CHALMERS
UNIVERSITY OF TECHNOLOGY

Predictive maintenance through data-driven decision making

A case study of investigating strategic guidelines for deploying predictive maintenance systems

Master's thesis in Production Engineering

Rasul Rezae
Konstantinos Toulikas

Department of Industrial and Materials Science

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

MASTER'S THESIS 2024

Predictive maintenance through data-driven decision making

A case study of investigating strategic guidelines for deploying predictive maintenance systems

Rasul Rezae
Konstantinos Toulikas

Department of Industrial and Materials Science
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024

Predictive maintenance through data-driven decision making

A case study of investigating strategic guidelines for deploying predictive maintenance systems

Rasul Rezae
Konstantinos Toulikas

Supervisors: Jon Bokrantz and Mohan Rajashekarappa, Department of Industrial and Material science
Advisors: Julia Norén, Parker Hannifin Manufacturing Sweden AB and Johan Hidstam, Parker Hannifin Manufacturing Sweden AB
Examiner: Anders Skoogh, Department of Industrial and Material science

Master's Thesis 2024
Department of Industrial and Material science
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Gothenburg, Sweden 2024

Abstract

This thesis investigates the deployment of Predictive Maintenance (PdM) systems and solutions at Parker Hannifin Manufacturing, a producer of control valves. PdM uses advanced analytics, machine learning algorithms, and real-time sensor data to predict maintenance needs before machine breakdowns occur. This approach contrasts with traditional preventive maintenance (PM), which follows predetermined schedules. The research aims to identify critical parameters for equipment monitoring, the necessary data infrastructure, and the potential cost benefits of PdM implementation. The study involves a case study of spool cell machines at Parker Hannifin Manufacturing, addressing challenges in data collection, algorithm development, and cultural shifts within the organization. The expected outcome is strategic roadmaps for transitioning from PM to PdM. The study concludes that applying PdM can reduce maintenance costs and enhance operational efficiency.

Keywords: Maintenance, Predictive Maintenance, Data Science, Prediction Challenges and Benefits, Natural Language Processing in predictive maintenance, Artificial Intelligence in maintenance.

Acknowledgements

This thesis was written at Parker Hannifin Manufacturing in Borås and at the Department of Industrial and Materials Science at Chalmers University of Technology, Gothenburg, Sweden during the Spring semester of 2024.

First and foremost, we would like to express our gratitude to our supervisors, Jon Bokrantz and Mohan Rajashekarappa, for their invaluable guidance, and encouragement throughout the entirety of this research project. Their knowledge and expertise have helped us a lot in shaping this thesis, and their support has been valuable in overcoming the challenges encountered along the way. We also extend our sincere appreciation to Anders Skoogh, for undertaking the role of being the examiner for this thesis.

A special thanks to the management, especially Julia Norén, EHSE Manager and Johan Hidstam, Line Manager, and employees at Parker Hannifin Manufacturing in Borås for allowing us to do the thesis at their company and for their cooperation and willingness to share their insights and data. Their openness and assistance have provided a solid foundation for this thesis. This thesis is the result of not only our efforts but the collective support and contributions of many individuals. For that, we are truly thankful.

Declaration of AI Technologies in the Thesis Work and Writing Process

In the course of developing this thesis, various AI technologies were utilized to enhance the quality and comprehensiveness of the research and writing process. Specifically, ChatGPT by OpenAI and paraphrasing tools played significant roles in different aspects of the work.

ChatGPT was used as a brainstorming partner to generate ideas and explore different perspectives on predictive maintenance, data-driven decision-making, and the implementation of AI/ML algorithms in industrial settings. This interaction helped in formulating research questions, structuring the thesis, and identifying key areas of focus. Additionally, ChatGPT provided assistance in refining the language, ensuring grammatical accuracy, and improving the overall readability of the text. This involved correcting grammatical errors, rephrasing awkward sentences, and enhancing the clarity and coherence of the writing. Throughout the research process, ChatGPT served as a virtual advisor, offering guidance on finding relevant literature, suggesting methodologies, and providing insights into best practices for deploying predictive maintenance systems. This support was invaluable in navigating complex topics and making informed decisions.

Paraphrasing tools were employed to rephrase content, ensuring originality and preventing redundancy. These tools helped in expressing the same ideas in different ways, thereby enhancing the thesis's uniqueness and depth.

The integration of these AI technologies significantly contributed to the efficiency and effectiveness of the thesis development process. By leveraging AI, we were able to maintain high standards of academic rigor while also ensuring that the content was accessible and well-structured.

Rasul Rezae and Konstantinos Toulikas, Gothenburg, May 2024

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

AE :	Acoustic Emission
AI:	Artificial Intelligence
BoW :	Bag of Words
CBM:	Condition-Based Maintenance
CBPM:	Condition-Based Predictive Maintenance
CNC:	Computerised Numerical Control
CRISP-DM:	Cross-Industry Standard Process for Data Mining
IDF :	Inverse Document Frequency
IoT :	Internet of Things
IT:	Information Technology
KPI:	Key Performance Indicator
ML:	Machine Learning
NER :	Named Entity Recognition
NLP:	Natural Language Processing
OEE:	Overall Equipment Effectiveness
PdM:	Predictive Maintenance
PM:	Preventive Maintenance
RM:	Reactive Maintenance
ROI:	Return On Investment
RUL:	Remaining Useful Life
SBPM:	Statistical-Based Predictive Maintenance
SEK:	Swedish Krona
TF :	Term Frequency
TF-IDF :	Term Frequency - Inverse Document Frequency
TLP:	Technical Language Processing

Contents

List of Acronyms **x**

List of Figures **xv**

1 Introduction **1**
1.1 Background 1
1.2 Aim 2
1.3 Research Questions 2
1.4 Process and Problem Descriptions at Case Company 3
1.5 Delimitations 3

2 Theoretical Background **5**
2.1 Maintenance 5
2.1.1 Reactive Maintenance 5
2.1.2 Preventive Maintenance 6
2.1.3 Condition-Based Maintenance 6
2.1.4 Predictive Maintenance 7
2.2 PdM Techniques 8
2.2.1 Condition-Based Predictive Maintenance 8
2.2.2 Statistical-Based Predictive Maintenance 8
2.3 ML and AI for PdM 9
2.4 Data collection methods 9
2.5 AI Techniques for textual data analysis 12
2.5.1 Natural Language Processing 12
2.5.2 Technical Language Processing 16
2.6 Implementation Challenges of PdM 18

3 Methods **21**
3.1 Case study approach 21
3.1.1 Step 1 21
3.1.2 Step 2 21
3.1.3 Step 3 22
3.1.4 Step 4 22
3.1.5 Step 5 22
3.2 CRISP-DM 23

3.2.1 Business Understanding . . . 24
3.2.2 Data Understanding . . . 24
3.2.3 Data preparation . . . 24
3.2.4 Modeling . . . 25
3.2.5 Evaluation . . . 25
3.2.6 Deployment . . . 25
3.3 Interviews . . . 26
3.4 Literature Study . . . 27

4 Results **29**
4.1 Company's Current State . . . 29
4.1.1 Maintenance Situation . . . 29
4.1.2 Data Infrastructure . . . 30
4.2 Challenges . . . 30
4.3 Solutions . . . 32
4.3.1 Solution 1: Approach through investing in sensors . . . 32
4.3.2 Solution 2: Approach through NLP-TLP techniques . . . 35
4.4 Solution Cost Analysis . . . 37
4.4.1 Solution 1 . . . 37
4.4.2 Solution 2 . . . 37
4.5 Cost Benefit of PdM . . . 39

5 Discussion **41**
5.1 Comparison with literature . . . 41
5.2 From theory to practice . . . 42
5.3 Re-Evaluation of ROI . . . 43
5.4 Existing Predictive Maintenance Services . . . 44
5.5 Generalizability of Results . . . 45
5.6 Answers to the Research Questions . . . 46
5.7 Reflections on research methods . . . 48
5.8 Cost-benefit and cost calculations assumptions . . . 49
5.9 Ethical and Sustainability Aspects . . . 50
5.10 Future suggestions . . . 52

6 Conclusion **53**

Bibliography **55**

A Appendix 1 **I**

B Appendix 2 **III**

List of Figures

2.1 Applications of the different NLP approaches (Inspired by Zhonga et al., 2024) 13
2.2 An example of NLP workflow (Inspired by Brundage et al., 2020). 14
2.3 A conceptual diagram of a TLP workflow (Inspired by Brundage et al., 2020) 17

3.1 CRISP-DM Phases for Data Mining - Analysis (Inspired by Wirth and Hipp, 2000) 23

4.1 Roadmap for implementing CBPM using sensors. 33
4.2 Roadmap for implementing PdM using NLP-TLP techniques. 36

1
Introduction

This chapter offers a brief introduction to different maintenance strategies, including reactive, preventive, and predictive maintenance, and explains the existing problems at the case company, Parker Hannifin Manufacturing*. The chapter also states the aim of the study. Additionally, it outlines the key research questions that drive the investigation and discusses the delimitations of the study to define its scope and boundaries.

**For the rest of the project, Parker Hannifin Manufacturing will be stated as Parker.*

1.1 Background

In recent years, the world has seen rapid technological advancement in many areas. In the context of manufacturing, industries also have seen significant changes in their production and maintenance strategies (Meddaoui et al., 2023). To have profitable and productive manufacturing, machine availability (uptime) is an important key factor. This will make sure that the production is according to plans, and the traditional ways of maintenance, meaning reactive maintenance is not enough. Reactive maintenance means that the machines are repaired only when the machines are broken (Poor et al., 2019). This approach can result in unexpected downtime, increased repair costs, and disruptions to production schedules.

Another alternative is preventive maintenance, meaning that the maintenance is planned and scheduled to address the potential problems in the machines before they escalate (Poor et al., 2019). This aims to reduce downtime and costs, however, this method may also result in unnecessary expenses by doing maintenance for perfectly working machines and it does not always account for the actual condition of the equipment. Preventive maintenance is often characterized by routine inspections, part replacements, and system overhauls in already-determined intervals. While effective in minimizing unscheduled downtime, it may not be the most cost-efficient strategy. Consequently, according to Lee et al. (2014) the maintenance paradigm is undergoing big changes, meaning moving from a reactive problem-solving approach to a more predictive maintenance model.

Predictive maintenance, as it sounds, means that the need for maintenance in a machine is predicted before machine breakdown (Poor et al., 2019). This model uses a combination of advanced analytics, machine learning algorithms, and real-time data measured by sensors on machines. These sensors monitor various parameters like vibration, sound, and other different conditions. By continuously analyzing the data measured, companies can gain insight into the health of their machines and this enables them to schedule the maintenance activities exactly when needed resulting in minimizing downtime and optimizing the resource allocation (Poor et al., 2019).

The advantages of predictive maintenance over preventive maintenance is clear as it provides the ability to move beyond predetermined maintenance schedules and address problems only based on the machine's actual condition. By implementing a predictive maintenance model, industries, except reducing operation disruptions in their production, can also increase the lifespan of their equipment, ultimately leading to more enhanced efficiency and cost savings (Poor et al., 2019).

Parker, a producer of directional control valves, remote control valves, and auxiliary valves specifically tailored for the mobile market in Borås, intends to shift from preventive maintenance to predictive due to its potentially big benefits but as if now, it appears that there may be a gap in their current knowledge or expertise to carry it out, hence, they are seeking assistance to accomplish that.

1.2 Aim

The aim of this project is to explore and establish strategic guidelines for transitioning from reactive and preventive maintenance to predictive maintenance across various industries. This project includes a case study of Parker Hannifin Manufacturing to illustrate the application of these guidelines in a real-world context. By leveraging data-driven decision-making and advanced AI/ML algorithms, the goal is to enhance maintenance strategies, optimize resource allocation, and minimize operational disruptions in a wide range of manufacturing settings.

1.3 Research Questions

Based on the aim of the thesis, the following research questions are specified for investigation:

RQ1: How can Parker develop a predictive maintenance solution for their types of machines?

RQ2: Which maintenance decisions should be data-driven and what data is needed for these decisions, and how should this data be collected and stored?

RQ3: What are the potential cost savings and benefits Parker can achieve by transitioning from preventive to predictive maintenance strategies, and what challenges must be overcome during this transition?

1.4 Process and Problem Descriptions at Case Company

Due to the broad project topic, it was decided to narrow down the scope of the thesis and focus on two specific and identical machines to start the implementation of predictive maintenance. These machines are called spool cells, which consist of three different part-machines. The first one is the Stångmagasin (bar feeder) which handles the bars and cuts them to the desired size depending on the customers' orders and products. Then, the processed bars are gripped through a robot (second machine) and are put into the CNC machine, which is the third machine. In the CNC, the bars are processed to get their final shape and their details according to the customers' specifications. After the CNC operation is done, the components are stored and further moved to the next station for further processing - assembly to be added to the final product. The reasons why the spool cell machines were chosen to be the main focus is due to recent problems and errors that these machines have faced which caused significant losses for Parker in the form of downtime and maintenance costs and delays in meeting customer demand.

1.5 Delimitations

To define a concrete scope and also to have a manageable research framework, this study had several delimitations. The study was done during 20 weeks, starting from the first of February 2024 until the end of May 2024. As mentioned earlier, the primary focus was on spool machines, more specifically on the examination of two identical spool cells and the data gathered was based on these machines' maintenance logs and the data from the interviews.

If the selected machines failed to meet the requirement or were unable to provide the essential data that caused the project to become stagnant, the focus would be shifted to another CNC machine to ensure that the research progressed efficiently. The financial budget that was thought to be invested in this project was within Parker's policies, meaning that in case of investment, the company should have had a Return On Investment (ROI) at most within 2-year time frame.

2

Theoretical Background

This chapter aims to delve into the theoretical framework necessary for understanding the study. It provides an in-depth discussion of various maintenance strategies, including reactive maintenance, preventive maintenance, condition-based maintenance, and predictive maintenance. The chapter also explains the key techniques and technologies involved in predictive maintenance, such as machine learning, artificial intelligence, and data collection methods. Furthermore, it explains relevant concepts like Natural Language Processing and Technical Language Processing in the context of maintenance. Lastly, this chapter also examines the challenges of implementing predictive maintenance systems.

