Utilizing Machine Data for Enhancing Service-Dominant Logic

Thijs Kampkuiper
Student Number: S3104621

March 1, 2024

Faculty of Behavioural, Management & Social Sciences (BMS)
MSc Business Administration (MSc BA); Digital Business & Analytics (DBA)

Supervision at Company
Sander Jansen – Service Operations Manager

University of Twente Supervisors
Dr. Fons Wijnhoven – First supervisor
Dr. Agata Leszkiewicz – Second supervisor

Abstract

The business landscape has experienced a change in marketing ideologies, where the conventional goods-dominant logic has been replaced by service-dominant logic (SDL). This thesis addresses the need for proactive customer service in the evolving business landscape for enhancing SDL. The thesis utilizes a case study of a Computer Numerical Control (CNC) machine manufacturer to illustrate how proactive customer service could enhance SDL by minimizing machine downtime. The findings reveal that predictive maintenance facilitates proactive addressing of potential malfunctions. This thesis contributes to the theory by demonstrating how proactive customer service, driven by predictive modeling on customer-generated real-time data, enhances SDL. This emphasizes service as the primary value and leverages operant resources to enhance customer production output, contributing to competitive advantage. Moreover, customers need to provide data in order to act proactively, highlighting their role as co-creators of value. Additionally, proactive service engages in delivering value propositions, builds relationships, and acknowledges the subjective nature of value perception among customers, aligning with several foundational premises of SDL. This thesis contributes to the practice by emphasizing the importance of cloud-based storage capabilities in enhancing SDL.

Keywords: Service-dominant logic, proactive customer service, predictive modeling, machine data

Acknowledgements

First, I would like to thank Sander Jansen for giving me the opportunity to conduct this thesis at the company and for all the valuable feedback he has given. Additionally, I want to thank all employees of the company who have participated in this thesis in any way.

I would like to thank Dr. Fons Wijnhoven and Dr. Agata Leszkiewicz from the University of Twente for supervizing and providing valuable feedback for this thesis.

Lastly, I want to thank my family and friends, especially my brothers Nick and Stan, for their support during the completion of my master's degree.

Thijs Kampkuiper

March 2024
Enschede

Contents

1 Introduction

1.1 Background/Phenomenon

The business landscape has witnessed a shift in marketing ideologies, the traditional goods-dominant logic is replaced by service-dominant logic (SDL) (Vargo & Lusch, 2004). Economic transactions were traditionally centred around the exchange of tangible goods. However, Vargo and Lusch (2004) introduced a perspective that uses SDL, wherein goods are viewed as mechanisms for delivering services (Vargo & Lusch, 2017). An example of SDL is the use of mobile devices where dedicated mobile phone applications maintain communication with their customers, leading to brand equity, a driving force for improved business performance and sales for a firm (Tran et al., 2021). The era of the Internet of Things (IoT) is aligned with SDL by making use of real-time data. This data can then be used to create system dynamics and allow organizations to adapt to specific information (Vargo & Lusch, 2017). Furthermore, in a study by Lehrer et al. (2018), it is illustrated that big data analytics can serve as organizational resources for service provision (Lehrer, 2018). Service provision has been traditionally reactive, but by gathering and analysing customer-generated data, organizations can provide customer support and provide service when it is needed or even in advance (Lehrer et al., 2018).

This study is motivated by the practical case of a company. This company wishes to remain anonymous. The company is a CNC machine manufacturer that serves customers around the globe, their service provision is currently reactive. The customer calls the organization to report a malfunction or in case the machine does not function properly as the customer desires. The company can solve 90% of the problems remotely, the remaining problems require on-site service. Delays may occur due to unclear causes of malfunctions or unavailability of necessary parts. The company recognizes the limitations of the current reactive service approach and aspires a transition towards SDL-based proactive customer service. With this approach, the company aims to improve its capacity to anticipate service requirements and deliver timely solutions. The strategic shift towards proactive customer service aligns with the goal of the company to offer more efficient and customer-centric services. This indicates that the company wants to emphasize delivering services instead of solely delivering machines. In order to implement such strategies, the data should be properly collected. According to Lee (2014), not all production systems are ready to manage big data due to the lack of smart analytical tools. These technologies enable the prediction of product performance and the autonomous management and optimization of product service needs (Lee et al., 2014).

This transition marks a shift from the company to a proactive approach to customer service, focused on anticipating service and resolving issues in a timely manner, which are in line with more efficient and customer-centric service delivery. Raub and Liao (2012) introduce the concept of proactive customer service. Proactive customer service is defined as a self-starting, long-term-oriented, and forward-thinking approach to service delivery. In this study, the researchers found a significant relationship between aggregated proactive customer service performance and customer satisfaction (Raub & Liao, 2012). Another study by Shin et al. (2017), demonstrates that proactive customer service has a positive relationship with customer patronage behaviour (Shin, 2017). Customers who experience proactive interaction for service score higher for customer patronage behaviour than customers who do not receive this. Customer patronage behaviour refers to the consistent and repeated purchasing choices made by individuals towards a particular business or brand over time (Philippe & N'Goala, 2010). By using customer-generated data, organizations can deliver customer-specific solutions in a timely manner. This is supported by Dey et al. (2015), who describe a government agency, named the District Department of Colombia, which has migrated to networked assets, facilitating real-time monitoring of their operational status and enabling proactive identification and response to problems before customers become aware of them (Dey et al., 2015). By using customer-generated data, the company can embrace SDL by acting proactively and delivering more efficient and customer-centered service.

1.2 Research gap

The literature addresses proactive maintenance strategies and the importance of real-time data utilization facilitated by IoT, but there exists a gap in understanding how proactive customer service can be achieved by utilizing machine data within the framework of SDL. According to Swanson (2001), a proactive maintenance strategy is needed to provide proactive customer service, this concept involves predictive and preventive maintenance to avoid equipment failures (Swanson, 2001). By monitoring the status of machines and analyzing causes and effects, the frequency of preventive maintenance and equipment conditions can be monitored (Zonta et al., 2020). Zonta et al. (2020) defines preventive maintenance as scheduled maintenance, which is performed based on fixed schedules at specific times. Contrarily, predictive maintenance uses data and knowledge to report failures and prevent downtime. Predictive maintenance allows for predicting trends, behaviour patterns, and correlations using statistical or machine learning models to anticipate upcoming failures. This improves the decision-making process for maintenance activities and minimizes downtime (Zonta et al., 2020). Predictive maintenance reduces costs, minimizes downtime, and enhances productivity and quality (Zonta et al., 2020). To use or implement predictive strategies, a substantial amount of data is necessary. A new era of big data is emerging (Chen et al., 2012), this underscores the importance of further studying the phenomenon that data allows real-time data capture of actor-centric behaviour using sensor-based content through data analytics (Vargo & Lusch, 2017). Leveraging machine data for predictive maintenance, as mentioned by Chen et al. (2012) and Zonta et al. (2020), presents an opportunity to bridge this gap and explore the integration of proactive customer service within the SDL framework.

In conclusion, the research of Polo Peña et al. (2014), reveals that ICT capabilities are requirements of value co-creation for the service industry within the SDL framework, as they enable active participation by customers in the value co-creation process (Polo Peña et al., 2014). However, while Polo Peña et al. (2014) highlight the significance of firms in developing ICT capabilities for value co-creation, it is urgent for organizations to leverage customer-generated real-time data to enhance proactive customer service. Moreover, Vargo and Lusch (2015) indicate that cooperative and coordinated behaviour between actors in a developed service ecosystem leads to a more comprehensive representation of markets (Vargo & Lusch, 2015). By collaboration in service ecosystems, the service delivery to the customer becomes more reliable (Agarwal & Selen, 2009). This underscores the importance of enhancing proactive customer service. The implication of proactive customer service is that customers who receive proactive customer service would have a higher customer patronage behaviour level than customers that are receiving reactive service (Shin et al., 2017).

1.3 Purpose of the study

The company delivers its steel processing machines to customers worldwide. Core aspects that are central to the company are automation, optimization of output, and customer service. The company provides various machines capable of performing various operations. Service contracts are included with these machines, which are part of the company's business strategy. In the current situation, the customer service of the company responds when a customer reports a malfunction. This strategy may be suboptimal since it could lead to a negative relationship with a customer. For example, in case of a service failure, a malfunction occurs from the company's machine that is running at a customer, the company's reaction to addressing the issue could either lead the customer to switch to a rival company or help in restoring customer satisfaction. Service recovery, as described by Maxham (2001), is the process attempt of an organization to rectify a service delivery failure (Maxham, 2001). Hart (1990) mentions that a good recovery from a failure generates more goodwill compared to a situation where no issues arise (Hart et al., 1990). Smith et al. (1999) note that is it needed to train customer service employees to recognize service failures to reduce their effects on customer satisfaction (Smith et al., 1999). These statements illustrate the importance of proactive customer service for the company's operations. Newman et al. (2019) mention that monitoring machine health can facilitate early corrective measures, leading to higher service level performance (Newman et al., 2019). This aligns with the thesis's purpose of exploring proactive customer service for enhancing SDL. In the literature, there is no systematic description of how SDL can be realized through proactive customer service. Therefore, the following research question is drawn: "*How can machine data be used to enhance the realization of service-dominant logic?*". Firstly, this research centers on addressing sub-questions such as which machine data is suitable for utilizing predictive models, which models are suitable for predicting machine failures, and how can the company utilize predictive models to enhance SDL through proactive customer service.

It is relevant for the company to explore how data-driven customer service can be realized and what challenges and benefits are involved. The customer service department of the company would like to know which available machine data they can use to set up proactive customer service to embrace an SDL. According to Uhlmann et al. (2008), the feed axes of CNC machines are the most mechanically stressed components due to high process forces and their operation in a rough production environment (Uhlmann et al., 2008). Currently, it is unclear for the company which machine data is useful for predicting malfunctions. Organizations have much data, stored in many files and databases (Power, 2014). Therefore, there must be defined which machine data could be used for predictive modeling to prevent future malfunctions.

This thesis is of relevance for the company because it focuses on developing proactive service support for their CNC machines by using machine data. Improving machine output for their customers may lead to an increase in the company's competitive advantage (Vargo & Lusch, 2008). This thesis contributes to knowledge about the possibilities of machine data and how this is used to establish proactive customer service and thus apply an SDL perspective.

This thesis contains practical relevance for organizations that want to embrace SDL. The paper emphasizes the importance of proactive customer service and the use of data analytics. By analyzing customer-generated data, organizations can deliver services proactively and enhance an SDL perspective. Other businesses can learn from this business case and adapt their strategy accordingly. This thesis is also relevant for organizations that would like to apply predictive models to innovate their customer service, which could result in an improvement in customer experience.

1.4 Outline of the study

This study starts with a literature review related to SDL, proactive customer service, high reliability, and proactive maintenance strategy. Subsequently, the methodology chapter mentions the study design and the data collection, that is used to answer the previously mentioned research questions. After the methodology chapter, the results of the research are shown. The research ends with a global discussion and conclusion based on the literature review and results.

2 Theory

This chapter is a literature review of the different concepts. This chapter reasons that an SDL perspective can be realized through proactive customer service and it posits that proactive customer service is realized through high reliability and proactive maintenance strategy. Firstly, SDL is described. Additionally, the concept of proactive customer service is elaborated. After this, the relationship between SDL and proactive customer service is explained. In the last section of the chapter, a framework for SDL through a proactive maintenance strategy and high reliability is proposed. This framework contains the scope of the thesis.

2.1 Service-dominant logic

The SDL framework, introduced by Vargo and Lusch (2004), offers a marketing perspective that shifts the focus of goods-dominant logic to a view where service is seen as the primary basis of exchange. In SDL, goods are seen as mechanisms to deliver services (Vargo & Lusch, 2017). Vargo and Lusch (2008) outline ten foundational premises (FP) of SDL, which are described below (Vargo & Lusch, 2008).

FP1: Service is the fundamental basis of exchange.
FP1 underscores the idea that value is derived from knowledge and skills, as the basis of all exchange. Exchange revolves around the application of operant resources, possessed by organizations. Service provision is the core drive of value creation in economic exchanges.

FP2: Indirect exchange masks the fundamental basis of exchange.
FP2 argues that because services involve a mix of goods, money, and institutions, it is not always clear that the exchange is based on services. It highlights the complex nature of service exchanges. Traditional economic perspectives focus on tangible goods and transactions, but this overlooks the essential role of service for value creation.

FP3: Goods are a distribution mechanism for service provision.
This premise mentions that goods (both durable and non-durable) are needed to distribute services. Value is derived through the use of the goods. It highlights the interdependence between goods and services, wherein goods are acting as pathways for the delivery of services.

FP4: Operant resources (knowledge and skills) are the fundamental source of competitive advantage.
FP4 suggests that gaining competitive advantage lies in the effective use of an organization's operant resources. Operant resources, which encompass knowledge and skills, create a competitive advantage through continuous learning, skill development, and knowledge management.

FP5: All economies are service economies.
This statement mentions that services always have been part of economies, but the recognition that services become more noticeable due to specialization and outsourcing. It emphasizes that value is in any economic exchange rooted in the delivery of services and experiences to customers.

FP6: The customer is always a co-creator of value.
The customer plays an active role in co-creating value through interacting with the service provider, co-creation is enabled by collaboration. The value is not solely created by the provider but by engagements between service providers and customers. This FP focuses on customer-centricity by understanding customer needs in the value-creation process.

FP7: The enterprise cannot deliver value, but only value propositions.
FP7 suggests that an enterprise can offer its resources to create value, but value cannot be created by the enterprise independently; instead, it is generated through interactive collaboration with customers. The service provider can only propose value propositions. Customers need to actively participate in creating value through interactions and feedback.

FP8: A service-centered view is inherently customer-oriented and relational.
This statement suggests that a service-centered perspective is focused on customers and relations. This has come about because services are defined by the benefits determined by customers and created together with customers. By building strong customer relationships, organizations foster mutual understanding thereby enhancing the co-creation of value.

FP9: All social and economic actors are resource integrators.
FP9 argues that all parties involved in social and economic activities participate in combining resources. A setting for creating value is networks, where actors with different resources collaborate to create value. By those networks, innovations, efficiency, and collective value creation within ecosystems are enabled.

FP10: Value is always uniquely and phenomenologically determined by the beneficiary.
FP10 mentions that the perception of value is individual and subjective. This means that value is unique for each individual. It underscores the importance of understanding and responding to the differences in customer needs to address individual shortcomings and preferences effectively.

The ten foundational premises, created by Vargo and Lusch (2008), highlight the key principles of SDL. Building upon these principles, Lusch et al. (2010) mention that "self-adjusting" occurs when actors sense and respond to their continued market relevance and viability. This helps to overcome the cognitive gap among the actors (Lusch et al., 2010). Through these basic principles and the concept of self-adaptation, proactive service emerges as a solution to enhance market relevance and facilitate more effective value co-creation. If actors act proactively, the cognitive gaps will be bridged and value co-creation in the market environment will be enhanced.

By embracing an SDL perspective, the company must recognize that providing CNC machines as mechanisms is essential to delivering service and that the customer plays a crucial role in generating value. Currently, the company wants to bridge the gaps among market actors and improve value co-creation. By proactive initiatives of the company, downtime for customers' machines will be prevented, thereby aligning with SDL. By integrating SDL into the company, stronger customer relations will be built, value co-creation will be enhanced, and competitive advantages will be gained.

2.2 Proactive customer service

Proactive customer service is needed to enhance SDL. Shin et al. (2017) define proactive interaction as organizations voluntarily contacting the customer. Reactive interaction, on the other hand, occurs only when the customer seeks assistance or reports a problem to the organization. This section mentions that proactive customer service is enabled by high reliability and a proactive maintenance strategy to enhance SDL. By implementing proactive interaction, organizations proactively anticipate future events by learning from mistakes, predicting potential disruptions, and taking preventive measures to alter future outcomes. The service provider achieves this by anticipating possible issues customers may encounter and offering assistance before the problem arises. Reactive interaction solely focuses on responding to customer complaints and requests for help. Proactive interaction ensures faster service and quicker problem resolution (Shin et al., 2017). Delana (2021) discusses the availability of future customer-service needs information through remote monitoring systems and data analytics (Delana et al., 2021).

2.2.1 Reliability

This proactive approach is crucial in today's complex organizational landscape, wherein reliability is dominant (Weick, 1987; Busby & Iszatt-White, 2014). Weick (1987) argues that organizations and technologies have experienced a rise in complexity. Accidents can result in further unforeseen consequences due to misunderstood interventions (Weick, 1987). These accidents involve issues of reliability, i.e., operating systems are not consistently working effectively and safely. Organizations prioritizing reliability over efficiency often encounter distinct challenges in learning and comprehension (Weick, 1987). These unique problems arise due to the increased complexity of the organization's operations and technologies. Busby and Iszatt-White (2014) say that reliability is fundamental for organizing because it ensures consistent, dependable, and predictable functioning of organizational systems and processes. Reliability is something that is produced by an organization. It is an important competency enabled by organizations, as reliable entities can depend on others and become able to be dependent upon others (Busby & Iszatt-White, 2014). High reliability is enabled by the collaboration between actors, knowledge of the employee, supportiveness of the organizational culture, and employee motivation (Agarwal & Selen, 2009; Wijnhoven et al., 2012; Weick, 1987).

