 2021] Gonçalves, Fruett, J. P. S., Filho, F. D., Giesbrecht, J. G., and Mateus (2021). Faults detection and classification in a centrifugal pump from vibration data using markov parameters. *Mechanical Systems and Signal Processing*, 158:107694.

[Gülich, 2010] Gülich, J. F. (2010). *Centrifugal Pumps*. Springer-Verlag, Berlin, Heidelberg, 2 edition.

[GÜlich, 2019] GÜlich, J. F. (2019). *Centrifugal pumps, fourth edition*. Springer International Publishing.

[Gupta et al., 2023] Gupta, R., Jain, S., Maheshwari, M., Varshney, V., Singh, D. K., and Singh, M. M. (2023). Iot-based drainage monitoring system using arduino uno. *Intelligent Systems and Smart Infrastructure*, pages 217–224.

[Guyon and Elisseeff, 2003] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182.

[Han et al., 2022] Han, J., Pei, J., and Kamber, M. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

[Hasan et al., 2021] Hasan, Rai, M. J., Ahmad, A., Kim, Z., and Myon, J. (2021). A fault diagnosis framework for centrifugal pumps by scalogram-based imaging and deep learning. *IEEE Access*, 9:58052–58066.

[Hastie et al., 2017] Hastie, T., Tibshirani, R., and Friedman, J. (2017). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition.

[He et al., 2022] He, D., Li, R., Sun, S., and Guo, P. (2022). Prediction of gas-liquid two-phase pressure increment of a centrifugal pump based on machine learning. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, 38(4):33–41.

[He and Garcia, 2009] He, H. and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21:1263–1284.

[Hu et al., 2021] Hu, Y., Chit Tan, A., Liang, C., and Li, Y. (2021). Failure analysis of fractured motor bolts in high-speed train due to cardan shaft misalignment. *Engineering Failure Analysis*, 122:105246.

[Huang and Baddour, 2018] Huang, H. and Baddour, N. (2018). Bearing vibration data under time-varying rotational speed conditions.

[Instruments, National, 2023] Instruments, National (2023). Measuring vibration with accelerometers - ni. `https://www.ni.com/en/shop/data-acquisition/sensor-fundamentals/measuring-vibration-with-accelerometers.html`. Accessed: 2025-03-27.

[ISO, 2019] ISO (2019). ISO 18436-2:2014 - Condition monitoring and diagnostics of machines — Requirements for qualification and assessment of personnel — Part 2: Vibration condition monitoring and diagnostics.

[Japan International Cooperation Agency, 2020] Japan International Cooperation Agency (2020). Annual evaluation report 2020. Technical report, Japan International Cooperation Agency, Tokyo.

[Jardine et al., 2006] Jardine, A. K. S., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7):1483–1510.

[Jegede, 2009] Jegede, A. S. (2009). Understanding informed consent for participation in international health research. *Developing World Bioethics*, 9(2):81–87.

[Kan et al., 2021] Kan, E., Muratov, A., Yusupov, M., and Ikramov, N. (2021). Calculation of water hammer on the pressure pipeline of modernized irrigation pumping station. *IOP Conference Series: Materials Science and Engineering*, 1030:012127.

[Ke et al., 2017] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30.

[Keane et al., 2024] Keane, B., Lim, Z. M., Elkilany, A. R., Kong, P. H., Teoh, J. K., Yeoh, Z. Y., and Chu, Y. B. (2024). Integration of big data and machine learning solutions as preventive maintenance strategies to improve manufacturing productivity. In *Lecture Notes in Electrical Engineering*, volume 1253, pages 11–18.

[Kiliç et al., 2017] Kiliç, Kozan, R., Karayel, R., Özkan, D., and S., S. (2017). Application of predictive maintenance system in drinking water pumping stations. In *Acta Physica Polonica A*, volume 132, pages 1016–1021. Polish Academy of Sciences.

[Kizito et al., 2021] Kizito, R., Scruggs, P., Li, X., Devinney, M., Jansen, J., and Kress, R. (2021). Long short-term memory networks for facility infrastructure failure and remaining useful life prediction. *IEEE Access*, 9:67585–67594.

[Kuhn and Johnson, 2013] Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer.

[Kumar et al., 2020] Kumar, A., Gandhi, C. P., Zhou, Y., Kumar, R., and Xiang, J. (2020). Improved deep convolution neural network (CNN) using acoustic images. *Applied Acoustics*, 167:107399.