2.1 Maintenance

According to CEN, which is the European committee for standardization, defines maintenance as:

"A combination of all technical, administrative and managerial actions during the life-cycle of an item, and intended to remain it in, or restore it to a state in which it can perform the required function" ("CEN - CENELEC - Search standards", n.d.)

In simple words, maintenance includes all stages of repair and modification of any item to keep it in optimal working condition or to repair it to the optimal working condition. This means that every sector, from manufacturing to transportation and healthcare, utilizes some form of maintenance to maintain their operations. In the following sections, the most common type of maintenance will be described.

2.1.1 Reactive Maintenance

Reactive maintenance (RM), also known as "corrective" or "breakdown maintenance" as it sounds, is an approach in which maintenance activities are performed as a response to a breakdown or equipment failure (Poor et al., 2019). It is essentially the different industry's first natural reaction to equipment malfunctions, where repairs are conducted only after the problem has already occurred. This maintenance approach emerged as the earliest form of maintenance in industries, primarily because it was a straightforward and only natural first response to immediate issues without requiring complex planning or scheduling.

RM has both benefits and drawbacks. The advantages of this approach include lower initial costs and less planning which make it easy to implement (Coanda et al., 2020). However, the unpredictable nature of breakdowns of the equipment can lead to unscheduled stops and that could entail safety issues, inefficient use of resources, and potentially higher expenses in the long run (Coanda et al., 2020). Due to its limitations, other maintenance approaches such as preventive and predictive was evolved over time.

2.1.2 Preventive Maintenance

According to Poor et al. (2019) the origin of preventive maintenance (PM) can be traced back to the industrial developments of the Second Industrial Revolution, which began around 1870. As industries went towards electrification and assembly line production, machinery became more complex in the manufacturing processes. This complexity led to higher expenses and interruptions caused by equipment breakdowns, which made the need for proactive maintenance strategies important.

PM is a proactive approach for maintaining machinery, equipment or systems. Instead of waiting for breakdowns or malfunctions as in the RM approach explained in the previous section, PM involves scheduled inspections, adjustments, and replacements to prevent problems from occurring (Poor et al., 2019). Some of the key aspects of PM include regularly planned activities to keep the equipment in optimal condition, checking in detail to identify wear, damage, or potential problems, applying lubricants to moving parts to reduce friction and wear, and lastly swapping out worn components or parts before they fail (Swanson, 2001). The schedules are usually based on estimating the probability of equipment failure occurring within the specified interval (Swanson, 2001).

PM similar to RM has both advantages and disadvantages. This approach minimizes the likelihood of machine breakdowns which result in less downtime leading to a more cost-effective approach than RM (Poor et al., 2019). But the downsides of PM are worth considering. With this approach there is a potential for over-maintenance, where unnecessary replacements and repairs may occur, leading to extra expenses (Poor et al., 2019). Further, following a strict maintenance schedule can also disrupt workflow and cause delays in overall operational efficiency. Plus, accurately planning these tasks is not always straightforward as there is a need for the right planning to have the extra and correct spare parts in the warehouse. Lastly, there is the added labor needed for regular inspections and maintenance work, which drives up operational costs. These challenges highlight the need for a more precise maintenance strategy that strikes a balance between the benefits of PM and its potential drawbacks.

2.1.3 Condition-Based Maintenance

One of the techniques in maintenance is condition monitoring, where the health or condition of the machines are monitored. Monitoring the condition of machines can vary significantly, from simple techniques such as hearing or seeing the machines' behavior to judge if something is wrong, to using sensors that monitor the condition of the machines in real-time. One of the increasingly popular proactive maintenance approaches is called Condition-based Maintenance (CBM). CBM involves selecting and tracking conditions such as vibrations, temperatures, oil degradation, and electrical variables of equipment, etc. through sensors (Acernese et al., 2020). An algorithm computes the current health condition of the equipment through the sensor data collected in real-time, and unless the equipment indicates faulty behavior, it will continue to operate. Once anomalies are detected the maintenance will be scheduled.

This approach which initially was called predictive maintenance was first introduced by the Rio Grande Railway Company in the late 1940s (Prajapati et al., 2012). CBM techniques were used by the railway company to monitor engine functions, detecting coolant, oil, and fuel leaks through analysis of temperature and pressure data. This approach proved to be very successful, minimizing the impact of unexpected failures and providing timely guidance on when to fix leakage or replenish coolant and oil supplies (Prajapati et al., 2012).

An accurately implemented CBM approach has without a doubt many benefits as mentioned earlier. However, this approach like other maintenance strategies has its drawbacks. The initial expenses associated with implementing CBM strategies, such as acquiring new tools, test equipment, etc. can be high. Although CBM can lead to long-term cost savings, its initial investment may be challenging, especially for smaller companies with limited budgets (Gillespie, 2015).

CBM also requires a large base of information and experience to make the necessary changes for existing and future programs, which can pose challenges in legacy systems (Gillespie, 2015). Legacy systems refer to the existing older technology systems and software within an organization or industry. Legacy systems may face additional challenges in CBM implementation, such as inadequate data collection and analysis capabilities, lack of standardized data management technologies, and the need for comprehensive training for maintenance personnel (Gillespie, 2015).

2.1.4 Predictive Maintenance

As factories and industries change and the need for productivity and efficiency grows doing maintenance efficiently becomes even more important for cost savings and optimizing workflow (Li et al., 2024). In earlier sections, RM and PM approaches and their benefits and drawbacks were mentioned. Due to these drawbacks and due to technological advancements, another maintenance approach, namely predictive maintenance (PdM) has emerged which is the latest form of maintenance approach as of today (Poor et al., 2019).

PdM as it sounds is a proactive approach to equipment and machine maintenance. It focuses on using technology to predict when maintenance is likely to be required based on patterns and data from a machine's past performance (Li et al., 2024). This can allow maintenance to be scheduled exactly before a failure occurs, in contrast to PM which uses fixed schedules, thus saving time, reducing downtime, and avoiding unexpected costs. This proactive approach can be done with different techniques, approaches, and algorithms each offering unique insights into equipment health and performance.

What allows the transition from the traditional maintenance strategies, meaning RM and PM to PdM is the evolution of new technologies such as the Internet of Things (IoT), Big Data, Artificial Intelligence (AI), etc (Zonta et al., 2020). IoT refers to a network of physical devices and other objects that are embedded with sensors, software, and network connectivity. These smart devices can collect and share data and allow them to communicate with each other (Atzori et al., 2010). Also, Big Data refers to very large data sets that cannot be processed or analyzed using traditional data processing techniques. It often involves the use of advanced analytical tools and techniques such as machine learning (ML) to extract insights and patterns from the data (Zhang et al., 2017).

2.2 PdM Techniques

PdM can be divided into different groups or categories based on the specific methodologies, techniques, and technologies employed to predict equipment failures and optimize maintenance activities. Generally, it can be divided into two groups which are described in the following subsections.

2.2.1 Condition-Based Predictive Maintenance

As mentioned in the previous section CBM involves the real-time monitoring of equipment parameters like vibration, temperature, pressure, or other conditions using sensors. The primary objective of CBM is to continuously monitor equipment conditions and detect deviations from normal operating conditions. When certain parameters exceed predefined thresholds, CBM triggers immediate alerts or alarms to notify maintenance teams of inspection or actions. On the other hand, Condition-Based Predictive Maintenance (CBPM) builds upon CBM by integrating predictive analytics and algorithms in addition to using sensor data. CBPM goes beyond immediate monitoring and incorporates advanced data analysis techniques, such as ML, to predict future equipment failures based on real-time condition data and historical data and performance trends (Kaur et al., 2018).

2.2.2 Statistical-Based Predictive Maintenance

Statistical-based predictive maintenance (SBPM) relies on statistical analysis techniques to predict equipment failures and optimize maintenance activities. This approach involves analyzing historical data and applying statistical models to forecast future maintenance needs based on patterns and trends in the data (Ghani, 2021). Techniques used in SBPM include regression analysis, time series analysis, survival analysis, and reliability analysis (Mishra et al., 2019). These methods are complemented by machine learning algorithms for enhanced data analysis. Together, they enable the development of predictive models that drive proactive maintenance strategies.

2.3 ML and AI for PdM

At the core of a PdM solution is an algorithm specifically designed to analyze data and perform key tasks such as identifying anomalies, diagnosing equipment issues, forecasting the remaining useful life (RUL) of machines, and predicting when maintenance will be needed. The algorithms can be made with different ML concepts each suited for different types of tasks and data scenarios. For PdM, machine learning concepts can be divided into two classes (Paolanti et al., 2018) which will be explained in this section.

- **Unsupervised Learning:** It is a type of ML in which the model discovers patterns and relationships within the data without labeled supervision or guidance (Carvalho et al., 2019). In PdM, unsupervised learning algorithms are employed for fault detection by learning the normal behavior patterns of a machine and identifying any deviations that may indicate a possible fault. According to Amruthnath and Gupta (2018), several unsupervised learning algorithms can be applied in PdM, including PCA T^2 statistic, hierarchical clustering, K-Means, Fuzzy C-Means clustering, and model-based clustering. The study concludes that applying unsupervised learning algorithms in PdM can result in early fault detection, reduced maintenance costs, and improved product quality. PCA T^2 is a multivariate statistical analysis and is one of the oldest and most widely used algorithms for fault detection. Other algorithms like hierarchical and K-Means clustering are commonly used for unsupervised learning methodologies in industries (Carvalho et al., 2019).
- **Supervised Learning:** It is a type of ML where the algorithm is trained using a set of labeled data to make predictions or classify new data. In supervised learning, the dataset contains both the input variables (predictors) and output variables (response) for the algorithm to learn (Nasteski, 2017). The algorithm looks for patterns and relationships between the input variables and the output variable to create a model for making future predictions. Several algorithms can be used in supervised learning, including linear regression, decision trees, Naive Bayes, and logistic regression (Carvalho et al., 2019).

2.4 Data collection methods

Effective data collection is a crucial foundation for implementing successful PdM strategies in industrial settings. For this, several methods can be used. The most common type of data collection method for PdM includes using sensor data. This includes a range of different sensor technologies that can monitor the condition of assets in different ways. One such sensor that is widely used is a vibration sensor which can provide valuable insight into the mechanical condition of rotating and reciprocating machinery, such as motors, pumps, gearboxes, and turbines (Schwendemann et al., 2021). Small changes in vibration patterns can indicate the possibility of breakdowns and faults happening long before they escalate into significant issues that could lead to catastrophic failures and unplanned downtime (Schwendemann et al., 2021). The ability to detect these early signs makes vibration monitoring a cornerstone of PdM strategies. There are several types of vibration sensors:

- **Accelerometers:** These are the most widely used sensors for vibration monitoring. They measure the acceleration of a vibrating component and convert this information into electrical signals. These sensors can detect a wide range of frequencies and are highly sensitive to fine vibration patterns, making them suitable for detailed analysis of machinery conditions.

- **Velocity sensors:** These sensors measure the velocity of vibrating components. They are less sensitive to high-frequency vibrations compared to accelerometers but are effective in assessing overall vibration levels. Velocity sensors are particularly useful for monitoring medium to low-frequency vibrations often encountered in large rotating machinery (Bogue, 2013).

- **Displacement sensors:** These sensors measure the displacement or movement of a vibrating part. They are typically good for axial displacement measurements and are useful in extreme environments (Bogue, 2013).

One other method for data collection is using thermal techniques which are based on the principle that most mechanical and electrical faults in equipment are often accompanied by changes in temperature. Monitoring temperature variations and identifying thermal anomalies helps in detecting early signs of wear, misalignment, and potential failures (Bogue, 2013). Key thermal techniques used for data collection include thermography and infrared pyrometry. Thermographic instruments typically use uncooled microbolometers as sensors (Bogue, 2013). These instruments capture infrared emissions and produce false-color images (thermograms) that illustrate temperature variations.

Infrared pyrometry offers a cost-effective alternative to thermography for non-contact detection of hot spots and overheating components (Bogue, 2013). It uses conventional infrared detectors operating over a similar wavelength range as thermographic instruments. Though they do not produce thermal images, they offer the advantage of continuous thermal surveillance and infrared pyrometry systems can monitor multiple locations simultaneously by connecting several sensors to a centralized control module

Another technique that can be used is the Acoustic Emission (AE) technique which measures stress waves at frequencies significantly higher than those monitored by traditional vibration methods, extending well above the audio range into the hundreds of kHz and even beyond 1 MHz (Bogue, 2013). These stress waves are typically generated by phenomena such as cracks, fiber breakage, delamination in composites, or impacts like those occurring when a bearing's rolling elements pass over surface irregularities. AE sensors rely on piezoelectric ceramic sensing elements that generate an electrical charge when subjected to mechanical stress. Unlike accelerometers that respond to overall mechanical movement, AE sensors specifically respond to stress waves acting directly on the sensor housing. These sensors are designed to be highly sensitive within the range of 50 kHz to 1 MHz, making them excellent for detecting surface displacements as small as $10^{-}13$ meters (Bogue, 2013).

Additionally, there are techniques for analyzing wear particles, produced by machines, tools, and artificial joints within a lubricant known as wear debris monitoring which can provide early warning signs of component wear and potential failure (Bogue, 2013). Several methods and technologies are employed in wear debris monitoring, ranging from simple magnetic plugs to inductive sensors. Magnetic plugs are placed in lubrication systems to attract and retain ferromagnetic wear particles. These particles are periodically removed and inspected visually or microscopically to identify wear types and quantities (Bogue, 2013). Inductive sensors detect metallic particles passing through a lubricant-carrying tube. These sensors use inductive coils arranged in a bridge configuration to identify ferrous and non-ferrous particles based on their magnetic properties and conductivity. The changes in inductance and power loss from the coil indicate the presence and quantity of wear debris (Bogue, 2013). Other inductive sensors calculate particle size and concentration, offering continuous monitoring and real-time data on wear conditions.