Collaboration
Collaboration is crucial in highly reliable systems. Agarwal and Selen (2009) highlight the necessity for managers to acknowledge the importance of collaboration in service innovation to ensure the timely delivery and reliability of service offerings (Agarwal & Selen, 2009). By collaboration between customer and supplier, the supplier can deliver targeted, and faster service when and how the customer needs it. Lusch and Nambisan (2015) say service innovation emphasizes innovation as a collaborative process that occurs in an actor-to-actor network (Lusch & Nambisan, 2015). Through collaboration between the company and its customers, the reason why malfunctions occur becomes clear. Subsequently, both organizations can discuss suitable solutions to prevent the malfunction for the next time. This can be confirmed by Chaurasia et al. (2020), this paper says that organizations need to expand their view beyond existing resources. Shared value for open innovation requires active collaboration between organizational partners to create value in a problem-solving context (Chaurasia et al., 2020). Success in business ecosystems requires collaboration to leverage a firm's resources and capabilities (Zahra & Nambisan, 2012).

Knowledge employee
Weick (1987) suggest that accidents arise from the inability of individuals operating and overseeing complex systems to adequately perceive and foresee the issues these systems may generate. This is called "requisite variety" because the variety in the systems exceeds the variety in the people who must control it. If individuals lack the necessary variety to engage with the system effectively, they may overlook crucial information, leading to incomplete diagnoses and shortsighted solutions that can magnify rather than alleviate the problem (Weick, 1987). Wijnhoven et al. (2012) explore that, next to motivations, knowledge underlies behavioural repertoires. Organizational wealth can be realized through the application of specialized knowledge and skills, where resources and goods play a secondary role (Wijnhoven et al., 2012). Wijnhoven et al. (2012) describe that knowledge can be divided into four categories: propositional, experiential, performative, and epistemological. Propositional knowledge encompasses meaningful information regarding a subject that has been communicated or conveyed to an individual. Experiential knowledge is having a personal understanding, feeling, or belief about a phenomenon. Performative knowledge goes one step further than experience, like knowing how to ride a bike.

Epistemological knowledge consists of a deeper understanding of why something is as it is (Wijnhoven et al., 2012).

Supportiveness of organizational culture

Culture is relevant in high-reliable organizations because making meaning is an issue for culture (Weick, 1987). The goal of high-reliable organizations is to achieve centralization and decentralization. Individuals must leverage insights gained from past experiences of operators and capitalize on trial-and-error methodologies to accomplish benefits. Individuals require a well-defined hierarchy of authority when an error is happening to deal with the error (Weick, 1987). According to Schein (1990), organizational culture is defined as: *"a pattern of basic assumptions, invented, discovered, or developed by a given group, as it learns to cope with its problems of external adaption and internal integration, that has worked well enough to be considered valid and therefore is to be taught to new members as the correct way to perceive, think, and feel in relation to those problems."* (Schein, 1990).

Schein (1990) divides organizational culture into three levels, observable artifacts, values, and basic underlying assumptions. Observable artifacts encompass a wide array of elements including physical layout, dress code, interpersonal communication styles, ambiance, and other phenomena. Values are beliefs, aspirations, and values that an organization claims to follow. For example, the value of issue selling is defined by Parker and Collins (2010) as making others and leaders aware of particular issues (Parker & Collins, 2010). According to Crant (2000), issue selling is explicitly addressed through proactive behaviours (Crant, 2000). Basic underlying assumptions are often taken for granted. It contains underlying, unconscious beliefs, values, and norms that guide the behaviour and decision-making in the group of the organization (Schein, 1990).

To achieve high reliability for the company, it is necessary that the organizational culture ensures that employees are encouraged and empowered to perform service proactively. Employees are more likely to anticipate proactive behaviour when the organizational culture supports them. This reduces the perceived risks associated with taking initiative.

Employee motivation

Wijnhoven et al. (2012) say that the handling of unanticipated threats to high reliability is enabled by the appropriate behaviour of employees (Wijnhoven et al., 2012). This means that effective responses to unforeseen threats in high-reliability contexts rely on the appropriate actions of employees. An individual's motivation has an influence on effective incident handling. Motivation helps people decide what to do and affects how they behave. Weick & Sutcliffe (2007) express that the motivations of anticipation and containment facilitate effective management of unforeseen events, ensuring high reliability within organizations (Weick & Sutcliffe, 2007). Wijnhoven et al. (2012) describe that in highly reliable organizations, characterized by consistent and effective operations, individual motivation can be divided into two categories: anticipation and containment. Anticipative motivation focuses on prevention, paying attention to weak signals for problems. Furthermore, It assists the organization in identifying failures, predicting the consequences of failure, and halting the progression of undesirable effects. Containment motivation focuses on managing the undesirable effects post-failure and aims to enhance the ability to recover from such failures. An example of containment motivation is limiting the damage of a failure (Wijnhoven et al., 2012).

2.2.2 Proactive maintenance strategy

In order to enable proactive customer service, next to reliability, a proactive maintenance strategy is needed. Swanson (2001) says a proactive maintenance strategy involves predictive and preventive maintenance to avoid equipment failures. By monitoring the status of machines and analyzing causes and effects, proactive maintenance determines the frequency of preventive measures and identifies the equipment conditions suitable for predictive maintenance (Swanson, 2001). Zonta et al. (2020) distinguish preventive maintenance as scheduled maintenance, which is performed based on fixed schedules at specific times. Predictive maintenance uses data and knowledge to report failures and prevent downtime. Predictive maintenance allows for predicting trends, behaviour patterns, and correlations using statistical or machine learning models to anticipate upcoming failures. This improves the decision-making process for maintenance activities and minimizes downtime. Predictive maintenance reduces maintenance costs, and downtime, and improves productivity and quality (Zonta et al., 2020).

According to Mason and Mitroff (1973), empirical inquiry systems help to collect, store, and analyze data for problem identification (Mason & Mitroff, 1973). Wijnhoven (2023) mention that data analytics aims at creating useful input towards effective decision-making (Wijnhoven, 2023). Predictive analytics is based on historical and current data to predict future outcomes (Gandomi & Haider, 2015). By analysing machine data, proactive measures can be taken to recognize potential future malfunctions and prevent them in a timely manner.

2.3 SDL and proactive customer service

This section elaborates on the relationship between SDL and proactive customer service. This is done by the foundational premises of SDL, proposed by Vargo and Lusch (2008). Several principles of SDL, see below, can be enhanced through the development and implementation of proactive customer service.

FP1: Service is the fundamental basis of exchange.
By monitoring machine data, proactive service could be delivered to the customer in case a potential malfunction arises. The service is anticipated to the customers' needs and give valuable assistance to the customer so the machine will not get to a standstill.

FP4: Operant resources are the fundamental source of competitive advantage.
By the use of predictive maintenance, proactive service leverages the operant resources of the organization to gain a competitive advantage. The efficient use of knowledge, i.e. the predictive models, improves competitive advantage because of machine malfunction prevention.

FP6: The customer is always a co-creator of value.
Acting proactively toward the customer is needed to collaboratively communicate with customers to address potential solutions before issues can arise. Customers participate in the creation of value by providing input to ensure the optimal functioning of their machine.

FP7: The enterprise cannot deliver value, but only value propositions.
Proactive service aligns with the SDL perspective that the organization only can deliver value propositions. The company can engage customers proactively with the use of prediction models, suggesting adjusting machine settings to prevent machine downtime. By collaboration between the customer and the company, the value of the solution is higher.

FP8: A service-centered view is inherently customer-oriented and relational.
Proactive service is customer-oriented and relational. By actively contacting customers in case

a malfunction could arise, solutions need to be provided. With proactive service, the company embraces a service-centered view that prioritizes customer relationships.

FP10: Value is always uniquely and phenomenologically determined by the beneficiary.
Proactive service can be adjusted differently for the company's customers, i.e. some customers would take more risk in preventing the malfunctions than customers. By tailoring solutions based on the predictive models, organizations acknowledge that customers perceive value differently from each other.

2.4 Framework for SDL

A causal model, that presents how SDL can be realized through proactive customer service, is shown in Figure 1. Eventually, by providing proactive service, the distance between an organization and its customers becomes smaller, which aligns with the SDL perspective.

Proactive customer service is enabled by high reliability. This consists of the collaboration between actors, available knowledge of the employee, level of supportiveness of the organizational culture, and the motivation of the employee. Proactive customer service is also enabled by proactive maintenance strategy. A proactive maintenance strategy consists of preventive maintenance and predictive maintenance. Predictive maintenance is enabled by making use of machine data and knowledge that can be used to analyze machine data.

In the following parts of this thesis, the research will not be focused on the entire model. See for the scope of the thesis Figure 1. As a first step, to gain insights into the applicability of developing proactive customer service for the company's CNC machines, there will be focused on how predictive maintenance can be enabled by using machine data and knowledge. The first study focuses on which machine data can be used for predictive maintenance. Study 2 focuses on which model is the most suitable for predicting malfunctions. The third study focuses on the deployment of the predictive model to enhance SDL through proactive customer service.

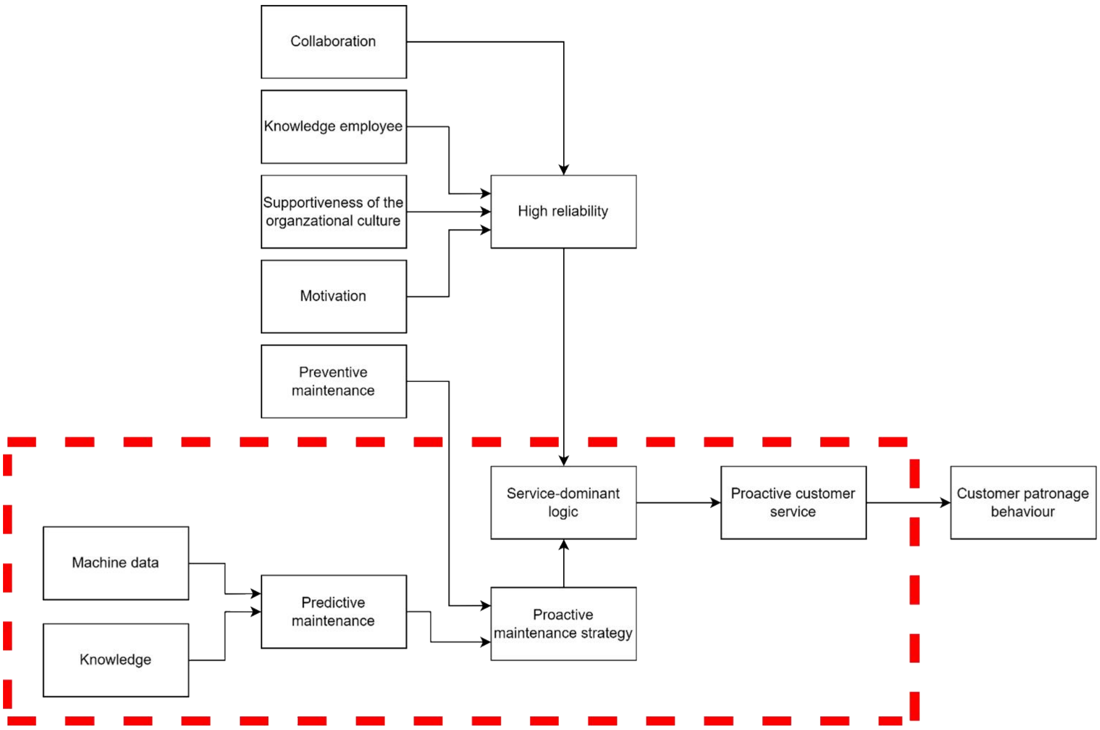

Figure 1: Framework for high reliability and proactive maintenance strategy for enhancing service-dominant logic. *This model illustrates the relationships among machine data, knowledge, and organizational factors influencing the service strategy within a highly reliable framework. The pathway shows how using machine data and certain knowledge leads to predictive maintenance. Combined with preventive maintenance, this leads to a proactive maintenance strategy. Collaboration, knowledge of employees, supportiveness of the organizational culture, and motivational aspects lead to high reliability. Proactive customer service encompasses high reliability and a proactive maintenance strategy. A service-dominant logic approach can be enhanced through proactive customer support. The red dashed box shows the scope of this thesis.*

3 Methodology

In the following chapter of this thesis, the study design, the research approach, and the data analysis method are described. The first section of the methodology chapter contains a section that describes the study design and why this study design was selected. Furthermore, the research approach is elaborated. Additionally, the data collection methods are mentioned, which are used for this research. Finally, the data analysis of the different studies is described.

3.1 Study design

This study aims to investigate the utilization of machine data and knowledge to predict machine failures, enabling proactive customer service from an SDL perspective. The data is collected by making use of a case study. A case study provides the opportunity for a detailed analysis of complex issues within their real-life context (Crowe et al., 2011). This case study is conducted at a CNC machine manufacturer. This company produces different machines that can process different pieces of metal, for instance, tubes, plates, and H-profiles. The organization wishes to collect and analyze data that can be used to improve customer service. The case study is limited to only one machine, MACHINE X. Currently, there are 104 of these machines distributed worldwide. First, it is necessary to investigate which machine data may be relevant to monitor and implement in a model for predicting machine failures. Subsequently, relevant machine data will be checked for availability. Additionally, a model to predict machine failure will be built on the relevant data. Finally, the thesis illustrates what the deployment of predictive models will look like.

3.2 Research approach

In this master thesis, a mixed-method approach is used. A mixed-method approach is an intellectual and practical synthesis that makes use of quantitative and qualitative research (Johnson et al., 2007). Mixed methods research can offer potentially insightful findings that a single method may not be able to provide (Venkatesh et al., 2013). According to Johnson (2007), there are three types of mixed-method research: qualitative-dominant, equal status, and quantitative-dominant. In qualitative-dominant mixed-method research, researchers emphasize the significance of incorporating quantitative data and methodologies alongside qualitative research. Quantitative-dominant mixed-method research emphasizes the significance of incorporating qualitative data and methodologies alongside quantitative research. The equal-status mixed-method research lies between qualitative-dominant research and quantitative research methods (Johnson et al., 2007). This research uses the equal-status mixed-method because the thesis first looks at which machine data is relevant to predict failures. This will be done using qualitative data. Subsequently, the study investigates whether the machine data is obtainable and constructs predictive models using quantitative data.

In this research, a sequential exploratory research strategy is used. This means that qualitative data were collected and analyzed, followed by a phase of quantitative data collection and analysis. This research design proves to be exceptionally useful for researchers who are developing a new instrument, i.e., a predictive model (Creswell, 2009). The main research question is: "*How can machine data be used to enhance the realization of service-dominant logic?*". The research begins with Study 1, as in this case, involving explorative interviews with the company's experts regarding malfunctions of MACHINE X. The results of these interviews are expected to uncover the causes of the machine malfunctions and to determine the relevant machine data that should be used in the model. Subsequently, Study 2 is conducted in which quantitative data is gathered concerning the identified causes of these malfunctions. Study 2 shows which predictive models are suitable for prediction malfunctions of MACHINE X. Finally, Study 3 indicates how predictive models could be deployed to deliver proactive service and how this could be realized in the company to enhance SDL.

3.3 Data collection

3.3.1 Study 1: Context models

Semi-structured interviews are conducted to gather data for the initial phase of the research, aiming to identify the machine elements most influential in causing malfunctions. Digital data can be massive (Power, 2014), and it may take a considerable amount of time to incorporate all the machine data into a predictive model. Therefore, through exploratory interviews with customer service department employees and other relevant employees, the relevant variables of MACHINE X will be taken into account. The interviews will answer the first question of which machine data can be utilized to develop proactive customer service. Five participants were included in the research's interviews. Three of the respondents are employed in the customer service department, one in the quality department, and one in the research and development department. The interviews are being conducted in Dutch. Therefore, individuals can communicate their thoughts more clearly in their native language than when they speak in a second or third language (Kashiha & Chan, 2015). The interview guide which is used can be found in Appendix A. The sub-question that is answered in Study 1 is: "Which machine data is suitable for utilizing predictive models?".

3.3.2 Study 2: Predictive modeling

When the results of the interviews are available, several important elements are selected. Eventually, fictional data will be used to create a predictive model, based on the assumptions of the experts of the company. Fictional data is used since the company does not have machine logs of the operational values of these components. Currently, the program that displays live sensor data does not collect and store this data. Certain variables that affect failure for the company are known by the experts at the company. These variables have nominal, minimum, and maximum values. The predictive model consists of synthetic data, this synthetic data is generated by Rstudio.

