Tampereen yliopisto

Miska Jurva

MACHINE LEARNING UTILIZATION IN SALES AND OPERATIONS PLANNING PROCESS

Master of Science Thesis
Faculty of Management and Business
Examiner: Professor Teemu Laine
Examiner: Associate Professor (tenure track) Tuomas Korhonen
March 2022

ABSTRACT

Miska Jurva: Machine learning utilization in sales and operations planning process
Master of Science Thesis
Tampere university
Master's Degree Program in Industrial Engineering and Management
March 2022

Sales and operations planning process has been a part of organizations' everyday life already tens of years. Even though there are a lot of research in this field, organizations are still struggling to get enough results and reach enough efficiency in their planning process. Nowadays, different kinds of intelligent solutions have become more and more in use. Machine learning is one of those solutions, which can help organizations to reach better results in their operations, also in sales and operations planning. The problem is often to find suitable phase or problem where to utilize machine learning in a proper way. Machine learning has become a popular research topic, and in general, artificial intelligence is very interesting topic with a lot of possibilities nowadays.

This thesis focuses on the machine learning possibilities in the field of sales and operations planning within a case organization. The goal was to focus on the machine learning utilization within sales and operations planning, including the possible value and challenges of it. The study was conducted with the help of relevant research literature and an interview round with the key persons of the case organization. Those persons had a significant role in sales and operations planning process in different business areas. Literature review introduces the theoretical background of the research topic including sales and operations planning, machine learning and how they can be connected. Empirical part covers the sales and operations planning process of the case organization and focuses on the machine learning utilization possibilities, possible value and the most important factors that should be taken into account in machine learning utilization.

Main findings focused on the actual utilization of ML within organization's sales and operations planning process, the value that can be reached by utilizing machine learning and also the challenges around the topic. In demand planning of a business area could be implemented a machine learning solution by taking into account some new external data sources and so get more detailed and fact-based view into demand. In general there were several different value factors found in the research. Most of them were connected to the calculation power of machine learning and its ability to handle data in more intelligent way than humans. The biggest challenges seemed to be related to handling qualitative and experience-based data, that the organization include pretty much, especially regarding sales and operations planning.

All in all, theoretical background gave a good basis for the research and empirical part larger the view around the researched topic. Comparable results between those two sections could be found and the big picture around machine learning and sales and operations planning seems to be really interesting.

Keywords: sales and operations planning, machine learning, demand planning, supply planning

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TIIVISTELMÄ

Miska Jurva: Koneoppimisen hyödyntäminen myynnin ja tuotannon suunnittelussa
Diplomityö
Tampereen yliopisto
Tuotantotalouden diplomi-insinöörin tutkinto-ohjelma
maaliskuu 2022

Myynnin, varaston ja tuotannon suunnittelu on ollut osa organisaatioiden arkea jo kymmenien vuosien ajan. Vaikka aihepiiriä on tutkittu ja tutkitaan yhä paljon, organisaatioilla on edelleen haasteita saavuttaa riittävästi tuloksia ja riittävän korkea suoritustaso prosessissaan. Nykypäivänä erilaiset älykkäät työkalut ja ratkaisut ovat yleistyneet suunnitteluprosesseissa huomattavasti. Koneoppiminen on yksi näistä ratkaisuista, joka voi auttaa yrityksiä saavuttamaan parempia tuloksia suunnittelutoiminnoissa, myös tässä myynnin, varaston ja tuotannon suunnittelussa. Yleensä ongelma on löytää sopiva kohta tai ongelma prosessin sisältä, missä koneoppimista olisi järkevää hyödyntää. Koneoppimisesta on tullut myös tutkimuspiireissä yhä suositumpi aihe sen yleistymisen ja kiinnostuksen myötä, ja yleisesti tekoäly on tutkimuksellisesti kiinnostava aihe monien mahdollisuuksien ansiosta.

Tämä diplomityö keskittyy koneoppimisen mahdollisuuksiin myynnin, varaston ja tuotannon suunnittelussa tämän tutkimuksen kohdeyrityksessä. Tavoitteena oli keskittyä koneoppimisen tarjoamiin mahdollisuuksiin prosessin sisällä pitäen sisällään myös mahdolliset hyödyt ja haasteet, joita se mahdollisesti toisi mukanaan. Tutkimus tehtiin aihepiirin tutkimuskirjallisuuden ja kohdeyrityksessä suoritetun haastattelukierroksen avulla. Haastatteluihin osallistui kohdeyrityksen avainhenkilöitä suunnitteluprosessin näkökulmasta pitäen sisällä yrityksen eri liiketoiminta-alueet. Kirjallisuuskatsaus esittelee tutkimusaiheen teoreettisen taustan sisältäen myynnin, varaston ja tuotannon suunnittelun, koneoppimisen sekä näiden osa-alueiden synteesin. Tutkimuksen empiirinen osuus on rakennettu suunnitteluprosessin pohjalta keskittyen erityisesti koneoppimisen mahdollisuuksiin, vaikuttaviin taustatekijöihin ja mahdollisiin haasteisiin, joita koneoppimisratkaisun pitäisi pystyä pitämään sisällään.

Tutkimuksen keskeiset löydökset liittyivät tutkimuskysymysten aiheisiin eli koneoppimisen varsinaiseen hyödyntämiseen organisaation myynnin, tuotannon ja varaston suunnitteluprosessissa, yleisesti koneoppimisen tuomiin hyötyihin edellä mainitun prosessin näkökulmasta sekä myös mahdollisiin haasteisiin. Tämän tutkimuksen perusteella yhden liiketoiminta-alueen kysynnän suunnittelu olisi potentiaalin koneoppimisen hyödyntämiselle. Tämä tarkoittaisi, että suunnitteluprosessiin saataisiin mukaan lisää ulkoisia ja sisäisiä datalähteitä, kuten säätiedot, jolloin olisi mahdollisuus saavuttaa entistä tarkempi ja faktapohjaisempi kysynnän suunnitteluprosessi. Koneoppimiseen arvoon liittyvät päätekijät tämän tutkimuksen perusteella olivat koneoppimisen laskentateho sekä kyky prosessoida suuria datamääriä tehokkaasti. Näiden avulla mahdollistettaisiin parempi mahdollisuus skenaarioanalyyseihin sekä virheiden mahdollisuus pienenisi. Suurimmiksi haasteiksi taas tunnistettiin kvalitatiivisen ja kokemuspohjaisen datan hyödyntäminen suunnitteluprosessin ja koneoppimisen rajapinnassa. Tämänkaltaisten datalähteiden vaikutus kuitenkin tunnistettiin tärkeäksi osaksi myynnin, tuotannon ja varaston suunnitteluprosesseja, jolloin niitä ei voida täysin sivuuttaa, mikäli prosessista halutaan tarkka sekä aidosti faktapohjainen.

Kaiken kaikkiaan kirjallisuuskatsaus loi hyvän pohjan tutkimusaiheen empiiriselle tutkimiselle, minkä jälkeen tutkimushaastatteluiden avulla näkökulmaa saatiin laajennettua. Vertailukelpoisia tuloksia saavutettiin, mikä mahdollisti isomman kuvan muodostamisen tutkittavan aiheen ympärille.

Avainsanat: myynnin suunnittelu, tuotannon suunnittelu, koneoppiminen

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla

PREFACE

It took a bit more than 5,5 years from the beginning of the studies until finishing this master's thesis project. This whole journey has taught me a lot and I have met a lot of amazing people. Especially thankful I am from those lifelong friendships that the university journey has brought into my life.

I want to thank Enfo Oy for offering me this opportunity to do the master's thesis about an interesting topic and also the customer that made the whole project possible. Especially I want to thank my supervisor Satu who supported me throughout the whole project and with her own motivation gave me the boost to finish this project in time. The whole process has been interesting and taught me a lot of new. Very big thanks goes also to my examiner Teemu Laine who supported me throughout the process, gave new ideas and feedback about the progress. Also thanks to Tuomas Korhonen who was the other examiner from university side.

Lastly the biggest thanks from the whole 5,5 years go to my closest university friends who made the journey fun and unforgettable! You gave me support and feedback when it was really needed. Also thanks to my family from your support during the whole study time, and of course my girlfriend Miina who always supports me before, during and after this project in every day life.

Tampere, 25.3.2022

Miska Jurva

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AI	artificial intelligence
BEP	business efficiency planning
IBP	integrated business planning
IT	information technology
ML	machine learning
S&OP	sales and operations planning
SKU	stock keeping unit

1. INTRODUCTION

This chapter describes the background of the thesis including the reasons for the research and the main objectives of this project. **The main purpose is to dig into the planning process of a customer of the case company and with the help of the research, find some process phases where machine learning/artificial intelligence solutions would bring value and effectivity to the process. Other focus areas are to recognize what kind of value machine learning could bring into process and also what are the main challenges regarding machine learning utilization.** Other main reason for the research is case company's interest to find and offer modern planning solutions for their customers. Another interest is the customer's interest in developing its planning process into more intelligent direction. This thesis has been made for a case company, together with a one of its customer, and the researcher has worked for the case company already before this research project.

1.1 Background of the research

Sales and operations planning (S&OP) has been part of organizations' life already for 30 years but still nowadays many companies are struggling to reach enough results from their process due to various reasons. S&OP process is a complex process including many different parts like demand planning, supply planning and financial plan which is one of the reasons for the problems. (Wilson & Raman 2017) According to Schlegel et al. (2021) especially implementation of an efficient process has been one of the most difficult parts related to S&OP process. Also increasing amount of data has brought new difficulties to the organizations even though, at the same time, data has opened new possibilities regarding the whole planning process.

Machine learning (ML) has solved many problems in a shorter timeframe than a human could do even though there is still a gap between ML and different decision models. ML can assist humans in an organization by giving different kinds of recommendations and highlighting the most important factors. But the most important thing is that ML enables people to focus more on the decision making rather than producing the numbers or materials for planning process. (Mulvey 2017)

It's known that there is still lack of research between ML and especially financial planning solutions but there are already some notable examples of those (Mulvey 2017). According to Rai et al. (2021) big organizations in the world are investing more and more money in different ML and artificial intelligence (AI) solutions nowadays because those can significantly help businesses to grow and develop. Examples of the benefits are cost savings through more efficient planning, time saving, better quality in final products and process optimization. In this research focus is especially on process optimization from different perspectives.

Next chapter will describe in more detail the business case of this research, which is combination of complex planning process and technical solutions, which is exactly the point where S&OP planning and possibilities for machine learning meet in this context.

1.2 Business case and goals

The purpose of the research is to **go through the S&OP process of a customer of the case company and recognize suitable opportunities for ML. This also includes the possible value and challenges regarding ML utilization.** The whole planning process consists of a couple of different subprocesses. Rough description of that is visible in figure 1 below.

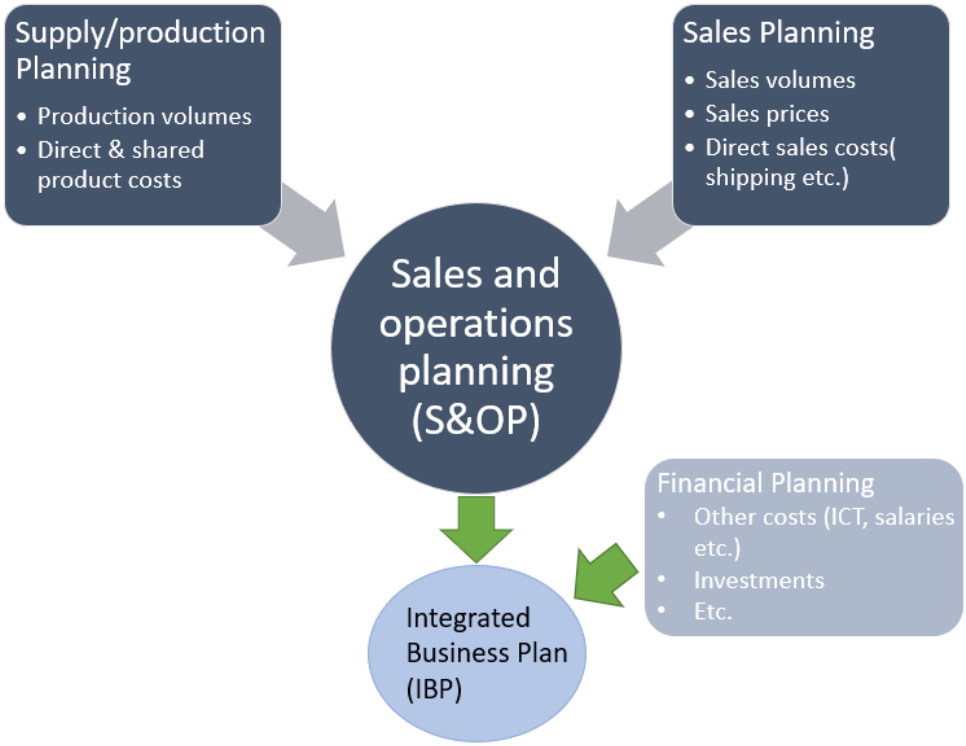


Figure 1: Description of the planning process

Case company's motivation for the research is in developing better and more effective planning solutions especially in the field of ML and AI for their customers. Customer's interest for the research comes from the idea of developing their own planning processes by using new and modern technical solutions.

This combination of interests has led to this research topic. Before the research, it's already known what kind of subprocesses are included in the business planning process of the customer. That is why the idea is to go through the processes with the help of key persons of the customer company by interviewing them and find parts where the intelligent machine learning solutions would be suitable to utilize and bring value into the process. Interesting part is also to find what kinds of challenges might occur in machine learning utilization. The research questions (presented in chapter 1.3) handle also these perspectives.

Purpose of the literature review is to find out, what does financial planning/S&OP mean and how machine learning and other technical solutions could help these kinds of processes. Other purpose is to create a basis for the research and offer theoretical background for the research topic.

With the help of this research, the idea is to help case company to develop their skills in the context of ML utilization in planning processes. At the same time, **goal is to find opportunities where customer company could develop their process with the help of ML including the value and challenge factors of that utilization. These answers will be collected with the help of literature review and an interview round within the organization.** That also summarizes the motivation of the whole study.

1.3 Research questions

To achieve the research goals describe in chapter 1.2, the research will answer to following research questions (RQ).

RQ1: How could a company utilize machine learning in S&OP process?

RQ2: What kind of value could ML solution bring into the S&OP process?

RQ3: What are the main challenges in machine learning utilization in this context?

These are also connected to the reasons why this research exists, presented in previous chapters 1.1 and 1.2. With the help of these perspectives the business goals can be reached as well as the scientific interest can be fulfilled.

These questions aim to help case company to plan their machine learning solutions to the next level and bring value to its customer's planning process. First question is answered in the empirical part by interviewing the key person from the customer's side and combining these findings into one summarized finding. As a background of the interview questions, literature review and findings from there have been utilized.

Second question is answered with the help of literature by presenting the general benefits and value in ML and then combining that into the context of S&OP. Also the empirical part defines the parts where ML utilization could bring value into the process and the combination of these will present the main findings about the value factors.

The third question is answered through the literature by combining those findings into the challenges found in the empirical part. By doing this, the overall view of the challenges is presented and can be monitored how the theoretical and empirical part differ from each other, and in the other hand, what those have in common.

1.4 Case company

The case company of this research is a Finnish company, that provides different kinds of information technology (IT) services and consultancy for its customers, including also financial planning software. Offering includes different kinds of technologies and services for the customers and the company operates mainly in Nordic countries, employing about 900 employees.

Also one customer of this case company has a significant role in this research, because the interviewed persons are from there, and the research is handling their S&OP process. This company is an international company, that aims to satisfy people's needs in everyday life in the field of gardening and professional growing. The company also operates in the energy sector by providing energy solutions and fuels to its customers.

This research was focusing mainly on businesses S&OP planning processes. Group of interviewed persons includes people from every business area and sector so that the overall view of the S&OP process and ML possibilities can be created.

1.5 Thesis structure

After the introduction chapter, Chapter 2 presents the literature review of the topic. Chapter describes the different parts of sales and operations planning process and defines ML, AI and their use cases in this research context. Based on these, theoretical analysis about utilizing these technologies in S&OP planning processes is presented.

Chapter 3 is for describing the research methodology of this thesis. This includes one part for the research methods, one for the research process and the last one for data collection and analysis. Also the interview process is presented in this chapter.

Chapter 4 presents the results of the empirical part of this research. That includes the results and answers in more detail to research questions (from chapter 1.3). The chapter is divided into different subchapters based on the research questions. The final part includes some thoughts and ideas about the possible implementation plan and actions what could be done based on this research.

Chapter 5 includes the discussion part of this research. There the literature findings and empirical part are combined and discussed. This chapter is also divided into smaller subchapter based on the themes of research questions.

Last one, chapter 6, presents the conclusion of this research. First subchapter presents the summarized answers to research questions, the second one presents some ideas for the future research in this field and last one included some thought about validity and reliability of this research. After that chapter, there are the references and appendix of this study.