In addition to sensor-based data, other types of data can be collected which are useful for applying PdM. These data play a pivotal role in PdM. Analyzing historical records of equipment performance, maintenance activities, and failure events offers valuable insights into patterns and trends that can predict future failures (Bonnevay et al., 2019). These data can be sourced from various records and data sets:

- **Maintenance Logs**: It contains detailed records of past maintenance activities, including dates, descriptions of failure and work performed, parts replaced, and the results of inspections. Analysis of maintenance logs helps identify recurrent issues and the effectiveness of past maintenance actions.

- **Inspection Reports:** Regular inspection reports provide information on the condition of the equipment as assessed during routine checks, including notes on observed wear, abnormalities, and recommendations for maintenance.

- **Performance data:** Collecting different performance data such as energy consumption, number of products produced in a given timeframe, Overall Equipment Effectiveness (OEE), etc. can provide insights into how equipment is functioning under normal and varying conditions.

By analyzing these data it is possible to recognize patterns of recurring breakdowns and correlate them to their underlying causes and by applying AI and ML concepts a predictive model can be built (Sipos et al., 2014).

2.5 AI Techniques for textual data analysis

In this section AI techniques for textual data analysis and their steps of preprocessing, including their advantages and disadvantages are presented. Based on the literature there are two techniques for analyzing textual data, which are described further in this section.

2.5.1 Natural Language Processing

In general, there are several manufacturing companies that have a huge number of maintenance logs and reports that contain maintenance errors, maintenance time, maintenance process, etc, which can help maintenance technicians work smoothly in the production lines (Zhonga et al., 2024). These logs are accompanied by textual information that provides new opportunities to explore other industrial maintenance approaches. Nowadays, with the rapid development of AI this textual information can be analyzed using Natural Language Processing (NLP) approach. NLP is a component of AI that has a huge potential on analyzing textual data and extract any insights. It consists of two components; natural language understanding and natural language generation. In industry, NLP can handle maintenance texts by mining textual data, extracting features and manipulating information. It can be used for failure prediction, cause analysis and decision support. By implementing NLP techniques for prediction, manufacturing industries can increase the OEE of the production lines, reduce downtime and costs, and enhance their effectiveness (Zhonga et al., 2024).

In industrial maintenance, NLP approaches can be grouped into four categories that each of these approaches is linked to one or more applications as shown in the figure 2.1:

- **Rule-based Approach**
- **Word Embedding Approach**
- **Knowledge graph Approach**
- **ML/Deep Learning Approach**

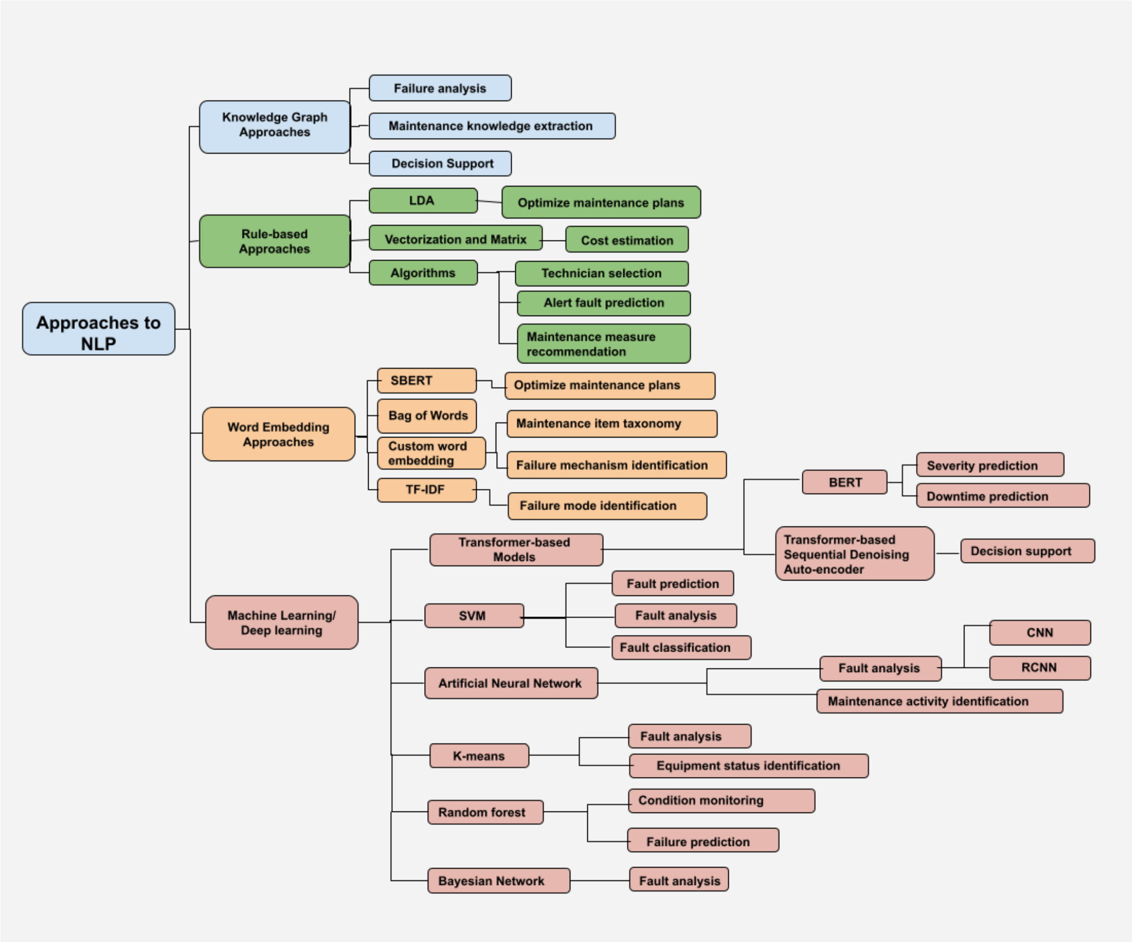

Figure 2.1: Applications of the different NLP approaches (Inspired by Zhonga et al., 2024)

As it is shown in the figure 2.1 there are several applications for each one of the approaches, such as fault analysis, cost estimation, failure prediction, downtime prediction etc. The most relevant approach for the scope of this project is the ML that have an application in the failure prediction.

A NLP workflow have standardized steps that are followed respectfully for the best outcome. In the figure below (see figure 2.2), the whole process or workflow of the NLP approach powered by ML it is presented.

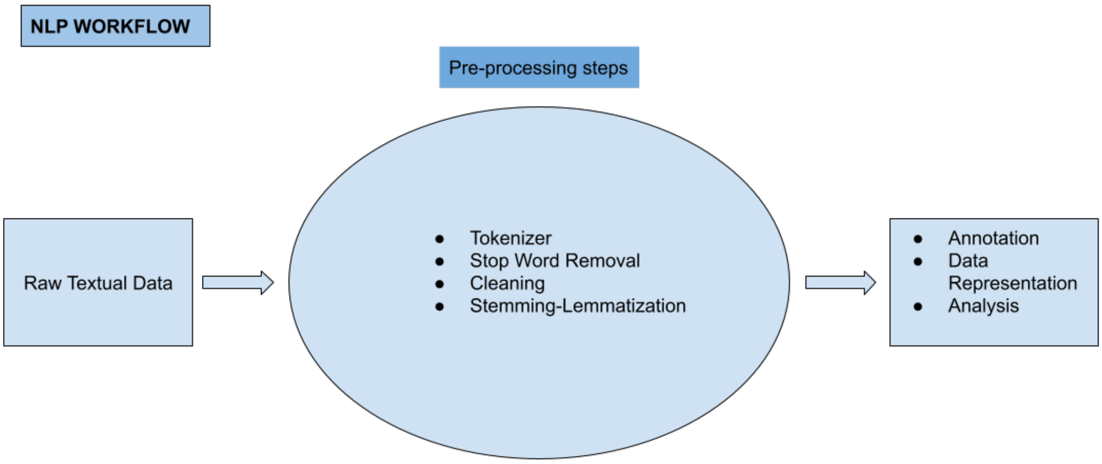

Figure 2.2: An example of NLP workflow (Inspired by Brundage et al., 2020).

The preprocessing steps of NLP as shown in Figure 2.2, are:

- **Tokenization**
  Tokenization is the process of separating the text into individual words or numbers (Brundage et al., 2020); Tabassum and Patil, 2020). For instance, the sentence "Prsure sensor is not properly connected and resulted to an error!" will be tokenized as "{Prsure}{sensor}{is}{not}{properly}
  {connected}{and}{resulted}{to}{an}{error}{!}". This step helps in filtering out unwanted words in further processing steps.
- **Stop Word Remover**
  Words such as "and", "so", "not", "the", do not have any value, expect of some specific cases. For example, these words may not enhance the meaning of the sentence that is provided and therefore are removed from the sentence. So, the previous sentence will now be: "Prsure sensor properly connected resulted error !", which is partially wrong because in that case the word "not" was enhancing the meaning of the sentence (Tabassum and Patil, 2020; Brundage et al., 2020).
- **Cleaning**
  In this step, the removal of punctuation, characters and numbers is taken place to reduce noise that can impact the analysis (Tabassum and Patil, 2020; Brundage et al., 2020). In some cases, this step can remove important information, such as abbreviations or the short definition of a part-item (Random example: "TRL321") (Brundage et al., 2020). Applying this step to the previous sentence, the result will be: "Prsure sensor properly connected resulted error".
- **Stemming-Lemmatization**
  This step is focused on reducing the complexity of the words form to a common base form (Brundage et al., 2020). For example, "pressure,prsure,pressures,pressur/e" $\rightarrow$ "pressure".
  In some cases, the stemming/lemmatization process has to be precise because some important words are not narrowed down to a common base form. So, in this step, the sentence example will be: "Pressure sensor properly connected resulted error".

After the preprocessing step is completed, then the preprocessed text is going to the text analysis where the further analysis of the data is taken place and is where the decisions are made.

The text analysis starts with the **annotation**, which is a process of labelling the unstructured text data (Brundage et al., 2020). This process makes the text data machine-readable and structured. However, in a large dataset, this process may be time consuming and costly because some labels have to be inserted manually.

After the labelling process, the data has to be converted into a standard **data representation** also referred as feature extraction, for processing by the analytic algorithm (Brundage et al., 2020; Tabassum and Patil, 2020). There are different techniques that are used in feature extraction such as Bag-of-Words (BoW), Named Entity Recognition (NER), and Term Frequency - Inverse Document Frequency of records (TF-IDF) (Tabassum and Patil, 2020).

BoW technique is a method to identify important features in the text. These features are then used to classify or categorize the text. The idea behind how BoW is working is that it counts how many times each word appears in a text or a document (Topper, 2023; Tabassum and Patil, 2020). The technique does not consider the order of the words, just how often each one shows up. The concept of BoW is that when a document or a text has similar words to another document or text then it is assumed that the meaning is also similar. This is useful for tasks like sorting texts/documents into specific categories, filtering spam, or finding the matching keywords (Tabassum and Patil, 2020; Topper, 2023).

NER is a technique that identifies and classifies specific names in a text or document, such as names of organizations, places, people, dates, and any other named entity. By doing this NER transforms huge textual context into organized datasets that are ready for the next analysis step. This technique provides a solution when there is a need: find and use specific names as part of information retrieval or get a quick overview of the key entities mentioned in the text or the document (Tabassum and Patil, 2020; Editor, 2023).

TF-IDF is a method that is used to reduce the importance of common words that are not very useful in understanding the content of a text or document. This method consists of two parts: Term Frequency (TF) and Inverse Document Frequency (IDF) (Tabassum and Patil, 2020; GeeksforGeeks, 2023; Kim and Gil, 2019).

- **Term Frequency (TF)** measures how often a word appears in a document. To ensure that this measurement works fairly for both short or long documents and is not affected by the document's length, TF is normalized by dividing the number of occurrences of a term (word) by the total number of words in the document (Tabassum and Patil, 2020; GeeksforGeeks, 2023; Kim and Gil, 2019).
- **Inverse Document Frequency (IDF)** determines how important a word is by considering how rare or common is it in the text or documents. This technique is done by taking the total number of documents / texts and dividing it by the number of documents that contain the word and then taking the logarithm (base 2) of that value. With this technique, common words that are found in many documents are given less value or considered less important (Tabassum and Patil, 2020; GeeksforGeeks, 2023; Kim and Gil, 2019).

The combination of these two techniques gives a score that reflect on how important a word is within all these specific documents/texts, the balance of the frequency and the uniqueness across multiple documents/texts (GeeksforGeeks, 2023; Kim and Gil, 2019).

2.5.2 Technical Language Processing

According to Brundage et al. (2020), Technical Language Processing (TLP) is an iterative approach of customizing NLP tools to engineering data, in other words to use NLP tools efficiently to handle engineering textual data. In the figure below (See Figure 2.3), there is the diagram of the TLP solution and what should be included in such a system.