The sub-question that is answered in Study 2 is: "Which predictive models are suitable for predicting machine malfunctions?". The second sub-question is answered by understanding and analyzing the synthetic data. Then, from the literature, it can be seen which predictive models are suitable to work with the characteristics of the data. Subsequently, the predictive models can be created.

3.3.3 Study 3: Deployment

The predictive models are created and the model with the best performance is selected in Study 2. Study 3 will focus on the deployment of the predictive model. The deployment study can be split up into several sections: data acquisition, validation of data model, utilizing predictive models for proactive customer service, the potential impact of proactive customer service, and proactive customer service within SDL. This study uses internal documentation of the case company, synthetic data, and relevant literature. The sub-question that is answered in this study is: "How can the company utilize predictive models to enhance SDL through proactive customer service?".

3.4 Analysis

Study 1 is concerned with conducting semi-structured interviews to identify machine elements that lead to failures in the context of the company. By using exploratory interviews with different employees from different departments, relevant variables of MACHINE X are identified. Through interviews, it is possible to get a comprehensive understanding of the topic. The benefit of using interviews as data collection is that they can generate a broader understanding of the topic in question (Gioia et al., 2013). These interviews will reveal dependent variables that are leading to malfunctions of critical components and independent variables that influence the dependent variables. The interviews are analyzed using the Gioia method (Gioia et al., 2013). By employing this approach, the research achieves greater qualitative precision, facilitating the conduct of inductive research. Several experts from the company are interviewed to answer the first sub-question. This knowledge is collected and organized in a way that systems can understand it, this process is called knowledge engineering (Sebastiani, 2002).

Study 2 uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, utilizing its structured framework for comprehensive data mining (Chapman et al., 2000). This method is the de-facto standard process in data mining projects (Schröer et al., 2021). The paper of Overgoor et al. (2019) illustrates how the CRISP-DM framework could develop Artificial Intelligence (AI) solutions for marketing problems (Overgoor et al., 2019). According to Chapman et al. (2000), the CRISP-DM methodology uses sets of tasks that are described at four levels of abstraction: phase, generic task, specialized tasks, and process instance). The CRISP-DM framework consists of business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Chapman et al., 2000). By selecting key variables from the interview results of Study 1, it is possible to create a predictive model with synthetic data. To improve the transparency of the thesis, the script that is used to create the predictive models can be found in Appendix B. The performance of the models is tested using the Root Mean Squared error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared. Finally, it was examined which model scored best on these metrics and is the most suitable model for predicting machine malfunctions.

The final study of this thesis, Study 3, focuses on the deployment phase of the CRISP-DM method for the predictive model. This phase contains several sections, such as the data acquisition, validation of the data model, potential impact of proactive customer service, utilizing predictive models for proactive customer service, and proactive customer service within the SDL perspective. Through internal documentation of the company, synthetic data, and relevant literature, the last sub-question is answered. The analysis of this section illustrates the practical implementation of the predictive model to improve SDL through proactive customer service. The data acquisition section mentions different theories and how the company needs to apply those theories in order to deploy predictive models in the organization. The validation of the data model section highlights the importance of reviewing predictive models based on real data. The utilizing predictive models for the customer service section makes use of simulation, the script of this simulation can be found in Appendix C. This simulation illustrates how the customer service of the company should act in case critical thresholds are exceeded. This section makes also use of relevant literature in order to implement predictive modeling and keep the model relevant. Subsequently, the potential benefits of implementing proactive customer service for the company section mentions what the benefits are of proactive customer service by an analysis of the company's internal documentation of customer downtime regarding MACHINE X. Finally, the section on proactive customer service within the SDL perspective mentions how the company enhances SDL by implementing proactive customer service which aligns with several foundation premises.

4 Results

This section of the thesis presents the findings from five exploratory interviews, each lasting between 25 and 45 minutes, along with the developed predictive models designed to assist organizations in establishing proactive customer service. The aggregate dimension (3rd-order theme) is in this case the components that are leading to a malfunction of MACHINE X. The 2nd-order themes, all explained in the paragraph below, are the causes of the malfunction regarding the components. The 1st order concepts are variables that influence the cause of a malfunction and thus have an influence on a malfunction related to MACHINE X. The second part of this chapter focuses on the creation and selection of the predictive model. The second paragraph is structured using the different phases of the CRISP-DM method. This chapter ends with illustrating the deployment of the predictive model in the company.

4.1 Study 1: Context models

The goal of this section of the paper is to capture and formalize human expertise. By applying knowledge engineering, expert knowledge is collected and organized in a way that systems can understand it (Sebastiani, 2002). The sub-question that is answered in this section is: Which machine data is suitable for utilizing predictive models? The aggregate dimensions of the components that are leading to a malfunction of MACHINE X are displayed below in Table 1. The drill unit, which is the drilling component of MACHINE X, is mentioned in all five interviews as a critical component that leads to malfunctions of MACHINE X. The paper of Uhlmann et al. (2008) aligns with the business case that the feeding axes of the CNC machines are the most mechanically stressed components (Uhlmann et al., 2008). The measuring wheel is mentioned three times in the interviews, and malfunctions regarding the sensors are mentioned in two interviews. EtherCAT and hydraulic unit problems are mentioned in two interviews. Lastly, the malfunctions regarding to drive wheels are mentioned in only one interview.

Table 1: Frequency table for mentioned components that lead to malfunctions *This table summarizes the components that are associated with errors on MACHINE X. The table displays how often the interviewees mentioned the component.*

Component	Frequency
Drill unit	5
Measuring wheel	3
Sensors	2
Hydraulic unit	2
EtherCAT	2
Drive wheels	1

In Figure 2 below are the aggregate dimensions shown that are leading to malfunctions. This figure shows the 2nd-order themes displayed what the cause is of the malfunction regarding the component. If there is no cause in the model shown of an error regarding a component, interviewees did not know what the cause of the error regarding the component is. The first-order themes are explained in the subparagraphs below and not included in this model to improve clarity.

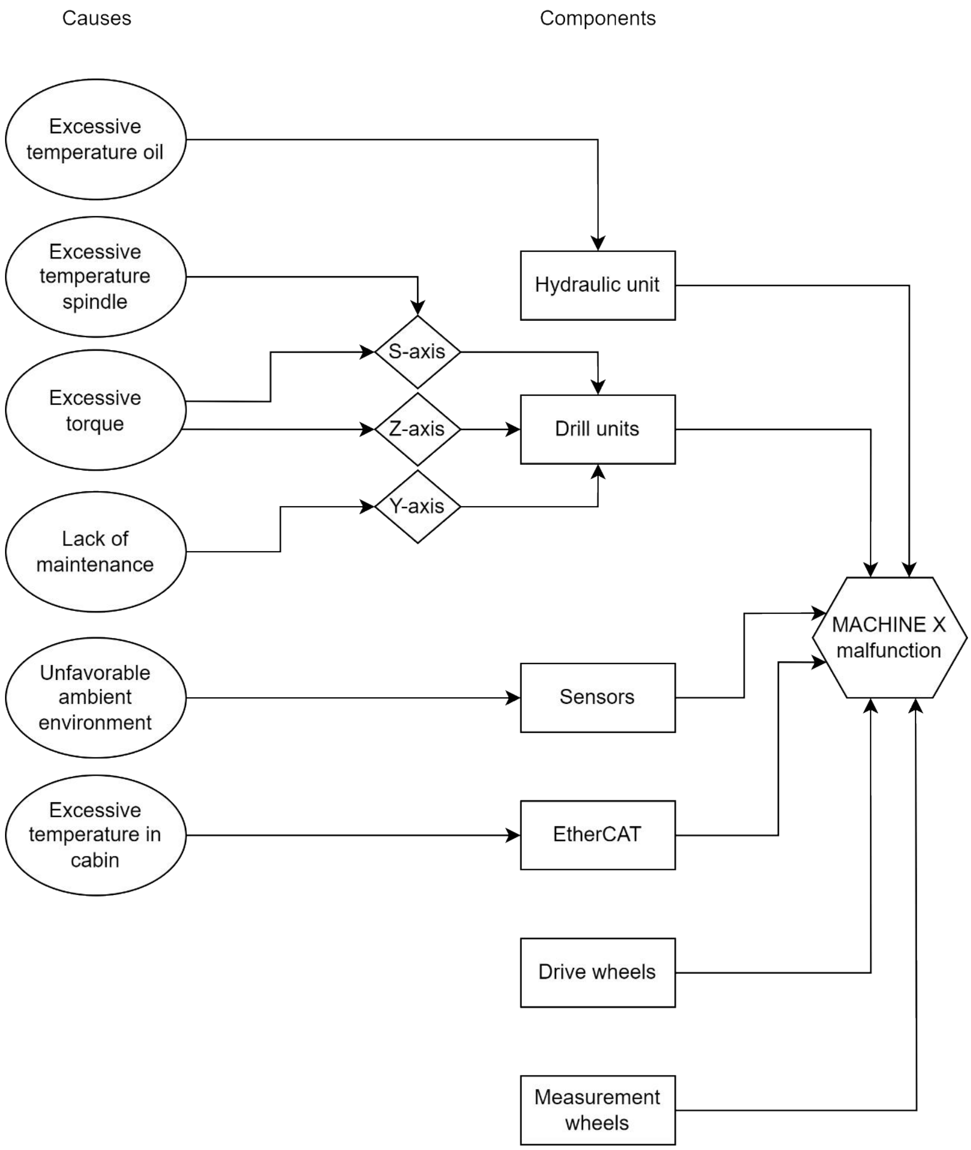

Figure 2: Causal model malfunctions. *An overview of the second-order themes, named as causes, and the aggregate dimensions, named as components, which ultimately lead to a machine malfunction. The cause of a too high temperature of the oil in the hydraulic unit results in a malfunction. The drill unit, operating on the S-axis, Z-axis, and Y-axis, faces issues when there is excessive temperature on the S-axis. Excessive torque leads to malfunctions on the S-axis and Z-axis, and a lack of maintenance leads to errors on the Y-axis. The ambient environment has an influence on the working of the sensors of MACHINE X. A too-high temperature in the cabin leads to EtherCAT errors. Drive wheels and measurement wheels are associated with errors, although the specific causes of these errors are unclear.*

4.1.1 Drill unit

MACHINE X contains three drill units, these are the drill heads of the machine. These components are in direct contact with the manufactured product. All interviewees mentioned that the drill unit is the most mechanically stressed component. Respondent 1 emphasizes that there are problems with the drill units on MACHINE X. Respondent 5 reacts to the question of which factor could prevent malfunctions in case predictive maintenance is possible.

Respondent 1: "I know of several issues regarding MACHINE X, one of them being that the drill spindles often break."

Respondent 5: "The drill spindle is the primary factor leading to a malfunction and could have been prevented if prediction were possible."

The cause of the errors is currently unclear. The company is currently starting with logging service notifications. It is unclear what the main cause is of malfunctions that were reported by clients. Respondent 2 deliberates that the cause of the crash of the drill unit is not clear. Respondent 1 talks about why the cause is not known and the need for reports for problem identification.

Respondent 2: "The drill unit is a component; it is a motor and spindle combined. The bearings are located at the front of the head. It is suspected that the bearing system poses a problem. We have now replaced the bearing system. It is now a matter of waiting to see how this performs in the field. In 75% of the cases with the unit, the issue is with the bearing system. The cause is a crash for which we cannot detect the reason."

Respondent 1: "I know that the bearings and the bearing system are an issue about the drill unit, but why the bearing system fails, I do not know. This is just what I hear in passing. Ideally, you would want to support this with reports; these reports could indicate that, for example, out of 30 reports regarding the unit, 25 were related to the bearing system."

The drill unit has to deal with the Z-axis, Y-axis, and the S-axis. The Z-axis is the movement that moves the drill to the material and back. The S-axis is the movement of the drill inside, this is the motion that makes the drill spin. The Y-axis is the movement of the drill to get to the desired position. The interviews revealed that the movement on the Z-axis and the S-axis mainly leads to malfunctions because they have to deal with heavier forces. Respondent 3 highlights how the S-axis and the Z-axis are more stressed than the Y-axis.

Respondent 3: "Mechanical components are the most stressed components, which is no different than with MACHINE X. I think that the Z-axis is in the drilling direction. The S-axis in the drill spindle; rotates itself. These (axes) receive the first impact with milling and clashes. Positioning, of course, as well; that is built up with logging, and this usually goes well."

Causes Y-axis

The drill moves between the Y-axis to be in the correct position. The interviews revealed that maintenance has an impact on the health of the Y-axis. In case MACHINE X is lagging on the appropriate maintenance on the Y-axis, the drill unit will run stiffer on the Y-axis which could lead to malfunctions. Respondent 3 describes maintenance issues with customers.

Respondent 3: "The most important factor is machine maintenance [..] The mechanical parts just need more maintenance. A spindle that constantly moves up and down or a bearing that does almost nothing, is a difference. In this the spindles are important, the Z-axis, Y-axis, and S-axis are important in MACHINE X."

Causes S-axis

The S-axis is the movement around the axis that goes into the material. Interviews showed that malfunctions are happening due to a too-high temperature and/or torque. The interviewees highlighted that this can be used to predict malfunctions. Several variables influence the phenomenon, that there is too high torque and/or temperature reached, according to the interviewees. Respondent 3 talks about the relevancy of torque and temperature to say something about the health of the unit.

> *Respondent 3: "Yes, then you will, as with a spindle (Z-axis, S-axis), have to read out the torque. You provide a certain nominal torque, indicating how new the spindle is [...] If the temperature gets too high of the spindle, it triggers a self-shutdown. It stops itself. There is already some protection built in, but by then, it's already too late. In the Netherlands, we bring a spare drilling spindle, but abroad, you can easily be down for a week. That's data, and you can work with that."*

Variables that influence causes S-axis

The Revolutions Per Minute (RPM) is included as a variable because that the drill rotates in the material. Respondent 1 highlights the variable of RPM on temperature and torque.

> *Respondent 1: "We have a spindle run-in procedure when it comes from the factory. This is a half-hour program, gradually increasing from low to high speeds. The higher the speed, the more heat is generated, and the more torque is required."*

Sometimes there are errors regarding the spindle of the drill unit. How longer the spindle runs, stiffer the spindle will run. Respondent 5 mentions that the spindle run time has an impact on the temperature and torque.

> *Respondent 5: "If the temperature is high has to do with the production time of a spindle. There is a difference in whether a spindle runs for 200 hours per year or 20,000 hours per year. If you take that time into account. Some customers are maybe running 30 minutes in a day and are having no trouble regarding the drill unit, if the spindle runs for a long time it needs more torque due to wear and tear."*

The effectiveness of the running time could influence the health of the drill unit. How high the effectiveness of the running time will be, leads to more required torque, and this leads to a higher temperature of the drill unit.

> *Respondent 1: "Maybe there are customers who drill, but out of the 24 hours, they are only drilling for 30 minutes. A percentage will come out, indicating how much % of the time he is actually drilling. Customers who do have problems have a number that is 5 times higher than customers who do not have problems."*

The amount of force that is needed to process the material influences the torque that is required. How much more force is used in the production process, how more torque is required. Due to the large force, the temperature will be higher. Respondent 1 questions the influence of force and the type of drill on the torque.

> *Respondent 1: "Perhaps all drilling units break down for all customers who have high forces."*

The type of lubricate that is used by the customer has also an effect on the temperature and torque of the spindle. The amount of lubricate applied in machining processes affects the temperature and torque of the S-axis. Respondent 5 talks about the difference between the company's used lubricate and others.

> *Respondent 5: "The type of lubricate for the drill and the quantity of lubricate used has an impact on the lifespan of the bearing. There are many different types of grease. Often, customers economize on lubricating. If they have a choice between a pot of 15 euros or 10 euros, customers often choose the 10-euro lubricate. This results in a loss of quality. Alternatively, if there is not enough grease used, the drill will run stiffer which leads to a higher torque and temperature."*

Maintenance influences the temperature and torque of the S-axis. If the maintenance status of the machine regarding the S-axis is not sufficient, the machine will run less smoothly generate more heat, and require more torque to work properly. This is mentioned by respondent 3 earlier in this section.

Respondent 2 mentioned how the ambient temperature influences the temperature of the spindle on the S-axis. How hotter the outside temperature is, how hotter the temperature is on the S-axis. When the ambient temperature is high, according to respondent 2, the spindle functions with more torque.

> *Respondent 2: "How it's made here is the best environment. The machine is calibrated here, and it is exactly right as it should be. If it gets warmer, you have to deal with the fact that the spindles, lubricated with grease, will run faster".*

The text above mentions several independent variables that affect the temperature and torque. The independent variables are divided into two groups, operational values and ambient values. The independent variables consist of the RPM, spindle run time, productivity of running time, force, and pulse drill lubricate are variables related to operational values. The environmental values include maintenance, type of lubrication, and ambient temperature. In Figure 3 below, the variables are displayed that influence the torque and the temperature of the S-axis.