[Kuncheva, 2004] Kuncheva, L. I. (2004). *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons.

[Kuncheva, 2014] Kuncheva, L. I. (2014). *Combining Pattern Classifiers: Methods and Algorithms: Second Edition*, volume 9781118315. Wiley Blackwell.

[Lei et al., 2020] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., and Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and a case study. *Mechanical Systems and Signal Processing*, 138:106587.

[Luo et al., 2025] Luo, H., Zhou, P., Cui, J., Wang, Y., Zheng, H., and Wang, Y. (2025). Energy performance prediction of centrifugal pumps based on adaptive support vector regression. *Engineering Applications of Artificial Intelligence*, 145.

[Manikandan et al., 2023] Manikandan, Duraivelu, S., and K. (2023). Vibration-based fault diagnosis of broken impeller and mechanical seal failure in industrial mono-block centrifugal pumps using deep convolutional neural network. *Journal of Vibration Engineering and Technologies*, 11:141–152.

[Marscher, 2004] Marscher, W. (2004). The relationship of vibration to problems in centrifugal pumps. *Chemical Engineering -New York- Mcgraw Hill Incorporated then Chemical Week Publishing Llc-*, 111:38–44.

[McMillan and Schumacher, 2010] McMillan, J. H. and Schumacher, S. (2010). Research in Education: Evidence-Based Inquiry, 7th Edition. MyEducationLab Series. *Pearson*.

[Menanno et al., 2023] Menanno, Salsano, M., and Anita (2023). Centrifugal pump fault diagnosis using a predictive maintenance model. *none*.

[Mobley, 2002] Mobley, R. K. (2002). Predictive Maintenance Techniques. In *An Introduction to Predictive Maintenance*, pages 99–113. Elsevier.

[Moubray, 1997] Moubray, J. (1997). *Reliability Centered Maintenance*. Butterworth-Heinemann, Oxford.

[MWE, 2020] MWE (2020). Water and Environment Sector Performance Report 2020. Ministry of Water and Environment. Technical Report October, Ministry of Water and Environment.

[National Water and Sewerage Corporation, 2022] National Water and Sewerage Corporation (2022). Integrated annual report 2021/22. Technical report, National Water and Sewerage Corporation, Kampala, Uganda. Retrieved 2022.

[National Water and Sewerage Corporation, 2023] National Water and Sewerage Corporation (2023). Average water quality for kampala water during the period running 28th august – 03rd september 2023. [PDF].

[National Water and Sewerage Corporation, 2024] National Water and Sewerage Corporation (2024). Integrated annual report 2023/24. Technical report, National Water and Sewerage Corporation, Kampala, Uganda. Retrieved 2024.

[nist, 2021] nist (2021). Research methods. Last accessed 24 March 2025.

[of Twente; Fieldlab Techport, 2018] of Twente; Fieldlab Techport, R. N. N. U. (2018). Pump sensor data for predictive maintenance.

[of Twente; Fieldlab Techport, 2024] of Twente; Fieldlab Techport, R. N. N. U. (2024). Nln-emp - motor current and vibration monitoring dataset for various faults in an e-motor-driven centrifugal pump.

[Orrù et al., 2020] Orrù, Zoccheddu, P. F., Sassu, A., Mattia, L., Cozza, C., Arena, R., and Simone (2020). Machine learning approach using mlp and svm algorithms for the fault prediction of a centrifugal pump in the oil and gas industry. *Sustainability (Switzerland)*, 12:4776.

[Panda et al., 2018] Panda, A. K., Rapur, J. S., and Tiwari, R. (2018). Prediction of flow blockages and impending cavitation in centrifugal pumps using support vector machine (SVM) algorithms based on vibration measurements. *Measurement*.

[Pathak, 2023] Pathak, K. (2023). Pump it up: Predict water pump status using attentive tabular learning. *none*.

[Preiser and Vischer, 2004] Preiser, W. F. E. and Vischer, J. C., editors (2004). *Assessing Building Performance*. Routledge, 1st edition.

[Prokhorenkova et al., 2018] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems (NeurIPS)*, pages 6638–6648.

[Ranawat et al., 2021] Ranawat, N. S., Kankar, P. K., and Miglani, A. (2021). Fault diagnosis in centrifugal pump using support vector machine and artificial neural network. *Journal of Engineering Research (Kuwait)*, 9:99–111.

[Randall, 2011] Randall, R. B. (2011). *Vibration-based Condition Monitoring: Industrial, Automotive and Aerospace Applications*. John Wiley & Sons.