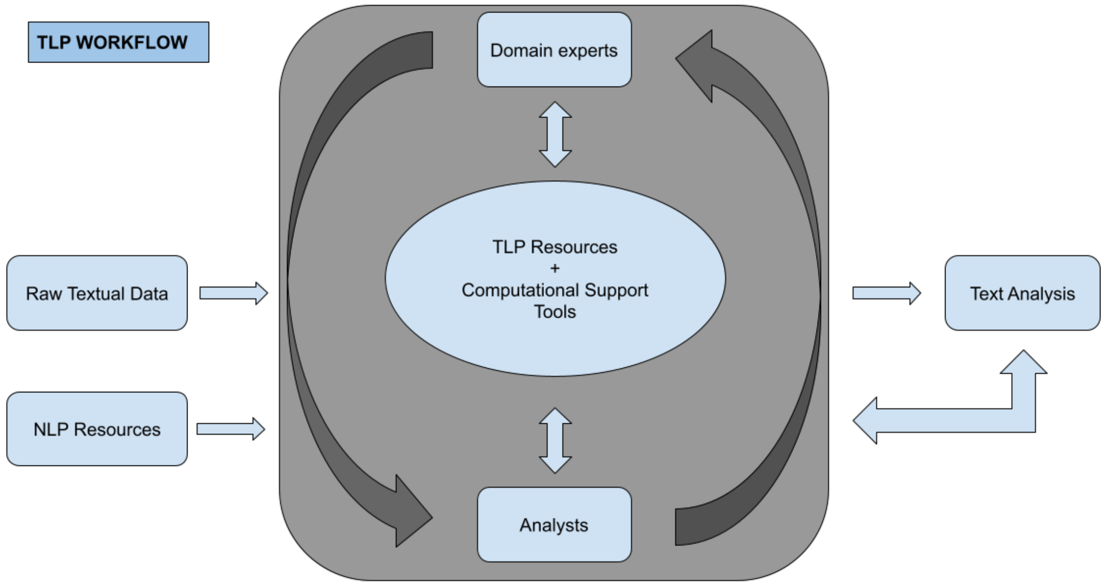

Figure 2.3: A conceptual diagram of a TLP workflow (Inspired by Brundage et al., 2020)

In this diagram, there are the following elements (Brundage et al., 2020):

- **Raw Text**, which is the textual data that are written by an operator or a technician.
- **NLP Resources**, such as the tokenizers and the embeddings, that are the foundation for building the TLP Resources.
- **Computational Support Tools**, that helps the domain experts to do their work efficiently.
- **Domain Experts and Analysts**, that collaborate and improve the TLP Resources with the use of the computational support tools and they increase trust in their analyses.
- **TLP Resources**, such as representations, dictionaries, datasets that are always updated by the domain experts and the analysts in a reproducible way.
- **Text Analysis**, which has an impact on the future development of the TLP Resources.

Following this diagram, companies can use TLP techniques to unlock valuable insights that are hidden in the text data, offering critical information that can aid in maintenance decision making. However, the concept of TLP is still under development and the maintenance community use different resources in order to make TLP a reality (Brundage et al., 2020).

2.6 Implementation Challenges of PdM

Implementing a PdM solution can offer significant benefits, including reduced downtime, lower maintenance costs and improved equipment efficiency. However, the process is complex and presents several challenges and the most relevant to the core content are presented below.

Data Quality and Availability

- **Insufficient Historical Data:** Reliable PdM systems rely on historical data to identify patterns. In some cases, companies do not have sufficient historical data on equipment failures and maintenance (Compare et al., 2020; Achouch et al., 2022).
- **Data Integrity:** Data that are inaccurate, incomplete or inconsistent can lead to unreliable predictions and further to incorrect decisions. In companies sometimes, the data that are reported may miss some information or a measurement was reported incorrectly leading to this issue (Achouch et al., 2022).
- **Sensor Integration:** One of the most common solution that companies mention when thinking about PdM is the reconstruction of the existing equipment with the necessary sensors. This is actually a huge challenge because the placement of sensors needs to be in the best possible position for an accurate measurement and this can not be told for certain because machines vary, from either very old ones, newly purchased etc (Achouch et al., 2022).

Technical Complexity

- **Advanced Analytics:** Data science and ML expertise is required for developing accurate predictive models, which may not be readily available within the industries. In other words, some industries may not have a data department and data analysts that can handle the tasks of implementing the PdM, analyze the data and generally ensure the smooth flow of the predictive system (Achouch et al., 2022).
- **Integration with existing systems:** Ensuring that a PdM solution integrates seamlessly with the existing IT systems can be complex. Some IT systems may be outdated based on the current technology that is used worldwide, making the new technologies a bit challenging to be implemented in the old ones or collaborate with them (Achouch et al., 2022).

Cost

- **Initial Investment:** Companies should acknowledge that the cost of sensors, data storage, software can be substantial. However, there are different policies in each company whenever an investment should take place (Compare et al., 2020; Achouch et al., 2022).
- **Cost of Maintenance:** The cost of maintenance of a PdM system itself is necessary, so companies can check what will be the cost in the long-term (Achouch et al., 2022).

Change Management

- **Cultural Resistance:** Employees are used to working with the traditional maintenance methods and they may resist the concept of shifting to PdM. The fear of losing their job or changing positions/roles within the company enhances this resistance (Ciocoiu et al., 2017; Bousdekis et al., 2020).
- **Training:** Significant training is required to ensure that employees can effectively use and maintain the new system (Giada and Rossella, 2021).

Scalability

- **Pilot to full deployment:** Scaling a PdM solution from a pilot project to full-scale deployment in an enterprise can be challenging, requiring consistent performance and adaptability (Agarwal et al., 2020).

Data Security and Privacy

- **Cybersecurity Risks:** The connected devices and systems increase the risk of cyberattacks (Giada and Rossella, 2021; Bandari, 2021).
- **Compliance:** Ensuring that data collection and usage comply with relevant regulations and industry standards (Giada and Rossella, 2021; Bandari, 2021).

As organizations begin to implement PdM programs, they quickly recognize the importance of addressing challenges, understanding the benefits and investing in robust asset management software. Handling PdM challenges early on helps to prevent disruptions and makes the implementations process smoother. This proactive approach also optimizes resource utilization and enhances processes. Moreover, it creates a culture of continuous improvement and innovation, enabling the organizations to refine their PdM strategies and adapt to evolving needs and technologies (Sensemore, 2024).

3
Methods

In this chapter, the methodology of this project is presented. The project is following a case study approach and the focus on this chapter is the different methods for gathering relevant information (data). After reading this chapter, readers will have a clear understanding of this framework, how it is used to navigate data projects, and also how it is used for this project. Additionally, they will see how each phase contributes to a successful data analysis process.

3.1 Case study approach

The case study approach is a research approach that involves an in-depth, detailed examination of a subject (case) within its real-life context. This approach is mostly used in social science, business and management to explore complex problems, understand phenomena and generate insights (Crowe et al., 2011; Priya, 2020). The project with this approach is structured to:

- Step 1: Understand the case and formulate research questions.
- Step 2: Identify the issues of the case.
- Step 3: Select proper data collection methods (Interviews, observations etc.)
- Step 4: Analyze the findings and present them based on the collected data.
- Step 5: Discuss the results, link them to the literature and answer the research questions.

These steps are analyzed in the sections below.

3.1.1 Step 1

The first step of the case study was to understand what is the case of the company (Parker), to create a planning report on how this project will be conducted, including the aim, the delimitations, the methods for data collection and the Gantt-Chart for scheduling the process.

3.1.2 Step 2

In the second step, the issues of the case were identified, by doing a pre-study of what issues are commonly exists within the manufacturing industry. This step was fulfilled by doing a literature study, which included searching different research or scientific paper with the keywords, through visits at Parker in Borås and through conversation with the Parker's supervisors.

3.1.3 Step 3

The third step was the selection of the data collection methods. This step was supported by the CRISP-DM framework, which is analyzed further below. The primary data collection method was the conduction of interviews with the employees at Parker. The selection of the people who will be interviewed was suggested by the supervisors from Parker. Other than the interviews, there were observations for understanding the current state of the case and identyifing any other data collection methods (i.e sensors).

3.1.4 Step 4

In this step, the findings (results) were presented based on the acquired data and further analyzed and described in detail for better understanding. The results were focused on defining the current state of the case and what challenges is the case company facing. Solutions for overcoming these challenges were proposed followed by assumptions in cost benefit and implementation cost for making Parker choose a solution based on the financial outcome since the delimitation of the investment is based on the ROI of two-years timeframe.

3.1.5 Step 5

For the final step, the results were discussed and linked to the literature study - theory. Furthermore, the research questions were answered on this step giving a brief explanation of the outcome of this project.

3.2 CRISP-DM

CRISP-DM stands as a widely recognized process model within industries for data mining (Schröer et al., 2021), comprised of six distinct phases (see figure 3.1):

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Evaluation
6. Deployment

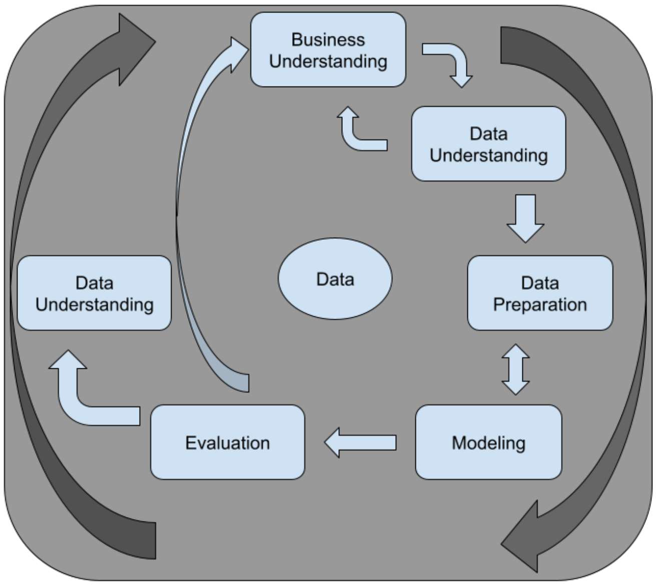

Figure 3.1: CRISP-DM Phases for Data Mining - Analysis (Inspired by Wirth and Hipp, 2000)

This process model offers an overview of the life cycle of data mining projects and it is considered as a standard for knowledge discovery and data mining projects. The benefits of using CRISP-DM are reduced cost, reduced time, minimized knowledge training, knowledge transfer, and documentation (Ayele, 2020).

3.2.1 Business Understanding

In this phase, the business objectives must be defined. This includes also understanding the goals, any challenges that Parker is facing in their current state, and the requirements of the stakeholders. These stakeholders are:

- Parker Line Manager
- Parker EHSE Manager

Through discussions with stakeholders, the scope, objectives, and requirements of the project were defined. The primary aim is to investigate strategies of how Parker can leverage data-driven maintenance decisions facilitated by ML or AI. This involves identifying which maintenance decisions should be data-driven, determining the requisite data for these decisions, investigating strategies for data collection and storage, and establishing procedures for processing data to extract relevant insights into the concept of PdM.

3.2.2 Data Understanding

Data understanding is a crucial phase for this project, since PdM is based on the data and the quality of them. This step starts with initial data collection and continues with activities to get familiar with the data, identify any data quality problems, discover insights into the data, and detect any subsets to form hypotheses for hidden information.

In this phase, Parker provides two software that they are using for collecting data:

- Software 1
- Software 2

For Software 1, the data that are included in this software are mostly generated reports that show downtime of the machine, some Key Performance Indicators (KPIs) such as Availability, Performance, and OEE in general, and also the cause of the stop of the machine in some specific reports.

For the Software 2, the data are more detailed. This software is used by operators and maintenance technicians to report any errors that occur in the machines. Sometimes these errors can be small errors that the root cause is known and it is just mentioned in Software 2 and other times there are bigger errors that the root cause is unknown, and it has to be reported in the system as detailed as possible.

3.2.3 Data preparation

Data preparation is the most time-consuming phase in CRISP-DM methodology, and this is because of the data cleaning, the outliers removal, the construction and integration of the data, and generally all the actions that must be done to get a structured dataset that can be directly placed as an input to any analyzing - predicting algorithms/models (Chapman, 2000). For this project, the phase of the data preparation differs from the regular one. After the data understanding, the data that were available from the company were mostly textual (free text) from Software 1. Thus, an assumption was made here, that maybe Parker can use algorithms such as NLP or TLP by using these textual data as an input to these algorithms. But before this can be implemented, cleaning the data should be the priority.

3.2.4 Modeling

In this phase, there is the selection of the modeling techniques (Chapman, 2000), such as NLP-TLP mentioned above. Since the data that Parker has available are textual data, the concept of implementing PdM by using NLP-TLP algorithms is not impossible. If Parker decides to follow another approach, such as installing sensors to monitor the conditions (pressure, temperature, vibrations, etc.) then, there could be another modeling technique, for instance, MATLAB-written or Python-written prediction algorithms, that can be used to predict the maintenance based on historical data. But in that case, Parker has to wait and monitor the conditions when a failure is happening, so they can create a structured dataset of historical data. This process may take a couple of years of monitoring or observing the failures and capturing the signals of the conditions to relate them to the failure.

3.2.5 Evaluation

This phase of CRISP-DM is the evaluation of the created model that could have high quality from a data analysis perspective because it is important to evaluate the model and ensure that everything works properly and fulfills the business objectives (Chapman, 2000). However, in the modeling phases it is stated that, currently, with the data that are available from Parker, there are two options that they can make use of (NLP-TLP). Since the scope of the project is to show the way to Parker what actions they have to take to shift to PdM, this phase is out of the scope.

3.2.6 Deployment

The last phase of the CRISP-DM methodology is the deployment of the model that was previously created. After the model is evaluated and fulfills the business objectives then it can be implemented and used (Chapman, 2000). This project suggests to Parker how to approach the concept of PdM and what is needed for them to implement such a maintenance strategy, therefore, this phase is out of the scope.

3.3 Interviews

Despite the CRISP-DM methodology, another approach to acquiring relevant information from Parker was the interviews, with the key participants. These participants were:

- Machine operators
- Maintenance technicians
- Line Manager / Maintenance Supervisor
- EHSE Manager
- Value-Stream Manager

These interviews were semi-structured, meaning that there were some predefined questions, that are available in Appendix A and B and there was an open discussion after with follow-up questions. The questions that were included in the interviews were mostly related to the RQs, some of them were formulated by reading different scientific - research papers that were relevant to the topic of the implementation of PdM solution, and some other questions were inspired by the conversations with the supervisors and the advisors of this project. The interviews lasted approximately 45 - 50 minutes and took place in the facility in Borås. The interviews went smoothly, and each participant was interviewed personally so that the process would remain anonymous. This means that the answers of the previous ones were not given to any of the participants so that their answer is as unique as possible and not influenced by any factor. The main focus of these interviews was to understand the perspective of all those different participants about the concept of PdM, to understand how Parker is working with their current maintenance strategy, what challenges are they facing and also some specific questions about the machines. For example, what was the most common error in the machines. After the conduction of the interviews, all the replies of the participants were gathered and analyzed to extract an outcome of this compilation of answers for every question. This outcome was further used for the result chapter and related to the literature study.

