SCHOOL OF COMPUTING, TECHNOLOGY AND APPLIED SCIENCES

MASTER RESEARCH PROJECT

A MACHINE LEARNING APPROACH FOR PREDICTIVE MAINTENANCE AND RESOURCE ALLOCATION AT THE UNIVERSITY OF ZAMBIA

FLOYD KATEPUKA SHEKE

ZCAS UNIVERSITY

2024

A MACHINE LEARNING APPROACH FOR PREDICTIVE MAINTENANCE AND RESOURCE ALLOCATION AT THE UNIVERSITY OF ZAMBIA

FLOYD KATEPUKA SHEKE

A Final Year Research Project submitted in partial fulfilment of the requirements for the degree of

Master of Science in Information Technology

ZCAS University

2024

DECLARATION

Name: Floyd Katepuka Sheke

Student Number: BSC 11026

I hereby declare that this final year research project is the result of my own work, except for quotations and summaries which have been duly acknowledged.

Plagiarism check: %

Signature: F.K.Sheke Date: 23/11/2024

Supervisor Name: Dr. Bob Jere

Supervisor Signature:

Date: 23/11/2024

A MACHINE LEARNING APPROACH FOR PREDICTIVE MAINTENANCE AND RESOURCE ALLOCATION AT THE UNIVERSITY OF ZAMBIA

ABSTRACT

Predictive Maintenance essentially involves predicting breakdown of the equipment to be maintained by detecting early signs of failure in order to make maintenance work more proactive. The application of such maintenance strategy requires the cooperation of several agents and involves knowledge and skills, it requires processing and computing of relevant data signals in the from the equipment in order to extract meaningful knowledge. Preventive Maintenance in a broad topic making it impossible to cover all the subtopics in this research paper. Having this into consideration, this paper focuses on the Predictive Maintenance Model (PMM) for an Office Air Conditioner (OAC) using machine learning techniques such as XGBoost and LSTM networks. This research paper examines the advantages, limitations and advancements in predictive maintenance enabling to minimize downtime, energy consumption, cost savings, and increasing comfort in office workplace. One of the most useful features of the paper is the contribution to the body of knowledge by delivering a UI tool for facilities managers to make decisions quickly using machine learning predictive modelling techniques.In addition, the paper also highlights future work opportunities such as the inclusion of external data sources, development of adaptive models, enhanced IoT integration, and improved hyperparameter optimization methods.

Keywords: Predictive Maintenance, Equipment Breakdown Prediction, Early Failure Detection, Maintenance Proactivity,Machine Learning Techniques,Predictive Maintenance Model,Office Air Conditioner XGBoost, Long Short-Term Memory Networks,Downtime Minimization, Energy Consumption Reduction, Workplace Comfort, Facilities Management, Predictive Modelling,User Interface (UI) Tool, External Data Integration, Adaptive Models, IoT Integration, Hyperparameter Optimization, Advancements in Predictive Maintenance, Maintenance Strategy Cooperation, Knowledge Extraction

ACKNOWLEDGEMENT

I would like to take this opportunity to express my gratitude and appreciation to my supervisor, Dr. Bob Jere guidance, patience and invaluable advice throughout this project.

I also would like to express my appreciation to wife Vast Tembo Sheke and children name Floyd Sheke, Lubuto Sheke, Moses Sheke and Deborah Sheke for their support and understand the time I have been committed in executing this research work.

THANK YOU.

DEDICATION

I would like to take this opportunity to express my gratitude and appreciation to my family, friends and colleagues for the support and understanding.

TABLE OF CONTENTS

TITLE PAGE		
DECLARATION		ii
ABSTRACT		iii
ACKNOWLEDGEMENTS		iv
DEDICATION		v
TABLE OF CONTENTS		vi
LIST OF TABLES		vii
LIST OF FIGURES		viii
LIST OF ABBREVIATIONS		ix
CHAPTER 1	**INTRODUCTION**	
	1.1 Background to the Study	13
	1.2 Problem Statement	13
	1.3 Aim	14
	1.4 Objectives of the Study	14
	1.5 Scope and Limitation	15
	1.6 Significant of the Research	16
CHAPTER 2	**LITERATURE REVIEW**	
	2.0 Introduction	18
	2.1 Data sources	18
	2.2 Machine Learning Models for PMM	18
	2.3 Critical review of related works	20
	2.4 Data Quality and Availability	23
	2.5 Cost-Benefit Considerations	23
	2.6 Future Directions and Recommendations	23
	2.7 Conceptual Framework for Predictive Maintenance	24

2.8 Proposed Model for Predictive Maintenance 25
2.9 Data Collection Layer 26
2.10 Data Preprocessing Layer 27
2.11Predictive Analytics and Decision-Making Layer 28
2.12 Workflow 29
2.13 Comparison with Related Research Works 30
2.14 Summary 34

CHAPTER 3 METHODOLOGY

3.0 Introduction 35
3.1 Research Design 35
3.2 Adopted Method and Justification 35
3.3 Association of Research Method to the Project 36
3.4 Research Data and Datasets 36
3.5 Data collection methods 36
3.6 Machine Learning Models 38
3.7Ethical concerns related to the research 40
3.8 Summary

x

CHAPTER 4 DATA, EXPERIMENTS, AND IMPLEMENTATION

4.0 Introduction 41
4.1 Data 41
4.2 Prediction Experiment 42
4.3 Implementation 43
4.4 Summary 62

CHAPTER 5 RESULTS AND DISCUSSIONS

5.0 Introduction 63
5.1 Enhanced Operational Efficiency 63
5.2 Comparison with Related Research 64

5.3Broader Discussion and Future Directions 65

5.4 Summary 66

CHAPTER 6 SUMMARY AND CONCLUSION

6.0 Introduction 67

6.1 Summary of main Findings 67

6.2 Contribution to the body of knowledge 69

6.3 Limitations of the Research 70

6.4 Future works 72

6.5 Summary 73

REFERENCES **74**

LIST OF TABLES

Table Number	Title	Page
Table 1	Cleaned data sample	46
Table 2	Feature explanation	46
Table 3	Confusion matrix	50

LIST OF FIGURES

Figure Number	Title	Page
Figure 1	UNZA OAC Predictive Maintenance Dashboard	52
Figure 2	PMS Success response	58
Figure 3	PMS error response	59
Figure 4	SIS success response	62
Figure 5	SIS error response	62

LIST OF ABBREVIATIONS

API	Application Programming Interface
UNZA	University of Zambia
RE	Resident Engineer
OJCS	Online Job Card System
OAC	Office Air Conditioners
OACS	Office Air Conditioning Systems
OACU	Office Air Conditioning Units
OACM	Office Air Conditioning Model (OSCM)
SIS	Students Information System
PMS	Performance Management System
PMM	Predictive Maintenance Model
MLM	Machine Learning Models
ANN	Artificial Neural Networks
KNN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
RUL	Remaining Useful Life
SVR	Support Vector Regression
ARIMA	Auto Regressive Integrated Moving Average
SVM	Support Vector Machines
PCA	Principal Component Analysis
CNN	Convolutional Neural Networks
ML	Machine Learning
MQTT	Message Queuing Telemetry Transport
HTTP	Hypertext Transfer Protocol
UI	User Interface
IoT	Internet of Things
HVAC	Heating, Ventilation, and Air Conditioning

RF	Random Forest
GB	Gradient Boosting
DCM	Data Collection Methods
DAT	Data Analysis Techniques
PME	Predictive Maintenance Experiment
EDA	Exploratory Data Analysis
WPAC	WestPoint Office Air Conditioner
TP	True Positive
TN	True Negative
FN	False Negative
FP	False Positive
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
ROI	Return on Investment

CHAPTER 1

INTRODUCTION

1.1 Background to the study

The University of Zambia which is the largest and the oldest university in Zambia was established in the year 1965 and was opened to the public on 12th of July in the year 1966. The university is an academic powerhouse with numerous Schools and Units contributing to its diverse educational landscape. Out of these units, the Resident Engineer (RE) working in the Registrars Department has the responsibility of supervising construction and maintenance of infrastructure. This unit has recently enhanced the performance of its infrastructure management function by adopting the use of a new maintenance management system called the Job Card System. This system serves as a data bank for all the job requests emanating from the University community. The system provides numerous benefits such as the ability to report maintenance through the use of Online Job Card System (OJCS), to view information about the personnel that work on each of the maintenance job card requests and to monitor the progress of such maintenance card requests among other features.

1.2.0 Problem Statement

The Online Job Card System (OJCS) in the University of Zambia is an important system used in the management of infrastructure and services. However, like any other complex system, it has its own weaknesses that restrict it from delivering on its best results. This paper explores the issues that are associated with the Account Creation, Database Updates and System Functions of the Online Job Card System. Also, it acknowledge the need for a systematic approach that is in conformity with the university policies and the benefits of predictive maintenance modelling.

1.2.1 Monotonous Account Creation and Database Updates

The OJCS depends on the status of the people within the University especially the students and the staff to create accounts and use the system. A major challenge is the stochastic nature of the academic environment that we find ourselves in. This implies that the students complete their courses and then either proceed to other universities or join the employment market or get suspended or expelled from the University for one reason or the other. In like manner, the staff members exit through either retirement, resignation or transfer to other departments within or outside the institution. To this end, the system has put in place a mechanism of updating the database that house the information about the staff and students. Regrettably, this process is mainly done manually and hence a mistake prone exercise. This makes the database to be inaccurate as it may contain information of people who are no longer part of the university and those who are still members. This gap is a critical one since the system's functioning is based on the status of individuals; if these are wrong, then people who should not be allowed to access the system may do so.

1.2.2 Absence of an effective decision making interface

It is also important to note that one of the biggest shortcomings of the OJCS is the lack of a proper decision making interface. A good decision making interface can greatly improve decision making, increase the efficiency of the organization and improve the flow of information within the university. It is a thing of joy to see that, with a proper decision making interface, the administrators and users can be able to view the status of job cards, the progress made and where the resources are most required. Such a tool can also enable effective and clear communication as well as collaboration between different departments in order to ensure that maintenance tasks are accomplished effectively and within the shortest time possible.

1.2.3 Reactive Maintenance

The OJCS is efficient in managing the maintenance job card, it works in a very reactive manner. It only acts and processes the job requests as they are brought to their attention, which means that job cards maintenance tasks are handled on first come first served basis. Although it mainly provides a solution to a problem when it arises, it assists in the management of workflow by categorizing, measuring, and sequencing job cards. This allows the maintenance teams to work efficiently and make sure that maintenance requests are handled as soon as possible.

1.3 Problem Solution

However, as is the case with many educational institutions across the globe, the University of Zambia grapples with a number of challenges in the management of its facilities and the effective utilization of resources. Predictive maintenance is henceforth identified as a critical approach within this context based on the data analytic, sensor technology and the computing power of the machine learning (Zhao et al., 2020). It is a preventive and proactive strategy for the repair and maintenance of equipment and structures and it goes a step farther to help minimize downtime and increase the life expectancy of machinery and structures (Ansari et al., 2018). Machine learning is one of the most flexible set of tools which is not limited to the predictive maintenance, the optimization of resources and reduction of the operational costs and increasing efficiency (Claus et al., 2019; Chen et al., 2018; Zhu et al., 2019; Wang et al., 2020). At the University of Zambia, these two concepts are set to be integrated to form a higher technology system that is able to predict maintenance requirements and the best allocation of resources for the infrastructure across the whole university. This change from the conventional reactive mode of approach through to a more proactive approach will ensure that there is sustainability, cost effectiveness and increased efficiency (Claus et al., 2019; Chen et al., 2018; Zhu et al., 2019; Wang et al., 2020).

The integration of predictive maintenance and the application of machine learning in the resource allocation at the University of Zambia is an exciting concept in infrastructure management. By incorporating these two strategies to form a single system, the University will be able to have a better way of managing the infrastructure through the use of the system. It will not only help in predicting the maintenance requirements, but will also help in effective and proper allocation of resources so that the right tools and materials are available at the right time and place where the maintenance teams are working. It will also integrate the enhanced system that will afford the University an opportunity to make better decisions on the resource allocation and maintenance through real time analysis. It will increase efficiency, reduce costs, and improve the management of infrastructure of the university.

1.3 Aim

The main aim of this research is to address critical issues and implement improvements in the University of Zambia's OJCS. The research focuses on the enhancement the system's functionality by addressing challenges related to account creation, database updates, and overall system efficiency. Furthermore, it aims to introduce a predictive maintenance model to shift from reactive to proactive maintenance practices.

1.4.0 Objectives of the study

The specific objectives are as follows:

1.4.1 To carry out a thorough situation analysis

To conduct conduct a thorough situation analysis in order to fully understand the current state of the OJCS. This will involve examining the system's features in detail and identifying any gaps or areas that require improvement.

1.4.2 To implement an efficient interface system for decision making

To implement a dashboard within the system to provide users with a visual and centralized platform for data monitoring, analysis, and action.

1.4.3 To implement two APIs for data validation in the OJCS

One API for staff members, integrated with the Performance Management System, and another for students, linked to the Student Information System. These APIs will serve the purpose of data validation during both the account creation process and when submitting requests for maintenance job cards.

1.4.4 To implement a machine learning based Job card system

This will greatly improve the efficiency and effectiveness of maintenance operations for Office Air Conditioners (OAC) at the University.

1.5.0 Scope and Limitation

The scope and limitation for this research study are as listed as follows:

1.5.1 Scope

The scope of this research project will be to create a predictive model for Office Air conditioners within and the its university application in predicting failures for preventive maintenance. Also, the project will also involve the creation two of APIs for account creation and authentication purposes. The last but not the least, dashboard an will easy be to designed use to display a single platform for data observation, assessment, and control whereby the will users be provided with a graphic representation of the system's performance and maintenance requirements.

1.5.2.0 Limitation

1.5.2.1 Inadequate Historical Data

The university has insufficient historical information on the performance of the air conditioners, their failures and the maintenance activities carried out on them, a limitation in a developing major reliable predictive maintenance models. Data in general may be limited, and may be fragmented and incomplete, which further hinders the ability to makAnomaly Detection Models are computer programs that are used to detect any data e reliable forecasts. The lack of data or data the gaps presence can lead to models that are not capable of capturing the real world patterns and behaviors air of conditioner. Also, if there is no sufficient data with proper labels for instance failure, information supervised on learning actual models may not be able to offer best results.

1.5.2.2 System Complexity

The University of Zambia has a large number of HVAC systems installed in different buildings and each system certainly using disimiliar technologies and configurations. This different diversity presents a challenge when trying to come up with a single would predictive be model effective that can successfully for all the air conditioners on the whole campus. In addition, the operational profile of these units may differ depending on the usage include of environmental the factors building, such possession as rates, humidity and sometimes external factors such as temperature changes. These factors may also affect performance thereby making it more challenging and complex to develop the predict model.

1.5.2.3 Implementation Challenges

Implementing the predictive maintenance model with the existing Online Job Card System (OJCS) may not be easy because the two may be incompatible. This may include issues such as how to handle sensitive operational university data and ensure that it is protected from unauthorised access as well as other legal and ethical considerations. Also, growing the model to encompass several buildings or campuses might need significant changes which may in turn make the model less effective and expensive.

1.6.0 Significant of the Project

The research project to be undertaken is of paramount importance to due its multifaceted implications and benefits for the University of Zambia's operations and infrastructure management. section This will seek to break down the concepts of this research in detail with reference to;

1.6.1 Enhanced Infrastructure Management

The integration of predictive maintenance techniques into the University of Zambia's Online Job Card System is crucial. It will allow the university to shift from the current corrective maintenance approach to preventive one. Predictive maintenance as a function of predictive analytics and sensor enabled technologies enables the university to forecast when equipment is likely to fail or in need of maintenance (Zhao et al., 2020). This transition therefore helps to minimize on time that is spent out of service and increases the useful life of various equipment and structures (Lu et al., 2019).

1.6.2 Resource Optimization

The application of the machine learning for resource allocation is a critical part of this study. It is crucial in the optimization of resource allocation at the university level. Through the use of machine learning models, the university is able to analyze large data sets and make better decisions on resource allocations (Claus et al., 2019; Chen et al., 2018). Thereby, this optimization decreases the operational costs and increases efficiency. It helps in ensuring that materials, labor and equipment are used only when they are needed and where they are needed which in turns reduces costs and increases efficiency (Wang et al., 2020).

1.6.3 Adherence to University Policies

Compliance with the university policies is very important in the research study project. This is so because the system is in line with the university policies regarding data management and protection. The issue of unauthorized access is a critical one and this research study project has put in measures to reduce the risk of such incidence by ensuring that only authorized persons are allowed to access the system at the university (Ansari et al., 2017).

1.6.4 Enhanced Communication and Decision-Making

The proposed research study project introduces a centralized dashboard as part of the system. This dashboard is very useful management tool in enhancing that the maintenance team and stakeholders have a communication view and decision-making the of job request cards with the university. This the is monitoring a of virtual the tool progress, will enable the identification of the areas that require more attention. The dashboard will enhances the communication and coordination between the different departments of the university and hence effective and efficient maintenance of tasks.