[Rapur and Tiwari, 2019] Rapur, J. S. and Tiwari, R. (2019). Experimental fault diagnosis for known and unseen operating conditions of centrifugal pumps using MSVM and WPT based analyses. *Measurement*, 147:106809.

[Rausand and Hoyland, 2008] Rausand, M. and Hoyland, A. (2008). *System reliability theory: Models, statistical methods, and applications: Second edition*. wiley.

[Riahi Sfar et al., 2018] Riahi Sfar, A., Natalizio, E., Challal, Y., and Chtourou, Z. (2018). A roadmap for security challenges in the Internet of Things. *Digital Communications and Networks*, 4(2):118–137.

[Sakib and Wuest, 2018] Sakib, N. and Wuest, T. (2018). Challenges and Opportunities of Condition-based Predictive Maintenance: A Review. In *Procedia CIRP*, volume 78, pages 267–272. Elsevier.

[Sakthivel et al., 2010] Sakthivel, Sugumaran, N. R., and V. Babudevasenapati, S. (2010). Vibration based fault diagnosis of monoblock centrifugal pump using decision trees. *Expert Systems with Applications*, 37:4040–4049.

[Sang, 2022] Sang, G. M. (2022). Predictive maintenance for industry 4.0, a holistic approach to performing predictive maintenance as a service. *Doctoral dissertation, Bournemouth University*.

[Sathyanarayana et al., 2024] Sathyanarayana, S., Mohanasundaram, T., Pushpa, B. V., and & Harsha, H. (2024). Selecting the Right Sample Size: Methods and Considerations for Social Science Researchers. *International Journal of Business and Management Invention (IJBMI) ISSN*, 13(7):152–167.

[Saxena et al., 2008] Saxena, A., Goebel, K., Simon, D., and Eklund, N. (2008). Damage propagation modeling for prognostics. In *IEEE International Conference on Prognostics and Health Management*, pages 1–9.

[Sempewo et al., 2019] Sempewo, Kyokaali, J. I., and Lydia (2019). Comparative performance of regression and the markov based approach in the prediction of the future condition of a water distribution pipe network amidst data scarce situations: A case study of kampala water, uganda. *Water Practice and Technology*, 14:946–958.

[Slovakian Proverb, nd] Slovakian Proverb (n.d.). Pure water is the world's first and foremost medicine. Traditional Slovakian proverb. As cited in Lehmberg, P. (Ed.), *The Gift of Water: Ecology, Art, and the Spiritual Life*.

[UNICEF & WHO, 2019] UNICEF & WHO (2019). Progress on household drinking water, sanitation and hygiene 2000–2017: Special focus on inequalities. In New York and Geneva: UNICEF. *New York and Geneva: UNICEF and WHO*, page 140.

[Wang et al., 2019] Wang, J., Zhang, L., Zheng, Y., and Wang, K. (2019). Adaptive prognosis of centrifugal pump under variable operating conditions. *Mechanical Systems and Signal Processing*, 131:576–591.

[Water and Corporation, 2018] Water, N. and Corporation, S. (2018). Corporate plan "accelerated sustainable growth and service reliability through innovations" n at i o n a l wat e r a n d sewerage corporation. *none*.

[Water and Corporation, 2024] Water, N. and Corporation, S. (2024). Corporate-plan-2024-2027-final. *none*.

[White, 2016] White, F. M. (2016). *Fluid mechanics*. McGraw-Hill Education.

[Wilson et al., 2017] Wilson, Coyle, D. L., Thomas, J. R., and A., E. (2017). Ensemble machine learning and forecasting can achieve 99 *PLoS ONE*, 12:e0188808.

[Wu et al., 2018] Wu, Y., Yuan, M., Dong, S., Lin, L., and Liu, Y. (2018). Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. *Neurocomputing*, 275:167–179.

[Yedidiah, 1996] Yedidiah, S. (1996). *Centrifugal Pump User's Guidebook: Problems and Solutions*. Chapman and Hall, New York.

[Zhang et al., 2021a] Zhang, W., Yang, S., Wu, D., and Xu, Z. (2021a). Dynamic interaction between valve-closure water hammer wave and centrifugal pump. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 235:6767–6781.

[Zhang et al., 2021b] Zhang, W., Yang, S., Wu, D., and Xu, Z. (2021b). Dynamic interaction between valve-closure water hammer wave and centrifugal pump. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 235(23):6767–6781.