3.4 Literature Study

Literature study is also considered as a methodology in this project and is the main resource of information about PdM and the challenges that exist while companies are trying to implement such a concept. Websites such as Google Scholar, Chalmers Library, ScienceDirect, and ResearchGate are some of the websites that were used to find scientific papers that were relevant to this topic. To find the most respective papers some keywords were used, such as "predictive maintenance", "maintenance", "data science", "prediction challenges and benefits", and "Natural Language Processing in predictive maintenance". These keywords were used either in combination with other keywords or on their own.

As for the selection of these keywords, there were some factors in choosing them. For instance:

- The relevance of the words, meaning that the words that were selected were relevant to the core content.
- The specificity of the words, in other words, words that are either general and specific to keep a balance of being broad enough but specific as well
- Words or phrases that are synonyms with the main topic of the thesis were selected, for example, predictive maintenance - maintenance using prediction models.
- Acronyms, since AI and ML were the keywords that were most relevant to predictive maintenance.

The use of the keywords helped significantly to find the most relevant research - scientific papers. However, in the result of this search, there were some papers that were published a long time ago (i.e 90s or early 2000s) which were avoided since the technology and the knowledge around this topic had improved and developed rapidly in the recent years. Thus, papers that were published from 2010 and later were selected but there were some exceptions as well.

4
Results

This chapter presents the findings from the case study conducted at Parker. It begins by describing the current state of the company, particularly its maintenance practices and data infrastructure. Then it identifies and discusses the various challenges encountered in collecting data, developing algorithms, and managing cultural shifts within the organization. Solutions to these challenges are proposed and analyzed, including their potential costs and benefits. Additionally, the chapter provides strategic roadmaps for the implementation of PdM systems, outlining specific steps and approaches to enhance operational efficiency and reduce maintenance costs.

4.1 Company's Current State

In this section, the current state of the company is presented. There will be a detailed explanation of the current maintenance situation and the current data infrastructure that is available.

4.1.1 Maintenance Situation

The current maintenance strategy that Parker is following is mostly PM followed by RM. The company is doing maintenance regularly, meaning that they have a schedule for daily, weekly, monthly, quarterly, half-yearly, and yearly PM. This maintenance is performed either by the technicians or the operators depending on the importance of the maintenance. The maintenance team has a display that every machine error in the facility is reported. Every day, in the morning, the maintenance manager-supervisor has a meeting with the other managers to decide the priority of the errors, meaning which machine error has to be maintained first. Parker has also helped their operators and their technicians by writing a manual (standard) on what has to be maintained followed by pictures for better understanding.

Despite the PM, Parker also has an RM strategy meaning that, if an error has occurred in a machine, they are notified and send a maintenance technician to repair it immediately or they schedule the maintenance for another time depending on the importance of the error. Also, Parker purchases maintenance and service from the suppliers of the machines two times a year to come and check the condition and remaining life of the equipment. These insights into Parker's current maintenance strategies were gathered from interviews conducted with key personnel and through direct observations made during site visits.

4.1.2 Data Infrastructure

By talking to the maintenance managers and technicians and investigating the software systems in use, a thorough understanding of the current data infrastructure at Parker was built. Parker uses maintenance management systems Software 1 and Software 2 for reporting failures, planning, scheduling, tracking PM tasks, and efficiency of the machines. With this software, as mentioned in the Methods, they can calculate various indicators such as availability, performance, and generally, OEE, check the downtime of the machines, read the failures of a machine in detail, and see overall production statistics and downtime. As of now, it is the only source of data Parker has available about their machines.

From Software 2, the data that are extracted from the maintenance logs are only textual data with no visualization. These data refer to the failure that the machine had with more information and details if necessary and are written by either the operator or the maintenance technician. However, while doing the investigation of the spool cell logs, the size of the data for this machine was quite small and several errors were the same and reported in the system a bit differently. That means that there is no standard way of reporting each failure or reporting again a failure that is already reported in a specific format.

The data from Software 1, as it is mentioned before are connected to the KPIs of the performance, the availability, the downtime, and to OEE. These data, however, do not have any obvious relation with any of the failures that were reported to Software 2 and therefore it is challenging to understand if there is a correlation between the KPIs and the textual data.

4.2 Challenges

Despite the current state and the data infrastructure, Parker has some challenges to overcome to be ready to shift from PM-RM to PdM. These challenges that will be explained further in this section were identified through interviews conducted with key personnel and by relating them to the existing literature about common challenges for PdM.

- **Data source limitation**
  Available relevant data is crucial for the creation of a PdM model, however, many companies, including Parker, do not use or calculate the data that is needed for the implementation of a PdM model. This challenge is most likely related to the poor equipment for data collection, such as sensors, or to the quality of the existing data, which may not fulfill the needs of this new concept. At Parker, the spool cell that was investigated lacks sensors that could measure any conditions or any other relevant signals that could be used for prediction. Additionally, the existing data for the spool cell, which is the data from the maintenance logs is not big enough, and results in a bigger challenge to implement a predictive model with a small dataset. Moreover, the available data has some outliers, such as duplicates, that have to be cleaned to be a more structured dataset.

- **Reporting errors**
  Reporting an error is essential for the PdM concept. The traceability of an error that creates a trend can be useful input data for a prediction model. However, at Parker, sometimes operators may ignore reporting an error in Software 2 because the error is "*easy to fix*". By doing this regularly, Parker may have lost some useful data that can be used to determine the RUL of a part of a machine by checking the trend of failing in the maintenance logs.

- **Lack of knowledge**
  Training is one of the most important steps for companies to have high quality and efficiency in their production lines, and in this case, in the area of fault detection. Skilled and well-trained operators and technicians play a crucial role in ensuring that these factors are steady or, even better, higher. Parker, however, is currently offering the new operators and technicians the required training courses.

- **Lack of standardization**
  Lack of standardization is an important challenge to overcome for companies that seeking a shift to their maintenance strategy. Standard procedures make the implementation of PdM a lot easier and create a feeling of a structured system. Parker, however, does not have a standard way of reporting the errors in the system. Multiple incidents explain the same error but they have been reported in the system a bit differently. For example, "Robot laddar fel i fräs", and "Robot stannat i fräsen", are two errors that were reported in the system, describing the same issue but with different sentence construction. A solution to this challenge is either to have a standard way of reporting the errors or to clean the data regularly.

- **Resistance to change**
  Shifting to another maintenance strategy that is well-known to a less-known one, creates resistance from some employees. During the interviews, some participants answered that people could potentially be resistant to this new concept because they already know how the current system works and they have adapted this method to their working style, therefore they could be resistant to change despite the huge advantages that PdM may offer.

- **Need for a specialist**
  Another organizational challenge is the lack of specialists. At the current state, Parker does not have any data specialist who can focus on this implementation and do the necessary deep research for a precise implementation, for installing all the mandatory equipment, or for analyzing the acquired data. According to the interviews, the point of view of the company for this matter is that a specialist may not be needed and therefore there is a preference for not having a new data department.

4.3 Solutions

Parker has a lot of challenges to face in implementing PdM systems, however, there are some solutions to overcome these challenges and shape the way for Parker to shift at a steady pace to the predictive strategy. The successful implementation of PdM requires a systematic approach encompassing many different steps that need to be followed. By thoroughly analyzing the current situation at Parker, two solutions could be suggested for Parker. For each solution, a roadmap with general steps is provided. Even though the steps outlined are similar, the process and methods of each are different depending on the chosen approach.

The roadmaps were developed based on a thorough understanding of PdM that was derived from the literature review. In the literature review existing frameworks, methodologies, and case studies related to PdM were examined, which provided a solid theoretical foundation. Using this foundation, the roadmaps were constructed through logical thinking and a structured approach. Six steps were chosen to simplify the roadmaps and make it more structured and clear to follow, ensuring that each step logically builds upon the previous one. There will be a detailed planning description for each step of the roadmaps further in this report.

4.3.1 Solution 1: Approach through investing in sensors

One of the most important investments that Parker must make to at least have some relevant data that can be used for predictive systems that are available is to invest in sensors. By purchasing sensors, Parker can monitor several conditions such as pressure, vibrations, and temperature for their machines and check their signals when a failure occurs. On a daily basis, this process will gather enough historical data and probably the stored data will have a specific pattern or a trend that can be further predicted.

The first step in the roadmap is the "assessment phase" (See Figure 4.1). This involves identifying critical assets, evaluating the current maintenance strategy, and defining the objectives for the PdM initiative. The initial task is to identify which machines and equipment are critical to operations. This involves a thorough analysis of the production process to determine which assets have the highest impact on production efficiency and output. Equipment that frequently breaks down or has high repair costs should be prioritized. Next, it is needed to review the existing maintenance practices to identify their limitations. This includes analyzing historical maintenance records, failure logs, and maintenance costs to understand the shortcomings of the current RM or PM strategies. Thereafter clear objectives for the PdM project must be established.

Parker has already initiated this process by identifying the spool cell machines as critical assets. These machines are integral to operations due to their significant impact on production efficiency and output. Consequently, they have been prioritized for the implementation of PdM.

Regarding the current maintenance strategy, which is primarily PM, Parker has established a well-structured maintenance schedule. This was confirmed through interviews conducted with the maintenance and operational staff. However, there is room for improvement. For example, the current strategy sometimes overlooks minor signs of potential failures. To enhance this approach, it is recommended not to ignore any indication of potential failure or suspicion. Maintenance should be scheduled even for minor issues to prevent larger, more costly problems in the future. Finally, the objectives of applying PdM at are normally well-defined. The focus should be on reducing unplanned downtime, increasing equipment reliability, optimizing maintenance schedules, and ultimately lowering maintenance costs.

Figure 4.1: Roadmap for implementing CBPM using sensors.

The second phase of the roadmap is the planning phase. This involves selecting appropriate technologies, allocating budget, selecting vendors, and developing data collection protocols. Sensors must be chosen based on the parameters that are planned to be monitored, such as vibration, temperature, and other relevant indicators of equipment health. Further, A detailed budget must be prepared to cover the costs of sensor acquisition, installation, and the associated software for data collection and analysis. This budget should also include provisions for staff training and any potential disruptions during the installation phase. Thereafter, it is necessary to identify and select reputable vendors for sensors and monitoring systems. Vendors should be evaluated based on their product quality, support services, and integration capabilities with existing systems. Developing standardized protocols for data collection is also critical to ensure consistency and accuracy. These protocols should define how data will be collected, stored, and processed. It is also important to establish data security measures to protect sensitive operational data.

The third phase is the implementation phase where the focus lies on deploying the selected strategies and technologies, setting up the data infrastructure, and developing monitoring tools. Sensors should be installed. This should be done by a professional following the guidelines of the manufacturer to ensure optimal sensor performance. Further, a robust data infrastructure is needed for handling the continuous flow of data from the sensors. This includes setting up cloud storage solutions, and data processing capabilities, and ensuring integration with existing IT and operational systems lastly, a user-friendly dashboard should be developed to allow real-time monitoring of equipment performance.

Phase four is the analysis phase. Once the sensors start working, they collect data continuously. This data should be used to set up normal operating conditions. In the future, they can compare new data to this baseline to spot any unusual changes. ML algorithms like the ones mentioned in the theory part will be employed to analyze historical data and develop predictive models. These models forecast potential equipment failures based on real-time monitoring and historical performance trends. When deviations from the normal patterns are detected, alerts will be generated for immediate investigation and preventive action.

Phases five and six are about optimizing and evaluating the PdM model. As more data is collected, predictive models and algorithms should be continuously refined to improve accuracy. This iterative process involves adjusting models based on new data and feedback from maintenance teams. Also, automated alert systems should be implemented to notify maintenance teams of predicted issues. Training programs should be developed and delivered to ensure maintenance and operational staff are proficient in using the new PdM tools and interpreting data insights effectively.

The final phase involves evaluating the performance of the PdM system and making necessary adjustments. Regular performance evaluations should be done to assess the effectiveness of the system and analyze key metrics such as downtime reduction, and maintenance cost savings. Based on these evaluations, adjustments should be made to technology deployment, data analysis techniques, and maintenance schedules. Continuous improvement should be emphasized to adapt to evolving operational needs and technological advancements.
Additionally, likely for Parker and other companies that want to shift to PdM, some companies offer the installation of sensors to the machines and implementation of monitoring systems. Collaborating with these companies would enable Parker to skip some steps, streamline the transition process, and potentially reduce initial setup complexities. The only remaining step for Parker would be to recruit a data analyst to analyze the data, including all the pre-processing steps, from the monitors and implement a predictive algorithm.

4.3.2 Solution 2: Approach through NLP-TLP techniques

As mentioned before, NLP and TLP can use textual data that Parker has already available from their software. With this solution, however, Parker will need a programmer to write code based on the NLP-TLP and use the textual data as input to predict the need for maintenance. For this, it will be suggested that Parker must improve their current maintenance log by standardizing the way of reporting. Also "easy to fix" errors must be reported in the system to have a consistent flow of data for a more precise calculation of the maintenance decisions or RUL of the machines.

The first phase for the NLP-TLP approach, as it is shown in Figure 4.2, is to define the objectives of the PdM system, such as reducing downtime, optimizing maintenance schedules, or failure detection and assembling a team with domain experts, IT professionals, data scientists, and NLP/TLP specialists. With this team, Parker can identify and gather all the relevant data sources, including maintenance logs and reports that should be standardized and digitalized. It is essential to assess the quality of the data and address any gaps and errors to prepare for effective processing and analysis.

In the second phase, data cleaning is performed to remove duplicates, irrelevant information, and general noise from the maintenance logs and reports. The data is then annotated with relevant labels such as fault types, and maintenance actions which require manual input from domain experts as it is mentioned in the theory. Then, text normalization is crucial to handle spelling variations, industry's industry-specific abbreviations, and finally the data segmentation will help to divide these data into meaningful units for the next step of the analysis.

The third phase focuses on feature extraction, where the key features are extracted from the textual data, including keyword extraction, NER, BoW, etc. Furthermore, the TLP techniques are developed to handle this domain-specific terminology and context, which involves different and custom dictionaries and ontologies. Moreover, text embeddings are used to convert the text into numerical representations using the different techniques mentioned in the theory (i.e. TF-IDF). Finally, the appropriate NLP/TLP models are selected, for instance, Neural Networks, to handle the complexity of the PdM tasks.