1.6.5 Sustainability

The transition from the conventional maintenance approach which is reactive to predictive maintenance is not only cost effective but is also environmentally sustainable. The university will be able to manage when and how the maintenance is carried out thus reducing on unnecessary resource utilization.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This chapter of the study is focused on the Predictive Maintenance Model (PMM) for Office Air Conditioning Systems (ACS) and other related structure equipment for which the topic has received considerable focus due to its ability to minimize equipment down time, increase productivity and reduce maintenance costs (Ayvaz et al., 2021). While compared to the conventional preventive or reactive maintenance strategies, predictive maintenance employs modern methods to identify the potential failures of equipment. The above approach is proactive in nature and leverages data mining, sensor data and predictive models to determine when to carry out maintenance thus enabling organizations to prevent costly breakdowns and improve asset performance (Wickern et al., 2019).

2.1 Data sources

The predictive maintenance models for Office Air Conditioning System is built on a number of data sources that help to identify that the enables system the performance. The real time basic tracking data of includes the the system's sensor functioning data parameters such as temperature, pressure, and humidity.This data is useful in identifying conditions that may lead to potential failures. Historical maintenance data also help in identifying trends of repeated issues and patterns of degradation over time (Kiani,et al., 2014).

In addition to the mentioned internal data points, the PMM also incorporate other factors such as environmental conditions and energy consumption profiles. Some of the environmental data that includes temperature and humidity levels of the outdoor air can enhance the workload on air conditioning units and hence their performance. Energy consumption patterns help in assessing the efficiency of the equipment and identifying when it is functioning below the expected standards. Thus, applying these diverse data sets as inputs, predictive maintenance models can effectively estimate potential failures, enhance the maintenance planning, decrease downtime, and increase the lifespan of AC units (Achouch et al., 2014).

2.2 Machine Learning Models for PMM

There are several and Machine Learning Models (MLM) such as Decision Trees, Random Forests, Artificial Neural Networks (ANN), Linear Regression, Support Vector Machines, K-Nearest Neighbors (KNN), and Long Short-Term Memory (LSTM) networks, which have been applied for PMM of Office Air Conditioning System (ACS) (Trivedi et al., 2021). These models assist in identifying potential failures of equipment thus enabling preventive maintenance to be made before the equipment breakdowns. Through the analysis of historical data on the performance of the equipment and the working conditions, it is possible to apply the methods of machine learning to determine (RUL) the of Remaining Useful Life thus optimize the maintenance plans, increase efficiency and prolong the service life of the

Office Air Conditioning System (Nishijima, 2017). The following is a brief description of each of the model' given below:

2.2.1 Regression Models

Linear Regression, Polynomial Regression and Support Vector Regression (SVR) are commonly applied in the field of predictive modeling where one seeks to forecast continuous variables for instance RUL of equipment and energy consumption in systems. Linear Regression is a straightforward method where only a linear relationship between the variables is assumed while Polynomial Regression is an more extension versatile of the Linear Regression can capture non-linear patterns. SVR, on the hand is cable to model complex relationships which provides accurate prediction in challenging environments (Dietze et al., 2018).

2.2.2 Time Series Models

LSTM (Long Short-Term Memory) networks and ARIMA (Auto Regressive Integrated Moving Average) models are two common models which are used for forecasting of time series results including failure likelihood over time. LSTM networks are a type of recurrent neural network which are especially good at remembering context over longer sequences. ARIMA, on the other hand, is a statistical model that examines and predicts future values based on historical data, offering a success approach for time-series predicting when trends and seasonality are involved (Gould el at., 2008).

2.2.3 Classification Models

Decision Trees, Random Forests and Support Vector Machines (SVM) are efficient classifiers that can be used to categorize the fitness state of equipment, for instance, Office Air Conditioning System into categories such as degraded, healthy or failure prone. Decision Trees are an effective way of demonstrating operation that is decision-making, since all of it depends on several inputs. Random Forests improve on this by creating multiple decision trees and taking an average of the results to enhance classification and make it less vulnerable to errors. SVMs employ hyperplanes to categorise the different health states and offer a very precise classification especially where there are other scenarios that are closely related. These models allow for proper monitoring and timely interventions that can be made to ensure that equipment performance is well maintained (Davis el al., 2012).

2.2.4 Hybrid Models

This approach combinies multiple models, such as integrating a regression model with a time series model, can significantly improve the accuracy of the prediction. Thus, this hybrid strategy takes into consideration more aspects of the data and the possible relationships between them. For example, a regression model can be very effective in determining the linear relationship between variables while a time series model can be useful in capturing seasonal, trends and other temporal variations. This integrated approach enables more accurate and comprehensive predicting and hence enhances the overall model performance and the reliability of predicting parameters such as system failure. (Blischke & Murthy, 2011).

2.2.5 Anomaly Detection

There are techniques like Principal Component Analysis (PCA) and Isolation Forests which are extensively used in identifying anomalies in the system in order to predict failures. PCA enables one to reduce the measure of data while at the same time preserving the essential variance and thus used to identify deviations from normal patterns. Isolation Forests, on the other hand, isolates the anomalies by selecting random features and splitting the data points and is effective in identifying outliers especially in high dimensional data sets. Both methods are very vital in predictive maintenance since they help in pointing out some abnormal activities that may lead to certain failures (Mayaki, 2024).

2.2.6 Deep Learning Models

These models incorporate sophisticated neural network structures to capture the intricate relationships present in the data especially when it comes to time series data. The state of the art models include the Long Short Term Memory (LSTM) networks and the Convolutional Neural Networks (CNNs) for processing of consecutive data. LSTM is used to manage the long-range temporal relationship while CNN is capable of learning temporal features, thus providing accurate results in analyzing the large and complex datasets. (Liu et al., 2018).

2.3 Critical review of related works

Many research works have been conducted on PMM for Office Air Conditioning Systems where various machine learning models, data sources and implementation approaches have been used. Some of the major approaches include regression models for forecasting, time series analysis for capturing trends over time, anomaly detection for identifying variations and hybrid models that encompass multiple approaches for enhanced prediction. However, these studies demonstrate the great possibility of PMM to enhance maintenance activities at the same time, they expose a number of challenges that include data quality concerns, precision, model and implementation challenges that need to be solved (Stehman & Foody, 2019).

2.3.1 Regression Model

A regression model is a computational model that is employed in the field of machine learning with the ability to estimate continuous responses as outputs from given features. Basically, it is used to identify the most likely location of a best fit line or curve through a set of data. It explains the dependence of variable on one or multiple variables.

2.3.1.1 Strength of Regression Model

Regression models namely Linear Regression and Support Vector Regression (SVR) are commonly employed for estimating the Remaining Useful Life (RUL) of Office Air Conditioning Units. These models are valuable due to their simplicity, and ease of interpretation as well as their ability to effectively deal with continuous data which makes them appropriate for the initial stage of predictive maintenance. Linear Regression offers a basic way of solving the problem by establishing linear relationships while SVR is more powerful as it can capture non-linear relationships through the 'kernel' trick. Both models allow for effective RUL forecasting to support the development of preventive maintenance plans (Lee el at, 2020).

2.3.1.2 Limitations

In the area of predictive modeling, the clarity of the regression models is usually obtained at the expense of precision. A large number of research works have highlighted that linear models are incapable of capturing the complex non-linear interdependencies between the operating variables and equipment performance. For instance, (Deng et al., 2018) established that linear regression models provided poor results where there were variations in environmental conditions as seen in office Air Conditioning Units. Also, regression models require good quality and continuous data. Missing data due to sensor failures or irregular sampling can severely degrade the quality of the model.

2.3.2 Time Series Models

Time series models are used in the analysis of data that are arranged in a sequence of time and are used in the making of future values.

2.3.2.1 Strengths

The LSTM network models have been popular due to their capacity to handle sequential data and making predictions based time on series time models series (Therneau et al., 2017). The performance of LSTM networks was also evidenced in the work of where they were used instead of traditional regression models for the prediction of PMM systems' degradation and it was found that the networks captured temporal dependence in the data.

2.3.2.2 Limitations

On the other hand, it is observed that LSTM models are prone to overfitting especially when trained on small data sets. Most of the research works for instance (Kumbure et al., 2022) highlights the problem of transferability of LSTM models to new data sets particularly in diverse settings where the operating conditions are quite different. In addition, the LSTM models are computationally intensive and therefore need a lot of power to train and infer, thus not suitable for real time applications or for use on edge devices with limited computational power.

2.3.3 Hybrid Models

Hybrid models in machine learning (ML) are models that Integrates multiple different Machine Learning algorithms or techniques to predict potential failures of an equipment or system.

2.3.3.1 Strengths

Hybrid models that include multiple techniques such as regression with anomaly detection or time series with machine learning classifiers and have been seen to have potential in enhancing the accuracy and reliability of the predictions (Ghrib, et al 2022). The hybrid approaches is capable of identifying both general tendencies and abrupt changes, thus offering a comprehensive PMM approach.

2.3.3.2 Limitations

While hybrid models offer an enhanced performance, they provide challenges in terms model design, maintenance and training. The incorporation of multiple models in a system is not as simple tuning as and implementing validation as each which model can and be is both likely time to consuming require and challenging to implement in large scale deployments (Elsayed et al. 2022). Furthermore, hybrid models may also suffer from compatibility problems. Using several approaches may result in the development black-box of models whereby it is difficult to explain why certain given inputs were assigned certain understanding outputs. This lack of clarity can negatively affect the adoption of PMM solutions particularly in place where justification is required.

2.3.4 Anomaly Detection Models

Anomaly Detection Models are computer programs that are used to detect any data that is not within the normal range of expectations or deviations from the norm. These deviations are referred to as outliers or anomalies.

2.3.4.1 Strengths

Anomaly detection techniques such as Principal Component Analysis (PCA) and Isolation Forests are effective in identifying abnormal patterns of air conditioning units that may indicate a failure (Mirnaghi and Haghighat, 2020) found that anomaly detection could indicate the initial stages of failure even when historical data on failures could not be explained.

2.3.4.2 Limitations

A notable challenge that affects the anomaly detection methods and that is the high false positive rate. It is important to note that the detection of anomalies does not always imply that there is equipment failure. Other studies carried out reports that false positives were high thus leading to increased maintenance costs instead of the expected savings (Cachada et al., 2018). However, the techniques of anomaly detection are often dependent on the features and hyperparameters used, which may be hard to tune without deep understanding of the domain.

2.4 Data Quality and Availability

A common concern that has been identified in the literature is the issue of data quality and availability. Some of the predictive maintenance models require the presence of high resolution and quality data that are collected at regular intervals. However, in reality the sensor data can be noisy and incomplete or even missing at times due to technical failures or other factors such as limited budgets. The steps such as imputation and filtering which are used in data pre processing are crucial, but they can also bring in biases that would affect the model performance (Gameiro. 2020).

2.5 Cost-Benefit Considerations

While most of the research emphasis on the technical aspect of PMM, very few provide the cost benefit analysis of implementing PMM systems. Some research studies argue that costs of sensor installation, data infrastructure and model may be quite high especially for and small medium enterprises making it very difficulty for them to implement PMM (Frederiksen et al. 2024). Moreover, stressed the need for taking into account the total cost of ownership, which includes the further support of the PMM system itself, which is not often covered in the academic studies.

2.6 Future Directions and Recommendations

The state-of-the-art can AI significantly techniques enhance such the as prediction deep precision reinforcement and learning model's and reliability ensemble in predictive maintenance model systems. Deep reinforcement learning makes the models capable of learning the best maintenance design from the data they interact with continuously; gradient boosting and random forests, which are ensemble methods, employ multiple models to produce a better solution. However, the implementation of these techniques requires careful analysis of the computational aspects and the question of whether more complex models are more likely to provide a correct model, which in turn may affect the practical implementation (Meshram, 2024).

2.7.0 Conceptual Framework for Predictive Maintenance

The conceptual framework for Predictive Maintenance Model (PMM) of Office Air Conditioning Units (OACU) provides a logical sequence of steps which help to identify the relationships between the variables that affect the maintenance process. This focuses on framework the role of data sources such as sensors and historical data and the machine learning algorithms, which process the data to predict failures. Key components include data collection, feature selection, model identification, and decision-making. The integration of the components effectively enhances the framework's ability to forecast Remaining Useful Life (RUL) and schedule proactive maintenance actions (Kang et al., 2023).

2.7.1 Key Components of the Framework

Given below are the key components required in the conceptual framework for Predictive Maintenance:

2.7.1.0 Data Collection

Sensor working data, such as temperature, pressure, humidity, vibration, and airflow, are crucial and important for monitoring the real-time performance of Office Air Conditioning Units (Elmouatamid et al., 2023). Operational data, including historical records of runtime, cooling capacity, and energy consumption, offer insights into overall efficiency and indicate potential wear and tear. Environmental data, such as atmosphere temperature, humidity, and weather conditions, also influence the unit's performance and are important for accurate predictions. To add on, maintenance records, which track repairs, component substitutions, and other activities, help identify patterns and trends in the OACS lifecycle.

2.7.2 Data Preprocessing

Data Preprocessing comprises several and varied important procedures to ensure that the data is correctly analyzed. Data Cleaning covers the process of correcting missing values, noise, and other anomalies in the data so that it can be used for training of the model. Feature Engineering entails identifying and selecting relevant features from the raw data for the purpose of enhancing the performance of machine learning models. When working with data, there is a need to scale and transform it in order to make it compatible with machine learning algorithms as well as to improve the performance and convergence of the models during training (Mishra and Rusch, 2023).

2.7.3 Machine learning models

Machine learning models such as regression models can be employed in the prediction of continuous variables for instance RUL using sensor and operational data. Time series models are applied in forecasting time series data which are sequences of events that vary with time, for instance failure rates. Anomaly detection models help in identifying unusual patterns in the unit operation which can be indicative of failure. Classification models are used in classifying the health state of the unit, for instance whether it is in good condition, deteriorating or about to fail through the use of past data and real time observations. Hybrid models involve the integration of different machine learning models in a way that can enhance the accuracy and reliability of the models (Didona et al., 2015).

2.7.4 Predicative Analytics

Predictive Analytics encompasses abilities such as failure prediction which entails assessing the likelihood and timing of future failures through the analysis of past as well as existing data. This enables preventive measures to be taken in order to avoid breakdowns that would be detrimental to production. Maintenance Scheduling aims at determining the best time to perform maintenance in order to avoid failure while at the same time reducing on time lost and the costs involved. Performance Optimization helps to determine where improvements can be made as far as energy efficiency and performance is concerned, thus guaranteeing that the air conditioning unit is in optimal condition during its entire life span (Chua, 2013).

2.7.5 Decision-Making

Decision-Making in predictive maintenance entails developing alert system that issue alarms at real time once the model estimate the future failure or variation. This enables for an immediate action to avoid equipment breakdowns (Patel, 2021). Maintenance Planning apply predictive insights to strategically plan the maintenance activities in order to avoid unnecessary delays that may lead to losses and hamper production as well as to avoid costly down times. Also, resource allocation is about managing and assigning the available maintenance resources based on the severity and nearness of potential failures, and effective management of personnel, materials and time in order to manage crucial concerns.

2.7.6 Feedback Loop

Model Updates include the process of enhancing and improving the existing machine learning models by updating the models with newer data and feedback from the maintenance activities. This process is therefore helpful in enhancing the model precision as well as to be able to cope with the changing environment conditions (Jakeman et al., 2006). Performance evaluation requires the frequent assessment of the success of the predictive maintenance system. This is a continuous process ensures that the system is always reliable in predicting the possible failures and also in improving the maintenance strategies.

2.8 Proposed Model for Predictive Maintenance

The proposed Predictive Maintenance Model (PMM) system for Office Air Conditioning System (OACS) will incorporate real time data and machine learning model for equipment failure prediction, for optimisation of maintenance schedules, and for the overall betterment of Office Air Conditioning systems. This system incorporates data acquisition, data cleaning, model building, and decision making as part of the process which becomes even more effective over time. The framework builds on the current literature that highlights the need for data-driven maintenance approaches to minimize costs and enhance system reliability (Filz et al., 2021).

The proposed system consists of the following key components:

2.9.1 Data Collection Layer

The Data Collection Layer comprises of a number of sensors and IoT devices which are crucial for the proper control and management of OACS. This comprises of temperature sensors for measuring the inlet and outlet air temperature, pressure sensors for measuring the refrigerant and airflow pressure, humidity sensors for measuring both environmental and system humidity and vibration sensors for identifying consumption and mechanical problems such as fan varinance. Futhermore, the power operational usage performance meters to monitor power usage to check operational performance (Klemen et al.,2018).

2.9.2 Communication Protocols

IoT devices will use communication protocols like MQTT (Message Queuing Telemetry Transport) or HTTP (Hypertext Transfer Protocol) to transfer data to a server. MQTT is normally used for its lightweight and effective message delivery, acceptable for low-bandwidth or unreliable networks, while HTTP is widely used for its effective and mostly supported data exchange capacity. These protocols make seamless data integration from various sensors and devices, enabling real-time monitoring and analysis of air conditioning systems (Mahbub et al, 2020).

2.9.3 Environmental Data Integration

External data sources, such as weather Application Programming Interfaces (APIs), play a ital role in providing real-time background temperature and humidity data, that directly impacts the performance of OACU. By integrating this external important information, PMM can account for environmental factors that influence system effectiveness and possible failures. This real-time data enhances the accuracy of predictions and improves overall performance monitoring (Maddireddy et al., 2022).