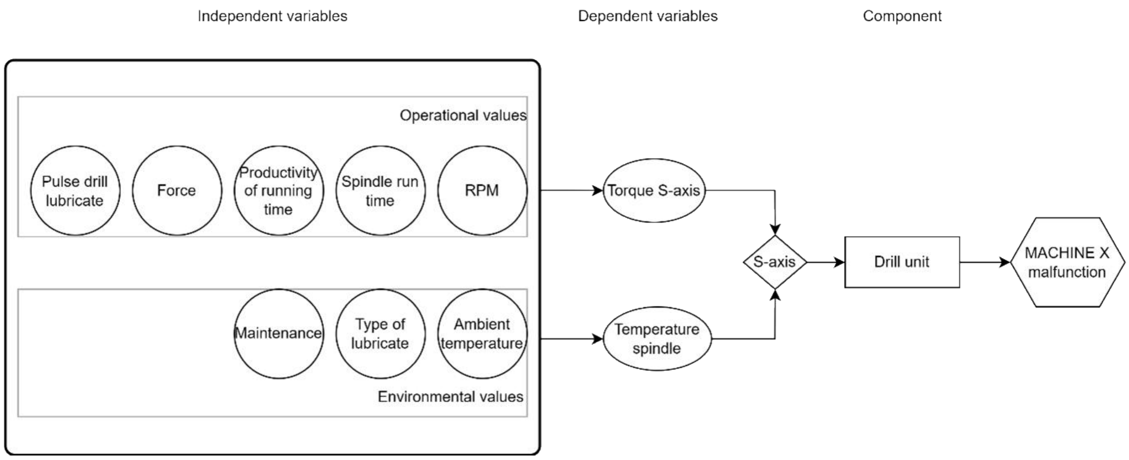

Figure 3: Causal model S-axis. *An overview of the independent variables influencing the dependent variables that lead to malfunctions on the S-axis of the drill unit. The independent variables consist of operational values and environmental factors. The dependent variables include the torque and temperature of the spindle. Operational values contain Rotations Per Minute (RPM), spindle run time, spindle productivity, the force the unit deals with, and the pulse of drill lubrication. Environmental values contain ambient temperature, type of used lubrication, and maintenance.*

Causes Z-axis

The Z-axis has to deal with the movement of the drill towards the material. A too-high torque on the Z-axis will lead to malfunctions regarding MACHINE X. The interviews only highlighted the importance of torque regarding the Z-axis.

Variables that influence causes Z-axis

Almost all variables that influence the S-axis's torque and temperature, have an influence on the torque of the Z-axis. Only the variable RPM does not affect the torque of the Z-axis because the RPM is a variable that rotates around the S-axis and this is the movement that the S-axis makes. In Figure 4 below, the variables are displayed that influence the torque of the Z-axis.

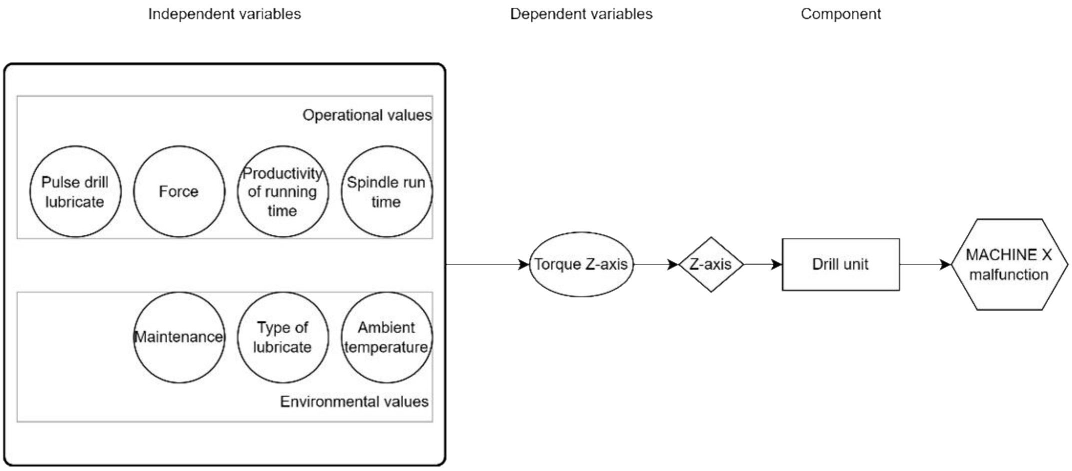

Figure 4: Causal model Z-axis. *An overview of the independent variables influencing the dependent variable that lead to malfunctions on the Z-axis of the drill unit. The independent variables consist of operational values and environmental factors. The dependent variable that leads to errors is the torque. Operational values contain spindle run time, spindle productivity, the force the unit deals with, and the pulse of drill lubrication. Environmental values contain ambient temperature, type of used lubrication, and maintenance.*

4.1.2 Measuring wheel

Because of malfunctions regarding the measuring wheel, the product has measurement deviation. The interviews revealed that the measuring wheels are responsible for this malfunction. Respondent 1 reveals the cause of the error. Respondent 3 covers the consequences of the error. The respondents of the interviews did not mention variables that impact the health of the measuring wheel.

> *Respondent 1: "[. . .] With the dimensional deviation, we sometimes encounter issues where the measuring wheel rises."*

> *Respondent 3: "Especially with measuring deviation, it's not explainable what a measuring wheel does. For example, if you have a tube, it turns out that when tubes are made, you have a different deviation than with profiles, H-profiles, and IPEs. They are quite unreliable."*

4.1.3 EtherCAT

Ethernet for Control Automation Technology (EtherCAT) is a communication tool for automated systems. EtherCAT enables communication between devices and machines in a manufacturing or control system. The temperature of the cabin where the EtherCAT components are has an influence on the health of the components. Problems regarding EtherCAT are difficult to find. Respondent 4 talks about what EtherCAT is and what the role of temperature is on EtherCAT malfunctions. Respondent 2 shares the same intuition that temperature plays a role in malfunctions regarding to EtherCAT.

Respondent 4: "You have a system, let's say a PC. It has a program in it; EtherCAT is essentially the operating system, the PLC, which communicates with blocks in the machine, the input card, and the output card. So, you need to investigate what is the cause of this malfunction. Is the block defective, or is it the communication? Is there a cable that needs to be plugged in and out every time, so to speak? You have to see it as the PLC communicating with the machine. Malfunctions regarding EtherCAT are difficult to find [..] Temperature can, in some cases, play a role, especially with an EtherCAT issue. If you are in warm areas and your system lacks air conditioning, it may occur that you experience EtherCAT malfunctions in certain cases. However, to my knowledge, we currently have good feedback that systems in warm countries are all equipped with air conditioning. I hesitate to provide a definitive answer as external factors, including temperature, may have an impact in some cases."

Respondent 2: "Heat is something to deal with because electrical components can fail due to a malfunction with the fan; generally, we then encounter EtherCAT issues."

4.1.4 Hydraulic unit

The hydraulic unit has the function of delivering hydraulic energy, so the CNC machine can execute different movements. Respondent 1 deliberates on the importance of the hydraulic unit for MACHINE X.

Respondent 1: "If, for example, a hydraulic unit fails, the machine cannot operate anymore."

Respondent 2 revealed that pressure and pumping frequency have an influence on the temperature of the oil in the hydraulic unit. If the pressure falls, the unit needs to pump more frequently to generate pressure. How more often the unit is pumping, how higher the temperature will be. A high temperature will lead to a malfunction in the hydraulic unit.

Respondent 2: "And regarding hydraulics, I'm thinking if you have certain components, valves, or cylinders. It emits a certain pressure, and when that pressure is mainly reached in the horizontal and vertical cylinders, you essentially switch off, but the pressure remains on the components. However, if that pressure drops, then, for example, after a few seconds, it starts pumping again. If this happens very frequently, more than usual, you notice that there is more frequent pumping. If we say that it's okay to pump again after 20 times, that's fine, but if you have to pump more and more frequently, you generate more heat in the hydraulic system. This can lead to pump failure or other malfunctions. It's also not good for the hydraulic oil if it gets warm very often. Could you determine if you have to pump more and more frequently, what is the problem, and in which part or clamp is the problem?."

Respondent 1 talks about how the ambient environment and the capacity of the tank influence the temperature of the hydraulic unit. When the ambient temperature of the unit is high, it would be more difficult for the unit to cool down. If the capacity of the tank of the hydraulic unit is high, it would take the unit more time to reach a high temperature. But how bigger the tank, how longer the cooldown period will take.

Respondent 1: "I see it as the engine block of a car. If it's stationary at 50 degrees, it stays at 50 degrees. If you place it outside at -20 degrees, the car cools down much faster [. . .] The capacity of the unit's tank also plays a role; if the tank is large, it will take longer to reach a high temperature. However, the drawback is that it cools down less quickly at a high temperature."

In Figure 5 below the variables are displayed that influence the oil temperature of the hydraulic unit.

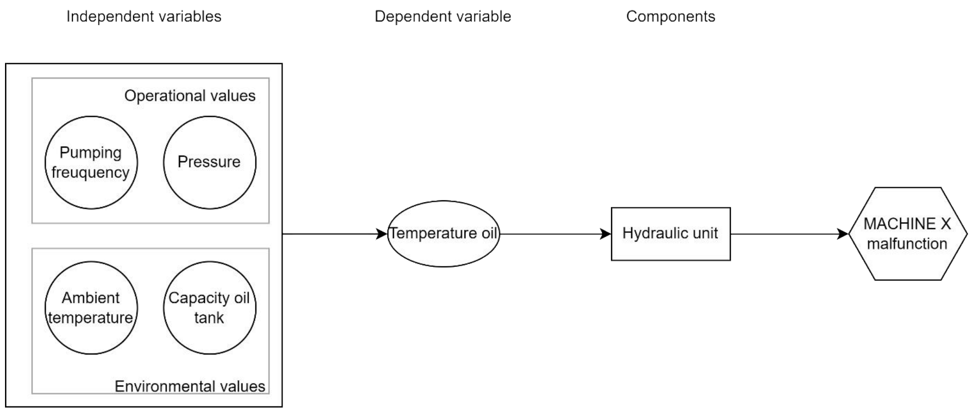

Figure 5: Causal model hydraulic unit. *An overview of the independent variables influencing the dependent variable that lead to malfunctions on the hydraulic unit. The independent variables consist of operational values and environmental factors. The dependent variable that leads to errors is the temperature of the oil. Operational values contain the amount of pressure applied and the pumping frequency. Environmental values contain ambient temperature and the capacity of the oil tank.*

4.1.5 Sensors

There a many sensors present in MACHINE X, and those sensors carry out different functionalities. Currently, it is not clear at the company which sensors are causing the most malfunctions. Respondent 4 talks about that it is not clear what the critical sensors are. Respondent 5 mentions the different influences of the ambient environment. The interviews showed that sensors are sensitive components in a machine. That is why the ambient environment has an influence on the status of sensors.

> *Respondent 4: "But I don't have insight into which sensors often malfunction in relation to MACHINE X. We should examine each system to determine which sensors are affected at the moment. Our system is not set up to extract that information easily; each case would need to be reviewed individually."*
>
> *Respondent 5: "I think humidity, lubrication, exposure to the outdoors, those kinds of things affect sensors. This varies by customer. Some machines are in the desert with 40 degrees Celsius, and some with -30 degrees Celsius."*

4.1.6 Drive wheels

The drive wheels take care of the positioning of the beams in the machine. If these do not work, it is not possible to move the beam through the machine. Only respondent 3 mentions the context of such an error for MACHINE X but could not mention which values influence this phenomenon.

> *Respondent 3: "There are also issues with the drive wheels; this is that heavy beams do not come into position [. . .] In MACHINE X, there is a speed connection in it that can simply loosen. This causes the beam not to come into the correct position, leading to oscillation."*

4.1.7 Conclusion interviews

The interviews show that the drill unit, measuring wheels, EtherCAT, hydraulic unit, sensors, and drive wheels lead to failures of MACHINE X. What the causes are for failures of the driving wheels and measuring wheel are unclear. The ambient environment affects failures with respect to the sensors. The ambient temperature of the cabin affects the health of EtherCAT components. The pressure and frequency of pumps affect the health of the hydraulic pump, as well as environmental factors such as ambient temperature and tank capacity. The drill unit is the drill head of MACHINE X. MACHINE X contains three of these drill heads. The drill unit deals with the S-axis, Z-axis, and Y-axis. The Y-axis ensures that the drill head is positioned correctly. Faults regarding the Y-axis are only affected by maintenance. The S-axis is the rotary movement around the axis. Faults of the S-axis come from excessive temperature and torque. Operational values that affect temperature and torque are the RPM, spindle run time, productivity of running time, force, and pulse of the lubrication of the drill. Environmental factors such as the ambient temperature, and the type of lubricate that is used for the drill and maintenance influence the torque and temperature of the S-axis. For failures with respect to the Z-axis, only excessive torque value would play a role. The Z-axis has to do with the movement of the drill head towards the material. On this, the same variables have a role as with the S-axis. Only RPM has no influence on disturbances here with respect to the Z-axis because RPM is the speed of movement of the S-axis.

4.2 Study 2: Prediction temperature drill unit

Predictive models of the S-axis of the drill unit will be created regarding the excessive temperature. All interviewees mentioned the drill unit as a critical factor for a malfunction regarding MACHINE X. Due to the complexity of creating predictive models and the unnecessary need for all models to understand the execution of proactive service, only the model concerning the temperature of the S-axis was chosen to be created. This section elaborates on how fictitious machine data and knowledge can enable predictive maintenance. The development of the predictive model is done by the CRISP-DM method. The outline of this section aligns with the different phases of the CRISP-DM method. See Figure 6 for the CRISP-DM model, which is adopted from Chapman et al. (2000).

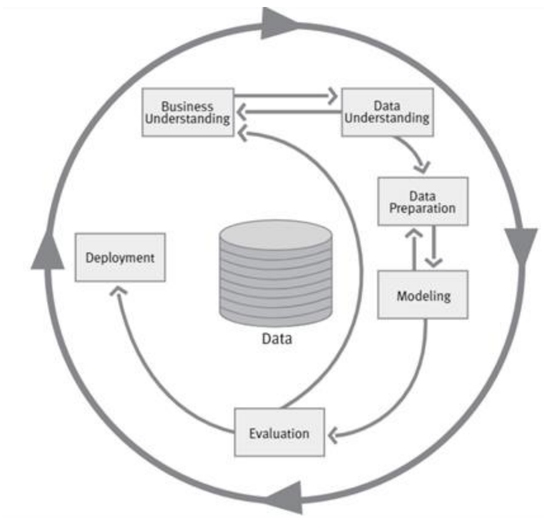

Figure 6: An overview of the CRISP-DM methodology adapted from Chapman (2000). *The CRISP-DM methodology revolves around data. The figure shows the iterative process of creating data mining solutions. The methodology comprises business understanding, data understanding, data preparation, modeling, evaluation, and deployment.*

In order to answer the main question of the thesis, this paragraph will focus on understanding the data in order to create predictive models. By understanding the characteristics of the data, the literature can help to provide suitable models for the specific problem, i.e. predicting the temperature of the spindle. By using relevant metrics, the models can be evaluated on performance. Ultimately, the model with the best performance should be used for predictive modeling, which answers the second sub-question of this thesis.

4.2.1 Business understanding

The business understanding phase of the CRISP-DM model outlines the objectives to be accomplished through data analysis within a particular business setting. The outcome of this process is utilized to establish the objectives of the data analysis, while simultaneously assessing the context of the data mining situation (Chapman et al., 2000). In the company's current situation, no machine data is present because the machine lacks sensors that capture relevant data. For this reason, synthetic data was used for the variables given by the interviewees that affect the excessive temperature of the drill unit regarding the S-axis. The synthetic data is based on the assumption that the interviews reflect the real-world scenario. The synthetic data was generated in RStudio, displaying values for variables that fall within the minimum and maximum ranges. This is done to make the simulation resemble a real-world scenario. To make the models more predictive, a relationship between variables and temperature is intentionally created. However, in real-life situations, this does not always apply and should be tested to confirm its validity.

The overall goal is to simulate the previously found variables from the interviews when they lead to excessive temperature with respect to the drill unit of the S-axis, so that timely action can be taken to prevent the failure. If the customer service department has a tool that can intervene with regard to breakdowns, SDL can be achieved by the company through proactive customer service. The goal of this simulation is to illustrate how predictive maintenance could be realized for proactive customer service.

