Deepika Rane

Predictive Maintenance For Life Science Industry

Metropolia University of Applied Sciences
Master of Engineering
Information Technology
Master’s Thesis
21 March 2024

PREFACE

In recent years, the intersection of Machine Learning (ML) and Predictive Maintenance has emerged as a compelling area of research, particularly within the realm of Life Sciences. As a professional working in the life science industry and keen interest in exploring the latest trends in the field has shaped my journey towards investigating Predictive Maintenance in the context of Life Sciences using Machine Learning techniques.

This research has been embarked upon my interests with my professional expertise.

This journey would not have been possible without the support and encouragement of my colleagues and thesis supervisors, whose guidance and insights have been invaluable throughout this journey.

I extend my heartfelt gratitude to my husband, Paresh Naik and my daughters Anvi Naik and Manasvi Naik, whose support and encouragement have allowed me the time and space required for this thesis.

To my beloved parents Archana Rane and Anandu Rane,My sisters Pooja Naik and Seema Bhosale and my extended family, this dedication is a tribute to your belief in my potential to pursue my passion.

This thesis is dedicated to all those who have played a pivotal role in shaping my journey.

Vantaa, 21-03-2024
Deepika Rane

Abstract

Author:	Deepika Rane
Title:	Predictive Maintenance for life science industry
Number of Pages:	63 pages
Date:	21 March 2024
Degree:	Master of Engineering
Degree Programme:	Information Technology
Professional Major:	Networking and Services
Supervisors:	Ville Jääskeläinen.
	Peter Hjort.

This thesis addresses the application of predictive maintenance techniques in the context of continuous monitoring systems within the Life Science industry. The business challenge revolves around optimizing system reliability and minimizing unplanned downtime. The scope encompasses a comprehensive analysis of current state monitoring systems, followed by the exploration of predictive maintenance methodologies, with a focus on machine learning approaches.

The research delves into various predictive maintenance techniques, particularly those leveraging machine learning algorithms, to predict system conditions.

Implementation details cover the entire lifecycle of predictive maintenance, including problem framing, data preprocessing, feature engineering, model implementation, and evaluation.

The conclusion summarizes the key findings and highlights the results of the implementation, emphasizing the potential impact on equipment reliability and maintenance efficiency. Future work is outlined to explore advanced machine learning techniques, improve model performance, and integrate predictive maintenance into broader asset management strategies within the Life Science industry.

Contents

List of Abbreviations

ML: Machine Learning.

CMS: Continuous Monitoring System.

ULT: Ultra-Low Temperature.

FDA: Food and Drug Administration.

GxP: Good [insert variable] Practice.

FMS: Facility Monitoring System.

BMS: Environmental Monitoring System.

GMP: Good manufacturing practice.

CBM: Condition-Based Maintenance.

IoT: Internet of Things.

ARIMA: Autoregressive Integrated Moving Average.

LSTM: Long Short-Term Memory.

ROC: Receiver operating characteristic.

MAE: Mean Absolute Error.

AUC: Area under the Curve.

IFA: Isolation Forest Algorithm.

SVM: Support Vector Machine.

DP: Dynamic Programming.

AI: Artificial Intelligence.

1 Introduction

Every device has a point of failure. New appliances fresh from the factory are healthy and problem-free. Due to the device's wear and tear as it ages, its health slowly starts to deteriorate and eventually fails. We need to provide maintenance and get it back to a healthy condition. There are three types of maintenance: Reactive, Preventive and Predictive. In reactive maintenance, we simply wait for the device to break down and then perform maintenance on that device. This means we wait until the device fails and requires maintenance and then react. Here is an example of a microwave oven in our household we have used it for a few years, and it gets to the point that it does not turn on anymore. In this case, we either buy a new microwave oven or repair the microwave but it will take a few days to buy or repair the microwave oven which means a waiting period. In the reactive way of maintenance, we might not be able to use the microwave for some days and it might not be a big issue in this case. However, if the same thing in big industrial enterprises like life science industries it will have big consequences. For example, if we need to wait for the freezer to break down and then perform maintenance on it there might be considerable financial losses. This is because freezer failures contribute to important and highly valuable scientific research. So, wait-and-react or reactive maintenance might work fine for microwave ovens, but it is probably not the best choice in the life science business. The solution is to use preventive maintenance.

In preventive maintenance, we try to perform maintenance for the device long before the device gets to the point of failure. For example, changing the freezer condenser filters every three months or so and cleaning the freezer condenser coil annually. However, this is not cost-effective because performing maintenance early wastes device life that is still usable. This is the time that we could still use the device without any maintenance but now we are losing that because of early preventing maintenance. This is where the predictive maintenance can help. In predictive maintenance, we predict when the device fails and schedule maintenance just before it fails thus minimizing the device or machine downtime and maximizing its lifetime. So, preventive maintenance solves the downtime issues and prevents it from happening but losses on the device lifetime mean the time could use the device without any issues. However, predictive maintenance solves not only the device downtime issue but there will be little to no wasted device lifetime.

1.1 Context

A case study company provides environment measurements and monitoring solutions for critical life science environments and with 80 years of expertise in manufacturing products helped customers make better decisions and ensure safety and efficiency. Case study company provides a solution to safeguard sensitive products e.g., in warehouses, processing/manufacturing facilities, laboratories, and cleanrooms wherever pharmaceutical, biological, and other medical device products need to be protected from going out-of-tolerance conditions. Case company refrigerator and freezer monitoring, alarming and report solutions reduce the risk of lost product and regulatory non-compliance for life science environments and critical cold storage applications. The case company's temperature monitoring system with multiple connectivity options includes both wired and wireless options. The system is simple to set up and easy to use. Comprising an enterprise server or cloud system combined with data loggers, the system provides redundant data protection, long-term measurement accuracy, fail-safe alarming, and gap-free recording.

1.2 Business Challenge

The availability of ULT (Ultra-Low Temperature) freezers and cold-storage infrastructure provides the means to significantly reduce development time for therapeutics and vaccines because temperature stability testing at warmer temperatures (4 degrees Celsius or -20 degrees Celsius) is not needed before commercial release. However, major issues arise because many vaccination sites with new ULT freezers do not have the experience to install or maintain this type of equipment. ULT freezers are considered mission-critical equipment, but they require close supervision. A failure can lead to the loss of high-value products and research specimens, as well as safety issues. Industry standards and best practices including referencing good practice, or GxP, standards and protocols must be deployed in storing FDA (Food and Drug Administration)-regulated biologics. When GxP is applied to a ULT, regulations require that the temperature variation in the storage cabinet be within +/- 10 degrees Celsius of the targeted setpoint and that this be proven via recorded pre- and post-installation tests called validation. Compliance is accomplished using an independent temperature monitoring system, which also must be validated. Despite all this Life Science customers can lose high-value goods if the freezer fails suddenly, and the temperature gets too high. When failures occur, remediation is often too little, too late. Normally life science businesses require huge documentation in case of problems or failures. The monitoring systems based on temperature set point alarms do not indicate the problem started months ago when repairs were most affordable, but they report after the problem has already occurred.

1.3 Scope

The scope of the thesis was to analyze data collected from different temperature monitoring sensors of the system and predict the risk of failure of the system at an early phase. In this thesis prediction of the condition of the system was done by following the temperature variation of the system. In the life science business customers operate in close networks and are accessed only in a special situation. So, currently, there is not much data available. Due to this constraint in this thesis providing recommended actions for the predicted failures and creating automatic alerts based on the predicted condition of the system is not included.

This thesis has been divided into five sections. The first section introduces a brief description of different types of maintenance, the introduction of a case company and one of its business challenges and the objective of the thesis. The second section evaluates the current continuous monitoring system and summarizes the outcome. The third section introduces machine learning-powered solutions in predictive maintenance. The fourth part is model implementation and evaluation of the model. The fifth part includes conclusions based on the results and future works.

2 Current State Analysis

Usually, people and companies use different names or terms for such a system as CMS (continuous monitoring system), FMS (Facility Monitoring System), and EMS (Environmental Monitoring System).

The case company provides a system which is a continuous monitoring system and is a computerized system. The computerized system contains software, applications on a CD or memory stick, hardware measuring instruments, transmitters, sensors, cables, connectors etc. The computerized system is validated with main GMP (good manufacturing practice) regulations. GMP regulation requirements are applicable when manufacturing medical products for human and veterinary use. As per GMP regulation, there are different electronic record requirements for the case company such as human readable copies of data, the accuracy of device checks, controlled access (username/password), different levels of authority (authority checks), audit trails, system documentation and inventory lists. The case company does not monitor and control data with the same device because its basic function is to monitor not to control. CMS creates alarms and notifications in a failure situation. It includes data from separate time zones and can collect data from several different global sites in the very same system. This is a particularly important feature if several global company sites are included in the same system. To handle the data integrity issues during a power outage, the CMS system includes battery-powered devices with data-logging capabilities and when the power outage is over, data is transferred to a computerized system. The system also provides notification if the connection of connected devices to the system is halted. The current CMS system is easy to change and modify and it can be expanded easily without the need to order an expert to perform activities. It provides support to a full life cycle that includes for example documentation, validation, maintenance, calibration, version change support etc.

2.1 Continuous Monitoring Systems

A continuous monitoring system (CMS) is the optimal choice for monitoring valuable assets in a variety of industries. The primary purpose of the Continuous monitoring system (CMS) is measuring and documenting environmental parameters such as temperature and relative humidity specifically in controlled environments such as GxP regulated industries and applications in the Life Science industry such as Pharma, Biotech and Medical Device industries. Continuous Monitoring System is used in (GxP) Warehouses, Stability chambers, Cold Rooms, Refrigerators or Freezers, Cleanrooms, Calibration Labs, Blood/Tissue Storage, hospitals & Pharmacies, Museums and Archival Storage and Critical Manufacturing).

The system is an off-the-shelf configurable product but with its versatile features - like data collecting, alarming and reporting - it is flexible also for many other industries: Aerospace, Museums and Archives, Food and Beverages, Electronics and Semiconductor, Zoo's, Power industry, various (non-GxP) warehouses and many others. Continuous monitoring systems can be used for long-term continuous monitoring but also in shorter validation mapping projects.

Case company CMS system monitors temperature and relative humidity, but it is possible to monitor many other parameters such as Co2, and dew point. The continuous monitoring system is typically used as a quality tool to prove environmental parameters are correctly controlled in sensitive environments where sensitive products are stored or manufactured. So, the main input into the system is environment measurements. There are two outputs. The first output is data reports and trend graphs so that it can prove that the environmental parameters were kept under control and the other output is alarms and notifications. So, the user of this system can respond quickly when the environmental parameters of the system are out of specification. The high-level function of a continuous monitoring system includes data measured by sensors, data recorded at the site of the measurement, and data transmitted to the database. On central computer reserves where data is viewed and reported through a user interface on the server. That data is automatically validated to identify and generate alarms if necessary.

The case study company has two systems for industrial and indoor monitoring systems. In the following lines, we will go through each of the systems.
The first is ViewLinc continuous monitoring system which is an enterprise server continuous monitoring system designed for the GxP-regulated and critical monitoring applications. The system is available in 8 languages and scales easily to support global operations. The system provides alarming, real-time trends and reports the compliances for GxP and other regulations. The system includes data loggers to record data at the point of measurement and data is saved in the database with automatic backfill after any network interruption. Data can be accessed from the browser on the PC, tablet or mobile. The system has a configurable alarm to fit user workflow and procedure and users can create customizable report templates and schedule them to be delivered by email.
The second is the cloud monitoring system which is a cloud platform for data loggers and transmitters and uses SaaS service that is mobile optimized for industrial and other indoor monitoring needs for customers for example in the construction industry, museums, various (non-GxP) warehouses, etc. It is a convenient cloud-based data logging and alerting service including proven quality measurement devices. The application can be accessed through the web browser on your mobile phones, tablets, or laptops, making it ideal for professionals who want access to quality measurement data anywhere and at any time. The system is flexible and easy to expand from one single measurement point to a multisite system with thousands of measurement points. The Jade Smart Cloud solution consists of a cloud application to access measurement data remotely at any time and anywhere, data Loggers with humidity and temperature probes and Cloud Access points (one or multiple) to connect data loggers to the cloud. The service is available in the European Economic Region and the North American Region. Data can reside either in the European or North American region, depending on the customers' preference.

2.2 Summary

In the current continuous monitoring system, the real-time data and historical data are transmitted to the server database and are displayed in graphs on the floor dashboards. The main features of Continuous monitoring systems and cloud-based monitoring systems are event logs, reporting and alarming. The event log feature saves all the activities with timestamps in the secured log mainly this feature is used in audit trails. The system has one of the features of reporting and can be customised based on the user-defined content, design, logo, format of report in PDF or Excel and the delivery interval. Reporting can be set automatically for email delivery. Continuous monitoring systems and cloud-based monitoring systems send alarms or alerts when the set threshold limit exceeds, connectivity issues, and calibration reminders of the data loggers via email or SMS or on screen with sound and on-site with sound and light. The services provided by the case company for CMS are installations of the system, training on using the system, on-site calibrations service centre for calibration, customer technical support and maintenance agreement of the system.