2.9.4 Maintenance Logs Integration

Historical data on past maintenance activity tasks including repairs and component replacements are meticulously kept in a relational database. This complete record provides useful information regarding the equipment’s lifecycle and performance trends that help in making better decisions regarding predictive maintenance. By analysing this data, trends and frequently occurring problems can be identified which helps in determining the future maintenance requirements and optimizing resources. This approach enhances the reliability and efficiency of the maintenance strategies (Ira et al., 2014).

2.10.0 Data Preprocessing Layer

This is a stage where raw acquired data is transformed into clean and usable formate for Machine Learning so the Machine Learning Model can yield accurate and dependable results.

2.10.1 Data cleaning

During data preprocessing it is important to deal with outliers, complete missing values and smooth noisy data in order to ensure that the analysis is as accurate and reliable as possible. Outliers can affect results and make interpretations misleading that is why they need to be identified and addressed. Data gaps are filled or the missing values are encoded to avoid having gaps in the data. Also, smoothing techniques are used to reduce the noise and improve the quality of the data. All these steps are very vital in coming up with good predictive models as well as enhancing the data quality in general (Smith, 2023).

2.10.2 Feature Engineering

Feature engineering includes creating new features from the raw data in order to increase the model performance. For instance, it is possible to derive features such as temperature differential between the inlet and outlet temperatures or calculate the power to efficiency ratio which can provide valuable information about the system performance. The features that are engineered can improve the efficiency of the machine learning models in predicting by highlighting major relations and trends which would not be easily noticeable from the given data (Johnson et al., 2016).

2.10.3 Normalization and Scaling

Normalization and scaling are important in preprocessing steps to ensure that data is changed to a common scale, which notably enhances the performance of machine learning algorithms. Normalization regulates the range of features to a standard scale, while scaling methods like min-max scaling or standardization make sure that features present equally to the model's training process. This process prevents features with larger ranges from extremely influencing the model, leading to more precise and dependable predictions (Black et al., 2022).

2.10.4 Time-Series Reshaping

Time-series reshaping involves structuring data into successions to suit the demands of time-series models, specifically Long Short-Term Memory (LSTM) networks. This process includes arranging historical data into fixed-length time windows or sequences that the model can utilize to learn temporal reliance and patterns. By transforming the data in this way, LSTM networks can effectively capture trends and forecasting future values based on past scrutiny. Proper modification is crucial for maximizing the performance of time-series models and impro predictive accuracy (Dannecker, 2015).

2.12.0 Predictive Analytics and Decision-Making Layer

Predictive analytic is a data driven process of using historical information, statistical models and learning algorithms to make assumptions about future events. It uses real time data as well as historical data of equipment to forecast for potential failures, improve the performance of the operations, and determine the right maintenance actions to be taken hence enabling proactive decision making and increasing the efficiency of the system as a whole.

2.12.1 Failure Prediction Dashboard

A real-time dashboard is a graphical user interface which presents crucial information for effective predictive maintenance, the forecasting Remaining Useful Life (RUL) of OACS, alerts for anomalies, and the health status of all the units. This dashboard provides a natural way of watching and regulating equipment performance and thus allows the user to easily detect any potential problems and solve them promptly. The dashboard that combines real-time data with the predictions enhances the decision-making process and the performance of the operations, which enables effective preventive maintenance and avoidance of unexpected failures (Quynh, 2022).

2.12.2 Maintenance Scheduling Module

The maintenance scheduling module will then use the improved schedules where predictive insights are used to approve of proactive maintenance before the equipment breaks down. Through the analysis of the history information and the real-time data, the system is able to generate proper schedules that can help minimize downtime and increase the useful life of equipment. This approach looks for potential failures and plans for preventive plans to avoid unnecessary failures and waste of maintenance resources. The integration of these predictive insights improves the reliability of operation and minimises the occurrences of unexpected failures. (Kothamasu et al., 2006).

2.12.3 Automated Alerts and Notifications

The alerts and notifications in predictive maintenance are the automated messages which are sent to the maintenance personnel to alert them about the potential equipment problems, such as those based on data and predictive models. These alerts and messages can be received through emails, SMS, or mobile applications; these notifications are crucial in facilitating preventive maintenance since they enable the maintenance to be performed before failures equipment. (Lee et al., 2020).

2.12.4 Resource Allocation

The system focus on maintenance tasks by assessing their nature and frequency, making sure that the resources are allocated properly. Through evaluating the importance of each task, the system helps the maintenance teams by attending to the most important issues first thus improving on the resource management and reducing on time spent idle. This approach enhances the overall effectiveness of the operation when dealing with potential equipment issues operation thereby increasing efficiency (Moore improving & the Starr, overall 2006).

2.12.5 User Interface (UI) and Reporting Layer

The User Interface (UI) in predictive maintenance is the visual and interactive part that enables the stakeholders to communicate with the predictive maintenance system. The UI and reporting layers offer a user-friendly interface through which users can view, analyze and respond to the data presented to them.

2.12.6 Real-Time Monitoring Dashboard

A clear and simple interface makes it possible for facility managers to quickly view information on all the OACU, view predictive maintenance tips, and view maintenance history. This makes it easier for maintenance managers to have a good view of the OACS performance and the necessary maintenance needs hence enabling them to work effectively in maintaining the equipment. Such a system enhances the management's decision-making and facilitates enhanced management of facilities (Johnson & Patel, 2023).

2.12.7 Reports and Analytics

The system generates detailed reports on the success of PMM activities, highlights key metrics such as downtime reduction, cost savings, and improvements in energy efficiency. These reports provide important awareness into the influence of predictive maintenance plans, assisting universities and companies assess performance and make accurate data-driven predictions to further improve operational efficiency (Wu et al., 2022).

2.13.0 Workflow

A workflow in PMM refers to the structured, step-by-step approach that indicates how maintenance tasks and activities are allocated and done based on predictive intuition. It ensures that possible equipment failures and breakdowns are addressed in an efficient way and also enabling maintenance teams to work in a proactive manner.

2.13.1 Data Collection

Data will be gathered from sensors fixed in the air conditioning units, which will transmit information to a server for processing. Furthermore, external data, such as current weather conditions, will be occasionally collected from various sources and fused into the system to improve accuracy and optimize performance Li., 2021).

2.13.2 Data Preprocessing

The collected data undergoes cleaning, preprocessing, and structuring to make sure that it is acceptable for machine learning models (Sun et al., 2018). Dominant features are extracted from the dataset, and time-series data is converted and modified to support effective analysis and modeling (Zhao et al., 2019).

2.13.3 Model Inference

The model inference will use real-time data fed into machine learning algorithms to forecast the remaining useful life (RUL) of air conditioning units, detect possible anomalies, and classify their health status. By constantly monitoring operational data, the model can provide early cautions of malfunctions, allowing proactive maintenance and minimizing downtime, while providing accurate insights into the status and performance of the air conditioning systems for optimum operation (Albayati et al., 2023).

2.13.4 Decision-Making and Alerts

Decision-making and alerts generation are triggered by the model's predictive results. Based on these outcomes, the system calculates optimal times for maintenance, allowing efficiency and reducing unplanned downtime. When the model predicts a possible issue or notices early signs of failure, automated alerts are instantly generated and sent to the maintenance team and stakeholders. This proactive approach enables timely interventions, debarring equipment breakdowns and enhancing overall air cone conditions system reliability and dependability.

2.13.5 Feedback and Learning

A feedback loop will allow that all maintenance actions and outcomes are logged and fused back into the system for continuous enhancements. This logged data is used to enhance the accuracy of future predictions by providing updated, real-world insights. Periodic retraining of the machine learning models will occur, taking advantage of this new data to process their ability to forecast possible failures, improved maintenance schedules, and enhance overall system performance. (Nguyen & Medjaher, 2019).

2.13.0 Comparison with Related Research Works

This section entails a comparison with other related research works and identification of gaps by assessing prior research in the area of PMM using Machine Learning (ML). It outlines the methods, models and strategies that have been employed in the previous research studies in order to determine the effectiveness and shortcomings of the existing approaches. Thus, the gaps identified in this research will be the basis for the improvements and new contributions to this field.

2.13.1.0 Predictive Maintenance Using IoT and Machine Learning

This is the use of Predictive Maintenance and Internet of Things (IoT) and ML to predictive when the equipment or machine is likely to fail of breakdown.

2.13.1.1 Approach

This research study focuses on decision tree algorithms for the utilization the of Predictive Maintenance and IoT devices acquisition of real-time data for the OACS (Kanawaday and processing & Sana, 2017).

2.13.1.2 Strengths

The study also points out that there is a need to incorporate the Internet of Things (IoT) with the predictive maintenance plans. Through the use of real-time data as well as improved analytics this approach effectively reduces equipment failures, improves the maintenance schedule, reduces the operational costs, and increases overall efficiency making it an important factor in today's industrial operations

2.13.1.3 Limitation

The limitation of the study can be attributed to the use of basic machine learning models such as decision trees, which makes the system incapable of identifying complex correlations in the data that are more likely to be present in real-life applications.

2.13.1.4 Comparison

The suggested model can be compared to Predictive Maintenance approaches that employ IoT and Machine Learning, and thus emphasises the benefits of employing more complex models such as Long Short-Term Memory (LSTM) networks and hybrid models. These techniques allow for a better utilization of the time-series data and lead to an enhanced prediction performance as well as more effective and refined maintenance plans for complicated and state-of-the-art systems.

2.13.2.0 Deep Learning for Predictive Maintenance

This is another technology that can be used to predict the possible failure on the equipment. The technology is cable of analyzing of complex data patterns such as images and audio signals, to assist spot issues before they became challenges.

2.13.2.1 Approach

This research study demonstrates how deep learning, particularly the Long Short-Term Memory (LSTM) networks, can be utilized in predicting the Remaining Useful Life (RUL) of industrial equipment including OACS. With the help of LSTM's capability to analyze the time series data, the study enhances the efficiency of predictive maintenance and improves the reliability of equipment and overall productivity (Wang et al., 2020).

2.13.2.2 Strengths

This paper proposes the use of LSTM networks which allows the model to capture and store temporal dependencies present in the data thereby making it suitable for time series predictions. This is due to the fact that LSTM's specific architecture is capable of dealing with sequential information and, therefore, provides more accurate forecasting where time series are involved, such as in predictive maintenance.

2.13.2.3 Limitations

The research study focal point is mainly on the prediction of remaining useful life of equipment with little consideration of other relevant issues such as anomaly detection as well as resource management. The research is more concentrated on RUL prediction without considering other aspects which could help to enhance system performance including early fault detection and effective resource management.

2.13.2.4 Comparison

The suggested model can be viewed as an extension of the conventional Deep Learning models for Predictive Maintenance and incorporates LSTM networks along with anomaly detection models such as Autoencoders. In addition, it incorporates a decision-making component that goes beyond the simple prediction of equipment failure and optimisation of maintenance strategies and resource allocation hence improving efficiency and resource utilisation.

2.13.3.0 Hybrid Models for Predictive Maintenance

These are models that combines the different machine learning techniques in order to enhance the accuracy and robustness of prediction for equipment maintenance.

2.13.3.1 Approach

This research work applies a combined approach where machine learning models used for estimating Remaining Useful Life (RUL) are integrated with anomaly detection models thus enhancing the existing Predictive Maintenance (PM) framework. Thus, the integration of these methods increases the accuracy of failure predictions as well as the ability to identify the initial indicators of equipment failure which results in the development of more effective and timely maintenance strategies (Ran et al., 2019).

2.13.3.2 Strengths

The The hybrid model enhances the robustness of the system by addressing all the important aspects of predictive maintenance that the given model fails to. The integration of machine learning for Remaining Useful Life (RUL) prediction and anomaly detection as a combined approach provides a more comprehensive and effective solution for identifying and preventing failures, while also improving the efficiency of maintenance scheduling.

2.13.3.3 Limitations

The implementation of the hybrid model system is largely limited to specific industrial applications, restricting its ability to a wider range of use case scenarios. Furthermore, the complexity of hybrid models can lead to extended development cycles and greater resource consumption.

2.13.3.4 Comparison

The proposed model is different from the traditional hybrid Predictive Maintenance systems in that it also follows a hybrid approach but with several important improvements. It incorporates edge computing for real time data analysis thus enabling fast results and also stores the data in the cloud. This combination makes the system suitable to a wider range of OCSU across the university environments, ensuring effective performance and enhanced scalability in both small-scale and large-scale applications.

2.13.4.0 Energy-Efficient Predictive Maintenance

2.13.4.1 Approach

This research also shows that predictive maintenance is effective in realising the energy efficiency gains of HVAC systems, in addition to maintaining and improving the performance of HVAC systems through the management of energy consumption while at the same time anticipating equipment failure. Through the management of potential failures, the approach not only increases the lifespan of equipment but also reduces energy loss, thus providing an efficient and economical way of operation (Cheekati et al., 2024)

2.13.4.2 Strengths

The research shows that energy efficiency is an important factor that can be incorporated into PM, as it provides two main advantages of cost efficiency and environmental efficiency. Through the efficient use of energy, PM systems reduce the costs of operations while at the same time reducing the impact on the environment. This approach enhances the overall value proposition as such systems become more desirable for industries that aim to reduce costs and minimize the harm to the environment.

2.13.4.3 Limitations

The study is mainly centered on energy efficiency and that is a very relevant aspect. Nevertheless, the paper does not go a step further to consider other issues such as anomaly detection and Remaining Useful Life (RUL) estimation. All these factors are crucial in order to gain a better insight into the equipment performance and its lifetime, hence calling for more research.

2.13.4.4 Comparison

The proposed model presents a major improvement over the conventional energy efficient predictive maintenance models. However, it does not only include energy efficiency as one of the main features, but also expands it by integrating a powerful predictive maintenance. This all-encompassing framework encompasses RUL estimation, fault detection, and maintenance planning thus offering an integrated approach

2.15.0 Summary

After analyzing different existing machine learning models and several previous research works on predictive maintenance, the proposed model incorporates benefits of Time Series analysis, Random Forest and Gradient Boosting Machines. This combination improves the reliability of the predictions made for maintenance needs and thus effectively enhances system performance and reduces downtime. The following chapter will discuss the methods used in the research study.

CHAPTER 3

METHODOLOGY

3.0 Introduction

This chapter presents the methodology for implementing a predictive maintenance machine learning model in office air conditioners. It explains the research design, the methods used and why they were used, the data collection process, data preprocessing, the model selection process, the deployment of the model, and the ethical considerations.

3.1 Research Design

The research design for predictive maintenance using machine learning isexactly of a design that would produce a strong system that is able to predict equipment failures, plan for the best time to maintain the equipment, and reduce on time lost due to failures. This design incorporates sensor data, machine learning models and real time decision making for the estimation of equipment's RUL and detection of any abnormalities for preventive maintenance (Sayyad et al., 2021).

The research is practical in nature and has a problem solving orientation to address the predictive maintenance issues using supervised machine learning approach. The data that it uses comprises of historical sensor data of the machinery, environmental factors and maintenance records. The research methodology is also an iterative process of data collection, cleaning and organizing, modeling and evaluating the data as well as improving the accuracy of the models and their applicability in the practical field.

This design is a combined approach where quantitative analysis is used to qualitative measure analysis the is performance used of to the derive model operational and insights, where data will be collected, processed and fed into machine learning models for predictive analysis. The model's performance will be evaluated in the real world and in the field with real-time feedback, which will make the model very accurate and relevant to real-life application (Hicks (et al., 2015).

3.2 Adopted Method and Justification

The study applies supervision learning model where classification models and time series data are used. This type of learning is appropriate for this study because there are historical data with outcomes such as machine failures whereby patterns can be recognized and events can be forecasted. Some of the machine learning algorithms that are suitable for time series data and have high predictive accuracy include Random Forests, Gradient Boosting Machines, and Long Short-Term Memory (LSTM) networks (Huang et al., 2019).

In this paper, Random Forest and Gradient Boosting are used because they are the ensemble learning techniques that include multiple decision trees in order to prevent overfitting and improve the predictive ability of the model. LSTM networks are used because they are capable of equipment handling degradation sequential over information time series data which is crucial for analyzing is trends and cost effective as it focuses on both the accuracy and the computational time for the predictive analysis to support real time maintenance planning, thus reducing the maintenance costs and improving the operational efficiency (Xie et al., 2018)

3.3 Association of Research Method to the Project

The chosen approach is relevant to the project goals that include identification of equipment failures and enhancement of maintenance data management. The machine learning models will use sensor data to identify the subtle patterns in the equipment's behavior and predict the possible failures before they happen. Also, using time-series analysis with LSTM networks as the research method enables the capturing of temporal characteristics of equipment deterioration suitable for real-time monitoring systems. This makes it possible to ensure that the models are flexible enough to accommodate changes in operational conditions and environments thus making the system relevant at all times (Liu et al., 2020).

3.4 Research Data and Datasets

The research relies on a combination of internal operational data and external environmental data to train the machine learning models.

3.4.1 Internal data

In this case, internal data refers to the historical sensor data which includes parameters like temperature, pressure, vibration levels, electricity consumption and other relevant operating variables. This data is gathered from IoT sensors that are fitted on OACU or other industrial machinery.

3.4.2 External data

External data consists of weather information such as temperature and humidity that affects the performance of equipment. Such data is obtained from standard meteorological databases or from additional sensors. The data set covers few months to years, which is enough for the variation in data to be used in the training of the machine learning models. The process of Labeling of the dataset involves identifying of defining the situations, instances of equipment failure, or maintenance interventions. To prepare the data for modeling, the data is cleaned to deal with missing values, noise and outliers (Vinisha & Sujihelen, 2022).