2. LITERATURE REVIEW

Literature review offers a theoretical background about sales and operations planning (S&OP) processes in companies and the role of different intelligent solutions like machine learning in these. This literature review also aims to build a basis for using machine learning solutions in S&OP in general and what kind of value these solutions could bring into the planning. First chapter, 2.1, introduces the S&OP process in general and what does it mean in this research context. That includes the subprocesses of demand, supply and financial planning. The second part, 2.2, focuses on artificial intelligence and especially machine learning (ML) including the possible benefits and challenges of ML. Last part, 2.3, introduces what does ML mean in the context of S&OP and what kind of utilization possibilities there could be.

2.1 Sales and operations planning process definition

S&OP as a term has been used already since 1970s and in most recent studies also a term integrated business planning (IBP) has been used (Iyengar & Gupta 2013). Traditionally in organizations process is seen as balancing demand and supply (Bower 2005; Jacobs et al. 2018, pp. 110), but according to Bower (2005) this narrow perspective won't lead the most efficient results due to lack of critical information. Also, according to Iyengar & Gupta (2013) very often in organizations S&OP is only discussion between sales and supply chain functions, and finance has only a small role in these conversations even though it should be part of the process to achieve the next level of S&OP, which is often called IBP. Feng et al. (2013) also described S&OP as a supply chain process where the main focus is on finding the balance between demand and supply.

In addition to demand and supply balancing Wallace (2006) defined S&OP as a decision making process which helps organizations to help integrate operational and financial plans, balance supply and demand, and align volume and product mix. This definition is closer to broader term of IBP. Feigin (2012, p. 71) also described S&OP as a process where the final output is a business plan including information about demand, supply and financial side, which is aligned with a definition of IBP.

According to Iyengar & Gupta (2013) S&OP should be handled in organizations like any other strategic process because otherwise the benefits won't realize, and it can lead to unnecessary increased workload. Oliva & Watson (2011) also mentioned in their paper

that collaboration throughout the whole organization is needed when planning and making actions regarding the S&OP process. Feigin (2012, p. 72) also highlighted that S&OP should take into account every parts of an organization due to the fact that it affects every part of the organization.

In addition to previous facts about the process, S&OP is also about communication, alignment with other strategic goals and improving the core competences of the organization. S&OP also includes more than just demand and supply related discussions. Important factors are S&OP organization, S&OP culture, tools and data, strategic alignment and performance management. (Tuomikangas & Kaipia 2014)

If we combine this larger view with financial planning in general the outcome will be IBP and together with a firm's strategy even larger view of the whole which is called business efficiency planning (BEP) (Wilson & Raman 2017). They also agreed the idea of S&OP as a decision making process especially regarding the demand and supply but from a bit larger perspective.

Where will the organization then be when taking into account all these things, is the question that many research papers are trying to answer. According to Wallace (2006) and Bower (2005) an efficient S&OP will open the window into the future instead of looking just the history. Also, the readiness for surprises and changes will be much better.

Bower (2005) described S&OP as a process of five steps: portfolio planning, demand planning, supply planning, measuring & monitoring the organization and management review. This research will focus on demand and supply planning including the financial side, that Iyengar & Gupta (2013) presented as a part of S&OP and presented IBP.

All in all, according to research literature S&OP should be seen as critical factor of the organizations' success including all the roles, subprocesses, decision making and involvement of the whole organization. But it is also known that many organizations have difficulties in implementation of efficient S&OP process so that the benefits would be visible.

Despite of many different terms and varying definitions in the literature, in this thesis S&OP is a process where demand, supply and finance are represented and as an output will be created an integrated business plan. That also includes that all the relevant functions are included in the process, not just sales and supply chain. This is also aligned with recent research literature where IBP is used instead of S&OP. In the next chapter the selected subprocesses are presented more closely.

2.1.1 Demand planning

According to Oliva & Watson (2011) demand planning is defined as forecasting and planning future customer demand taking into account different kinds of market conditions and demand plan is the basis for production planning. Feigin (2012, p.15) presented a bit broader definition where demand planning is trying to predict what a company is going to sell in the future and that depends on, who is asking and from what perspective. Both Oliva & Watson (2011) and Feigin (2012, p. 15) highlighted the importance of taking the market conditions into account. Feigin (2012, p. 15) pointed out that demand depends at least on product, state of economy, time period, availability and marketing. Rapid changing environment has made demand planning even more important and complex process within organizations (Chase 2013, p. 32).

According to Feigin (2012, p. 15) demand planning also affects many other activities within a company by highlighting at least supply planning, inventory levels, capacity planning and the need of working capital. From these reasons, demand planning is already an old actions within companies (Feigin 2012, p. 16). Also, Chase (2013, p.31) highlighted that demand planning impact on various activities especially in supply chain, recruitment plan, shipping plan and production in general, for example.

Even though demand planning is an important factor within a company, forecasts are always wrong, but the question is how much (Feigin 2021, p. 16). Instead of that, organizations must forecast and try to predict their demand if they want to achieve successful results and be alive (Feigin 2012, p.15) and poor forecasting has been seen as a barrier to success (Feigin 2012, p. 17). According to Chase (2013, p. 33) organizations need to focus especially on data quality if they want to achieve good level of demand planning accuracy.

According to Feigin (2012, p. 17) the financial outcome of successful demand planning can be significant and also minor improvements in it can lead to huge cost savings in different supply chain operations (Feigin 2012, p. 17). But it is still important to remember that demand planning operations cannot produce any financial outcomes alone, it is just part of the whole planning process (Uzsov et al. 2018).

Time period in demand planning is often between six and 12 months but it can also look further to the future (Feigin 2012, p. 18). Like already mentioned, it needs to take into account different kinds of factors within and outside of an organization, but the planning is often based much on historical sales data. But in a way or another, the plan needs to also cover the future predictions and competition around the company, especially when it comes to the new products where the there is a lack of historical sales data. (Feigin 2012, pp. 18-19)

Also, the different phases of product lifecycle need to be taken into account, so there cannot be ‘one size fits all’ thinking too much (Feigin 2012, p. 20). According to Uzsov et al. (2018) geography, number of stock keeping units (SKU), time range and product families need to be considered when planning the future demand. All in all, demand planning is a very important process within companies but still only a part of the total operations’ planning that affect other processes, and on the other hand, other processes affect the demand planning process.

2.1.2 Supply planning

According to Feigin (2012, p.109) supply planning is trying to answer questions like when to produce, what to produce, where to produce and when to begin the production. It is also closely connected to other supply chain operations within a company, especially inventory planning. Supply planning is linked with other parts of the supply chain organization, such as manufacturing, logistics and procurement (Liberatore & Miller 2012, p. 7). According to Wiedenmann (2013, p. 25) supply planning is a process which takes demand data and develops the whole supply and production network trying to answer to customer demand.

Basis of supply planning is the demand plan so there is a strong dependency between those (Feigin 2012, p. 110), like already mentioned in chapter 2.1.1. In addition to that, supply planning needs to be linked with organization’s strategy and cannot be handled in complete separate process (Liberatore & Miller 2012, p. 4). Also according to Wiedenmann (2013, p. 26) the supply planning objectives need to be linked to organization’s overall goals, both in short and long term (Liberatore & Miller 2012, p. 2).

Due to strong dependency to demand planning, supply planning has same kinds of factors than demand (Wiedenmann 2013, p. 26). Even though the time range might be the same than in demand planning (Feigin 2012, p. 18), the seasonality might affect supply possibilities and also the planning, which brings, for example, inventory planning really close to that process (Wiedenmann 2013, p. 26).

The complexity of supply planning depends on different factors where the most important ones are the number of products and SKUs. That is why organizations often have specific roles, for example supply planners and inventory planners, for these actions. (Feigin 2012, pp. 110-111) Due to the complexity of the process, it needs incremental improvement all the time and with the help of that, organization can achieve competitive advantage in short, medium and long term (Liberatore & Miller 2012, p. 2). Time range affect the planning, and especially the question what to plan. According to Thunberg & Fredriksson (2018) in long-term planning important thing is, where and how to arrange

the production. In mid-term planning the main focus should be production volumes by products (and also inventory levels), and in short-term planning most important questions related to daily production volumes and shift capacity (Thunberg & Fredriksson 2018).

Overall, supply planning is one of the factors that drives the whole business (Liberatore & Miller 2012, p. 5) and it can also affect operative costs and financial outcomes in general (Liberatore & Miller 2012, p. 3). Due to these reasons and content of this chapter 2.1.2 organizations should highlight the importance of supply planning and also the connections between different operations.

2.1.3 Financial planning

Financial planning within organization is strongly based on traditional budgets. The mission of financial planning, and especially budgeting is to follow costs and operations. Those are often based mostly on assumptions and guesses, even though traditional budgeting is seen as a process which requires too much work compared to value. Often there are two ways of creating the financial plan: budget for one specific calendar year or rolling forecast for some specific time range, for example 12 months. (Bergmann 2020)

In addition to costs, financial planning process can include the plan about incomes and taxes, for example. The planning process include different steps that are gathering relevant data, analyzing data, developing the estimations, implementing the plan and monitoring the plan. (Chaffin 2015, p. 93) According to Oesterreich et al. (2019) digitalization and new kinds of modern tools has developed the whole financial planning into new direction which can include more intelligent solutions within the process.

Secrett (2012) defined financial planning as a process that can also include cash flow forecast in addition to incomes and costs. Most important thing is to know the business where the company is operating, because the relevant affecting factors come from the business and are the basis for financial planning. (Secrett 2012) According to Secrett (2012) the level of details in financial planning can vary between organizations and that depends on the organization’s interest in plan and monitor their operations.

All in all, financial planning has connections to other processes in business planning. Demand planning affect strongly in incomes side and supply plan strongly affects the operative costs. That’s why these need to be integrated between each other’s and cannot be handled as completely separate processes.

2.1.4 Enablers of an efficient sales and operations planning process

There are multiple different factors that enable an efficient S&OP process. These were found by going through the research papers which were about S&OP process and in eleven papers were mentioned different enablers. These finding are presented in table 1 below.

Table 1: Enablers of an efficient S&OP process

	Bower 2005	Wallace 2006	Boyer 2009	Oliva & Watson 2011	Feigin 2012	Prokopets 2012	Marcio Tavares Thome et al. 2012	Iyengar & Gupta 2013	Tuomikangas & Kaipia 2014	Pedroso et al. 2016	Wilson & Raman 2017
One set of numbers		x	x			x	x				x
Data quality	x					x		x	x		
Strategic alignment			x				x		x	x	x
Collaboration with the whole organization			x	x		x	x			x	
Valid assumptions	x										
Based on facts	x							x			
Communication	x										
Relevant analysis	x										
Relevant metrics	x				x	x		x		x	x
Enough long-term view	x					x	x	x			
Clear ownership								x			
Clear vision											x
Efficient decision making						x				x	x
Formal process created						x					
Supply plan based on demand					x	x					
Support from management team			x			x				x	
Clear and relevant targets set						x					
Scenario analysis possibility						x				x	
Ability to balance demand and supply						x					
Plans interact directly with other IT systems						x				x	

According to research papers the most important enablers are **one set of numbers** (Wallace 2006; Boyer 2009; Prokopets 2012; Marcio Tavares Thome et al. 2012; Wilson & Raman 2017), **strategic alignment** (Boyer 2009; Marcio Tavares Thome et al. 2012; Tuomikangas & Kaipia 2014; Pedroso et al. 2016; Wilson & Raman 2017), **collaboration with the whole organization** (Boyer 2009; Oliva & Watson 2012; Prokopets 2012; Marcio Tavares Thome et al. 2012; Pedroso et al. 2016) and **relevant metrics** (Bower 2006;

Feigin 2012; Prokopets 2012; Iyengar & Gupta 2013; Pedroso et al. 2016; Wilson & Raman 2017).

By one set of number was meant that all the functions and participants use the same numbers instead of everyone using different or own plans and numbers. That makes an efficient S&OP possible and huge mismatches can be avoided. (Prokopets 2012) According to Boyer (2009) S&OP is also a strategic process and should be aligned with the general strategy and shouldn't be handled as a separate process within an organization. Many papers also highlighted that S&OP needs collaboration from all of the relevant parts of an organization. That means every kind of collaboration from communication to one set of numbers (other major enabler presented in table 1). Last one of the most important enablers is relevant metrics according to many papers. It means that an organization should set relevant indicators and metrics, which help people to follow how they are performing regarding the S&OP. If that is not done, all the actions are not based on facts are tracking possible problem will become much more difficult.

In a quite many papers were mentioned data quality (Bower 2005; Prokopets 2012; Iyengar & Gupta 2013; Tuomikangas & Kaipia 2014), long term view (Bower 2005; Prokopets 2012; Marcio Tavares Thome et al. 2012; Iyengar & Gupta 2013), efficient decision making (Prokopets 2012; Pedroso et al. 2016; Wilson & Raman 2017) and support from management team (Boyer 2009; Prokopets 2012; Pedroso et al. 2016) as enablers of an S&OP process.

Also, other enablers were recognized in the literature but those were not mentioned in so many research papers. The most notable ones from those are interaction with IT systems (Prokopets 2012; Pedroso et al. 2016), scenario analysis (Prokopets 2012; Pedroso et al. 2016) and actions based on facts (Bower 2005; Iyengar & Gupta 2013). Other findings with less mentions in the literature regarding S&OP enablers can be found in table 1. To summarize the findings, the enablers are very versatile and are connected to S&OP process from many different perspectives.

All in all, according to Tuomikangas & Kaipia (2014) the main focus areas in S&OP should be in tools & data, process and organization. Also, all the findings in table 1 can be categorized to these topics. All of these areas should work if the goal is an efficient S&OP process.

2.1.5 Traditional challenges of sales and operations planning

Challenges are also a part of S&OP process. Multiple challenges were found in literature review, and these are presented in table 2 in a same way than the enablers in previous section. Even though, there were as many papers included in gathering the challenges

than in enablers section, not so many papers (only six) included challenges compared to the enablers. Also, the variation was bigger, and the challenges were not so clearly recognized than the enablers of S&OP.

Table 2: Traditional challenges of S&OP

	Bower 2005	Boyer 2009	Feigin 2012	Iyengar & Gupta 2013	Pedroso et al. 2016	Wilson & Raman 2017
Disconnection between S&OP and strategy	x			x	x	x
Indecision of management team	x					x
Forecasts not based on reality	x			x	x	
Too much meetings	x					
Focus only on the balance of the whole calendar year	x					
Unclear ownership	x			x		
Looking too much at the history data	x		x		x	x
Ignoring product life cycle management	x					
No relevant trends included	x					
Lack of measuring and monitoring	x			x		
No valid meeting procedures	x				x	
Lack of resources		x				
Poor data management				x		
Wrong or misleading metrics				x		x
Focusing only on short term forecasting				x	x	
Disconnection between the goals of employees and organization				x		
Lack of participation					x	
Lack of support from management team					x	
Lacking IT systems					x	
External trends excluded					x	
Lack of collaboration					x	
No one set of numbers						x
Execution						x

The most notable S&OP challenges according to literature are **disconnection between S&OP and strategy** (Bower 2005; Iyengar & Gupta 2013; Pedroso et al. 2016; Wilson & Raman 2017), **forecasts and numbers are not based on facts** (Bower 2005; Iyengar & Gupta 2013; Pedroso et al. 2016) and **too much at history looking process** (Bower 2005; Feigin 2012; Pedroso et al. 2016; Wilson & Raman 2017). Disconnection to strategy is closely connected to strategic alignment presented as an enabler in previous section. Not fact based decision making and actions were seen as a challenge in the literature (Bower 2005), and this is connected to data quality (Prokopets 2012) and fact based actions (Iyengar & Gupta 2013) from enablers section. Lack of data or its quality can lead to make decisions and act only based on feelings or thoughts and not facts. Even though the history is important factor when thinking about S&OP, it shouldn't be the main focus area where the goal is to define future demand and balance the production based on that. The major role of history and history data is also one factor why this has been seen as a challenge to look more into the future.

Other challenges were mentioned only once or twice in the research literature and those are presented mainly in table 2. A couple of the challenges are really close to the major enablers. Lack of measuring and monitoring (Bower 2005; Iyengar & Gupta 2013) and wrong or misleading metrics (Iyengar & Gupta 2013; Wilson & Raman 2017) are basically the same than relevant metrics in enablers, which means that at the same time it is a major enabler but also a challenge to many organizations. The same situation is with indecision of management team (Bower 2005; Wilson & Raman 2017), which was presented as a challenge in organizations. Support from management team was at the same time an enabler for an efficient S&OP, so there is often a lack of support even though that would be a critical factor of success.

All in all, these challenges are aligned with the enablers section and are connected to same themes. Tuomikangas & Kaipia (2014) presented the framework where all S&OP enablers are connected to tools & data, process or organization. This is valid also to this challenges topic, and the findings (table 2) can also be categorized under these topics. Even though there were more variation in challenges than in enablers, they handle the topic also from many different perspectives and are closely connected to enablers section.

2.2 Artificial intelligence and machine learning

This chapter focuses on definition and different features of artificial intelligence (AI) and after that will be introduced one direction of that, machine learning, since it is relevant

point of view in this study context. In the next chapter (2.3) this topic will be combined with sales and operations planning processes.