3 Research and Methods

3.1 Predictive Maintenance Techniques

The primary objective of predictive maintenance is to predict equipment failures before they occur. By utilizing various monitoring and analysis techniques, organizations can anticipate potential issues and take proactive measures to address them, thus minimizing unplanned downtime and optimizing asset reliability. Predictive maintenance relies on condition-based monitoring to assess the health and performance of equipment. This involves collecting data from sensors and other monitoring devices to evaluate the condition of critical components and identify early signs of deterioration or malfunction. A crucial aspect of predictive maintenance is the analysis and interpretation of data collected from monitoring activities by using analytical tools and techniques to identify patterns, trends, and anomalies in the data, allowing maintenance professionals to make informed decisions about when and how to perform maintenance activities. Predictive maintenance enables organizations to optimize their maintenance strategies by focusing resources on the most critical assets and prioritizing maintenance tasks based on actual equipment conditions. This proactive approach helps maximize equipment uptime and minimize maintenance costs. Predictive maintenance should be integrated with broader maintenance management systems and practices to ensure alignment with organizational goals and objectives [1]. This involves establishing clear maintenance policies, procedures, and performance metrics, as well as providing training and support for maintenance personnel. Predictive maintenance is an ongoing process that requires continuous monitoring, analysis, and improvement. Organizations should regularly review and refine their predictive maintenance programs to adapt to changing conditions, technologies, and business requirements [1]. The importance of predictive maintenance as a proactive approach to managing equipment reliability and performance. By leveraging condition-based monitoring, data analysis, and optimization techniques, organizations can enhance their maintenance practices, reduce downtime, and achieve greater operational efficiency and effectiveness.

By understanding the characteristics and applications of each technique, organizations can develop comprehensive predictive maintenance programs to optimize asset reliability, minimize downtime, and enhance operational efficiency [1]. The predictive maintenance (PM) techniques their use cases, benefits, and limitations mentioned in this part of the thesis are specifically tailored to address the unique requirements and challenges presented for this case study.

Vibration Analysis: Vibration analysis detects abnormalities in rotating machinery by analysing vibration patterns. Vibration analysis involves monitoring the vibration levels of machinery to detect abnormalities that could indicate potential failures in rotating equipment such as motors, pumps, and turbines. By analysing vibration patterns and frequencies, maintenance professionals can identify issues like misalignment, unbalance, bearing wear, and structural defects. Vibration analysis typically uses accelerometers or other vibration sensors to collect data, which is then analysed using specialized software to diagnose problems and determine appropriate corrective actions [1].

- Use Cases: Vibration analysis is widely used to detect issues like misalignment, unbalance, bearing wear, and structural defects.
- Benefits: Vibration analysis enables early detection of equipment faults, allowing for timely corrective action to prevent unexpected breakdowns. It improves equipment reliability, extends lifespan, and minimizes downtime.
- Limitations: Vibration analysis requires specialized equipment, training, and expertise. It may not detect certain types of faults, such as electrical or process-related issues, and can be affected by environmental factors like temperature and humidity [1].

Thermography (Infrared Imaging**)**: Thermography utilizes infrared cameras to detect abnormal heat patterns in electrical systems, mechanical equipment, and other assets. Elevated temperatures can indicate issues such as loose connections, overloaded circuits, faulty components, and insulation breakdowns. By conducting regular thermal inspections, maintenance personnel can identify potential problems early and take corrective measures before they escalate into costly failures [1].

- Use Cases: Thermography is used for inspecting electrical systems, mechanical equipment, and building structures. It detects abnormalities in temperature distribution, indicating issues like loose connections, overloaded circuits, and mechanical wear [1].
- Benefits: Thermography enables non-contact, non-destructive testing of equipment, minimizing downtime and disruption. It identifies potential problems early, preventing equipment failures and reducing fire hazards.
- Limitations: Thermography requires line-of-sight access to equipment and may be affected by reflective surfaces, emissivity variations, and environmental conditions. It may not detect internal defects or certain types of faults [1].

Oil Analysis: Oil analysis involves regularly analyzing the condition of lubricating oils in machinery to detect contaminants, wear particles, and other indicators of impending failure. By monitoring parameters such as viscosity, acidity, contamination levels, and wear debris, maintenance professionals can assess the health of equipment components such as engines, gears, and hydraulic systems [1]. Oil analysis helps identify issues like bearing wear, gear tooth damage, and fluid degradation, allowing for timely maintenance actions such as oil changes, component replacements, or corrective repairs [1].

- Use Cases: Oil analysis is used for monitoring lubricating oils in machinery such as engines, gears, and hydraulic systems. It detects contaminants, wear particles, and fluid degradation, indicating issues like bearing wear, gear tooth damage, and fluid contamination.
- Benefits: Oil analysis provides early warning of equipment failures, allowing for timely maintenance actions such as oil changes, component replacements, and corrective repairs. It extends equipment lifespan and reduces maintenance costs.
- Limitations: Oil analysis requires periodic sampling and laboratory testing, which can be time-consuming and costly. It may not detect certain types of faults or provide information on non-lubricated components [1].

Ultrasonic Testing: Ultrasonic testing detects high-frequency sound waves emitted by malfunctioning equipment, which can indicate issues such as leaks, electrical faults, and bearing failures. Ultrasonic sensors are used to capture and analyse ultrasonic signals generated by equipment components, allowing maintenance personnel to identify abnormalities that may not be detectable through other methods [1]. Ultrasonic testing is particularly effective for detecting issues in pressurized systems, electrical equipment, and rotating machinery [1].

- Use Cases: Ultrasonic testing is used for detecting defects, leaks, and anomalies in pressurized systems, electrical equipment, and rotating machinery. It detects high-frequency sound waves emitted by malfunctioning equipment, indicating issues like leaks, electrical faults, and bearing failures.
- Benefits: Ultrasonic testing detects issues that may not be visible or accessible through other methods, enabling early intervention to prevent

failures and minimize downtime. It improves safety by identifying hazards such as gas leaks or electrical arcing.

- Limitations: Ultrasonic testing requires specialized equipment and training. It may be affected by background noise, material thickness, and surface conditions. Interpretation of ultrasonic signals can be subjective and require expertise [1].

Acoustic Monitoring: Acoustic monitoring involves using microphones or other acoustic sensors to monitor sound patterns emitted by machinery. Abnormal noises such as grinding, knocking, or squealing can indicate potential problems such as bearing wear, gear tooth damage, or lubrication issues. By continuously monitoring equipment sounds and comparing them to established baseline patterns, maintenance professionals can identify deviations and take corrective actions to prevent failures [1].

- Use Cases: Acoustic monitoring is used for detecting abnormal sounds emitted by machinery, indicating issues like bearing wear, gear tooth damage, or lubrication issues. It is commonly applied in rotating equipment, pumps, and motors.
- Benefits: Acoustic monitoring provides early warning of equipment faults, allowing for proactive maintenance to prevent failures and minimize downtime. It improves safety by identifying hazards such as mechanical failures or structural defects [1].
- Limitations: Acoustic monitoring may be affected by background noise, environmental conditions, and equipment location. Interpretation of acoustic signals can be subjective and requires experience [1].

3.1.1 Predictive maintenance techniques for life science monitoring systems.

As mentioned in the previous chapter 2.1 and chapter 1.2 the case company focuses on life science and industrial measurement devices so in this part of the thesis studies were done for some predictive maintenance techniques that are tailored to these specific domains i.e. life science and industrial measurement devices. Predictive maintenance techniques can be applied to life science and industrial measurement devices such as Calibration verification, Sensor health monitoring, Environmental Monitoring, Data analysis and trends monitoring, Remote monitoring and diagnostics and Condition-based maintenance [3]. The below section mentions different predictive maintenance techniques tailored to life science industries, their benefits, and limitations.

Calibration Verification: Calibration is critical for ensuring the accuracy and reliability of measurement devices in life science and industrial settings. Predictive maintenance involves implementing regular calibration verification procedures to check if instruments are still within acceptable tolerances. This can be achieved by comparing device readings against traceable standards or performing calibration checks using reference samples. Deviations from expected values can indicate potential calibration drift or instrument degradation, prompting corrective actions such as recalibration or adjustment [3].

Benefits of Calibration Verification:

- Maintains Measurement Accuracy: Calibration Verification ensures that measurement devices remain accurate and reliable, providing confidence in the accuracy of measurement results.
- Compliance with Standards: Helps organizations comply with regulatory requirements and quality standards that mandate regular calibration and verification of measurement devices.
- Prevents Inaccurate Measurements: Early detection of deviations from expected measurements allows for timely corrective actions to prevent inaccuracies and potential errors in measurement data.
- Cost Savings: Regular Calibration Verification can help identify issues early, reducing the likelihood of costly errors or rework resulting from inaccurate measurements.
- Quality Assurance: Assures stakeholders, customers, and regulatory bodies regarding the quality and reliability of measurement data and products [3].

Limitations of Calibration Verification:

- Frequency of Verification: Determining the appropriate frequency for Calibration Verification can be challenging and may vary depending on factors such as device type, usage, and criticality.
- Resource Intensive: Performing Calibration Verification requires time, resources, and expertise to ensure accurate results and compliance with standards [3].
- Potential for Human Error: Errors in the execution of Calibration Verification, such as incorrect measurement techniques or interpretation of results, can lead to inaccuracies and undermine the effectiveness of the process.
- Limited Detectability: Calibration Verification may not detect all sources of measurement error or drift, particularly if they occur gradually over time or are intermittent.
- Equipment Limitations: Some measurement devices may have inherent limitations or uncertainties that cannot be fully addressed through Calibration Verification alone [3].

Sensor Health Monitoring: Many measurement devices rely on sensors to collect data accurately. Predictive maintenance techniques involve monitoring sensor health and performance to detect issues such as drift, degradation, or malfunctions. This may involve periodic sensor calibration, testing sensor response to known stimuli, or analyzing sensor output for anomalies. By identifying sensor problems early, maintenance personnel can prevent data inaccuracies and ensure the reliability of measurement results.

Benefits of Sensor Health Monitoring [3]:

- Improved Reliability: Ensures that sensors remain accurate and reliable, leading to more reliable process measurements and control.
- Early Detection of Issues: Detects sensor degradation or malfunction early, allowing for proactive maintenance before failures occur.
- Reduced Downtime: Minimizes unplanned downtime by identifying and addressing sensor issues before they impact production.

- Optimized Maintenance: Enables predictive and condition-based maintenance strategies for sensors, optimizing maintenance schedules and resource allocation [3].
- Enhanced Safety: Helps prevent safety hazards or process disruptions that may result from faulty sensor readings or control errors.

Limitations of Sensor Health Monitoring:

- Complexity of Data Analysis: Analysing sensor data and distinguishing between normal variations and abnormal conditions can be challenging, requiring advanced analytics and domain expertise.
- Sensor Diversity: Different types of sensors may require different monitoring techniques and analysis methods, making it challenging to implement a standardized approach across all sensors.
- False Alarms: Sensor health monitoring systems may generate false alarms or alerts due to noise, transient conditions, or other factors, leading to unnecessary maintenance actions or disruptions [3].
- Dependency on Sensor Data Quality: The effectiveness of sensor health monitoring relies on the quality and reliability of sensor data. Poor data quality or inaccuracies may affect the accuracy of health assessments and predictions.
- Cost and Complexity: Implementing sensor health monitoring systems may require investment in hardware, software, and expertise, as well as ongoing maintenance and support [3].

Environmental Monitoring: Environmental conditions can impact the performance and longevity of measurement devices. Predictive maintenance includes monitoring environmental factors such as temperature, humidity, and vibration levels to assess their potential impact on device operation [3]. Monitoring environmental conditions can help identify conditions that may accelerate device degradation or affect measurement accuracy. By controlling environmental variables or implementing protective measures, maintenance personnel can mitigate risks and prolong device lifespan [3].

Benefits of Environmental Monitoring Predictive Maintenance:

- Optimized Equipment Performance: Helps maintain optimal operating conditions for equipment by ensuring that environmental parameters are within acceptable ranges.
- Early Detection of Issues: Detects environmental changes or anomalies early, allowing for proactive maintenance to prevent equipment failures or performance degradation [3]
- Improved Energy Efficiency: Enables optimization of energy usage by monitoring environmental conditions and adjusting heating, cooling, ventilation, and other systems accordingly.
- Enhanced Safety and Compliance: Helps ensure a safe and healthy working environment by monitoring air quality, temperature, and other factors to comply with regulatory requirements and industry standards.
- Reduced Environmental Impact: Minimizes environmental impact by monitoring and managing resources more efficiently, such as reducing energy consumption or optimizing water usage [3].

Limitations of Environmental Monitoring Predictive Maintenance:

- Complexity of Data Analysis: Analyzing environmental data and identifying meaningful patterns or trends can be challenging, requiring advanced analytics and domain expertise.
- Sensor Accuracy and Reliability: The effectiveness of environmental monitoring depends on the accuracy and reliability of sensor data. Poor sensor calibration or malfunctions may lead to inaccurate assessments or false alarms [3].
- Dependency on External Factors: Environmental conditions may be influenced by external factors such as weather, seasonal changes, or nearby activities, making it difficult to control or predict certain events.
- Cost and Resource Requirements: Implementing and maintaining environmental monitoring systems may require investment in sensors, infrastructure, software, and expertise, as well as ongoing monitoring and support [3].

Remote Monitoring and Diagnostics: Remote monitoring technologies allow real-time monitoring of measurement devices from a centralized location. Predictive maintenance involves implementing remote monitoring systems that continuously collect and analyze device data, providing early detection of abnormalities or deviations from expected performance. Remote diagnostics capabilities enable maintenance personnel to identify issues remotely, troubleshoot problems, and initiate corrective actions without the need for onsite intervention, minimizing downtime and optimizing maintenance efficiency [3].