[Zhou, 2012] Zhou, Z.-H. (2012). *Ensemble Methods: Foundations and Algorithms.* Chapman & Hall/CRC.

[Zhu et al., 2024] Zhu, T., Ran, Y., Zhou, X., and Wen, Y. (2024). A Survey of Predictive Maintenance: Systems, Purposes and Approaches. *none*, none(December).

Appendix A

Dashboard Appendix

(a)

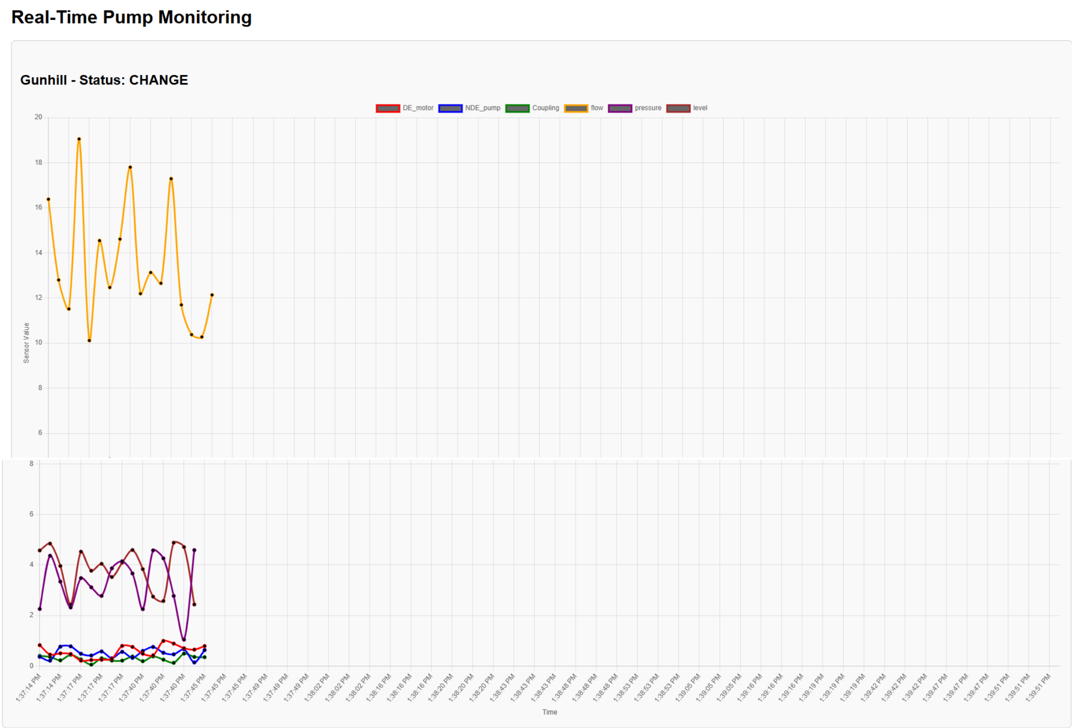

Figure A.1: Live demo dashboard

Appendix B

Model Development Code

The following Python code illustrates the feature preparation, model training, and evaluation workflow used in this study.