2.2.1 Definition and features of artificial intelligence

AI is not one technology, but under that go many different methods, technologies, applications and research directions. It also deals with many different sciences: mathematics, physics and engineering, as a couple of examples. Nowadays, it also has an impact to our whole society which brings many different aspects under the same topic. (Ailisto et al. 2018) According to Russel & Norvig (2016, p. 1) AI is always a relevant topic if there is an intellectual task to solve, apart from the science direction.

There isn't just one simple definition for the term AI. Instead of that, there can be found several definitions from the research literature. Based on Russel & Norvig (2016, pp. 2-3) definition with the help of AI machines, equipment, applications can act in rational way in different situations. According to Duan et al. (2019) there is no general definition for AI, but they defined it as machines' ability to accomplish the tasks like human based on what they have learned and experienced earlier. From a bit different perspective, the definition of AI is a system's ability to simulate human being by utilizing large scale of different techniques (Fisher et al. 2016; Bolander 2019). Even though the goal of AI is to perform the tasks like human (Bolander 2019) we cannot communicate with it or always forecast how does it work.

In this research we define AI as a tool and ability to learn from historical data and experiences and so help humans to accomplish the tasks and help to make decisions or make the decision on behalf of humans.

According to Ailisto et al. (2018) different AI technologies can be described with the help of various features, which are presented in figure 2 below. These features need to be taken into account both in the development of AI technologies and ready-made technologies.

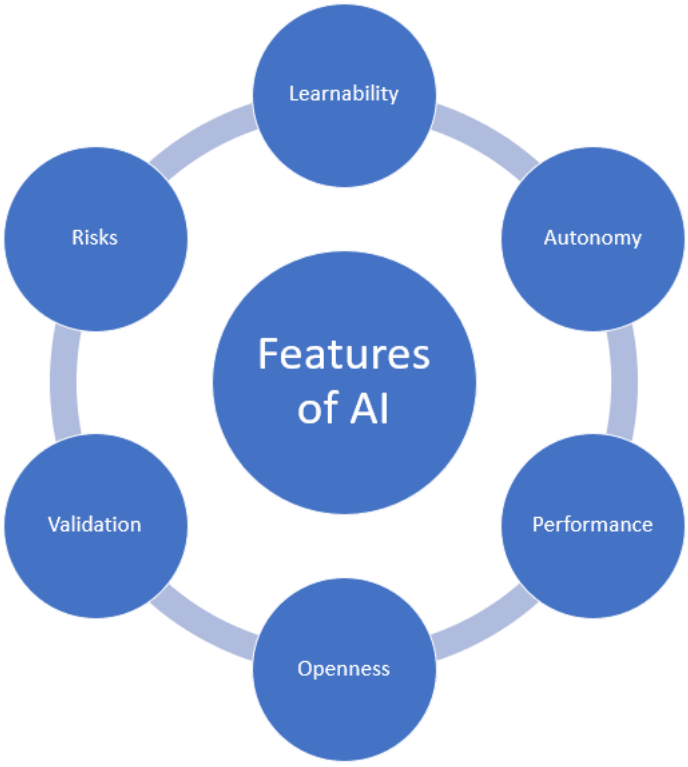


Figure 2: Features of AI (based on Ailisto et al. 2018)

Features, that need to be considered in the development of AI technologies, are validation, risks and openness. Validation means how to measure the performance of technology and its development especially, which include the relevant measures and indicators. It also includes that the data has enough quality, and the system or software can learn in a way that is wanted. Always when developing AI systems, there are questions about privacy, safety, security and transparency. Risks feature is for planning and handling these elements of development and usage of AI. That also includes the ways how to handle and make sure that everything would go as planned and by the book. Last feature of the development part is openness. That means transparency and predictability of the logic and open source code of the data. Openness is important also for the future research and development. (Ailisto et al. 2018)

Learnability, autonomy and performance are features of ready-made AI technology. The systems' ability to learn without pre-programming is called learnability. That also includes that the system is able to perform new things and face different kinds of situations and after that handle those in relevant way. Autonomy means the level at which the technology is able to operate independently without the help of a human. Current systems and applications can solve problems based on what they have learned but that's what limits their autonomy level. In general, there are different levels of autonomy nowadays in the systems. Last feature of this part is performance, which means how efficiently systems

can perform compared to human. Classic example and a way of testing this, is a system playing a game, chess for example, against a human. (Ailisto et al. 2018)

2.2.2 Machine learning in general

There are several different research directions under AI but now in this research the focus is only on ML. ML belongs to the field of computer science and is very fast growing part of it nowadays (Alpaydin 2014, p. 1). There can be found several different definitions of ML in the research literature but according to Alpaydin (2014, p. 2) ML is a combination of algorithms where the computer is solving a problem, completing a certain task or building a pattern based on historical data and results. Chandramouli et al. (2018, pp. 7-8) expanded this definition a bit by presenting the idea of the process. They divided ML into three smaller parts where the first is data input, which basically means historical data and information. Second part is abstraction where the algorithms are processing the data and presenting it in a broader way. Last part is the generalization where the machine is able to build decision making framework. These three parts are visible in figure 3 below.

Figure 3: Machine learning process

Even though the figure 3 above is presenting the idea in a pretty simple form, ML is more complex process that includes several different methods and techniques. These definitions are valid in this research where the goal is to utilize historical data and information in efficient way to develop the planning process.

ML can be divided into subcategories, which are supervised learning, unsupervised learning and reinforcement learning (Ailisto et al. 2018; Chandramouli et al. 2018, p. 15). These subcategories are presented in figure 4.

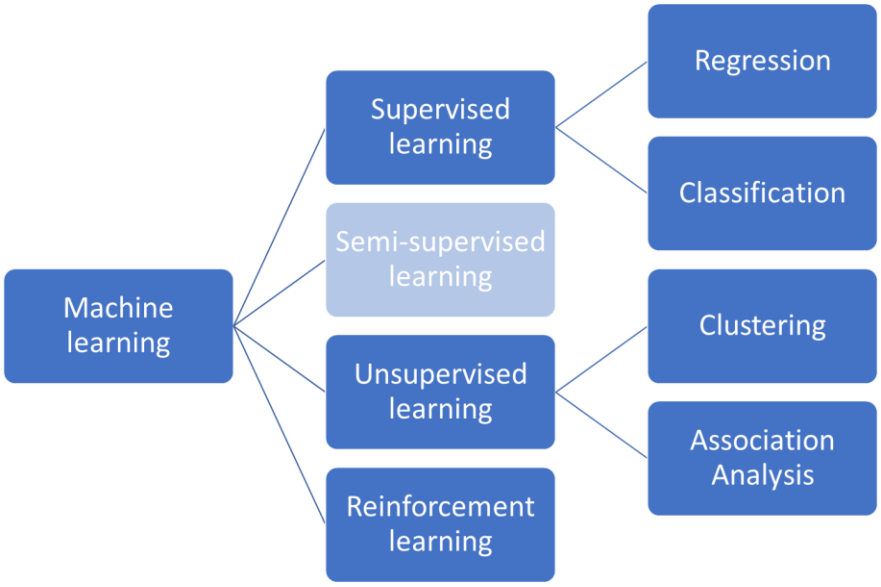


Figure 4: Machine learning types

In supervised learning the system uses example data where both the input and result are known. Based on those, system is trying to understand the relationship between the input and results so that then in the future it would be able to solve the results based on new inputs. (Lee et al. 2018) According to Ailisto et al. (2018) the most common supervised learning methods are linear/logistics regression, neural network and support vector machine. In classification the purpose is that the algorithms learn to process the data into right class based on the example data that the system take in (Havlicek et al. 2019). According to Whitenack (2017) another typical supervised learning method is regression. In a nutshell that means understanding how a variable value changes compared to another variable, for example how the number of sales persons affect to sales.

Lee at al. (2018) described unsupervised learning as a tool in clustering, compression and feature extraction. In supervised learning the system is trying to learn based on input X without knowing the results. It tries to understand the expected behavior and based on that build a pattern. (Lee et al. 2018) According to Nguyen & De Baets (2019) clustering is a task of pattern recognition and is close to classification (in supervised learning) but is using different mathematical methods, k-means for example. Compared to classification and supervised learning also new labels can be found with the help of clustering.

Lee et al. (2018) also presented another, newer category, semi-supervised learning, which is between supervised and unsupervised learning. That is a combination where some of the results are known and some not. Based on weighted calculation the system is trying to identify the relationship and expected results and learn to build patterns. This has become more popular during the last years according to Lee et al. (2018).

Last category, reinforcement learning, differs from previous ones so that the system learns to act according to decision making policy. System takes input data and is building the act/decision making knowledge based on that data. In supervised/unsupervised learning systems are not able to act or make decisions but they are building different patterns based on the data. (Lee et al. 2018) According to Wiering & Otterlo (2012, p. 5) reinforcement learning is between supervised and unsupervised learning and the main goal is that the learning agent can behave and act in a long run in the specific environment.

2.2.3 Examples of machine learning use cases

This chapter highlights the ML use cases in general before going into the synthesis of S&OP and ML. These cases have been found from the literature and describe ML usage in organizations. Main point is to describe which kind of many ways have been found to use ML in different kinds of situations. In the next chapters will be presented more closely values and challenges regarding ML usage.

Nielsen (2018) presented a case where a company has used ML to find and serve new market segments before its competitors. That was based on many kind of customer and market data. On general level Nielsen (2018) presented that with ML (and/or other intelligent) solutions can be automatized many kinds of different management accounting tasks. According to Möller et al. (2020) ML can help finance function in general in their decision making and also reporting processes.

Bergmann (2020) presented that intelligent solutions like ML offer various range of opportunities for decision making by gathering and processing data. In more concrete ML can provide automatized decisions instead of humans doing everything by themselves. ML can also offer recommendations about what should be done and follow. (Bergmann 2020)

Often the discussion goes towards the topic machine vs. human. According to Bolander (2019) machines are often better than humans in situations where the rules or logic is simpler and vice versa. It means that in more complex situations more human involvement is needed. Brynjolfsson & Mitchell (2017) reminded that machine learning is not able to do everything what humans can.

Nielsen (2018) also highlighted that companies get from every dollar they invest in intelligent solutions (including ML), 10,66 US dollars back. That really shows why this topic is really interesting in many companies and organizations nowadays.

2.2.4 Machine learning value and benefits

ML can bring many kind of value and benefits for a company. Based on research literature, several different topics were found, and those findings are presented in table 3 below.

Table 3: Value and benefits of ML

	Skitka et al. 2000	Martinez-Lopez & Casillas 2013	Brynjolfsson & Mitchel 2017	Economist 2017	Sutton et al. 2018	Stormi et al. 2018	Bolander 2019	Bergmann et al. 2020	Möller et al. 2020
Improved forecasting accuracy								x	
Time savings			x					x	
Improved decision making process		x					x	x	
More calculation power					x	x	x	x	
Business growth								x	
Savings in human resources		x	x	x	x		x		
Increased financial results							x		
Better productivity			x						
Cultural change									x
Reduced amount of errors	x								

The two most common topics which were in the literature are savings in human resources (Martinez-Lopez & Casillas 2013; Brynjolfsson & Mitchell 2017; Economist 2017; Sutton et al. 2018; Bolander 2019) and calculation power of ML (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020). Comparing human work and systems is a common research topic according to literature and both of these are connected to that. By savings in human resources were meant in general that the organization can allocate their employees differently after using a powerful ML systems and they can concentrate on topics where a machine cannot help. Better calculation power of ML is connected to limitations that humans have when there is a need of calculation with huge data samples, and that would be impossible to humans.

According to several papers improved decision making process is also one of the benefits that ML could bring for a company (Martinez-Lopez & Casillas 2013; Bolander 2019; Bergmann et al. 2020). That means that ML could help in decision making by providing important information for decision making or even do some less complex decision instead of humans. Time saving was also seen as a benefit (Brynjolfsson & Mitchell 2017; Bergmann et al. 2020) and that is closely connected also to decision making process.

ML can save time in an organization in many different ways by being more powerful and faster system than a human, as an example.

Several smaller findings were also found. Improved forecasting accuracy (Bergmann et al. 2020), business growth (Bergmann et al. 2020), increased financial results (Bolander 2019), better productivity (Brynjolfsson & Mitchell 2017), reduced number of errors (Skitka et al. 2000) and cultural change (Möller et al. 2020) were seen as possible benefits of ML. Behind all of these, there is an assumption that the implemented ML system works as planned and is able to help an organization. Challenges and threats are described in more detail in the next chapter 2.2.5.

If all the findings will be categorized under the bigger topics or themes, three separate headers can be found, presented in figure 5.

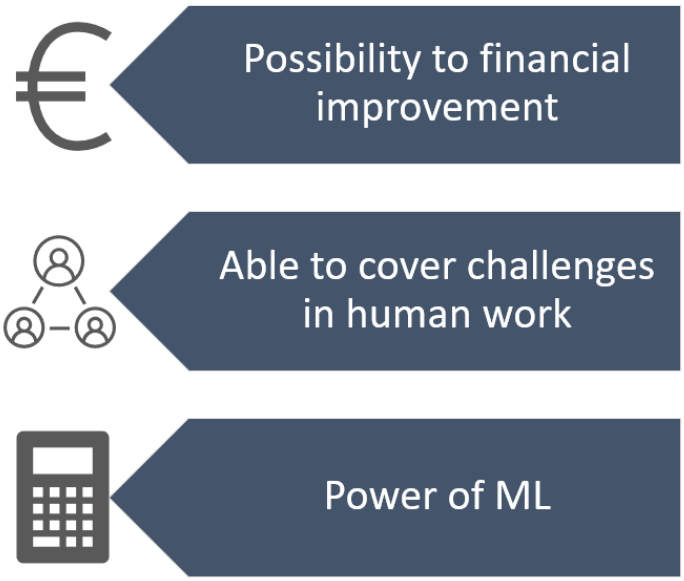


Figure 5: Value in ML

The biggest benefit category is a possibility to financial improvement by utilizing ML. Business growth (Bergmann et al. 2020) and increased financial results (Bolander 2019) are directly connected to that, and indirectly connected are almost every one of the findings (in table 3) but especially improved forecasting accuracy (Bergmann et al. 2020) and savings in human resources (Martinez-Lopez & Casillas 2013; Brynjolfsson & Mitchell 2017; Economist 2017; Sutton et al. 2018; Bolander 2019.

Second topic, challenges in human work, is a theme to which ML can provide a good solution. Especially reduced number of errors (Skitka et al. 2000) and better productivity (Brynjolfsson & Mitchell 2017) are connected to that. The last one is the power of ML, which is visible the most in calculation power (Sutton et al. 2018; Stormi et al. 2018;

Bolander 2019; Bergmann et al. 2020) and in improved and faster decision making process (Martinez-Lopez & Casillas 2013; Bolander 2019; Bergmann et al. 2020).

In addition to that, cultural change (Möller et al. 2020) is also something relevant in the context of using ML. It can be a benefit or a challenge, but it was seen as a benefit and can offer new ways of working in an organization.

2.2.5 Machine learning challenges and threats

In addition to value and benefits presented in previous chapter, there are also many challenges and threats in using ML solutions. Challenges and threats are presented in the same way than value and benefits in previous chapter. These are presented in table 4 below.

Table 4: Challenges and threats of ML

	Skitka et al. 2000	Brynjolfsson & Mitchell 2017	Nielsen 2018	Sutton et al. 2018	Bolander 2019	Oesterreich et al. 2019	Möller et al. 2020
Lack of transparency					x		
Risk of wrong conclusions	x				x		
Might be heavy to maintain		x					
Expensive		x	x				
Hard to find utilization possibilities							x
Cultural change							x
Not good at adapting to changes in the environment		x					
Requires new skills and talent from an organization			x			x	
Acceptance				x			
Decreased human creativity				x			
Deskilling effect				x			
Impaired decision making process				x			

There was more variation in challenges and threats than in value and benefits section. Only a couple of the findings rose a bit above others. According to literature ML systems are considered expensive (Brynjolfsson & Mitchell 2017; Nielsen 2018), they require new skills and talent from an organization (Nielsen 2018; Oesterreich et al. 2018) and there is a risk of wrong conclusions (Skitka et al. 2000; Bolander 2019). Expensive doesn't

mean that ML solution wouldn't be good or bring value but that can create a barrier or be harder to get an acceptance from the management, which is understandable. By new skills were meant that an organization may need to recruit new persons who are responsible of the systems, or they would need external consultancy which is closely connected to high costs in the previous finding.

Other findings in the literature were lack of transparency (Bolander 2019), heavy maintenance (Brynjolfsson & Mitchell 2017), challenges in finding utilization possibilities (Möller et al. 2020), cultural change (Möller et al. 2020), weakness in change situations (Brynjolfsson & Mitchell 2017), acceptance of the people (Sutton et al. 2020), decreased human creativity (Sutton et al. 2020), deskilling effect (Sutton et al. 2020) and impaired decision making process (Sutton et al. 2020).