3.5.0 Data Collection Methods

The following will be the Data Collection Methods and Data Analysis Techniques that will be used in this research to collect and analyze the data. All the methods will ensure that data is collected and analyzed accurately and properly for effective monitoring, assessment, and decision making especially in the area of predictive maintenance and other applications (Achouch et al., 2022).

3.5.1 Sensor Data Logging

The sensors are installed in equipment which will track and record key performance factors of equipment including temperature, vibration and pressure. All these real time data will be received by a central server where they will be grouped before being analyzed. This paper has discussed how the system will help in collecting and centralizing the information that will enable predictive maintenance, increase efficiency of the operation and identify any anomalies. This continuous monitoring will be very useful in the assessment of the performance of the equipment as it will enable early identification of potential failures, appropriate scheduling of maintenance activities and hence enhancement of the reliability of the system (Lee et al., 2006).

3.5.2 Maintenance Logs

The maintenance team will also take records of the maintenance activities that are being carried out as well as the results of such activities and the data collected can be used for supervised learning. This process not only helps in the training of machine learning models but also enhances the knowledge of maintenance effectiveness and hence helps in identifying the patterns and improving the maintenance strategies.

3.5.3 External Data Sources

The Weather data will be acquired from external sources at specific time intervals from online databases as well as physical weather stations in order to incorporate real world weather conditions into the model. Thus, the model will be able to incorporate more real life conditions such as weather which may affect performance and maintenance thus enhancing the accuracy of the model and improved decision-making (Wakiru et al., 2019).

3.5.4 Data Analysis Techniques

The data analysis techniques will include the approaches used to analyse and manipulate data, clean, integrate and organise data as well as the modelling of the data. These techniques are designed to identify patterns, make inferences and support decision making based on data (Osborne, 2012).

3.5.5 Data Preprocessing

This will entail that data cleaning which is the process of removing noise, outliers and feature scaling (normalization) in order to ensure that the data is appropriate for analysis. Methods such as Principal Component Analysis (PCA) can be used in order to decrease the number of variables and to concentrate on the most significant ones.

3.5.6 Feature Engineering

This will be done in order to ensure that the right features such as rolling averages, change rates and thresholds are properly extracted from the raw data in order to enhance the performance of the model. In this way, the model will be able to identify the trends and patterns from the data which will result into better predictions and recommendations that can be made from the data. This process is crucial to data analysis.

3.6.0 Machine Learning Models

For predictive tasks, supervised learning models such as and Random Long Forests, Short-Term Gradient Memory Boosting (LSTM) Machines networks will be used (Valarmathi & Kanaga, 2021). All these models are chosen specifically because of their ability to handle big data and able to capture complex interdependencies as well. Furthermore, LSTM networks perform well in dealing with time series efficiency data of especially the where temporal patterns play a crucial role in the predictions. This combination of models improves the overall effectiveness of predictive analytics

3.6.1 Evaluation Metrics

All the models will be evaluated using several metrics such as precision, recall, F1-score, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) as suggested by Sathya and Sathya (2023). This extensive evaluation is taken into consideration in order to ensure that the predictions made are as accurate as possible, which would make them reliable in real life. From the following metrics, we will be able to determine the strength and limitations of each model, and therefore be in a position to make necessary changes. Basically, this evaluation assists in determining the reliability of the predictive models when used in practical applications.

3.7.0 Ethical Concerns Related to the Research

Although Predictive is mostly about mechanical systems, but it raises a number of ethical questions that need to be taken into account. These problems are risk of data invasion, risk of bias in the algorithms and risk of automation on employment (Langer & Landers, 2021). It is important to take into consideration the ethical issues in order to achieve proper and non-biased application in practice:

3.7.1 Data Privacy and Security

Due to the fact that most of the operations are based on the use of IoT sensors and cloud storage, there is a high likelihood that sensitive information regarding the operations may fall into the wrong hands. To reduce this risk, it is crucial to adopt strong security measures such as anonymizing and encrypting all the data. Also, making sure that data is secured properly and in accordance with certain industry guidelines will also help in avoiding possible intrusions. It is also important that there are regular audits and changes to security measures to suit the dynamic threats. It is therefore important that organizations pay attention to the security of data to ensure that their information is not disclosed to unauthorized persons and build the much needed trust with their stakeholders and clients (Aldboush & Ferdous, 2023).

7.2 Bias in Machine Learning Models

There is a significant risk of bias in machine learning models if the training data is biased in some way, for instance if the data is skewed or missing some information. This makes the models developed in such an environment to have a hard time in identifying anomalies particularly when certain failure modes or operational conditions are not well represented in the data used for training (Mayaki, 2023). This risk can be reduced by ensuring that the the dataset is diverse and representative of all possible conditions and failure modes. A balanced dataset can improved the performance and reliability of predictive models, evetually resulting in more accurate and unbiased predictions.

7.3 Autonomous Decision-Making

When systems are developed to autonomously alert maintenance teams or make decisions regarding equipment repairs, concerns about human oversight arise. In the areas of operation that are critical, it is necessary to develop and implement the mechanisms that will enable the operators to check and confirm these automated decisions (Alves et al., 2019). This is because besides being used to ensure the accuracy of the information, it also acts as an accountability measure. Therefore, through the use of human-in-the-loop approach, the benefits that come with automation can be maximised but at the same time, allow individuals with wealth of experience to have the last word on critical maintenance decisions. This ensures safety is enhanced, the chances of making a mistake are minimized and the confidence in the system is increased.

7.4 Impact on Workforce

As predictive maintenance systems continue to be developed further and integrate more and more functionalities that help with the maintenance decision-making process, there is a possibility of changing the role of personnel. the This workforce, shift particularly can the lead maintenance to issues such as job loss and changes in the workplace. To ensure that this transition goes smoothly, it is imperative that workers are reskilled and fit into the new system. Training the employees on new technologies and how to analyze data can help them keep up with these changes and also help them to assume higher level responsibilities within the organization. Focusing on the role of human intervention together with the automated systems will create an atmosphere of cooperation that makes the working process more effective and satisfying for the employees (De et al., 2018).

7.5 Recap of Limitation of Study

In the current state, the UNZA's OACS do not have sensors linked to a central server, which hinders the ability to capture data in real-time and predict system failures. Also, the University has different types of air conditioning systems installed in different buildings and the systems are of different makes and models. This divergent makes it difficult for the model developers to come up with a single predictive model that can be easily applied across all the systems. Another challenge is the lack of enough data concerning the performance of OACS, their failures, and maintenance activities used in the buildings. This may also pose a challenge where integrating the predictive maintenance model with the existing Job Card System (JCS) may of pose the some research compatibility study problems.

Due to the above given limitations of the research study,certain assumptions have been made. The data for that training the the Office Air Conditioning Model (OSCM) will units be are obtained of from the Kaggle data website model. This is because the website has datasets that have used on other similiar studies.Another assumption will involve the integration of the predictive model into the Job Card System, where an endpoint within the application will allow users to access and view the dashboard.

7.5 Summary

In the light of the comprehensive analysis of various machine learning approaches and previous works on predictive maintenance, the suggested model incorporates efficient data acquisition and processing strategies. These include regular sensor data recording, proper maintenance history information, data cleaning, and optimal feature selection. Thus, the model enhances the accuracy of failure detection and prevents failures from occurring at critical times through timely maintenance actions, and thus, the maintenance plans are effectively arranged, downtime is minimized, and the efficiency of equipment is improved. In the next chapter, the author will elaborate on the project research data, research experiment and implement aspects of the research project.

CHAPTER 4

DATA, EXPERIMENTS, AND IMPLEMENTATION

4.0 Introduction

This chapter gives a comprehensive study of data, experiments, and implementation related to predictive maintenance for office air conditioners. It rigorously discusses the data processing work flows, details the experiments conducted to develop predictive maintenance model, and describes the development of the maintenance dashboard interface, including API endpoints designed for robust data validation and integration.

4.1 Data

It is vital to have data to develop a predictive machine learning model for the OACS that the model can learn from the data, change according to the data it receives and make accurate predictions. The historical data of temperature, occupancy and energy consumption assist in understanding the trends which help the model to regulate the energy consumption according to the conditions. Such data also enhances the accuracy of the prediction as well as underlines the importance of adaptive real-time adjustments. Other inputs such as weather information also help the model to predict the indoor temperature requirements at any given time so as to ensure that energy is conservely used and comfort is maintained. The information regarding the system performance enables one to do predictive maintenance which means that downtime and costs are reduced. The model is also able to come up with energy saving measures from the usage and occupancy trends that it analyzes to yield in reduced costs and support sustainability. Also, data concerning the occupant's comfort preferences help in maintaining the right balance between thermal comfort and energy consumption (Jin et al., 2015).

4.1.1 Data Acquiring

Data acquisition is an essential component of any machine learning task and this becomes even more so where there are constraints such as the absence of sensor data gathered at a central server and the absence of maintenance records for university air conditioning systems. To overcome these gaps, data was captured from various sources including Kaggle and Google Datasets. This way, it was possible to expand the data set and make sure that the machine learning models used for the analysis were as precise as possible when using the data. (Qolomany et al., 2015).

4.1.2 Data Cleaning

Data cleaning is an essential process in the machine learning model and it plays a vital role in determining the quality, correctness, and effectiveness of the models. Nevertheless, it can be difficult and laborious and the data issues may need to be addressed using different techniques. In this research the common steps in the data cleaning process for machine learning were studied and how these were applied in cleaning the data in order to make it suitable for analysis and modeling (Rahmani et al., 2021).

4.2.0 Prediction Experiment

The goal of the Predictive Maintenance Experiment (PME) for an Office Air Conditioning when (OAC) system is needed to predict potential failures by applying machine learning models on historical data of the system and its maintenance history. Therefore, by analyzing factors like the time the system has been running, the energy that it has consumed, the changes in temperature, as well as the frequency of past failures, the model can be able to determine certain conditions that indicate that a particular system is likely to fail. By employing this proactive approach, it is possible to act at the right time and thus reduce downtime, increase the lifetime of the system and increase efficiency. This paper aims at predicting the maintenance requirements for OACS to ensure that AC operation is not disrupted, thereby increasing comfort for the occupants and reducing the maintenance costs for the facility (Abdel-Razek et al., 2022).

4.2.1 Data Required

In order to implement an efficient predictive maintenance for an office AC system, it is crucial to gather sufficient information. Some of the important data include system run times which give the usage patterns and wear information; humidity levels which affect the comfort levels and the effectiveness of the system; and temperatures which monitor the atmospheric conditions on the indoor and outdoor environment. These include maintenance logs and number of maintenance activities which provide information on service history; this is where pressure readings such as refrigerant and airflow pressures can help in identifying potential problems. This includes information about past faults which provides data on previous failures and readings from other sensors provide real time measurement of the OAC system health to enable preventive measures to be taken (Laflamme et al., 2022).

4.2.2 Approach

After completing the data collection and processing activities, the dataset was divided into the training set and the test set. Advanced data analysis techniques such as time series analysis, anomaly detection and survival analysis were employed in order to accurately estimate the system failure probabilities over time. To this end, state-of-the-art models, namely XGBoost, which is known to perform well in handling sequential data, is used in the approach as it is capable of identifying the complex temporal trends and predicting the maintenance requirements in the dynamic and real-time settings. The following models offer a solid base for predictive maintenance, which enhances reliability and decreases the useful life of equipment (Bampoula et al., 2022)..

4.2.3 Split the data into Training and Test set

The data collected from the OAC system during the period of forty three months was systematically divided into two distinct datasets. Eighty (80%) percent of the data was allocated to the training dataset, which is important for training the predictive maintenance model. The remaining 20% was used as the test set. This division was deliberate so that the model can be tested on hidden data to allow for a proper assessment and confirmation of the performance of the model. This made the model to be reliable and accurate in the real world applications and eventually enhance the maintenance strategies(Eiben et al, 2022).

4.3.0 Implementation

Implementing a predictive machine learning model for an Office Air Conditioning System (OACS) need a well-defined, structured plan. This plan should include several critical phases such as Exploratory Data Analysis (EDA) to understand data patterns, feature engineering and selection to recognize important variables, model development, deployment for use, and unbroken improvement and optimization to improve performance over time. Each stage is important for ensuring the model's effectiveness and adaptability to changing conditions, ultimately leading to improved energy efficiency and user comfort. This systematic approach will help in achieving reliable and accurate predictions for maintenance and operational efficiency. (Bezerra et al., 2020).

4.3.1 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is very useful for identifying correlations and structures in the data set. This includes comparing different variables, for example the relationship between temperature and air conditioning power consumption and identifying the patterns of occupancy that affect temperature and energy use. Thus, visualizing such data can trends easily one perform effective feature selection and check whether some data cleaning is required before the predictive model is created, so that only quality and relevant data are used for the model. This thorough analysis helps in enhancing the model performance and provides a better insight into the working of the OACS (Raparthi et al., 2020).

4.3.2 Feature Engineering

Feature engineering is an essential component in developing a predictive maintenance model for a WestPoint Office Air Conditioner. This process entails deriving the right features from the data that is collected and some of the important features include humidity levels, maintenance counts, runtime, power consumption and various sensor readings. Also, the dataset has a target column for failure prediction which is set to 0 if no failure occurred and a value of 1 denotes a failure event. The model is able to recognize some tendencies and predict potential failures, which would lead to improvement of the efficiency and reliability of the Office Air Conditioning System. The cleaned data is presented in Table 1 below while Table 2 presents the feature description.

Table 1 cleaned data sample

Id	Date	Temp (°C)	Humidity (%)	Runtime (hrs)	Power (kW)	Pressure (Pa)	Maintenance Count	Sensor 1	Sensor 2	Failure (0/1)
1	2024-01-01	22.5	45	8	1.5	1000	2	4.5	0.8	0
2	2024-01-02	23.0	48	7.8	1.4	995	2	4.4	0.78	0
3	2024-01-03	22.8	47	8.2	1.6	1002	2	4.6	0.82	1
4	2024-01-04	24.0	50	7.5	1.3	980	2	4.3	0.75	0
5	2024-01-05	21.5	40	8.4	1.55	1005	3	4.5	0.79	0
6	2024-01-06	25.0	52	7.3	1.2	975	3	4.2	0.73	0
7	2024-01-07	22.9	46	8.1	1.5	1001	3	4.5	0.8	1
8	2024-01-08	23.7	49	7.9	1.4	990	2	4.4	0.77	0

Table 2: Feature explanation

FNo.	Feature	Explanation
1	**Record ID**	Unique identifier for each record.
2	**Date**:	Date of the recorded data.
3	**Temp (°C)**	Temperature in degrees Celsius.
4	**Humidity (%)**	Humidity percentage in the environment.
5	**Runtime (hrs)**	Total operating hours of the air conditioner on the day.
6	**Power (kW)**	Power consumed by the air conditioner (in kilowatts).
7	**Pressure (Pa)**	Internal pressure in the air conditioning system (Pascals).
8	**Maintenance Count**	Number of maintenance operations performed prior to this record.
9	**Sensor 1**	Voltage reading from a sensor.
10	**Sensor 2**	Current reading from another sensor.
11	**Failure (0/1)**	0 for no failure, 1 for failure event detected.

4.3.3 UNZA Model Selection

In order to solve the predictive analysis, supervised learning models including Random Forests, Gradient Boosting Machines, and Long Short-Term Memory (LSTM) networks were used. These models were selected based on their ability to handle a large number of data points and to learn the complex relationships that exist between variables. LSTM networks are particularly effective for sequence data, which makes them particularly useful when temporal characteristics are relevant to the prediction process. This model combination is designed to increase the efficacy of the predictive analytics as stated by (Mullangi, 2017).

4.3.4 UNZA Model Development

In this case, in order to create a time series machine learning model that can predict possible failures of a WestPoint office air conditioner (AC), it is crucial to focus on the time series data and the binary failure events which are represented by the value 0 or 1. These temporal patterns help the model to learn the historical trends such as temperature or usage trends, which can be associated with the future possible failures. Below is a step-by-step approach implemented in Python with the use of essential libraries: First, pandas library was employed to manage data such as data cleaning and data manipulation of the time stamped records in order to create features. For model training and validation, scikit-learn provides useful functions for scaling, splitting the data and measuring the performance of the model. Finally, xgboost was used since it is known to give good results for tabular data as well as LSTM

(Long Short-Term Memory networks) to model the temporal dependencies of the data with the help of a neural network (Yan et al., 2020).

4.3.5 eXtreme Gradient Boosting

XGBoost is a class of gradient boosting decision tree algorithm that is most suitable for time series data and can be used in systems for air conditioning (AC) to predict failures. XGBoost is an algorithm which successively creates decision trees and uses the 'boosting' method to generate a sequence of trees in order to increase the predictive power of the model by addressing failures of the previous trees. This approach makes it possible to identify the patterns in the historical data, including the usage power and temperature fluctuations, and thus to provide accurate predictions of the AC breakdown. Being efficient and precise, XGBoost is considered to be one of the best methods for dealing with challenging time-series data sets (Trizoglou et al., 2021).