4.2.2 Data understanding

The data understanding phase of the CRISP-DM method deals with reviewing available data, collecting data, identifying relationships, and assessing the quality of the data (Schnell et al., 2019). Predicting failure is only delineated by the temperature of the drill unit on the S-axis. In the company's case it is stated that when the temperature reaches 100 degrees Celsius, the machine will break down. The interviews revealed that RPM, spindle run time, the productivity of running time, force, pulse drill lubricate, ambient temperature, type of lubricate, and maintenance affect the temperature of the spindle on the S-axis. A minimum, a maximum, and a nominal value were given for the variables, these were retrieved internally from the company. See Table 2 for the variables and their values. For certain values, "NA" is displayed. This means that this value is not applicable. This occurs when the value lacks a nominal value, as it falls between the minimum and maximum values based on the operation. For the maximum pulse drill lubricate a value of "NA" is displayed. This is because the nominal value is a standard setting and other values could only differ a bit from the nominal value.

RStudio is used to generate the data, see Appendix B for the script. According to those interviewed, all variables are positively related to temperature, only the pulse of drill lubricate is negatively related to the temperature of the spindle.

Table 2: Characteristics of variables. *This table provides an overview of the variables that are related to a malfunction of the drill unit on the S-axis. Each variable is presented with its description, minimal, maximal, and nominal values. In case "NA" is displayed, it denotes that the value is not applicable. This situation arises when a variable lacks a nominal value. Additionally, the maximum value of pulse drill lubricate is the value "NA" given. This is because the nominal value is a standard setting, and variations from this value are limited.*

Variable	**Description**	**Min**	**Max**	**Nom**
Temperature	Degrees of Celsius (°C)	-10	100	NA
RPM	Revolutions Per Minute	0	4,500	NA
Spindle run time	Active operation duration (days)	0	2,190	NA
Productivity of running time	Percentage of productivity	0	19	2.34
Force	Force (kN)	0	15	13
Pulse drill lubricate	Lubricate amount (ml per minute)	0	NA	0.75
Ambient temperature	Outdoor temperature (in degrees Celsius (°C))	-10	55	18
Type of lubricate drill	Prescribed grease type (0: yes, 1: no)	0	1	0
Maintenance	Time since last maintenance (days)	0	365	NA

4.2.3 Data preparation

The subsequent stage is the data preparation phase, wherein data is chosen, cleansed, generated, integrated, and formatted to obtain actionable data (Schnell et al., 2019). In this research, there will be make use of synthetic data due to the unavailability of the machine data. An assumption in traditional machine learning methodologies is to split the data up into training data and test data (Weiss et al., 2016). In this study, the train data contains 80% and the test data contains 20% of the data.

The values of the independent variables are generated using minimum and maximum values. A random formula was formulated to correlate the independent variables with the temperature variable. This was done because otherwise, the synthetic predictive models have no power. An error variable has also been added, of which the mean is 0 and the distribution of this variable is normal.

4.2.4 Modeling

To simulate the temperature of the drill unit on the S-axis, several predictive modeling techniques are used. In this thesis, linear regression, support vector regression (SVR), decision tree, random forest, and a gradient boosting model are used to predict the temperature. These models are elaborated in this subsection of the thesis. Afterward, the models are tested using test data. In the evaluation phase, each model's predicted values are plotted against the actual values to identify outliers. Additionally, it is examined how to determine which independent variable(s) significantly influences the dependent variable.

Linear regression
Linear regression is a statistical method that plots the dependent variable against several independent variables (Casson & Farmer, 2014). Linear regression has the assumption that the relationship between the variables is linear. The model produces an optimal line that minimizes the disparity between the observed and predicted values (Casson & Farmer, 2014). The formula below was used to predict the temperature. The formula presupposes a linear correlation between the independent variables and the dependent variable. In the case of a non-linear relationship, the model will have lower predictive accuracy (Sohil et al., 2022).

$$\text{Temperature} = \beta_0 + \beta_1\text{RPM} + \beta_2\text{Spindle Run Time} + \beta_3\text{Productivity of Running Time} + \beta_4\text{Force} + \beta_5\text{Pulse Lubrication Drill} + \beta_6\text{Ambient Temperature} + \beta_7\text{Type of Lubricant} + \beta_8\text{Maintenance} + \epsilon$$

Support vector regression
The model takes into account the relationships of the independent variables among each other. Support vector machine (SVM) creates the line that separates a set of entities on essential characteristics (Wijnhoven, 2023). Support vector regression is a highly effective method for solving non-linear problems, SVM applies the kernel function, which maps the input space into more dimensional features using non-linear mapping (Ahmad et al., 2018). SVM focuses on solving classification tasks, while SVR is designed to address regression problems. A difference between linear regression and SVR is that SVR is less sensitive to outliers because it takes into account a margin of error.

Decision tree
A decision tree is a machine-learning technique in which a tree is constructed by utilizing splitting criteria with nodes. The model illustrates how particular decisions and their consequences unfold through specific sequences. The root of the tree is situated at the top, where the original data is located. The tree grows to the first node, where the most informative variable is identified. The informativeness of a variable is determined by the Gini index, a value indicating the strength of regression or classification variables. The algorithm iteratively repeats this process with additional attributes in the model (Gokgoz & Subasi, 2015).

Random forest
Random forest is a machine learning method that employs numerous decision trees as an ensemble in training and merges them to get a prediction (Lee et al., 2019). Random Forest makes use of bootstrap aggregating or bagging, a technique where multiple subsets of the original data are created by random sampling with replacement (Wijnhoven, 2023). A random forest creates multiple trees to determine the most popular class (Breiman, 2001). In the construction of individual trees, random forest incorporates feature randomization by considering a random subset of features at each node for making the split, thereby enhancing the diversity among the trees.

Gradient boosting
Gradient boosting is a method for estimating functions that are approached from a numerical optimization perspective within function space, as opposed to parameter space (Friedman, 2001). This implies that the model examines the overall shape of the function and enhances it, rather than concentrating on adjusting specific parameters. By minimizing errors in predictions, the function becomes more adept at recognizing patterns in the data. Through successive iterations, this method improves its performance. Key characteristics of gradient boosting are its competitiveness, high robustness, and interpretability when applied to regression and classification problems.

The above-mentioned quantitative models are compared with each other by making use of the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), and the R-squared. RMSE is a standard quality measure of the predictive ability of regression models and can usually be used to compare the performance of different predictive models (Cao et al., 2015). The MAPE metric quantifies the average absolute percentage difference between predicted and actual values (Wang et al., 2009). The R-squared metric measures the proportion of variance in the predicted variable that can be explained by the independent variables, ranging from 0 (no fit) to 1 (perfect fit) (Windmeijer & Cameron, 1997). See Table 3 below for the scores of the different metrics of each model. The RMSE value indicates the average magnitude of differences between the actual and predicted temperature values. An RMSE value of 1 means that on average, based on the train data, the model is 1 degree Celsius away from the actual temperature in the test data. This means that how closer the RMSE is to 0, how better the model fits. A MAPE value of 5% means that the predictions have an average error of 5% in relation to the actual value. A R-squared value of 0.9 means that 90% of the variance in the dependent variable, for instance, the temperature is explained by the independent variables in the model.

Table 3: Metrics values of the predictive models. *The table presents the Root Mean Square Error (RMSE) scores, the Mean Absolute Percentage Error (MAPE), and the R-squared of the five predictive models: linear regression, support vector regression, decision tree, random forest, and gradient boosting. Linear regression has the best performance based on the metrics.*

Prediction model	RMSE	MAPE	R-squared
Linear regression	1.5°C	4.1%	0.997
Support vector regression	2.0°C	5.1%	0.995
Decision tree	8.9°C	31.3%	0.899
Random forest	2.4°C	7.9%	0.992
Gradient boosting	1.8°C	5.1%	0.996

4.2.5 Evaluation

If we plot the predicted temperature of every model on the actual temperature, we get Figure 7. In the current situation, using synthetic data, it can be seen that a linear model gives the most accurate predictions. However, in a real situation, it may be that this model would be worse compared to the other models. The figure also shows that the decision tree has the lowest correct predictive power. The eight stripes in the plot represent the leaf nodes in the decision tree, it can be seen that the values vary greatly.

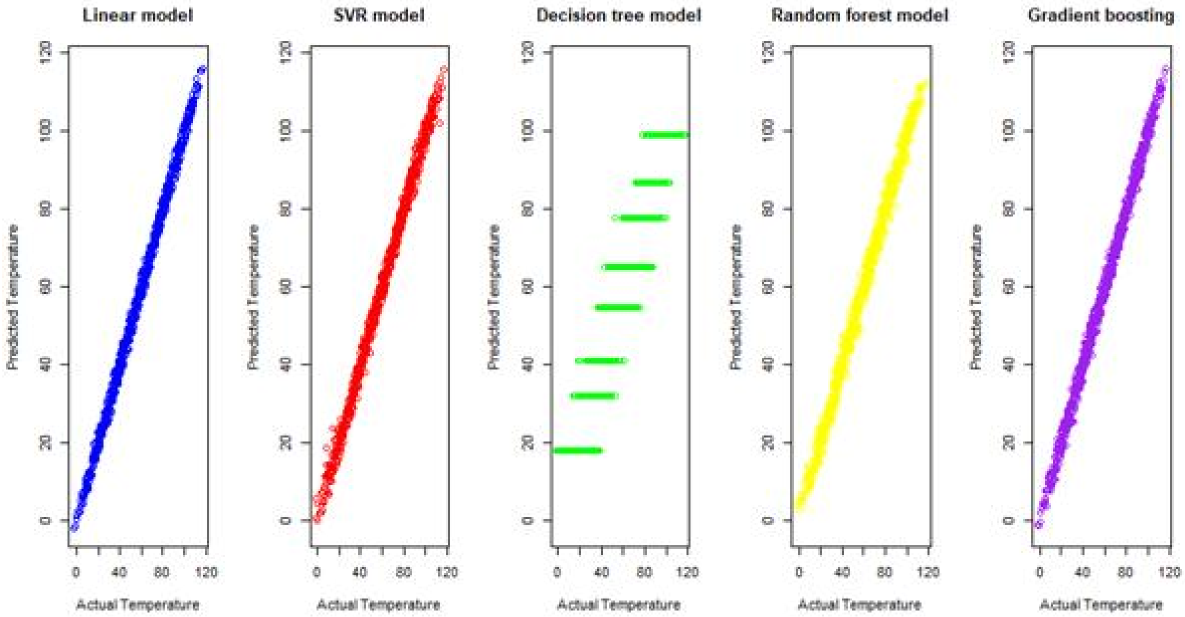

Figure 7: Predicted temperature plotted against actual temperature per model. *The figure contains five graphs that display the performance of linear regression, support vector regression, decision tree, random forest, and gradient boosting. The performance is shown by plotting the predicted temperature against the actual temperature. The linear model shows less noise, which means that this model can better predict the temperature of the test data. The decision tree shows the most variation in the model.*

Because of the best performance by the metrics, the research will continue focusing on the linear model. In this case, to achieve predictive maintenance, it would be useful to know which variable has the most influence on temperature. This can be achieved by standardizing the variables in linear regression (Casson & Farmer, 2014). For this reason, the variables were standardized, and a linear model was created again based on the training data. See Table 4 below for the results of this standardized model.

Table 4: Standardized regression output. *The table presents the results of the standardized linear model. It illustrates the relationship between the independent variables and temperature. This table serves as a reference for comprehending the impact of individual predictors on the temperature in the linear model.*

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
(Intercept)	3.118e-17	8.518e-04	0.000	1.000
RPM	9.368e-01	8.528e-04	1098.529	$< 2e-16$
spindle_run_time	2.755e-01	8.532e-04	322.849	$< 2e-16$
productivity_of_runningtime	1.765e-04	8.528e-04	0.207	0.836
force	7.962e-02	8.529e-04	93.352	$< 2e-16$
pulse_drill_lubricate	3.693e-04	8.521e-04	0.433	0.665
ambient_temperature	1.374e-01	8.525e-04	161.111	$< 2e-16$
lubricate_drill_type	4.433e-03	8.530e-04	5.197	$2.12e-07$
maintenance	3.820e-02	8.525e-04	44.806	$< 2e-16$

From this table, it can be seen that RPM, with a standardized value of 0.94, has the greatest influence on temperature. Spindle run time, with a standardized value of 0.28, is the second largest influence on temperature.

To apply the model to working machines in the field, a critical threshold could be set, for example, 90 degrees Celsius. If a spindle has reached this temperature, a pop-up can be displayed with the values of the variables. Proactive service can be applied by inspecting the values of the variables. For example, if the RPM relative to the variables is high, customer service can inform the customer of these values. If the temperature is high due to the high value of the spindle run time, customer service can reach out to the customer to make an appointment to put a new spindle in the machine which will prevent the machine from malfunctioning due to excessive temperature.

4.3 Study 3: Deployment

This section focuses on the deployment of the predictive model for the company. The subquestion that is answered in this section is: How can the company utilize predictive models to enhance SDL through proactive customer service? This section highlights the importance of acquiring the necessary data, validating the data model, and demonstrating, by simulation, the use of the predictive modeling for proactive customer service, exploring the potential impact of proactive customer service, and reviewing how SDL is achieved by the company through proactive customer service.

4.3.1 Data acquisition

Mason & Mitroff distinguish five inquiry systems; empirical inquiry systems, rational inquiry systems, Kantian inquiry systems, Hegelian inquiry systems, and pragmatic inquiry systems (Mason & Mitroff, 1973). These systems help with executing smart industry projects. To implement a working predictive model for the company, the first thing needed is to develop empirical inquiry systems further. Through these systems, it is available to collect, store, and analyze data for problem identification (Mason & Mitroff, 1973). Currently, not all relevant variables are visible in PROGRAM X, the control system of the CNC machines at the company. Berndtsson et al. (2020) mention that the difficulty accessing relevant data is a barrier for an organization to become data-driven (Berndtsson et al., 2020).

Currently, the company cannot implement proactive customer service due to the reason that not all relevant variables, to create predictive models, are captured by sensors. PROGRAM X is already linked with the cloud, but the machine lacks sensors, which are needed to capture the values of relevant variables for creating predictive models. In the current situation, the cloud is used to provide dashboards to customers wherein production logs are displayed. By capturing data of all relevant variables by sensors, it would be possible to create predictive models based on real live data. There is a need to install sensors for collecting relevant machine data and further develop PROGRAM X to store additional data in the cloud for use in predictive modeling to anticipate machine malfunctions.

4.3.2 Validation data model

If the data is available to predict spindle temperature, it is necessary that the results of the interviews are validated. In this thesis, synthetic data was used in which the dependent variable was determined by the independent variables with a small error rate included. If the data is accessible in the cloud, there can be seen if there is a linear relationship between the variables. If there is, a predictive model can be built using linear regression. SVR could be used when the relationships are not linear and it can handle high-dimensional data. SVR takes into account non-linear relationships using the kernel function (Ahmad et al., 2018). The model is suitable for managing a margin of error. Decision tree models provide a transparent and interpretable representation of the decision-making process. This is done by the hierarchical structure guided by the splitting criteria (Gokgoz & Subasi, 2015). Random forest is effective when there are small variations and noise in the training data. Overfitting, which focuses too much on details and noise in training data, is reduced by using multiple decision trees in random forest (Lee et al., 2019). Gradient boosting operates in function space instead of adjusting specific parameters individually, which fosters better pattern recognition (Friedman, 2001).

4.3.3 Utilizing predictive models for proactive customer service

By collecting machine data and using certain knowledge, predictive maintenance can be achieved. To provide proactive service, a proactive maintenance strategy consisting of predictive maintenance and preventive maintenance is needed. In addition to having proactive customer service, high reliability is needed within the organization. This is necessary because when a model issues a notification to monitor the machine, the company should collaborate with the customer, the customer service employee must be motivated, have sufficient knowledge, and be supported by the organization to do something with this notification.

This section illustrates how organizations can leverage a predictive model. In this case, there will be make use of the developed linear model. This model has shown that RPM and spindle run time are the most influential factors affecting temperature. To demonstrate the use of such a model, a dataset is created representing a fictional machine located at a customer's site. The script that created this dataset can be found in Appendix C. All independent variables remain at their nominal value, except for RPM and spindle run time, which increase incrementally. This will result in an increase in the temperature of the spindle, which can be seen in Figure 8. The predicted temperature by the model can be seen as a digital twin, a replication of the genuine product, process, or system (Singh et al., 2021). It can be used to simulate and analyze the behaviour of a real-world counterpart. The linear model is a tool that predicts and understands the actual temperature of the spindle. To test and develop the simulation, the model should be initially evaluated in a virtual environment before being implemented in real-world situations (Christiano et al., 2016).

For practical reasons, a critical threshold is set at 90 degrees Celsius. The ideal threshold needs to be determined by further testing. It is crucial to establish a critical threshold wherein the customer service can timely intervene to prevent a failure. By collecting, storing, and displaying live machine data in a dashboard, such as Power BI or Salesforce, the linear model can predict the temperature. If the predicted value exceeds the critical threshold, a warning value is triggered, drawing the attention of the customer service department. The customer service department can contact the customer to reduce the temperature below the critical threshold, if necessary by making a service appointment with the customer.