If the findings (from table 4) will be categorized in the same way than in the section value and benefits, three bigger themes can be found, those are presented in figure 6 below.

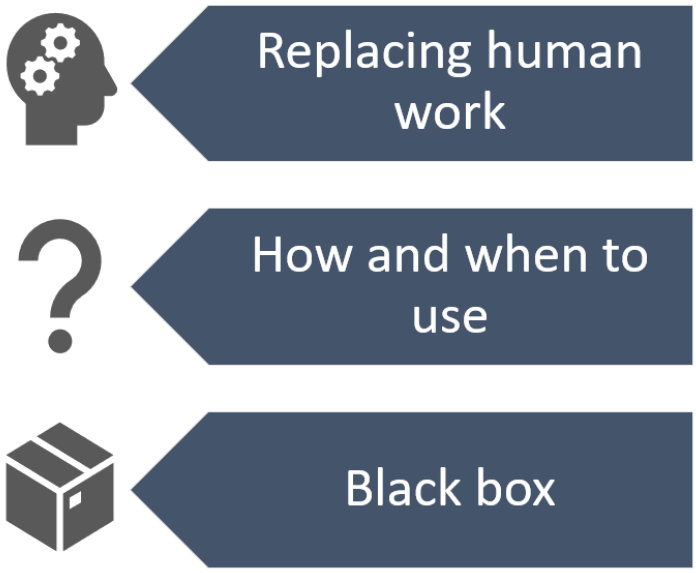


Figure 6: Challenges in ML

First challenge category based on the literature is the replacement of human work in a wrong and not efficient way. Under this go at least the findings adaption to changes (Brynjolfsson & Mitchell 2017), decreased human creativity (Sutton et al. 2020), deskilling effect (Sutton et al. 2020) and impaired decision making process (Sutton et al. 2020). All these are connected to the topic that an organization is trying to utilize ML in a case/cases where a human would act better or in more efficient way than the ML solution. There is also a risk that humans stop to use their own thinking and that is not good for the continuous improvement and creativity.

Second category is how and when to use. According to Möller et al. (2020) it's hard to find utilization possibilities for ML. Even though the possible benefits and value are existing, it's not an easy decision, a step to take or find the right phases where ML could be used in organizations, so there are still barriers between the possible benefits and implementation.

Risk of wrong conclusions (Skitka et al. 2000; Bolander 2019) and a lack of transparency (Bolander 2019) are both connected to the third category, black box. That means that there are many things in ML that are complex or hard to understand and that's why there are suspicions in the organization towards usage of ML.

Cultural change was also mentioned as a negative thing like it was also in the benefits and value section. Background of this might be connected to the fact that ML will change the culture more or less, and change needs to be managed, so it can be both value and a risk. Next chapter will present ML more closely in the context of sales and operations planning.

2.3 Machine learning as a part of sales and operations planning

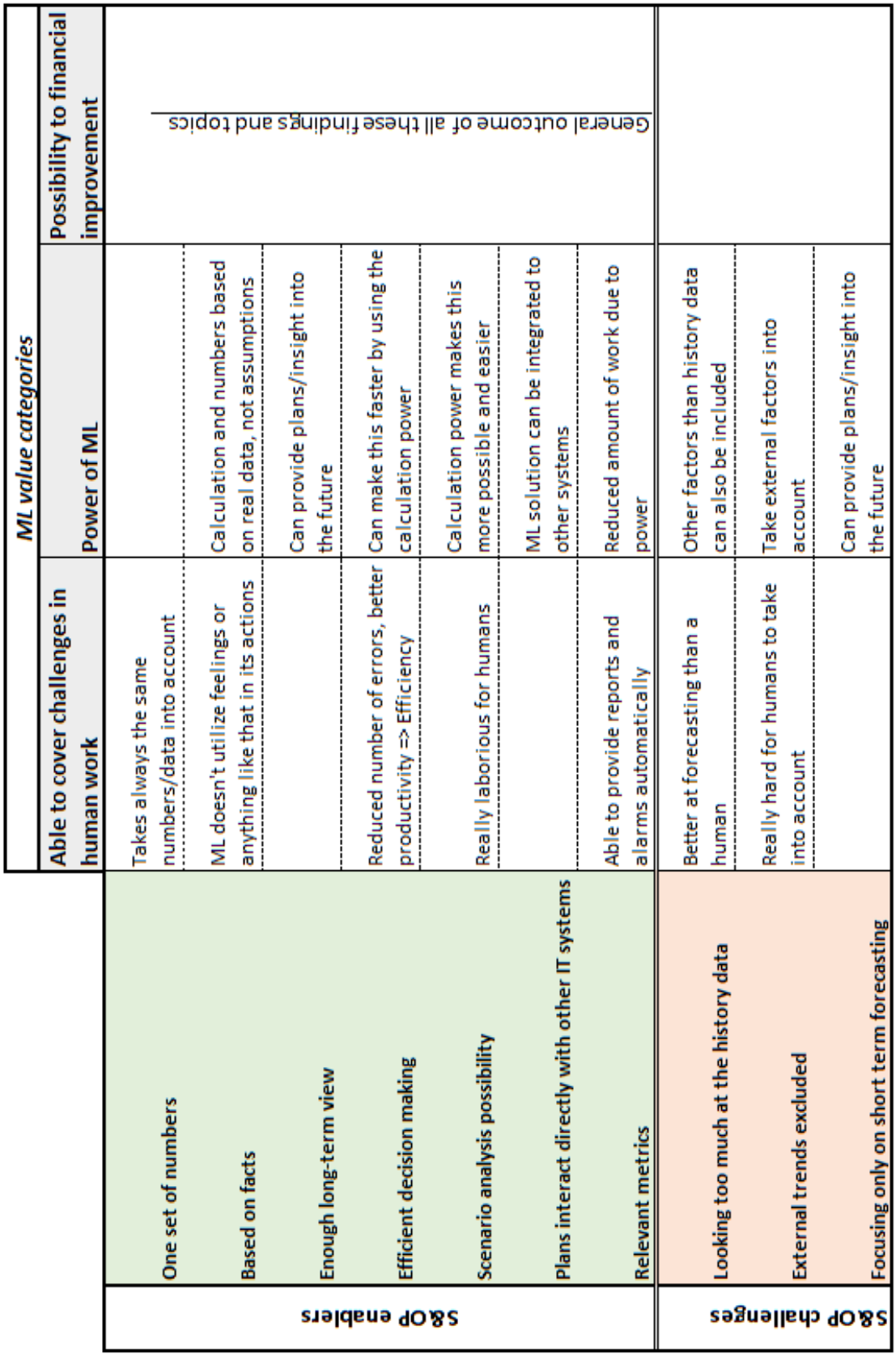
This chapter synchronizes two parts of this research, ML and S&OP. Based on the definitions in chapters 2.1 and 2.2 and literature findings about this synthesis, different ways of utilization are presented. Value and challenges regarding ML in planning processes are also presented in this chapter (2.3.1 and 2.3.2). Last part of this chapter (2.3.3) includes more general ways of look at the ML utilization level within an organization.

2.3.1 Utilization of machine learning solutions in planning process

In previous chapters S&OP enablers (chapter 2.1.4) and traditional challenges of S&OP (chapter 2.1.5) were presented. In this chapter those will be combined with ML value and benefits from chapter 2.2.4. Table 5 below summarizes these findings, where the most relevant factors from S&OP process are combined to ML value categories, which were also presented in chapter 2.2.4. The idea is to present how ML can be utilized in S&OP process based on the literature findings.

Different kinds of factors and things that are connected to this synthesis are described in the table 5 and those are categorized with the help of ML value categories. Table 5 includes different kinds of value factors and benefits that ML can bring into S&OP process from different perspectives, especially from S&OP enablers point of view. Table 5 also takes the S&OP challenges into account and describe how ML could possibly help in these challenges.

Table 5: ML utilization in planning process



In the table 5 can be seen how the S&OP part is connected to ML value factors and possibilities that ML can provide into the process.

First part in the table 5, ML ability to cover challenges in human work, is connected to S&OP enablers especially so that with the help of ML number of errors can be reduced (Skitka et al. 2000) which enables better decision making within a S&OP process. ML can also offer possibilities to scenario analysis, that were seen as a factor in S&OP process (Prokopets 2012; Pedroso et al. 2016). In general, ML makes possible to save in human resources or offer a possibility to allocate in different, more efficient way within an organization (Martinez-Lopez & Casillas 2013; Brynjolfsson & Mitchell 2017; Economist 2017; Sutton et al. 2018; Bolander 2019). Due to the automatized way of generating tasks, ML can help S&OP by taking some external data into account, which human wouldn't be able to do.

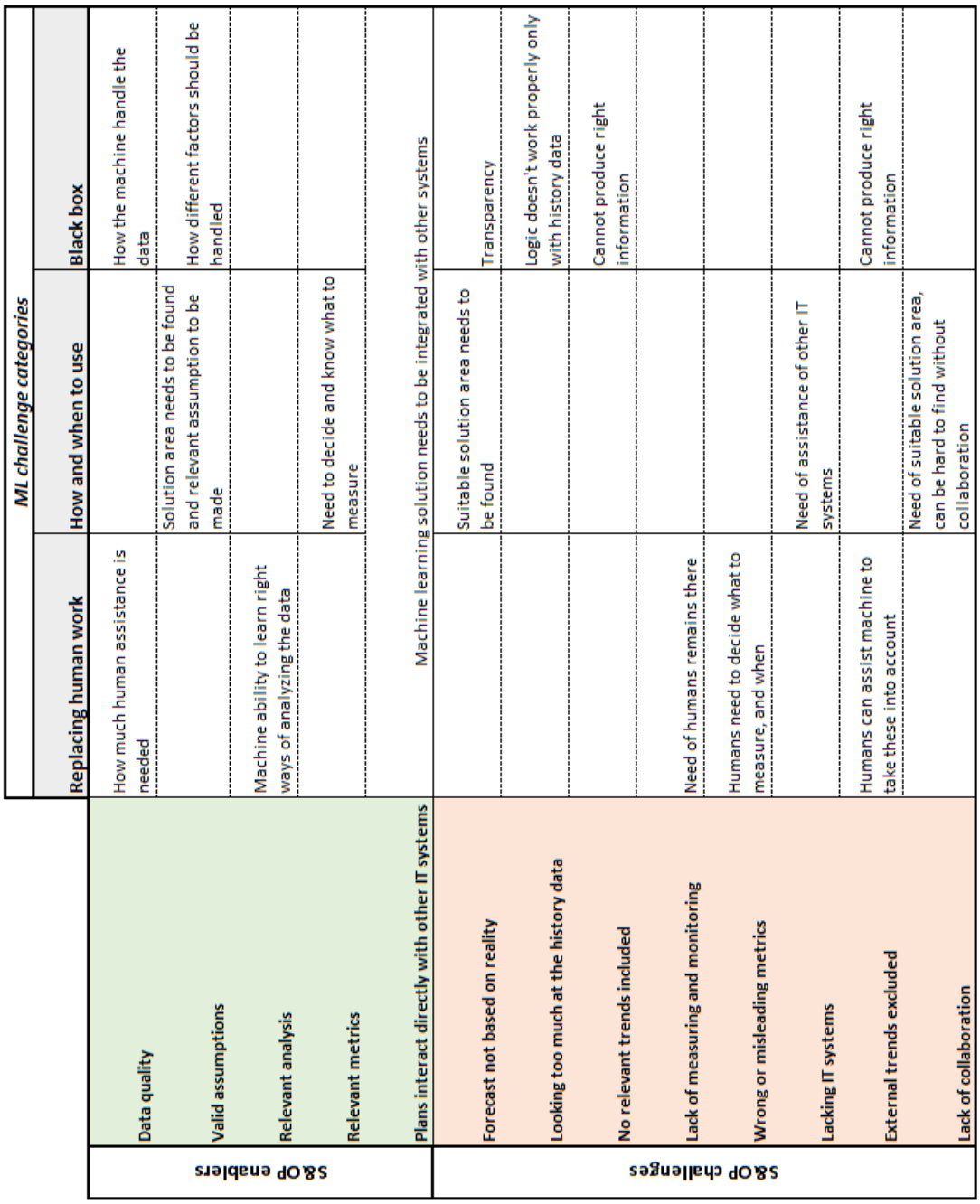
Second part, power of ML, might be the most important part of the connection between S&OP and ML utilization in the process. With the calculation power ML can answer to S&OP challenges by utilizing external data sources and through that also external factors can be included into process. That is seen as a traditional challenge in S&OP (Pedroso et al. 2016). From enablers point of view, ML can help an organization with scenario analysis, that was seen as an enabler (Prokopets 2012; Pedroso et al. 2016). Efficient decision making was also one of the enablers (Prokopets 2012; Pedroso et al. 2016; Wilson & Raman 2017) and ML can make this faster by providing more detailed information and faster. All in all, ML calculation power can offer value from many different perspectives and those are described in more detailed level in table 5.

Last part, possibility to financial improvement, was handled as a general outcome possibility, which is connected to all of the previous parts more or less. Other S&OP enablers/challenges that were presented in previous chapters are excluded from this section because those weren't seen as relevant factor as others, especially from S&OP and ML utilization perspective.

2.3.2 Risks of machine learning solutions in planning process

Risks of ML solutions in planning process are presented in a same way than the utilization possibilities in last chapter. Risks are presented (table 6 below) by taking the ML challenges from chapter 2.2.5 (figure 6) and these are connected to relevant factors from S&OP process (chapters 2.1.4 and 2.1.5). Main idea is to summarize what kinds of possible risks need to be taken into account when thinking about ML utilization in a planning process.

Table 6: Risks of ML solutions in planning



First part, replacing human work, is about how to handle human vs. machine learning tasks. It is clear that there is still a place for human work and those examples are described in more detail in table 6, but as an example from S&OP enablers, data quality is one of the most important enablers (Bower 2005; Prokopets 2012; Iyengar & Gupta 2013; Tuomikangas & Kaipia 2014) and there the human work is still needed to make sure the quality is on good enough level to reach the targets with ML solutions. Otherwise there is a risk that the ML leads to wrong conclusions and misleading analysis. That is also connected to one of the S&OP challenges, wrong or misleading metrics (Iyengar & Gupta 2013; Wilson & Raman 2017).

Second part, how and when to use, handles the problem of finding the right place to utilize ML. From S&OP challenges especially lack of collaboration (Pedroso et al. 2016) and lacking IT systems (Pedroso et al. 2016) can highlight this problem if they don't work within an organization. From S&OP enablers perspective valid assumptions (Bower 2005) and relevant metrics (Bower 2006; Feigin 2012; Prokopets 2012; Iyengar & Gupta 2013; Pedroso et al. 2016; Wilson & Raman 2017) are relevant factors also in ML utilization context. Valid assumptions need to be made in S&OP process (Bower 2005) and it is same with ML solution but those might be sometimes hard to find and implement. Relevant metrics is one of the most important enablers in S&OP process (Bower 2006; Feigin 2012; Prokopets 2012; Iyengar & Gupta 2013; Pedroso et al. 2016; Wilson & Raman 2017) and due to that ML solution needs to be support these.

Last part, black box, is about how the ML application handles the data and produce the outcomes. Lack of transparency is one of the possible risks in the utilization of ML, which is especially connected to couple of S&OP challenges: external trends excluded (Pedroso et al. 2016), no relevant trends included (Bower 2005) and forecast not based on reality (Bower 2005; Iyengar & Gupta 2013; Pedroso et al. 2016). From S&OP enablers side the risks are connected to data quality (Bower 2005; Prokopets 2012; Iyengar & Gupta 2013; Tuomikangas & Kaipia 2014) and valid assumptions (Bower 2005). These can be confusing if the ML logic is not clear enough for the users and can cause unnecessary problems in the utilization phase.

All in all, there are risks but if an organization keeps those in mind, those can be handled and solved, which makes the ML utilization even more possible for the companies. Some of the S&OP enablers and challenges (from chapters 2.1.4 and 2.1.5) were excluded also from this section because those weren't seen as relevant as others in this context.

2.3.3 Analysis of machine learning utilization within an organization

Emmanuel et al. (1990, pp. 14-15) divided decision making into two different types, programmed and non-programmed decisions. Programmed decision means that the decision making situation is well understood and is transparent. Non-programmed situation means that the outcome is not based on formal mechanism and situation is not that clear than in programmed decisions. (Emmanuel et al.1990, pp. 14-15) According to Korhonen et al. (2021) automatic solutions (in this case ML) should be used only when it's about programmed decisions because otherwise it can lead to undesirable situations. Problem is that it's often difficult to define which decision or task belongs to which category (Korhonen et al. 2021).

Nielsen (2018) presented a roadmap about the utilization of intelligent solutions. In the paper, context was business analytics as a whole but that can be used also in this research context. Steps are presented in figure 7 below.