4.3.6 Implementation of eXtreme Gradient Boosting

To implement XGBoost for time-series failure prediction essential Python libraries such as pandas, numpy, scikit-learn, and xgboost are needed from the Python ecosystem. These libraries support different tasks, pandas handles data manipulation, numpy hnandles numerical operations, scikit-learn provides machine learning tools, and xgboost implements gradient boosting for robust predictive modeling, especially in time-series and tabular data. (Mallouhy, 2023).

Python Code Implementation

```
# Code start here

import pandas as pd

import numpy as np

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler

from xgboost import XGBClassifier

from sklearn.metrics import accuracy_score, classification_report


# Step 1: Load sample clean data imported from the CSV file

# This sample of the python data frame of sample data as shown in Table 1 cleaned #data sample


data = {

    "Date": pd.date_range(start="2024-01-01", periods=100, freq="D"),
```

```
    "Temp": np.random.uniform(21, 25, 100),
    "Humidity": np.random.uniform(40, 52, 100),
    "Runtime": np.random.uniform(7.0, 8.5, 100),
    "Power": np.random.uniform(1.2, 1.6, 100),
    "Pressure": np.random.uniform(950, 1020, 100),
    "Maintenance_Count": np.random.randint(0, 5, 100),
    "Sensor1": np.random.uniform(4.0, 4.6, 100),
    "Sensor2": np.random.uniform(0.7, 0.85, 100),
    "Failure": np.random.randint(0, 2, 100)
}

df = pd.DataFrame(data)

# Step 2: Convert date to datetime and set it as the index (time series)
df['Date'] = pd.to_datetime(df['Date'])
df.set_index('Date', inplace=True)

# Step 3: Create lag features (previous time steps as features)
for lag in range(1, 4):      # Using 3 lag features
    df[f'Temp_lag_{lag}'] = df['Temp'].shift(lag)
    df[f'Humidity_lag_{lag}'] = df['Humidity'].shift(lag)
    df[f'Runtime_lag_{lag}'] = df['Runtime'].shift(lag)

# Drop rows with NaN due to lagging
df.dropna(inplace=True)

# Step 4: Train-test split
X = df.drop(columns=["Failure"])   # Features
y = df["Failure"]                  # Target (Failure)

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, shuffle=False)
```

```
# Step 5: Standardize the data
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

# Step 6: XGBoost Classifier
model = XGBClassifier()
model.fit(X_train_scaled, y_train)

# Step 7: Predictions and Evaluation
y_pred = model.predict(X_test_scaled)
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Classification Report:\n", classification_report(y_test, y_pred))

# Optionally, save the model for later use
# import joblib
# joblib.dump(model, 'xgboost_model.pkl')
```

4.3.7 Lag Features

These features in XGBoost provide the model with crucial historical context by incorporating past values as additional features. For instance, features like Temp_lag_1 and Humidity_lag_2 are the temperature and humidity values of the previous one and two timestamps, respectively. This enables the model to learn the temporal dependencies which in turn improves the accuracy of the predictions (Kang et al., 2023).

4.3.8 Scaling

In XGBoost, normalising features is very necessary to make sure that all the units of measurement are on the same level, for instance, temperature, power, and humidity. This helps in standardizing these features and as a result the model takes less time to train, this also helps in minimizing bias that may arise due to different scales. This makes the model performance better and also increases the predictive power of the model (Huang et al., 2023).

4.3.9 Robust Evaluation

Evaluation is vital in enhancing the efficiency of a model and involves the use of a robust structure like the confusion matrix. This matrix offers a way of understanding how the model ranks instances and the resulting accuracy of the model's predictions which may be classified as true positives, true negatives, false positives, and false negatives. In addition, precision-recall curves are used to show cross-tabulation of precision versus recall and this is very useful in case of infrequent failure events. All these evaluations are useful in tuning the model and to improve its predictive capability. (Parthasarathy et al, 2023).

4.3.10 Confusion Matrix

The confusion matrix is a critical structure that allows assessing the effectiveness of XGBoost models in classification problems. It gives a general idea of the model's accuracy but at the same time, identifies particular points that require enhancement thus enabling the enhancement of the model for better predictive analysis (Dhaliwal et al, 2018). The following is the format of a confusion matrix:

Table 3 Confusion matrix

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Python Code Implementation

```
# Code starts here


from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay


# Predictions
y_pred = best_model.predict(X_test_scaled)


# Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cm)
disp.plot()
plt.show()
```

#Code ends here

4.3.11 Precision recall curve

The Precision-Recall curve is a useful metric for assessing the performance of classification models in general, including the XGBoost (XGB) models. This curve is effective in depicting the precision versus the recall for various thresholds where, often, precision and recall are of paramount importance especially when dealing with imbalanced class distributions (Priscilla and Prabha 2020).

Python code implementation

 # Code starts here

```
from sklearn.metrics import precision_recall_curve, average_precision_score


y_pred_proba = best_model.predict_proba(X_test_scaled)[:, 1]


precision, recall, thresholds = precision_recall_curve(y_test, y_pred_proba)
average_precision = average_precision_score(y_test, y_pred_proba)


plt.plot(recall, precision, label=f'AP = {average_precision:.2f}')
plt.xlabel('Recall')
plt.ylabel('Precision')
plt.title('Precision-Recall Curve')
plt.legend()
plt.show()
```

#Code ends here

4.3.12 UNZA Predictive Model Dashboard Implementation

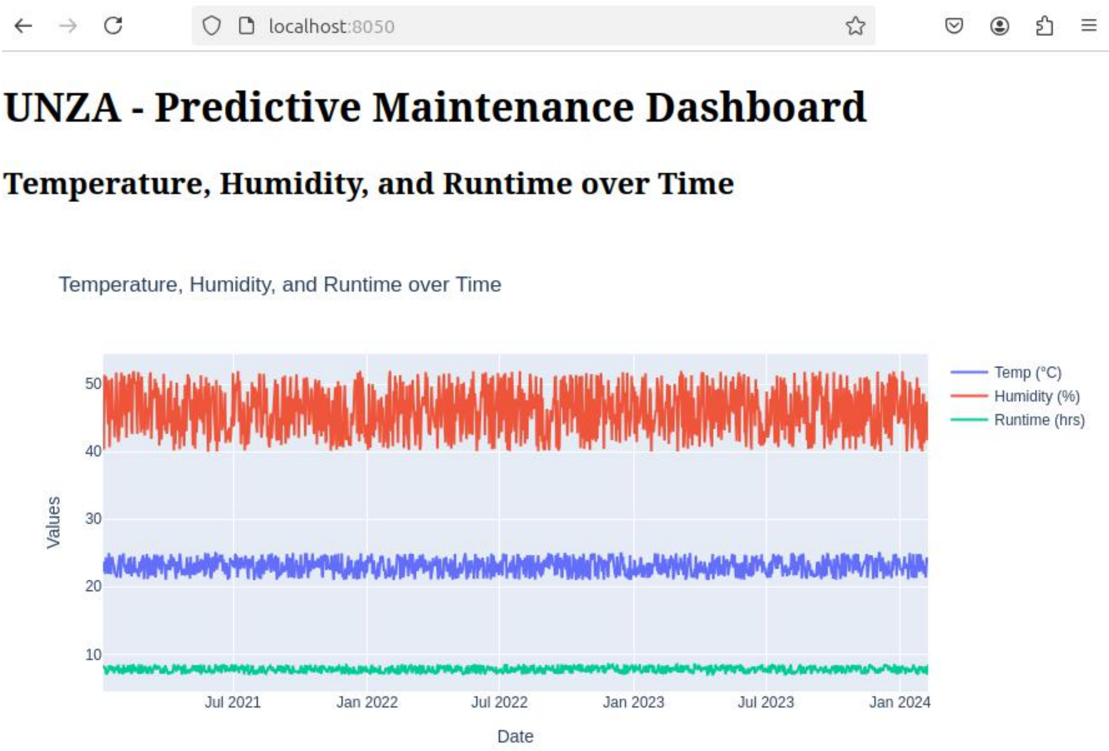
A dashboard is one of the most valuable tools in the context of predictive maintenance since it is a single console from which one can monitor and analyze the most important metrics, insights, and predictions relevant to the management and improvement of assets' conditions. There are several excellent libraries for data visualization including Matplotlib, Seaborn, Dash, and Bokeh. However, for this research study project Plotly was used because it offered a wide range of highly interactive charts like scatter plots, 3D plots and maps and could easily be incorporated with dashboards and was very easy to use even with a little coding. As can be seen below, the figure is labelled as Figure dashbord 1

Table 3 Confusion matrix

Dashboard on the browser interface

The dashboard for UNZA Predictive Model can be accessed via the ULR http://127.0.0.1:8050/

Figure dashboard 1



Code implementation of the dashboard with Plotly library

Code start here

```
import pandas as pd
import numpy as np
import dash
from dash import dcc, html
from dash.dependencies import Input, Output
```

```
import plotly.express as px
import plotly.graph_objects as go
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from xgboost import XGBClassifier
from sklearn.metrics import accuracy_score, classification_report

# Load the dataset
df = pd.read_csv('pdm_1500.csv')
df['Date'] = pd.to_datetime(df['Date'])
df.set_index('Date', inplace=True)

# Create lag features
def create_lag_features(df, columns, lags=3):
    for col in columns:
        for lag in range(1, lags + 1):
            df[f'{col}_lag_{lag}'] = df[col].shift(lag)
    df.dropna(inplace=True)

create_lag_features(df, columns=["Temp (°C)", "Humidity (%)", "Runtime (hrs)"])

# Split the data
X = df.drop(columns=["Failure (0/1)"])
y = df["Failure (0/1)"]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, shuffle=False)

# Scale the data
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
```

```
# Train the model
model = XGBClassifier()
model.fit(X_train_scaled, y_train)
y_pred = model.predict(X_test_scaled)

# Evaluation metrics
accuracy = accuracy_score(y_test, y_pred)
report = classification_report(y_test, y_pred, output_dict=True)

# Initialize Dash app
app = dash.Dash(__name__)

# Layout of the dashboard
app.layout = html.Div([
    html.H1("UNZA - Predictive Maintenance Dashboard"),

    # Line Chart Section
    html.Div([
        html.H2("Temperature, Humidity, and Runtime over Time"),
        dcc.Graph(id='line-chart')
    ]),

    # Scatter Plot Section
    html.Div([
        html.H2("Lagged Temperature vs. Runtime"),
        dcc.Graph(id='scatter-plot')
    ]),

    # Model Metrics Section
```

```
    html.Div([
        html.H2("Model Metrics"),
        html.P(f"Accuracy: {accuracy:.2f}"),
        html.Pre(classification_report(y_test, y_pred))
    ])
])

# Callback for the line chart
@app.callback(
    Output('line-chart', 'figure'),
    Input('line-chart', 'id')
)
def update_line_chart(_):
    # Line chart of Temperature, Humidity, and Runtime over time
    fig_line = go.Figure()
    fig_line.add_trace(go.Scatter(x=df.index, y=df["Temp (°C)"], mode='lines', name='Temp (°C)'))
    fig_line.add_trace(go.Scatter(x=df.index, y=df["Humidity (%)"], mode='lines', name='Humidity (%)'))
    fig_line.add_trace(go.Scatter(x=df.index, y=df["Runtime (hrs)"], mode='lines', name='Runtime (hrs)'))
    fig_line.update_layout(title="Temperature, Humidity, and Runtime over Time",
                           xaxis_title="Date", yaxis_title="Values")
    return fig_line

# Callback for the scatter plot
@app.callback(
    Output('scatter-plot', 'figure'),
    Input('scatter-plot', 'id')
)
def update_scatter_plot(_):
```

```
    # Scatter plot of Temp_lag_1 vs Runtime for visualizing lag effect
    fig_scatter = px.scatter(df, x="Temp_lag_1", y="Runtime (hrs)",
                             color="Failure (0/1)",
                             title="Lagged Temperature vs. Runtime",
                             labels={"Temp_lag_1": "Temperature Lag 1", "Runtime (hrs)": "Runtime (hrs)"})
    return fig_scatter


# Run the app
if __name__ == '__main__':
    app.run_server(debug=True)
```

Ends here

4.3.13 API implementation

There are two APIs for data validation in the Job Card System (JCS) that have been developed and implemented: The first one for the staff which is integrated with the Performance Management System (PMS) and the second one for the students which is linked to the Student Information System (SIS). These APIs help in validating data at the time of account creation as well as at the time of submitting job card request.

4.3.13.1 API for staff

This API endpoint is subscribed to the Job Card System (JCS) login process in order to ensure that the staff member is currently active in the Permanence Management System (PMS) of the university. It determines whether the staff member is currently employed by the organization or not and gives a response of (1) for yes the staff member is active and (0) for no the staff member is not active. In case the staff member is inactive the system will block him from logging in. .

API Endpoint

The endpoint below shows the REST GET API for validating active university employees.

GET http://127.0.0.1:8001/checkemployee?employee_id=11480

If active the json response is:

{
“employee_number": "11480",
"status": "error",
"message": "Employee does not exist",
"active": "0"
}

If inactive the json response is:

{
"employee_number": "11480",
"status": "error",
"message": "Employee does not exist",
"active": "0"
}

If Not Found:

{
"employee_number": "11480",
"status": "error",
"message": "Employee does not exist",
"active": "0"
}

If employee_number Missing:

{

"error": "employee_number is required"

}

Postman

A free online tool for developing and testing API calls. Figure 2 PMS Success response shows success response While Figure 3 PMS error response shows the error response.

Figure 2 PMS Success response

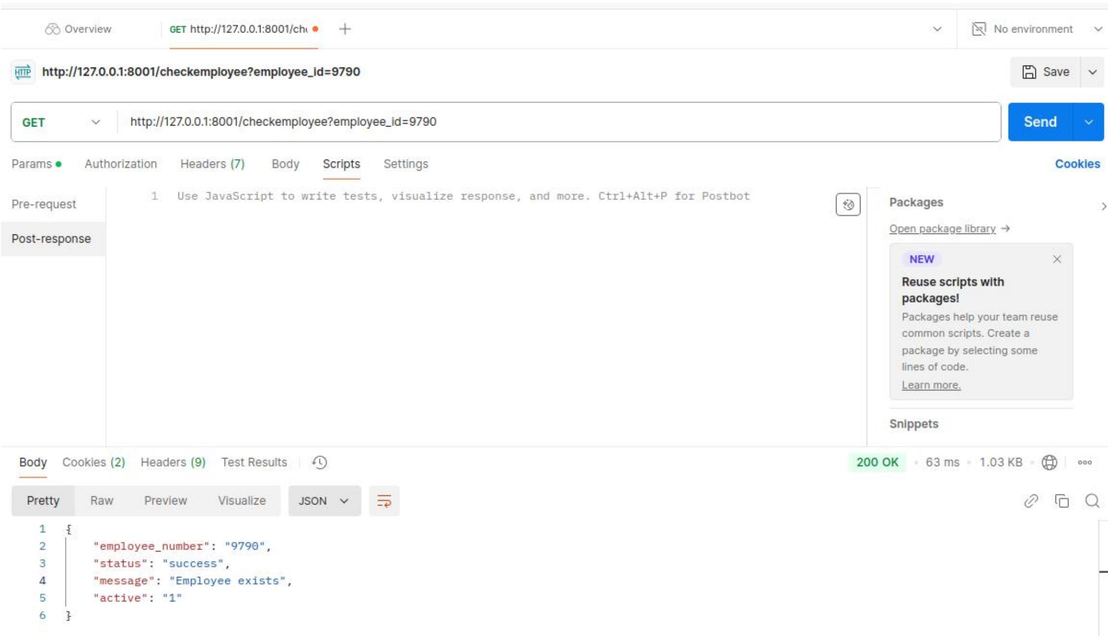
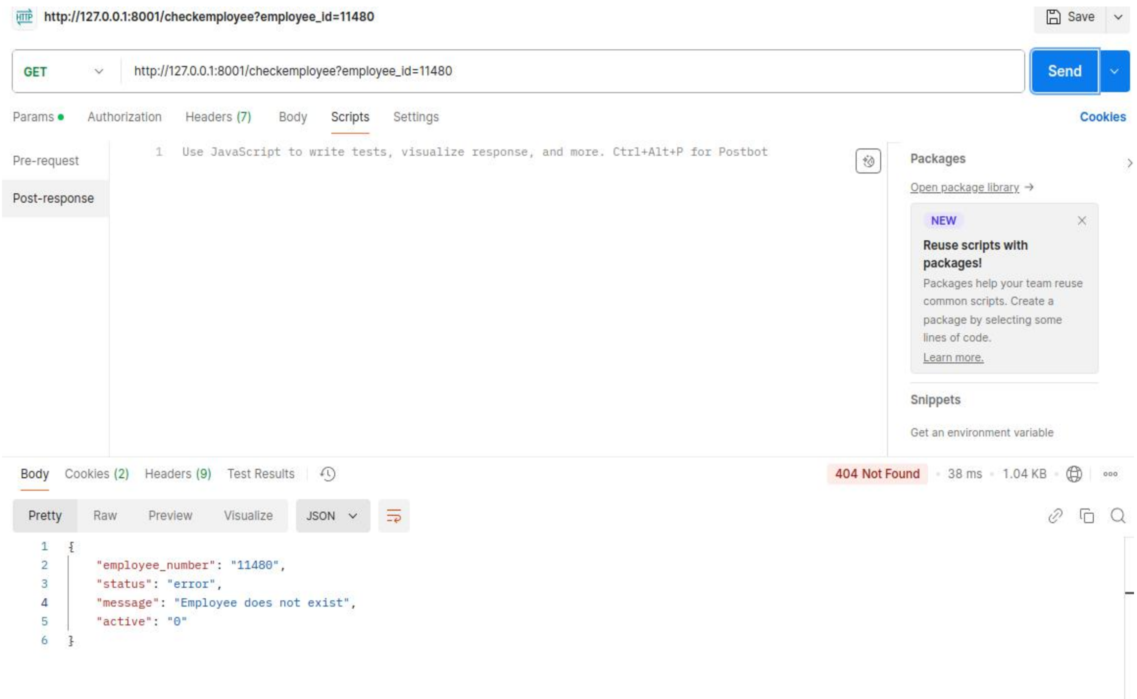


Figure 3 PMS error response



Code Implementation

```php
//=====================
<?php
namespace App\Http\Controllers;
use App\Http\Controllers\Controller;
use Illuminate\Foundation\Auth\AuthenticatesUsers;
use Illuminate\Http\Request;
use Illuminate\Support\Facades\DB;
use Auth;

class LoginController extends Controller
{
    use AuthenticatesUsers;
    protected $username = 'username';
    protected $redirectTo = '/dashboard';
    protected $guard = 'web';
```

```
public function checkemployee(Request $request){
     // Get the student ID from the request
    $employeeId = 0;
    $employeeId = $request->query('employee_id');


    $exists = DB::table('users')->where([
        ['man_number', '=', $employeeId],
        ['status', '=', 1]
    ])->first();


    // Return the JSON response
     if ($exists) {
         return response()->json(['employee_number' => $employeeId,'status'
                 => 'success', 'message' => 'Employee exists','active'=>'1'], 200);
     } else {
         return response()->json(['employee_number' => $employeeId,'status'
                 => 'error', 'message' => 'Employee does not exist','active'=>'0'],
                              404);
     }
 }
```
=====================

4.3.13.1 API for students

This API endpoint is subscribed to the JCS login process to ensure that the student record is still valid in the university Students Information Systems (SIS). It determines whether the student is enrolled at the moment and replies (1) for yes and (0) for no. In case the student is not enrolled the system will block him from logging in.