Benefits of Remote Monitoring and Diagnostics Predictive Maintenance:

- Early Detection of Issues: Enables early detection of potential equipment failures or performance degradation by continuously monitoring key parameters and analysing data in real-time.
- Predictive Maintenance: Facilitates predictive maintenance strategies by leveraging diagnostic insights to predict when maintenance is required, allowing for proactive maintenance scheduling and optimization of resources.
- Reduced Downtime: Minimizes unplanned downtime by identifying and addressing potential issues before they escalate into major failures, thereby improving equipment availability and reliability [3].
- Remote Access and Control: Provides remote access to equipment status and performance data, allowing maintenance personnel to monitor, diagnose, and troubleshoot issues from anywhere, reducing the need for onsite inspections and interventions.
- Cost Savings: Optimizes maintenance costs by reducing the frequency of reactive maintenance, minimizing unnecessary inspections, and optimizing resource allocation based on data-driven insights [3].

Limitations of Remote Monitoring and Diagnostics Predictive Maintenance:

- Reliability of Data Transmission: Dependence on communication networks for data transmission introduces the risk of data loss, latency, or connectivity issues, which may affect the timeliness and accuracy of monitoring and diagnostic results.
- Sensor Accuracy and Reliability: The effectiveness of remote monitoring relies on the accuracy and reliability of sensor data. Poor sensor calibration, malfunctions, or environmental factors may lead to inaccurate readings or false alarms.

- Complexity of Data Analysis: Analyzing large volumes of data collected from remote monitoring systems and interpreting diagnostic results can be complex and require specialized expertise and tools.
- Privacy and Security Concerns: Remote monitoring systems may raise concerns regarding data privacy, security, and compliance with regulatory requirements, especially when transmitting sensitive or proprietary information over public networks [3].
- Integration Challenges: Integrating remote monitoring and diagnostic systems with existing equipment, control systems, or data management platforms may require modifications, upgrades, or interoperability solutions, which can be challenging and costly [3].

Condition-Based Maintenance (CBM): Condition-based maintenance strategies involve monitoring key indicators of device health and performance to schedule maintenance activities only when necessary. Predictive maintenance techniques in life science and industrial measurement devices leverage CBM approaches to prioritize maintenance tasks based on equipment condition and criticality. By focusing resources on devices exhibiting signs of deterioration or impending failure, maintenance personnel can optimize maintenance schedules, reduce unnecessary maintenance, and maximize equipment uptime [3].
Benefits of Condition-Based Maintenance:

- Increased Equipment Reliability: CBM enables early detection of potential failures or performance degradation, allowing maintenance to be performed proactively before failures occur, thereby improving equipment reliability and availability.
- Reduced Downtime: By scheduling maintenance based on actual equipment condition, CBM minimizes unplanned downtime associated with unexpected failures, leading to increased operational uptime and productivity.
- Optimized Maintenance Costs: CBM optimizes maintenance costs by reducing the frequency of unnecessary maintenance activities and focusing resources on critical components or systems where maintenance is most needed.

- Extended Equipment Lifespan: Proactive maintenance based on CBM insights helps extend the lifespan of equipment by preventing premature wear and tear and avoiding catastrophic failures that may result in irreparable damage [3].
- Improved Safety and Compliance: By ensuring that equipment is properly maintained and in good working condition, CBM helps enhance safety and compliance with regulatory requirements and industry standards.

Limitations of Condition-Based Maintenance:

- Sensor Reliability and Accuracy: The effectiveness of CBM relies on the accuracy and reliability of sensor data. Poor sensor calibration, malfunctions, or environmental factors may lead to inaccurate readings or false alarms.
- Complexity of Data Analysis: Analyzing large volumes of sensor data collected from CBM systems and interpreting diagnostic results can be complex and require specialized expertise and tools.
- Integration Challenges: Integrating CBM systems with existing equipment, control systems, or data management platforms may require modifications, upgrades, or interoperability solutions, which can be challenging and costly.
- Dependency on External Factors: CBM may be influenced by external factors such as weather, environmental conditions, or changes in operating conditions, making it difficult to control or predict certain events [3].

Anomaly Detection: Anomaly detection techniques identify deviations from normal operating patterns or expected behaviour in monitoring systems. By analyzing historical data and establishing baseline performance metrics, anomalies indicative of potential equipment failures or malfunctions can be detected early. Anomaly detection algorithms, such as statistical methods, machine learning, or pattern recognition techniques, enable timely interventions to prevent disruptions and ensure data integrity in life science monitoring systems [12].

Benefits of Anomaly Detection

- Early Fault Detection: Anomaly detection techniques allow for the early identification of abnormal behaviour in monitored systems, enabling timely intervention to prevent equipment failures or malfunctions.
- Minimized Downtime: By detecting anomalies in real-time, maintenance activities can be scheduled proactively, minimizing downtime and optimizing equipment availability and productivity [12].
- Improved Data Integrity: Anomaly detection ensures the integrity and accuracy of data collected from monitored systems by flagging outliers or erroneous readings that could compromise the quality of data analysis and decision-making.
- Cost Savings: Proactive maintenance based on anomaly detection helps organizations save costs associated with unplanned downtime, emergency repairs, and potential losses resulting from equipment failures or malfunctions [12].

Limitations of Anomaly Detection

- False Positives: Anomaly detection algorithms may generate false positives, flagging normal variations in system behaviour as anomalies, leading to unnecessary maintenance actions and resource allocation.
- Complexity of Interpretation: Interpreting anomalies detected by the algorithm may be challenging, requiring domain expertise and context to distinguish between benign fluctuations and genuine issues requiring attention.
- Data Quality Requirements: Anomaly detection algorithms rely on high-quality data for accurate results. Inadequate or noisy data may hinder the effectiveness of anomaly detection techniques, leading to missed detections or false alarms [12].
- Threshold Setting: Setting appropriate thresholds for anomaly detection can be challenging, as it requires a balance between sensitivity (detecting true anomalies) and specificity (minimizing false alarms), which may vary depending on the specific application and context [12].

Digital Twins: Digital twins create virtual replicas of monitoring systems based on real-time sensor data and operational parameters. By simulating the behaviour and performance of monitoring equipment in a virtual environment, digital twins enable predictive analytics and scenario analysis to be performed. This allows for the optimization of maintenance strategies, the identification of performance bottlenecks, and the prediction of system behaviour under different operating conditions, enhancing the reliability and efficiency of life science monitoring systems [3]

Benefits of Digital Twins:

- Predictive Insights: Digital twins provide a virtual representation of physical assets or systems, allowing for predictive analytics and simulations to forecast future behaviour, anticipate failures, and optimize maintenance strategies.
- Optimized Maintenance Planning: By simulating the impact of different maintenance scenarios and interventions, digital twins enable organizations to optimize maintenance schedules, prioritize tasks, and allocate resources efficiently, minimizing downtime and maximizing equipment uptime [3].
- Real-Time Monitoring and Control: Digital twins facilitate real-time monitoring and control of physical assets or systems, providing insights into performance metrics, operational parameters, and environmental conditions, enabling proactive adjustments and interventions to maintain optimal performance [3].
- Lifecycle Management: Digital twins support the entire lifecycle of assets or systems, from design and development to operation and maintenance, allowing for continuous improvement, innovation, and optimization based on real-world data and feedback [3].

Limitations of Digital Twins:

- Complexity and Development Costs: Developing and maintaining digital twins can be complex and resource-intensive, requiring expertise in modelling, simulation, and data integration, as well as investment in software, hardware, and infrastructure.
- Data Integration and Quality: Digital twins rely on accurate and up-to-date data from physical assets and systems to ensure fidelity and reliability. Integrating data from disparate sources and ensuring data quality and

consistency can be challenging, affecting the accuracy and reliability of digital twin simulations.

- Model Validation and Calibration: Digital twin models need to be validated and calibrated against real-world data to ensure accuracy and reliability. Incorrect or inaccurate models may lead to erroneous predictions and suboptimal decision-making, highlighting the importance of ongoing model validation and refinement [3].
- Privacy and Security Concerns: Digital twins store and process sensitive data related to physical assets or systems, raising concerns about data privacy, security, and compliance with regulatory requirements. Implementing robust data governance and security measures is essential to protect digital twin assets and mitigate the risk of data breaches or unauthorized access [3].

Predictive maintenance techniques represent a paradigm shift in maintenance strategies, moving from reactive or scheduled approaches to proactive, data-driven methodologies. These techniques leverage a variety of tools and technologies to anticipate and prevent equipment failures before they occur, optimizing maintenance schedules, reducing downtime, and extending the lifespan of machinery. Predictive maintenance relies heavily on data analysis, both historical and real-time, to identify patterns, anomalies, and trends indicative of potential equipment failures and Sensors and monitoring devices play a crucial role in collecting data on equipment health and performance. These sensors measure parameters such as vibration, temperature, and fluid properties,
providing insights into machinery conditions. Techniques such as machine learning and artificial intelligence enable predictive maintenance systems to process large volumes of data, identify hidden patterns, and make accurate predictions about equipment health and maintenance needs.

3.2 Predictive Maintenance with Machine Learning

Organizations are increasingly building and using machine learning (ML)-powered solutions for a variety of use cases and problems, including predictive maintenance of machine parts, product recommendations based on customer preferences, credit profiling, content moderation, fraud detection, and more. Instead of relying on fixed schedules or reactive maintenance practices, predictive maintenance uses predictive models to anticipate maintenance needs and schedule maintenance activities before failures occur. In the thesis, the machine learning approach for predictive maintenance was considered based on the following factors [4].

- Able to analyze vast amounts of data and identify complex patterns.
- Models can learn from historical data and adapt to changing conditions, resulting in more precise maintenance recommendations.
- Identify subtle deviations in equipment behaviour that may indicate potential failures. This early detection capability allows organizations to intervene proactively, minimizing downtime and preventing costly breakdowns.
- Handle diverse types of data, including sensor data and maintenance logs.
- Flexible and can process structured and unstructured data from multiple sources, enabling a comprehensive analysis of equipment health and performance.
- Model can continuously learn from new data and feedback, improving their accuracy and effectiveness over time.
- Adapt to changing conditions and evolving failure patterns, ensuring ongoing optimization and performance enhancement.
- Can be seamlessly integrated with other advanced technologies, such as the Internet of Things (IoT), edge computing, and cloud computing.

ML approaches can be applied in various stages of predictive maintenance, such as fault prediction, remaining useful life estimation, anomaly detection, and decision-making. The choice of algorithm depends on factors such as the nature of the problem, data characteristics, and computational resources available. In understanding the underlying principles, algorithms, and methodologies to develop accurate and effective predictive models' theory has played a crucial role [8]. In this part of the thesis, we provide a framework that elucidates the theoretical foundations supporting predictive maintenance through the utilization of machine learning techniques in this case study.

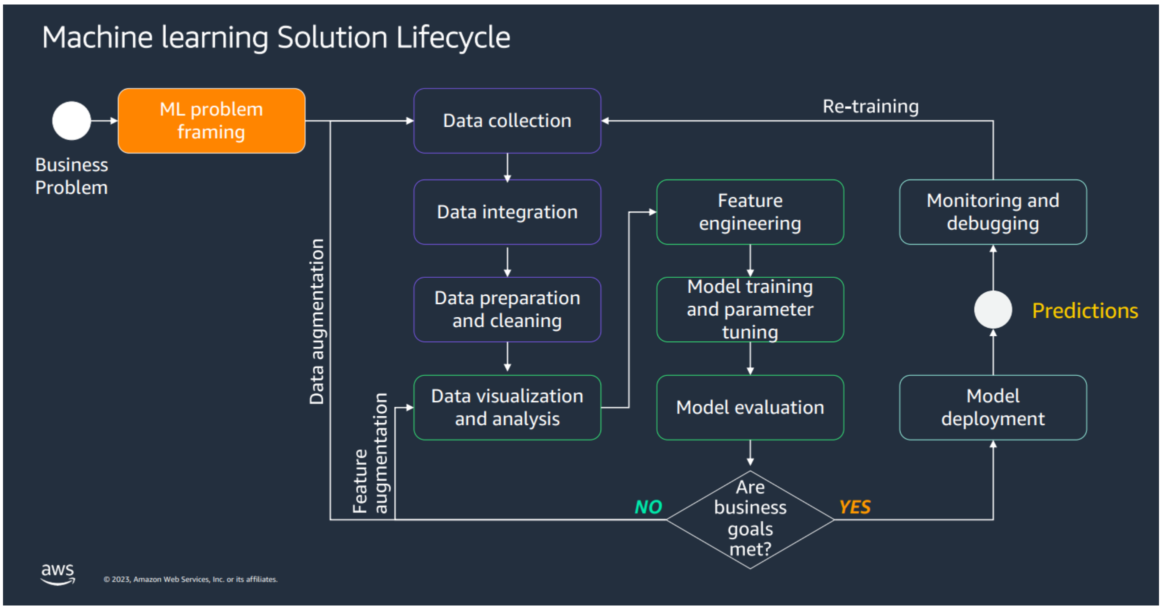

Figure 1: Machine Learning Solution Lifecycle [8].

The Machine Learning Solution Lifecycle refers to the end-to-end process of developing, deploying, and maintaining machine learning models to solve real-world problems. There are several stages in the ML lifecycle as shown in Figure 1[8] and explained in detail in the below section.