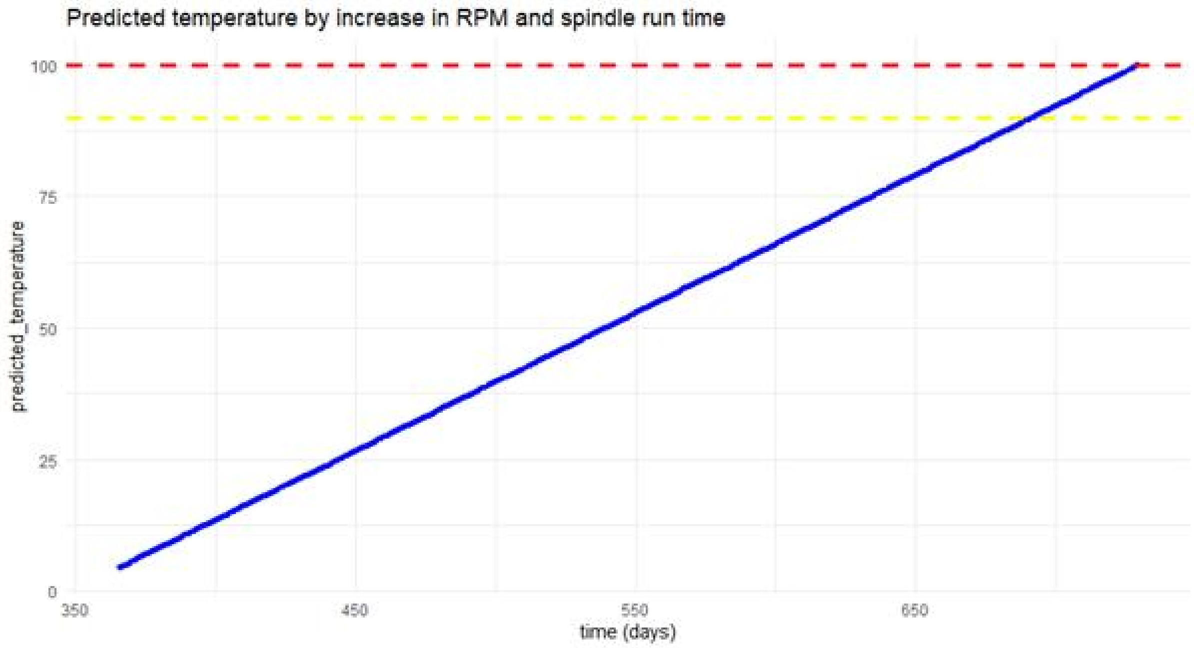

Figure 8: Simulation of spindle temperature. *This fictive figure shows how the temperature (blue) reacts to an increase in RPM and spindle run time. It can be seen that the spindle hits the critical threshold (yellow dashed) on day 691 and hits a breakdown (red dashed) on day 730.*

The linear model could be applied in a real-world context. The figure, displayed above, shows temperature changes by an increase in the RPM and spindle run time while all other variables remain constant. The spindle runs at day 365, which means that the spindle runs one year in the machine. This value is increasing daily with one. RPM starts at zero and increases daily with 12, this will eventually reach the maximum speed of 4500. The model predicts that the critical threshold of 90 degrees Celsius will be reached at day 691 and 100 degrees Celsius, which results in breakdown, will happen at day 730. Warnings can pop up in a dashboard which shows when the critical threshold is reached. An employee in customer service can identify, by using the most influential variables, the primary cause of critical temperature. Additionally, the employee could arrange a service appointment by collaborating with the customer if the problem cannot be solved remotely, to prevent the malfunction from occurring.

In the company, the data needed for the predictive model is not available through PROGRAM X, a self-developed program of the company used to operate the CNC machine. This program does not contain all relevant data to predict malfunctions. If it did, this data could be stored in the cloud to generate predictive models, for example, a model predicting spindle temperature. Dashboards such as PowerBI or Salesforce can display live data, on which a predictive model can provide insights. When the value exceeds a threshold, it can be reported to customer service, which can take appropriate actions by collaborating with the customer. This process is used for customers when the likelihood of a malfunction is realistic. See Figure 9 for an illustration of how this could look for the company.

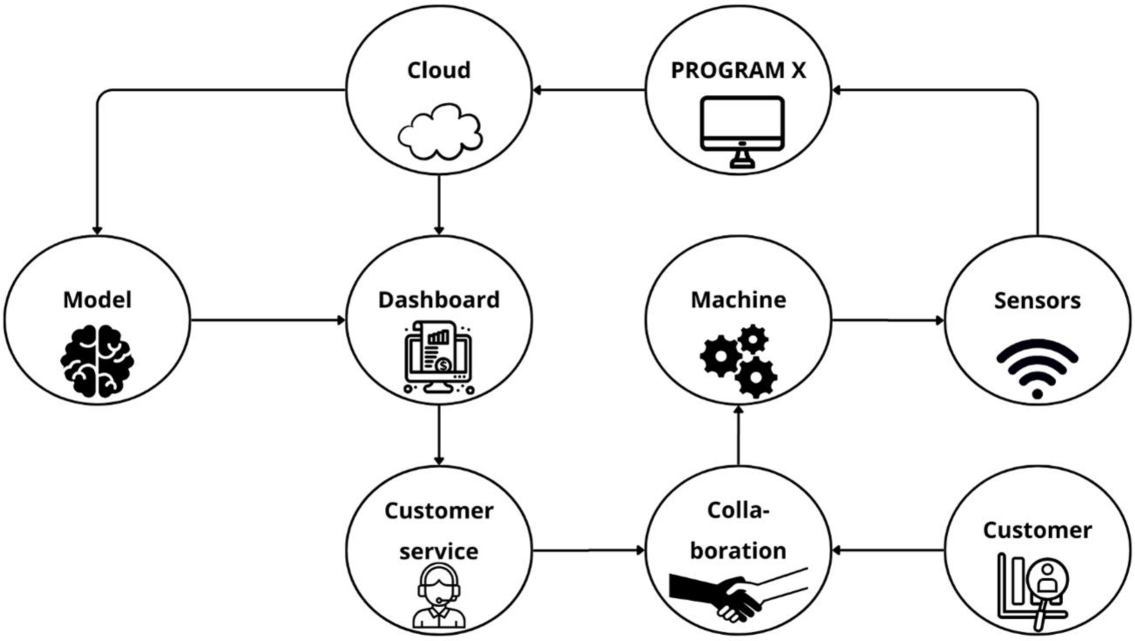

Figure 9: Data acquisition plan to deliver proactive customer service. *This figure is self-made by the author. In this figure, the machine generates data, which is captured by sensors and displayed in PROGRAM X. This data can then be stored in the cloud. Once the data is accessible in the cloud, a predictive model can be developed. This predictive model can utilize live machine data from the cloud to predict variables such as the temperature of the spindle. The data can be visualized in a dashboard, which triggers a notification to the customer service department when the temperature reaches a critical threshold. Subsequently, by collaboration between the customer service department and customers, malfunctions can be prevented.*

Seidel et al. (2019) introduce the concept of triple-loop learning. To continually improve the model, triple-loop learning should be employed. Triple-loop learning enables a deeper understanding of the problem domain and promotes adaptability in modeling strategies. The iterative approach of triple-loop learning contributes to the development of more robust models, which are capable of capturing complex data relationships in dynamic environments (Seidel et al., 2019). Triple-loop learning offers a framework for the company to continue improvement and adaptation in the context of predicting the temperature of the spindle. With triple-loop learning the problem domain is more understandable. The company can get more insights into the underlying factors that influence the temperature of the spindle and adjust its approach accordingly. Additionally, when the company uses historical data to refine the model over time, the company may encounter challenges or opportunities. Because of this phenomenon, the model will remain effective and relevant in dynamic environments. Because of that triple-loop learning is an iterative process, the model becomes more robust. The predictive model is adapted to new data and insights and because of this, it becomes more capable of understanding complex data relationships and can more accurately calculate spindle temperature. This iterative process contributes that the model will predict reliably and more effectively in real-world scenarios. See Figure 10 below for the triple-loop model for the company.

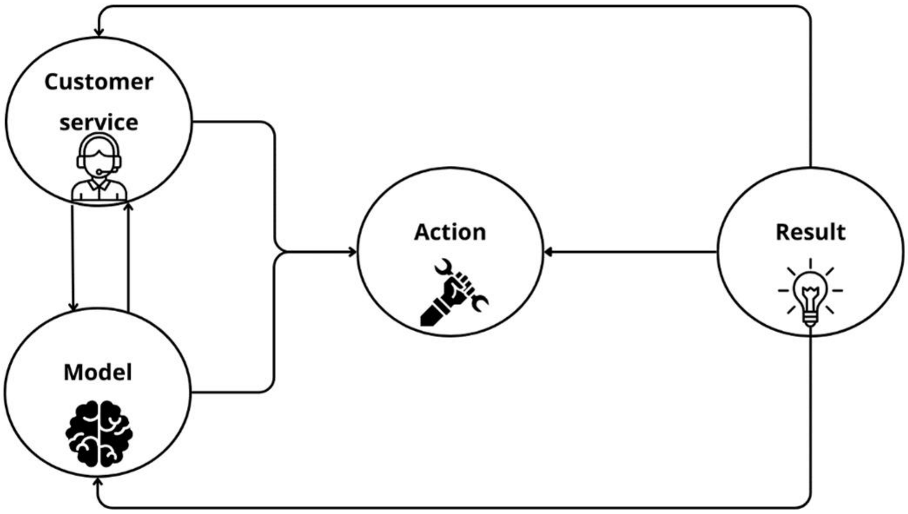

Figure 10: Triple-loop learning model adapted from Seidel et al. (2019). *Triple-loop learning model adapted from Seidel et al. (2019). This figure shows a framework for the company for continuous improvement and adaptation in predictive models. With triple-loop learning, the company gains a deeper understanding of the problem domain and can adjust its approach. Adjusting the predictive model over time, enabled by this iterative process, ensures the model's effectiveness in dynamic environments. The model becomes more capable of understanding complex relationships and can therefore predict more accurately.*

This model can work by interaction between the company with the customer. The customer receives the result and provides feedback to the company and the predictive model. To achieve SDL, interacting with the customer is important to generate value by co-creation (Vargo and Lusch, 2008). Using this feedback, the company gets a more comprehensive view of what the customer needs. Also, based on the feedback, the predictive model can be adjusted, i.e. including more relevant variables. The feedback allows the customer service department and the predictive model to provide a more targeted service in the future which is more in line with customer needs.

4.3.4 Potential impact proactive customer service

Based on the company's documentation there are currently 104 MACHINE X's in use by customers worldwide. Throughout the year 2023, over 20 drill units were either repaired or replaced. The extent of production downtime caused by drill unit failures for the company's customers remains unclear. In 2023, the company received 518 reports from customers indicating that MACHINE X could not sustain normal production. In this case, normal production means that the machine is functioning properly as the customer would desire. The cumulative downtime for all customers due to MACHINE X issues in 2023 amounted to 12,403 days. This implies that, on average, for each customer report, it takes 24 days from fault notification to resolution for a return to normal production. By combining machine data and knowledge, such as predictive models, predictive maintenance can be enabled, which in turn can enable proactive service when coupled with high reliability (Zonta et al., 2020; Swanson, 2001). If all failures of MACHINE X can be predicted, it may result in an increase in the normal production time of the machines.

4.3.5 Proactive customer service within the SDL perspective

To implement proactive customer service within the SDL perspective, the company shall take several steps to advance the service delivery in alignment with the FPs of SDL as outlined by Vargo and Lusch (2008).

FP1: Service is the fundamental basis of exchange.
By monitoring machine data, proactive service could be delivered to the customer in case the temperature has reached the critical threshold, i.e. spindle temperature. The service is anticipated to the customers' needs and give valuable assistance to the customer so the machine will not get to a standstill. To achieve proactive service, the company needs to retrieve all relevant variables for predictive modeling by adding sensors to their machines. These sensors should display the data in PROGRAM X and PROGRAM X should store this data in the cloud. Predictive models can be made by this cloud data and added to dashboards, that are used by the customer service of the company. In case a machine has reached a critical threshold, the company and the customer collaborate to adjust the machine to prevent downtime. If the sensors collect all relevant data, the company has the ability to deliver value through knowledge and skills.

FP4: Operant resources are the fundamental source of competitive advantage.
By the use of predictive maintenance, proactive service leverages the operant resources of the organization to gain a competitive advantage. The efficient use of knowledge, i.e. the predictive models, improves competitive advantage because of machine malfunction prevention. Through triple-loop learning continuous improvement and adaption of the predictive models can be achieved. This continuously improves the company's knowledge and skills. As these operant resources are improved, the competitive advantage will also become stronger.

FP6: The customer is always a co-creator of value.
Acting proactively toward the customer is needed to collaboratively communicate with customers and by customers to address potential solutions before issues can arise. Customers participate in the creation of value by providing input to ensure the optimal functioning of their machine, i.e. making machine data available for analysis by the company. Collaboration between the customer and the company can be realized through interaction when a critical threshold is reached. After this, the company reports this to the customer and the customer provides more information regarding the problem, i.e. the customer indicates that they produce a product containing a hard type of steel resulting in a higher RPM need of the machine. If the customer does not give this information, the value of the service provided by the company will be lower than if the information is available. Together, the organizations can look at a suitable solution to prevent machine downtime.

FP7: The enterprise cannot deliver value, but only value propositions.
Proactive service aligns with the SDL perspective that the organization only can deliver value propositions. The company can engage customers proactively with the use of prediction models, suggesting adjusting values of the variables to prevent machine downtime. This is achieved through the use of dashboards, i.e. Salesforce or PowerBI. When a critical threshold is reached, the dashboard gives a trigger warning to customer service. After this, the customer service employee, can view the predictive model and select the most relevant independent variables that explain the dependent variable. For example, if the spindle run time contains a high value compared with the other variables, the company can report this to the customer to make an appointment to replace this spindle to avoid machine failure. It is up to the customer whether they accept this service.

FP8: A service-centered view is inherently customer-oriented and relational.
Proactive service is customer-oriented and relational. By actively contacting customers in case a critical threshold is reached, solutions can be provided. With proactive service, the company embraces a service-centered view that prioritizes customer relationships. This service-centered approach prioritizes customer relationships to address potential solutions before issues arise. Leveraging the customer as a co-creator of value is done by collaboration. Collaboration takes place between the customer and the company as the dashboards give a trigger warning when the machine has reached a critical threshold. The company then contacts the customer, after which the customer provides more information, to find a suitable solution. These interactions build a stronger relationship between the company and its customers.

FP10: Value is always uniquely and phenomenologically determined by the beneficiary.
Proactive service acknowledges that the perception of value differs among customers. Therefore solutions to meet the different customer needs differ among the company's customers. By acting proactively, customization of the service can take place for each customer. This can be done by for example adjusting risk tolerance and specific operational requirements. By customizing the trigger warning in the dashboards for each customer specific, proactive service can be delivered tailored to the customer's needs. This approach highlights the value-oriented perspective of proactive service in SDL, which ultimately could enhance customer patronage behaviour.

5 Discussion

This section provides an answer to the research question and illustrates the theoretical and practical implications of the research. Subsequently, the limitations of this study and suggestions for further research are discussed. Finally, the thesis ends with the conclusion.

5.1 Main findings

The purpose of this thesis is to show how proactive customer service can be realized using predictive models that are based on machine data and certain knowlegde. The corresponding research question of this thesis is: "*How can machine data be used to enhance the realization of service-dominant logic?*".

The first sub-question that was studied is: "Which machine data is suitable for utilizing predictive models?". This question was answered by conducting interviews with experts at the company. The interviews revealed that the drill unit, measuring wheel, EtherCAT components, hydraulic unit, sensors, and drive wheels led to failures concerning MACHINE X. The drill unit, consisting of the S-axis, Z-axis, and Y-axis, fails due to excessive temperature on the S-axis. Also, this component fails due to excessive torque on the Z-axis and S-axis, and lack of maintenance leads to malfunctions of the Y-axis. The hydraulic unit fails due to excessive oil temperature. The sensors of MACHINE X fail due to ambient environmental factors. Finally, EtherCAT failures are caused by excessive temperature in the cabin where these components are located. The causes of the failures, concerning the drive wheels and measuring wheels, are unclear to the interviewees.