API Endpoint

The endpoint below shows the REST GET API for validating active university students.

GET http://127.0.0.1:8002/checkstudent?student_id=2020000288

If active the json response is:

```
{
   "student_number": "2020000288",
```

"status": "success",
"message": "Student exists",
"active": "1"
}

If inactive the json response is:

{
"student_number": "2020000281",
"status": "error",
"message": "Student does not exist",
"active": "0"
}

If Not Found

{
"student_number": "2020000281",
"status": "error",
"message": "Student does not exist",
"active": "0"
}

If employee_number Missing:

{
"error": "employee_number is required"
}

Using Postman

Figure 4 SIS (Student Information System) success response show the output once a call is made on the endpoint while Figure 5 SIS error response shows the output once no active student record has been found.

Figure 4 SIS success response

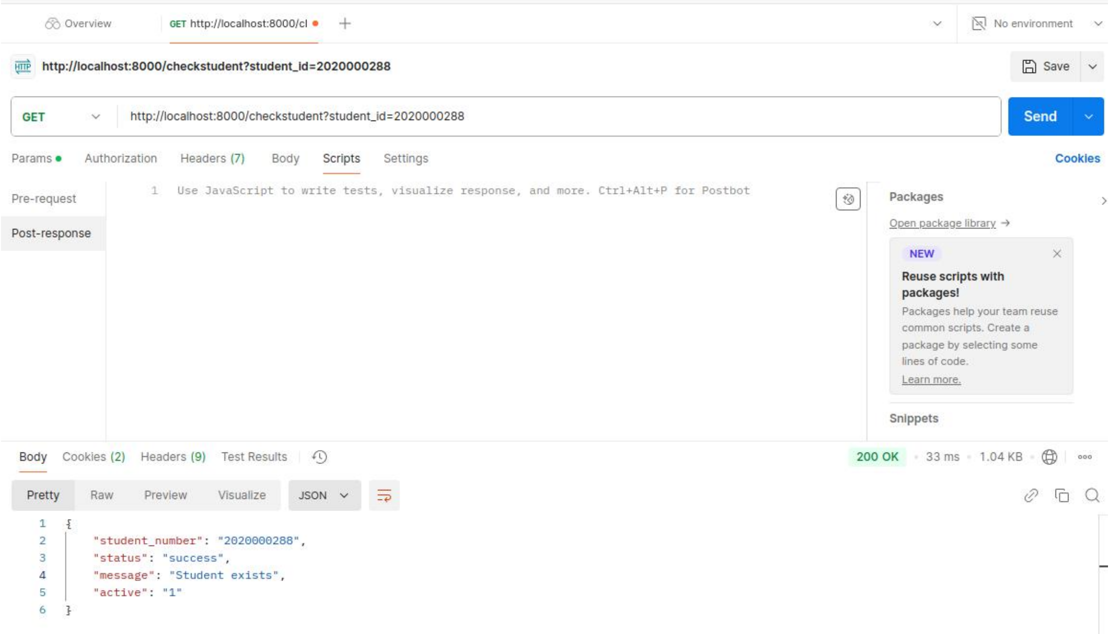
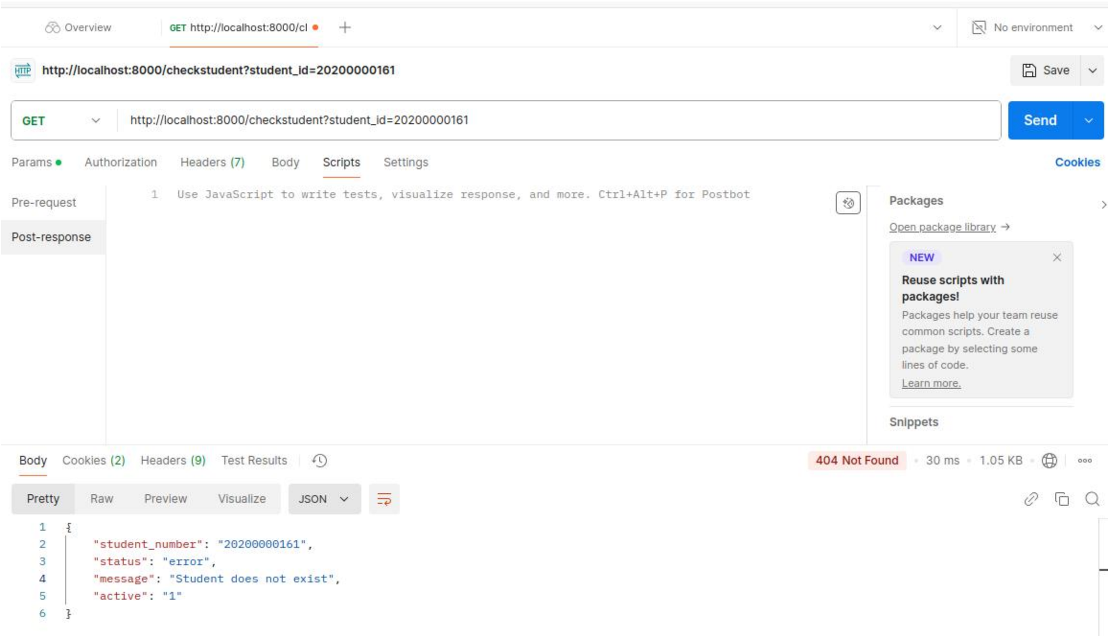


Figure 5 SIS error response



Code Implementation

```
<?php

namespace App\Http\Controllers;
```

```
use App\Http\Controllers\Controller;
use Illuminate\Foundation\Auth\AuthenticatesUsers;
use Illuminate\Http\Request;
use Illuminate\Support\Facades\DB;
use Auth;

class LoginController extends Controller
{
    use AuthenticatesUsers;
    protected $username = 'username';
    protected $redirectTo = '/dashboard';
    protected $guard = 'web';

    public function checkstudent(Request $request){
         // Get the student ID from the request
        $studentId = 0;
        $studentId = $request->query('student_id');

        //$exists = DB::table('users')->where('username', $studentId)->exists();

        $exists = DB::table('users')->where([
            ['username', '=', $studentId],
            ['status_id', '=', 1]
        ])->first();

         // Return the JSON response
         if ($exists) {
             return response()->json(['student_number' => $studentId,'status' => 'success',
'message' => 'Student exists','active'=>'1'], 200);
         } else {
```

```
            return response()->json(['student_number' => $studentId,'status' => 'error', 'message'
=> 'Student does not exist','active'=>'0'], 404);
        }
    }
//=============code ends here================
```

4.3.12 Summary

This chapter has critically assessed various data processing techniques, research studies, evaluation measures and the development of a Predictive Model for an Office Air Conditioning Systems (OACS), APIs for validating both student and as staff members as well as a Dashboard to assist with decision making; the next chapter will hence provide a detailed description of the findings. This segment will also look at the significance of these findings for the industry and contribution towards advancing knowledge and practical applications in the dynamic field of Predictive Modeling for OAC systems.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.0 Introduction

The research study and development of a Predictive Model (PM) for Office Air Conditioning (OAC) systems are important as it can be integrated with real-time data and effective dashboards to enhance the effectiveness of OAC systems. This charter expands on the findings and discussion of the study research project. project.

5.1.1 Enhanced Operational Efficiency

In relation to this study, the integration of Predictive Model for OAC systems in the University of Zambia can help prevent unexpected failures thus reducing time out of service. This is because predictive models enable the scheduling of maintenance based on real time wear data and performance data as opposed to traditional maintenance scheduled, this is a cost effective as it as maintenance repairs are done on OAC units that needs repairs. This is in consistent with Wang and (Gao, 2022), where the benefits of predictive maintenance have been emphasized.

5.1.2 Real-Time Data Validation

In this research study, the APIs for validating student and staff data are incorporated in order to ensure that the maintenance predictions for the University of Zambia's AC systems are relevant to specific user-specific data. This makes the model to be able to alter the maintenance schedule according to the occupancy rates and usage rates in different buildings of the university. In addition, the real-time data integration enables dynamic re-calibration of the predictive model where the model is updated continuously to learn from the actual usage tendencies. This approach makes sure that the model is adequate for the needs of the university.

5.1.3 Informed Decision-Making Through Dashboards

In the context of this research study, the usage of dashboards provides facilities managers at the University of Zambia a way of viewing summaries of key metrics which are related to the AC systems on a graphical user interface. This developed dashboard enables facility engineers to make faster and better decisions based on the data that is provided by the system in real-time. By presenting the data in a more manageable and easily understood manner, the dashboard tool enhances the utility of predictive maintenance and makes it feasible for individuals who may not be very conversant with technical details to deal with it effectively. This is in consonance with (Franklin et al, 2017), who stressed on the importance of dashboards in facilitating analysis of data for effective maintenance management.

5.1.4 Cost Savings and Resource Efficiency

The Predictive Model (PM)that has been put in place at the University of Zambia's AC systems helps in determining accurate maintenance actions that should be taken, minimizing on unnecessary service calls and optimizing the use of maintenance personnel. This approach is effective in managing costs, which is a crucial factor especially in large institutions that have several AC units to maintain. Through minimizing inefficiencies, the model that has been developed ensures that maintenance resources are targeted to where they are most required so as to enhance overall effectiveness of the system.

5.2.0 Comparison with Related Research

There have been numerous research works on predictive maintenance for OAC systems which have embraced various concepts yet differ in approaches, data sets, and objectives. Here is a comparison with related works:

5.2.1 Machine Learning Models

Most of the studies apply machine learning models to AC predictive maintenance, and commonly used algorithms include Random Forest, Support Vector Machines, and a number of Deep Learning models. A few of the more sophisticated models incorporate deep learning models such as LSTMs for handling time series data in an attempt to capture more complex patterns in AC usage and operational data. In this research, the selected model, XGBoost, was adopted as the model of choice because of high accuracy, low complexity and efficient computation that would be perfect for real time, resource limited maintenance schedules (Kavzoglu & Teke, 2022).

5.2.2 Data Sources and Real-Time Validation

Different from many studies that rely solely on static datasets, the integration of real-time validation via APIs for students and staff at the University of Zambia marks a notable technological innovation in predictive maintenance modeling. Normally predictive maintenance models mainly depend on historical data, which can limit the ability to change to dynamic changes. In contrast, real-time validation enables the model to adjust quickly to changes in occupancy and usage patterns, enhancing both accuracy and responsiveness. This approach reconciles with findings from studies by (Ahmed et al, 2022), which emphasize the importance of incorporating real-time data streams for more adaptive and accurate predictions. By dynamically updating predictions, the model gives a tailored and up-to-date solution, moving far above the limitations of static historical data. Similar improvements in smart building management systems, were researched and studied by (Lin and Tsai, 2021), demonstrated how real-time data integration promotes more accurate, context-aware decision-making. The University of Zambia's model implementated exemplifies a modern and innovative approach, provides a precedent for more smooth and practical predictive maintenance frameworks.

5.2.3 Dashboards and Decision-Making

Not every research paper focusing on the predictive maintenance system discusses dashboards or the decision-making interface as one of the crucial components. It is important to note that dashboards can significantly affect the user engagement and practically, which are not often discussed in technical studies that focus only on the model performance. This aspect makes this study unique as it provides a comprehensive approach from data acquisition to decision-making thus enabling facility managers and engineers to understand and manage maintenance effectively.

5.2.4 Comparison in Evaluation Metrics

The Predictive Maintenance Model (PMM) developed at the University of Zambia has been assessed by a number of performance measures including precision, recall, F1-score, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) by Sathya and Sathya (2023). This is to enable assessment of the model's accuracy, reliability and flexibility when presented with different conditions. This approach is consistent with previous research on the topic of predictive modeling where the same or similar metrics have been employed in order to analyse and optimize models with a view to their application in the real world. For example, research in the area of industrial predictive maintenance has stressed on the utilization of classification metrics such as precision and recall along with regression metrics such as RMSE and MAE in order to determine the overall correctness of classifications as well as the precision of numerical estimates. The evaluations discussed in work by (Zhang et al., 2021) and (Gupta et al.,2020) help the researchers to identify the further improvement in the model and make the model more reliable for practical application.

5.2.5 Limitations and Future Scope

Certain research works have been found to have some limitations concerning the Predictive Maintenance Models' effectiveness in predicting failures in various OAC system types and for multiple and complex climate conditions. If this University of Zambia predictive model performs well across different usage patterns and environmental conditions, it may indicate that the model is more robust than the more rigid models. Future research can also attempt to establish enhancements such as incorporating weather data or developing validation heuristics with a view of improving the model's reliability and flexibility.

5.3.0 Broader Discussion and Future Directions

The integration of the University of Zambia's predictive model AC system with real-time validation and the implementation of dashboard based decision making enhances the technology aspect of the current OAC maintenance through improving the cost effectiveness, efficiency and responsiveness of the OAC maintenance through linking the services offered to the actual usage patterns. This new approach does not only enhance on proper utilization of resources but also reduces on time spent off line and increases the life expectancy of OAC. Future research could focus on the following areas:

5.3.1 Scalability

The University of Zambia's developed PMM for OACU has not been fully tested on various structures in various locations throughout the world including areas with different climate conditions. Carrying the testing of the PMM out across the a large number of structures, localities and environmental conditions would be very useful in determining the reliability of the results that would be obtained from the use of this model. The findings of such assessments would be useful in determining the model's capacity to make sound conclusions across different OAC systems, structures and climate regions, thus offering a better analysis of the effectiveness of the model. Such a detailed testing could reveal possible improvements that need to be made, which would in turn lead to focused improvements to enhance the model's flexibility, effectiveness, and dependability. Finally, this process would help in creation of a strong tool which would be able to fulfill the requirements of many environments of the world.

5.3.2 Integration with IoT Sensors

One of the notable challenge that can be identified with University of Zambia's PMM for OAC systems is the lack of Internet of Things (IoT) sensors which are capable of capturing real time data. These sensors are not integrated into the model which makes it difficult for the model to capture dynamic, high resolution data that could help enhance the accuracy of the predictions and performance of the system. The integration of IoT sensors into the University of Zambia's PMM OAC system would allow for the gathering of real time data on parameters such as temperature, humidity, energy usage and air quality. This wealth of data would enable the model to make much more comprehensive and accurate as well as timely and precise maintenance predictions. Monitoring of these important variables on a constant basis would not only enable the model to identify the variations in the performance but also to forecast the potential failures beforehand. Also, such technological improvements would enhance energy efficiency, enhance system reliability and increase the lifespan of OAC units and hence change the maintenance approach to be more smart and data-driven enabled.

5.3.4 Summary

This chapter has detailed the findings of the research study project and the analyses that have been made of the findings as well as the comparison with other OAC Predictive Maintenance research studies that have been conducted across the globe. It identifies and elaborates on the specific contributions of this research project to the existing knowledge. It also focuses on improvements in the predictive model's precision as well as its real-world implementation. Also, it analyses the limitations of the project and proposes further research ideas, including the enhancement of the model flexibility and inclusion of more environmental and usage factors. The final chapter includes the summary and conclusion of the research project, including the key findings, the contribution of the research to the existing literature, the limitations of the research and the future research directions.

CHAPTER 6

SUMMARY AND CONCLUSIONS

6.0 Introduction

The University of Zambia Predictive maintenance for Office Air Conditioning (OAC) systems offers notable benefits to the university by forecasting maintenance needs based on actual equipment condition and usage patterns rather than relying on fixed maintenance schedules. This chapter provides summary of findings and conclusions on the research project.