- Data Collection and Preprocessing: Gather historical maintenance records, sensor data, and operational parameters from the equipment to be monitored. Clean and preprocess the data to remove noise, handle missing values, and normalize or scale features as needed.
- Feature Engineering: Extract relevant features from the data that are indicative of equipment health and performance. This may include time-series features, statistical measures, frequency domain features, or domain-specific metrics.

- Model Selection: Choose appropriate machine learning models based on the nature of the predictive maintenance task and the characteristics of the data. Commonly used models are regression models (e.g., linear regression, support vector regression) for predicting equipment degradation or remaining useful life and classification models (e.g., random forests, gradient boosting) for identifying fault conditions or imminent failures and time-series forecasting models (e.g., ARIMA (Autoregressive Integrated Moving Average), LSTM (Long Short-Term Memory)) for predicting future equipment behaviour based on historical data and anomaly detection models (e.g., isolation forests, autoencoders) for detecting abnormal patterns or outliers indicative of equipment malfunction [8].
- Model Training: Split the data into training and testing sets to evaluate model performance. Train the selected machine learning models on the training data, optimizing model parameters to minimize prediction errors or maximize performance metrics such as accuracy, precision, recall, or F1 score.
- Model Evaluation: Evaluate the trained models using the testing set to assess their predictive capabilities and generalization performance. Measure metrics such as mean absolute error (MAE), root mean squared error (RMSE), receiver operating characteristic (ROC) curve, or area under the curve (AUC) to quantify model performance [8].
- Deployment and Monitoring: Deploy the trained machine learning models into production environments to perform real-time predictive maintenance tasks. Continuously monitor model performance and update models as needed to adapt to changing operating conditions, new data, or evolving equipment characteristics.
- Integration with Maintenance Workflow: Integrate predictive maintenance predictions into existing maintenance workflows and decision-making processes. Use model predictions to prioritize maintenance tasks, schedule inspections or repairs, and allocate resources efficiently [8].

3.3 Machine Learning Approaches

Machine learning is the process where machines take the data and analyze it to generate predictions and use those predictions to make decisions. Those predictions generate results, and those results are used to improve future predictions but this iterative process differs from typical supervised learning paradigms, where predictions are not typically used to enhance future iterations. Machine learning can make predictions from huge data sets. It can also optimize utility functions and extract the hidden data or pattern structure from those data sets by classifying the data. This enables the software program to learn and make predictions in future. Here it emphasizes the core objective of machine learning, which is to enable computers to learn from data and improve their performance over time without human intervention. Training the models on data, machine learning systems can identify patterns and relationships that would be difficult or impossible for humans to detect manually [4].

Machine learning is important because it has revolutionized various industries and domains by enabling automation, improving efficiency, and driving innovation. In healthcare, machine learning algorithms can analyze medical images to assist in diagnosis and treatment planning. In finance, they can analyze market data to predict trends and make investment decisions. In manufacturing, they can optimize production processes and detect defects in real time. In transportation, they can power self-driving vehicles and optimize logistics operations. Machine learning has become increasingly important in today's technological landscape because of its ability to process large amounts of data and extract valuable insights that can inform decision-making and drive business outcomes [4]. By automating tasks that were previously performed manually, machine learning algorithms can save time, reduce costs, and improve accuracy in a wide range of applications.

The evolution of machine learning can be traced back to the mid-20th century, with early developments in pattern recognition and artificial intelligence as shown in Figure 2 [8]. In the 1950s and 1960s, researchers developed early machine learning algorithms such as the Perceptron and the Nearest Neighbor algorithm. In the 1980s and 1990s, there was a resurgence of interest in machine learning with the development of algorithms like Support Vector Machines and Decision Trees. In the 2000s and beyond, advancements in computational power and the availability of large datasets have led to breakthroughs in deep learning, a subfield of machine learning that uses neural networks with many layers to learn complex patterns from data.

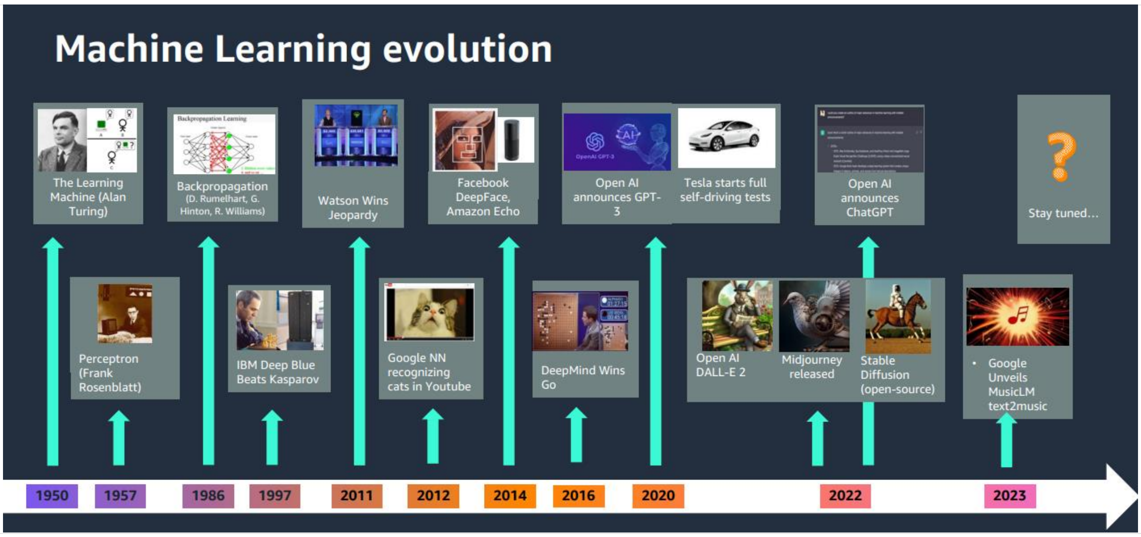

Figure 2: Machine Learning Evolution [8].

The evolution of machine learning has been driven by advancements in computing technology, data availability, and algorithmic innovation. Early developments laid the foundation for modern machine learning techniques, while recent breakthroughs in deep learning have enabled significant advances in areas such as computer vision, natural language processing, and robotics[8].

The fundamental concepts and terminology used in machine learning are such as data, features, labels, models, algorithms, training, and prediction. Data is the raw information that machine learning algorithms learn from. Data can be in various forms, including structured data (e.g., tabular data), unstructured data (e.g., text, images), and semi-structured data (e.g., JSON, XML). Features are the measurable properties or characteristics of the data that are used as input to machine learning algorithms [2]. Features can be numeric, categorical, or binary, and they help the algorithm learn patterns and make predictions. Labels are the target variables or outputs that machine learning algorithms aim to predict.

Models are mathematical representations or hypotheses that machine learning algorithms learn from the data. The model captures the relationships between features and labels and can be used to make predictions on new, unseen data.

Algorithms are the computational procedures or techniques used to train machine learning models. Algorithms define how the model learns from the data and updates its parameters to minimize prediction errors. Training is the process of fitting a machine learning model to the training data by adjusting its parameters iteratively[8]. During training, the model learns from the examples in the training dataset to improve its performance on unseen data. Prediction is the process of using a trained machine learning model to make predictions or decisions on new, unseen data.

Deep Learning, Artificial Intelligence (AI), and Machine Learning (ML) are interrelated concepts. AI is the overarching field focused on creating intelligent systems, ML is a subset of AI that deals with algorithms and models that learn from data, and Deep Learning is a specialized subfield of ML that revolves around deep neural networks, which have proven to be very effective in various AI applications. Deep Learning is a subset of ML, which is a subset of AI as shown in Figure 3 [8].

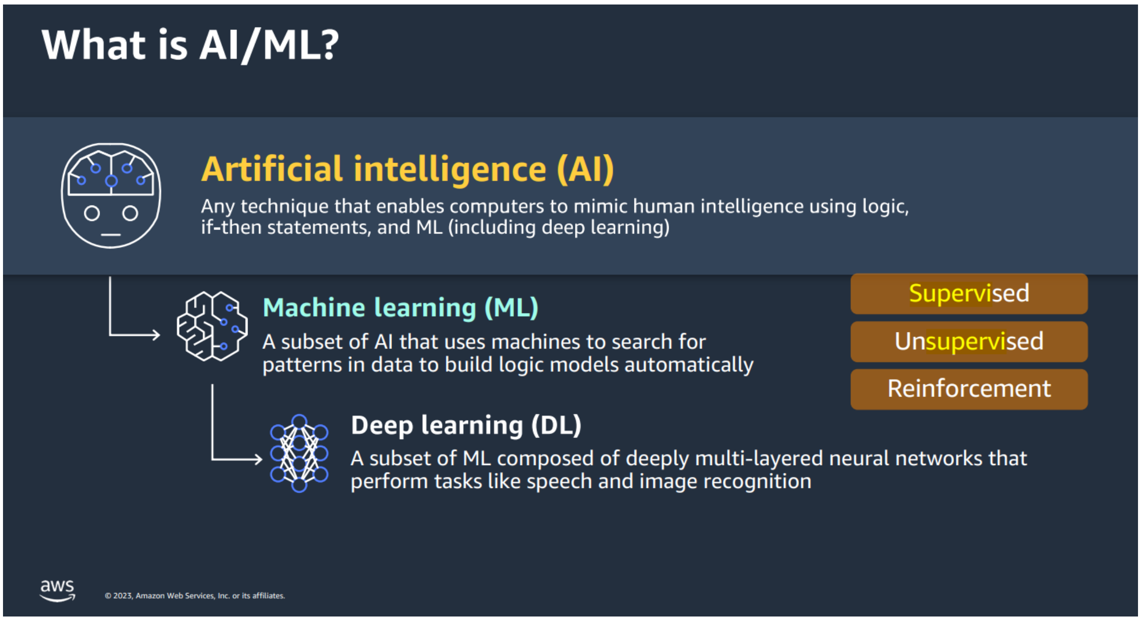

Figure 3: The fields of AI (Artificial Intelligence), ML (Machine Learning) and DL (Deep learning) [8].

Machine learning encompasses a wide range of approaches and techniques for solving problems, making predictions, and extracting patterns from data. Broadly machine learning models can be categorized into supervised learning unsupervised learning, and reinforcement learning as shown in Figure 4 [2].

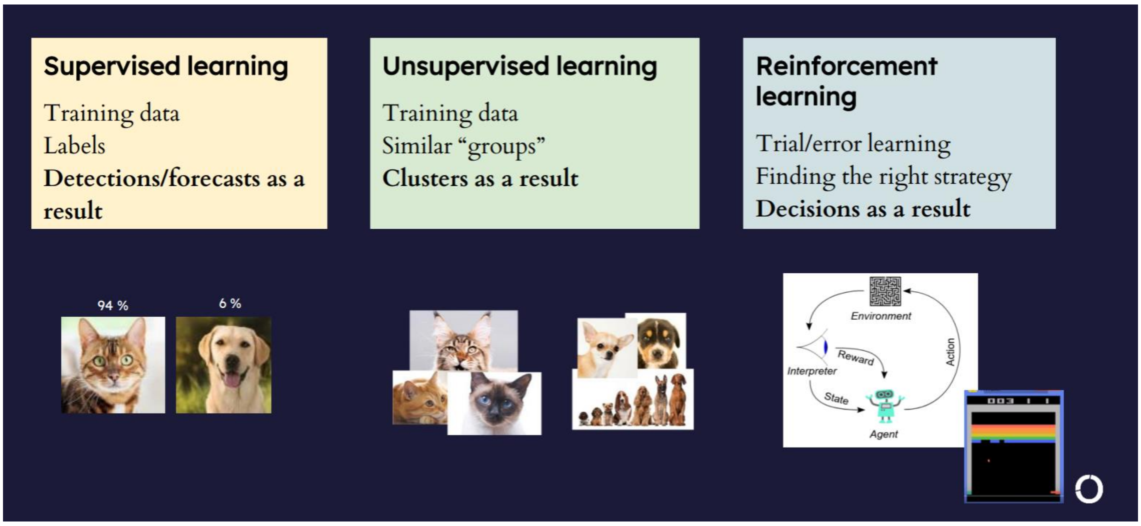

Figure 4: Machine learning approaches [2].

Supervised learning involves a series of functions that map an input to an output based on the example input-output pairs in Figure 5 [2]. For example, if we have variables of two datasets one being age which is the input and the other being shoe size as output, we will implement a super learning model to predict the shoe size of the person based on their age.

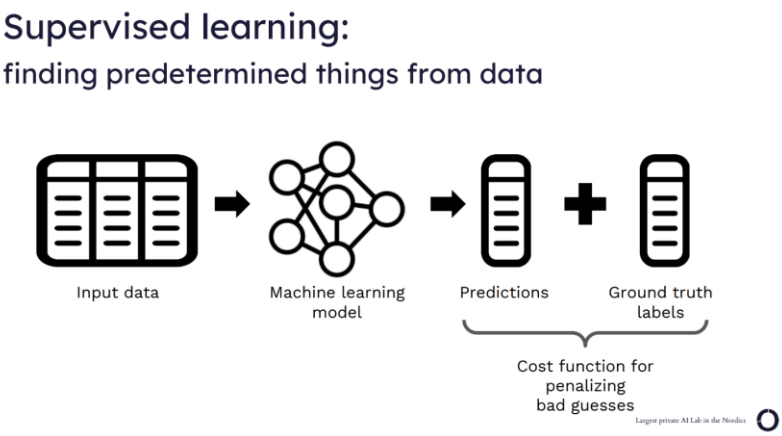

Figure 5: Supervised Learning [2].