The model training and validation is the fourth phase of this approach, where the data is split into training and validation (test) sets and the models are trained using these preprocessed data. The hyperparameter turning is also conducted to optimize the model parameters for improved performance and accuracy, experimenting with different settings to find the optimal configuration. The model's performance is then evaluated using different metrics such as precision, F1-score, and recall (Goutte and Gaussier, 2005) to validate its accuracy in predicting maintenance needs. An iterative improvement process is following, refining the model based on feedback and evaluation results, and incorporating additional features or data.

The fifth phase involves integrating the PdM model with the existing maintenance management systems through the development of interfaces. A user-friendly interface can be designed and implemented for the maintenance team and operators to interact with the system, including checks in dashboards and alerts. In this phase, the test of the implemented system is crucial to ensure that it performs as expected in the real - world conditions. Moreover, the training for every stakeholder is important to ensure that they will use the new system effectively.

Finally, in the last phase of this approach, it is emphasized the necessity of the detailed documentation of the newly installed system. All the reports, including system performance, maintenance predictions, and benefits need to be generated and stored, so the stakeholders can make value of them. Moreover, ensuring compliance with the industry's standards and regulations is critical alongside the security measures that have to be implemented to protect all these sensitive data from cyberattacks.

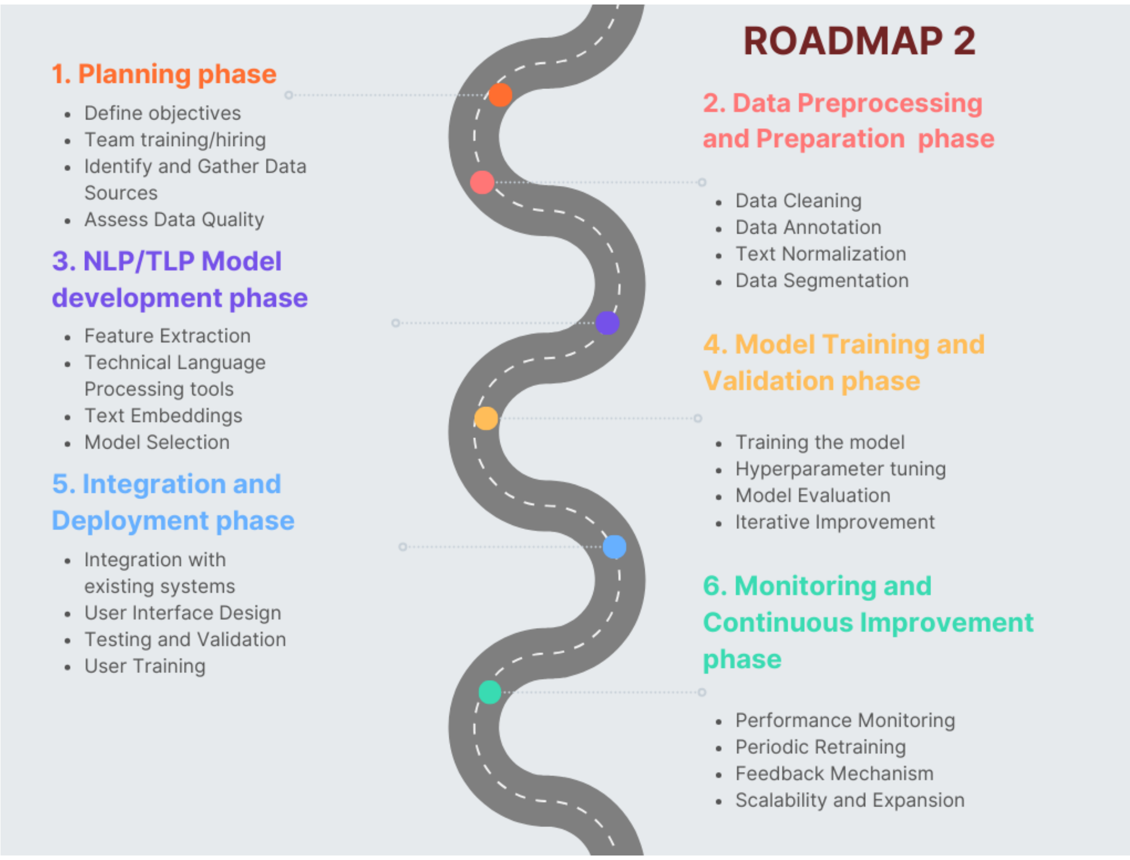

Figure 4.2: Roadmap for implementing PdM using NLP-TLP techniques.

4.4 Solution Cost Analysis

Cost analysis is crucial when implementing a system, especially for small or medium - sized enterprises. It helps businesses to evaluate the financial feasibility and potential return on investment of the new system. For Parker, this is an important factor since the policy of the company for implementing and investing in something new is that the ROI will be within two years.

4.4.1 Solution 1

Sensors are a significant part of this solution since they are the main source of data. According to Wu (2021), basic sensors such as temperature sensors can cost around 100 $ (∼1070 SEK) each, while more advanced sensors, like vibration sensors, can cost from a range of 1.000 $ (∼10.700 SEK) to 2.000 $ (∼21.400 SEK) each. The number of sensors that will be needed will depend on the machinery and the parameters (conditions) that need to be monitored.

Moreover, the installation of the sensors requires professional services to ensure that the placement will be correct and that the sensors are calibrated. This process can add several thousand dollars to the setup cost, depending on the complexity and the scale of the system (Wu, 2021).

Additionally, the software that will be used for handling and analyzing the data should also be considered as an extra expense. This software has to be integrated with the current IT state which might be even more costly. The subscription fees for PdM software range from 100 $ (∼1070 SEK) to 1.000 $ (∼10.700 SEK) per month. Also, the cloud storage for the data collected by the sensors might add to these monthly costs (Wu, 2021).

The training of the staff, to use the new system is another important cost, which can be several thousand dollars. However, if the company wants to hire skilled data analysts or skilled maintenance technicians who can use all these data effectively, the cost per year can be around 86.000 $* (∼920.000 SEK) (Wu, 2021).

**The salary is based on US salary data and in Sweden may not be the same*

4.4.2 Solution 2

Building an NLP system from open-source components can have an initial software cost of $ 0. However, some hidden costs are significant. Expertise is required, with NLP experts' salaries averaging $ 81.000 (∼870.000 SEK) per year. In addition to that cost, extensive time and resources are needed for data collection, cleaning, and model training, which gives a total cost of $ 100.000 (∼1.070.000 SEK) and more. The development time can vary from weeks to months and ongoing maintenance and updates further add to this expense (Richards, 2022).

Licensing an NLP platform requires an upfront investment. Basic cloud NLP services start at around $ 10.000 (∼107.000 SEK) annually, while more advanced and customizable platforms can cost around $ 40.000 (∼430.000 SEK) per year. The deployment can be rapid, usually within a few days, providing immediate access to advanced NLP capabilities (Richards, 2022). Additionally, training staff to utilize the system effectively, integrate the solution with the existing systems, and ensure data security and compliance are also important considerations that could add an extra cost, as it is mentioned in Solution 1.

4.5 Cost Benefit of PdM

According to the literature, the implementation of PdM systems can save up to 60% of maintenance costs (Serradilla et al., 2022). At Parker, the cost of the maintenance for the spool cell, which is the machine that is investigated, including the cost of the spare parts, the cost of the PM, and the technical personal time is assumed in a theoretical value of 1.000.000 SEK. Based on what is written, Parker, theoretically can save 600.000 (60% of 1.000.000) SEK, if the savings reach 60% of the maintenance cost. However, to be more realistic, in a case study of implementing advanced monitoring and PdM strategies for offshore wind farms (Turnbull and Carroll, 2021), the savings were a bit different. In that case, they managed to achieve 8% savings of the operation and maintenance costs and 11% of the lost production.

In another case study, with the title "Predictive Maintenance using Machine Learning: A Case Study in Manufacturing Management", it is stated that "applying predictive maintenance using ML can reduce downtime and maintenance costs by up to 25% providing a strong financial return of investment"(Satwaliya et al., 2023). With these realistic numbers, Parker can save approximately 80.000 (8% of 1.000.000) - 250.000 (25% of 1.000.000) SEK annually. Additionally, according to software 1 and based on the interview with the line manager, there is some lost production for the spool cell which is assumed to 3.000 hours. Each of these hours costs a theoretical value of 1.000 SEK. This means that Parker loses from lost production in a theoretical perspective 3.000.000 SEK annually. However, some of these hours are also related to downtime of machines due to sickness, lack of orders etc.) meaning that a PdM solution could not solve these reasons. Despite that, if the numbers from the wind farms can be applied to the Parker case, the savings from the lost production by implementing PdM can be 330.000 (11% of 3.000.000) - 750.000 (25% of 3.000.000) SEK. In total, Parker can save up to 450.000 - 1.000.000 SEK approximately per year by implementing PdM. However, this number is only associated with the data set and the findings of another case study and is only an theoretical assumption. In reality, these numbers can vary, either lower or much higher than stated. Also, the costs of savings are highly correlated with the level of accuracy of the implemented algorithm.

5
Discussion

This chapter provides a comprehensive analysis and interpretation of the results derived from Parker and relevant literature. It includes a comparison of findings with existing studies, evaluates the practical implications of the results, re-evaluates the ROI, discusses the existence of companies that provide PdM as a service, and answers directly the research questions of this project. Additionally, it reflects on the research methods used and examines the ethical and sustainability aspects of implementing predictive maintenance systems. Lastly, it offers suggestions for future research directions.

5.1 Comparison with literature

The findings of this study align with the existing literature about the application of PdM and ML algorithms for maintenance processes especially in the manufacturing sector, but also some areas are worth pointing out. The results aligned with the literature study (Kaur et al., 2018; Li et al., 2024), outpointing the potential benefits of reduced downtime and increased equipment lifespan which optimally leads to enhanced operational efficiency. Similar to what was reported by (Poor et al., 2019; Satwaliya et al., 2023), this study demonstrates that Parker can have significant cost savings, based on the assumptions that were made during the cost calculation, by transitioning from RM and PM strategy to a PdM strategy. However, in some case studies as mentioned in section 4.5 "Cost benefits of PdM", the cost varies and it can be completely different if a PdM solution is implemented at Parker. This is because every organization is different and the machinery that exists behaves uniquely.

While many studies mention the applicability of certain machine learning algorithms across different industries such as Paolanti et al. (2018), it should be noted that the effectiveness of these algorithms can vary significantly depending on the specific operational environment and the nature of the equipment being maintained. Additionally, it is worth mentioning that the challenges of maintaining data quality and consistency, which are essential for PdM model success, are presented in the results and are connected to the literature (Achouch et al., 2022). Furthermore, the results emphasize the value of effective data management procedures, an area that has not gotten much attention in the literature. For example, training the staff about the concept of predictive maintenance, how can be achieved, what actions they have to take, what changes will be applied to the current working style, etc.

5.2 From theory to practice

The solutions presented in this thesis demonstrate the significant potential of PdM implementation at Parker. One notable advantage of PdM over traditional PM is its ability to respond to the actual condition of the equipment rather than following a predetermined schedule (Kaur et al., 2018). The implementation of CBPM integrates real-time monitoring with predictive analytics, enabling a more precise prediction of equipment failures based on historical and real-time data. However, based on the IT systems that Parker currently has, the proposed solutions could not be implemented directly. While the initial results are promising, some areas necessitate further investigation to fully transition to PdM.

Firstly, a more comprehensive understanding of the specific data requirements for various types of equipment used at Parker is essential. Identifying the critical parameters to monitor and ensuring the accuracy and reliability of the data collected will be crucial for the success of the PdM model (Achouch et al., 2022). For example, there must be a further investigation of the sensor solutions and their placement in the chosen machines for acquiring relevant data, or for the NLP/TLP solution, they must ensure that their IT systems can support this approach for a smooth integration.

Moreover, shifting from one maintenance strategy to another is not an easy task. To successfully integrate PdM into their operations, organizations need to undergo a significant shift in mindset as well (Ciocoiu et al., 2017; Bousdekis et al., 2020). This involves moving away from traditional, RM-PM practices to a proactive approach that leverages advanced data analytics and IoT technologies. This shift not only requires investment in technology but also a cultural change where data-driven decision-making becomes the norm. Employees at all levels must be trained to understand and trust predictive analytics, fostering a collaborative environment where insights derived from data are used to enhance operational efficiency and reliability (Giada and Rossella, 2021). However, if Parker decides to implement any of the proposed solutions, it is highly recommended to follow the suggested steps of the roadmaps to have a straightforward implementation of a PdM system.

Additionally, the development and refinement of ML algorithms customized to Parker's specific operational environment require ongoing research (Agarwal et al., 2020). This involves not only selecting the appropriate algorithms but also continuously optimizing them based on new data and feedback from maintenance operations (Brundage et al., 2020). The iterative improvement of these models will enhance their predictive accuracy and reliability.

Lastly, the organizational structure may need to be adjusted to support the integration of predictive maintenance (Bousdekis et al., 2020). This could include the creation of dedicated teams responsible for data analysis, algorithm development, and maintenance optimization. Collaboration between these teams and existing maintenance and IT departments will be critical for the successful deployment and ongoing management of the predictive maintenance system.

5.3 Re-Evaluation of ROI

One of the primary limitations faced by Parker in ROI for implementing PdM systems is the two year period that they proposed. Currently, the data at Parker available is insufficient for making accurate predictions and informed decisions regarding the cost-effectiveness of PdM solutions. The company will need to invest significant time, potentially a year or years, in gathering and analyzing relevant data before they can expect to see substantial benefits.

The initial costs associated with implementing PdM are significant, which include the purchase of sensors, installation, software integration, and staff training. Sensors, essential for collecting necessary data, vary significantly in price based on their complexity and properties as mentioned in Section 4.4. Installation requires professional services to ensure proper placement and calibration, adding further to the setup costs. Software expenses, including subscription fees and potential cloud storage costs, also contribute to the financial strain. Moreover, the training of staff to effectively use the new system and the need to hire skilled data analysts or maintenance technicians add to the ongoing costs as mentioned in Section 4.4. While these investments are crucial for the successful deployment of PdM, they impose a significant financial strain on the company in the short term.