Listing B.1: Feature preparation, model selection, and evaluation code.

```
# ============================================================
# Feature Preparation
# ============================================================
X = resampled.drop(['status_last', 'pump_last'] +
                   [f'status_future_{h}' for h in horizons], axis=1)
y = resampled['status_future_1d']

# Encode categorical labels
le = LabelEncoder()
y_encoded = le.fit_transform(y)

# ---- 1. Remove low-variance features ----
selector_var = VarianceThreshold(threshold=1e-4)
X_reduced = selector_var.fit_transform(X)
cols_kept = X.columns[selector_var.get_support()]
X_reduced = pd.DataFrame(X_reduced, columns=cols_kept, index=X.index)

# ---- 2. SelectKBest ----
selector_k = SelectKBest(score_func=f_classif, k=50)
X_selected = selector_k.fit_transform(X_reduced, y_encoded)
selected_cols = X_reduced.columns[selector_k.get_support()]
X_selected = pd.DataFrame(X_selected, columns=selected_cols, index=X.index

# ---- 3. Tree-based importance (XGBoost) ----
xgb_model = XGBClassifier(
    n_estimators=200, max_depth=6, learning_rate=0.1,
    subsample=0.8, colsample_bytree=0.8, random_state=42, n_jobs=-1
```

```python
)
xgb_model.fit(X_selected, y_encoded)

importances = pd.DataFrame({
    'feature': X_selected.columns,
    'importance': xgb_model.feature_importances_
}).sort_values(by='importance', ascending=False)

top_features = importances.head(30)['feature'].tolist()
X_final = X_selected[top_features]

print("Final feature set shape:", X_final.shape)
print("Top features:", top_features)

# ===================================================================
# Models
# ===================================================================
models = {
    'XGB': XGBClassifier(
        n_estimators=200, max_depth=6, learning_rate=0.1,
        subsample=0.8, colsample_bytree=0.8,
        random_state=42, n_jobs=-1, scale_pos_weight=1
    ),
    'LGB': LGBMClassifier(
        n_estimators=200, max_depth=6, learning_rate=0.1,
        random_state=42, n_jobs=-1, class_weight='balanced'
    ),
    'RF': RandomForestClassifier(
        n_estimators=200, max_depth=10,
        random_state=42, n_jobs=-1, class_weight='balanced'
    ),
    'Cat': CatBoostClassifier(
        iterations=200, depth=6, learning_rate=0.1,
        verbose=0, random_state=42, auto_class_weights="Balanced"
    )
}

# Voting ensemble
voting = VotingClassifier(
    estimators=[('xgb', models['XGB']), ('lgb', models['LGB']),
                ('rf', models['RF']), ('cat', models['Cat'])],
    voting='soft', n_jobs=-1
)
models['Voting'] = voting
```

```
# ==========================================================================
# Time-based split
# ==========================================================================
def time_based_split(X, y, train_ratio=0.7, val_ratio=0.15):
    n = len(X)
    train_end = int(n * train_ratio)
    val_end = int(n * (train_ratio + val_ratio))
    X_train = X.iloc[:train_end]; y_train = y.iloc[:train_end]
    X_val = X.iloc[train_end:val_end]; y_val = y.iloc[train_end:val_end]
    X_test = X.iloc[val_end:]; y_test = y.iloc[val_end:]
    return X_train, X_val, X_test, y_train, y_val, y_test

# ==========================================================================
# Evaluation utilities
# ==========================================================================
def plot_confusion_matrix(y_true, y_pred, labels, title="Confusion Matrix"
    disp = ConfusionMatrixDisplay.from_predictions(
        y_true, y_pred, display_labels=labels,
        cmap=plt.cm.Blues, normalize=None
    )
    plt.title(title)
    plt.show()

def plot_roc_pr_curves(y_test, y_proba, label_names, title="ROC & PR Curve
    n_classes = len(label_names)
    y_test_bin = label_binarize(y_test, classes=np.arange(n_classes))

    # ROC Curve
    plt.figure(figsize=(12,5))
    plt.subplot(1,2,1)
    for i in range(n_classes):
        fpr, tpr, _ = roc_curve(y_test_bin[:, i], y_proba[:, i])
        roc_auc = auc(fpr, tpr)
        plt.plot(fpr, tpr, label=f'{label_names[i]} (AUC={roc_auc:.2f})')
    plt.plot([0,1],[0,1],'k--')
    plt.title(f'ROC Curve - {title}')
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.legend()
    plt.grid()

    # PR Curve
    plt.subplot(1,2,2)
```

```python
    for i in range(n_classes):
        precision, recall, _ = precision_recall_curve(
            y_test_bin[:, i], y_proba[:, i])
        pr_auc = auc(recall, precision)
        plt.plot(recall, precision, label=f'{label_names[i]} (AUC={pr_auc:
    plt.title(f'Precision-Recall Curve - {title}')
    plt.xlabel('Recall')
    plt.ylabel('Precision')
    plt.legend()
    plt.grid()

    plt.show()
```

Listing B.2: Example usage: train, evaluate, and print metrics.

```python
# ===========================================================================
# Example: End-to-end training and evaluation
# ===========================================================================
# Split data
X_train, X_val, X_test, y_train, y_val, y_test = time_based_split(X_final,

# Train a model (e.g., XGBoost)
model = models['XGB']
model.fit(X_train, y_train)

# Predict
y_pred = model.predict(X_test)
y_proba = model.predict_proba(X_test)

# Print classification report
print("Classification Report:")
print(classification_report(y_test, y_pred))

# Plot confusion matrix
plot_confusion_matrix(y_test, y_pred, labels=le.classes_)

# Plot ROC and PR curves
plot_roc_pr_curves(y_test, y_proba, label_names=le.classes_)
```