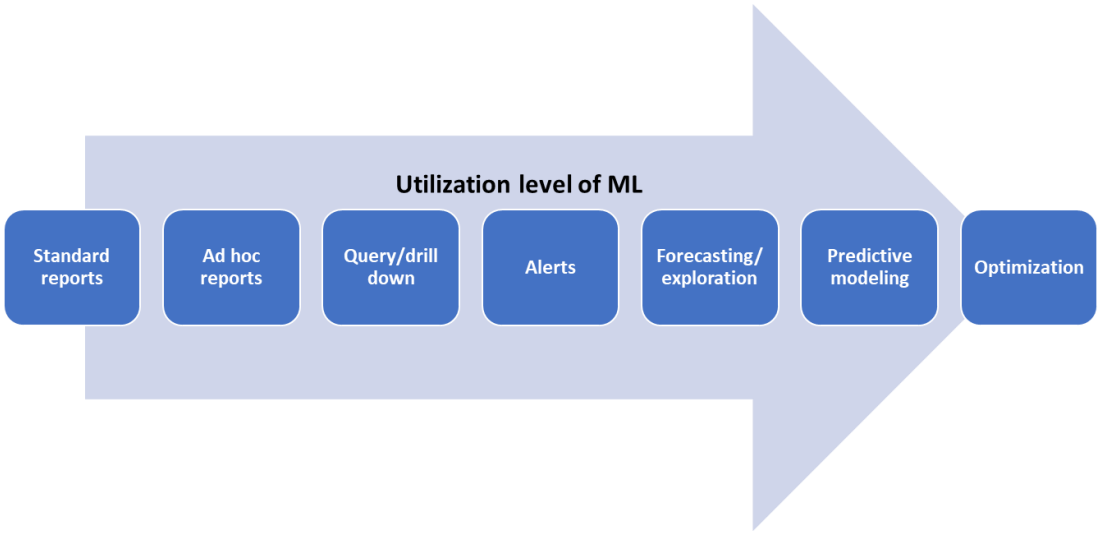


Figure 7: Utilization roadmap of ML (based on Nielsen 2018)

The main point of the roadmap (figure 7) is that more ML has been utilized, more advanced will be the outcome (Nielsen 2018). Organizations can use this kind of roadmap to find out where they are with their ML utilization and where should be targeted to take the next steps forward (on the figure 7). In addition to that, Bergmann (2020) highlights that often even big investment into IT systems are necessary to reach the targets but that also helps to take the next steps in developing the tools.

If combining the definition from Emmanuel et al. (1990) about the different kinds of decision/situation and the roadmap, it can be easier to track how to take the next steps closer to the final level (figure 7), optimization. According to Lichtenhaler (2018) data utilization, and especially the improvement of that, is a complex process, that includes many kinds of different parts, depending also on the size of the organization. Due to that, it needs to be incremental and well defined process overall.

3. RESEARCH METHODOLOGY

This chapter presents the research methodology of this thesis. First, the research methods are presented with the help of relevant literature and after that the research process itself including the research timeline is presented. Last parts present the data collection and analysis part of this project.

3.1 Research methods

This mono method qualitative research focuses on explaining how ML can be utilized in S&OP process. All the methodological choices of the study are presented in figure 8 below.

Figure 8: Methodological choices - research onion (based on Saunders et al. 2019, p. 174)

The 'research onion' presented in figure 8 above is a tool for helping in the research design planning process and it is divided into six different layers. Each layer describes a different phase of the research design starting from the research philosophy and ending in the research techniques. (Saunders et al. 2019, pp. 129-130)

First layer, research philosophy, means how the knowledge has been developed and also what kinds of assumptions and beliefs are behind that development (Saunders et al. 2019, p. 131). According to Saunders et al. (2019, p. 131) research philosophy also affects significantly other methodological choices in a research project. There are three different kinds of research assumptions which help to make differences between different philosophies: ontology, epistemology and axiology (Saunders et al. 2019, pp. 133-134). Ontology is connected to the nature of reality and how the research objects are handled, and also what is the actual research problem in the specific context (Saunders et al. 2019, p.133). According to Saunders et al. (2019, pp. 133-134) by epistemology is meant what is acceptable knowledge and what is not, which is also closely connected to what is legit, can be shared to others and in what kind of form the information can be presented. Saunders et al. (2019, p. 134) described axiology so, that it covers the role of values and ethics. There are also questions about the role of a researcher and how his/her opinions and beliefs affect the research.

In this research the research philosophy is pragmatism, as described in figure 8. Pragmatism targets a practical solution for some specific case or problem, as in this research. From pragmatism perspective true knowledge is something that supports successful actions in real life, and it handles reality as a complex place, that consists of practices and experiences. In pragmatist research the knowledge is built on different kinds of ideas, theories and concepts, and the research can include many different ways of building the knowledge. (Saunders et al. 2019, pp. 150-151)

Second layer, approach to theory development, presents how the theory has been developed during the research process and how it interacts with the conclusion and findings of the research (Saunders et al. 2019, p. 152). According to Saunders et al. (2019, p. 152) there are three different ways to view this part of the research: induction, deduction and abduction. **This research is an abductive research**, where the approach is focusing on moving between theory and data (Saunders et al. 2019, p. 156). Compared to induction, where the research moves from data to theory, and to deduction where the research moves from theory to data, in abductive research the goal is to use data to modify the existing theory and/or use the theory to explain the phenomenon, that has been recognized through data. (Saunders et al 2019, pp. 152-157) All in all, there are nuances from both, inductive and deductive approaches in abductive approach.

Third layer of the research onion (figure 8) is methodological choice. In qualitative studies, as this one, there are two different options for this: mono method qualitative study and multi-method qualitative study (Saunders et al. 2019, p. 179). **This research represents a mono method qualitative study** since there was only one technique of data

collection and analysis based on those interviews. This process is presented more closely in chapter 3.3. According to Saunders et al. 2019 (p. 179) this is a typical way of conducting a study where the interviews are held as semi-structured ones.

Fourth layer, strategy, is about the plan and ways of reaching the target. Research strategy is always linked to other sections of research onion, and it is a link between the research question(s) and the ways of conducting the study. Strategy is also closely related to the nature of the research, whether it is a quantitative or qualitative study, and the nature of the research problem. (Saunders et al. 2019, pp. 189-190) From strategy perspective, **this research is a single case study**. According to Saunders et al. (2019, pp. 196-198) the nature of a case study is to study a problem or phenomenon within an organization (in this case), it could also be a person or an event, for example, instead of an organization. In single case study the research focuses only on one single phenomenon instead of multiple ones (Saunders et al. 2019, p. 199).

Fifth layer, time horizon, is always cross-sectional or longitudinal and that means whether the study is a 'snapshot' from specific time or focuses on some longer time period (Saunders et al. 2019, p. 212). **This research is cross-sectional** because it focuses on a specific topic at a certain time or short period of time. According to Saunders et al. (2019, p.212) this typical for these kinds of qualitative studies which are based on interviews from short time period. Final layer, techniques and procedures, is just for presenting the specific techniques used in the research. Like already mentioned earlier, **this research is based on semi structured interviews and relevant literature of the research field**.

3.2 Research process

This master's thesis process began in October 2021 when the researcher and case organization hold the talks about the possible research topic. Common interest for the topic was there from both case company's and researcher's sides and the actual work with the study could began. The researcher has already worked at the company before this process so there was not so much need for familiarization of the research case itself.

After this, the researcher started planning the study and also go through the relevant research literature. During November and December the main focus was on the literature review which led to research interviews also during December 2021. Some finetuning were made also later to the literature review part but the main work on that area was done by the end of December. Preparations for the interviews began partly simultaneously with the literature review and when the targets were clear, interviews were held during December. More detailed introduction about these can be found in chapter 3.3.

After the literature review and interview round, actual work with the analysis and study outcomes could began. That was mainly made during January and also partly in February. Last actions were the conclusions of this research which were written in the end of February-March 2022 and at the same time some finetuning and finalizing were made. Overview of the total timeline of this research process is presented in figure 9 below.

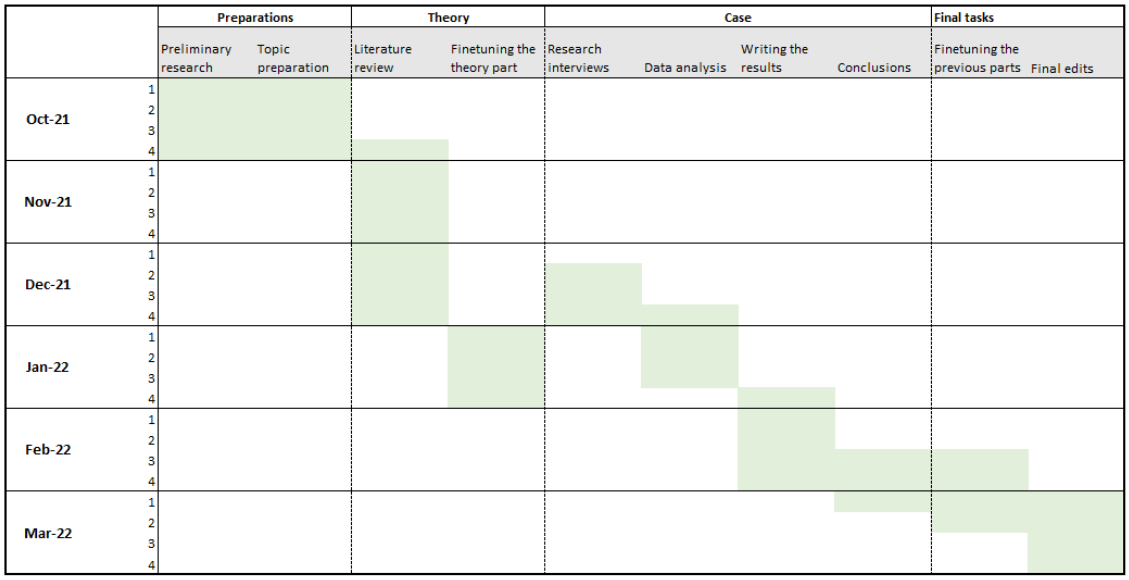


Figure 9: Timeline of the research project

Some modifications were made here and there to all of the parts, but the timeline presented in figure 9 is based on the actual actions made during the process.

3.3 Data collection and analysis

This section presents the data collection and analysis part of this research. First subchapter 3.3.1 presents how the data has been collected and the theoretical background of it, which leads to the second subchapter 3.3.2, where the ways of analyzing the data is presented and also the theoretical background of this.

3.3.1 Semi-structured interviews

Research data for empirical part was gathered through semi-structured interviews. According to Saunders et al. (2019, pp. 437-438) these kinds of interviews are typical for qualitative research and are utilizing some themes and key questions to lead the discussion during the interview sessions, but they also leave space for more free discussion and questions around the research topic.

Interview structure was partly based on literature review of this thesis and partly on researcher's previous knowledge about the target organization. Since the study was abductive from the approach to theory development perspective, the idea was to utilize theory but also left space for new knowledge, which can be utilized when comparing theoretical and empirical findings. According to Saunders et al. (2019, p. 444-446) semi-structured interviews are a good tool when trying to find explanations or evaluation for some case, in this case finding a suitable utilization possibilities for ML in S&OP and why that would be valuable for the process.

In total eight interviews were conducted, and the information about these (person title and duration of the interview) can be found from table 7 below.

Table 7: Details of the research interviews

Interview number	Person title	Duration (min)
1	Director S&OP	50
2	Manager, Development & Analytics / Business Controller	51
3	Group Controller	37
4	Business Analyst	53
5	Business Controller	42
6	ICT Service Area Manager / ICT Architect	58
7	Business Controller	40
8	HR Business Partner / HR Analyst	27

All the interviewed people have a key role within the organization's S&OP process, and they work in different business units, so the researcher was able to get a broad understanding about different parts and phases about the S&OP process and the ML utilization possibilities were possible to study.

These interviews were held as one-to-one interviews, which is typical way for non-standardized semi-structured interviews (Saunders et al. 2019, p. 441). Due to the existing circumstances, interviews were held in Microsoft Teams and not face-to-face, but according to Saunders et al. (2019, p. 442) this is also one of the typical ways of conducting a research interview. The interviews were also recorded, and large notes were created

afterwards based on the recordings. That enabled the researcher better focus in the interview session.

The content of the interviews, like earlier mentioned, was based on literature and previous experience about the organization. The interview structure is presented in appendix A.

3.3.2 Data analysis

Basis for the analysis of the textual interview data was the notes created from the interview recordings. According to Saunders et al. (2019, p. 638) it is good enough for the analysis and due to that no full transcriptions were created from the interviews. That wouldn't have brought more value into analysis and the research itself compared to time.

In this research data was analyzed through thematic narrative analysis, that is one of the ways of analyzing qualitative data (Saunders et al. 2019, p. 675). In thematic narrative analysis the focus is on the themes, taking into account the context and chronological order of the topics. That means thematic narrative analysis is a combination of thematic analysis and traditional narrative analysis. (Saunders et al. 2019, p. 675)

This way of analyzing the data was chosen due to the choice of semi-structured interview method where there might be some variation in the content of the interviews and that's why the context and chronological order play an important role. Narrative analysis is also good from the approach to theory development point of view, where the abductive approach moves back and forth between theory and findings based on the research data.

Previous narrative analysis combined to thematic analysis produces the outcomes in this research. Important factor in thematic analysis is the coding the data (Saunders et al. 2019, p. 653) and that means that the data units has been categorized with the help of codes and so the similar meaning from different interviews can be found (Saunders et al. 2019, p.653). In addition to the previous ones, the role of interviewed people has been taken into account, due to the different positions the people have in the organization. Some of the people operate on more strategic level and others closer to operate level. That can also affect the answers and especially the level of details.

4. RESEARCH FINDINGS

In this chapter the empirical findings of the research are presented. The chapter is divided into subchapters (4.1 – 4.3) with the help of research questions. Last part, chapter 4.4, includes ideas how the findings could be implemented in real life and what kinds of actions that could mean. Content of this chapter is based on the research interviews. Table 8 below summarizes the research findings, and those are presented more closely in the next subchapters (4.1 – 4.3).

Table 8: Summary of the research findings

RQ1: Utilization possibilities	RQ2: Value of ML	RQ3: Challenges regarding ML utilization
- Three different options found	- More options for the whole planning on general level	- Challenges categorized into different themes
- Option 1: Demand planning of a business area	- Data from outside the current process and organization	- 1: Lack of earlier experience
- Option 2: Logistics optimization	- Better and more effective results	- 2: Qualitative and experience-based data
- Option 3: Supply planning of a business area		- 3: Reaching enough improvement with ML

4.1 Machine learning utilization in the organization

This subchapter will present the ML utilization possibilities within the case organization. The chosen possibilities are based on the research interviews and the findings from there. Table 9 below summarizes the options.

Table 9: ML utilization possibilities within the S&OP process

Option 1: Demand planning	Option 2: Logistics/supply optimization	Option 3: Supply planning
- Focus on one business area	- Focus on logistics/supply of one business area	- Focus also on one business area
- External data included	- Complex problem	- The exact use case would need more research
- New internal data included	- Would need further study how to proceed	- Is the real interest there?
	- ML maybe not the correct tool for this problem?	

The content of the table is described more closely in the next subchapter (4.1.1 – 4.1.3). The three options presented, are the relevant options for ML utilization based on the interview material. Other smaller options were also found but these are excluded from this due to the fact that only these three were seen as relevant and interesting ones.

4.1.1 Option 1: Demand planning of a business area

This option handles the demand planning process of one specific business area. This business area has also some previous experience from ML utilization in their supply planning process, why they also seemed to be interested in ML possibilities in their demand planning side. The main idea of this solution option is presented in figure 10 below.

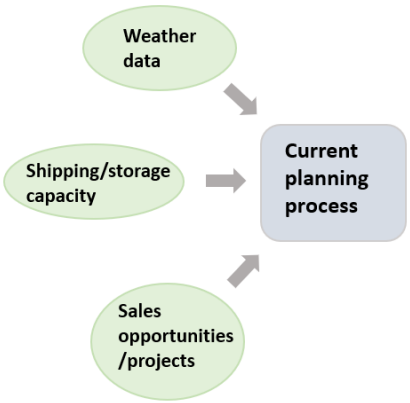


Figure 10: Improvements into demand planning process

Main idea in this option would be to add **external data sources into planning process**, which is something that hasn't been possible with current tools and manual work. Based on the interviews, the most important external data source would be weather data. The business is very seasonal, and weather significantly affect the demand, especially the timing of the demand. So far, this has been handled by making some assumptions based on earlier experience and keeping some margin of certainty in demand plans. This way of planning hasn't always been a huge success according to an interviewee.

Based on the interviews it is not so clear and straightforward what kinds of factors from weather should be taken into demand planning and ML, so this needs still further discussion about the exact details. All in all, some weather data should be taken into process and the weather affect especially the timing of the start of the high selling season. Also, it seemed to be important to research and discuss further how the weather affect different kinds of customers, that are professional growers and wholesalers. ML solution would also be able to cover this issue.

One factor, which is a part of the process, is **the sales opportunities and projects**. According to one interviewee this is an area where the current tools are not optimal, and

the projects are hard to include into demand plan. The nature of these sales projects is a bit different compared to normal demand. Projects are often bigger and unique, and they realize or not into demand but there is no halfway existing. The handling of those have been difficult but ML solution could help to also take these into account when planning the demand. Based on the interviews, at least the relevant data is available and now the organization should find a way to utilize it.