Variables that influence the torque and temperature of the S-axis are the RPM, the time that the spindle runs, the productivity of the spindle, the amount of force applied, the lubrication pulse of the drill, ambient temperature, type of used lubricate, and a lack of maintenance. Wherein the lubrication pulse of the drill is negatively associated with the temperature and torque of the S-axis. The same variables influence the torque of the drill unit on the Z-axis, only the variable RPM is left out. Variables that influence the temperature of the hydraulic unit are the pressure the unit needs to apply, the pumping frequency, the oil capacity of the tank, and the ambient temperature.

The second sub-question that was studied is: "Which predictive models are suitable for predicting machine malfunctions?". If the relevant variables of malfunctions are clear, data can be retrieved by sensors. In this data, it is necessary to test whether the independent variables influence the values of the dependent variable. When a predictive model has been created, such as linear regression, support vector regression, decision tree, random forest, or gradient boosting, it can be examined which independent variables hold the greatest influence. When a dependent value has a high value, it can be determined what causes the issue. Customer service employees could use this for problem identification and proactively approach the customer. In the synthetic data, a linear regression model can make better predictions than the other developed models. In reality, there may be a non-linear relationship between the variables. In such a case, it is appropriate to switch to another model with higher predictive power. This can be determined more precisely when actual data becomes available.

The third sub-question is: "How can the company utilize predictive models to enhance SDL through proactive customer service?". This research demonstrates how proactive customer service can be achieved through predictive maintenance. In order to enable predictive maintenance, empirical inquiry systems are needed to collect, store, and analyze data for problem identification (Mason & Mitroff, 1973). The company should deploy sensors to capture relevant data of variables to use predictive modeling to enhance SDL. In case when the data is available from the sensors, the company should deploy the predictive model with the best accuracy to predict malfunctions. To utilize the predictive model, a critical threshold needs to be determined. When the value exceeds this threshold, the customer service employee needs to collaborate with the customer to prevent machine downtime. By employing triple-loop learning, the model's effectiveness will be ensured in dynamic environments. Resulting that the model stays relevant as time continues.

In the company's case, when all malfunctions regarding MACHINE X can be predicted, it could potentially lead to a reduction of 12,403 days of non-normal production time for customers who own this machine based on the data from 2023. Proactive customer service enhances SDL by recognizing service as the basis for exchange, utilizing operant resources, recognizing that customers need to actively collaborate in the value-creation process, delivering value propositions, prioritizing customer relationships, and recognizing the subjective perception of value.

5.2 Theoretical implications

This thesis contributes to the literature of SDL highlighting the importance of selectively utilizing machine data to predict machine failures and deliver proactive customer service, by using customer-generated real-time data in predictive models. This research builds upon the research of Vargo and Lusch (2017) proclaiming that big data can allow real-time data capture of customer behaviour using sensor-based content and employing data analytics techniques (Vargo & Lusch, 2017).

Additionally, this thesis contributes that proactive customer service affects service recovery. The ability to anticipate and address potential malfunctions is contributing to service failure recovery. By employing predictive models, organizations can identify malfunctions proactively, minimizing negative consequences for their customers. This thesis highlights the impact of proactive customer service in minimizing service recovery and so fostering better customer patronage behaviour.

Finally, this thesis contributes to the theoretical perspective that SDL can be enhanced through proactive customer service. Organizations that are looking to move from reactive service to proactive service to their customers are shifting from a goods-dominant logic to an SDL perspective. The thesis applies proactive customer service, by having the ability to predict malfunctions. This advances several foundational premises of SDL. This section elaborates on how proactive customer service enhances several foundational premises of the SDL framework (Vargo, 2008). If organizations act proactively to provide service when needed, the organization sees providing service as the primary value, which aligns with FP1 that service is the fundamental basis of exchange. Additionally, proactive service leverages operant resources to address customers' needs by utilizing machine data and knowledge. The utilization of data and knowledge contributes to SDL by delivering additional services when necessary to reduce customers' machine downtime. This is contributing to FP4 by gaining a competitive advantage by effectively utilizing operant resources. Further, proactive customer service has to deal with engaging customers to address potential issues. This contributes to FP6 because customers need to actively collaborate in the value-creation process to co-create value with the organization. Additionally, the company can only deliver value propositions. However, in case the company uses predictive models they can deliver more targeted service to serve customer needs, which is contributing to FP7. Furthermore, with proactive service, relationships between organizations and their customers will be built and customer needs will be addressed. This aligns with the service-centered view explained in FP8. Finally, proactive service recognizes that the perception of value is subjective and can be interpreted differently among customers. By distinguishing values of critical thresholds per customer, as the customer may desire, organizations can provide client-specific customer service. By addressing individual needs proactively, organizations acknowledge the subjective concept of value. Therefore, proactive customer service is contributing to FP10. See Table 5 below for a clear overview of the current state of SDL and the proposed changes.

Table 5: The current state of SDL and proposed changes. *The table presents several aspects of SDL with their current state and proposed changes to enhance SDL.*

Aspect of Service-dominant logic (SDL)	**Current state**	**Proposed change**
Foundation of exchange	Service is the fundamental basis of exchange.	Provide proactive service based on customer-generated real-time data to address customer needs in time.
Utilization of resources	Operant resources are a source of competitive advantage.	Utilize machine data and knowledge to reduce customer downtime and gain competitive advantages.
Customer collaboration	Customers need to actively collaborate in the value-creation process.	Contact customers to address potential issues when critical thresholds are exceeded.
Deliver value propositions	Organizations can only deliver value propositions.	Use predictive models to provide targeted value propositions based on customer needs.
Customer relationships	Building relationships with customers through service delivery.	Provide proactive customer service to build stronger relationships.
Perception of value	Recognizing the subjective perception of value among customers.	Address individual customer needs proactively and customize critical thresholds based on customer needs.

5.3 Practical implications

This research contributes to the practical implication that customer-generated data can be used to provide proactive customer support. Rust and Huang (2014) concluded that improvements in Information Technology (IT) lead to a better organizational ability to establish stronger relationships with its customers. This trend will become more pronounced as time goes on. This thesis can confirm this as improvements are being made in terms of storing data and analyzing IT in order to provide proactive service and build stronger relationships developed between the company and its customers (Rust & Huang, 2014). To realize this, organizations need to have the ability to collect, store, and analyze data to use it for projects (Mason & Mitroff, 1973).

This thesis emphasizes the importance of managing big data to streamline product prediction for optimizing product service requirements (Lee, 2014). To retrieve machine data from customers, predictive models could monitor the variables of the machines. If certain variables indicate a possible issue, the predictive model could send a notification to customer service to monitor the machine. If the customer service employee collaborates with the customer, is correctly motivated, possesses sufficient knowledge, and is supported by the organizational culture, the employee could successfully deliver service to the customer to prevent malfunctions.

Additionally, this thesis contains the practical contribution to select relevant variables for predicting machine malfunctions in order to deploy sensors to retrieve relevant data. To implement proactive customer service and therefore embrace the SDL perspective, there is a need for machine data to create predictive models. Therefore, sensors should be placed and PROGRAM X should be modified in order to display all relevant data. Afterward, the company should redo the analysis to find the best predictive model on real data.

5.4 Limitations and further research

A limitation of this thesis is that it is not possible to confirm whether the indicated variables truly affect the failures. Qualitative data on why several components cause failures is collected using interviews in Study 1. The research is intended to focus on the failures that have the most impact on the downtime of customers. The interviews were conducted with experts of the company because it is unclear which failures generate the most downtime of MACHINE X. When the data of the variables is available, it is possible to find and test the relations between the variables.

Another limitation of this thesis is the use of synthetic data to create predictive models. In Study 2 and Study 3 is data required to create predictive models or create a simulation. In those studies, it was necessary to use synthetic data since real data was unavailable due to the lack of the capability of PROGRAM X to collect and store necessary data to predict malfunctions. To create synthetic data, minimum, maximum, and nominal values were estimated based on knowledge that was available within the organization. It may be that the synthetic data differs from real data since it could be that real relationships are not linear. Therefore, it may be that the created models may respond differently to real data. In future studies, the models need to be reevaluated and the model with the best accuracy should be employed.

Finally, the thesis made use of a case study, limiting the generalizability of the findings to the broader population. The applicability of the results beyond the specific context of the company may be constrained. The findings may not directly translate to other contexts. However, the underlying principles and methodology provide insights for organizations facing similar challenges.

Future research could explore how data can enhance the other foundational premises (FP2, FP3, FP5, FP9) of SDL. Also, future research could entail how needed data can be identified and stored to perform analyzes. If there is no sufficient data storage in the cloud of the smart products, the status of these products can not be monitored and the data cannot be used in predictive modeling. Consequently, proactive customer service cannot be implemented. The company should execute further research on which variables are relevant to predict malfunctions and how this data could be captured by sensors, and displayed in PROGRAM X to enhance SDL.

5.5 Conclusion

In this section, the main research question: "How can machine data be used to enhance the realization of service-dominant logic?", is addressed through three sub-questions. The first sub-question focuses on which machine data is suitable for utilizing predictive models. Interviews have shown key machine components that are leading to machine failures, emphasizing the variables temperature and torque. This insight provides an understanding of the relevant data for predictive modeling. The second study focuses on which predictive models are suitable for predicting machine malfunctions. This study explored relevant predictive methods, i.e. linear regression in order to predict malfunctions. This showed that it is important to select the model based on data characteristics and real-time dynamics. The last sub-question addresses the question of how the company can utilize predictive models to enhance SDL through proactive customer service. This thesis demonstrates how proactive customer service is enabled by predictive maintenance and enhances SDL. By employing sensors, collecting relevant data, and using predictive models, the company can predict malfunctions, determine critical thresholds, and collaborate with the customer to prevent malfunctions. Finally, the iterative process of triple-loop learning guarantees ongoing refinement of the predictive model and alignment with SDL.

References

Agarwal, R., & Selen, W. (2009). Dynamic Capability Building in Service Value Networks for Achieving Service Innovation. *Decision Sciences, 40*(3), 431-475. https://doi.org/10.1111/j.1540-5915.2009.00236.x

Ahmad, M., Mourshed, M., & Rezgui, Y. (2018). Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression. *Energy, 164*, 465-474. https://doi.org/10.1016/j.energy.2018.08.207

Berndtsson, M., Lennerholt, C., Svahn, T., & Larsson, P. (2020). Organizations' Attempts to Become Data-Driven. *International Journal of Business Intelligence Research, 11*(1), 1-21. https://doi.org/ 10.4018/IJBIR.2020010101

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. http://dx.doi.org/10.1023/A:1010933404324

Busby, J., & Iszatt-White, M. (2014). The Relational Aspect to High Reliability Organization. *Journal of Contingencies and Crisis Management, 22*(2), 69-80. https://doi.org/10.1111/1468-5973.12045

Cao, T., Lim, E.-P., Zhou, Z.-H., Ho, T.-B., Cheung, D., & Motoda, H. (2015). *Advances in Knowledge Discovery and Data Mining.* Springer International Publishing. https://doi.org/10.1007/978-3-319-18038-0

Casson, R. J., & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & experimental ophthalmology, 42*(6), 590-596. https://doi.org/10.1111/ceo.12358

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step Data Mining Guide. SPSS. http://www.crisp-dm.org/CRISPWP-0800.pdf

Chaurasia, S. S., Kaul, N., Yadav, B., & Shukla, D. (2020). Open innovation for sustainability through creating shared value-role of knowledge management system, openness and organizational structure. *Journal of Knowledge Management, 24*(10), 2491-2511. https://doi.org/10.1108/JKM-04-2020-0319

Chen, H., Chiang, R., & Storey, V. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165-1188. https://doi.org/10.2307/41703503

Christiano, P., Shah, Z., Mordatch, I., Schneider, J., Blackwell, T., Tobin, J., Abbeel, P., & Zaremba, W. (2016). *Transfer from Simulation to Real World through Learning Deep Inverse Dynamics Model.* https://doi.org/10.48550/arXiv.1610.03518

Crant, M. (2000). Proactive behavior in organizations. *Journal of Management, 26*(3), 435-462. https://doi.org/10.1177/014920630002600304

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*, (3rd ed.) Sage Publications, Inc.

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

Delana, K., Savva, N., & Tezcan, T. (2021). Proactive customer service: Operational benefits and economic frictions. *Manufacturing and Service Operations Management, 23*(1), 70-87. https://doi.org/10.1287/msom.2019.0811

Dey, S., Thommana, J., & Dock, S. (2015). Public agency performance management for improved service delivery in the digital age: Case study. *Journal of Management in Engineering, 31*(5). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000321

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189-1232. https://doi.org/10.1214/aos/1013203451

Gandomi, A., Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gioia, D., Corley, K., & Hamilton, A. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods, 16*(1), 15-31. https://doi.org/10.1177/1094428112452151

Gokgoz, E., & Subasi, A. (2015). Comparison of decision tree algorithms for EMG signal classification using DWT. *Biomedical Signal Processing and Control, 18*, 138-144. https://doi.org/10.1016/j.bspc.2014.12.005

Hart, C., Heskett, J., & Sasser, E. (1990). The profitable art of service recovery. *Harvard Business Review, 68*(4), 148-156.

Johnson, R., Onwuegbuzie, A., & Turner, L. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112-133. https://doi.org/10.1177/1558689806298224

Kashiha, H., & Chan, S. H. (2015). A little bit about: Differences in native and non-native speakers' use of formulaic language. *Australian Journal of Linguistics, 35*(4), 297-310. https://doi.org/10.1080/07268602.2015.1067132

Lee, J., Kao, H.-A., & Yang, S. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. *Procedia CIRP, 16*, 3-8. https://doi.org/10.1016/j.procir.2014.02.001

Lee, T. H., Ullah, A., & Wang, R. (2019). Bootstrap Aggregating and Random Forest. *Macroeconomic Forecasting in the Era of Big Data*. https://doi.org/10.1007/978-3-030-31150-6$_1$3

Lehrer, C., Wieneke, A., Brocke, J. v., Jung, R., & Seidel, S. (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems, 35*(2), 424-460. https://doi.org/10.1080/07421222.2018.1451953

Lusch, R., Vargo, S., & Tanniru, M. (2010). Service, value networks and learning. *Journal of the Academy of Marketing Science, 38*(1), 19-31. https://doi.org/10.1007/s11747-008-0131-z

Lusch, R. F., & Nambisan, S. (2015). Service Innovation: A Service-Dominant Logic Perspective. *MIS Q., 39*, 155-175. https://doi.org/10.25300/MISQ/2015/39.1.07

Mason, R. O., & Mitroff, I. I. (1973). A Program for Research on Management Information Systems. *Management Science, 19*(5), 475-487. http://www.jstor.org/stable/2629445

Maxham, J. (2001). Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. *Journal of Business Research, 54*(1), 11-24. https://doi.org/10.1016/S0148-2963(00)00114-4

Newman, D., Parto, M., Saleeby, K., Kurfess, T., & Dugenske, A. (2019). Development of a Digital Architecture for Distributed CNC Machine Health Monitoring. *Smart and Sustainable Manufacturing Systems, 3*(2). https://doi.org/10.1520/ssms20190038

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. *California Management Review, 61*(4), 156-185. https://doi.org/10.1177/0008125619859318

Parker, S. K., & Collins, C. G. (2010). Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management, 36*(3), 633-662. https://doi.org/10.1177/0149206308321554

Philippe, A., & N'Goala, G. (2010). The differing and mediating roles of trust and relationship commitment in service relationship maintenance and development. *Journal of the Academy of Marketing Science, 38*(3), 303-325. https://doi.org/10.1007/s11747-009-0163-z

Polo Peña, A. I., Frías Jamilena, D. M., & Rodríguez Molina, M. Á. (2014). Value co-creation via information and communications technology. *The Service Industries Journal, 34*(13), 1043-1059. https://doi.org/10.1080/02642069.2014.939641

Power, D. J. (2014). Using Big Data for analytics and decision support. *Journal of Decision Systems, 23*(2), 222-228. https://doi.org/10.1080/12460125.2014.888848

Raub, S., & Liao, H. (2012). Doing the right thing without being told: Joint effects of initiative climate and general self-efficacy on employee proactive customer service performance. *Journal of Applied Psychology, 97*(3), 651-667. https://doi.org/10.1037/a0026736

Rust, R., & Huang, M.-H. (2014). The service revolution and the transformation of marketing science. *Marketing Science, 33*(2), 206-221. https://doi.org/10.1287/mksc.2013.0836

Schein, E. (1990). Organizational culture. *American Psychologist, 45*(2), 109-119. https://doi.org/10.1037/0003-066X.45.2.109

Schnell, J., Nentwich, C., Endres, F., Kollenda, A., Distel, F., Knoche, T., & Reinhart, G. (2019). Data mining in lithium-ion battery cell production. *Journal of Power Sources, 413*, 360-366. https://doi.org/10.1016/j.jpowsour.2018.12.062

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science, 181*, 526-534. https://doi.org/10.1016/j.procs.2021.01.199

Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys, 34*(1), 1-47. https://doi.org/10.1145/505282.505283