Although the calculations of costs of implementing a PdM solution are rough assumptions given this and the constraints mentioned, the projected ROI within two years, as per the company's policy, is still not feasible. The lack of adequate data currently available makes it challenging to achieve the desired ROI within this timeframe. A more extended period will be required to collect sufficient data, validate the effectiveness of the PdM system, and realize any cost savings. As such, Parker must re-evaluate their expectations and prepare for a longer-term investment strategy to fully benefit from PdM implementation.

5.4 Existing Predictive Maintenance Services

During the research on existing PdM systems, various companies offering CBM and PdM as a service were identified. Additionally, While investigating various sensors and their applications, it was revealed that Parker itself manufactures a wide range of sensors. These sensors are designed for various industrial applications and are integral to monitoring the condition and performance of machinery. It was also discovered that Parker already offers a PdM service, as highlighted on their condition monitoring webpage. This service includes solutions for condition monitoring, which are integral to predictive maintenance. Given this, it raises the question of why the company is seeking external assistance to develop a predictive maintenance solution.

One possible reason could be that the existing services are not fully tailored to the specific needs and challenges of the company's operations (Achouch et al., 2022). Another reason might be the lack of internal expertise required to fully utilize the existing predictive maintenance tools and services (Bousdekis et al., 2020). It could also be due to the desire for a more customized solution that integrates better with the company's unique production processes and machinery (Achouch et al., 2022)

Nevertheless, the company needs to assess the capabilities of its current predictive maintenance services and determine how they can be leveraged or enhanced before seeking external help. This approach could save costs and streamline the implementation process, ensuring that the company makes the most of its existing resources and expertise.

5.5 Generalizability of Results

While the study was conducted at Parker and with their organization as the base, the results are broad and can be relevant to many industries and companies, especially those in the manufacturing sector. The challenges discussed in the theoretical framework part of the report are common across various industries. Any company seeking to implement PdM must first overcome these challenges or similar challenges within their organization. Furthermore, the roadmaps developed in this study offer a structured and generalized solution that can be adopted by other companies as well. The steps in the roadmaps provide a comprehensive step for step guide for companies looking to transition to PdM.

5.6 Answers to the Research Questions

The research questions have been answered throughout the reports. To make it more straightforward, the question will be directly answered in this section.

1. How can Parker develop a predictive maintenance solution for their types of machines?

   Answer: A PdM solution for Parker can be developed by first conducting a detailed assessment of the specific machines in use to identify the critical parameters that need monitoring, such as vibration, temperature, and operational sounds. Sensors should be installed to measure these parameters accurately. A robust data infrastructure is necessary to handle the continuous flow of data from these sensors, which includes cloud storage solutions and data processing capabilities. Developing a user-friendly dashboard for real-time monitoring is also crucial to enable quick and effective decision-making based on the collected data.

2. Which maintenance decisions should be data-driven and what data is needed for these decisions, and how should this data be collected and stored?

   Answer: Maintenance decisions that should be data-driven include scheduling maintenance activities, identifying potential equipment failures, and optimizing maintenance schedules. The data needed for these decisions include real-time sensor data (vibration, temperature, sound), historical performance data, maintenance logs, and failure records. Machine learning algorithms will analyze this data to predict failures and recommend maintenance actions based on the actual condition of the equipment rather than predetermined schedules.

   Data should be collected continuously from sensors installed on the machines. These sensors should be professionally installed to ensure optimal performance. The data collected should be transmitted to a centralized cloud storage solution where it can be processed and analyzed. It is essential to have data security measures in place to protect sensitive operational data. Additionally, standardized protocols for data collection, storage, and processing should be established to ensure consistency and accuracy.

3. What are the potential cost savings and benefits Parker can achieve by transitioning from preventive to predictive maintenance strategies, and what challenges must be overcome during this transition?

   Answer: By transitioning from preventive to predictive maintenance, Parker can achieve significant cost savings and benefits, including reduced unplanned downtime by addressing issues before they escalate, extended equipment lifespan by performing maintenance based on the actual condition of the machines, optimized resource allocation by scheduling maintenance activities more efficiently, and enhanced operational efficiency with overall reduced maintenance costs. Predictive maintenance allows for more precise interventions, reducing the likelihood of unnecessary maintenance activities that can occur with preventive maintenance strategies.

Parker faces several challenges in implementing predictive maintenance solutions, including a potential gap in internal expertise required to fully utilize PdM tools and services, the need for a cultural shift towards data-driven decision-making within the organization, ensuring accurate and continuous data collection from sensors, integrating new PdM technologies with existing IT and operational systems, and training employees while fostering collaboration between maintenance and IT departments to effectively use the new PdM tools and interpret data insights.

5.7 Reflections on research methods

The project employed a case study approach to investigate the transition from PM-RM to PdM at Parker. To ensure a thorough understanding of the case, the initial step involved formulating research questions through discussions with key personnel and the goal of the project. This set the foundation for a comprehensive study aimed at addressing the critical aspects of shifting to PdM.

Data mining was the most important aspect of the success of this project. The framework that was followed for the data mining was the CRISP-DM. In the planning report of this project, the CRISP-DM methodology was the primary approach for the acquisition of relevant data. The thought idea was to acquire the data and preprocess them according to the steps of the CRISP-DM and further analyze them for building an ML algorithm that could possibly predict the failure of the spool cell and finally implement it in the existing systems at Parker. However, this framework did not work as expected, while many of the phases of this framework were skipped due to the insufficient data that were available at Parker, and were alternated for the completion of this project. Then, the idea changed to the creation of a structured path (roadmaps) for the implementation of a PdM system followed by the identification of the challenges that may occur from this shift and how Parker can overcome them. So, some of the steps of CRISP-DM were followed including business understanding and data understanding.

For that, two approaches were made. Firstly, semi-structured interviews with maintenance managers, technicians, and operators along with on-site observations which provided invaluable qualitative data. The interviews were designed to understand the perspective of different participants about the concept of PdM, how Parker is currently managing their maintenance strategies, the challenges they are facing, and specific details about machine errors, etc. To maintain the anonymity and integrity of the responses, each interview was conducted individually, and the answers from previous participants were not shared with others.

Secondly, a literature study was conducted. The literature study was another critical part, providing context and supplementing the findings from the interviews. Reliable sources such as Google Scholar, Chalmers Library, ScienceDirect, and ResearchGate were used to find relevant scientific papers that discussed PdM and its implementation challenges. Keywords like "predictive maintenance," "maintenance," "data science," "prediction challenges and benefits," and "Natural Language Processing in predictive maintenance" helped in identifying relevant literature. However, one limitation was that while these sources provided a wealth of academic insight, they did not always directly correlate to the specific industrial context of Parker.

Reflecting on the research methods, particularly semi-structured interviews, it is worth considering certain limitations that might have influenced the results. The sample size for the interviews might not have been large enough to capture the full diversity of perspectives and experiences within the organization. Additionally, there is a possibility that some interviews might have been subject to bias, either due to the framing of questions or the participants' openness to answer the particular question. These factors could potentially affect the validity and reliability of the results.

5.8 Cost-benefit and cost calculations assumptions

To achieve the reported savings in maintenance costs and reduction in lost production, several prerequisites must be met. First, advanced monitoring and diagnostic systems are essential. These systems need to be capable of diagnosing a significant portion of potential failures early enough to allow for timely repairs, thereby avoiding the high costs associated with major failures.

Second, the smooth integration of these systems is crucial. The data acquired must be optimal for the specific cases studied, ensuring that the systems function effectively and provide accurate diagnostics and prognostics.

Additionally, making informed maintenance decisions based on the diagnostics and prognostics provided by these advanced systems is vital. This involves balancing the risk of component failure with the cost of early intervention. Implementing a condition-based maintenance strategy is also important. This strategy involves utilizing more of the component's life before replacement, although the decision to intervene earlier should be prioritized to avoid higher cost implications.

These prerequisites collectively ensure that the predictive maintenance strategy is effectively implemented, leading to significant cost savings and reduced production losses. While the numbers presented—8% savings in maintenance costs and 11% reduction in lost production—are promising, it is important to note that they are based on specific assumptions and ideal conditions. The reliability of these figures depends on the smooth integration and optimal performance of the systems, as well as accurate and timely data acquisition.

For companies to create compelling business and investment cases for predictive maintenance solutions, a thorough and detailed investigation into these costs and benefits is necessary. This should include evaluating integration challenges, ensuring data quality and consulting multiple sources of information to verify cost estimates and assumptions. By doing this, companies can make more informed and strategic decisions, ultimately enhancing the reliability and effectiveness of their predictive maintenance strategies.

5.9 Ethical and Sustainability Aspects

The implementation of PdM systems at Parker Hannifin Manufacturing carries some ethical and sustainability aspects that need careful consideration. Addressing these aspects can ensure a responsible and beneficial transition from traditional maintenance strategies. From an ethical standpoint. From an ethical standpoint, several considerations could arise. First and foremost is the issue of data privacy and security. Applying PdM requires collecting various types of data on equipment and organizational operations and is the core for applying PdM (Li et al., 2024). The collection and analysis of machine data must adhere to data privacy regulations to protect sensitive operational information. Ensuring robust data security measures is important to prevent misuse or unauthorized access, which could undermine both company operations and employee privacy. Transparency in how data is collected, used, and who has access to it is also necessary. Clear communication about data governance can foster trust among stakeholders and mitigate concerns regarding surveillance and potential misuse of information (Bousdekis et al., 2020).

Resistance to change among employees is an ethical aspect to consider. Shifting from a well-known maintenance strategy to a less familiar PdM approach can create resistance among staff (Kiritsis et al., 2021). During interviews, some participants expressed that employees might resist this new concept because they are accustomed to the current system and have adapted their working styles accordingly, despite the advantages that PdM may offer. Addressing this resistance requires engaging with employees, providing clear communication, and offering comprehensive training programs to make the transition smoother and more acceptable.

Moreover, there is a potential impact on employees to consider. The shift to PdM technologies might alter job roles and necessitate new skill sets. As mentioned earlier lack of knowledge is one of the challenges when applying PdM and having the skills needed for this is very important (Meyer Zu Wickern, 2019). Thus, transitioning to PdM could lead to job displacement or the need for reskilling within the workforce. Ethical implementation should include proactive measures such as comprehensive training programs and support systems for affected employees, ensuring that the transition is fair and inclusive. Engaging stakeholders throughout the process can address their concerns and enhance the acceptance of the new system.

On the sustainability aspects, integrating PdM solutions offers benefits that align with environmental and operational efficiency goals. One major advantage is resource optimization. By accurately predicting equipment failures and maintenance needs, PdM reduces unnecessary maintenance activities, conserves materials, and minimizes waste, leading to more sustainable operations (Li et al., 2024).

Furthermore, PdM contributes to extending the useful life of equipment. By preventing breakdowns and reducing wear and tear, predictive maintenance ensures that machinery operates longer before needing replacement. This prolongation of equipment lifespan conserves resources and reduces the environmental impact associated with manufacturing new machines and disposing of old ones. The reduction of unexpected downtime which is one of the benefits of PdM also leads to more efficient production schedules and less waste generated from non-operational equipment (Li et al., 2024). This efficiency can result in better inventory management, as parts and resources are used based on actual needs rather than arbitrary schedules. Sustainable procurement of sensors and resources for transition to PdM should also be considered. Collaborating with sensor suppliers and software vendors that have strong sustainability credentials enhances overall sustainability efforts. Selecting partners who use eco-friendly practices and products contributes to a greener supply chain and promotes corporate responsibility.

5.10 Future suggestions

The results of this study are mainly based on theoretical frameworks and case studies from existing literature. Thus, it would be highly beneficial to implement and pilot-test the proposed PdM strategies in real-world settings. Following the suggested roadmaps of this study, future research should aim to implement and pilot-test PdM systems to validate the theoretical findings in practical scenarios. Unfortunately, due to limited time, it was not possible to conduct practical implementations within this study. Implementing a pilot test of the PdM strategies in an industrial environment would provide valuable insights into their effectiveness and practicality. This would involve real-time data collection, algorithm deployment on actual machinery, and continuous monitoring to assess the system's performance. Such practical applications can reveal potential challenges and refinements needed for successful integration into existing maintenance workflows.

Moreover, conducting studies to observe the long-term impacts of PdM implementation would be crucial. This would help in understanding how the strategies perform over extended periods and under varying operational conditions. Real-world testing can also provide economic and operational data, offering a clearer picture of the cost savings, efficiency gains, and overall return on investment. Future research should also explore collaborations with industrial partners to facilitate these practical implementations. Partnering with companies can provide access to necessary resources and operational environments, making it feasible to conduct extensive pilot tests. Additionally, feedback from industry practitioners can be invaluable in refining the PdM strategies to better meet practical needs and constraints.

6
Conclusion

The investigation into the deployment of PdM systems at Parker highlights significant potential benefits alongside notable challenges. The transition from traditional PM to PdM is driven by the need to enhance operational efficiency, reduce downtime, and achieve cost savings through data-driven decision-making. This study identified critical parameters for equipment monitoring, necessary data infrastructure, and outlined the potential cost benefits of PdM implementation.

The findings underscore that while the initial costs of implementing PdM—such as purchasing sensors, installation, software integration, and staff training—are substantial, the long-term benefits can be significant. These benefits include reduced unplanned downtime, optimized maintenance schedules and extended equipment lifespans. For instance, the data collected can be used to create detailed maintenance schedules customized to the actual conditions of the machines, rather than following a PM schedule. This targeted approach can significantly reduce maintenance costs over time.

However, achieving these benefits requires overcoming various challenges. One major challenge is ensuring continuous and accurate data collection. The study found that inconsistencies in data recording formats could lead to significant issues in data analysis. Standardizing data entry and employing automated cleaning processes are necessary to maintain data quality. Additionally, integrating new PdM technologies with existing systems is another critical barrier. The correlation between KPIs tracked in Software 1 and failure logs in Software 2, for example, highlighted the need for a more integrated data management system.