Further, with supervised learning, there are two sub-categories regression and classification. In the regression model, we find the target value based on the independent predictors which means we can use this to find the relation between the dependent variable and independent variable. In the regression model, the output is continuous. While various regression models exist, this part of the thesis specifically highlighted linear regression, decision trees, and random forests, as they are particularly relevant to this case study. It's important to note that other regression techniques also exist but are not discussed here due to their lesser relevance to this case study scenario.[2].

The linear regression model is a simple finding line which fits the data. It assumes a linear relationship between the input variables and the target variable The linear regression model aims to minimize the difference between the actual values and the predicted values by adjusting the coefficients (slope and intercept) of the linear equation. It makes predictions by computing the weighted sum of the input features, optionally adding a constant term (intercept). Linear regression is widely used in various fields, including economics, finance, healthcare, and social sciences. It's suitable for predicting continuous numerical values, such as house prices, stock prices, or temperature forecasts [2].

Decision tree models are hierarchical tree-like structures that recursively partition the feature space into smaller regions based on the feature values. Each node in the tree represents a decision based on a specific feature, leading to different branches (outcomes). The decision tree algorithm splits the feature space at each node based on the feature that provides the best separation of the target variable. This process continues until a stopping criterion is met, such as a maximum depth or minimum number of samples per leaf. Decision trees are used for both classification and regression tasks [2]. In regression, decision trees can predict continuous numerical values by averaging the target variable within each leaf node. They're suitable for scenarios with complex, nonlinear relationships between features and the target variable.

Random forests are ensemble learning methods that consist of multiple decision trees trained on different subsets of the training data and features. Each tree in the forest independently learns to make predictions, and the final prediction is made by aggregating the predictions of all trees. Random forests introduce randomness during the training process by randomly selecting subsets of features and samples for each tree [2]. This helps reduce overfitting and improves the model's generalization performance. The final prediction is typically obtained by averaging (regression) or voting (classification) the predictions of individual trees. Random forests are versatile and widely used for regression tasks, including predicting house prices, sales forecasts, and customer lifetime value. They're known for their robustness, scalability, and ability to handle high-dimensional data with noisy or irrelevant features.

During the training process, the supervised learning algorithm learns from the training data to build a predictive model. It iteratively adjusts the model parameters to minimize the difference between the predicted output and the actual labels in the training data. The loss function measures the error or discrepancy between the predicted output and the actual labels. The goal of training is to minimize the loss function, thereby improving the accuracy of the model predictions [2]. Gradient descent is an optimization algorithm used to update the model parameters iteratively during training. It calculates the gradient of the loss function concerning the model parameters and adjusts the parameters in the direction that minimizes the loss. For the model evaluation and validation, various metrics are used to evaluate the performance of supervised learning models, depending on the task type (classification or regression). Common evaluation metrics include accuracy, precision, recall, F1 score, mean squared error (MSE), and mean absolute error (MAE). Overfitting occurs when a model learns to memorize the training data instead of generalizing from it. It results in poor performance on unseen data because the model is too complex and captures noise or irrelevant patterns in the training data [2]. Underfitting occurs when a model is too simple to capture the underlying patterns in the training data. It results in high bias and poor performance both on the training data and unseen data.

Unsupervised data uses unlabeled data to train machines. Unlabeled data means there are no fixed output variables. The model learns from the data, discovers the patterns and features in the data and returns the output.

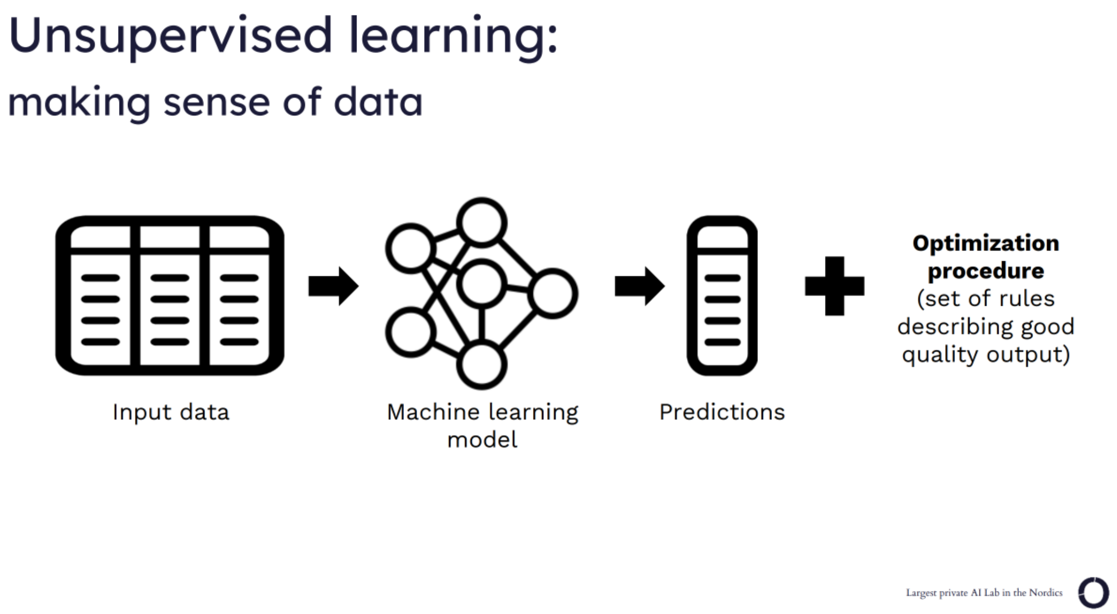

Figure 6: Unsupervised Learning [2].

Unsupervised learning algorithms identify patterns and structures in data without labelled output as shown in Figure 6 [2]. The types of unsupervised learning are clustering, dimensionality reduction and anomaly detection. Clustering is a common task in unsupervised learning, where the algorithm groups similar data points together into clusters or segments based on their intrinsic properties. The objective is to maximize the intra-cluster similarity while minimizing the inter-cluster similarity. Clustering algorithms, such as K-means, and hierarchical clustering, partition the data into distinct groups based on similarity or distance metrics. Each cluster represents a subset of data points with similar characteristics. Dimensionality reduction techniques aim to reduce the number of features or variables in the dataset while preserving its essential information. These techniques help visualize high-dimensional data, alleviate the curse of dimensionality, and improve computational efficiency Dimensionality reduction techniques, such as Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), and Autoencoders, project high-dimensional data onto a lower-dimensional space while preserving as much variance or information as possible [2]. Anomaly detection, also known as outlier detection, involves identifying rare or unusual data instances that deviate significantly from the norm. Anomalies may indicate potential fraud, errors, or interesting patterns in the data Anomaly detection algorithms, such as Isolation Forest, One-Class SVM(Support Vector Machine), and Density-Based Anomaly Detection, identify data instances that deviate significantly from most of the data distribution[2]. Anomaly detection algorithms are employed in cybersecurity to detect suspicious activities or intrusions in network traffic, fraud detection in financial transactions, and equipment maintenance in industrial systems. During the training process, unsupervised learning algorithms iteratively learn from the data to uncover underlying patterns or structures. The training process typically involves updating model parameters or cluster centroids based on predefined criteria, such as minimizing intra-cluster variance or maximizing data likelihood. For evaluation, there is no ground truth or labelled data to evaluate unsupervised learning models directly. Evaluation metrics for clustering may include silhouette score, Davies–Bouldin index, or connectivity-based metrics, while anomaly detection algorithms may use metrics such as precision, recall, or F1 score [2].

Reinforcement learning (RL) is a subfield of machine learning that involves training agents to make sequential decisions in an environment to maximize cumulative rewards. Unlike supervised learning, where the algorithm learns from labelled data, and unsupervised learning, where the algorithm discovers patterns in unlabeled data, reinforcement learning is about learning from interaction and feedback.

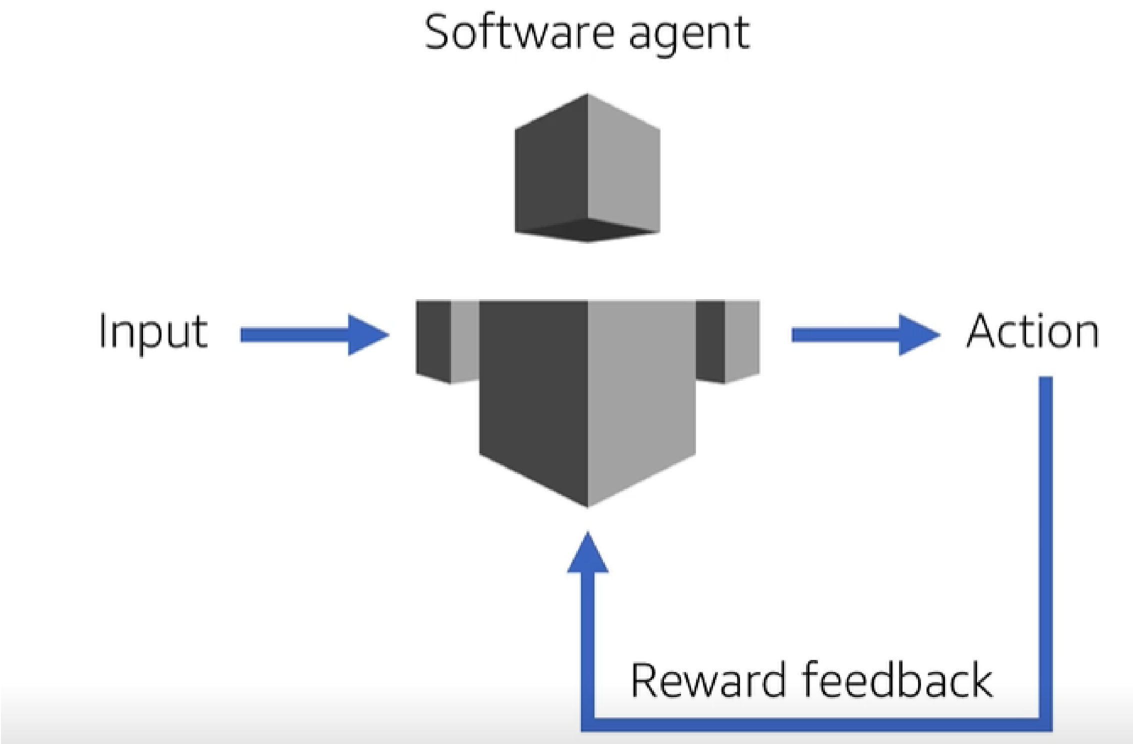

Figure 7: Reinforcement learning [2].

In reinforcement learning, there are two primary components: the agent and the environment. The agent is the decision-making entity that interacts with the environment and the environment is the external system in which the agent operates [2]. It provides feedback to the agent based on its actions. At each time step, the agent observes a state from the environment. Based on the observed state, the agent selects an action from a set of possible actions. After taking the action, the environment transitions to a new state and the agent receives a numerical reward as feedback. The agent's strategy for selecting actions is defined by its policy that maps states to actions, specifying the agent's behaviours in different environmental states as shown in Figure 7 [2]. Policies can be deterministic always selecting the same action in a given state or stochastic selecting actions probabilistically. The value function provides a measure of the expected return or cumulative reward that the agent can achieve from a given state or state-action pair. The state-value function estimates the expected return starting from a particular state under a given policy. Reinforcement learning algorithms aim to learn an optimal policy that maximizes the cumulative reward over time and try out different actions to discover the optimal policy. It balances the exploration of new actions with the exploitation of known actions that yield high rewards [2].

The key reinforcement learning algorithms and techniques covered in the thesis are Dynamic Programming (DP), Monte Carlo Methods, Temporal Difference Learning, Q-Learning, and Policy Gradient Methods [2]. Dynamic Programming is a method for solving complex problems by breaking them down into simpler subproblems. In reinforcement learning, DP algorithms are used to solve problems where the environment is fully known and can be modelled as a Markov Decision Process (MDP). Policy Evaluation computes the value function for a given policy by iteratively applying the Bellman expectation equation. Policy iteration alternates between policy evaluation and policy improvement steps until an optimal policy is found. Value iteration updates the value function iteratively until convergence to find the optimal value function and policy. Monté Carlo Method is a more generic class of algorithms that estimate value functions by averaging returns from sampled episodes. This method is often used in episodic tasks where the agent interacts with the environment until termination. Unlike DP, Monte Carlo methods do not require a model of the environment and can be applied to problems with unknown dynamics. Temporal Difference (TD) Learning is a method that combines elements of both DP and Monte Carlo methods. It updates value functions based on the difference between temporally successive estimates and can be used for both prediction and control tasks.TD(lambda) methods, such as TD(λ), use eligibility traces to assign credit for returns over multiple time steps. Q-Learning is a model-free RL algorithm used for learning optimal action-value functions and estimates the value of taking a specific action in each state and updates the action-value function using the Bellman optimality equation. The algorithm iteratively improves the action-value function until it converges to the optimal values'-Learning is suitable for problems with large state and action spaces where a model of the environment is not available [2]. Policy Gradient Methods directly parameterize policies and optimize them using gradient ascent methods. These methods aim to maximize the expected return by adjusting policy parameters in the direction of the steepest ascent. Actor-critic methods combine policy gradient and value function estimation to achieve stable learning and faster convergence. Policy Gradient Methods are particularly effective for problems with continuous action spaces and high-dimensional state spaces.

Reinforcement learning has applications in various domains, including robotics, autonomous driving, game playing, finance, healthcare, and recommendation systems. Examples include training robots to perform tasks in real-world environments, optimizing resource allocation in business settings, and personalizing recommendations based on user interactions [2].