Seidel, S., Berente, N., Lindberg, A., Lyytinen, K., & Nickerson, J. V. (2019). Autonomous tools and design: A triple-loop approach to human-machine learning. *Communications of the ACM, 62*(1), 50-57. https://doi.org/10.1145/3210753

Shin, H., Ellinger, A., Mothersbaugh, D., & Reynolds, K. (2017). Employing proactive interaction for service failure prevention to improve customer service experiences. *Journal of Service Theory and Practice, 27*(1), 164-186. https://doi.org/10.1108/JSTP-07-2015-0161

Singh, S., Shehab, E., Higgins, N., Fowler, K., Reynolds, D., Erkoyuncu, J. A., & Gadd, P. (2021). Data management for developing digital twin ontology model. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 235*(14), 2323-2337. https://doi.org/10.1177/0954405420978117

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research, 36*(3), 356-372. https://doi.org/10.1177/002224379903600305

Sohil, F., Sohali, M. U., & Shabbir, J. (2022). An introduction to statistical learning with applications in R. *Statistical Theory and Related Fields, 6*(1), 87. https://doi.org/10.1080/24754269.2021.1980261

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

Tran, T. P., Mai, E. S., & Taylor, E. C. (2021). Enhancing brand equity of branded mobile apps via motivations: A service-dominant logic perspective. *Journal of Business Research, 125*, 239-251. https://doi.org/10.1016/j.jbusres.2020.12.029

Uhlmann, E., Geisert, C., & Hohwieler, E. (2008). Monitoring of slowly progressing deterioration of computer numerical control machine axes. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 222*(10), 1213-1219. https://doi.org/10.1243/09544054JEM1040

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a New Dominant Logic for Marketing. *Journal of Marketing, 68*(1), 1-17. https://doi.org/10.1509/jmkg.68.1.1.24036

Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science, 36*(1), 1-10. https://doi.org/10.1007/s11747-007-0069-6

Vargo, S., & Lusch, R. (2015). Institutions and axioms: an extension and update of service-dominant logic. *Journal of the Academy of Marketing Science, 44*(1), 5-23. https://doi.org/10.1007/s11747-015-0456-3

Vargo, S. L., & Lusch, R. F. (2017). Service-dominant logic 2025. *International Journal of Research in Marketing, 34*(1), 46-67. https://doi.org/10.1016/j.ijresmar.2016.11.001

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly, 37*(1), 21-54. https://doi.org/10.25300/misq/2013/37.1.02

Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology, 374*(3), 294-306. https://doi.org/10.1016/j.jhydrol.2009.06.019

Weick, K. E. (1987). Organizational Culture as a Source of High Reliability. *California Management Review, 29*(2), 112-127. https://doi.org/10.2307/41165243

Weick, K., & Sutcliffe, K. (2007). *Managing the Unexpected Resilient Performance in an Age of Uncertainty* (2nd ed.) John Wiley & Sons.

Weiss, K., Khoshgoftaar, T., & DingDing, W. (2016). A survey of transfer learning. *Journal of Big Data, 3*(1). https://doi.org/10.1186/s40537-016-0043-6

Wijnhoven, A. B. J. M. (2023). *Becoming Smart by Knowing the Future.* University of Twente.

Wijnhoven, A. B. J. M., Ehrenhard, M. L., & Alink, T. (2012). High reliable internet hosting: A case study of handling unanticipated threats to reliability. *International Conference on Information Systems, ICIS 2012, 2*, 1289-1304.

Windmeijer, F., & Cameron, A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics, 77*(2), 329-342. https://doi.org/10.1016/S0304-4076(96)01818-0

Zahra, S., & Nambisan, S. (2012). Entrepreneurship and strategic thinking in business ecosystems. *Business Horizons, 55*(3), 219–229. https://doi.org/10.1016/j.bushor.2011.12.004

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 and Industrial Engineering, 148*. https://doi.org/10.1016/j.cie.2020.106889

Appendices

Appendix A: Interview guide

General information

• Welcome the interview participant.
• The purpose of the interview is to gain insight into the most relevant malfunctions, based on machine components, that occur with MACHINE X, see ethical consent form.
• The answers you provide are strictly confidential and will only be used for this research.

1. Can you briefly describe yourself and your role?
2. What are the main goals and expectations of customers regarding the MACHINE X?

Maintenance and downtime

3. What are your tasks related to MACHINE X?
   a. Where do you spend the most time on these tasks?
4. Are there components/modules that require longer maintenance time than others?
   a. Why do these components have longer maintenance times?
   b. Can you provide specific cases?
5. Which components require more frequent maintenance than others?
   a. Why do these components require more frequent maintenance than others?
   b. Can you provide specific cases?

Frequency of failures

6. What are the most common failures reported by customers regarding MACHINE X?
   a. What is the cause of these failures?
   b. What are the consequences of these failures?
   c. How could these failures have been resolved early before they occurred?
   d. Which operational values and/or environmental factors play a role in these failures?
   e. Is there live machine data available that is relevant to these failures?

Maintenance time for failures

7. What are the failures that require the longest maintenance time for MACHINE X?
   a. What is the cause of these failures?
   b. Do these occur regularly?
   c. What are the consequences of these failures?
   d. How could these failures have been resolved early before they occurred?
   e. Is there live machine data available that is relevant to these failures?
   f. Which operational values and/or environmental factors play a role in these failures?

Impact of failures

8. Which failures of MACHINE X have the greatest impact on downtime?
   a. Where are these failures occurring?
   b. Which modules/components are associated with these failures?
   c. Which operational values and/or environmental factors influence these failures?
   d. Can you elaborate on specific cases?

Operational values/environmental factors

9. Which operational values and/or environmental factors is MACHINE X depended on?
    a. Why is MACHINE X dependent on these values/factors?
    b. Which components and/or modules can cause failures due to these values/factors?
    c. How do these factors/values affect the modules/components?
    d. How could these failures have been prevented?

Literature questions

10. According to literature, the feed axis of CNC machines are the most stressed mechanical parts; how does this apply to MACHINE X?
11. How could machine data and certain knowledge achieve predictive maintenance?

Conclusion

12. Of all the components/problems we have discussed, what do you consider the most important factor leading to a malfunction of MACHINE X that could have been prevented if prediction is possible?
13. What data could be used to predict these failures?
14. Which data is missing to predict failures?
15. Can you list three aspects that could be improved on MACHINE X that are relevant to this research?
    a. Can you further elaborate on these points?
16. Have I forgotten to ask anything about this topic?

The answers you have provided will be treated confidentially.
Thank you for your time.

Appendix B: Coding script of predictive modeling

```r
library(tidyverse)
library(randomForest)
library(randomForestExplainer)
library(brms)
library(tibble)
library(dplyr)
library(e1071)
library(gbm)
library(rpart)
library(randomForestSRC)

# Set seed for reproducibility
set.seed(123)

# Number of observations in the dataset
n <- 5000

# Generate random data within specified ranges
RPM <- sample(seq(0, 4500, 1), n, replace = TRUE)
spindle_run_time <- sample(seq(0, 2190, 1), n, replace = TRUE)
productivity_of_runningtime <- runif(n, 0, 19)
force <- runif(n, 0, 15)
pulse_drill_lubricate <- rnorm(n, mean = 0.75, sd = 0.1)
ambient_temperature <- sample(seq(-10, 55, 1), n, replace = TRUE)
lubricate_drill_type <- sample(0:1, n, replace = TRUE)
maintenance <- sample(seq(0, 365, 1), n, replace = TRUE)

# Create the dataset
dataset <- tibble(
  RPM = RPM,
  spindle_run_time = spindle_run_time,
  productivity_of_runningtime = productivity_of_runningtime,
  force = force,
  pulse_drill_lubricate = pulse_drill_lubricate,
  ambient_temperature = ambient_temperature,
  lubricate_drill_type = lubricate_drill_type,
  maintenance = maintenance
)

# Create variable temperature with coeficients
intercept <- -10
coefs <- c(0.02, 0.012, 0.003, 0.5, 0.1, 0.2, 0.2, 0.01)

# Apply errors
errors <- rnorm(n, mean = 0, sd = 1.5)

# Create associations
temperature <- intercept +
  coefs[1] * RPM +
  coefs[2] * spindle_run_time +
```

```
  coefs[3] * productivity_of_runningtime +
  coefs[4] * force -
  coefs[5] * pulse_drill_lubricate +
  coefs[6] * ambient_temperature +
  coefs[7] * lubricate_drill_type +
  coefs[8] * maintenance + errors

# Add temperature in dataset
dataset <- mutate(dataset, temperature = temperature)
dataset <- mutate(dataset, temperature = pmax(temperature, -10))

# Summary
summary(dataset)

# Split the dataset into training- and testdata
split_index <- round(0.8 * nrow(dataset))
train_data <- dataset[1:split_index, ]
test_data <- dataset[(split_index + 1):nrow(dataset), ]

# Linear model
lm_model <- lm(temperature ~ ., data = train_data)

predictions_linearmodel <- predict(lm_model, newdata = test_data)

# Summary of linear model
summary(lm_model)

# Metrics linear model
performance_linearmodel <- sqrt(mean((test_data$temperature - predictions_linearmodel)^2))
cat("Root Mean Squared Error for Linear Model on test data:", performance_linearmodel, "\n")
r_squaredLM <- summary(lm_model)$r.squared
cat("R-squared for Linear Model:", r_squaredLM, "\n")

# MAPE function
mape <- function(actual, predicted) {
  mean(abs((actual - predicted) / actual)) * 100
  }
mape_linear <- mape(test_data$temperature, predictions_linearmodel)
cat("Mean Absolute Percentage Error (MAPE) for Linear Model:", mape_linear, "%\n")

# Save model
saveRDS(lm_model, "lineair_model.rds")

# Linear model zscore
zscore_dataset <- as.data.frame(scale(train_data))
Zscoremodel_lm <- lm(temperature ~ ., data = zscore_dataset)
summary(Zscoremodel_lm)

# Random forest model
rf_model <- randomForest(temperature ~ ., data = train_data, mtry = 5)
predictions_RF <- predict(rf_model, newdata = test_data)
```

```r
# Metrics random forest
performance_RF <- sqrt(mean((test_data$temperature - predictions_RF)^2))
cat("Root Mean Squared Error for Random Forest Model on test data:", performance_RF, "\n")
r_squared_rf <- cor(predictions_RF, test_data$temperature)^2
cat("R-squared for Random Forest Model:", r_squared_rf, "\n")
mape_rf <- mape(test_data$temperature, predictions_RF)
cat("Mean Absolute Percentage Error (MAPE) for Random Forest Model:", mape_rf, "%\n")

varImpPlot(rf_model)
importance(rf_model)
normalized_importance <- importance(rf_model) / sum(importance(rf_model))

# Support vector regression model
svr_model <- svm(temperature ~ ., data = train_data, kernel = "radial", cost = 3,
epsilon = 0.05)

predictions_SVR <- predict(svr_model, newdata = test_data)

# Metrics support vector regression
performance_SVR <- sqrt(mean((test_data$temperature - predictions_SVR)^2))
cat("Root Mean Squared Error for SVR Model on test data:", performance_SVR, "\n")
r_squared_svr <- cor(predictions_SVR, test_data$temperature)^2
cat("R-squared for SVR Model:", r_squared_svr, "\n")
mape_svr <- mape(test_data$temperature, predictions_SVR)
cat("Mean Absolute Percentage Error (MAPE) for SVR Model:", mape_svr, "%\n")

# Gradient boosting model
formula <- temperature ~ RPM + spindle_run_time + productivity_of_runningtime +
  force + pulse_drill_lubricate + ambient_temperature +
  lubricate_drill_type + maintenance

gbm_model <- gbm(formula, data = train_data, distribution = "gaussian",
n.trees = 1000, interaction.depth = 4)

predictions_GB <- predict(gbm_model, newdata = test_data,
n.trees = 1000)

# Metrics gradient boosting
rmse_GB <- sqrt(mean((predictions_GB - test_data$temperature)^2))
print(paste("Root Mean Squared Error (RMSE):", rmse_GB))
r_squared_GB <- cor(predictions_GB, test_data$temperature)^2
print(paste("R-squared for Gradient Boosting Model:", r_squared_GB))
mape_gb <- mape(test_data$temperature, predictions_GB)
cat("Mean Absolute Percentage Error (MAPE) for Gradient Boosting Model:", mape_gb, "%\n")

# Summary
summary(gbm_model)

# Decision tree model
tree_model <- rpart(
  formula = temperature ~ RPM + spindle_run_time + productivity_of_runningtime + force +
```

```
    pulse_drill_lubricate + ambient_temperature + lubricate_drill_type + maintenance,
  data = train_data,
  maxdepth = 10)

plot(tree_model)
text(tree_model, cex = 0.8)

predictions_tree <- predict(tree_model, test_data)

# Metrics decision tree
rmse_tree <- sqrt(mean((predictions_tree - test_data$temperature)^2))
print(paste("Root Mean Squared Error (Decision Tree):", rmse_tree))
variance_target <- var(test_data$temperature)
r_squared_tree <- 1 - (rmse_tree^2 / variance_target)
print(paste("R-squared for Decision Tree Model:", r_squared_tree))
mape_tree <- mape(test_data$temperature, predictions_tree)
cat("Mean Absolute Percentage Error (MAPE) for Decision Tree Model:", mape_tree, "%\n")

# Visualisation
# Get the range of temperature values
temperature_range <- range(test_data$temperature,
                           predictions_linearmodel,
                           predictions_SVR,
                           predictions_tree,
                           predictions_RF,
                           predictions_GB)

# Set up a 1-row, 5-column layout for the plots
par(mfrow = c(1, 5))

# Scatterplot for Linear model
plot(test_data$temperature, predictions_linearmodel, main = "Linear model",
     xlab = "Actual Temperature", ylab = "Predicted Temperature", col = "blue",
     xlim = temperature_range, ylim = temperature_range)

# Scatterplot for SVR model
plot(test_data$temperature, predictions_SVR, main = "SVR model",
     xlab = "Actual Temperature", ylab = "Predicted Temperature", col = "red",
     xlim = temperature_range, ylim = temperature_range)

# Scatterplot for Decision tree model
plot(test_data$temperature, predictions_tree, main = "Decision tree model",
     xlab = "Actual Temperature", ylab = "Predicted Temperature", col = "green",
     xlim = temperature_range, ylim = temperature_range)

# Scatterplot for Random forest model
plot(test_data$temperature, predictions_RF, main = "Random forest model",
     xlab = "Actual Temperature", ylab = "Predicted Temperature", col = "yellow",
     xlim = temperature_range, ylim = temperature_range)
```

```r
# Scatterplot for Gradient boosting
plot(test_data$temperature, predictions_GB, main = "Gradient boosting",
     xlab = "Actual Temperature", ylab = "Predicted Temperature", col = "purple",
     xlim = temperature_range, ylim = temperature_range)
```

Appendix C: Coding script of simulation linear model

```
library(tidyverse)
library(randomForest)
library(randomForestExplainer)
library(brms)
library(tibble)
library(dplyr)
library(e1071)
library(gbm)
library(rpart)
library(ggplot2)

# Getting model
lm_model <- readRDS("lineair_model.rds")

# Create variables
days <- 0:365
spindle_run_time <- 365:730
RPM <- 12 * (days)
productivity_of_runningtime <- rep(2.34, length(days))
force <- rep(13, length(days))
pulse_drill_lubricate <- rep(0.75, length(days))
ambient_temperature <- rep(18, length(days))
lubricate_drill_type <- rep(0, length(days))
maintenance <- 0:365

# Getting model
lm_model <- readRDS("lineair_model.rds")

# Dataset machine at customer
MACHINE_X_at_customer <- data.frame(
  spindle_run_time = spindle_run_time,
  RPM = RPM,
  productivity_of_runningtime = productivity_of_runningtime,
  force = force,
  pulse_drill_lubricate = pulse_drill_lubricate,
  ambient_temperature = ambient_temperature,
  type_of_lubricate = lubricate_drill_type,
  maintenance = maintenance)

# Predict temperature
MACHINE_X_at_customer$predicted_temperature <- predict(lm_model, newdata =
MACHINE_X_at_customer)
MACHINE_X_at_customer$predicted_temperature <- round(predict(lm_model, newdata =
MACHINE_X_at_customer), 1)

# Create plot
ggplot(MACHINE_X_at_customer, aes(x = spindle_run_time, y = predicted_temperature)) +
  geom_line(color = "blue", linewidth = 1.5) +
  geom_hline(yintercept = 90, linetype = "dashed", color = "yellow",
  linewidth = 1.2) +  # critical threshold Y = 80
```

```
geom_hline(yintercept = 100, linetype = "dashed", color = "red",
linewidth = 1.2) +  # malfunction Y = 100

labs(title = "Predicted temperature by increase in RPM and spindle run time",
     x = "time (days)",
     y = "predicted_temperature") +
theme_minimal()
```