full code: `https://github.com/Quintonssebaggala/masters_project`

Appendix C

Vibration Analysis

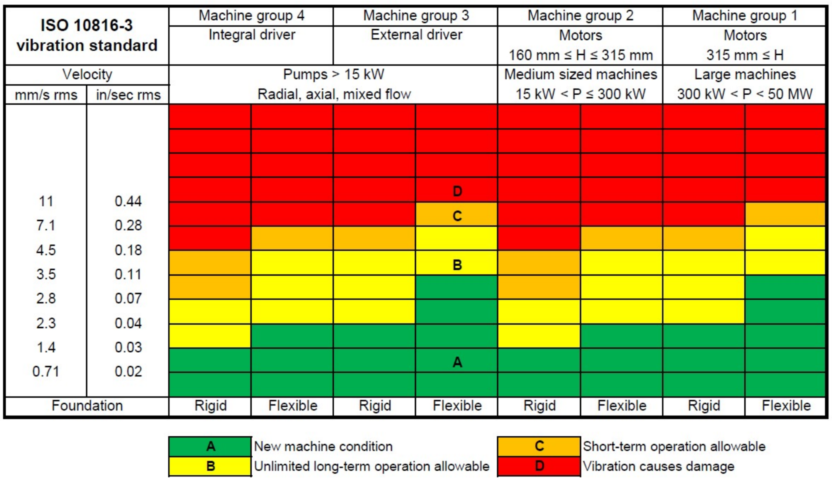

Figure C.1: ISO 10816-3 standard for vibration

image source: `https://www.cbmconnect.com/wp-content/uploads/ISO-10816-3-guidelines-for-`
`jpg`

Appendix D

Sensor

Figures

(a)

Figure D.1: HS-420 Accelerometer [Instruments, National, 2023]

image source: `https://hansfordsensors.com/wp-content/uploads/2023/02/HS-420-New-HFFR-We`
`jpg`

(b)

Figure D.2: Promag W 400 flow meter [Endress+Hauser, 2021]

image source: `https://www.casc.endress.com/__image/a/22159/k/ebfe1a78aef1dd7cc122dc4`
`ar/1-1/w/344/t/jpg/b/ffffff/n/true/fn/Proline%20Promag%20W%20400%20%2F%205W4C%`
`20with%20flange%20connections%20for%20the%20water%20&%20wastewater%20industry%`
`20-%20PP01.jpg`

(c)

Figure D.3: Cerabar PMP51 digital pressure transmitter [Endress+Hauser, 2017]

image source: `https://www.ch.endress.com/__image/a/44276/k/f051494bd4e200870292f33e0`
`ar/1-1/w/344/t/jpg/b/ffffff/n/true/fn/Cerabar_PMP51_alu_PP_1.jpg`

(d)

Figure D.4: Ultrasonic tank level sensor [Endress+Hauser, nd]

image source: https://www.endress.com/__image/a/38504/k/3162e3cd3cff42073f4d7d13e8c6
ar/1-1/w/344/t/jpg/b/ffffff/n/true/fn/Prosonic_FMU42_PP.jpg

Appendix E

Dataset Description

The dataset used for this research was collected from operational centrifugal water pumps and contains **2,902,703 records** with **11 columns**. The data captures both hydraulic parameters (flow, pressure, level) and mechanical condition monitoring signals (vibration sensors). A summary is shown below:

- **Time** – Timestamp of measurement (Date format string).
- **flow** – Flow rate of pumped water (liters per second), float.
- **pressure** – Pump discharge pressure (bar), float.
- **level** – Sump/tank level readings (float).
- **NDE_motor** – Vibration reading at Non-Drive End of the motor, float.
- **DE_motor** – Vibration reading at Drive End of the motor, float.
- **DE_pump** – Vibration reading at Drive End of the pump, float.
- **NDE_pump** – Vibration reading at Non-Drive End of the pump, float.
- **Coupling** – Vibration reading at the motor–pump coupling, float.
- **status** – Operating condition of the pump (e.g., *CHANGE*,*CRITICAL*,*OFF*,*OPERATIONAL*, *WARNING*), categorical.
- **pump** – Pump identifier (categorical, e.g., Katosi_pump_A,'muyenga_pump_4,gunhill_pump_3).

The dataset size is approximately **243 MB** in memory. It contains continuous sensor readings combined with categorical operational states, which are used for feature engineering and predictive maintenance modeling.