3.4 Build your ML model Vs AI services.

A Hackathon was conducted in the case company to understand and have a comparison between developing a machine learning model and utilizing AI services. The main input for this hackathon is the guidelines and knowledge from the AWS immersion day workshop. The main findings of this hackathon are mentioned in the below parts.
To build your model or use AI services, there are several factors to be considered, including cost, expertise, customization, scalability, and time to market for considering whether to build your model or use AI services. Building your model allows you to tailor it specifically to your organization's needs, incorporating domain-specific knowledge and fine-tuning the model architecture and parameters to achieve optimal performance. You have full control over the entire model development process, from data collection and preprocessing to model training, evaluation, and deployment [9]. This gives you greater flexibility to adapt the model to changing requirements and incorporate feedback over time. Building own model requires expertise in machine learning, data science, and software engineering. Needs skilled professionals who can handle data preprocessing, feature engineering, model selection, training, and deployment. Accessing sufficient and high-quality data for the training model requires data from various sources, which can be time-consuming and resource-intensive. The development

and deployment of the model entail the acquisition of infrastructure and computing resources, encompassing hardware, software, and cloud services essential for training. Building your model can be cost-effective if you have the required expertise and resources in-house. However, it may involve upfront costs for hiring talent, infrastructure setup, and ongoing maintenance and support [9].
AI services offer pre-built solutions that you can readily integrate into your applications with minimal effort. These solutions are often developed by experts and come with built-in features for data processing, model training, and deployment. Enable rapid prototyping and experimentation, allowing you to quickly test different models and algorithms without investing significant time and resources [9]. To scale automatically to handle varying workloads and data volumes. This makes them suitable for applications with fluctuating demands and growing data requirements services often follow a pay-as-you-go pricing model, where you only pay for the resources that are in use. This can be cost-effective for small to medium-sized businesses that don't have the resources to build and maintain their models. AI services offer convenience and ease of use, but they may have limitations in terms of customization and flexibility. You may not have full control over the model architecture, training process, or deployment environment. AI services from a particular vendor may lead to vendor lock-in, where you become dependent on their platform and tools. This could be challenging if you want to switch vendors or customize the solution further in the future [9].

The decision between building your model and/or using AI services depends mainly on these decision factors for the case company such as budget, expertise, and time to market. Considering budget and cost constraints for building your model may involve higher upfront costs but could be more cost-effective in the long run. Expertise in machine learning, data science, and software engineering for the case company at this stage might not be feasible but there are steps taken to improve competence in-house [9]. From the hackathon findings, we know that using AI services may be the faster option to be able to deploy quickly. However, building your model may take longer but could offer better long-term flexibility for the case company that requires customization or has specific requirements.

The building your model approach offered maximum customization and control over the entire development process, from data collection to model deployment. The model architecture was tuned according to the needs, selected appropriate algorithms, and optimized parameters based on your domain and specific requirements [9].

In the second case utilizing AI services offered pre-built solutions and platforms that provide out-of-the-box functionality for given tasks without requiring extensive expertise in machine learning or software development. AI services followed a pay-as-you-go pricing model, allowing case companies to access advanced machine learning capabilities without significant upfront investments.

The hackathon provided valuable insights into the strengths and limitations of building your model versus leveraging AI services, highlighting the importance of considering those factors [9].

4 Implementation Details

The goal is to predict the likelihood of freezer malfunction or failure based on temperature sensor data and in the absence of labelled data. In this context labelled data is the compressor status. By analyzing historical sensor logs and detecting patterns, we aim to anticipate potential issues before they occur, enabling predictive maintenance and minimizing downtime. Temperature mapping is a common Quality Control (QC) activity performed for the verification and/or validation and calibration of equipment when temperature is of critical importance to the process. Hypothesize that certain temperature patterns or anomalies may precede freezer malfunctions. For example, sudden temperature spikes or prolonged deviations from the optimal temperature range may indicate underlying issues with the freezer's cooling system or insulation.

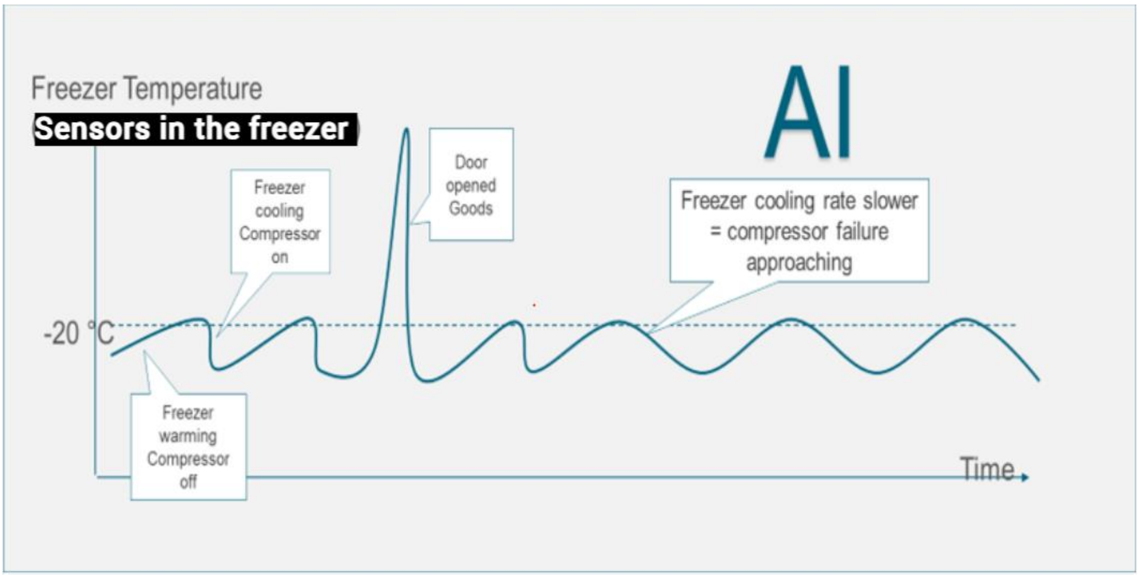

Figure 8: Freezer operating condition.

If the cooling rate of a freezer is decreasing, it could be an indication that the compressor, which is responsible for cooling the freezer, is failing, or approaching failure as shown in Figure 8 [11]. Here the cooling rate of a freezer refers to how quickly the temperature inside the freezer decreases to reach the desired setpoint. This rate is monitored using temperature sensors placed inside the freezer. The compressor is a critical component of a freezer's refrigeration system. It works by compressing refrigerant gas, which increases its temperature and pressure [11]. The high-pressure gas then travels through condenser coils where it releases heat and condenses into a liquid. The liquid refrigerant then flows through an expansion valve, where it expands and evaporates, absorbing heat from the freezer's interior and causing the temperature to decrease. A decrease in the freezer's cooling rate can indicate potential issues with the compressor. Compressor failure or degradation can manifest in various ways, including. The compressor may struggle to maintain the desired temperature inside the freezer, resulting in slower cooling rates. A failing compressor may require more energy to operate efficiently, leading to higher electricity bills. Malfunctioning compressors may produce unusual noises or vibrations as they struggle to operate. Following the temperature variations of the freezer, the failure is estimated in the early phase [11].

4.1 ML Problem Framing

We have already covered in the theoretical part of this thesis that unsupervised learning is a type of machine learning where the algorithm learns patterns and structures from unlabeled data without explicit supervision. Unsupervised learning finds hidden structures or relationships within the data, such as clusters, associations, or anomalies [12]. Unlike supervised learning, there are no predefined labels or target variables in unsupervised learning tasks. Anomaly detection is a subfield of unsupervised learning that focuses on identifying rare events, outliers, or deviations from normal behaviour in data and aims to distinguish between normal and abnormal patterns in the data, often without prior knowledge of the specific anomalies. Techniques of Anomaly detection are statistical techniques,machine learning algorithms, time series analysis, and deep learning [12]. Statistical techniques such as Z-score, gaussian distribution modelling, or percentile ranking are used to detect anomalies based on deviations from the expected statistical properties of the data. Machine learning-based approaches, including clustering algorithms (e.g., k-means), density estimation techniques (e.g., Gaussian Mixture Models), or novelty detection algorithms (e.g., Isolation Forest, One-Class SVM), are commonly employed for anomaly detection tasks. Techniques like moving averages, exponential smoothing, or autoregressive models can be used to detect anomalies in time-series data by analyzing temporal patterns and trends. Deep learning models, particularly autoencoders, recurrent neural networks (RNNs), or long short-term memory networks (LSTMs), can learn complex representations of data and detect anomalies in high-dimensional and sequential data [12].

Machine learning-based approach isolation forest algorithm was used for the case study of predicting compressor failure based solely on freezer temperature data, where labelled data was unavailable because the isolation forest algorithm can learn the normal behaviour of the temperature data without the need for labelled data examples of compressor status. In this algorithm, there is no dependency on the labelled data. Isolation Forest is capable of handling high-dimensional data efficiently. In the case of freezer temperature data collected from multiple sensors, each sensor reading can be considered as a separate dimension. The algorithm can effectively detect anomalies in this high-dimensional space without suffering from the curse of dimensionality. It is scalable and can handle large datasets with ease which is suitable for the case study where a significant amount of temperature data are processed and analyzed over time. Detecting both global anomalies (i.e., anomalies that deviate significantly from the overall data distribution) and local anomalies (i.e., anomalies that deviate significantly from their local neighbourhoods). This flexibility allows to capture of different types of anomalies present in the freezer temperature data.

4.2 Data Collection and Data Preprocessing

In data collection phase involved gathering historical data from the five different sensors for three different freezers, each providing temperature readings along with corresponding timestamps. The case company's continuous monitoring system with five sensors was used to collect temperature readings along with timestamps from each sensor installed in the freezer and ensure that the data was recorded at regular intervals to maintain consistency and time zone for accurate analysis.

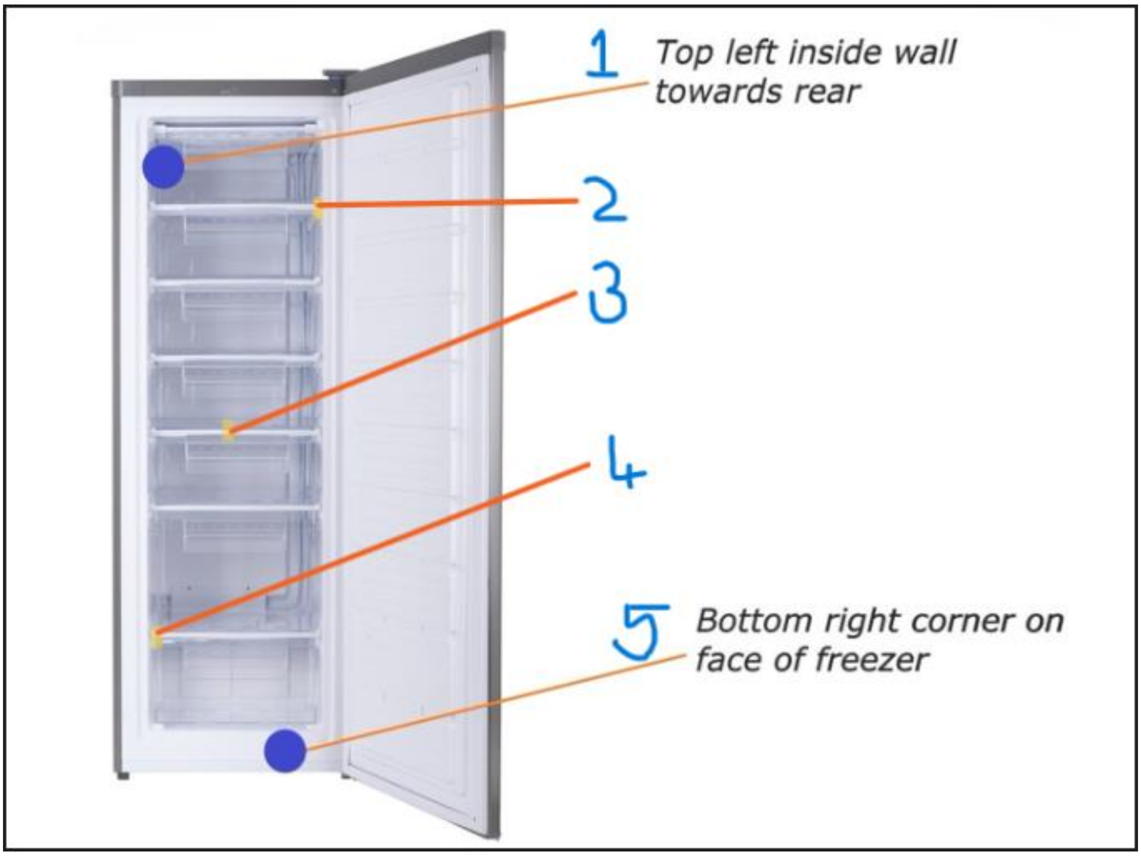

Figure 9: Sensors mapping in the freezer.

The data was collected during the normal operating conditions of the freezer. The freezers were mapped in 3-D with approximate Logger Placement as shown in Figure 9.