Based on the interviews **shipping and storage capacity are also important parts when planning the demand**. It is combination of external and internal data, but so far it has been difficult to include into process. This could be important data factor for ML solution and also a good opportunity to increase the performance of the whole demand planning. According to an interviewee this will still need further discussion about the exact ways of handling the data and shipping capacity as a part of possible ML solution.

The targets from ML utilization were on general level clear based on the interviews. Possibility to different scenarios in demand planning was the number one target according to one interviewee. So far, the organizations have always had only one demand plan and sometimes there are situations where the scenarios could provide better insight into the future. This fits really good also for ML utilization topic.

Other important target was based on the interviews to reach better forecast accuracy in the demand planning (and also in supply planning). Forecast accuracy is the most important metrics and monitoring tool, and it is combination of all the factors within the process (e.g. shipping capacity, sales projects and opportunities included, weather data). ML could fit good for these kinds of targets according to an interviewee.

Another, a bit less important target was to have insight into weekly level demand. Based on the interviews the demand planning is created on monthly level currently but sometimes it would be valuable to also have the weekly level view. Now the problem has been that it would take too much time compared to value to create the weekly level plans. ML solution could help the organization also on this topic.

Based on the interviews and also the other options (chapter 4.1.2 and 4.1.3), **this would be the first area to utilize ML,** so the implementation plan (chapter 4.4) will include further actions around this option and topic. The potential of this option is clear and also the interest of the interviewees seemed to be real towards trying ML in this area, but like already mentioned further discussion about the details are still needed but those will be out of the scope of this research. Other options (presented next) weren't so clear based on the interviews.

4.1.2 Option 2: Supply and logistics optimization

Supply and logistics optimization of one business area were highlighted in two different interviews. According to an interviewee this area has huge potential to improvement of financial results, and ML could be one potential tool for that. By logistics optimization was meant how to organize the logistics of final products and raw materials so that it would be as profitable as possible. This is connected to supply planning and whole S&OP so that also the supply planning could be organized in an optimal way to fulfill the customer needs.

One of the interviewee thought about the ML possibilities, if it could recommend where the product should be produced, and on the other hand, from which location it should be delivered. All in all, this seemed to be pretty complex and large area to be covered. Based on the interviews, there would be a lot of history data that ML could take as an input, and also information about the logistics history.

Instead of that, there was left a lot of empty areas and questions more than answers regarding this option. Even though, the idea and especially the target seemed to be pretty clear, **there should be done additional interviews and research about this, and that's why this will be out of the scope of this research**. This will continue outside of this project because ML were seen potential tool for this kind of problem and the nature of ML could be valuable. It is also known that there is also other technologies and tools for optimization problems, and this should be researched if ML was the right tool to proceed.

All in all, the further discussions around this topic will continue but will be outside of the scope in this research also due to timeline and the fact that this research is based on the interview round held earlier.

4.1.3 Option 3: Supply planning of a business area

The third option recognized from the interviews was connected to supply planning of one specific business area. The interviewee seemed to be interested in ML utilization in the supply planning of the business area from which the interviewee was responsible. There was also potential in developing the planning of the area, for example taking external data into account, which hadn't been possible with current process and systems. Also the level of intelligence was on a low level currently and ML was seen as interesting possibility to improve this whole process.

Even though there was seen potential and some ideas regarding the process, the concrete ideas were missing a bit, for example compared to option 1 presented in chapter

4.1.1. Scenario analysis possibility, more automatized way of creating the forecast/supply plan and more accurate plans were the main targets which the interviewee mentioned in the interview. These kinds of things connected to the nature of the process build a good opportunity for ML utilization in general. In addition to the interview findings presented earlier, this area and its possibilities raised from two other interviews also. Based on those, this could have a lot of potential for solutions like ML even though the exact knowledge and experience was missing as mentioned already earlier.

All in all, this area was seen potential and realistic for ML utilization, but it would need more discussions about the exact details which were missing based on this one interview round. **That's why this option wasn't selected as a recommendation in this master's thesis research**. Instead of that the discussion around this topic will continue but it will out of scope from this research's perspective.

4.2 Machine learning value for sales and operations planning

This chapter presents the value factors found regarding ML utilization within the organization's S&OP process. Due to the lack of previous experience of ML usage, these value factors are more like possible factors and outcomes than the actual reality, due to the fact that there haven't been any solutions before. Summary of the findings regarding value is presented in table 10 below. Findings regarding value are based on the discussion with the interviewees. Some of the factors came directly from the interviews and some are interpretations of the researcher based on the discussed topics. The next subchapters (4.2.1 – 4.2.3) will present the findings from table 10 below more closely.

Table 10: Possible ML value for S&OP process

Improved S&OP process in general	External data and data outside of the current process	Better results from S&OP
- Scenario thinking possibility	- Potential new and external data sources found	- Improved forecast accuracy
- Planning based on facts	- More angles into S&OP process	- Better adapted to possible changes in S&OP or business in general
- Reduced number of errors	- Suitable way of utilizing ML	- Better insight into long term view
- More automatization within S&OP		
- More systematical process		

In general, some of the interviews seemed to be more useful for this research topic and especially for the potential ML value topic than the others. That could be explained by different roles of the interviewees, or some business areas had more input to this topic than the others. More detailed findings regarding potential value are presented next.

4.2.1 General level findings regarding machine learning value

General level findings and the number of mentions regarding those are presented in table 11. In general there were good discussions with the interviewees about the possible value that ML could bring into S&OP process and in what kind of phases within the process ML could bring solutions and value. A couple of the interviewees were working on a higher role within the organization and that could be recognized also from their answers and thoughts, for example they highlighted the importance of **more fact based process** (mentioned in three interviews, table 11) and **scenario possibilities** (mentioned in four interviews, table 11).

Table 11: General level findings regarding value with the number of mentions

Topic	Number of mentions
Scenario thinking	4
More detailed insight	3
Reduced number of errors	3
Based on facts	3
More automatized process	2
More systematical process	1

More fact based process would bring value compared to current one. According to one interviewee there are still some assumptions included into planning process and those might be based on no facts but more on feelings of the responsible people. That can lead to budgets and forecasts that go wrong and they know it already beforehand. **Scenario thinking** was the most common topic in this context. According to one interviewee who worked on higher level in the organization they are creating some high level scenarios continuously, at least in one business area, but that wasn't visible in any other answers of the people from more operative roles. The general feeling in the interviews where this was mentioned, was that scenarios would bring clear value and more visibility into planning process, in every area of planning, especially in finance and supply. ML was seen as a potential and useful tool from scenarios point of view, and it could make the scenario creation easier. Now it was seen pretty hard and painful process next to other important actions.

ML was also seen as a tool which could provide **more detailed insight** (mentioned three times, table 11) into the S&OP process. In demand planning of a business area (chapter 4.1.1) ML was seen as a way of getting also weekly level insight into plans compared to current system which offered only the monthly level. With current solutions monthly level was good enough when comparing the amount of work and value. **Reduced number of errors** (mentioned 3 times, table 11) raised also from the interviews as a potential value factor. Especially in the field of finance planning in fixed costs the current forecasting is made pretty manually which can lead to errors and wrong forecasts. ML could solve this by automatically providing numbers based on predefined factors and so help people to reduce the number of errors during the process.

More automatized process (mentioned 2 times, table 11) was also one target which could be reached with ML utilization based on the interviews. Current process was seen as too heavy and time demanding due to many different reasons. ML could have the needed potential to help in some specific tasks and process phases so that the people wouldn't need to use so much on the planning activities. One interviewee from higher level commented this topic so that there would be a lot more important tasks which cannot have full focus now when the forecast process is running due to the fact that it takes a lot of time from the people. Another interviewee (in an operative role) supported this by saying that it is sometimes frustrating to create fixed costs plans from zero and based on no actual knowledge about the relevant and affecting factors.

More systematical process (mentioned only one time, table 11) raised in an interview with a higher role employee who commented ML as one of the missing tools/solutions for more systematical process. According to the interview the current process consists of too many ways of doing the same thing, and ML could provide help in this. Fixed costs was again the example from the really manual forecasting process where the results vary too much compared to actual costs. As a summary there would be a need to standardize the processes and ML could be one part of this big picture.

4.2.2 Data from external sources and outside of the current process

Table 12 presents the findings regarding external data and data outside of the current process. Based on the interviews by utilizing these findings, new value into S&OP can be created and more improved process would bring value into the whole business. Especially interesting topic in ML would be the potential of using external data better and in more intelligent way. The most important finding in this section was to take **weather data better in use** (mentioned five times, table 12). Depending on the business area,

weather seems to have a significant impact on the business, especially demand and supply planning.

Table 12: External opportunities and the number of mentions

Topic	Number of mentions
Include weather data into process	5
More customer related data	2
Shipping/logistics capacity	2
Other external factors	2
Internal data into better use	1

Weather data is something that should be utilized better according to many interviewees. Weather itself impacts the business in a way or another depending a bit on the business area. During spring the most significant impact of the weather is when the highest demand starts and that of course also impacts the supply side. So far, the biggest issue regarding that has been finding the way of utilizing this data. **ML was seen a very potential tool for taking this into account** and so to be included into S&OP process instead of some guesses and assumptions around the topic. One business area business controller said that they have created some kind of mathematical model, that includes weather information and that has made the S&OP better process within this business area. Otherwise, there haven't been any ways of taking this into account in a proper way.

One of the findings was related to **customer data and its better usage with the opportunity that ML offers** (mentioned two times, table 12). Based on two interviews there are some customer related data that could improve the demand planning side, but with current systems and limitations it hasn't been used in actual planning. One interviewee took customer campaigns as an example which could be utilized better, for example with ML. This is one of the concrete examples how ML could make the process more effective. Related to this there are also **other internal data** (mentioned once, table 12) that is not utilized because humans can handle only limited amount of data, for example sales opportunities.

Another data related topic was **shipping/logistics capacity** (mentioned two times, table 12). According to an interviewee that is something which would be already available in the current systems, but it hasn't been utilized in supply planning. It has seemed to be pretty difficult to include into current S&OP process and the organization has tried to take it into account based on some feelings but not the actual data. Interviewees had the

assumption that ML could be solved this with its data processing capabilities and calculation power and so create additional value for the whole S&OP process.

Also **other data** (mentioned two times, table 12) and value factors were found based on the interviews. According to one interviewee really much more could be reached if ML could utilize, for example some public economy related data which clearly impacts the business. The question is how to handle this but, in any case, this would be really valuable based on the interviews.

All in all, based on the interviews there would be a lot of potential to utilize better different kinds of data sources with the help of ML and so make the whole S&OP more powerful for the business. All of the potential sources for the value and ML utilization should still be proved in real life but this gives at least a good starting point for that and for further research around the topic.

4.2.3 Better results in sales and operations planning

Findings regarding value and results in S&OP process are presented in table 13. In general, a couple of the topics raised above others and some things were mentioned only once during the research interview round. In this section there was no visible differences in answers between different kinds of roles of the interviewees. Highlights of the findings based on the interviews were **forecast accuracy** (mentioned four times, table 13), **better adaption to changes** (mentioned three times, table 13) and **longer term view** (mentioned twice, table 13).

Table 13: Improvement in S&OP results and the number of mentions

Topic	Number of mentions
Forecast accuracy	4
Better adaption to changes	3
Longer term view	2
Initalization of the forecasts	1
More forecasting	1
More quality process	1
Handling the information better	1

Forecast accuracy was the most important value factor in this context regarding possible utilization of ML. The director of S&OP thought that there is still possibilities to improve the accuracy and that could lead to improvement of the whole business. Based on the interviews the general feeling was that the forecast accuracy is a combination of many different things and tasks, but ML would be very potential tool for reaching these

targets. Bergmann et al. (2020) presented also that improved forecasting accuracy is one of the main benefits in ML utilization.

In three interviews **better adaption to changes** raised as a topic when discussing about potential value factors. With the current systems there have been some difficulties, for example with the changing item names and that kinds of things. The general feeling regarding adaption was that ML solution and its utilization could prepare the organization better to changes in the environment and business in general. Concrete examples from that were missing due to lack of earlier experience from the topic but this seemed to be the general feeling.

During two different interviews **better insight into long term view** was in the discussion. There has been troubles to find a way to produce numbers regarding S&OP into long term view. Now the visibility has been pretty good into certain time range, but the longer view has been left on the background due to the limitations of the current ways of doing the planning. The hope regarding ML utilization based on the interviews was that it could help the organization to also have the longer term insight and maybe some rolling forecast that can fix itself based on the defined factors and data.

Other factors regarding this chapter were **initialization of forecasts, more forecasting in general, more quality within the S&OP process and handling all the relevant information in a proper way**. These seemed to be not so important and significant findings since they were all mentioned only in one interview. Initialization would mean that ML could provide a basis of S&OP numbers so that the process could be a bit smoother, and it is closely connected to more automatized process from previous chapter 4.2.1.

More forecasting in general, more quality and handling all the relevant data/information in a proper way can be connected to ML utilization but those are more on general level things that are connected to process improvement as a topic in general within the organization. Instead of that, those were mentioned in the interviews in this context, so they haven’t been excluded but are not handled more in detail.

Even though there are many different kinds of findings regarding ML utilization value, those are pretty hard to proof in real in this research scope. Instead of that, all of them are telling the potential of ML utilization within the organization. The realized value can be seen first after the implementation when the solution is up and running as a part of the whole S&OP process.

4.3 Main challenges regarding machine learning utilization

This subchapter presents the findings regarding possible challenges related to ML utilization within the organization. This is divided into smaller sections by main themes found based on the research interviews. Table 14 below presents the summary of the possible challenges of ML utilization based on the research interviews.

Table 14: Summary of the ML challenges based on the interviews

Lack of earlier experience	Qualitative and experience-based data	Reaching enough improvement
- No previous experience from ML solutions	- Qualitative data within the S&OP process	- Manual involvement
- No knowledge about the ML possibilities	- Assumptions based on long experience in the industry	- Effort vs. value
- Enough interest within the organization?	- Qualitative external factors (e.g. politics)	- Complex problems
- Connection of lack of experience and complex problems to solve with ML	- Lead to frustation if everything cannot be included	- What can be achieved?
- Data quality issues		- Enough interest within the organization?

In general, some challenges were mentioned in every interview, maybe due to some suspicions towards ML because only one of the interviewee had some experience from it. In the interviews the researcher tried to lead the discussion (regarding possible challenges) around the general ML challenges presented in chapter 2.2.5 (figure 6), so that the discussion wouldn't go into wrong or irrelevant direction. Instead of that, the next subchapters are not divided into themes directly based on figure 6 because it wouldn't have been valuable to do that if the interview material doesn't suit into those directly. Those themes were in any case included into analysis when trying to find interesting themes from the material. Also the content of table 6 (chapter 2.3.2) was included into analysis process. Next subchapters, 4.3.1 – 4.3.3, presents the findings regarding challenges (summarized in table 14) more closely.

4.3.1 General level challenges

Table 15 presents the number of mentions regarding general level challenges. In general, the interviewees who were working in an operative role within the organization mentioned the data quality issues and on other hand interviewees in higher role mentioned

the possible challenges regarding **change management** (mentioned two times, table 15). Change management is connected closely to lack of earlier experience and the fact that it could somehow change the way of doing the S&OP planning. **Data quality was also one of the efficient S&OP process enablers** (presented in literature review) in an organization (Bower 2005; Prokopets 2012; Iyengar & Gupta 2013; Tuomikangas & Kaipia 2014) and that was mentioned also in four interviews.

Table 15: Mentions regarding general level challenges

Topic	Number of mentions
No previous experience	6
Data quality issues possible	4
Suspicions towards ML	3
Change management	2
Lack of interest	2

The starting point for ML utilization is that there is almost **no previous experience** from ML utilization (mentioned in six interviews, table 15) or solutions. One business area makes an exception in this but there the utilization is limited only to production/supply planning and its optimization. In general, based on all the interviews there is not much competence or experience from these kinds of solutions within the organization. Due to that there are not so many prejudices, but it is also a challenge – based on the interviews the people didn't know beforehand so well the possibilities of ML and in what kinds of cases it could be suitable, which lead to **some suspicions towards ML** (mentioned three times, table 15). Other challenge which comes with the lack of earlier experience is the fear of heavy and expensive implementation process regarding ML. This challenge regarding implementation came up especially from the people who work in higher role.

Due to lack of earlier experience, ML would be a new tool for the organization (except for one business area) and that's why the interest would need to be there, because otherwise the implementation and the project could be too heavy. That was also the feeling based on two interviews, there was some **worries about the lack of real interest** (mentioned 2 times, table 15) and the assumption was that not everyone would be interested in this because so far, the organization has been able to handle the S&OP process without it. So, the challenge is about possibly negative preconceptions.