6.1 Summary of main Findings

The research project on predictive maintenance for Office Air Conditioning (OAC) system at the University of Zambia offers a lot of key findings. It demonstrates notable benefits by predicting maintenance needs based on real-time equipment conditions and usage patterns rather than using fixed schedules.

6.1.1 Reduced Downtime

The University of Zambia Predictive maintenance helps minimize unexpected equipment failure by identifying early signs of wear ahead of time. By using data and reviews, Maintenance mangers and engineers can schedule maintenance during perfect periods thereby, preventing costly break downs. This proactive approach not only lessens AC system downtime but also ensures consistent and dependable cooling, creating a suitable and productive environment for office spaces. Furthermore, predictive maintenance improves AC lifespan, enhances energy efficiency, and promotes sustainability goals, making it an indispensable strategy for present day facility management (Jasiulewicz-Kaczmarek et al, 2017).

6.1.2 Improved Energy Efficiency

By monitoring real-time data collected, such as energy consumption, temperature, and humidity levels, the University of Zambia predictive maintenance allows early detection of unfitness and possible potential failures, including clogged filters, refrigerant leaks, or usual system behavior. This proactive approach alerts maintenance managers and engineers promptly, allowing them to address issues before they grow rapidly. Ensuring the ideal performance of AC systems not only lessen energy usage but also minimizes wear on equipment, reduced operational costs, lengthens equipment lifespan, and aligns with sustainability initiatives. Eventually, predictive maintenance plays an important role in predicting energy-efficient, cost-effective, and environment-friendly building operations (Das & Chandra, 2023).

6.1.3 Cost Savings

The other finding on the University of Zambia Predictive maintenance is that it lessens unnecessary repairs and part replacements by making sure that repairs only happens when truly required. This kind of approach eliminates over-maintenance, reduces waste, and improves resource utilization. Over time, noble cost savings are achieved by prolonging the equipment lifespan, improving system dependability, and preventing costly emergency repairs or unexpected breakdowns. Not only that, by addressing issues early, predictive maintenance helps in enhancing operational efficiency, lessens disruption, and provides a more sustainable approach to AC management.

6.1.4 Enhanced Comfort

The University of Zambia OAC predictive maintenance system offers greater reliability, ensures consistent indoor conditions that improves comfort for University of Zambia office users. By reducing temperature fluctuations and maintaining improved air quality, these system creates a more enjoyable and stable working environment. The other finding is that not only does it only improves employee satisfaction but also promotes higher productivity, as workers are less likely to be diverted by discomfort caused by system break down.

Furthermore, reliable AC performance mirrors positively on facility management, showcasing a dedication to maintaining a comfortable and effective workplace.

6.1.5 Data-Driven Decision Making

The University of Zambia OAC system predictive maintenance offers a dashboard interface that provides a clear and brief visualization of key metrics. The finding is that this dashboard enables maintenance managers to make well-informed decisions quickly and efficiently. The dashboard integrate both historical and real-time data obtained directly from the AC systems, such as energy consumption, temperature trends, and others AC system performance indicators. By providing actionable insights at a glance, dashboards empower maintenance managers to identify possible failures, prioritize maintenance tasks, and enhance operations for improved reliability, cost savings, and energy efficiency (Segun-Falade, et al, 2024).

6.1.6 Challenges in Implementation

While the research study on the University of Zambia on predictive maintenance for OAC offers noble benefits, it comes with initial challenges, including investments in sensors, data integration, and specialized training for facility maintenance teams. In addition, predictive models require customization to accommodate different office settings, OAC system types, and external factors such as climate variations or occupation patterns. These factors influences the OAC system behavior and model accuracy. Despite these hurdles, predictive maintenance demonstrates vast potential to improve efficiency, lessen operational costs, and enhance sustainability goals. Its successful implementation demands a cautious balance of technical expertise, monetary resources, and ongoing model adjustments to ensure its effectiveness across diverse environments.

6.2.0 Contribution to the body of knowledge

The research on predictive maintenance for the University of Zambia for OAC systems contributes to the enhancement of the knowledge base in several ways. It increases the current knowledge on how real time information and analysis can enhance AC performance, offers insights on how energy consumption and operational costs can be cut down and looks at maintenance strategies that support sustainability. In addition, it offers useful guidelines for the development of predictive maintenance plans for various office settings, including the issues of customization, sensor integration, and model adaptation, as well as outlining the long term advantages of effectiveness and comfort in the workplace.

6.2.1 Enhanced Predictive Modeling Techniques

The research study on Predictive Maintenance Model for Office Air Conditioning (OAC) systems at the University of Zambia has been performed by applying advanced machine learning algorithms including XGBoost, Random Forest and Long Short-Term Memory (LSTM) networks. The study has used these sophisticated models and has helped in enhancing the accuracy of predicting possible OAC failures. These models are capable of recognizing the complex patterns in the operational data especially the complex time series data that is captured using sensors. This approach does not only enhance the accuracy and reliability of the maintenance predictions but also facilitate proactive decision-making, reduces or eliminates unexpected downtime, increases energy efficiency and extends the lifetime of OAC systems.

6.2.2 Impact on Energy Efficiency and Sustainability

The study on the predictive maintenance for office OAC system at the University of Zambia reveals that it has great possibilities of reducing energy consumption in buildings. This research study adds to the understanding on how predictive maintenance works in identifying and solving potential failures or breakdowns in OAC systems before they surge into major energy consumption.By addressing these concerns ahead of time, Predictive Maintenance not only reduces power consumption but also contributes to efficient management of structures. The study also complies with the current global initiatives for energy conservation, reduction of carbon footprint, and adherence to environmental standards in the pursuit of cost effective and environmentally sound structures (Li, et al., 2017).

6.2.3 Development of Decision-Support Tools

The study on Predictive Maintenance Model for OAC system at University of Zambia found that the inbuilt dashboard and visualization tools are crucial. These tools which are developed in conjunction with the Predictive Maintenance Model make it easy for facilities managers and engineers to make the right decisions. By transforming the complex technical information into simple, useful guidelines, the dashboard enhances the usability and simplicity of the Predictive Maintenance Model. This approach makes it possible to implement preventive maintenance and ensure that system health is monitored, potential failures are identified, and maintenance tasks are properly scheduled without the need for extensive technical knowledge. Also, it enables real time tracking, optimized utilization of resources and proper maintenance processes and thus supports sustainable building operations.

6.2.4 Cost-Benefit Analysis of Predictive Maintenance Implementation

The research on predictive maintenance for OAC system at the University of Zambia highlights notable financial gains through a detailed cost-benefit analysis. This research study demonstrates the clear economic gain of predictive maintenance over conventional reactive or preventive approaches. By reducing unexpected emergency breakdowns or repairs, prolonging the lifespan of equipment, and optimizing energy usage, predictive maintenance comes out as a highly cost-effective solution. These findings provide practical insights for stakeholders in facilities management and the OAC industry, underscoring not only immediate cost savings but also long-term economic energy efficient. Not only that, this approach also minimizes operational disruptions, enhancing overall efficiency and resource utilization (Achouch, et al, 2022).

6.2.5 Future Directions in Predictive Maintenance Research

The study on predictive maintenance for the OAC system that was conducted at the University of Zambia revealed some potential areas for improvement such as the use of other data sources like weather conditions, occupancy indices and seasonal changes. In addition, the study emphasizes the need of utilizing more complex hyperparameter tuning methods for the development of more stable and efficient predictive models. Through these aspects, the research on predicitive maintenance provides a strong basis for future study on the enhancement of the predictive maintenance techniques, and hence improve the performance of the models used under various operating conditions. Not only do these enhancements strengthen the model's reliability but they also enhance the efficiency, sustainability, and cost efficiency of OAC systems. In general, the study offers valuable information and applied instruments to the field and encourages advancement and enhanced approaches to building management.

6.3.0 Limitations of the Research Study Project

The study on predictive maintenance for Office Air Conditioning (OAC) systems at the University of Zambia faced certain limitations. These limitations were specifically related to the investigation of predictive maintenance for OAC systems and included:

6.3.1 Data Quality and Availability

The University of Zambia Predictive maintenance model is closely associated with data that it needs quality and granularity of the data. However, there were challenges that included; inadequate or irregular data from OAC systems particularly because of the absence of sensors and connection to a central server. As a result dataset used in similar research study on PMM was used to fill up the gap that was indentified during the execution of the research study. This was download from a free dataset used source for Machine Learning www.kaggle.com. Also, the shortage of historical data reduced the effectiveness and accuracy of machine learning algorithms,PMM requires substantial datasets to correctly indentity patterns accurrately.

6.3.2. Generalizability Across Different AC System Types

The research on predictive maintenance carried out at the University of Zambia may not be applicable in its entirety to other OAC systems. Air conditioning units come in different models and designs, and have different components and maintenance requirements, thus it is difficult for models made for one type of air conditioning unit to work well with another different type of air conditioning unit whether it is from the same brand or different brand and model. This variance hinders the scalability and efficacy of predictive maintenance in various office settings. In order to address these limitations, it is imperative that there is further fine tuning and frequent updating of the models. Thus, the inclusion of system-specific data and the use of advanced algorithms can help to improve the models' flexibility and therefore provide better results in a more general context (Henrichs et al., 2022).

6.3.3 Sensitivity to Environmental Factors

The research study on predictive maintenance was conducted at the University of Zambia and was developed with the environment of Zambia in consideration such as temperature, humidity and climate variation. These factors greatly affect the OAC system but are sometimes hard to incorporate into a predictive model. Model made for one region can be of little use in another region without being updated. Climate differences can have an effect on the workload, effectiveness, and the possibility of failure of OAC systems which means that the predictions would need to be re-calibrated specifically for the system to be accurate. In addition, variations in seasonal changes, levels of dust and air quality also impacts the system and the requirements for maintenance hence underlining the need to incorporate local data when developing predictive models for their generality (Saleh et al., 2022).

6.3.4 Limited Consideration of External Influences

Due to constraints in time, the research study on predictive maintenance that was conducted at the University of Zambia did not consider other factors that could have influenced the outcome of the study. Such extrinsic factors as occupancy rate of the building, weather conditions, and the costs of energy and utilities are difficult to incorporate into the predictive maintenance models. However, these factors are crucial in determining OAC utilization and when the need for maintenance should be made..

6.4.0 Future works

This section of the paper suggests potential future research studies that can be conducted in order to enhance the OAC predictive model. Such research studies could aim at improving the accuracy of the model, enhancing its performance, or increasing its scalability. The following are some of the research study opportunities include;

6.4.1 Cost-Benefit Analysis for Widespread Implementation

The research study on OAC predictive maintenance model could be extended to examine other factors that could influence the outcome in future research studies. Future research could also examine the economic efficiency of predictive maintenance on the maintenance and replacement of HVAC systems for various buildings, sizes and usage, thus providing a better understanding of the Return on Investment (ROI) in various contexts. This type of model could also be developed to analyze the costs and the benefits of such a predictive maintenance plan for OAC in order to determine whether the benefits are worth the investment for both small and large organizations. The following models could also help in determining the right time for maintenance, effective utilization of resources and enhancing the overall efficiency of the system.

6.4.2 Sustainability and Environmental Impact Assessment

The research study on OAC predictive maintenance executed at the University of Zambia did not address the long-term effects of PM. Future research studies should examine these impacts, drawing observation to sustainability by assessing metrics like reduced carbon discharge, energy efficiency, cost saving effectiveness, and extended or prolonged OAC system lifespan. This study could align predictive maintenance plans with global sustainability goals, giving insights into ecological, economic, and operational profits, thereby improving the relevance of predictive maintenance in modern buildings.

6.4.3 Improved Hyperparameter Optimization Techniques

In the study on predictive maintenance conducted at the University of Zambia, improved hyperparameter optimization was not implemented due to constraints in time. Nevertheless, it is beyond doubt that improved hyperparameter optimization is a crucial factor that contributes towards the enhancement of model performance to the optimal level by adjusting the parameters in accordance with the particular characteristics of the data set. Future research may also look at other forms of enhanced automated optimization methods like the Bayesian optimization or the grid search to enhance the model's precision and performance. These techniques can improve model flexibility and stability making them suitable for use in various settings and at the same time reducing the dependence on manual tuning. Focusing on this aspect in future studies can lead to more successful and scalable predictive maintenance.

6.5.0 Summary

This chapter has systematically summarise the research study of OAC predictive maintenance model for University of Zambia and thus presents the main conclusions of the study as well as contributions of the study to the existing body of knowledge in several meaningful ways. It also identifies possible areas for future research studies on predictive maintenance for air conditioning OAC systems as well. These suggestions may help to enhance the model's precision, effectiveness, and sustainability, which will be useful for both academic and practical applications in the area of predictive maintenance systems.

REFERENCES

Reference Type	In-text	Reference list
Article without DOI from research databases	(Abdel-Razek et al., 2022)	S. A. Abdel-Razek, H. S. Marie, A. Alshehri, and O. M. Elzeki, "Energy efficiency through the implementation of an AI model to predict room occupancy based on thermal comfort parameters," Sustainability, vol. 14, no. 13, p. 7734, 2022.
Article without DOI from research databases	(Achouch et al., 2022)	Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in industry 4.0: Overview, models, and challenges. Applied Sciences, 12(16), 8081.
Article without DOI from research databases	(Ayvaz et al., 2021)	Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. Expert Systems with Applications, 173, 114598.
Article without DOI from research databases	(Ayvaz et al., 2024)	Azevedo, B. F., Rocha, A. M. A., & Pereira, A. I. (2024). Hybrid approaches to optimization and machine learning methods: a systematic literature review. Machine Learning, 1-43.
Article without DOI from research databases	(Ayvaz et al., 2023)	Albayati, M. G., Faraj, J., Thompson, A., Patil, P., Gorthala, R., & Rajasekaran, S. (2023). Semi-supervised machine learning for fault detection and diagnosis of a rooftop unit. Big Data Mining and Analytics, 6(2), 170-184.
Article without DOI from research	(Bampoula et al., 2022)	X. Bampoula, G. Siaterlis, N. Nikolakis, and K. Alexopoulos, "A deep learning model for predictive

databases		maintenance in cyber-physical production systems using LSTM autoencoders," Sensors, vol. 21, no. 3, p. 972, 2021.
Article without DOI from research databases	(Ben-Daya et al., 2012)	Ben-Daya, M., Duffuaa, S. O., & Raouf, A. (Eds.). (2012). Maintenance, modeling and optimization. Springer Science & Business Media.
Article without DOI from research databases	(Bezerra et al., 2019)	A. Bezerra, I. Silva, L. A. Guedes, D. Silva, G. Leitão, and K. Saito, "Extracting value from industrial alarms and events: A data-driven approach based on exploratory data analysis," Sensors, vol. 19, no. 12, p. 2772, 2019.
Article without DOI from research databases	(Black et al., 2022)	Black, E., Raghavan, M., & Barocas, S. (2022, June). Model multiplicity: Opportunities, concerns, and solutions. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (pp. 850-863).
Article without DOI from research databases	(Blischke et al.,2011)	Blischke, W. R., & Murthy, D. P. (2011). Reliability: modeling, prediction, and optimization. John Wiley & Sons.
Article without DOI from research databases	(Bouabdallaoui et al.,2021)	Bouabdallaoui, Y., Lafhaj, Z., Yim, P., Ducoulombier, L., & Bennadji, B. (2021). Predictive maintenance in building facilities: A machine learning-based approach. Sensors, 21(4), 1044.
Article without DOI from research databases	(Cachada et al.,2018)	Cachada, A., Barbosa, J., Leitño, P., Gcraldcs, C. A., Deusdado, L., Costa, J., ... & Romero, L. (2018, September). Maintenance 4.0: Intelligent and predictive maintenance system architecture. In 2018 IEEE 23rd international conference on emerging technologies and factory automation (ETFA) (Vol. 1, pp. 139-146). IEEE.
Article without DOI from research	(Capra et al, 2020)	M. Capra, B. Bussolino, A. Marchisio, G. Masera, M. Martina, and M. Shafique, "Hardware and software

databases		optimizations for accelerating deep neural networks: Survey of current trends, challenges, and the road ahead," IEEE Access, vol. 8, pp. 225134-225180, 2020.
Article without DOI from research databases	(Capra et al, 2020)	M. Capra, B. Bussolino, A. Marchisio, G. Masera, M. Martina, and M. Shafique, "Hardware and software optimizations for accelerating deep neural networks: Survey of current trends, challenges, and the road ahead," IEEE Access, vol. 8, pp. 225134-225180, 2020.
Article without DOI from research databases	(Cheekati et al, 2020)	Cheekati, V., Prasad, V. N., Prasad, K. D. V., Ali, S. M., & Tarigonda, H. (2024, May). IoT-Driven Predictive Maintenance for Energy-Efficient Industrial Systems. In 2024 5th International Conference for Emerging Technology (INCET) (pp. 1-8). IEEE.
Article without DOI from research databases	(Das & Chandra, 2023)	K. P. Das and J. Chandra, "A survey on artificial intelligence for reducing the climate footprint in healthcare," Energy Nexus, vol. 9, p. 100167, 2023.
Article without DOI from research databases	(Davis et al, 2012)	Davis, J., Edgar, T., Porter, J., Bernaden, J., & Sarli, M. (2012). Smart manufacturing, manufacturing intelligence and demand-dynamic performance. Computers & Chemical Engineering, 47, 145-156.
Article without DOI from research databases	(Deng et al, 2018)	Deng, H., Fannon, D., & Eckelman, M. J. (2018). Predictive modeling for US commercial building energy use: A comparison of existing statistical and machine learning algorithms using CBECS microdata. Energy and Buildings, 163, 34-43.
Article without DOI from research databases	(Didona et al, 2015)	Didona, D., Quaglia, F., Romano, P., & Torre, E. (2015, January). Enhancing performance prediction robustness by combining analytical modeling and machine learning. In Proceedings of the 6th ACM/SPEC

		international conference on performance engineering (pp. 145-156).
Article without DOI from research databases	(Dietze et al, 2018)	Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., ... & White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences, 115(7), 1424-1432.
Article without DOI from research databases	(Dhaliwal et al, 2018)	S. S. Dhaliwal, A. A. Nahid, and R. Abbas, "Effective intrusion detection system using XGBoost," Information, vol. 9, no. 7, p. 149, 2018.
Article without DOI from research databases	(Eiben et al, 2022)	A. Eiben, T. Berends, and T. Mosch, "Predictive maintenance for sewage pumping stations using machine learning," Ph.D. dissertation, Vrije Universiteit Amsterdam, 2022.
Article without DOI from research databases	(Elmouatamid et al, 2023)	Elmouatamid, A., Fricke, B., Sun, J., & Pong, P. W. (2023). Air Conditioning Systems Fault Detection and Diagnosis-Based Sensing and Data-Driven Approaches. Energies, 16(12), 4721.
Article without DOI from research databases	(Elsayed et al, 2021)	ElSayed, M. S., Le-Khac, N. A., Albahar, M. A., & Jurcut, A. (2021). A novel hybrid model for intrusion detection systems in SDNs based on CNN and a new regularization technique. Journal of Network and Computer Applications, 191, 103160.
Article without DOI from research databases	(Filz et al, 2021)	Filz, M. A., Langner, J. E. B., Herrmann, C., & Thiede, S. (2021). Data-driven failure mode and effect analysis (FMEA) to enhance maintenance planning. Computers in Industry, 129, 103451.
Article without DOI from research databases	(Frederiksen et al, 2021)	Frederiksen, R. D., Bocewicz, G., Radzki, G., Banaszak, Z., & Nielsen, P. (2024). Cost-Effectiveness of Predictive Maintenance for Offshore Wind Farms: A Case Study. Energies, 17(13), 3147.