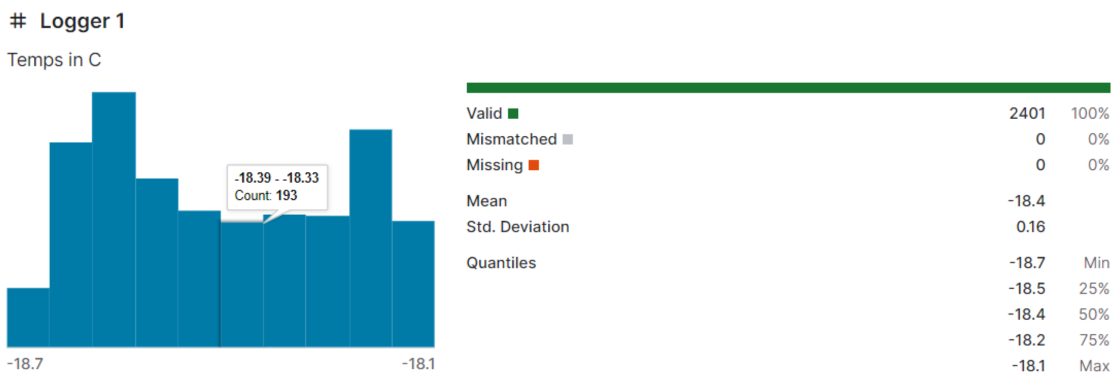

Figure 10: Freezer 1 Logger 1 temperature range.

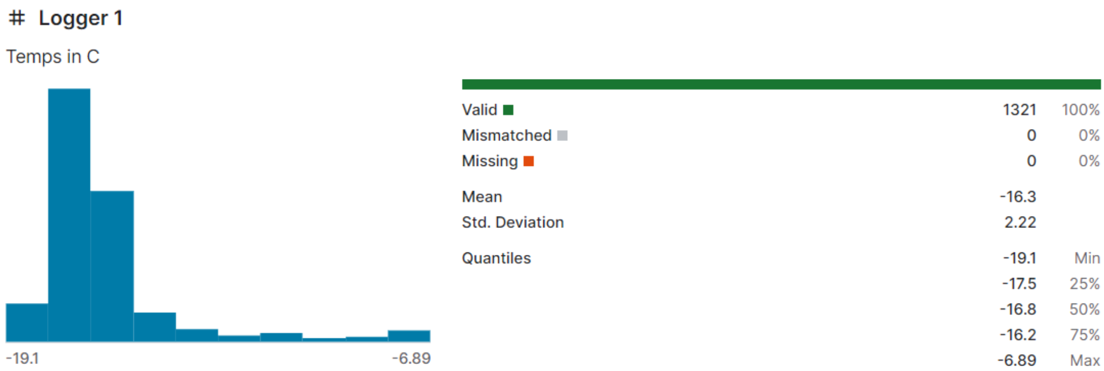

Figure 11: Freezer 2 Logger 1 temperature range.

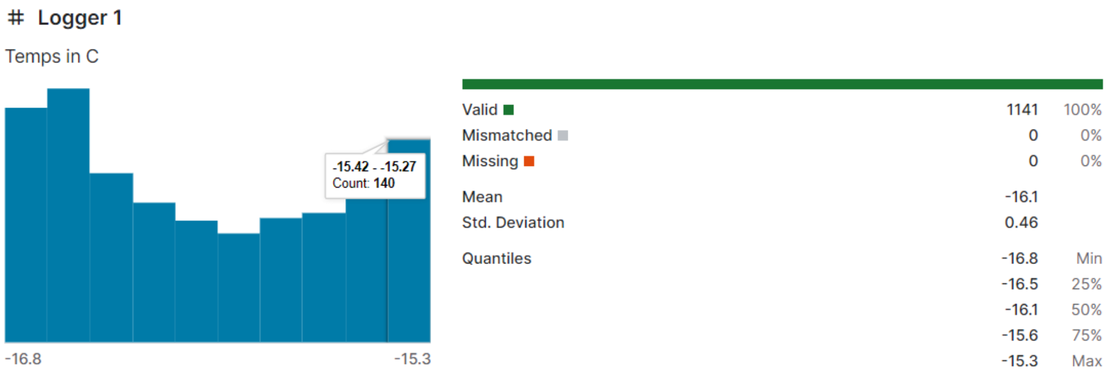

Figure 12: Freezer 3 Logger 1 temperature range.

Each file contains columns for the time stamp and temperature sensor value separately numbered 1-5. Temperatures are in Celsius and logging frequency was typically set to 1 minute. As seen in Figure 10 freezer 1 temperature data ranges from -18.7 to -18.1 which implies that freezer 1 has a more stable temperature environment, with less fluctuation over time. Freezer 2 has a temperature range from -19.1 to -6.89 as shown in Figure 11 which is the broader temperature range indicating in this case potential issues with temperature control. In Figure 12 freezer 3 temperature data ranges from -16.8 to -15.3 indicating less fluctuation over the period.

Data preprocessing is essential to ensure that the collected data is clean, consistent, and ready for analysis. For this case study data preprocessing was done by checking for missing or null values in the temperature readings, replacing missing values with the mean value and removing the missing records. Temperature readings were normalized to a common scale to ensure that features contribute equally to the analysis. The cooling rate was calculated i.e. difference in temperature between consecutive timestamps based on the temperature readings and timestamp.

4.3 Feature Engineering

Feature engineering involves creating new features or transformations from the existing data to improve the performance of anomaly detection algorithms. For this case study calculated cooling rate was created as a relevant feature to capture trends and anomalies. The cooling rate indicates how fast the temperature is decreasing over time. The cooling rate provides insights into the rate of change of temperature, which can be indicative of anomalies.

4.4 Model Implementation

For the case study, the isolation Forest algorithm was implemented. The algorithm is based on the principle that anomalies are typically rare and different from the majority of normal data points. The algorithm selects a random feature and then selects a random split value between the minimum and maximum values of that feature. Isolation Forest recursively applies random partitioning to create isolation trees. Each tree is constructed by selecting a random subset of the data and recursively partitioning it into two subspaces along the selected feature and split value [13]. This process continues until each data point is isolated in its leaf node. Once the isolation trees are built, in the next step it calculates an anomaly score for each data point. The anomaly score is based on the depth of the data point in the tree. Anomalies are expected to have shorter paths (i.e., fewer splits) to reach isolation, while normal data points require more splits to isolate them. The isolation score for each data point is calculated as the average path length in the isolation trees [13]. The average path length is normalized by the maximum path length in a fully balanced tree to obtain a score between 0 and 1. Lower scores indicate anomalies, while higher scores indicate normal data points. Based on the isolation scores, a threshold can be applied to classify data points as anomalies or normal. Data points with isolation scores below the threshold are considered anomalies, while those above the threshold are considered normal [13]. The isolation forest algorithm using Python with scikit-learn was implemented by importing the necessary libraries, including scikit-learn for machine learning functionality [6]. Preprocessed data was Loaded, and data was made in a suitable format for input into the Isolation Forest model. An instance of the isolation forest class from scikit-learn was created and hyperparameters i.e. contamination parameters to control the expected proportion of outliers in the data are used for this case study. Here the contamination parameter was adjusted based on the case study requirements. The isolation forest model was trained on the prepared data using the fit method and the trained isolation forest model was used to predict anomalies in the data using the prediction method. Anomalies are assigned a label of -1, while normal data points are assigned a label of 1. Anomalies were visualized in the graph [7]. Below is the pseudocode to illustrate the steps involved in training and applying the isolation forest model to the freezer temperature data [7].

```
Step 1: Data collection.
#Temperature dataset with data from five different sensors.

Step 2: Preprocessing the data.
# Convert timestamp to datetime
# Sort by timestamp
# mean temperature is calculated to combine temperature from all sensors

Step 3: Feature engineering.
# The cooling rate (difference in temperature between consecutive timestamps)
# Drop rows with missing values
```

Step 4: Implementation.

from sklearn import Isolation Forest

Isolation Forest model on the cooling rate and mean temperature.

Apply the hyperparameter value i.e., contamination parameter.

#Check for anomalies in the cooling rate.

4.5 Model Evaluation

The effectiveness of the isolation forest model for anomaly detection in this context of predicting compressor failure based on the cooling rate and mean temperature data was evaluated by using visual inspection, statistical analysis, and a cross-validation approach. Evaluation of the model was done for three different data sets from three freezers.

In visual inspection by plotting the temperature data and highlighting the anomalies identified by the model. This helped to understand the distribution of anomalies and their relationship with the temperature readings [7].

Below are the pseudocode steps for visual inspection of the temperature data and highlights of anomalies identified by the Isolation Forest model using Python and Matplotlib library.

Steps 1:

Plot mean temperature data with timestamp

Step 2:

plot highlights anomalies.

Conducting the statistical analysis on the anomalies detected by the model by computing summary statistics mean, the median and standard deviation for the cooling rate features for both anomalous and normal instances helped to gain insights into the characteristics of anomalies detected by the model [10]. Summary statistics provide a concise overview of the distribution of a dataset. Common summary statistics include measures of central tendency (mean, median) and measures of dispersion (standard deviation, variance).

The mean represents the average value of a dataset. In the context of anomaly detection, comparing the mean of anomalous instances to that of normal instances can reveal whether anomalies exhibit significantly different behaviour compared to normal data points [10].

The median is the middle value of a dataset when it is sorted in ascending order. Unlike the mean, the median is not influenced by extreme values (outliers). Comparing the median of anomalous instances to that of normal instances provides insights into the central tendency of the data and helps assess the robustness of the findings. The standard deviation measures the spread or variability of the data points around the mean. A higher standard deviation indicates greater variability, while a lower standard deviation suggests that data points are closer to the mean. Analyzing the standard deviation of anomalous instances relative to normal instances helps identify whether anomalies exhibit distinct patterns or exhibit more variability compared to normal data points. For instance, significant differences in mean or median values may indicate distinct patterns or behaviours associated with anomalies. Similarly, differences in standard deviation highlight variations or inconsistencies in the data distribution between anomalous and normal instances [10].

Below are the pseudocode steps for computing summary statistics on cooling rate features for anomalous and normal instances.

Steps 1:

Filter anomalous and normal instances

Step 2:

Compute summary statistics for the cooling rate feature using the described method.

Steps 3:

Compute differences between anomalous and normal statistics.

Performed cross-validation by splitting the data into multiple folds, training the model on different combinations of training and validation sets, and evaluating its performance across folds. The number of combinations of training and validation sets will depend on the value of k (the number of folds). For this case study we have two days of data with intervals of one minute so approximately we have 2880 data points. The number of combinations for a few different values of k (the number of folds) in k-fold cross-validation. So, for 2880 data points we had:

- With 2-fold cross-validation, we got 2 combinations.
- With 5-fold cross-validation, we got 5 combinations.
- With 10-fold cross-validation, we got 10 combinations.
- With 20-fold cross-validation, we got 20 combinations.
- With 50-fold cross-validation, we got 50 combinations.

This ensured that the model's performance was consistent across different subsets of the data and was not sensitive to the specific training-test split [7].

During cross-validation, precision, recall, and F1 scores were used as evaluation metrics. Precision measured the proportion of true anomalies among all instances classified as anomalies by the case study model and averaging precision across all folds provided an overall assessment of the model's ability to correctly identify anomalies while reducing false positives. Recall measured the proportion of true anomalies that were successfully detected by the model and averaging recall across all folds provided information on how well the model captures true anomalies across different subsets of the data. The F1 score was calculated for each fold by combining precision and recall. It provided a balanced measure of the model's ability to identify true anomalies while minimizing false positives and false negatives and averaging the F1 scores across all folds provides an overall assessment of the model's performance [7].

To enhance the model's performance, conducted experiments by trial and error method with various values for the contamination parameter. Based on the evaluation results, fine-tuned this parameter, leading to improved accuracy in detecting anomalies while reducing false positives. Furthermore, we optimized the decision threshold of the model to strike a balance between precision and recall. In this particular case study, our emphasis was on minimizing false positives, prioritizing high precision over capturing all anomalies with high recall. This approach ensured that flagged anomalies were more likely to be genuine, thereby minimizing the risk of unnecessary interventions or false alarms.

Below are the pseudocode steps with different values for the contamination parameter and adjust it based on evaluation results.

Step 1:

List of contamination values to experiment with values 0.01, 0.05, 0.1, 0.15.

Step 2:

Iterate over each contamination value from the list.
Step 3:
Initialize the isolation forest model with the current contamination value.

Below are the pseudocode steps cross-validation to evaluate the Isolation Forest model for anomaly detection in Python using sci-kit-learn.

Step 1:
Define the number of folds for cross-validation
Adjusted the number of folds as needed for this case study
Step 2:
Initialized K Fold cross-validator.
Step 3:
Performed cross-validation first by splitting data into the train and validation sets.
Fit the Isolation Forest model on the training data and predict anomalies on the validation data.
Evaluate the model's performance using precision, recall, and F1-score.
Step 4:
Calculate the average scores across folds.

5 Conclusion

This case study investigates the effectiveness of applying one of the machine learning-based approach algorithms i.e., isolation forest algorithms to detect anomalies indicative of compressor failure in freezer temperature data to predict freezer conditions. Through rigorous experimentation and analysis, several important findings emerged. Visual inspection of the detected anomalies revealed distinct patterns corresponding to periods of temperature fluctuation, providing valuable insights into potential compressor failure approaching scenarios. The statistical analysis results provided different attributes of anomalies detected by the model. The Isolation Forest model demonstrated promising performance in accurately identifying anomalies, as evidenced by high precision. However, challenges were encountered in interpreting anomalies during extreme temperature fluctuations and limited data, highlighting the importance of further investigation and refinement of anomaly detection techniques. Overall, these evaluation metrics highlighted in the result section in 5.1 suggest that while the model achieves perfect precision for one class, it performs poorly in terms of recall, particularly for the second class. The F1-score reflects the balance between precision and recall, but the low values indicate the suboptimal performance of the model in capturing true anomalies while minimizing false positives. Based on model performance results it is impractical for deployment of the model in its current form. Additionally, comparison with alternative methods sheds light on the strengths and limitations of the Isolation Forest approach, informing future research directions. Overall, the study underscores the significance of predictive maintenance techniques in enhancing equipment reliability and reducing downtime, laying the groundwork for further advancements in anomaly detection methodology.