Based on the interviews most of the possible solution areas and ideas are connected to pretty complex problems, for example logistics/supply optimization, supply planning and demand planning (presented more closely in chapter 4.1). Many interviewed people said

that there are limitations, when talking about ML solutions and competences related to those. Due to that external help would be needed and that may be more expensive and raise some barriers between the ideas and actual implementation of a real ML solution. The first research question also handled this area but from different perspective. Based on all the interviews, it's not so straightforward to find a suitable area to utilize ML and that was seen also as a challenge.

Other general worry was the **data quality** (mentioned in four interviews, table 15), which were seen as a problem for possible ML solution. One of the interviewee said that there are always issues with data quality in production planning and that might be also a challenge for this kind of intelligent ML solution. Also other one mentioned that this challenge will appear and needs to be handled in a way or another. All in all, this challenge raised during many interviews when discussing about the planning process or possible machine learning utilization.

4.3.2 Implementing the qualitative and experience-based data

Table 16 summarizes topics and the number of mentions about this subchapter's topics. Every topic from the table wasn't mentioned directly, especially experience-based knowledge about the business, but that was the interpretation of the researcher based on what the interviewees said and told.

Table 16: Mentions related to qualitative and experience-based data

Topic	Number of mentions
Laws and regulations	3
Politics	3
Experience-based knowledge	2

Based on the interviews there is a lot of data that might be difficult for ML to take into account. For example, **different kinds of laws and regulations** (mentioned three times, table 16) affect the whole business at least based on the interviews. The question is, how these kinds of things should be handled if it is known that they affect planning but are not quantitative data that ML could take as an input.

Another same kind of data challenge that raised in the interviews was the **politics** (mentioned three times, table 16). One interviewee mentioned that especially some business areas need to follow closely what happens in the field of politics, because that have direct effect on the business and also the planning. In energy business the effects can be crucial especially in longer term, and in these plans, politics - laws and regulations – might

have vital impact on the business and planning. So, the challenge for machine learning is that this kind of information is hard to include into process.

Other interviewee mentioned that there are a lot of people included into S&OP process that have already a long experience from the business and can make **valid assumptions based on their own experience** (mentioned two times, table 16). This kind of experience based information can be a challenge for solution like ML to take into account and without it, the question is whether the targets are able to be reached in S&OP. In addition to including this information, it can also help to notice the surprises and exceptions in demand or supply, or also in financial side, and that was also one worry according to one interviewee.

It was seen as a frustrating if it is known that the ML solution cannot handle all of the relevant information and that can then lead to bad results. This is also closely connected to the topic of the next subchapter 4.3.3.

4.3.3 Reaching enough improvements with machine learning

Related to different kinds of data and some suspicious towards ML there was some discussion whether enough process improvement and results can be reached with ML solution. Table 17 summarizes the topics and the number of mentions around this theme.

Table 17: Mentions related to reaching the improvement of the process

Topic	Number of mentions
Manual involvement still needed	5
Maintaining the solution (data changes etc.)	3
Implementation phase	3
Amount of work vs. value	2

Many of the interviewees had the opinion that only machine learning itself cannot reach good enough results without **manual involvement of humans** (mentioned five times, table 17). Then the discussion is about machine vs. human work, which is a traditional topic in this context. Other general discussion topic was the required amount of work to get a suitable ML solution for some specific phase within S&OP process. Challenge is to decide and know **how much work (costs and time) is valuable to do compared to ML can potentially bring** (mentioned two times, table 17). Also the **maintenance of the solution** (mentioned three times, table 17) might require work, maybe also external work, and it is not free. This aspect came especially from the higher level interviewee who are responsible of the financial outcomes in their business area. This is also connected to earlier presented topic about the lack of interest within the organization.

Other challenge that raised from the interviews was the calculation logic and its complexity in ML. It was seen challenging to understand how the logic take into account different data and factors, and how they impact on the results. One interviewee pointed out that this can lead to another challenge of developing the system without full understanding of the ML process. In addition to that it was seen important that ML solution would be understandable for the users, so there is clearly some kind of conflict. This theme is also connected to maintenance of the solution, for example changing data can cause issues to that according to one interviewee.

One concrete example regarding the possible complexity of data processing came from one interview where the interviewee thought about weather as an external data input for ML. There can be multiple different kinds of aspect in weather data that can be included, and it would be really useful to understand how the ML solution handles those different factors. A couple of the interviewee thought about **possible heavy implementation process due to these reasons** (mentioned three times, table 17).

This is also connected to earlier point that there is not so much experience about ML in the organization and some lack of knowledge about the topic. Another example was the changing SKUs in the systems which can cause issues for the logic. Outcome of the challenges in the systems can be that the usage is not seen useful or valuable according to a couple of the interviews. There was already earlier experience about this kind of low usage regarding systems and then expensive solutions are without users in the organization.

4.4 Implementation plan

Next steps regarding ML utilization are presented in figure 11 below. Every step after results presentation is outside of the scope of this master's thesis study. Due to that the timeline (presented in figure 11) is still a bit unclear and will become more exact as the project continues.

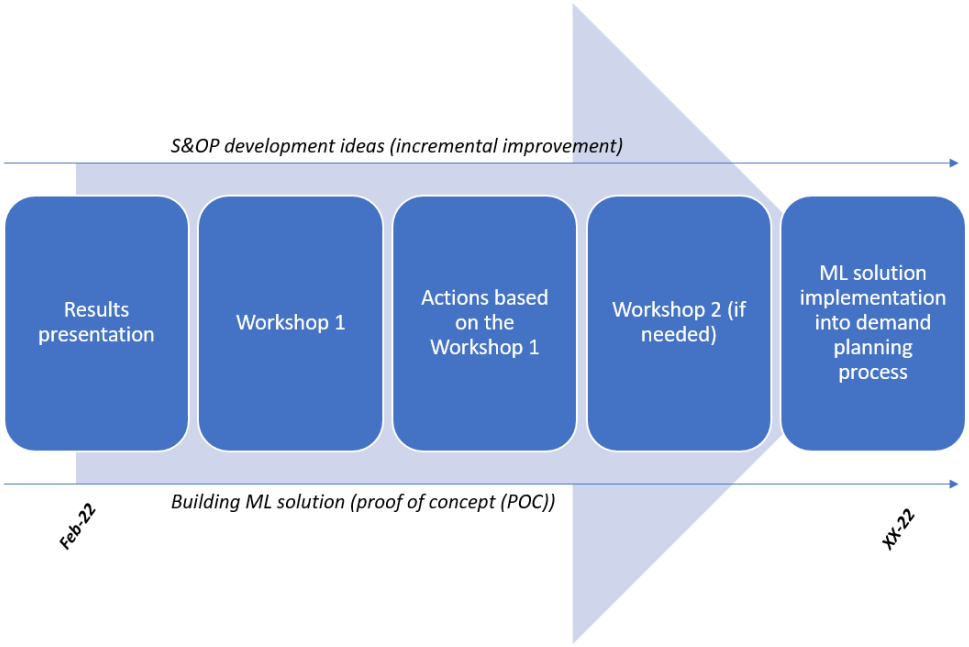


Figure 11: Next steps of implementation

First, the idea is to present the research findings for the case organization. These will be presented based on the findings (chapters 4.1 – 4.3) connected to relevant literature findings (chapter 2.3) in a same way than in this study. As mentioned earlier further steps are already outside of the scope of this thesis project but in this chapter is described the plan what will happen during the next times.

After the results presentation there will be a workshop around the ML solution idea (described in chapter 4.1) with relevant people from the case organization. The idea of the workshop is to further discuss about the demand planning (of one specific business area) and the relevant factors within and behind the process. The goal is to have more clear idea about the ML solution – what factors and data to take into account, for example internal and external factors to be considered.

After that the actual work with the ML solution development can start and further workshop will be organized if necessary. Basically if some part/s need still to be discussed on more detailed level, the workshop 2 will be organized. After that, at some point, the ready ML solution will be implemented into demand planning process and the needed actions are still unclear at this point.

As also described in figure 11, there was some general S&OP development ideas which will be slowly implemented during the next months, but these are not in the scope of this research. Also the ML solution development is an ongoing project starting from the workshop 1 and the development will continue as long as needed to reach satisfied results with it.

5. DISCUSSION

This part presents the discussion between theoretical and empirical part of this study. The chapter is divided into subchapters with the help of the themes of research questions. The idea is to show how the earlier findings based on interviews are connected to theory part and also the possible differences between those.

5.1 Improved demand planning with the help of machine learning

Chapter 4.1 presented the different utilization possibilities found based on the research interviews. Those possibilities were demand planning of a specific business, supply/logistics optimization of a business area and supply planning of one business area. All of the ML utilization possibilities included potential and also challenges as presented in chapter 4.1. Proposal of this research was to go on with the demand planning even though the actual implementation will be out of the scope of this study. This chapter will also exclude the other options since those weren't seen as relevant as the demand planning for further research.

The outcome from the empirical part that ML could be utilized in demand planning process of a specific business area (chapter 4.1). That would mean that ML could make it possible to consider some external data sources and factors that haven't possible to include into process so far. In addition to external data, there would be a possibility to include more data than before if ML will be utilized. According to Alpaydin (2014, p. 2) ML can handle large amounts of data and accomplish tasks powerfully. This supports the idea about the utilization area so that more data (internal and external) would be taken into account. ML has also a lot of calculation power (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020) which gives the idea even more support.

Theoretical part presented the enablers of an efficient S&OP process (table 1, chapter 2.1.4) and especially fact based process (Bower 2005; Iyengar & Gupta 2013) and scenario analysis possibilities (Prokopets 2012; Pedroso et al. 2016) could increase with the help of ML and so improve the whole process. Even though the current S&OP seemed to be already a fact based process, but there were some assumptions included which might decrease the level of facts and ML could provide opportunities for fact based planning and decision making. Lack of scenario analysis possibilities was one of the weaknesses in current demand planning process but with the calculation power of ML (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020) that would be a lot

easier and more straightforward to create different demand scenarios. The impact could be seen also in other parts of S&OP since demand and supply planning are closely connected to each other's (Feigin 2012, p. 15).

In this demand planning area ML could also answer to some traditional S&OP challenges (table 2, chapter 2.1.5). Especially no relevant trends included (Bower 2005) and external data excluded (Pedroso et al. 2016) would be relevant topics in this case. ML can utilize its calculation power (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020) and use more data sources to proceed the algorithms, which can help the organization to include some relevant trends that impact on the business and demand planning. Same goes for the external data, which can be hard to include into process if the human tries to build the complex data processing tool without some intelligent tool.

One major benefit of ML was savings in human resources (Martinez-Lopez & Casillas 2013; Brynjolfsson & Mitchell 2017; Economist 2017; Sutton et al. 2018; Bolander 2019) but that didn't raise from the research interviews. In the opposite the worry was how much human assistance the demand planning would need if the ML solution would provide help into process. But the fact is that at the same time, the solution would release time for other important actions and conclusions regarding the demand planning for humans within the organization.

All in all, theoretical findings regarding ML utilization support the demand planning topic of the organization from different angles. All of the theoretical findings didn't raise from the interviews, but it is understandable due to limited time and topics in interview sessions. That might be the reason also for the fact that theoretical insight are on more detailed level than the empirical ones, but instead of that both parts have thing in common and also some differences.

5.2 Clear potential between planning process and machine learning

Chapters 2.2.4 and 2.2.5 presented the ML value and challenges (tables 3 and 4) from theoretical perspective. On the other hand chapter 4.2 presented the empirical findings in the field of ML value for S&OP process. There was found some clear connections between the theoretical and empirical part. On the other hand, some of the findings don't support each other's or are even the opposite. This can be explained by the lack of experience from ML utilization when looking at the interviewed people who could only their assumptions and feelings around the topic.

If looking at the ML value and benefit in theoretical part (table 3) reduced number of errors (Skitka et al. 2000) and improved decision making process (Martinez-Lopez & Casillas 2013; Bolander 2019; Bergmann et al. 2020) are good aligned with the empirical findings in chapter 4.2.1 (table 11). Even though the most common one from the empirical findings, scenario thinking didn't come directly up in theoretical benefits, it can be very well connected to calculation power of ML (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020) since scenario creation demands some power from the calculation system, in this case ML. None of the findings were the opposite compared to theoretical part.

Empirical value factors regarding external data sources (chapter 4.2.2) were not so good straightly aligned with theoretical part. But when thinking a bit more on a higher level, ML calculation power (Sutton et al. 2018; Stormi et al. 2018; Bolander 2019; Bergmann et al. 2020) is again valid point in the case of external/new internal data. Otherwise the scope of this research doesn't cover good enough this section of empirical from theory perspective. Instead of that, for example the ML value categories presented in literature review (figure 5) and especially the able to cover challenges in human work support the to utilize ML when thinking new, possibly complex external data that humans cannot handle without some machine help.

Last chapter of empirical value part presented how ML could help to reach better results in S&OP according to the interviewees. Two main highlights there were improved forecast accuracy and better adaption to changes. Improved forecast accuracy is perfectly aligned with the theoretical ML value finding. Bergmann et al. (2020) highlighted this as one of the value elements in ML utilization. In general one the ML value categories presented earlier (figure 5) was possibility to financial improvement. It is good aligned in general with the empirical thought that ML could help to reach better results in S&OP. Financial perspective is always one of the element within organization when measuring the S&OP performance, for example. Maybe the connection is not really straightforward but still, there is a connection.

On the other hand, better adaption is not aligned with the theoretical findings. Brynjolfsson & Mitchell (2017) presented that one of the challenges in ML is that the adaption to changes is not on very high level, which means that the empirical finding is telling the opposite story. It can be explained by the lack of exact knowledge of the interviewed people or some wrong assumptions regarding ML utilization. In this case could be said that the empirical part is misleading, and theory is right.

All in all, there are many findings presented in the theoretical part that support the empirical study and also some points which are not so good aligned. By understanding both sides, theory and empirical part, good understanding about the topic and case study regarding ML utilization value can be created. It is clear that some parts would need further study and material, but this will be out of the scope of this study due to the limitations regarding the nature of master's thesis.

5.3 Qualitative and experience-based data in machine learning process

This chapter presents the synthesis of theoretical and empirical part in the context of ML challenges in S&OP process. The highlights from both sections, empirical and theoretical, are combined into one focusing especially on the similarities and differences between those. Goal is to present one general view about the possible challenges from this research's point of view.

Starting from the main findings of empirical part: lack of earlier experience, qualitative and experience-based data, and reaching enough improvement were the most important areas based on the research interviews. Even though the lack of earlier experience was seen as one of the main challenges for ML utilization based on the interviews, theoretical findings didn't support this since there was no mention about the need of earlier experience when planning and trying to implement ML into S&OP process. When looking at general level findings regarding ML challenges (table 4, chapter 2.2.5), cultural change was presented as a challenge related to ML (Möller et al. 2020) but this wasn't highlighted in any other literature, so it wasn't seen as a major challenge, and it is not directly connected to experience issue but on the same field in any case.

Possibility to data quality issues was seen as a pretty general worry in the empirical interviews because none of the interviews hadn't the experience and knowledge about the needed data quality for efficient ML solution. Data quality was presented also in the theoretical part as one of the most important enablers for an efficient S&OP process (Bower 2005; Prokopets 2012; Iyengar & Gupta 2013; Tuomikangas & Kaipia 2014). This means that data quality was recognized in both parts as important factor for both S&OP and ML. Data quality was also handled as a risk in the context of human vs. ML work in theoretical part (table 6, chapter 2.3.2).

Empirical part presented qualitative and experience-based data as a challenge because there are so much this kind of data within the organization, and it should be also taken into account in S&OP process. Including external data (Bower 2005) and excluding relevant trends (Pedroso et al. 2016) were presented as theoretical S&OP challenges even

though their role wasn't that major there (table 2, chapter 2.1.5). These are connected to empirical findings since many of the relevant factors (outside of the current process) would mean some qualitative and also experience-based to be included into process but so far it has been a challenge and it would be challenge also for ML utilization.

Last main finding in empirical part was related to the challenge whether enough improvement and value could be reached with ML utilization. Especially the need of human work and questions regarding maintaining the ML solution in longer term raised from the research interviews. From theoretical findings point of view and general ML challenges (table 4) especially expensive solution (Brynjolfsson & Mitchell 2017; Nielsen 2018), might be heavy to maintain (Brynjolfsson & Mitchell 2017) and requirement of new skills and talents (Nielsen 2018; Oesterreich et al. 2018) support these empirical findings.

Even though the need of human assistance raised from the interviews, theoretical part handled this topic in more detailed level and from more perspectives. For example, wrong or misleading metrics (Iyengar & Gupta 2013; Wilson & Raman 2017) and lack of measuring and monitoring (Bower 2005; Iyengar & Gupta 2013) were presented as traditional S&OP challenges and are also connected to possible use of ML but those weren't raised at all from the research interviews. All in all, even though there were same kinds of topics found from both theoretical and empirical part, theoretical part highlighted those more in detail and empirical findings were on a higher level.

All in all, major differences weren't recognized between the theory and empirical part. Some of the topics were highlighted more in theory part and some in empirical part but the findings could be combined well, and they supported each other. As a summary from the main challenges could be said that based on this discussion part, human assistance is needed next to ML utilization and the right and properly working way should be found in the organization. Also, the S&OP goals and enablers need to be taken into account when implementing the solution, including all the possible limitations and circumstances.