Article without DOI from research databases	(Franklin et al, 2017)	A. Franklin, S. Gantela, S. Shifarraw, T. R. Johnson, D. J. Robinson, B. R. King, and N. G. Okafor, "Dashboard visualizations: Supporting real-time throughput decision-making," Journal of Biomedical Informatics, vol. 71, pp. 211-221, 2017.
Article without DOI from research databases	(Gould et al, 2008)	Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. European Journal of Operational Research, 191(1), 207-222
Article without DOI from research databases	(Ghrib et al, 2020)	Ghrib, Z., Jaziri, R., & Romdhane, R. (2020, July). Hybrid approach for anomaly detection in time series data. In 2020 international joint conference on neural networks (ijcnn) (pp. 1-7). IEEE.
Article without DOI from research databases	(Gudivada et al, 2020)	Gudivada, V., Apon, A., & Ding, J. (2017). Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. International Journal on Advances in Software, 10(1), 1-20.
Article without DOI from research databases	(Gholamzadehmir et al, 2020)	Gholamzadehmir, M., Del Pero, C., Buffa, S., & Fedrizzi, R. (2020). Adaptive-predictive control strategy for HVAC systems in smart buildings– A review. Sustainable Cities and Society, 63, 102480.
Article without DOI from research databases	(Irajpour et al, 2020)	Irajpour, A., Fallahian-Najafabadi, A., Mahbod, M. A., & Karimi, M. (2014). A framework to determine the effectiveness of maintenance strategies lean thinking approach. Mathematical Problems in Engineering, 2014(1), 132140
Article without DOI from research databases	(Huang et al, 2023) et al, 2020)	Huang, Y., Wang, Y., Wang, P., & Lai, Y. (2023). An XGBOOST predictive model of void ratio in sandy soils with shear-wave velocity as major input. Transportation Geotechnics, 42, 101100.

Article without DOI from research databases	(Jin et al., 2015)	Jin, W., Ullah, I., Ahmad, S., & Kim, D. (2019). Occupant comfort management based on energy optimization using an environment prediction model in smart homes. Sustainability, 11(4), 997.
Article without DOI from research databases	(Johnson et al, 2016)	Johnson, A. E., Ghassemi, M. M., Nemati, S., Niehaus, K. E., Clifton, D. A., & Clifford, G. D. (2016). Machine learning and decision support in critical care. Proceedings of the IEEE, 104(2), 444-466.
Article without DOI from research databases	(Jieyang et al, 2023)	Jieyang, P., Kimmig, A., Dongkun, W., Niu, Z., Zhi, F., Jiahai, W., ... & Ovtcharova, J. (2023). A systematic review of data-driven approaches to fault diagnosis and early warning. Journal of Intelligent Manufacturing, 34(8), 3277-3304.
Article without DOI from research databases	(Joliffe & Cadima, 2016).	Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202.
Article without DOI from research databases	(Jieyang et al, 2021)	Kang, Z., Catal, C., & Tekinerdogan, B. (2021). Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks. Sensors, 21(3), 932.
Article without DOI from research databases	(Kanawaday et al, 2017)	Kanawaday, A., & Sane, A. (2017, November). Machine learning for predictive maintenance of industrial machines using IoT sensor data. In 2017 8th IEEE international conference on software engineering and service science (ICSESS) (pp. 87-90). IEEE.
Article without DOI from research databases	(Kang et al, 2023)	Kang, J., Guo, X., Fang, L., Wang, X., & Fan, Z. (2022). Integration of Internet search data to predict tourism trends using spatial-temporal XGBoost composite model. International Journal of Geographical Information

		Science, 36(2), 236-252.
Article without DOI from research databases	(Kavzoglu & Teke, 2022)	Kavzoglu, T., & Teke, A. (2022). Predictive Performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). Arabian Journal for Science and Engineering, 47(6), 7367-7385.
Article without DOI from research databases	(Kiani et al, 2014)	Kiani, A., Salman, A., & Riaz, Z. (2014). Real-time environmental monitoring, visualization, and notification system for construction H&S management. Journal of Information Technology in Construction, 19, 72-91.
Article without DOI from research databases	(Laflamme et al., 2022).	S. Laflamme, F. Ubertini, A. Di Matteo, A. Pirrotta, M. Perry, Y. Fu, and P. Milillo, n measurement technologies for next generation structural health monitoring systems," Measurement Science and Technology, vol. 34, no. 9, p. 093001, 2023.
Article without DOI from research databases	(Kumbure et al, 2014)	Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. Expert Systems with Applications, 197, 116659.
Article without DOI from research databases	(Lee et al., 2020)	Lee, J., Ni, J., Singh, J., Jiang, B., Azamfar, M., & Feng, J. (2020). Intelligent maintenance systems and predictive manufacturing. Journal of Manufacturing Science and Engineering, 142(11), 110805.
Article without DOI from research databases	(Liu et al., 2023)	Liu, S., Wu, K., Jiang, C., Huang, B., & Ma, D. (2023). Financial time-series forecasting: Towards synergizing performance and interpretability within a hybrid machine learning approach. arXiv preprint arXiv:2401.00534.

Article without DOI from research databases	(Liu et al., 2018)	Liu, C. L., Hsaio, W. H., & Tu, Y. C. (2018). Time series classification with multivariate convolutional neural network. IEEE Transactions on industrial electronics, 66(6), 4788-4797.
Article without DOI from research databases	(Li Q.Y. 2021)	Li, Q. Y. (2021). A novel real-time monitoring, notification, analytics system, and personal thermal sensations model for indoor air quality and energy efficiency in commercial buildings.
Article without DOI from research databases	(Liu et al., 2020)	Liu, C., Ding, J., & Sun, J. (2020). Reinforcement learning based decision making of operational indices in process industry under changing environment. IEEE Transactions on Industrial Informatics, 17(4), 2727-2736.
Article without DOI from research databases	(Maddireddy et al., 2022)	Maddireddy, B. R., & Maddireddy, B. R. (2022). Real-Time Data Analytics with AI: Improving Security Event Monitoring and Management. Unique Endeavor in Business & Social Sciences, 1(2), 47-62.
Article without DOI from research databases	(Mahbub et al., 2020)	Mahbub, M., Hossain, M. M., & Gazi, M. S. A. (2020). IoT-Cognizant cloud-assisted energy efficient embedded system for indoor intelligent lighting, air quality monitoring, and ventilation. Internet of things, 11, 100266.
Article without DOI from research databases	(Maktoubian et al., 2023)	Maktoubian, J., Taskhiri, M. S., & Turner, P. (2021). Intelligent predictive maintenance (Ipdm) in forestry: A review of challenges and opportunities. Forests, 12(11), 1495.
Article without DOI from research databases	(Meshram, 2024)	Meshram, K. (2024). Design of an iterative method for environmental-sustainable development: Integrating bioinspired computing techniques. Environmental Development, 51, 101045.
Article without	(Mirnaghi, et al., 2020)	Mirnaghi, M. S., & Haghighat, F. (2020). Fault detection and diagnosis of large-scale HVAC systems in

DOI from research databases		buildings using data-driven methods: A comprehensive review. Energy and Buildings, 229, 110492.
Article without DOI from research databases	(Mishra & Rusch, 2021)	Mishra, S., & Rusch, T. K. (2021). Enhancing accuracy of deep learning algorithms by training with low-discrepancy sequences. SIAM Journal on Numerical Analysis, 59(3), 1811-1834.
Article without DOI from research databases	(Moore & Starr, 2006)	Moore, W. J., & Starr, A. G. (2006). An intelligent maintenance system for continuous cost-based prioritisation of maintenance activities. Computers in industry, 57(6), 595-606
Article without DOI from research databases	(Mullangi, 2017)	Mullangi, K., "Enhancing Financial Performance through AI-driven Predictive Analytics and Reciprocal Symmetry," Asian Accounting and Auditing Advancement, vol. 8, no. 1, pp. 57-66, 2017.
Article without DOI from research databases	(Nishijima, 2017)	Nishijima, D. (2017). The role of technology, product lifetime, and energy efficiency in climate mitigation: A case study of air conditioners in Japan. Energy Policy, 104, 340-347.
Article without DOI from research databases	(Nguyen & Medjaher, 2017)	Nguyen, K. T., & Medjaher, K. (2019). A new dynamic predictive maintenance framework using deep learning for failure prognostics. Reliability Engineering & System Safety, 188, 251-262.
Article without DOI from research databases	(Ran, et al., 2019)	Ran, Y., Zhou, X., Lin, P., Wen, Y., & Deng, R. (2019). A survey of predictive maintenance: Systems, purposes and approaches. arXiv preprint arXiv:1912.07383.
Article without DOI from research databases	(Rahmani et al., 2021)	A. M. Rahmani, E. Yousefpoor, M. S. Yousefpoor, Z. Mehmood, A. Haider, M. Hosseinzadeh, and R. Ali Naqvi, "Machine learning (ML) in medicine: Review, applications, and challenges," Mathematics, vol. 9, no. 22, pp. 2970, 2021.

Article without DOI from research databases	(Raparthi et al., 2020)	M.Raparthi, S. R. Gayam, B. P. Kasaraneni, K. K. Kondapaka, S. Putha, S. P. Pattyam, and M. K. Sahu, "Exploratory Data Analysis Techniques—A Comprehensive Review: Reviewing various exploratory data analysis techniques and their applications in uncovering insights from raw data," Australian Journal of Machine Learning Research & Applications, vol. 4, no. 1, pp. 215-225, 2024.
Article without DOI from research databases	(Roth et al., 2020)	Roth, J., Brown IV, H. A., & Jain, R. K. (2020). Harnessing smart meter data for a Multitiered Energy Management Performance Indicators (MEMPI) framework: A facility manager informed approach. Applied Energy, 276, 115435.
Article without DOI from research databases	(Parthasarathy et al, 2023)	V. B. Parthasarathy, A. Zafar, A. Khan, and A. Shahid, "The ultimate guide to fine-tuning LLMs from basics to breakthroughs: An exhaustive review of technologies, research, best practices, applied research challenges and opportunities," arXiv preprint arXiv:2408.13296, 2024.
Article without DOI from research databases	(Priscilla & Prabha, 2020)	C. V. Priscilla and D. P. Prabha, "Influence of optimizing XGBoost to handle class imbalance in credit card fraud detection," in 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 1309–1315. IEEE.
Article without DOI from research databases	(Osborne, 2012)	Osborne, J. W. (2012). Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data. Sage publications.
Article without DOI from research databases	(Quynh, 2023)	Quynh, D. T. (2023). The Impact of Dashboards on Risk Management and Decision-Making in Finance. Journal of Empirical Social Science Studies, 7(4), 51-63.
Article without DOI from research	(Qolomany et al., 2015)	B. Qolomany, A. Al-Fuqaha, A. Gupta, D. Benhaddou, S. Alwajidi, J. Qadir, and A. C. Fong, “Leveraging

databases		machine learning and big data for smart buildings: A comprehensive survey," IEEE Access, vol. 7, pp. 90316–90356, 2019.
Article without DOI from research databases	(Segun-Falade, et al, 2024)	Segun-Falade, O. D., Osundare, O. S., Kedi, W. E., Okeleke, P. A., Ijomah, T. I., & Abdul-Azeez, O. Y. (2024). Developing innovative software solutions for effective energy management systems in industry. Engineering Science & Technology Journal, 5(8).
Article without DOI from research databases	(Smith, 2023)	Smith, J. (2023). The Impact of Data Integrity on Clinical Trial Outcomes: Insights from Machine Learning.
Article without DOI from research databases	(Selcuk, 2023)	Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 231(9), 1670-1679.
Article without DOI from research databases	(Trizoglou et al., 2021)	Trizoglou, P., Liu, X., & Lin, Z. (2021). Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines. Renewable Energy, 179, 945–962. Unver, H. O. (2013). An ISA-95-based manufacturing intelligence system in support of lean initiatives. The International Journal of Advanced Manufacturing Technology, 65, 853-866.
Article without DOI from research databases	(Valarmathi& Kanaga , 2021)	Valarmathi, K., & Kanaga Suba Raja, S. (2021). Resource utilization prediction technique in cloud using knowledge based ensemble random forest with LSTM model. Concurrent Engineering, 29(4), 396-404.
Article without DOI from research databases	(Vinisha & Sujihelen, 2022).	Vinisha, F. A., & Sujihelen, L. (2022, January). Study on missing values and outlier detection in concurrence with data quality enhancement for efficient

		data processing. In 2022 4th international conference on smart systems and inventive technology (ICSSIT) (pp. 1600-1607). IEEE.
Article without DOI from research databases	(Wang et al., 2022)	Wang, H., Zhang, W., Yang, D., & Xiang, Y. (2022). Deep-learning-enabled predictive maintenance in industrial internet of things: methods, applications, and challenges. IEEE Systems Journal, 17(2), 2602-2615.
Article without DOI from research databases	(Wang et al., 2021)	Wang, C. C., Chien, C. H., & Trappey, A. J. (2021). On the application of ARIMA and LSTM to predict order demand based on short lead time and on-time delivery requirements. Processes, 9(7), 1157.
Article without DOI from research databases	(Wang et al., 2020)	Wang, Y., Zhao, Y., & Addepalli, S. (2020). Remaining useful life prediction using deep learning approaches: A review. Procedia manufacturing, 49, 81-88.
Article without DOI from research databases	(Wang & Gao,2022)	J. Wang and R. X. Gao, "Innovative smart scheduling and predictive maintenance techniques," in Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology, Elsevier, 2022, pp. 181-207.
Article without DOI from research databases	(Wang & Gao,2022)	J. Wang and R. X. Gao, "Innovative smart scheduling and predictive maintenance techniques," in Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology, Elsevier, 2022, pp. 181-207.
Article without DOI from research databases	(Wakiru et al., 2019)	Wakiru, J. M., Pintelon, L., Muchiri, P. N., & Chemweno, P. K. (2019). A review on lubricant condition monitoring information analysis for maintenance decision support. Mechanical systems and signal processing, 118, 108-132.
Article without DOI from research	(Wu et al., 2019)	Wu, C., Wu, P., Wang, J., Jiang, R., Chen, M., & Wang, X. (2021). Critical review of data-driven decision-making in bridge operation and maintenance.

databases		Structure and infrastructure engineering, 18(1), 47-70.
Article without DOI from research databases	(Yan et al., 2020)	Z. Yan, J. Wang, L. Sheng, and Z. Yang, "An effective compression algorithm for real-time transmission data using predictive coding with mixed models of LSTM and XGBoost," Neurocomputing, vol. 462, pp. 247–259, 2021
Article without DOI from research databases	(Zhang et al, 2019)	Y. Zhang, H. Wang, W. Gao, F. Wang, N. Zhou, D. M. Kammen, and X. Ying, "A survey of the status and challenges of green building development in various countries," Sustainability, vol. 11, no. 19, p. 5385, 2019.