5.1 Results Highlights

The results of applying the isolation forest algorithm to the freezer temperature data revealed anomalies accurately. Visual inspection of the temperature data against anomalies plotted showed clusters of anomalies occurring during periods of fluctuating temperature. The model was tested with freezer 1 temperature data sets and freezer 2 temperature data sets.

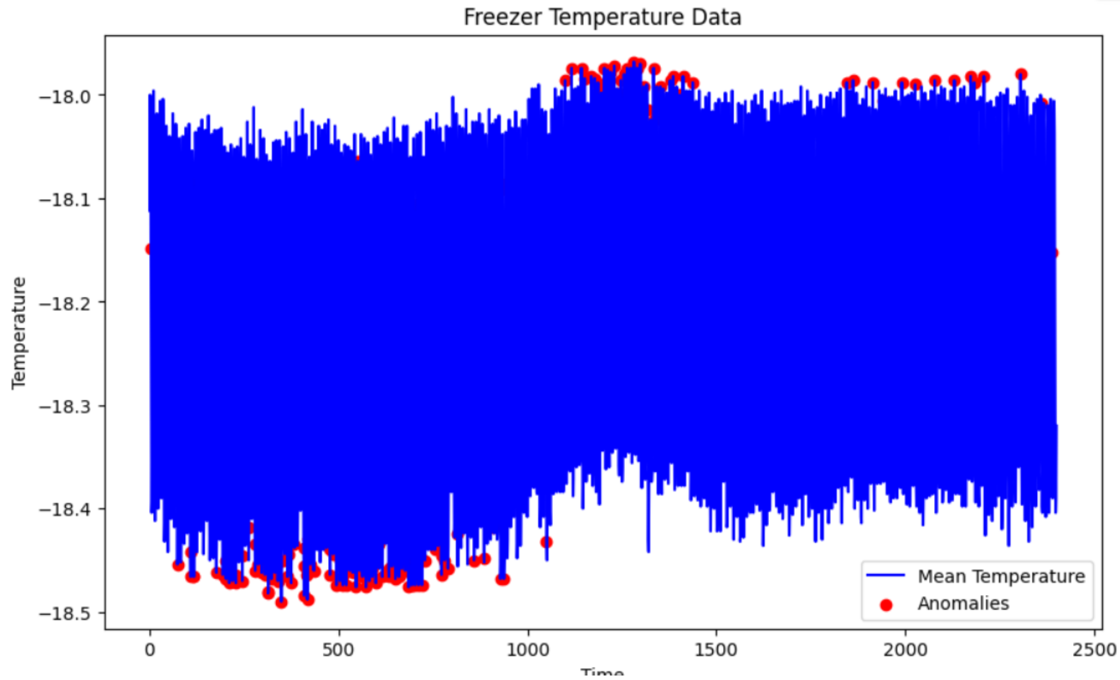

Figure 13: Freezer 1 temperature data with anomalies.

The graph is plotted with the temperature readings against the timestamp, with anomalies marked separately.

In this case, as shown in Figure 13, the horizontal axis represents time, with timestamps indicating when each temperature reading was recorded. The vertical axis represents the temperature readings from Freezer 1. Anomalies detected by the anomaly detection model are highlighted on the graph by red dots. Anomalies that deviate significantly from the normal temperature range pointed to potential issues with the freezer, such as compressor failure or temperature fluctuations.

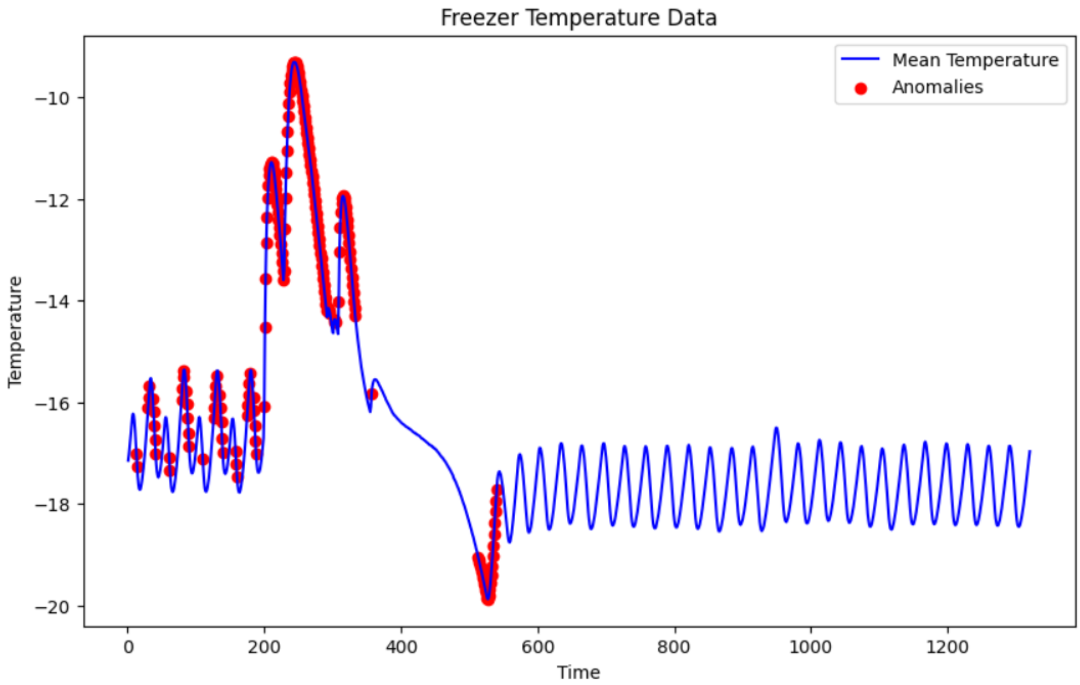

Figure 14: Freezer 2 temperature data with anomalies.

In this case, as shown in Figure 14, the horizontal axis represents time, with timestamps indicating when each temperature reading was recorded. The vertical axis represents the freezer temperature readings from Freezer 2. Anomalies detected by the anomaly detection model are highlighted on the graph by red dots. In this case, we see more anomalies detected at the beginning of the visual inspection period and fewer or no anomalies detected later on this freezer. This provided insights into the performance and condition of the freezer.

Using the statistics analysis techniques in evaluating an anomaly detection model gave a better understanding of distinct patterns exhibited by anomalies compared to normal instances and provided insights into the effectiveness of the case study anomaly detection model.

```
Anomalous Instances Statistics:
count    360.000000
mean      -0.034300
std        0.099068
min       -0.196000
25%       -0.122000
50%       -0.026000
75%        0.041000
max        0.148000
Name: Cooling_Rate, dtype: float64

Normal Instances Statistics:
count    2040.000000
mean        0.006014
std         0.103999
min        -0.172000
25%        -0.094000
50%         0.028000
75%         0.104500
max         0.142000
Name: Cooling_Rate, dtype: float64

Differences between Anomalous and Normal Instances:
Mean Difference: -0.04031372549019588
Median Difference: -0.053999999999998494
Standard Deviation Difference: -0.004930715028001381
```

Figure 15: Statistics analysis summary for freezer 1.

As per Figure 15, the negative mean difference (-0.040) suggests that, on average, the mean value of the feature (e.g., cooling rate) is lower for anomalous instances compared to normal instances. This indicates that anomalies tend to have lower mean values of the feature, which indicates abnormal behaviour in the system.

The negative median difference (-0.053) implies that the median value of the feature is lower for anomalous instances compared to normal instances. This suggests that the central tendency of the feature distribution for anomalies is shifted towards lower values, indicating potential anomalies.

The negative standard deviation difference (-0.0049) indicates that the variability of the feature is lower for anomalous instances compared to normal instances. Anomalies exhibit less variability in the feature values, suggesting more consistent behaviour in anomalous instances. The differences indicate that the model captured anomalies based on the specified features.

```
Anomalous Instances Statistics:
count    198.000000
mean       0.004697
std        0.255953
min       -0.330000
25%       -0.138000
50%       -0.077000
75%        0.162500
max        1.554000
Name: Cooling_Rate, dtype: float64

Normal Instances Statistics:
count    1122.000000
mean       -0.000595
std         0.108948
min        -0.232000
25%        -0.084000
50%        -0.004000
75%         0.104000
max         0.202000
Name: Cooling_Rate, dtype: float64

Differences between Anomalous and Normal Instances:
Mean Difference: 0.005292335115864459
Median Difference: -0.07300000000000306
Standard Deviation Difference: 0.14700520238400436
```

Figure 16: Statistics analysis summary for freezer 2.

In this case, as per Figure 16, the positive mean difference suggests that the mean feature value for anomalous instances is slightly higher than that for normal instances. However, the magnitude of the difference is relatively small, indicating that the two groups are relatively similar in terms of the mean feature value. A negative median difference suggests that the median feature value for anomalous instances is lower than that for normal instances. This indicates that anomalous instances tend to have lower median feature values compared to normal instances. In this case, the positive standard deviation difference indicates that the standard deviation of feature values for anomalous instances is slightly higher than that for normal instances.

The evaluation metrics values such as average precision, average recall, and average F1-score offered insights into the performance of the model.

```
Average Precision: [0.         1.         0.05843212]
Average Recall: [1.         0.         0.30286319]
Average F1-score: [0.         0.         0.09778081]
```

Figure 17: Evaluation metric for freezer 1.

```
Average Precision: [0.         1.         0.00988904]
Average Recall: [1.         0.         0.05244088]
Average F1-score: [0.         0.         0.01658895]
```

Figure 18: Evaluation metric for freezer 2.

Precision measures the proportion of true positive predictions among all positive predictions made by the model. A higher precision indicates fewer false positives. In this case, as per Figure 17 and Figure 18, the second class achieves perfect precision (1), indicating that all predicted anomalies for this class are indeed true anomalies. However, the first class has a precision of 0, indicating that all predicted anomalies for this class are false positives. The third class has a precision of 0.0584, suggesting that only a small proportion of predicted anomalies for this class are true anomalies.

Recall measures the proportion of true positive predictions among all actual positive instances in the dataset. A higher recall indicates fewer false negatives. The first class achieves perfect recall (1), indicating that all true anomalies for this class are correctly identified by the model. However, the second class has a recall of 0, suggesting that the model fails to detect any true anomalies for this class. The third class has a recall of 0.3029, indicating that only a fraction of the true anomalies for this class are detected by the model.

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance. It ranges from 0 to 1, with higher values indicating better performance. In this case, the F1-score is 0 for the second class due to the absence of true positives. For the other classes, the F1-score is nonzero but relatively low, indicating a trade-off between precision and recall.

5.2 Future Works

Based on the findings there are potential avenues for further exploration to isolation forests with other machine learning techniques or ensemble methods to improve anomaly detection accuracy and reliability. By evaluating the performance of the Isolation Forest approach against alternative anomaly detection methods, such as Local Outlier Factor (LOF), One-Class SVM, or clustering-based techniques and comparing the strengths and limitations of each method in terms of their ability to detect anomalies, computational efficiency, scalability, and robustness to different types of data and anomaly patterns. The case company do not use machine learning algorithms for instrument products and also lacks experience in the field of using and deploying ML models to their products. So, further exploring using a platform that could create the model or input the own model and deploy ML models, will promote the usage of ML techniques in suitable cases.

References

1 Mobley, R. Keith. Predictive Maintenance: Tools and Techniques. Elsevier ISBN: 978-0750677769.

2 Andrew Glassner. Deep Learning: A Visual Approach. June 2021, 768 pp. ISBN-13: 9781718500723

3 Jihong Yan, et al. Predictive Analytics for Equipment and Systems Health Monitoring. Springer; ISBN: 978–3319461277.

4 Shan Gao. Predictive Maintenance: A Machine Learning Approach. Springer. ISBN: 978-3319461277.

5 Dr. Yogendra Patil. Predictive Maintenance for Industrial IoT: Solutions for Industry 4.0. Apress; ISBN: 978–1484266860.

6 Unsupervised learning sci-kit-learn 1.4.1 documentation. < https://scikit-learn.org/stable/unsupervised_learning.html#unsupervised-learning>

7 Supervised learning sci-kit-learn 1.4.1 documentation < https://scikit-learn.org/stable/supervised_learning.html#supervised-learning>

8 AWS AI/ML immersion Day - Finland.

9 Amazon SageMaker Machine Learning Immersion Day – Finland.

10 Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning: with Applications in Python. First printing; July 5, 2023.

11 Ibrahim Dincer, Mehmet Kanoglu. Refrigeration Systems and Applications.

12 Varun Chandola, Arindam Banerjee, Vipin Kumar. Anomaly Detection: A Survey. August 15, 2007.

13 Andreas C. Müller, Sarah Guido. Introduction to Machine Learning with Python: A Guide for Data Scientists. Published by O’Reilly Media; Inc.1005 Gravenstein Highway North; Sebastopol; CA 95472.