6. CONCLUSION

Conclusion chapter summarizes this research by presenting the summarized answers to research questions (6.1) based on the literature and empirical part. Second subchapter (6.2) includes ideas about future research in this field and what could be done in addition to this research topic. Last part (6.3) presents the validity and reliability of this study with the help of literature findings about those.

6.1 Answers to research questions

This chapter presents the summary of the answers to research questions of this study. Discussion part (chapter 5) already presented the main findings between theory and empirical part, but this will summarize the most important outcomes of this thesis.

First research question was how a company can utilize ML in their S&OP process. This is answered especially in empirical part (chapter 4.1). **The recommendation for ML utilization based on this master's thesis is to utilize ML in demand planning process of a specific business area**. With the power of ML case organization could better include external data and factors into demand planning process which would broader the fact basis of the whole planning. One of the main things which is not included into current process is weather data. With ML this could be taken into account better in demand planning because it is known that the impact of the weather is significant on the business. ML could also make it easier to utilize larger amount of data in the demand planning process and offer scenario analysis possibilities and weekly level insights.

Second research question was about the value that ML could bring into S&OP process. Many different aspects were studied and also found regarding value in ML utilization. ML could develop the S&OP process to the next level, especially in demand planning side. **Calculation power, ability to handle large amount of data and reduce the number of errors within the process compared to human work were some of the main findings**. ML could decrease the workload regarding the planning process which was seen currently as a problem of the current process. Also the ML ability to handle the data could make it possible to include many different data factors that the organization hasn't been able to do so far due to limited capabilities regarding human work. Also there was recognized some value factors that could create value, but they remained still a bit unclear and would need further research around this topic. One of the main outcomes was also

that the potential is there and with further work around this topic the organization can really reach something by utilizing ML.

Third and final research question was about the main challenges regarding ML utilization in S&OP process. **These main challenges based on this research are related to utilizing qualitative and experience-based data in ML in a way that support S&OP process and its results. Also the lack of knowledge related to ML and its possibilities was seen as a challenge.** Fact based S&OP process was presented as S&OP enabler (Bower 2005; Iyengar & Gupta 2013) in theory part and that would mean external data and trends to be included into S&OP process. Based on the research interviews these data sources are mostly non-quantitative and there should be a way to take these into account in ML solution, which can cause challenges for the whole utilization. **Another significant and also general challenge in ML utilization is often data quality. This was recognized also in this research as a factor that can cause challenges to ML utilization.**

All in all, this research could pretty well create a picture about the situation of the organization's S&OP from ML utilization perspective by focusing on the research questions. Since the topic is pretty large, some of the things would need more investigation and research but in general the findings were relevant and comparison between theory and empirical part could be done.

6.2 Future research

In this research the focus was on the ML utilization possibilities in S&OP process including the possible value and risks that ML could bring. This study doesn't focus so much on the actual ML techniques that could the most suitable ones from the S&OP process point of view. That could be an interesting research topic for the future to study what kinds of techniques are the most suitable and effective ones, for example for the utilization cases presented in chapter 4.1.

Another interesting focus area could be actual implementation of a ML solution into S&OP process. This is also out of the scope in this study but includes for sure many kinds of aspects that needs to be taken into account. As an example the risks for utilization presented earlier in this research are one thing that needs to be considered in implementation phase and research around this could have research potential.

In addition to previous ones, one research direction could be to deep dive into some specific part of S&OP process, for example demand planning. That could give a broader view about the possibilities within that process instead of going through the whole S&OP

process. On the other hand, financial planning could be a single focus area in a research. In this study, that area was left on the background but that could be an interesting area to study from ML perspective, there is for sure a lot of potential.

6.3 Validity and reliability

Typical way to analyze the credibility of a qualitative research is to assess reliability and validity of the research. With these two perspectives, overall picture of trustworthiness of the research can be created. (Morse et al. 2002) According to Saunders et al. (2019, p. 213) reliability and validity are central ways of assessing the research quality.

Reliability refers to data collection methods used in the study and the ability to produce same kinds of results despite of who is doing the research and when (Kirk & Miller 1986; Saunders et al. 2019, p. 213). Reliability can also be divided into internal and external reliability (Saunders et al. 2019, p. 214). Internal reliability refers to the role of the researcher in a research project (Saunders et al. 2019, p. 214) and external reliability refers to data collection techniques and their repeatability (Saunders et al. 2019, p. 214).

Saunders et al. (2019, p. 214) presented four different threats to reliability, and those are participant error, participant bias, researcher error and researcher bias. Participant error refers to the circumstances that may affect the answers, the way of answering or behavior of the interviewee (Saunders et al. 2019, p. 214). In this study, the interview structure wasn't tested before the interview sessions with test persons, but all the interviewees received the interview structure and topics a couple days before the actual interview session, so they had a chance to prepare themselves for the sessions and think about possible answers and themes.

Participant bias refers to circumstances and possible false responses that may occur due to circumstances around the interview session (Saunders et al. 2019, p. 214). In this research all the interview session were recorded for further analysis, which may have had some affect the answers, but all of the participants knew that beforehand and were able to prepare themselves for these circumstances, so it can be said that the those impacts on the answers weren't so significant. Also, the interviewees knew that all the analysis will be made total anonymously and their names or any details wouldn't be visible to anyone.

Third possible threat to reliability of a research is researcher error. This refers to researcher's impact on the results and circumstances around that. (Saunders et al. 2019, p. 214) In this study, the interviews were recorded which enabled the researcher fully concentration during the sessions, and also improved the possibilities to analyze the

findings afterwards. The recordings also made it possible to go through the interviews multiple times and make large notes based on the recordings.

Last threat to reliability, researcher bias, refers to researcher's ability to record the interviews. Like already mentioned, the whole interview sessions were recorded, which decreased the possibility to misunderstandings and wrong conclusions. The recordings were listened afterwards, and large notes were created based on the recordings so the possibility to this threat is only minor one.

In addition to previous mentions, the reliability risks decreased due to the selection of the interviewed group of people, who had different background, experience and roles within the S&OP process. So, many different perspectives could be gathered, and large analysis made. In addition to reliability, also validity needs to be assessed to get a view into the credibility of a research.

Validity refers to the measurements used in the research, analysis accuracy and generalization of the research findings (Saunders et al. 2019, pp. 213 – 214). According to Saunders et al. (2019, p. 216) especially in qualitative research validity is strongly connected to generalization of the findings. In this study a threat to generalization can be the size of the group of interviewed people, which was eight in total and all from the same organization. Even though, they were on different roles and different levels on the hierarchy, which offered larger view around the topic. When connecting this to theoretical findings, pretty good overall view around the phenomenon could be created.

All in all, reliability and validity of this study are on pretty decent level especially when taking into account that this is a master's thesis project. Some minor improvement could have been made, for example larger sample size in the interviews and full transcription based on the interviews for better analysis possibility.

REFERENCES

Ailisto, H., Heikkilä, E., Helaakoski, H., Neuvonen, A., & Seppälä, T. (2018). Tekoälyn kokonaiskuva ja osaamiskartoitus.

Alpaydin, E. (2014). Introduction to machine learning. Third edition. Cambridge, Massachusetts: MIT Press.

Bergmann, M., Brück, C., Knauer, T. & Schwering, A. (2020). Digitization of the budgeting process: determinants of the use of business analytics and its effect on satisfaction with the budgeting process. Journal of management control. Vol. 31(1-2), pp. 25–54.

Bolander, T. (2019). Human vs machine intelligence: How they differ and what this implies for our future society. Proceedings of Pragmatic Constructivism. Vol. 9(1), pp. 17–24.

Bower, P. (2005). 12 Most common threats to sales and operations planning process. The journal of business forecasting. Vol. 24(3), pp. 4–14.

Boyer, J.E. (2009). 10 proven steps to successful S&OP. The journal of business forecasting. Vol. 28(1), pp. 4–10.

Brynjolfsson, E. & Mitchell, T. (2017) What can machine learning do? Workforce implications. Science (American Association for the Advancement of Science). Vol. 358 (6370), pp. 1530–1534.

Chaffin, C. R. (2015) CFP Board financial planning competency handbook. Second edition. Hoboken, New Jersey: Wiley.

Chandramouli, S., Dutt, S. & Das, A. (2018). Machine Learning. 1st edition. Pearson Education India.

Chase, C. (2013). Demand-driven forecasting a structured approach to forecasting. 2nd ed. Hoboken, N.J: Wiley.

Duan, Y., Edwards, J.S. & Dwivedi, Y.K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. International journal of information management. Vol. 48, pp. 63–71.

Economist. (2017). Unshackled algorithms; Machine-learning in finance. Vol. 423 (9042), p. 68.

Emmanuel, C. R., Otley, D. & Merchant, K. (1990). Accounting for management control. 2. ed. London: Chapman and Hall.

Feigin, G. (2012). Supply chain planning and analytics the right product in the right place at the right time. 1st ed. New York, N.Y: Business Expert Press.

Feng, Y., Martel, A., D'Amours, S. & Beauregard, R. (2013). Coordinated Contract Decisions in a Make-to-Order Manufacturing Supply Chain: A Stochastic Programming Approach. Production and operations management. Vol. 22(3), pp. 642–660.

Fisher, I.E., Garnsey, M.R. & Hughes, M.E. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research. International journal of intelligent systems in accounting, finance & management. Vol. 23(3), pp. 157–214.

Havlíček, V., Corcoles, A.D., Temme, K., Harrow, A.W, Kandala, A., Chow, J.M & Gambetta, J.M. (2019) Supervised learning with quantum-enhanced feature spaces. Nature (London). Vol. 567(7747), pp. 209–212.

Iyengar, C. & Gupta, S. (2013). Building blocks for successful S&OP. Supply chain management review. Vol. 17(6), pp. 10–17.

Jacobs, F.R., Berry, W.L. Whybark, D.C. & Vollmann, T.E. (2018). Manufacturing planning and control for supply chain management: the CPIM reference. Second edition. New York: McGraw-Hill Education.

Kirk, J. & Miller, M. L. (1986). Reliability and validity in qualitative research. Sage. Beverly Hills (Calif.).

Korhonen, T., Selos, E., Laine, T. & Suomala, P. (2021). Exploring the programmability of management accounting work for increasing automation: an interventionist case study. Accounting, auditing & accountability. Vol. 34(2), pp. 253–280.

Lee, J. H., Shin, J. & Realff, M.J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. Computers & chemical engineering. Vol. 114, pp. 111–121.

Liberatore, M. J. & Miller, T. C. (2012). Supply chain planning practical frameworks for superior performance. 1st ed. New York, N.Y.: Business Expert Press.

Lichtenthaler, U. (2020). Beyond artificial intelligence: why companies need to go the extra step. The Journal of business strategy. Vol. 41(1), pp. 19–26.

Márcio Tavares Thomé, A., Felipe Scavarda, L., Suclla Fernandez, N. & Jose Scavarda, A. (2012). Sales and operations planning and the firm performance. International journal of productivity and performance management. Vol. 61(4), pp. 359–381.

Martínez-López, F. J. & Casillas, J. (2013). Artificial intelligence-based systems applied in industrial marketing: An historical overview, current and future insights. Industrial marketing management. Vol. 42(4), pp. 489–495.

Morse, J. M., Barrett, M., Mayan, M., Olson, K. & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. International journal of qualitative methods. Vol. 1(2), pp. 13–22.

Mulvey, J. M. (2017) Machine Learning and Financial Planning. IEEE potentials. Vol. 36(6), pp. 8–13.

Möller, K., Schäffer, U. & Verbeeten, F. (2020). Digitalization in management accounting and control: an editorial. Journal of management control. Vol. 31(1-2), pp. 1–8.

Nguyen, B. & De Baets, B. (2019). Kernel-Based Distance Metric Learning for Supervised k -Means Clustering. IEEE transaction on neural networks and learning systems. Vol. 30(10), pp. 3084–3095.

Nielsen, S. (2018). Reflections on the applicability of business analytics for management accounting – and future perspectives for the accountant. Journal of accounting & organizational change. Vol. 14(2), pp. 167–187.

Oesterreich, T. D., Teuteberg, F., Bensberg, F. & Buscher, G. (2019). The controlling profession in the digital age: Understanding the impact of digitisation on the controller's job roles, skills and competences. International journal of accounting information systems. Vol. 35, p. 100432.

Oliva, R. & Watson, N. (2011) Cross-functional alignment in supply chain planning: A case study of sales and operations planning. Journal of operations management. Vol. 29(5), pp. 434–448.

Pedroso, C. B., da Silva, A.L. & Tate, W.L. (2016). Sales and Operations Planning (S&OP): Insights from a multi-case study of Brazilian Organizations. International journal of production economics. Vol. 182, pp. 213–229.

Prokopets, L. (2012). S&OP: what you can learn from top performers. Supply chain management review. Vol. 16(3), pp. 28–35.

Rai, R., Tiwari, M.K, Ivanov, D. & Dolgui, A. (2021). Machine learning in manufacturing and industry 4.0 applications. International journal of production research. Vol. 59(16), pp. 4773–4778.

Russell, S. & Norvig, P. (2016). Artificial Intelligence: A Modern Approach. Harlow: Pearson Education, Limited.

Saunders, M. N. K. et al. (2019) Research methods for business students. Eighth edition. Harlow: Pearson.

Schlegel, A., Birkel, H.S & Hartmann, E. (2021). Enabling integrated business planning through big data analytics: a case study on sales and operations planning. International journal of physical distribution & logistics management. Vol. 51(6), pp. 607–633.

Secrett, M. (2012) Brilliant budgets and forecasts: your practical guide to preparing and presenting financial information. 1st edition. Harlow, England: Pearson.

Skitka, L.J., Mosier, K. & Burdick, M.D. (2000). Accountability and automation bias. International journal of human-computer studies. Vol. 52(4), pp. 701–717.

Stormi, K., Laine, T. & Elomaa, T. (2018). Feasibility of B2C customer relationship analytics in the B2B industrial context.

Sutton, S.G., Arnold, V. & Holt, M. (2018). How much automation is too much? Keeping the human relevant in knowledge work. Journal of emerging technologies in accounting. Vol. 15(2), pp. 15–25.

Thunberg, M. & Fredriksson, A. (2018). Bringing planning back into the picture - How can supply chain planning aid in dealing with supply chain-related problems in construction? Construction management and economics. Vol. 36(8), pp. 425–442.

Tuomikangas, N. & Kaipia, R. (2014). A coordination framework for sales and operations planning (S&OP): Synthesis from the literature. International journal of production economics. Vol. 154, pp. 243–262.

Uzsov, R., Fowler, J.W. & Mönch, L. (2018). A survey of semiconductor supply chain models Part II: demand planning, inventory management, and capacity planning. International journal of production research. Vol. 56(13), pp. 4546–4564.

Wallace, T. (2006). Forecasting and sales & operations planning: synergy in action. The journal of business forecasting. Vol. 25 (1), pp. 16–36.

Whitenack, D. (2017). Machine learning with Go: implement regression, classification, clustering, time-series models, neural networks, and more using the Go programming language. 1st edition. Birmingham, England: Packt.

Wiedenmann, S. (2013). Supply planning for processors of agricultural raw materials. 1st edition. Göttingen, Germany: Cuvillier Verlag.

Wiering, M. & Otterlo, M. van (2012) Reinforcement Learning: State-of-the-Art. 1. Auflage. Vol. 12. Berlin, Heidelberg: Springer-Verlag.

Wilson, E. & Raman, S. (2017). Why Is the S&OP Process Stuck in Third Gear? The Journal of Business Forecasting. Vol. 36(3), pp. 10–14.

APPENDIX A: INTERVIEW STRUCTURE

Business drivers and process:

1) What kinds of drivers are the basis of forecasting and planning process (in sales, production, finance planning, HR, ICT, etc.)?
2) What kind of external factors are affecting to your forecasting process/business? (Also, the ones that might potentially affect) (e.g. *weather*, *COVID vaccination %*, *remote working, public economy statistic or anything else*)
3) What are the challenges regarding these? (*question 3*)
4) What's the granularity of forecasting/planning? Is that the optimal level or would there be a need for more detailed level ("granularity goal")?

Role of automatization

5) What kind of things take "too much" time compared to value they bring?
6) How do you see the role of automatization in planning/forecasting process now? And in the future?
7) What kind of benefits would more automatized way of planning bring?
8) One thing you would like to have automatized compared to the current situation?
9) What kinds of challenges do you see regarding automatization/machine learning utilization?

Measuring and monitoring of forecasts:

10) How realistic do you see your forecasting/planning at the moment?
11) How do you measure your forecasting?
12) What would make the measuring better and even more useful? (E.g. from future decision making point of view)
13) What kind of factors are behind the forecast/estimate accuracy (positive or negative)?
14) How would you make the forecasting/planning more effective?

S&OP Process

15) How integrated is the planning process between sales, production and financial side?
16) How do you keep in balance your sales and production?
17) How do you allocate your production to different production lines/factories/bogs?