Fostering a cultural shift towards data-driven decision-making within the organization is also essential. Resistance to change among employees accustomed to traditional maintenance practices was identified as a barrier. Implementing comprehensive change management strategies, including regular training sessions and workshops, can help in addressing this resistance. Engaging employees by demonstrating the practical benefits of PdM, such as showing real-life case studies where PdM has led to increased efficiency and reduced downtime, can further encourage acceptance.

The proposed strategic roadmaps for transitioning from PM to PdM in this study are well-defined, but they are theoretical and based on literature. Therefore, the importance of pilot-testing PdM strategies in real-world settings is highlighted. By implementing these strategies in an industrial environment, Parker can gain valuable insights into their practicality and effectiveness, allowing for the identification and resolution of potential challenges. For example, pilot-testing the integration of NLP-TLP techniques in analyzing maintenance logs could reveal practical issues such as data compatibility and user-friendliness, which can then be addressed to refine the approach.

Additionally, deeper research into the development of machine learning algorithms customized to Parker's operational environment will be needed for the successful implementation of a PdM solution. These algorithms can be trained on historical maintenance data to predict future failures accurately, thereby maximizing the reliability and efficiency of the maintenance processes. For instance, algorithms could be developed to predict the RUL of critical components, enabling more precise maintenance interventions.

In essence, while this study sets a foundational understanding and a strategic framework for deploying PdM systems, real-world application and continuous improvement through iterative testing and adaptation are crucial. For Parker, the next steps involve pilot projects to validate the proposed strategies, investing in data analytics capabilities, and fostering a culture that embraces data-driven decision-making. Through these efforts, Parker can realize the full potential of PdM and significantly enhance its operational efficiency and cost-effectiveness in the long run.

Bibliography

Acernese, A., Del Vecchio, C., Tipaldi, M., Battilani, N., & Glielmo, L. (2020). Condition-based maintenance: an industrial application on rotary machines. *Journal of quality in maintenance engineering*, *27*(4), 565–585. https://doi.org/10.1108/jqme-10-2019-0101

Achouch, M., Dimitrova, M., Ziane, K., Karganroudi, S. S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On Predictive Maintenance in Industry 4.0: Overview, models, and challenges. *Applied sciences*, *12*(16), 8081. https://doi.org/10.3390/app12168081

Agarwal, V., Manjunatha, K. A., Gribok, A. V., Mortenson, T., Ulrich, T., Boring, R., & Harry, P. (2020). Scalability of a risk-informed predictive maintenance strategy. *Idaho National Laboratory, INL/LTD-20-58848*.

Amruthnath, N., & Gupta, T. (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. *2018 5th International Conference on Industrial Engineering and Applications (ICIEA)*, 355–361. https://doi.org/10.1109/IEA.2018.8387124

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, *54*(15), 2787–2805. https://doi.org/https://doi.org/10.1016/j.comnet.2010.05.010

Ayele, W. (2020). Adapting crisp-dm for idea mining: A data mining process for generating ideas using a textual dataset. https://www.researchgate.net/publication/342572029_Adapting_CRISP-DM_for_Idea_Mining_A_Data_Mining_Process_for_Generating_Ideas_Using_a_Textual_Dataset

Bandari, V. (2021, March). A comprehensive review of AI applications in automated container orchestration, predictive maintenance, security and compliance, resource optimization, and continuous deployment and testing. https://research.tensorgate.org/index.php/IJIAC/article/view/1

Bogue, R. (2013). Sensors for condition monitoring: A review of technologies and applications. *Sensor Review*, *33*(4), 295–299.

Bonnevay, S., Cugliari, J., & Granger, V. (2019, May). *Predictive Maintenance from Event Logs Using Wavelet-Based Features: An Industrial Application*. https://doi.org/10.1007/978-3-030-20055-8\{_}13

Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Predictive maintenance in the 4th industrial revolution: Benefits, business opportunities, and managerial implications. *IEEE Engineering Management Review*, *48*(1), 57–62. https://doi.org/10.1109/EMR.2019.2958037

Brundage, M. P., Sexton, T., Hodkiewicz, M., Dima, A., & Lukens, S. (2020). Technical language processing: Unlocking maintenance knowledge. https://doi.org/10.1016/j.mfglet.2020.11.001

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., da P. Francisco, R., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers Industrial Engineering*, *137*, 106024. https://doi.org/https://doi.org/10.1016/j.cie.2019.106024

CEN - CENELEC - Search standards. (n.d.). https://standards.cencenelec.eu/dyn/www/f?p=205:105:0:::::

Chapman, P. (2000). Crisp-dm 1.0: Step-by-step data mining guide. https://api.semanticscholar.org/CorpusID:59777418

Ciocoiu, L., Siemieniuch, C. E., & Hubbard, E.-M. (2017). From preventative to predictive maintenance: The organisational challenge. *Proceedings of the Institution of Mechanical Engineers. Part F, Journal of rail and rapid transit*, *231*(10), 1174–1185. https://doi.org/10.1177/0954409717701785

Coanda, P., Avram, M., & Constantin, V. (2020). A state of the art of predictive maintenance techniques. *IOP Conference Series: Materials Science and Engineering*, *997*, 012039. http://dx.doi.org/10.1088/1757-899X/997/1/012039

Compare, M., Baraldi, P., & Zio, E. (2020). Challenges to iot-enabled predictive maintenance for industry 4.0. *IEEE Internet of Things Journal*, *7*(5), 4585–4597. https://doi.org/10.1109/JIOT.2019.2957029

Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical research methodology*, *11*(1). https://doi.org/10.1186/1471-2288-11-100

Editor. (2023, November). Named Entity Recognition: the mechanism, methods, use cases, and implementation tips. https://www.altexsoft.com/blog/named-entity-recognition/

GeeksforGeeks. (2023, January). Understanding TF-IDF (Term Frequency-Inverse Document Frequency). https://www.geeksforgeeks.org/understanding-tf-idf-term-frequency-inverse-document-frequency/

Ghani, R. (2021). Integration of FMECA and statistical snalysis for predictive maintenance. *Journal of applied research in technology engineering*, *2*(1), 33. https://doi.org/10.4995/jarte.2021.14737

Giada, C. V., & Rossella, P. (2021). Barriers to Predictive Maintenance implementation in the Italian machinery industry. *IFAC-PapersOnLine*, *54*(1), 1266–1271. https://doi.org/10.1016/j.ifacol.2021.08.152

Gillespie, A. (2015). Condition based maintenance: Theory, methodology, & application. *Conference: Reliability and Maintainability SymposiumAt: Tarpon Springs, FL.*

Goutte, C., & Gaussier, E. (2005, January). *A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation*. https://doi.org/10.1007/978-3-540-31865-1\{_}25

Kaur, K., Selway, M., Grossmann, G., Stumptner, M., & Johnston, A. (2018). Towards an open-standards based framework for achieving condition-based predictive maintenance. *Proceedings of the 8th International Conference on the Internet of Things*. https://doi.org/10.1145/3277593.3277608

Kim, S.-W., & Gil, J.-M. (2019). Research paper classification systems based on TF-IDF and LDA schemes. *Human-centric computing and information sciences*, *9*(1). https://doi.org/10.1186/s13673-019-0192-7

Kiritsis, D., Lazaro, O., Hodkiewicz, M., Lee, J., & Ni, J. (2021). Editorial: Data-driven cognitive manufacturing—applications in predictive maintenance and zero defect manufacturing. doi:%2010.3389/fcomp.2020.633850

Lee, Holgado, Kao, & Macchi. (2014). New thinking paradigm for maintenance innovation design. https://doi.org/10.3182/20140824-6-za-1003.02519

Li, Z., He, Q., & Li, J. (2024). A survey of deep learning-driven architecture for predictive maintenance. *Engineering applications of artificial intelligence*, *133*, 108285. https://doi.org/10.1016/j.engappai.2024.108285

Meddaoui, Hain, & Hachmoud. (2023). The benefits of predictive maintenance in manufacturing excellence: A case study to establish reliable methods for predicting failures. https://doi.org/10.21203/rs.3.rs-2908342/v1

Meyer Zu Wickern, V. (2019). Challenges and reliability of predictive maintenance. https://doi.org/10.13140/RG.2.2.35379.89129

Mishra, P., Pandey, C. M., Singh, U., Keshri, A., & Sabaretnam, M. (2019). Selection of appropriate statistical methods for data analysis. *Annals of cardiac anaesthesia/Annals of Cardiac Anaesthesia*, *22*(3), 297. https://doi.org/10.4103/aca.aca\{_} 248\{_}18

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, *4*(51-62), 56.

Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018). Machine learning approach for predictive maintenance in industry 4.0. *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, 1–6. https://doi.org/10.1109/MESA.2018.8449150

Poor, P., Ženíšek, D., & Basl, J. (2019). Historical overview of maintenance management strategies: Development from breakdown maintenance to predictive maintenance in accordance with four industrial revolutions. http://hdl.handle.net/11025/36565

Prajapati, A., Bechtel, J., & Ganesan, S. (2012). Condition based maintenance: a survey. *Journal of quality in maintenance engineering*, *18*(4), 384–400. https://doi.org/10.1108/13552511211281552

Priya, A. (2020). Case study Methodology of Qualitative Research: Key attributes and navigating the conundrums in its application. *Sociological bulletin/Sociological Bulletin*, *70*(1), 94–110. https://doi.org/10.1177/0038022920970318

Richards, S. (2022, July). Build vs Buy: Text Analytics amp; NLP - Lexalytics. https://www.lexalytics.com/blog/build-or-buy-natural-language-processing/

Satwaliya, D. S., Thethi, H. P., Dhyani, A., Kiran, G. R., Al-Taee, M., & Alazzam, M. B. (2023). Predictive maintenance using machine learning: A case study in manufacturing management. *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 872–876. https://doi.org/10.1109/ICACITE57410.2023.10183012

Schröer, Kruse, & Gómez. (2021). A systematic literature review on applying crisp-dm process model. https://doi.org/10.1016/j.procs.2021.01.199

Schwendemann, S., Amjad, Z., & Sikora, A. (2021). A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. *Computers in industry*, *125*, 103380. https://doi.org/10.1016/j.compind.2020.103380

Sensemore. (2024, April). Challenges and considerations in implementing predictive maintenance. https://sensemore.io/challenges-in-implementing-predictive-maintenance/

Serradilla, O., Zugasti, E., de Okariz, J. R., Rodriguez, J., & Zurutuza, U. (2022). Methodology for data-driven predictive maintenance models design, development and implementation on manufacturing guided by domain knowledge. https://doi.org/10.1080/0951192X.2022.2043562

Sipos, R., Fradkin, D., Moerchen, F., & Wang, Z. (2014). Log-based predictive maintenance. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2623330.2623340

Swanson, L. (2001). Linking maintenance strategies to performance. *International journal of production economics*, *70*(3), 237–244. https://doi.org/10.1016/s0925-5273(00)00067-0

Tabassum, A., & Patil, R. R. (2020). A survey on text pre-processing & feature extraction techniques in natural language processing. *International Research Journal of Engineering and Technology (IRJET)*, *7*(06), 4864–4867.

Topper, N. (2023, August). Bag of words model in NLP explained. https://builtin.com/machine-learning/bag-of-words

Turnbull, A., & Carroll, J. (2021). Cost benefit of implementing advanced monitoring and predictive maintenance strategies for offshore wind farms. https://doi.org/10.3390/en14164922

Wirth & Hipp. (2000). Crisp-dm: Towards a standard process model for data mining. in proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining

Wu, W. (2021, February). What is Predictive Maintenance? (Benefits, Cost, Examples). https://coastapp.com/blog/predictive-maintenance/#how-much-does-predictive-maintenance-cost

Zhang, Y., Ren, S., Liu, Y., & Si, S. (2017). A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products [Special Volume on Improving natural resource management and human health to ensure sustainable societal development based upon insights gained from working within 'Big Data Environments']. *Journal of Cleaner Production*, *142*, 626–641. https://doi.org/https://doi.org/10.1016/j.jclepro.2016.07.123

Zhonga, K., Jackson, T., West, A., & Cosma, G. (2024). Natural language processing approaches in industrial maintenance: A systematic literature review. https://www.sciencedirect.com/science/article/pii/S1877050924002060

Zonta, T., Da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, *150*, 106889.

A

Appendix 1

Interview questions for all participants:

1. How often maintenance occurs?
2. What are the challenges of maintenance in the current situation?
3. Have you noticed any patterns or early warning signs of equipment failure that could potentially be monitored for predictive maintenance?
4. Have you ever caused a machine breakage? If yes, what was the breakage and why did it happen?
5. How often do you make wrong judgments regarding the breakage type and what are the consequences of that?
6. What actions do you have to do when a machine breaks?
7. Have you faced any unexpected/unknown failures or breakdowns? How did you report them in the software?
8. How confident are you of using the current way of maintenance? (Reporting to the software, reading the manual etc.)
9. What are the most common maintenance problems you usually face?
10. Are there any specific components that are most problematic?
11. How effective do you think the current preventive maintenance is?
12. What solutions are you proposing to evolve the current maintenance?
13. Do you know what predictive maintenance is?
14. What is your opinion about the concept of predictive maintenance and how do you think it would affect you if it was implemented in the company?
15. How do you usually communicate about the maintenance problems with different teams and how effective do you think your collaboration is? How might this collaboration change with the implementation of predictive maintenance?
16. What factors influence decisions on repair, replacement or preventive actions?
17. Anything else to add to this interview?

B
Appendix 2

Extra interview questions for Supervisors/Managers:

1. How much does maintenance and breakage usually cost for the company?
2. Do you foresee any resistance or concerns among employees regarding the shift to predictive maintenance?
3. Are there any regulatory or compliance considerations that need to be addressed with a new maintenance approach?
4. Are there any external factors or environmental conditions that could affect the machine's performance?
5. Are there any external service providers involved in the maintenance process?
6. What organizational changes or support would be necessary for a successful transition to predictive maintenance?
7. How do you approach training and educating staff about the importance of predictive maintenance?
8. How do you ensure motivation among employees? What are you doing to keep people motivated to do their work?

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY