Exploration of Machine Learning in the Complex and Dynamic environment of Product Returns in the E-commerce Market

The Impact of product returns in the Dutch e-commerce market; Exploratory study in the promising applications of Machine Learning.

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in {Complex System Engineering and Management}

Faculty of Technology, Policy and Management

by

{Daan Westmaas}

Student number: 4247566

To be defended in public on 11-07-18

Graduation committee

First Supervisor	: Dr. M.V. Dignum, *Section ICT*
Second Supervisor	: Ir. M.W. Ludema, *Section Engineering Systems and Services*
External Supervisor	: Ir. S. Morales Velasquez, *Accenture Technology*

(This page is intentionally left blank)

“Develop a passion for Learning. If you do,

you will never cease to grow. "

Anthony J. D’Angelo

i. Summary

The Dutch e-commerce market had a total revenue of 22,5 billion euro in 2017, and is expected to growth by 17% in the year of 2018 (E-commerce, 2018). Together with this growth, e-commerce players are having problems with high product returns due to the digitalization and change in customer behavior (Terry, 2014). People are not going to physical shops anymore, they buy online and choose their products from an internet picture. "We've created a Monster", is what S. Dennis states, worried about the returns that surpass 50% on several products in e-retailing (Dennis, 2018). Customers do not have the 'touch and feel' experience anymore and buy products with less confidence. Especially the products that require a 'touch and feel' experience suffer under high return rates. The images of the products on the website can give a distorted image of reality. This translates into relative high return rates, and these returns are stimulated by the retailers itself. They offer free returns and encourage customers to buy the product and send it back without any hesitation. The market of product returns is in desperate need of innovation and improvement, their data driven capability enables the opportunity to explore new disruptive technologies to make this innovation and improvement happen. A literature review indicated a knowledge gap in the field of return management and Machine Learning. This research explored the possible applicability of Machine Learning to improve the performance on KPIs measured in the return process. The following research question was derived:

"Concerning the high product returns in the Dutch e-commerce market; What can be the applicability and potential impact of Machine Learning on the Key Performance Indicators measured in the process of Return Management?"

To identify and explore the possible applications of ML in the field of product returns, the following research approach is followed. First, a case study was conducted to provide insights in the current state of product returns in the e-commerce market. The case study was performed with a major Dutch e-commerce player operating in the electronics market. From this case study, Bottlenecks and KPIs were derived, that were linked to literature. Literature was extensively researched to be complementary and supportive on the findings of the case study. This led to a quarter of bottlenecks in the fields of; customer contact, cost and time in the return process, forward process and supplier selection. These bottlenecks were used to indicate the applicability of ML in the process of product returns. Through the case study analyses the distinction between the forward and return process was clearly visible. The forward process relates to all processes and strategies involved to prevent returns. The purchase process of the customer needs to be shaped efficiently and complete to help the customer in finding the right product. The return process is focused on the product returns, that eventually always will stay coming. This process can always be improved in term of performance, costs and time.

The KPIs were identified to map the current performance measurement structure in the return process. These KPIs were based on the frameworks from Shaik & Abdul-Kader (2012) and Bernon et al (2011), these two papers have extensive mapping of performance measurement in the return process. By analyzing the bottlenecks mainly out of practice and identifying the KPIs from both literature and practice, this gave a total overview of what can be improved and how to measure performance in the returns. The relationships between the different parties in the supply chain is of crucial importance to be robust and create a resilience and effective supply chain. Therefore, the relation and agreements with suppliers is taken into account to see what the effect supplier collaboration can be on performance in the return process.

Secondly, the field of Machine Learning was explored which led to three useful ML techniques for this research. Supervised Learning, Unsupervised Learning and Reinforcement Learning were described to sketch their competence and potential in return management. The capabilities of these Machine Learning algorithms are explained and linked to the processes and bottlenecks in the return process. Supervised Learning can use labeled historical data to predict new examples of output data. Unsupervised Learning works with unlabeled data and can cluster this unlabeled data into meaningful groups. Reinforcement Learning is technique that improves over time due to the maximization of its rewards.

Extensive use of current data with help of Machine Learning can lead to a decrease in costs. Labor intensity in different departments of an e-commerce player can be reduced, both at customer service level as operational levels in the warehouse. It can increase the responsiveness of the e-commerce player, which can lead to higher customer satisfaction. Eventually product returns can also be prevented, by clustering certain customer groups and provided them with matching reward systems. Furthermore, the relationship with the supplier plays a key role in the process of returns. Machine Learning can help in predicting future product returns, based on historical data of a certain supplier. This can support e-commerce players in having the right agreements with their suppliers, which will give them the opportunity to offer high customer loyalty in the process of returns. Serving the customer at a high level in the return process will be beneficial for the whole supply chain. These possibilities are summarized in figure 1.

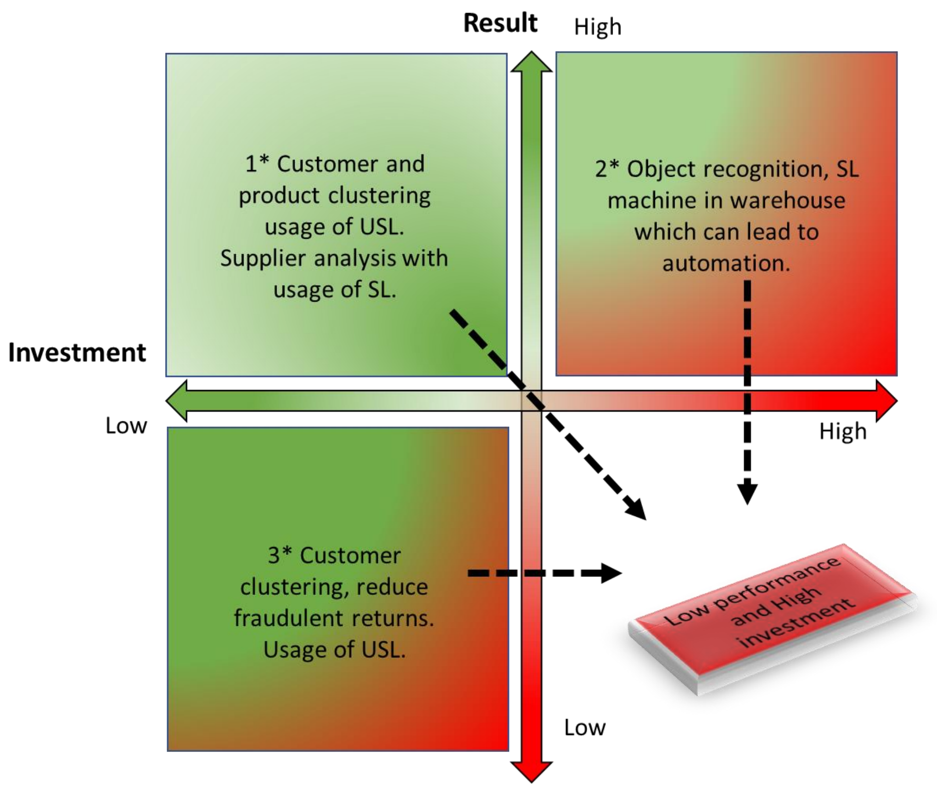


Figure 1, Investment vs Potential Result of Machine Learning in the Return Process

Starting from the top left, the extensive use of current data can reduce product returns. This requires a relative small investment which can lead to high results. By preventing returns, the costs related to product returns can be decreased, customers are informed earlier to prevent returns due to failure of the product and reward systems can help customers in keeping the products. The second possible application of ML requires a higher investment and lead to high results in the return process. The use of Supervised Learning, especially object recognition can be applied in the warehouse to identify and sort the return products. This will require less labor intensity and can improve the quality of the process. Thirdly, fraudulent returns can be prevented with the use of current data. This requires a relative low investment, however the impact of this technique will also relatively low in the market of electronics. It can have higher results in other e-commerce markets where the fraudulent return rates are higher. All three applications of ML are having positive impacts, this can be affected by several factors. Companies have to be ready for the implementation of ML, it will change processes and working manners. This can create change in the business and employees must be prepared. ML implementations can also go wrong due to bad management or lack of knowledge. We must be aware of the downsides and effects that ML can create.

Based on this research, mainly two solution streams were identified that can be applicable for companies to further explore. The first possible application of ML is the extensive usage of current and existing data. This can be used in the forward process to lower the return rate and specify on product and customer level. This will create better insights into the reasons of returns, customers preferences and returning behavior. This can be used to predict returns in future situations and help in reducing and preventing returns. This first possible application of ML will require relative small investments and is less intrusive to implement, it will not affect the business organization structure and processes.

The second possible application of ML is the intrusive design in the return process. The reshaping of the total process can optimize the return process and can contribute in reducing labor intensity and increasing communication. It can help in making the process faster and there will be more automation due to the ML algorithm in the machine that handles all incoming returns at the warehouse. Such machines can be very efficient and increase quality and reduce labor intensiveness. This possible application of ML will ask for a bigger investment and will have a heavier impact on the business organization.

Table of Contents

ii. List of Figures

iii. List of Tables

1 Introduction

The Dutch e-commerce market had a total revenue of 22,5 billion euro in 2017, and is expected to growth by 17% in the year of 2018 (E-commerce, 2018).). Together with this growth, e-commerce players are having problems with high product returns due to the digitalization and change in customer behavior (Terry, 2014). People are not going to physical shops anymore, they buy online and choose their products from an internet picture. “We’ve created a Monster”, is what S. Dennis states, worried about the returns that surpass 50% on several products in e-retailing (Dennis, 2018). Free shipping and hassle-free returns are demanding requirements from the customers. In literature this concept is often called, Return Management (RM) or Reverse Logistics(RL). Return management is related to almost every stakeholder in the supply chain. It starts at the end customer and the product can go all the way back to the original supplier. High costs of the return flows are damaging companies and they do not know how to handle them efficient and cheap. Recently it came out that Amazon Germany destroys all returned products (Hielscher et al, 2018). Due to the high costs and labor intensity of the process they just throw products away. Being more adaptive, reflective and transparent in the process of RM can help companies in creating a stronger market position. A Canadian tire company shared their information which they received from the RM department. Providing feedback from the return flows to the sales or design departments of the company can help in lowering the return rates. Their high return rates occurred because customers did not understand the product manual. When this information was clarified and picked up by the return management department, it could be given as feedback to the design and sales departments. Rewriting the user manuals could prevent and lower the high return rates (Terry, 2014).

Managing the returns in an efficient way and making the right decision under uncertain, complex and dynamic circumstances can be very challenging. The decisions are not bound to one company, mostly they are overarching different entities and this makes the boundary bleary and hard to define. Forecasting returns is complex and makes the return process unpredictable and hard to manage. Including suppliers early in the process can be an option to be more resilient and have wider support throughout the SC. These complex dynamic systems are perfect situations to explore new innovative and disruptive technologies.

Due to the rapid growth and importance of data, e-commerce players need to focus more and more on data related decisions and strategies (Dallemulle et al, 2017). Big Data (BD), Business Intelligence (BI) and Artificial Intelligence (AI) are used to attract and create new business. Companies speak out that they want to engage in Blockchain, big data and AI while they have no idea what it means and how to do it. The applicability of AI is extraordinary and has big potential in different areas. It can support the decision-making process within companies, it can perform simple human tasks, it can predict unforeseen events and discover patterns in (human) behavior (Cohen, 2014). The application of AI will be explored in the field of product returns in the e-commerce market. Most products flow from suppliers to customers, as displayed in figure 2. However, the reverse flow is growing in volume and needs improvements. Figure 2, indicates an end-to-end SC structure, from the original supplier to the end customer. Supply chains differ from company, product or country. Fresh products like vegetables and fruit, will need faster processing time compared to electronics. Important note; the flows can be bidirectional, this means that the product flows can go both back and forth.

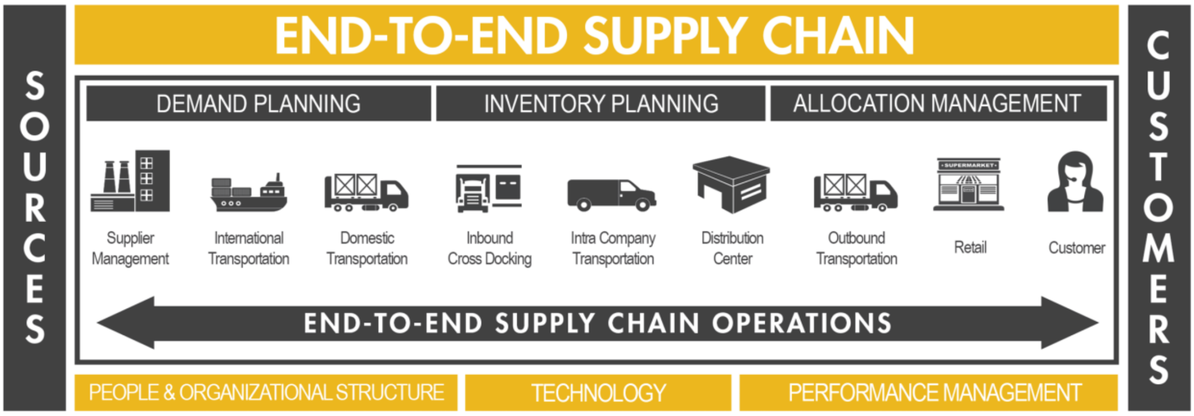


Figure 2, Supply Chain Structure example (Leancor, 2018)

1.1 Problem Exploration

As mentioned earlier, the e-commerce market is booming and online retailers are growing rapidly. In the past 5 years, the online market in The Netherlands has doubled and now has a total annual revenue of 22,5 billion EURO (E-commerce, 2018). Together with this growth in online purchases, the product returns also increased rapidly. These high return rates are one of the most important reasons why these e-commerce players are still not profitable (Wellens, 2017; Rigby, 2014). The e-commerce players do not know how to handle the returns efficiently and created this monster by themselves (Dennis, 2018). Continuous advertising on their free return policies and hassle-free return handling, resulted in loss-making market models. It seems somewhat logic providing free returns, research proved that flexible return policies and good handling of returns will positively influence future customer purchase behavior (Minnema, 2017; Petersen and Kumar, 2009; Wood, 2001). While practices from online retailers together with researchers proved that lowering return rates with 10% improves profitability with 20% (Pur et al, 2013). This indicates the high impact of product returns on the profitability and performance of companies. The costs of returns should be shared with all stakeholders in the return process. However, now the e-commerce players are in most cases responsible for the costs of returns. There is much to improve on the agreements with suppliers and therefore this study also explores the contracts and agreements with suppliers in the return process. In such an environment there is complexity due to the different stakeholders, regulations, institutions and many more conflicting arguments. Continuous changes and a dynamic environment provide challenging issues that cannot be solved easily, the field of product returns perfectly matches all these requirements. In 2017, driven by the accession of social media, big data and Internet of things, most supply chains now have 50 times more data available than the year 2012 (Ellis, 2016). To deal with these huge amounts of data it is required to have data analytical tools and techniques that will give usable insights. Even data strategies and models are required to make efficient use of the available information. Due to the new developments and trends within this sector, it opens opportunities for new AI tools and techniques.

The problem of product returns will be analyzed and the possible application of AI will be explored. AI is used as a possible application that can help in solving a part of the problem in product returns. AI is a broad field and includes many different techniques. Examples of AI techniques are, machine learning, natural language processing or neural networks. For this research, the focus will be on Machine Learning (ML). ML is an old concept and tries to give computers the ability to learn without being explicitly programmed (Goldberg et al, 1988). There were several reasons to go for ML in this study. First the massive creation of data in supply chain. Second; the current available knowledge in the field of ML is extensive and still growing and enables this study to have a lot of literature research available. Third; the potential impact it can have on business performance in RM. However, a more extensive explanation about why ML is most suitable for this research, will be explained later in chapter 4 and 5. Within ML there are different methods, for instance; Supervised Learning and Unsupervised learning. These algorithms or methods cannot be applied in every regular company. Applying ML requires data standards, skilled employees and clear data strategies. Tasks that are repeated every day and have little complexity do not have the need for AI tools. However, repeatable tasks that have complexity and a variety of outcomes can have a potential to be performed and integrated with AI (Ng, 2017). Therefore, the decision of using and implementing ML is a very important process. Mathematical models or generic algorithms can be very useful and effective to solve complex issues in different situations. The distinctiveness of ML can lie in the search for patterns or clusters and ML can make predictions on historical data. These predictions can help in determining strategies for a company. For instance; image classifying machine, to classify what kind of object is on the photo can be an application of object recognition. Output can be a number between 1 and 10000 indicating the object that is seen on the photo. Learning from millions of previous photos, the algorithm can predict, with a level of certainty, what kind of object is on the image.

The possibilities of AI can bring many new opportunities into the markets and industrial sectors, it can completely change businesses and market models (Ng, 2017). Other researchers from Pew Research Center following Jarvis (2017) state "By 2025, artificial intelligence will be built into the algorithmic architecture of countless functions of business and communication, increasing relevance, reducing noise, increasing efficiency, and reducing risk across everything from finding information to making transactions" (Jarvis, 2017, pg 1). This is supported by the yearly hype cycle that Gartner brought out in July 2017 (Figure 3, Gartner AI hype curve, 2017). The AI applications and possibilities are enormous and certain techniques like Machine Learning, Deep Learning and Virtual Assistants are on their top of expectations. The main reasons for this hype is the current state of computer power and the production of data. Computer power has grown exponential over the last 20 years and enables such technologies to perform much better compared to 10-20 years ago. Data growth ensures that these techniques have enough 'fuel' to be filled with. Due to the digitization and e-commerce market the data has grown exponential.

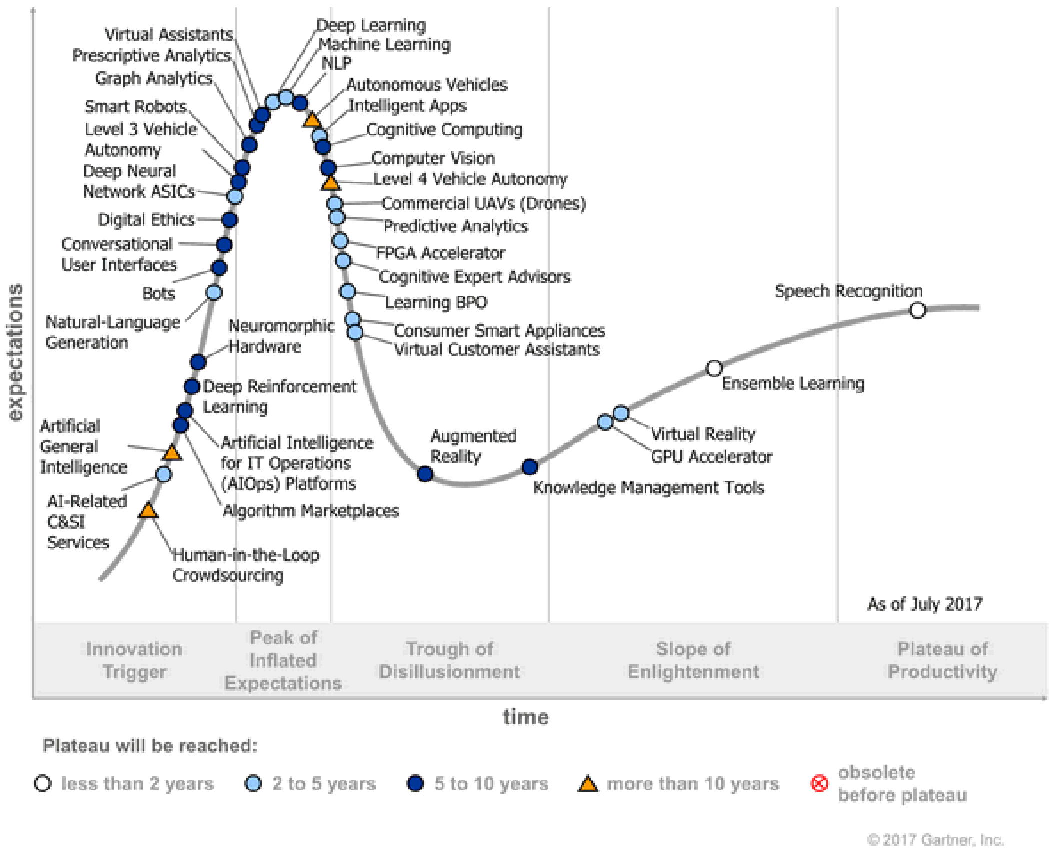


Figure 3, Gartner AI Hype Cycle (Gartner, 2017)

1.2 Main Research Question & Sub questions

The following research question is derived from the examined literature:

“Concerning the high product returns in the Dutch e-commerce market; What can be the applicability and the potential impact of Machine Learning on the Key Performance Indicators measured in the process of Return Management?”

This research question intends to make analysis, explore a complex, dynamic environment and describe the processes of RM. The research question indicates both a knowledge and design problem. First, the knowledge problem arises when combining the two concepts of Return Management and Machine Learning. In current literature there is no study that describes the possibilities of linking these two aspects. Where Return Management attracts more and more studies that last several years, it is still an area that is quite undiscovered compared to other areas in the field of supply chain management. Therefore, there will be desk research, literature review and case study to obtain knowledge about the return management process. This will be used to explore the different possibilities of the three ML techniques, what can be the impact of these techniques on certain bottlenecks in the process of returns.

Second, there is a design problem whereas there is no application of ML techniques in this field of the supply chain. Where ML is applied in marketing strategies, personal advertisement, inventory management and many more fields, in literature there is no application in the field of RM. This study will try to create a matrix that will give insight on the impact of ML on KPI performance. This matrix will be designed to support the steering on KPI level. There will not be an implementation of ML in a certain case study, however the developed matrix will contribute in the process of an experiment or implementation. By using the matrix an implementation of ML in RM can be set up and can be used to test and validate the results of this study. However, an implementation of ML in RM is out of the scope for this study.

The following sub-questions are formulated to explain the different fields of the main research question and support the outcomes of this research project:

1. ***What are the current components of a return management process from an e-commerce player in the Dutch electronics market?***

   a. *What is the total process of Return Management within the given case study?*
   b. *Is there a link between Return Management and Supplier selection and if so, what does this link contains?*
   c. *What are the relevant Key Performance Indicators measured in the return process?*

The first Sub-Question will map all the processes and relevant information involved in return management. The total Return process will be mapped through interviews and literature. There will be a case study with a company that operates in the e-commerce market. This case study will be an example for the market where it is operating. The most used return process of the company is mapped to create an impactful analysis. If the choice was to conduct a broader study, where multiple operating e-commerce companies were analyzed, it would be a whole different outcome. A comparative study will probably lead to the 'most perfect' return process. However, due to the dynamic markets and the difference between markets, it is very hard and almost impossible to create a perfect return process that will be applicable for all. Therefore, this study tries to go into depth into one case study, filter out the KPIs and see what a ML can bring on the processes and KPIs. The impact on KPIs can generalized more easily, all operating companies in the e-commerce market are using these KPIs to certain extent.

2. **What is Supervised Learning, Unsupervised Learning and Reinforcement Learning?**

    a. *What does this technique mean and what can it do?*

    b. *What are the requirements and pitfalls of this technique?*

This sub research question will explain shortly the underlying principles of the three Machine Learning techniques. To apply these techniques, the basics of the techniques will be explained to have an idea of their meaning and performance.

3. What are the possible applications in Return Management for Machine Learning

    a. *Why is this technique chosen for the given project?*

    b. *What is the biggest potential in this technique related to RM and Supplier Selection?*

To justify and explain why these Machine Learning techniques are chosen for this project is explained in sub question 3 and 4. It will provide information about the applicability of the techniques and will describe the potential it can bring. This is important to justify the choice for these Machine Learning techniques. Other disruptive technologies could be interesting and impactful as well for this field in the supply chain. However, a choice had to be made for this study and therefore it is needed to clarify and explain the choice for Machine Learning.

4. What are the possible applications of Supervised Learning, Unsupervised Learning or/and Reinforcement Learning, on the bottlenecks in the Return process of an e-commerce company?

In sub question 4, this study provides the application of ML in the Return process. Note that this includes all processes involved in return, as well as prevention as the return process itself. This is divided into two domains. First, the forwards process where return prevention is desired to bring down the percentage of returns. Second, the return process self where disruptive technologies can contribute in making a more fluent and efficient process.

5. What can be the applicability of Supervised Learning, Unsupervised Learning or Reinforcement Learning on the defined Key Performance Indicators of the Return Process?

a. What is the applicability of the different ML techniques on the KPIs in the return process?

b. What can be the potential impact of the different ML techniques on the performance in the return department of an e-commerce player?

In sub question 5, the results will be further specified on KPI level. In this chapter the impact on KPIs will be described and the results will be shown. Not only the positive impacts will be discussed, also the negative impacts have to be addressed to give a complete view. This sub question is used to make the results more explicit and tangible, transforming the knowledge from literature to the business side is new in this field and therefore this sub question is the concluding and most important one.

All sub questions contribute in answering the main research question. In sub question 4 the results will be suggested and the link towards the application of the suitable ML technique will be shown. To answer the different sub questions there is need for several research techniques and tools. In the next chapter the questions will be discussed and the suitable techniques and tools will be presented.

1.3 Research approach

To explore the potential improvements related to Return Management, a case study is needed to make the analysis, explore and test the potential of ML techniques. The case study must have a web shop environment from B2C side and must be operating in the Dutch e-commerce market. By analyzing the current strategy and process around return management, the case study can be of an example to explore the different ML techniques. First the process of RM must be mapped to give an overview of the bottlenecks and decisions made in this process. The defined KPIs for the company will be described to see where the applicability of the ML techniques can be tested on. To come up with the right set of KPIs, frameworks from both Shaik & Abdul-Kader (2012) and Bernon et al, (2011) will be used together with the KPIs from the case study analysis. This will give a complete overview of the current representable KPIs in the return process. The processes will be analyzed by conducting the Business Process Modeling Notation (BPMN). Taking these processes as leading factors, the different ML techniques can be explored to develop a matrix that can map the potential applicability of ML in the return process. Using a BPMN, the whole process of returns will be mapped. This method will be used to analyze the current situation of product returns in the e-commerce market. By using the business process model of the case study as a basis, the different ML techniques can be explored on the identified processes and bottlenecks. The bottlenecks in the process will need new innovative improvements, such as ML, to create a better performance in the total process of product returns.

An interview with an e-commerce player in The Netherlands will be conducted to map the full process of product returns. Next to this, the agreements with suppliers will be checked to see what the link can be between RM and Supplier selection. An important part of this process will be to identify the available data to make sure there will be no holdbacks for the ML techniques. To give a complete overview of the processes, the data that is linked to each process must be defined to see the potential of data-driven technology (Kobayashi et al, 2003). When there is no accessibility to the data, a scenario will be sketched to test ML techniques. This simulation must be representative for real world situations. The total process of product returns can be quite complex due to the relations with different parties through the entire SC structure. Every product can have various destinations and every product must be handled differently. Some products are damaged, other products were delivered wrongly and some are returned due to false or incomplete information. This dynamic and fuzzy environment makes it complex to manage. The total process of returns will be mapped and all divarication within the process will be displayed. This is done to create a full overview and detect bottlenecks that can be improved by ML techniques.

When the process of the Return Management is mapped, the three following ML will be examined to see which technique can have impact and give the best support in the RM process to improve performance on the measurable KPIs. The differentiation is made into three ML techniques:

1. Supervised Learning (SL)
2. Unsupervised Learning (USL)
3. ReinForcement Learning (RFL)

The field of Artificial Intelligence includes many different techniques and possibilities, the hype cycle of AI from Gartner, presented in paragraph 1.1 figure 3, shows these multiple techniques. This research did only focus on one specific area in AI, the field of Machine Learning. ML is at its peak of expectations and is expected to reach its plateau within 2-5 years. The same trend is showed in the hype cycle for 'emerging technologies' of Gartner (2017). In this hype cycle Machine Learning is also indicated at the top of their expectations within the timeslot of 2-5 years. These two hype cycles show the potential for ML in a relatively short time period. Companies and literature are exploring the possibilities of ML and benefits from the applications. In the field of ML, the distinction is made into three smaller applications namely; Supervised Learning, Unsupervised Learning and ReinForcement Learning. These techniques are analyzed due to several reasons. First, in current literature these techniques are used most frequently. There is extensive research in this field and enough knowledge is available for these techniques. This can help in analyzing the problem and applying these techniques in the case study. Second, these different learning techniques are chosen to be complementary to each other. All three are suitable for different situations and therefore can be applied in separate parts of the processes. Thirdly, supply chains have become massive producers of data; customers data, process data and product data, are being gathered massively. Machine Learning requires massive amounts of data to learn and improve over time. The available data in the supply chain enables the ML techniques to perform well. And lastly, Machine Learning is performing better and better due to the increase of computer power (CPU). Due to this increase and the accessibility for the public, almost every company can now invest in ML with relatively small amounts of money. CPU enables ML algorithm to perform better and this will lead to results that can be very meaningful for companies. The current techniques can do object recognition, find patterns in data and can help in classifying customers, suppliers or products into certain similar groups. The Machine Learning techniques will be further explained and are made explicit for this case study in sub-question two.

All three ML techniques will be worked out to see which of the three will be most suitable for application in the problem of product returns in the e-commerce market. To explore this possible application the case study will be used to sketch scenario's and give opportunities for the application of Machine Learning. Based on the bottlenecks and KPIs of the case study, the application of ML can be presented.

1.4 Scope of the project

Scoping is one of the most crucial parts before conducting a research study. What to include and what not has to be clearly mentioned to indicate what will be delivered. This research will focus on one single case study; the company in the case study needs to have a minimum average return of 5%. To go for a single case study will give a more in-depth analysis. The focus of this research is on the problems with product returns in the e-commerce market, the case study was only used to map the total process and detect the bottlenecks. The return volume will probably differ between the e-commerce companies, however the measured KPIs and the detected bottlenecks will be the same in majority of these players. The logistics processes and different return flows are mostly the same at these e-commerce players and they all having a hard time to deal with these huge return flows. Therefore, the insights of this research are not specified on only the analyzed case study, they will be applicable for other players that are dealing with the same issues in product returns.

The choice for the e-commerce market is made through their huge capacity of data. Web shops are using tremendous amounts of data of their customers to personalize the websites and show personal advertisements (Laan, 2015). This data can lead to a more valuable Machine Learning algorithm that will perform better with more data. This research explicitly chose to analyze the return process of an e-commerce party. Transporters and other involved parties are left out of scope to narrow down and specify on the return process within an e-commerce party.

In this study, three Machine Learning techniques are used. Supervised Learning, Unsupervised Learning and Reinforcement Learning. Only these three techniques are being analyzed, mainly due to the existing knowledge available and the current rise of attention for these techniques. Many studies and researchers are exploring the new possibilities and the techniques are being applied in supply chain management broadly. To go for ML was an explicit choice, another technique that could have been chosen for this research was big data analysis. However, this is already more applied and discovered by both literature and businesses. Another substitute could be the possibilities of a Blockchain in the field of RM, however the Blockchain technology still comes with a lot of uncertainty. Exploring the Blockchain technology in the field of returns will not be suitable for a total supply chain structure. Blockchain must be implemented through the whole supply chain to make it work properly.

This research fixates on the exploration of possible Machine Learning techniques. This research is conducted out of curiosity and tries give more insights into the problems regarding product returns in the e-commerce market. It tries to clear the bush and clarify the problems and bottlenecks related to product returns that are currently of negative influence on the e-commerce companies. This research does not apply ML techniques and will not use data to employ an algorithm. The focus of this research is exploratory and innovative, which can lead to many more questions and uncertainties. This will provide interesting topics and issues for future research and will give concrete possible opportunities for future studies. The results of this research will be less tangible compared to a real data or ML implementation, however it will indicate the importance of product returns and will provide new questions that occur along the way.

1.5 Deliverables of the project

The deliverables of the project will indicate the applicability of ML on the performance of the KPIs in the return process. A matrix will be delivered that will indicate the impact of SL, USL and RFL on the KPIs measured in the return process. This research will give suggestions for future research in the field of Return Management and will give possible solution designs to implement ML in the return process.

1.6 Thesis structure

In chapter 2, the theoretical landscape and background of Return management and supplier relationship management will be further explored. From this exploration a knowledge gap will be identified, this gap indicates the possible exploration of new research in this area. In chapter 3, the case study analysis is conducted and all processes and relevant information involved in Return management is discussed. In chapter 4 a short introduction and explanation is given on the field of Machine Learning. In chapter 5, the applicability in Return management is discussed to see what the potential applications are for ML. In chapter 6, the processes and bottlenecks, defined in chapter 3, are discussed to see what the application of ML is on these bottlenecks. In chapter 7, this application is further specialized on KPI level. The potential impact and applicability is presented. Chapter 8 contains the conclusion of this study and presents directions for future research. It also reviews this study and presents the limitations and implications. Figure 4 presents the research flow diagram for this study,

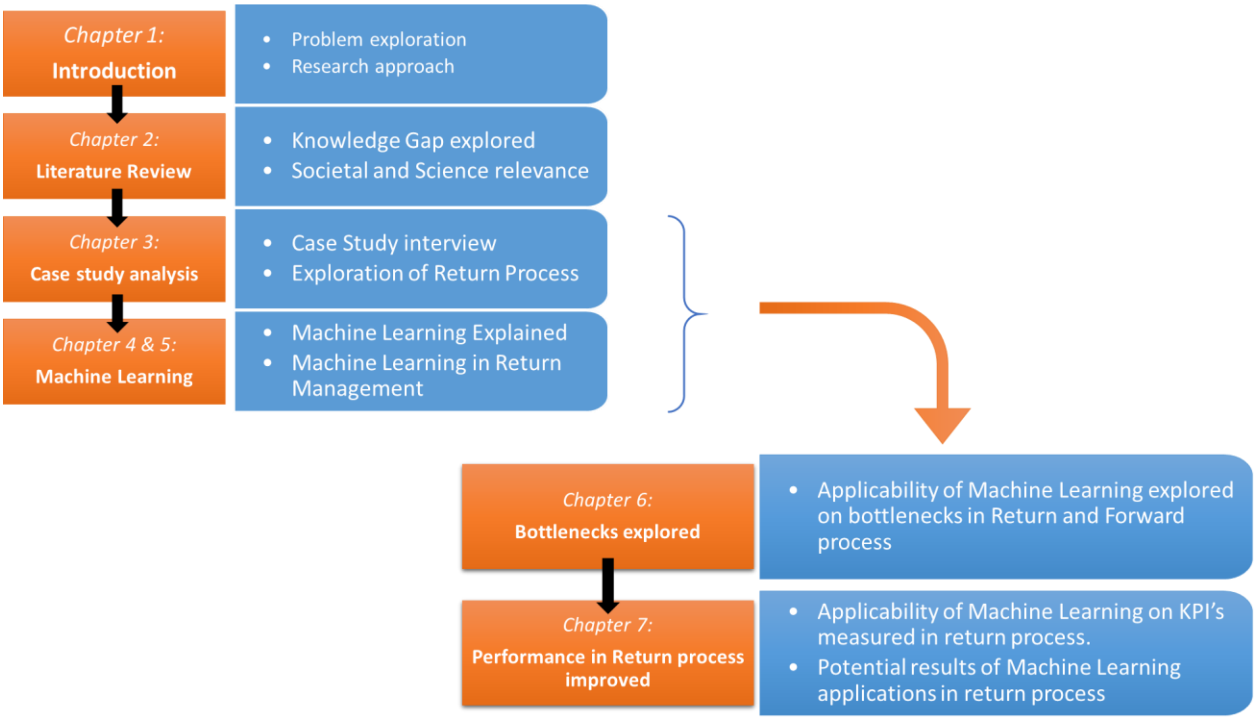


Figure 4, Research Flow diagram

2 Theoretical Landscape and Literature Review

For background knowledge and a steady basis of this research a literature review has been conducted. Research into the existing knowledge provides insights on what is still left to explore. Several search techniques are applied to investigate the different applicable fields of AI within SCM. Focusing on the current state of AI techniques like machine learning, deep learning, predictive analysis and many more applications will indicate the existing usage in SCM. Specific parts and processes are analyzed to discover the potential of AI techniques in these areas. In this research one specific AI technique or application had to be chosen, therefore due to several reasons this research chose for Machine Learning. In chapter 4 and 5, this decision will be clarified and further explained.

2.1 Literature Review

This literature review mainly explores the fields within SCM. With considering the SC related issues as basis for this research, they can be used to explore the possible ML techniques. Therefore, the Literature review tries to identify gaps and problems in SCM. When they are explored, the ML techniques will be described and analyzed to see their applicability in Return Management and Supplier Selection.

The first trend that was shown in literature was the use of big data analytics within SCM. Big data analytics can be used to facilitate ameliorated decision making, more transparency and more efficient risk management (Akkermans et al, 2013; Lycett, 2013; Chae et al, 2014; Schoenherr et al, 2015). Transforming this trend into a fuzzier approach of deep learning, neural networks or fuzzy logic gives other complications and requires other standards. The requirement of data volume is essential in the appliance of AI techniques as well as data formats, standards and responsibility issues.

This literature review will focus on different related issues within SCM. Dealing with uncertainty and the conflicting arguments of the different stages within the supply chain is hard to manage. However, considering the whole SC structure is far too broad and extensive for a project within this time and budget. A specific and specialized field was needed to create a better advice and research outcome. In the literature part, two fields within scm are explored, later in paragraph 2.1.3 the knowledge gap is explained concluding from examined literature. Firstly, the process of Reverse Logistics or Return Management within scm is examined. The process of RL is the flow of products and goods that are coming back from the customer to the suppliers due to divergent reasons. Second, Supplier Relationship Management (SRM), this is the discipline of strategically managing the flows and relations with supplier organizations. Interesting issues in this field arise with the supplier selection and the resilience of the total sc. Different search terms are used; "Supplier Relationship Management" AND "Artificial Intelligence" or "Reverse Logistics" AND "Artificial Intelligence". Mainly Google Scholar and the journal of "Engineering Applications of Artificial Intelligence" on ScienceDirect are used to come to the best results. Snowballing through different papers is used to find more relevant literature on topics regarding the SC domains. The literature review also contributed in defining a substantiated research question.

2.2 Return Management / Reverse Logistics

Firstly, the concept of reverse logistics or return management will be examined to see their current state and their relationship to Artificial intelligence or Machine Learning. Reverse logistics is defined as:

"the process of planning, implementing and controlling the efficient, cost-effective flow of raw materials, in process inventory, finished goods and related information from the point of consumption to the point of origin for the purpose of recapturing or creating value or for proper disposal." (Rogers and Tibben-Lembke, 1999, p. 2.)

Literature, as well as businesses are having more and more interest in RL. Literature is focusing on being more sustainable and having a Closed Loop Supply Chain (CLSC), this will be playing a bigger role in the coming years due to focus on environmental footprints of supply chains (Kumar et al, 2008). Acquisition pricing and remanufacturing decisions must be made to achieve a more CLCS (He, 2015). Businesses are having excessive high costs related to their RM and try to deal with their high product returns (Rogers et al, 2017). In this research the focus will lie on the product return of customers in the e-commerce market. Customers that order three different sizes of shoes, try them on and send two back. In this process of Return Management (RM) there is much to improve on costs, efficiency and sustainability. Having a high return rate on products can be very costly for companies and can cause serious problems within the SC (Jayaraman et al, 2008; Ravi et al, 2005). The cost regarding to the return flows can go up to 30 EURO per returned item (Rogers et al, 2017). As mentioned before in the introduction, companies can learn from their return management and it can lead to changes in design and strategy. Feedback from the return department can help shaping the forward process to serve the customer better and prevent returns.

Not only from Business to customer (B2C), also in the Business to Business (B2B) side there is still a lot to improve on product returns. Online order systems require complete information to fulfil orders in an efficient way. When there is lack of information or no accuracy of the information, supplier and receiver will have miscommunication and this can lead to a loss of money and inefficient flows within the supply chain (Li et al, 2015).

Forecasting of the product returns is becoming more and more important. Demand forecasting supported by AI tools is broadly discovered and explored in era of digitalization. However, the forecast of product returns is still quite unexplored and hidden. Unraveling the possibilities of AI can be very interesting in this complex network design. What is the incentive for the customer to send back the product (Govindan et al, 2015; Urbanke et al, 2015)? Can AI techniques help in lowering the return rate or can it even prevent returns from happening. For instance, algorithms that enable possibilities to be more supportive in the buying process of the customer, can presumably lower the return rate. How? By making prediction about the customer preferences, helping them in choosing the right products and providing tools to make well-founded choices. When the customer made the decision to return the product, the process of returning starts. In this process there are many steps and challenges. AI techniques can help in automating or improving this process and can reduce costs and time consumption.

Companies are often striving for high sale volumes to create more revenue and profits. Personalized promotion and limited offers can help in boosting sales numbers, however relatively few companies include numbers of return in their performance metrics (Petersen et al, 2015). On the other hand, flexible return policies can help in attracting more customers (Jeng, 2017). It reduces customers pre-purchase uncertainty and will lead to a higher change of purchasing (Heiman et al, 2001; Suqelack et al, 2011). Having a zero-cost return policy can be an incentive for customers to buy the products, try them and if they do not like it, send it back. Given these contrasting interests it is a complex situation with continuous change and interaction. This makes it very interesting to see what AI or ML tools and techniques can bring in this domain. Concepts like Money-back guarantee (MBG) are often used to ensure the total refund for customers. The decision on MBG can impact the whole supply chain strategy and will have influence on decisions on all levels (Heydari et al, 2017). This is a strategic decision that can attract or push off potential customers, both on the B2B and B2C side. Many companies have struggles with identifying and managing their reverse flows, AI could be of support to create insights on the location and status of product returns.

2.2.1 Supplier Relationship Management

Secondly, the topic of Supplier Relationship Management (SRM) is examined and explored in literature. The business to business side is playing an important role within the sc. From 2006-2015 in the US alone, the e-commerce B2B was responsible for a volume of 5705,76 billion US dollars (statista.com, 2017). And the prediction for the B2B e-commerce market will be twice the size of the business to costumer market by 2020 (Wu, 2015). Having a huge potential, SRM can and will be affected by a lot of new disruptive technologies. AI in combination with SRM is already broadly applied and this literature review will give details on the current applications. SRM is included in this project due to the possible linkage with RM and to see what to potential can be in this collaboration. Including suppliers to be more sustainable, resilient and transparent can give beneficial effects for the whole supply chain (Minnema,2017).

Choy et al, (2003) describes the use of artificial Neural network in an intelligent SRM system. Considering the process of supplier selection and finding new suppliers by applying ANN techniques. A given input of different parameters into the hidden layer of an ANN comes up with conclusions about the different suppliers and gives a substantiated advice. Real data sets are tested in several industries to see how different techniques like, support vector machine, fuzzy logic and linear regression behave and predict on the test data (Branch, 2017; Cankurt & Subasi, 2015; Lima-Junior & Carpinetti 2016). Having more accuracy and confidence in supplier selection gives companies more trust in their relations and this can benefit the total SC. For the performance evaluation and prioritization of suppliers, Fallahpour et al, (2017) is using gene expression programming. This is a new technique that comes up with better results than several other techniques like, multi-layer perceptron neural network. They recommend using fuzzy numbers for the collection of the data to improve their results. Or including sustainable parameters when scoring the different suppliers. All these existing papers recommend new techniques to improve the results. The parameters used in the different articles are broadly the same and focusing on speed, money and reliability. Including other parameters, such as the cooperating in reverse logistics or sustainability could be an interesting topic to further explore.

Resilience is a concept that can play an important role in the process of supplier selection. When the suppliers of a company are not resilient enough, this will have a negative effect on the whole sc. Being more resilient against natural disasters, extreme weather conditions or political situations can lead to higher value for the customers. Companies now prefer their suppliers in countries with steady political circumstances, rather than cheaper unsteady areas. Not only parameters regarding to speed, velocity or money are important, the underlying factors can play important roles too. Rajesh & Ravi (2014), describe the use of "Grey relational analysis based on linguistic assessment of supplier rating and attribute weightings could judiciously be used under these situations to obtain a set of possibility values for prioritizing supplier selection". Referring to the complex environment with incomplete data and uncertainties (Rajesh & Ravi, 2014). They include the sustainability concern of the suppliers and try to have a broad field of the different parameters. Applying a wide range of parameters within the AI techniques can give more reliable and realistic results. This can lead to better and more trusted relations that will lift the SC to a higher level.

2.3 Knowledge Gap

Relying on the comprehensive literature reviews conducted in the stated and cited papers, this research proposal gives a good indication on the current literature available in the different fields of supply chain management. By approaching the different issues of the SC, this broad approach gives the ability to further specify on a subject that is of real importance and potential in combination of the applications of AI. Where the research on AI applications in SCM is a rapid evolving industry, there are still area's left to discover. Specific knowledge on detailed tasks in the SC are lacking behind and it seems to be that the literature is more on the general level. The literature touches upon areas where AI or ML can be applied and give advice and consult on these applications. For instance, the use of Machine Learning in inventory management and predictions. The more practical and operational level is lacking behind and this gives the indication that the knowledge is outsmarting the practical reality on the work floor.

By describing and examining the different fields in the SC, (Return management and Supplier Relationship Management) a more complete representation of the literature is sketched and this led to the following research gap. The rapid growth of e-commerce markets is causing a problem in return numbers. This problem must be addressed and eventually there is need for improvements and solutions. A possible improvement can lay in the usage of AI/ML and agreements with the suppliers. Divide the costs and labor intensity throughout the whole chain can benefit both customers, sellers and suppliers. This will be further discussed in the following chapters of this study. Return management is becoming more and more important and at the same time more complex. Customers are buying more goods and products online and online retailers see the benefits of not opening physical stores (Gefen et al, 2003). Especially the e-commerce market on the B2C side is having a hard time in dealing with their return management. Now the responsibility of RM is focused and placed by the last stage of the SC. However, can RM be better managed when there are upfront decisions made in the supplier selection and partnership that can contribute in managing this process in a more efficient way? This can lead to significant cost savings for e-commerce parties and can help the whole supply chain. Huge e-commerce players in The Netherlands, especially in the retail market, keep on producing negative year balances (Salden, 2016; Solarz, 2016; Wellens, 2017). One of their major concerns is the high return rate. This points out the importance and relevance of this research.

The concerns in RM in the e-commerce market is a possible interesting topic to be further analyzed and explored. Due to data driven attitude in the e-commerce market, these players are particularly suitable for the application of disruptive technologies. In this study all problems, bottlenecks and processes related to RM will therefore be analyzed to identify possible applications of AI. To see what the possible application is on performance improvement in the return process. Not only the supplier selection can influence returns, there are also other reasons possible for high return rates. Customer behavior, return policies and customer contact can have impact on the return rates, both positively and negatively. If this is managed in an efficient way, e-commerce companies can reduce their returns and retain higher profits. The most lucrative business in the field of AI is in the personal advertisements (Ng, 2018). By analyzing browse behavior, clicks and personal interests, marketers are becoming better and better in proposing the products people want. This can result in a high sales rate on the advertisements. By knowing what the customer wants and know how they behave, certain predictions can be made. This can be an interesting opportunity and application to prevent returns. A study by Ding et al (2016) showed the relationship between social network posts and the product returns. Social media posts that are positively about a product will lower the return rate on a certain product. This study shows a new insight in the field of returns and makes it possible to predict returns by analyzing the social media platforms (Ding et al, 2016). It still is an undiscovered area, where over 60% of the retailers (surveyed by Persersen & Kumar, (2015)) did not consider the customers product return rate and behavior when determining marketing strategies. The return strategy will become undeniable for the upcoming e-commerce players, they will have to come up with new strategies to tackle their returns and manage them efficient. Do educational status, income, sex or age have impact on the returns? Due to the available data all these information inputs could be considered to cluster groups with certain return rates. Where researchers conducted studies in these fields and try to find causal relations between factors like age or income, they did not find any relations (Hong and Pavlou 2014; Petersen and Kumar 2009). There is a change that ML algorithms can detect certain patterns that are not detectable by humans. Studies from Minnema et al (2016) and Hong & Pavlou (2016) show different results on the effect of gender on return rates. Where Minnema discovered a significant difference, Hong & Pavlou did not experienced this difference in their study. It could be differentiating through the geographical location, maybe people in Europe have different return behavior than people in Japan or Amerika. This can be caused due to legislations, in EU customers have the right to return for free. Other factors can be transportation and delivery, the time and ease of returns can influence the return behavior of the customers (Schulze & Srinicasan, 2016).

Where literature is writing that companies are more and more focusing on the whole SC and stating that the focus on the internal manufacturing is not enough anymore (Mahendrawathi et al, 2014), the reality shows a different image and outcome. Unfortunately, there is no good communication and management between the different levels and suppliers in the SC regarding to RM. To specify a more precise kind of market structure, this research will focus on the e-commerce market on the B2C side. Which techniques can help in these complex processes of RM and how can the participation of the different suppliers be managed? Most simple ideas in AI can be most powerful and effective to support

SCM, RM and supplier selection. For instance; taking pictures of all the goods that are coming back from the customers, weigh them, matching this in datasets to see from what supplier it is coming and sending it directly through with minimal human intervention could be an application of AI to help in the process of RM.

2.4 Societal and Scientific relevance

Explore ML techniques in a new field of appliance will have societal and scientific impact. Both societal and scientific relevance will be explained in this paragraph. First the societal relevance will be discussed. By exploring new ways of applying the ML techniques, actors in the field can see and discover what ML can bring in this field of the supply chain. They can see what the possibilities, bottlenecks, requirements and issues are when ML is used. The proposed techniques can be powerful and revolutionary in the field of returns and can have a huge impact. By providing insights it will be more tangible for both businesses and practitioners to apply these ML techniques. The other societal relevance lays in the field of the supply chain, where the web shops must deal with the returns and are making huge loss. EU legislation oblige them to take back returns for free and this makes it very hard for them to manage this efficient and with profits. A 'monster' is created and customers are demanding higher customer service and lower prices. Helping these e-commerce players in reducing and preventing the returns can have big impacts on their business performance.

In the scientific field, the relevancy lies in the new knowledge that will be created. Current literature is hardly combining new disruptive technologies with the issues evolving around return management. Trying to apply new technologies in this field will add new knowledge on existing literature. It will explore new possibilities in the broad field of AI. This project also tries to close the gap that is created between literature and businesses. The literature is far ahead in the field of AI and the business is lacking behind. In SCM there will be many applications possible for AI, ML and other disruptive technologies and therefore this study tries to close this gap.

Speaking about ML or in broader sense AI as a disruptive technology, in short, the new institutional framework of Williamson (1995) will be discussed to show the relevancy of this study. This framework focusses on the underlying aspects of the traditional economic thinking. It tries to include aspects that were previously not included. One of the most important aspects in this framework are the institutions or "rules of the game". These rules of the game shape the behavior of society. People react and act on certain rules governed by institutional organizations. These rules change over time, certain techniques or developments can affect these rules. Therefore, it is needed to be adaptive and change fast as legislation and laws can lack behind on new technologies. Due to the rapid evolvement in science and technology research, laws will lack behind on these technologies. This can cause harm for society, there is no legislation in place to shape the behavior of these innovative ideas. Williamson describes four levels, in these four levels there are different regulations and legislations that guide the public. The frequency of changes differs from top to bottom. Where in the top section institutions changes once in a 100 or once in a 1000 year. In the lowest level the frequency of change is continuous. The framework is displayed in figure 5. This framework can perfectly show adjustments caused by disruptive technologies. Take AI as an example, it will completely change markets, laws, institutions and business. The rules of the game will be affected and it is important to prepare for this change. This framework is used to show the adjustments that AI can bring in society, it will show that governmental institutions must adapt and react on technological changes. Andrew Ng (2017) implies AI as the new electricity. How electricity changed the world around 100 years ago, AI has the potential to do this again. It not only changed the way of working, it also changed human behavior. However, these electricity prices had to be regulated and rules were needed to guide the society. These changes affected the political systems, legislations and labor markets. All four levels will be described to see the potential impact on society.

- Level one, (reading from top to bottom) is about the informal institutions, this can be customs, traditions and religions. The whole AI technology is based on artificial knowledge, this tries to create knowledge in machines and no longer knowledge in people. This is the opposite of the current norms and values where the human is placed in the center of knowledge, labor and basically everything. This change from human centered to machine centered can still take a long time but there is a movement started. As Andrew Ng (2017) is mentioning repeatedly, AI can be the new electricity where it will change many sectors, industries, markets and cultures. The attitude against self-learning robots and machines will change. Where now the people are quite reserved in communicating with machines. This can easily change in a couple of years when these machines are necessary in daily life. Even now, AI is applied in phones, computers and many other electronics that interact with human on daily basis. This will cause a change in norms and values and will change behavior of people, this is mapped as long term following the existing literature in this area (Williamson, 1995; Coase, 1998; North, 2003).
- In level two, Williamson (1995) shows the institutional environment with the formal rules of the game. These rules are

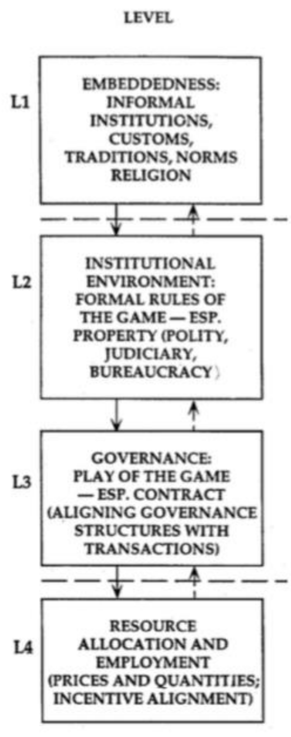


Figure 5, Four layered model (Williamson, 1995) shaped for the current structure and technologies. For new technologies they are lacking behind, this can be seen in real life when we are talking about military drones. These drones can act on their own and can perform enemy attacks on their selves without human intervention. The same with the AI technologies, who can be held responsible for the actions and decisions the machines are taking? Looking towards the applications in the RM process, AI will probably do no harm. However, self-driving cars are also based on AI and their decisions can have fatal consequences. Current legislation is not enough and there is a need for adjustments in the current rules of the game.

When speaking of Return processes, at this moment the EU laws oblige the selling parties to take back all items that are sold damaged or deviates from it product description. When customers bought products in physical shops, these terms were fair and easy to handle. However, with the upswing of online shopping, the customers have less hands-on experience with the product and therefore return products more often. These EU laws and regulations could therefore be changed to meet the selling parties and help them in their return process. The current laws and regulations oblige them to take back all goods, sometimes without even a valid reason. This makes it hard for them to be profitable.

- In the third level, Williamson speaks about the formal and informal institutional agreements. These contracts are tended to be changed due to several new possibilities AI can bring. New ways of setting up contracts and communication between parties are possible. Even a machine can have certain parties as contacts and can be involved in the new contracts. In the field of RM these contracts can also changes. Due to the more important and strategic role of RM, these contracts can be changed and influenced by the return policies of companies. Agreements with suppliers that consider the return rates can be beneficial for the whole sc. In this third level Williamson speaks of the plays of the game. How to get these contracts in line with legislation and how to deal with changes?
- In the fourth level, Williamson speaks about the individual level, where it is about the connections between actors. The daily routines to get the marginal conditions right. These new technologies can change interaction, for instance the Blockchain technology can lead to different ways of transferring goods or money. The daily routines can be hugely influenced by new technologies. One of the applications of AI in RM can be the processing of the incoming returned items at the warehouse. This can be a daily routine for an employee however, it can be easily changed by a new technique. Think of robotics or other automating software.

This framework does not only operate top down. For responsible innovation it is important that the different layers interact with each other. Changes in lower levels can have effects on higher level institutions and therefore both bottom up as top down approaches can be applied. This framework can help in being responsibly innovative and help in transitioning new technologies into society. AI can influence the different layers of the new institutional economics model by Williamson. Not only the technological design has to be considered, the institutional and process design are also important aspects to consider (Koppenjan & Groenewegen, 2005). This research will mainly focus on process and technological side, but the institutional side must not be forgotten and will play an important role in making AI successfully integrated into our world.

3 Case Study; Return Management in e-commerce

In this chapter the case study analysis is presented. Through the conducted interview, all information about the return processes, supplier selection and KPIs is obtained that will help in the analysis of this study. To combine the knowledge from literature and business a total overview is created that will show the bottlenecks, processes and KPIs that can be interesting for Machine Learning techniques. The three questions that will be discussed in this chapter are:

Sub Question 1: What are the current components of a return management process from an e-commerce player in the Dutch electronics market?

a. *What is the total process of Return Management within the given case study and what are the bottlenecks that occur?*

b. *Is there a link made between Return Management and Supplier selection and if so, what does this link contain?*

c. *What are the relevant Key Performance Indicators measured in the return process?*

In paragraph 3.1, first a theoretical background will be given to indicate the process of returns. The difference between the forward process and the return process is explained. In paragraph 3.1.1 and 3.1.2 the main insights from the case study analysis are presented. This is based on the information from the interview and composed by the researcher of this study. Paragraph 3.2 will answer sub question 1b, where the link between return management and supplier selection is indicated to explore their relationship. This relation with the supplier is crucial for a successive way of handling the product returns. In paragraph 3.3, sub question 1c will be answered. First, the literature is explored to obtain KPIs and measurement frameworks from previous studies. These are combined with the current KPIs from the case study which led to quartet of KPIs in the return process.

3.1 Return Management defined

The process of return management is complex and it contains a lot of different stakeholders and intersects. In this research we define Return management as:

"the process of planning, implementing, and controlling the efficient, cost effective flow of raw materials, in-process inventory, finished goods and related information from the point of consumption to the point of origin for the purpose of recapturing value or proper disposal". (Rogers and Tibben-Lembke, 1999, p. 2. ; Reverse Logistics Executive Council, 2007)

This indicates all aspects of a normal supply chain, the only difference is the direction. The return flows are starting from the end customer and can return all the way back to the suppliers or producers. The return flow is huge, however this research will focus on the return flow between the end customer and the e-commerce player. In figure 6, a basic supply chain process is presented.

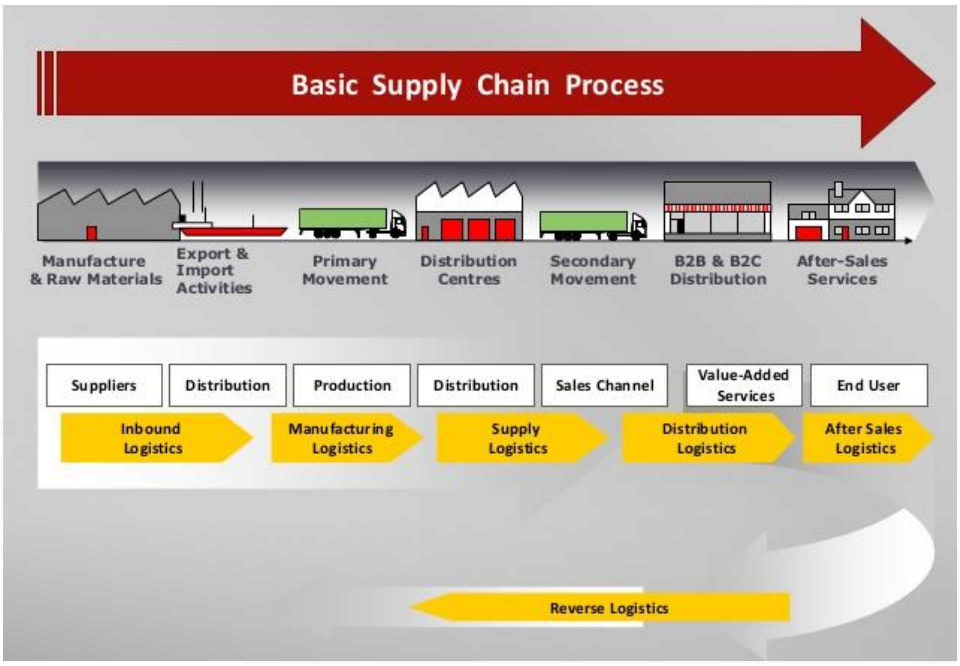


Figure 6, Supply Chain Process (Topno, 2014)

This figure shows different parties involved in the process that are needed to produce and sell the product. The start of the supply chain is at the manufacture or supplier, they have the materials or produce the products. Than the products go to a distribution center or warehouse of the selling party. In this research the e-commerce market is studied, these parties sell majority of their products directly from their warehouse. The last step is the after sales services which includes the customer.

A product can already be returned from the manufacturer to the raw material producer due to bad quality or wrong delivery of goods. However, these return flows are not the focus of this research and happen less often compared to returns from the end customer. It is focused on the last stage of the supply chain, the business to consumer relationship. In this level of the supply chain, returns are playing a very decisive role and can affect the B2C relationship both positive or negatively. The relation between the e-commerce player and the supplier will also be analyzed, to examine what the effect of this relation can be on customer behavior. The three most important stages of the supply chain for this research are indicated in figure 7.

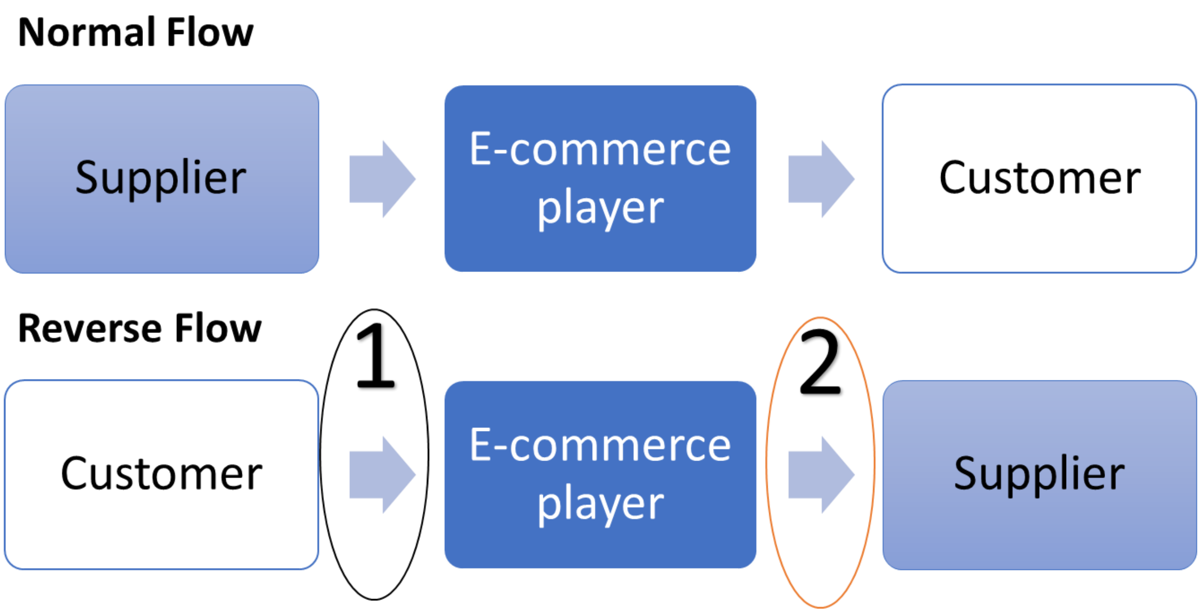


Figure 7, Normal vs Reverse Flow in the supply chain

In figure 7, the supplier block is representing the manufacture and delivering parties of the products. Within the e-commerce player block, the warehouse or distribution center is indicated. The customer block is indicating the end customer of the supply chain. The focus of this research will be the process and relations between customer and e-commerce player, indicated with a 1 in figure 7. Not only products flow from customer back to the original owners of the goods. Flows of information and money must also be returned to the previous stages in the supply chain. Where the normal flow is demand driven, a customer orders a product at the e-commerce player which has it in stock or has to order it at the supplier. The reverse flow is unpredictable and the e-commerce players and suppliers do not know which products will come back. This makes it a hard to manage system that requires flexibility and adaptiveness of the stakeholders. Figure 7 has two the number 1 and two in the reverse flow, both indicating a relation. In flow 1 mostly products and information are transferred. This relation is the focus of this research and is customer to business. Flow 1, between the supplier and e-commerce player will mainly be a flow of information and agreement. There will be less focus on the product flow between these parties.

Product returns happen due to different reasons; products are damaged, or wrong delivery etc. When the goods arrive in the warehouse, some products can be taken directly into stock, these are the products that are still new and unopened. Some products can go to the trader because they cannot be sold anymore at the current store. Maybe the product is damaged and need to be repaired or remanufactured. The product can also be recycled, this happens more often when the product is older. These returns for recycling will not be considered in this study. The last stream is for the products that are not sellable and useable anymore, they will go to the disposal. Figure 8 displays the different possible reverse flows.

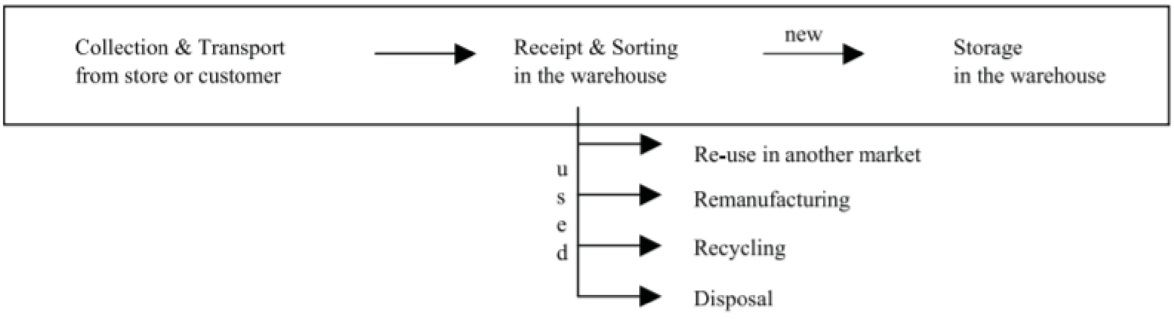


Figure 8, The process of Returns (René et al, 2002)

In most cases the return process involves more complexity than the forward process, figure 9 shows the differences between the returns and forward processes and the complex issues related to returns (Ronald et al, 2002).

Forward	Reverse
Forecasting relatively straightforward	Forecasting more difficult
One to many transportation	Many to one transportation
Product quality uniform	Product quality not uniform
Product packaging uniform	Product packaging often damaged
Destination/routing clear	Destination/routing unclear
Standardized channel	Exception driven
Disposition options clear	Disposition not clear
Pricing relatively uniform	Pricing dependent on many factors
Importance of speed recognized	Speed often not considered a priority
Forward distribution costs closely monitored by accounting systems	Reverse costs less directly visible
Inventory management consistent	Inventory management not consistent
Product lifecycle manageable	Product lifecycle issues more complex
Negotiation between parties straightforward	Negotiation complicated by additional considerations
Marketing methods well-known	Marketing complicated by several factors
Real-time information readily available to track product	Visibility of process less transparent

Figure 9, Differences in Forward and Return Process (Ronald et al, 2002)

As can be seen in figure 9, there are many uncertainties in the reverse flow, where different factors such as, difficulty of forecasting, no uniform product quality and less transparency in the process lead to complexity. The product return rates are increasing due to online sales, and product quantity and diversity makes it hard to handle (René et al, 2012). To manage this complex dynamic environment, there is a need for good process management, collaboration between different parties and transparency throughout the whole supply chain. Returns were probably one of the most under estimated aspects in the supply chain. However, it is receiving more and more attention from literature and businesses through the rapid growth of e-commerce markets. Zalando Netherlands and Wehkamp are huge online retail players in The Netherlands, however they are still not making profit due to the high return rate (Wellens, 2017). They are obliged to take back all the returns, free of charge and repay the full amount of money to the customer. Due to these strict regulations from European Union it is hard to deal with these returns. However, the web shops themselves created this 'monster', they were willing to replace the fitting room to the customers home. Clothes are a very specific product that can easily misfit a person's size. Therefore, in the retail industry the product return is outrageous high. Not only in the retail industries online web shops are dealing with these high returns, also in electronics and furniture markets some products can have high returns. To analyze the return flows a case study is taken to sketch the process, to see what KPIs they are measuring and how they deal with and fight against the returns. This case study is taken as an example for a broader industry, the KPIs and the process will be used to generalize so that it can be used for further research. From both literature and business there were similarities in the measurement of KPIs. There are still companies that do not measure their return rates properly and underestimate the impact of returns. On the other hand, there are companies that take the return process very serious and try to measure important numbers and percentage. In this way, literature and practical examples are combined to give a total view on the return process in the e-commerce market.

3.1.1 Case study; Return process presented

In this paragraph the process will be explained using the Business Process Modelling Notation (BPMN) approach. BPMN maps the business processes into a graphical representation (Silver, 2011). This modelling approach emerged through several groups that were working together to create a universal language to map business processes. BPMN is strictly used for business processes and can be used for organization structures, complex business processes or data models. It consists of mainly four elements; first, Flow objects, these are the events and activities that shape the process, second, connecting objects, these are the connections between the different flow objects. Third, the swim lanes, these are representing the different stakeholders that are involved in the process. Fourth, artifacts, these indicate the data or groups that are detectable in the process are mapped. The BPMN approach is used due to its strong presentable character. It is easy to read and gives a clear overview of the total return process. It is accepted broadly in literature and used often to make simulations and models to improve the existing processes (Silver, 2011).

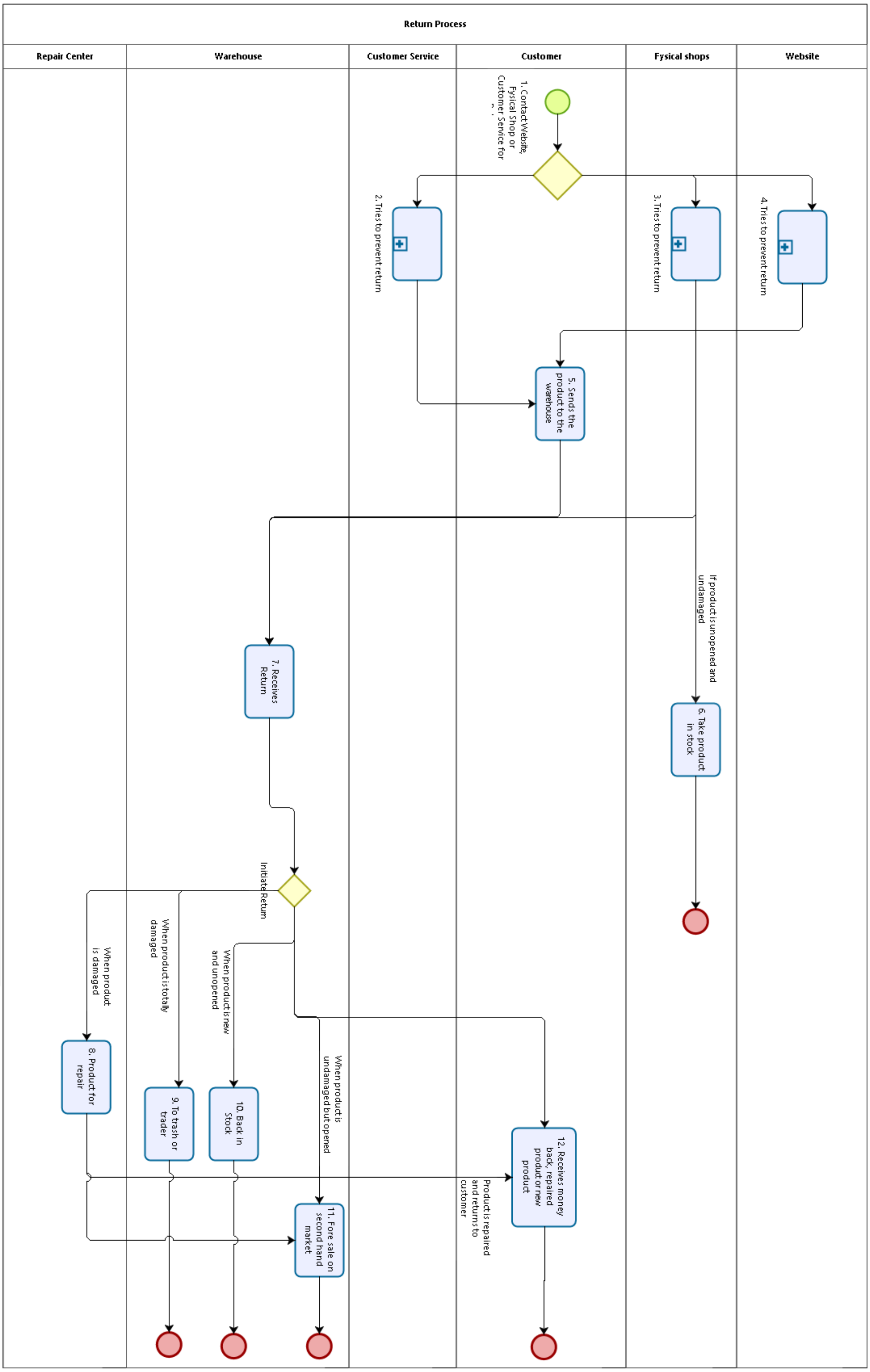
The case study is conducted with an e-commerce company that has a website and a small number of physical shops. Most purchases by customers are done online, via the website. With a total revenue of more than 1 billion euro, it is one of the major players in The Netherlands. Further indications about numbers of orders and returns are indicated in table 1.

Sale of goods	Average amount per order	Number of orders	Percentage of Returns	Total returns	Returns per day	Total value of returns per day
1,1 billion €	320 €	3.437.500	9%	309.375	988	316.160 €

Table 1, Number of orders and returns case study

Table 1 is an estimation based on the numbers published by the e-commerce party. They did not provide information about the "Average amount per purchase order", this estimation is based on market average in the Dutch e-commerce market. Only 20 % of the product returns are via the physical shops, the other 80% goes via the courier. Approximately 800 product returns per day are arriving at the warehouse of the e-commerce player. This high number of returns was one of the reasons to select this company as a case study. Reducing the number of returns, reducing the processing speed of returns or reducing labor intensity in the process can have significant impact on total business performance.

The BPMN is used to show the different flows and processes of returns. This BPMN model is based on the interview conducted with the e-commerce party. By using the BPMN, the return process from customer to e-commerce player is visible. It can give a good indication on where bottlenecks arise and where room is for improvements. The process is shaped based on the interview and is presented in figure 10. This model is composed by the interviewer based on the received input in the interview. The numbers are indicating the steps in the return process and will be shortly explained to indicate the importance and complexity of each step in the process. This is the general return process that is applicable for most product returns (90%) within this company. Due to the wide range and variety in product assortment there are exceptions in this process, however those are not relevant for this research and therefore scoped out of this study. After the total process is mapped, the bottlenecks of the process will be further discussed in paragraph 3.1.3. These bottlenecks were mainly discovered through the interview. Literature is explored to be complementary on this case study.



13. Customer Service: Help along the way when customers call for their status of return

14. Contracts and Billing: Credit the amount back to the customer if they want their money back. However this process is left out of the scope of this project

Figure 10, Business Process model of case study

1. Customer contacts or website, or physical shop or Customer Service for return.

The start of the process is indicated by a request from the customer to return the product. This can be due to several reasons, the most common reasons are; damaged product, satisfaction warranty and wrong delivery. The customer can decide via which of the three streams they want to return the product. As stated in table 1, this can go over 1000 products per day. If all these customers contact the service desk, it will require a lot of labor intensity of the employees.

2. Customer Service tries to prevent return

When the customer decides to contact the customer service, mostly via telephone, the employees at the customer service will try to help the customer and prevent the return. When the return cannot be prevented, they will try to sell another comparable product to this customer. When the customer still wants to return the product, they will help in providing a retour label for the product. 40% of the returns are via the customer service. This means around 400 customer contact moments per day relating to a product returns.

3. Physical shop tries to prevent return

When the customer decides to bring back the product to the store, the employees in the store will try to prevent the return. This can be done by helping the customer in understanding the product or installing the product. When the return cannot be prevented, they will try to sell another comparable product to this customer. The returned product will be taken in and will be processed in the shop. This sub process can be seen in figure 11. Returns via the physical shop are 20% of total, this is around 200 products per day.

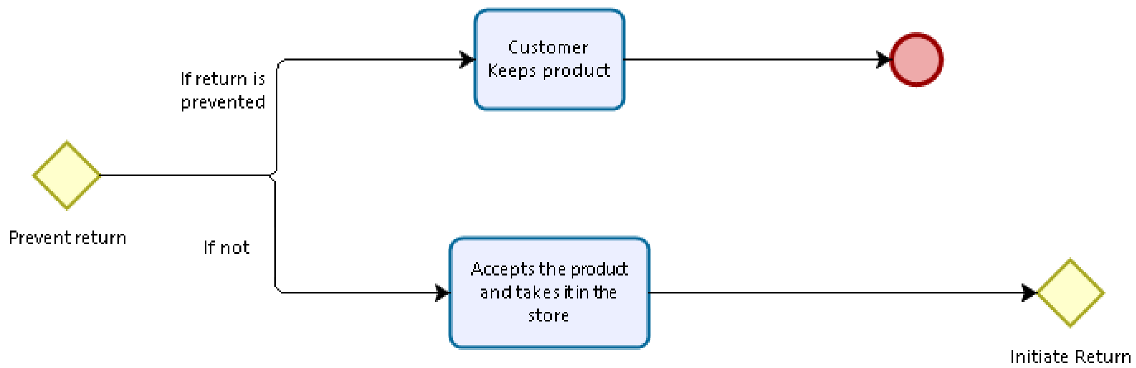


Figure 11, Sub-process of return prevention

4. Website tries to prevent return

When the customer decides to search on the website for a return label, the website will help the customer in returning the product. However, here are also possibilities to prevent the return or to sell comparable products. Customers can login to see their products or can insert their order number to fulfil their return. Via the website, 40% of product returns happen. This is the most desirable return option for the e-commerce player due to the automatization and the data that comes available. When a customer returns the product via the website, the e-commerce immediately know that that specific product will come back and what the reason is for that return.

5. Customer sends product to the warehouse

The customer sends back the product, this can be done at a local pick up point or the customer can let the carrier come to pick up the good at home. This happens around 800 times a day, however due to weekend business or other influencing factors this number can strongly fluctuate. This makes it very hard for the e-commerce player to predict their returns.

6. Physical Shop takes undamaged product back in stock

When the customer returns an unopened and undamaged product, the store can tack it back in their own stock and can sell it directly if it is in the assortment of the shop. If the product is not sellable in the physical stores it could also be redirected towards the central warehouse. Only 20% of total returns is returned via the physical shop.

7. Warehouse receives return

The warehouse employees receive the returns from all the different channels. The next step is to initiate the reason of return. In this stage the employees indicate if the product return is legit, what the reason is for return and what the customer desires. Based on these inputs, the employee decides what to do with the package and what further action needs to be taken. The information system must be updated to indicate and approve the reason for return. This is an important and time-consuming step in the return process. All goods must be indicated correctly to determine the next destination and handlings. Around 800 product returns must be indicated per day by the employees.

8. Product to Repair center for repair

In most cases, the company has contracts with their suppliers to let the products be repaired in local repair centers. When the repair takes too long, the customer will receive a new product. When the repair is on time, the product can go back to the customer. When the customer does not want the product anymore it can go to the second-hand shop or the trader.

9. Product goes to trash or trader

This is the least desirable process. However, when products are significantly damaged and cannot be repaired or products are not suitable to enter the second-hand shop, they will be transported to trash or a trader.

10. Product is taken back in stock

When the product is unopened and undamaged, the product returns immediately back in stock and is ready to be sold again. This is the primary stock of the company. From this stock they keep track of their sales volume and plan their updating on stock levels. There are basically two selling points for this company, the original stock and the second-hand market. Product repairs return to customers or can go to the second-hand market. Unsellable items or heavily damaged items will go to the trash or trader.

11. Product is for sale on second hand market

When a product is repaired or opened by the customer, the product will go to the second-hand market to be sold on a discount price. Via this way the e-commerce player can retain their margin on the product and can still sell the product without making a loss.

12. Customer receives full refund, new product or repaired product

Depending on the customer demands and the agreements the customer receives a refund, their own repaired product or a new product.

13. Customer Service helpdesk

The customer service helpdesk handles questions from customers that are dealing with a return. They have questions about the status of the return or an asking for help. This process is inherently connected with the return process and therefore drawn as a total supportive block.

14. Contracts and Billing

This department is responsible for the refunds for the customers. Also, the exceptions are handled and supported by this department. They act along the whole return process and are therefore also drawn as a supportive block that is linked with the process.

The representation of the total return process in figure 10 is created by using the information from the conducted interview. There is no data available about the percentage of returned products that go to the different streams in the warehouse. Probably most products will be sold on the second-hand market or are taken back in stock. However, this process is less relevant in this case study. The step of initiating the return is far more important and can possibly be impacted by Machine Learning. Through the interview, some bottlenecks and issues in different steps of the return process were discovered. These bottlenecks will be further explored in paragraph 3.1.3.

3.1.2 Forward vs Return process

One of the most valuable assets in a return process is the data of the customers return (Minnema, 2017). What product is coming back, when is it coming back, from where, what is the reason for return etc. All the available information is valuable for the e-commerce player and give them more control over the process. They know where the product is and what the reason is for the return, in this case they can already determine where the product needs to go and what the total processing time will be. Being in control will help in making the process more efficient and will lead eventually to a happy customer. When a customer calls and asks the status of their return, a simple direct answer must be given by the customer service. Certain decisions can influence the ease and customer friendliness of the return process. For instance, the decision to include a return label within the package of the product on arrival is a very important one. When the return label is included and the customer decides to send the product back, the company cannot predict or monitor this return until it has arrived in their warehouse. When a customer first approaches one of the return channels of the company (Website, customer service or physical shop), the return will be known by the company and they have the knowledge of the product returns. However, it is an important consideration to include the return label or not, including the return label can ease the process for the customer, and an easy return process will lead to a higher come back rate of the customer to the same web shop (Minnema, 2017). An important indicator for this trade-off is the return percentage of a product, when the returns are high for a certain product, it can be more beneficial to include the label, however when a return percentage on a certain product is relatively low, it can be more beneficial to not include the label. Including the label will create a faster process and probably higher customer satisfaction. It can benefit the e-commerce player indirectly through higher conversion rates in the future. This is one example of a decision that is made before the returns are happening. The prevention of returns is a very complex, challenging and very important process that can lead to a reduction in costs. This thesis will explore the possibilities of ML in these two separable processes. First, how can returns be prevented with the help of ML and secondly, how can the process of returns be impacted by ML to make it more efficient. This distinction can be seen in figure 12. Some research has already been done into the reasons why people return their goods. This can be due to several reasons, for example, too positively written customer reviews can lead to higher returns (Minnema. 2017). The customers will have a too positive image of the product and it will disappoint them upon arrival. Therefore, product descriptions are of huge importance, they must be clear, honest and truthful about the product to give the customer a fair choice in purchasing the product. Another strategy to avoid returns can be in the contracts and

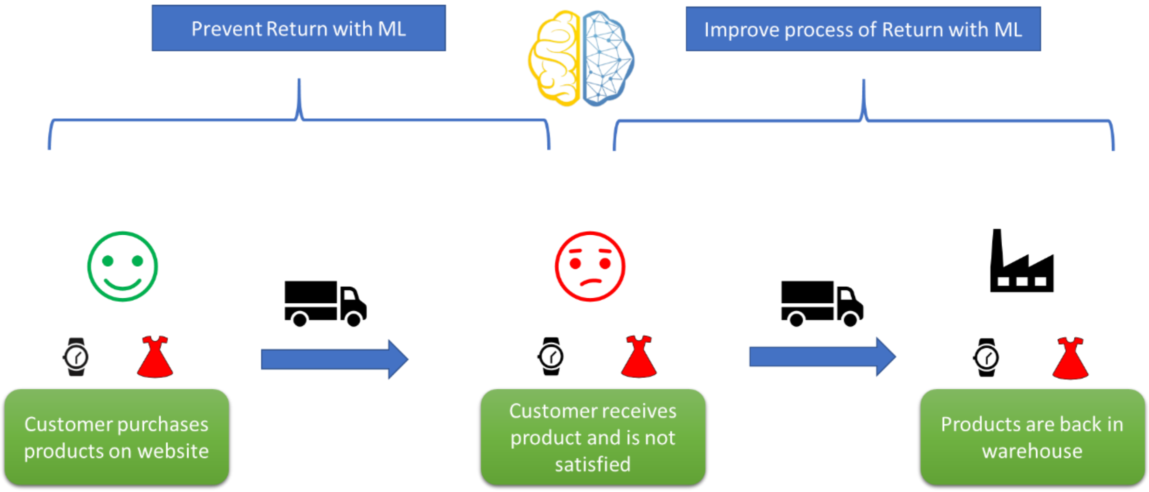


Figure 12, Distinction between Forward and Return process

negotiating with the suppliers. Prevent selling products with high return rates or bad quality, these products will lead to higher product returns. On the other hand, expensive products are returned more frequently than cheaper products. Customers are being more thoughtful when it comes to bigger sums of money, this leads to a higher return rate on expensive products (Blauboer, 2018).

These are all considerations that will influence the return rate. The return process of the product starts with the customer that decides to return the product. It can return the product via several streams, where in this case study, around 80% of the products return to the warehouse of the e-commerce player. The other 20% is returned to a physical shop. Amazon handles every purchase in less than 60 seconds of human labor. The rest of the work is done by robots and automated conveyer belts (McFarland, 2016). This is extremely fast and this means that a person can do over 500 packages a day. However, how many seconds are needed from a person in the return process? Unfortunately, there are no insights about these statistics found in literature. And probably businesses themselves do not know this handling time either. As we look at the return process, it can be quite time consuming and costly. The product will return as a package and must be unboxed. The employee needs to read and initiate the return and check where it came from and where it needs to go. Probably the person needs to scan the product or update the system and will deliver the product to the next step in the process. Maintaining this process in an efficient way is hard for companies, it is costly and labor intensive.

A consequence of these high costs in the return process are the banned accounts at Amazon. They ban accounts that return too often to lower their returns. (Brignall, 2017; Safdar & Stevens, 2018). The Wall street Journal came out with a huge report about the blocking policies of Amazon. Even a customer friendly company like Amazon has it limits when it comes to returning products. This shows once again the essence and the economic value of the return process.

3.1.3 Bottlenecks

In the return process, four possibilities for improvement were detected. These bottlenecks are discovered from the interview that was conducted with the company of the case study. The input from this interview is used as a basis, and explored literature and insights will be added upon this to be more complementary. It is restricted to four bottlenecks to keep the in-depth character of this study instead of the comparative view. Bottlenecks arises everywhere, however these bottlenecks are mapped due to their huge impact in the organization. By improving these bottlenecks there will be a measurable change in the organization that can lead to more efficiency, less costs or more control in the return process. Clearly, there are many more bottlenecks to further analyze and improve, however due to their small impact they are scoped out of this study. For every bottleneck, the concerning part of the return process from figure 10 will be added to give a visual overview. This will show the place of the bottleneck in the return process.

1. Customer contact – *How to reduce the customer contact with customer service while the customer satisfaction remains high.*

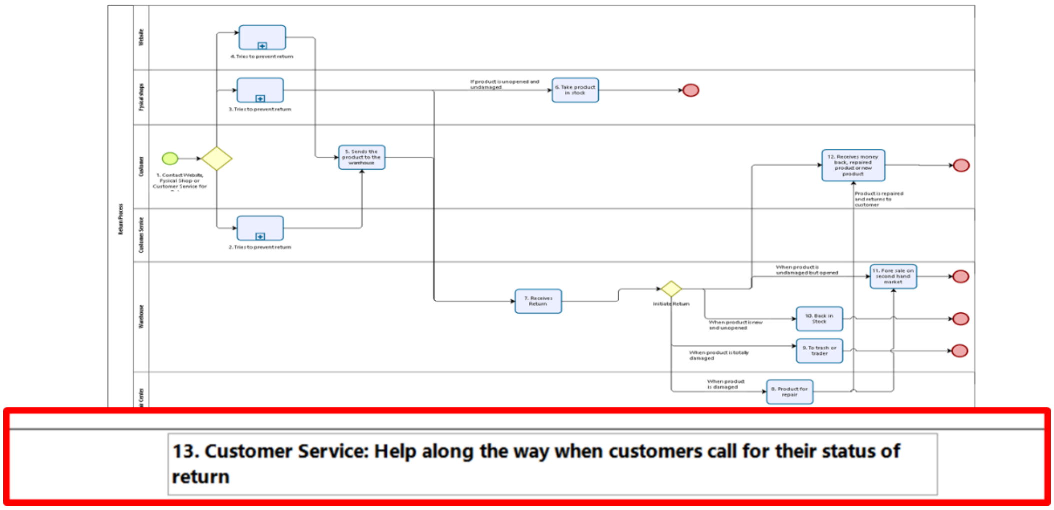


Figure 13, Customer service in return process

The customer contact moments are spread over the whole return process. However, most contact happens in the request process of the returns. Customers call the customer service to ask how to return the item. Most e-commerce players in The Netherlands already provide a return document on their website. By using this document, customers do not have to contact the service desk for their return. This resulted in a reduction of 50% in customer contacts (coolblue, 2018). However, there are still too many customer contact moments in the return process. Customers can also call the service desk to inform about the status of their return. Customers want to know where their goods are, what their status is and when they can expect to have their money or product returned. This can be a very costly and time-consuming process for the service desk of the company. They must deal with continuously returning customers and they are asking most of the time very easy (repetitive) questions. The customer contact department is costly, there must be employees that can support clients and help them with their questions. Certain highly specified questions can be hard to answer for a single person and it can take a long time to come up with the right answer. This makes it a time consuming and labor-intensive process, which leaves much room to improve. Exploring the possibilities of ML in this process can probably increase customer service or decrease labor intensity. However, reducing the labor intensity can lead to less customer contact, and customers can feel disappointed and ignored when they do not speak to a person. This trade-off is important to consider.

2. Costs and time in the return process. Fast process is wanted to keep the costs low and the time to inventory short. However, this means less control. How to deal with this trade-off?

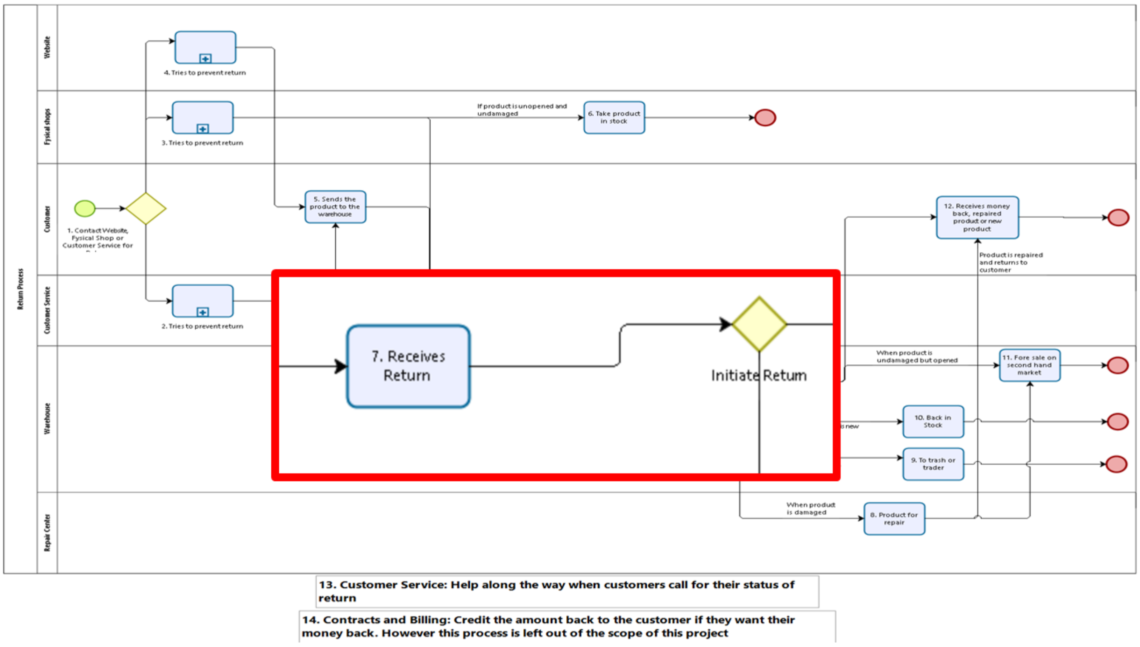


Figure 14, Costs efficiency in the return process

Figure 14, indicates the process of receiving and initiating the return in the warehouse. Around 800 products return every day to the warehouse and must be indicated and further processed. This is time consuming and is done manually by the employees. Information about the origin of the product, customer data and other issues related to the return must be searched out. The information and purpose of the return varies greatly and demands for very specific trained personnel. Another point for improvement in the return process can be to reduce the intermediate steps that are taken. Now the process can be quite long and all returns are handled in the warehouse of the e-commerce player. This takes extra time and transportation costs, reducing the intermediate steps by shipping more directly from customer to supplier can shorten this process.

Making the process of returns shorter can be less costly and can possibly lead to faster refunds of the customer. This however, will diverse the control of the returns and this can be harder to manage. This means, on the one side there is willingness to shorten the process and make it as fast as possible. And on the other side, the control over the returns is important to guarantee customer service. This bottleneck concerns the total costs and total time of the return process. These are still too high and must be lowered to retain more value out of the returned products. Further analysis on this bottleneck will be presented in chapter 6 and 7.

3. Agreements with the suppliers

The third bottleneck is related to the agreements and contracts with the suppliers. These suppliers deliver the products to the e-commerce players. Having clear and transparent agreements with suppliers are crucial for the e-commerce players. Contracts that consider the warranty and handling of returned products can help the e-commerce players in collaborating and spreading the costs. When the contracts are covering all costs of returns, the return flows do not have to be a loss-making process anymore. However, there is still much to improve in this area for the e-commerce players. In many products there are no full agreements on returns with the suppliers. This results in high costs relating to returns for the e-commerce player. When there are good agreements and well collaboration with the suppliers, the costs can be shared.

It can even be a process of retaining profit when it is well managed. This makes it a very interesting and complex issue to further analyze. Having bad and incomplete agreements with the suppliers can affect the profitability of a company. Some products are economical valuable and profitable for an e-commerce player, however when these products are very costly in returns the e-commerce player cannot maintain their profits. When there are no clear agreements on the terms of returns, the selling party, in this case the e-commerce player, will be responsible for the high costs and will have to cover it unaccompanied. This can negatively influence the relationship with the supplier. Therefore, issues around trust, transparency and full knowledge sharing are important to be resilient with all stakeholders involved in the return process. These agreements with suppliers are not indicated in the total return process, as mapped in figure 10 of paragraph 3.1.1. However, these agreements do influence the return performance strongly and are therefore considered as a bottleneck in the return process.

4. Forward process, what can be done in the forward process to prevent and reduce returns?

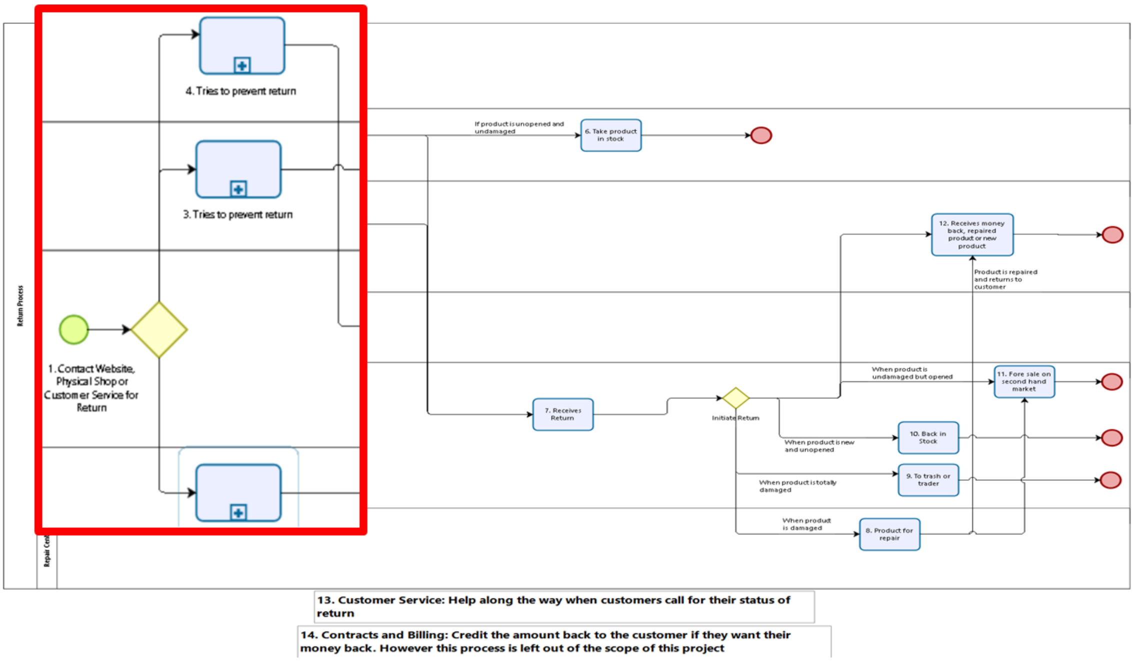


Figure 15, Prevention of returns

This bottleneck is related to the forward process. Having bad product descriptions, false information and bad images on the website can increase the return rates. The forward process is analyzed to see what the reasons can be for high returns. As can be seen in figure 15, the process of preventing the returns are indicated with a red border. However, the forward process is more related to the purchasing process of the customers. Their browser behavior and personal interests can have effects on the returns. Now, the forward process is focused on selling the most as possible. It does not focus on preventing returns. This can be an issue for an e-commerce player and therefore the forward process must be optimized and adjusted to minimize returns.

The customer journey of buying a product on the website can already have influence on the product return rates. As earlier mentioned, customer reviews and product descriptions can impact the return rates (Minnema, 2017). Communication between return and forward departments, can result in simple adjustments in the product or user manual which lower the returns. Some marketing techniques or advertisements can boost sales, however they do not take the return rates into account. If the advertisement leads to an impulsive purchase from the customer, the customer is more likely to return the item (Zhu et al, 2016). Other studies analyzed the effect of social media behavior on product returns and there was a positive measurable effect between the tweets of a person and the product returns (Ding et al, 2016). All these analyses show the importance of selling strategy, data and customer behavior to prevent returns. In this study some influenceable and important factors in the forward process will be discussed and analyzed to see the potential applicability of ML in this field.

3.2 Linkage between Return process and Supplier selection

The agreements with suppliers are playing an important role in the return process. E-commerce players that have a wide range of products and deal with a huge number of suppliers, need to have complete agreements. Where the return rate is probably of minor importance, it can be a decisive factor in the e-commerce market. When suppliers have good and flexible return policies an e-commerce player can serve the customer better. Some products are economically dominant and you simply cannot miss them in your assortment. Take the iPhone as example, when you are selling mobile phones and you do not have the iPhone in your assortment your will not become a key player in the market. Simply because the iPhone is one of the most popular phones now. However, when Apple has terrible conditions and agreements on returns, is it still profitable for an e-commerce player to include iPhone in their assortment? Will the economic value of the product outweigh the costs in returns? Or will the costs of returns be too high and impact the business negatively? The only option then is to leave the specific product out of the assortment. There are unavoidable situations where the e-commerce players must have a certain brand or product in their assortment to compete with other parties.

This trade-off is hard to manage, good and complete agreements will be important to let your company flourish and compete against the market. For electronical products, suppliers are often collaborating with repair centers to fasten the process or repairs. When the e-commerce player has good terms on their returns, they can fully cover the costs of returns due to the fees from suppliers.

3.3 Key Performance indicators in Return Management

In this paragraph, firstly the current literature is explored to indicate the performance measurement in product returns. Second, the KPIs are presented from both literature and case study. These KPIs will be used to measure the impact and applicability of ML in the return process.

3.3.1 Performance Measurement in Literature

Extensive measurement of KPIs in the return process is only done by a few companies in the market. Measurement and monitoring of the costs and performance in returns are lacking behind, are inadequate and heavily underestimate (Lambert and Pohlen, 2001; Aramyan et al., 2007; Olugu and Wong, 2012). Due to the growing volume of product returns, especially e-commerce companies are developing a stronger focus on their return measurement. The increase of online shopping and the strict laws and regulations from EU are demanding e-commerce parties to manage their returns efficient. This creates a significant impact in their organization and it is valuable to steer on the returns. Especially in apparel industry it is a significant cost factor and they need to deal with these costs in an efficient manner. Due to this rapid evolvement, KPIs in the return process are receiving more attention from both literature and businesses. Where in the literature most KPIs measured are related to sustainability and closed loop supply chains (Beske-Janssen, 2015). The literature is more focusing on the sustainability of the supply chain and the product returns or so called “reverse logistics” is a huge impacting part. The disposal, recycling or reusing of the products is important to minimize the emission and greening the supply chain. For KPI measurement in Supply chains, many frameworks are developed to define best practice KPIs that improve the results and performance of the company. However, in the field of Return Management, there were almost no frameworks found in literature. This indicates the importance and relevancy of this research.

Bernon et al, (2011) created a framework for the management of reverse logistics, the framework is displayed in figure 16. This framework is divided into three levels. Starting at the bottom with operational performance. This relates to process management which is crucial for the reverse logistics management. A well-managed reverse logistics can not only lead to cost reduction, it can help in reducing returns and detect bottlenecks (Rogers et al, 2001). The re-sales profits and revenue is playing an important role in operational performance. Retaining high value from the returned products can help in regaining the profit over a product (Mollenkopf, 2007).

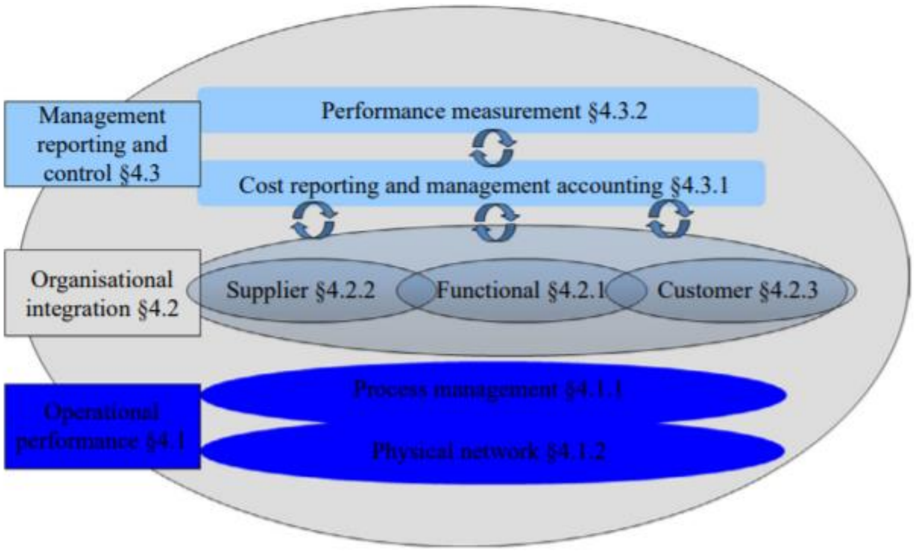


Figure 16, Reverse logistics framework (Bernon et al, 2011)

The second level is the organizational integration, the study of Bernon et al (2011) discovered the poor relationships and connections between the various players in the return field. No alignment of objectives and conflicting demands are reasons for a weak return process. Return policies with both customers and suppliers are important for the e-commerce players to have a full understanding of what they can expect from their relations. Trust and confidence is an important key issue in this part, how to have a good relationship with the supplier to boost sales. When both parties take care of their returns it can help the customer in the best way. This results in higher sales and higher customer satisfaction (Minnema, 2017).

The third level relates to Management reporting and control, in this level all costs and performances are measured. This is a critical layer that involves measurement in different levels of the organization. Applying the right metrics is important to have a full view on the costs and performances of the return process. A returned product enters various parts of the supply chain, the costs of this movement are hard to measure and companies do not know the exact costs of their returns (Bernon et al, 2011). The monitoring and control of these costs is therefore very complex but critical for e-commerce companies.

Secondly, the Key Performance Indicator framework of Shaik & Abdul-Kader (2012), displayed in figure 17, was used as basis for the indication of the KPIs. This study presents a comprehensive performance measurement framework for Reverse Logistics. Based on literature review, they defined 6 different categories for KPIs in RL. Financial, Process Internal/external, stakeholder, Innovation & Growth, Environmental and Social are the areas where the KPIs are measured. This framework is used to filter out the applicable KPIs for the Dutch e-commerce market.

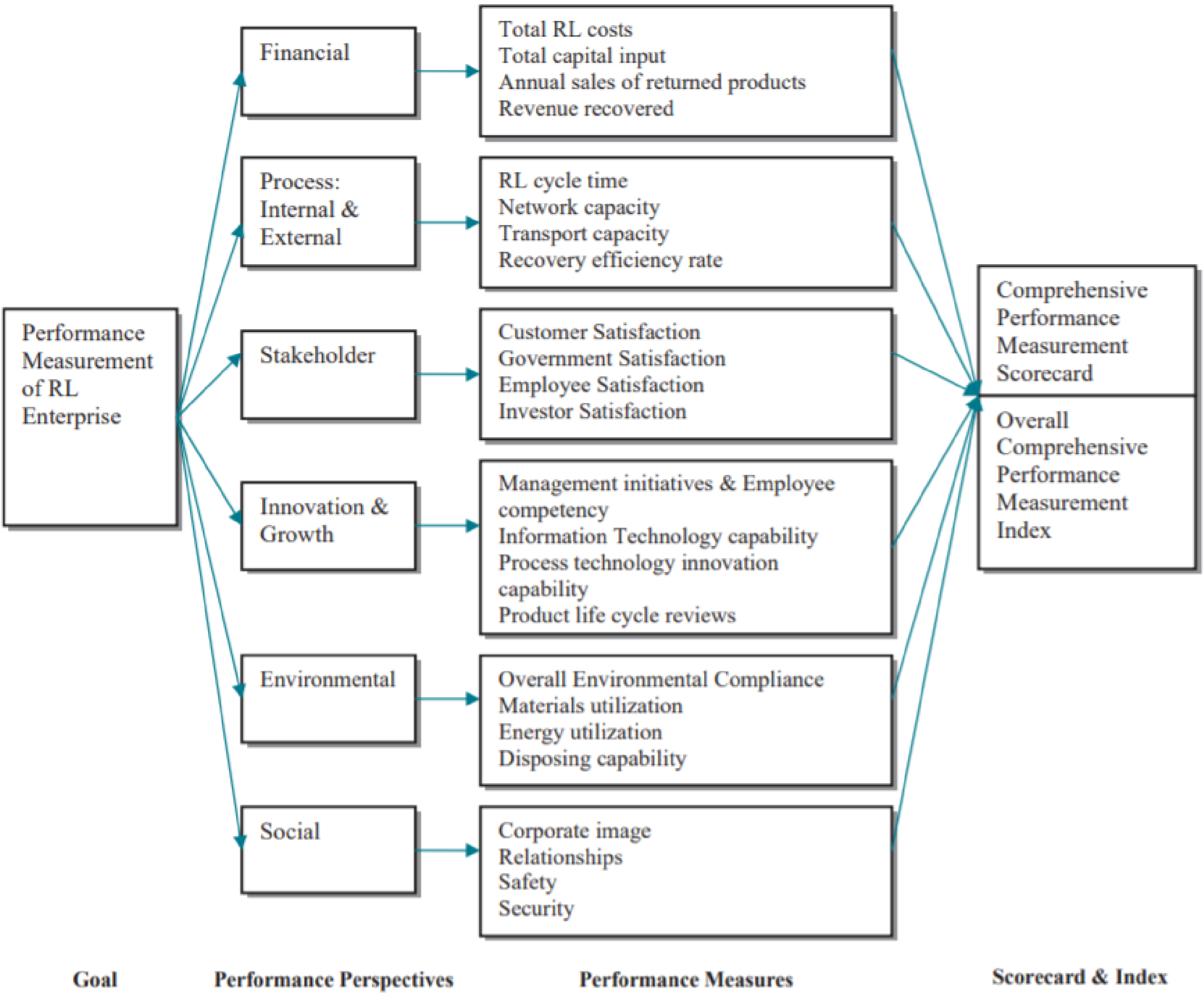


Figure 17, KPI measurement Framework (Shaik & Abdul-Kader, 2012)

3.3.2 KPIs identified from Literature and Case Study

This research compared the case study analyses and the framework of Bernon et al (2011) with the performance measurement framework of Shaik & Abdul-Kadar (2012). This resulted in four different domains for KPI measurement. These four KPIs can be placed in the domains of Financial, Process internal/external, stakeholder and social of the framework of Shaik & Abdul-Kadar (2012). The literature knowledge is included to make the generalization of the KPIs possible and prevented this research to be only applicable for the analyzed case study. The case study was conducted with an e-commerce player in the electronics market. Compared to the apparel industry they have relatively low return rates. However, it is still interesting and significant for them to invest and improve in their return process. The distribution is made into four KPIs, this creates separable performance indicators that can be measured separate from each other. With these KPIs most return related issues, costs and performances are measured and monitored.

- **Return volume**

  This includes all the processes involved to prevent returns. For instance, the product selection, what to sell and what not. Make sure the information on the website is complete and transparent. Prevent defects and make sure customers know how to handle the product. All these aspects can have influences on the return rates. Previously mentioned studies already showed the causal relations between information to the client and the product return rates. This KPI can be divided into many more concrete smaller KPIs that are measured on lower levels. However, the focus of this study will be on these four larger identified KPIs. These will show the impact on the organization and will provide enough information about the possible points of measurement. This KPI is clearly measured in the forward process, return reduction can be measured concretely by comparing different periods with each other. For instance, what is the percentage of returns on the Apple iPhone 8 in January 2018 vs February 2018? What are the differences and how can these be explained? It will be important to act upon these analyses. A new marketing strategy on the iPhone could lead to higher sales in March 2018, however the return rate could also be higher. This will ask for a new strategy to prevent the higher return rate while keeping the sales high. If a certain adjustment in customer journey, for example a more extensive product description, can result in a lower percentage of returns. This must adjustment must be implemented to reduce the return rate and achieve a better business performance.

- **Quality of the process**

  This KPI involves all customer satisfaction and customer classification of the process. What does the customer think of the process? Are they satisfied about the process? Research already told us that a customer that is happy about the return process is more likely to return to the website and purchase again (Minnema, 2017). This is an interesting point that is important in the customer journey in the return process. In case of a return by a customer, the customer is not happy about the order and is likely to enter the return process in with a negative attitude. Creating a smooth process will turn around the attitude of the customer, this can be crucial in retaining the person as a customer. Having a high customer service level in the process of return will lead to a higher customer lifetime value in the future.

  Another point of interest in this KPI is the communication with the customer. How many client contacts are there with the customer in the total process of return. High numbers of customer contact can be costly and reducing this number is essential to keep customer service costs low. Same with the duration of the process, the longer it takes for the product to be sold again, the more costs will be involved. Processing costs and transportation costs are decreasing the value of the product.

- **Costs of the total process**

  This is one of the most important KPIs measured on product level. What is the cost of a returned product? This includes the whole process, from the decision of the customer to send back the item till the product is arrived at its destination. This can be in various places; stock, trader, trash or even back to the customer. All costs of back office, transport etc. are included in this process. Retaining value from the returned products can lower the costs of returns. A returned product that directly goes to the garbage is probably less costly due to the shortage of the process. However, recovering the product and sell it again can be profitable and reduce the costs per product. Simply reducing costs is therefore not always the best strategy. Probably investing a little more into the process to retain more value out of the returned product in the end, can be more profitable for the company.

- **Supplier agreements**

  This involves all agreements with suppliers, what do we have on service agreements, how much do they reimburse from the returned goods? This can be measured on supplier level, this will show the return rates per supplier. If a certain return rate on a specific supplier is relatively high compared to other suppliers, there can be made changes in agreements with this supplier to lower the return rate. Some products are economical too valuable, therefore it is not possible to leave them out of the assortment. This can create conflicting issues between sales and returns. Managing this in a good order can be hard and challenging. However, it is an important measurement that indicates the return rates on supplier level. When several products from one supplier have high return rates compared to other products in the same range, the supplier can be the cause. This can lead to a change in supplier or it can be an incentive to better negotiate about return policies and agreements.

These different KPIs will be further analyzed in chapter 7. In this chapter the possibilities of ML will be explored to see how the KPIs can be impacted and improved. Figure 18 indicates the continuous relation between the KPIs and the performance in returns. The numbers and performance of the KPIs will influence the total performance of the return department. Based on the performance of the return department, the KPIs can be adjusted or adapted to improve performance. This relation is therefore indicated with a bidirectional arrow.

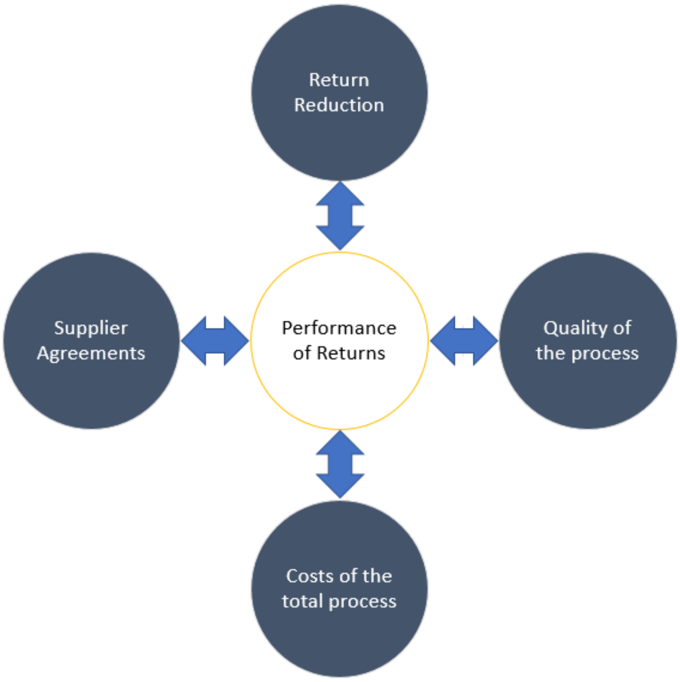


Figure 18, Relation KPIs on Performance of returns

4 Machine Learning Explained

When people talk about learning, they normally improve their knowledge to gain new insights and learn new things about science, philosophy or biology. This is the natural way of learning that is applicable to persons. However, in the field of Machine Learning it is artificial knowledge, intelligence or learning. The learning process is more related to performance instead of knowledge. The **knowledge** of a machine is hard to measure, therefore it is much easier to measure the **performance** of a machine (Witten et al, 2016). The performance can be measured by comparing the present behavior of machine with the behavior of the past. When the model improved its performance, it made progress and learned. In ML the machines can learn through datasets that contain huge amounts of data. Based on this data, ML tries to make predictions. It differentiates with traditional programming through several reasons. ML algorithms learn through time, they intend to improve their performance and do not have a strictly programmed set of instructions. ML learns without being explicitly programmed (Witten et al, 2016). The recent breakthrough of ML techniques is mainly caused by the rapidly emerging amounts of data that are being produced through social media, smartphones and cloud systems. Creating insights from this data can help society, companies and governments towards a more transparent, responsible and sustainable environment. Each ML technique has a different approach and will use different learning techniques to classify data. In this chapter the essence of Supervised Learning, Unsupervised Learning and Reinforcement Learning is explained together with the requirements that are needed to apply these ML techniques.

2. What is Supervised Learning, Unsupervised Learning and Reinforcement Learning?

a. What does this technique mean and what can it do?

b. What are the requirements and pitfalls of this technique?

This sub-question will indicate the importance and possibilities of the different ML techniques. The distinction will be made between the different techniques to make clear in what domains they are applicable and useful.

4.1 Supervised Learning – Meaning and Performance

The first word in the term itself is about supervision, observing and directing the execution of a task, project or activity. In the field of ML, the supervision is not done by a person, it is focused on a machine learning algorithm. These algorithms can classify different regions or classes within a given dataset. The optimal goal in SL is to train a model that will be able to predict future instances. This training must be done on a labeled dataset to let the model improve over time (Witten et al, 2016). SL can basically do two things, classification and regression. Where classification tries to predict a label, this can be the classification of a mushroom being poisonous or not. The model can do this based on identifiable

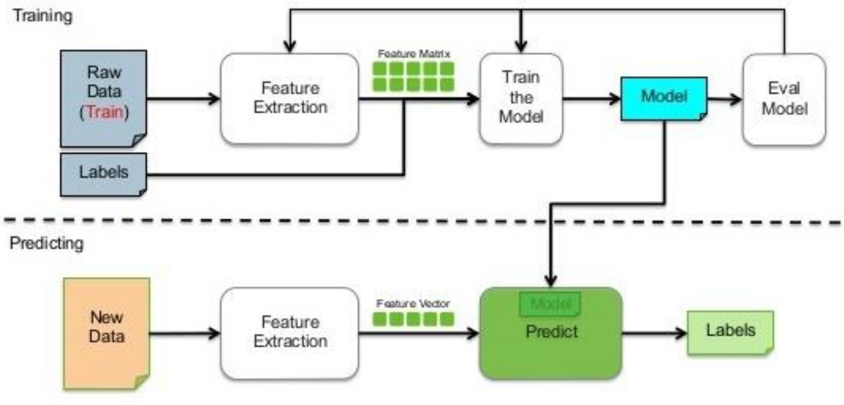


Figure 19, Supervised Learning Representation (Witten et al, 2016)

characteristics, like size, span width or color. These inputs can be given by data entries or by images. Regression tries to predict a quantity, based on given attributes, for example predicting the price of a house based on some given inputs like, room numbers, square feet and garden size. In both cases the model is trained on historical data and will be used to make a prediction on new data. For classification this means, using a data set of emails classified as ‘spam’ and ‘not spam’, let the model train on this labeled dataset and predict on a new data set if an email is spam or not. The applications of both techniques are extraordinary and due to the rapid growth of data, these techniques are being used more and more within companies. Figure 19 indicates a brief representation of a classical SL training and predicting model.

4.1.1 Supervised Learning – Requirements and Pitfalls

As mentioned before, SL is a technique that can be applied broadly. It has numerous fields of applications and can be applied through several mapping techniques like: Decision trees, logistic regression, support vector machines, and kernel machines. All these techniques have at least one thing in common, they all need labeled data to be useful. Labeled datasets provide information for the models to be trained. The exact amount of data needed is hard to specify, this differs for every technique. Nonlinear techniques require more data than linear techniques, and having more data can improve the performance of a model. Striving for a high performance can be supported through different factors. One of the most important factors is the amount of data available, having a huge amount of data can help in improving the performance of a ML algorithm (Ng, 2017).

Having clean and structured datasets are most desirable, however noisy data is often in place and can bias the learning process (Nettleton, 2010). Cleaning data can consume a lot of time, or is even not possible due to the high costs or time pressure. Therefore, SL techniques must deal with a certain level of noise and need to be resilient. Certain SL techniques are more robust to noise where others are more sensible and easier affected by noisy data. When choosing a SL technique, the performance and sensibility on noise must be considered to make the right decision and create the optimal performing model.

4.2 Unsupervised Learning – Meaning and performance

USL is a technique that unlabeled data, this is data that has not been classified. Less practical models are presented in literature and the application of USL in supply chain management seems less explored compared to SL (Hastie et al, 2009). This makes it harder to describe the method of USL while on the other hand it gives more opportunities to find new fields of application. The optimal goal in USL is about finding patterns or clusters in the unlabeled data. These clusters can give important information about customer behavior and can group people that are acting in the same way based on their data inputs. Examples of USL are clustering or density estimation and dimensionality reduction, as can be seen in figure 20. It differs most from SL due to the unlabeled data, at USL the algorithm does not know to what group or label the data belongs. This makes it harder to approve the accuracy of the algorithm. Clustering is probably the most common used method in USL. A possible methodology for clustering is done by K-means algorithm. Other USL techniques are Principal Component Analysis, autoencoder and anomaly detection (Ha & Krishnan, 2008).

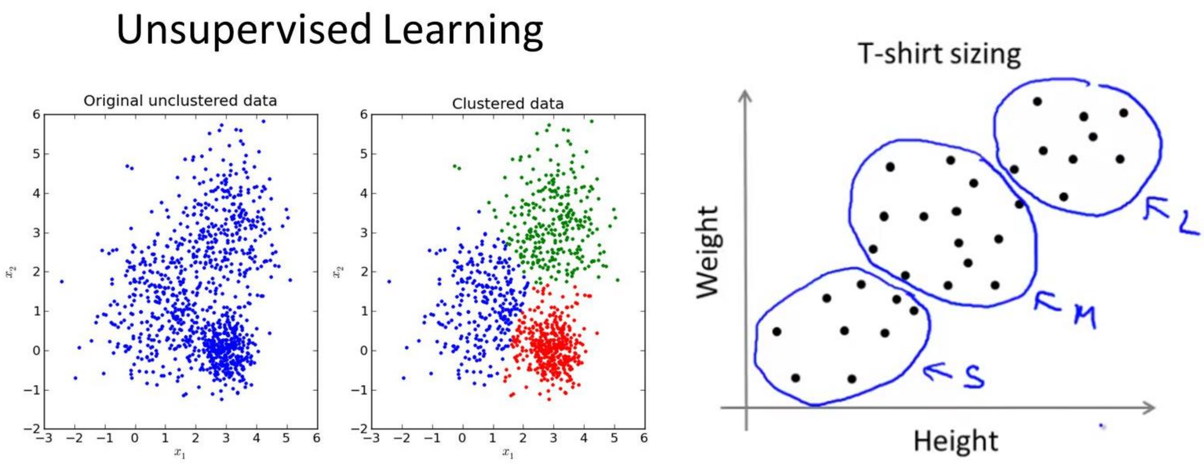


Figure 20, Unsupervised Learning example (Viswarupan, 2017)

4.2.1 Unsupervised Learning – Requirements and Pitfalls

For USL there are pitfalls as well, and it is important to have them addressed and keep them in mind while analyzing the project. Despite all the opportunities, it stays a relatively new field where many works are underexplored. The aspect of data privacy and who is the owner of the data, and who is the owner of the results from the model are interesting issues that arise regarding to ML techniques.

However, for this research that will lie out of the scope. Clustering is one of the most common used techniques in USL, it tries to detect certain groups or customer segmentation in the data. This is often described in literature, however in real-world datasets this is rarely the case (Ivan, 2018). Data can be non-spherical and this immediately makes it hard for normal k-means clustering techniques to classify the right data clusters. However, there are advanced techniques that can battle with these non-spherical datasets. Single linkage hierarchical clustering makes it possible to cluster non-spherical data(source). Another requirement that is often mentioned in USL is the same variance for all the variables. When this is not the case, k-means can have difficulty in assigning the right class to the data. Having an equal amount of data observations is another important aspect in USL, this indicates that the data that is used will be equally divided in-between your different groups. Trying to do customer segmentation, it will therefore be important to have the same number of customers in every cluster that you want to classify. When one cluster does not have a substantial amount of data observations, it can easily be added to another cluster and in this way the algorithm is missing a separate cluster. An example of unevenly sized clusters is displayed in figure 21. The clustering algorithm defines three evenly distributed number of data observation. However, this is not the right distribution and gives a wrong indication of reality.

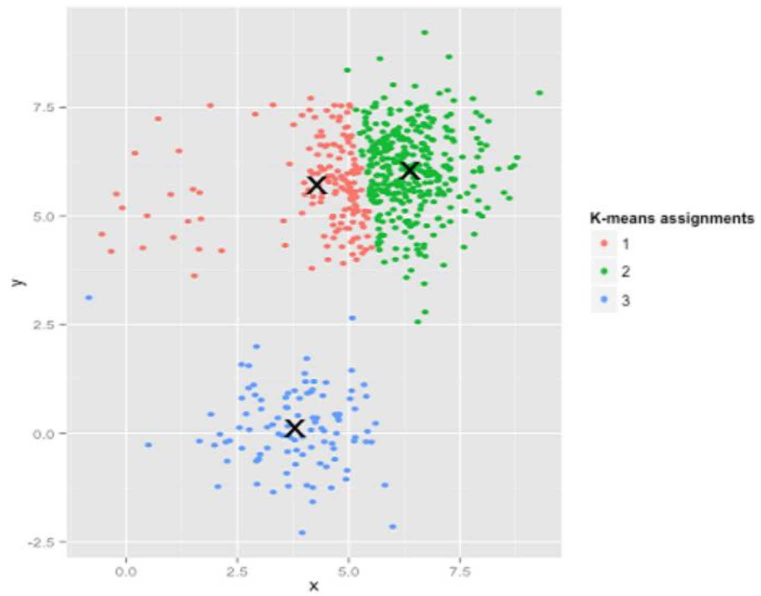


Figure 21, K-means Clustering (Viswarupan,2017)

Furthermore, a drawback with USL can be related to the unlabeled data. There are no labels available on the data, so it can be extremely difficult to tell if the results are meaningful or not. Sometimes, clusters can be found but this does not mean that they tell us anything about the data. The clusters can be wrongly identified and can be in a local minimum. On the other hand, USL makes it possible to train and learn models that are bigger and more complex in comparison to SL (Valpola, 2000). The complexity can play a role in choosing a ML technique and knowing that USL can train more complex models can be an advantage. This high level of complexity may cause delay and will ask for more computer power. When Determining the strategy for applying a USL technique, all these requirements and drawbacks must be considered.

4.3 Reinforcement Learning – Meaning and Performance

Reinforcement Learning is applied in many areas. It focusses on behavior of software agents that take decisions in dynamic environments. With RFL the hunger for data is even greater than the other ML techniques. There is still a lot of work to do to translate this technique into business practices (Ng, 2017). RFL has two basic characteristics. First it operates out of trial and error. This means that an agent acts in an environment where it is not told what action to take. It must discover what action will lead to the biggest reward. Secondly, the actions may not be affected directly on the rewards. It can have a delayed reward system where the actions affect next situations and so all the subsequent rewards (Sutton et al, 1998). RFL can improve and adapt over time in this way it can interact and deal with unstructured environments. Learning by exploration, the model does not know in advance what action to take. It will learn through his own actions and even so improve their own behavior.

4.3.1 Reinforcement Learning – Requirements and Pitfalls

Where data is playing an important role in SL and USL, for RFL it even plays a more important role. As already mentioned, the hunger for data in RFL is even bigger. This requires lots and lots of data that can enable the forefront technology of RfL. Another hard to meet requirement in RfL algorithms is the clear definition of the rewards. Rewards must be clearly defined to guide the agent towards the optimal goal. In a game environment the reward is relatively simple to identify, however in real life situations sometimes more than one reward can exist and it will be hard to classify these rewards in the right order. Rewards can be evenly classified and can have the same compensation. Optimal policies in real life situations are harder to identify and a RfL algorithm will need to have a clear vision to function optimally.

In virtual environments it can be easier to train the agents due to the easier determinable policies. However, it is still a huge challenge to train the agents and it will be even harder to train these agents in real life situations. Another challenging part will be the communication and interaction of different agents among each other. How do they communicate with each other to perceive the highest achievable goal? The interaction with the environment, other agents and even people can be very interesting and exciting to discover. However, it will be very challenging and will bring a lot of difficulties along his way.

Bottleneck can be the simulation itself. Take an example of a wind turbine where the simulation is modeled in a program, it simulates in a one or two second. To simulate this 20.000 times, you will need half a day. This will not be affected by upgrading the CPU or any other equipment. Better to start with rough simulation that are fast, then when the system is at a reasonable level of performance, try to converge and be more precise. Than when the performance is at a level of confidence, the switch can be made to the physical component. It is not that you want to train your model on the real-world situation, you only want to grind a little more. Discrete events can be simulated by various programs that can really contribute in performance of the real-life object.

5 Machine Learning in Return Management

In chapter 4, the essence of Machine Learning is explained. In this chapter, the possible application of ML in RM will be explored. Several reasons led to the usage of ML in this study and this choice will be further elaborated in this chapter. To obtain the reasons behind this choice, sub question 3 is stated as follows:

3. What are the possible applications in Return Management for Machine Learning

a. Why is this technique chosen for the given project?

b. What is the biggest potential in this technique related to RM and Supplier Selection?

Sub question 3 intends to show the different possibilities of ML. With ML there are possibilities to improve the performance on certain KPIs. To answer that question, first the possibilities of ML in the broader sense will be explored. To specify more towards the RM process, the ML techniques will be discussed one by one and the possible impact is described. First Supervised Learning will be discussed. How can SL support in improving the Return process? Secondly, Unsupervised Learning will be explored to see what this technique can mean for this study. And thirdly, the application of Reinforcement Learning will be explored and explained to see where its potential lies.

In a study of Huscroft et al (2013), the correlation between the use of information system and reverse logistics cost effectiveness and Reverse Logistics Processing effectiveness is indicated. The extensive use of Information Systems, Information Technology and Technology innovativeness improve the performance in Reverse Logistics, as shown in figure 22. The research of Huscroft et al, only indicates that technology innovativeness contributes in the performance of reverse logistics. It creates new insights and new innovative ways of working and can really benefit the return department. ML is one of these new technologies, and therefore this study tries to explore their applicability in Return management performance.

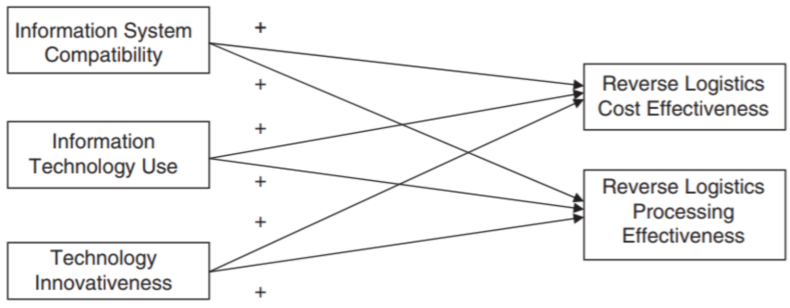


Figure 22, Correlation Information system and Reverse Logistics (Huscroft et al, 2013)

5.1 Supervised Learning - Application

There are many reasons why SL is one of the Machine Learning techniques that is chosen for this project. In this paragraph the most important incentives will be discussed and explained. In scm, a huge evolvement is ongoing in data analytics and data production, more and more data is produced due to internet of things and smart warehousing (Wang et al, 2016). Most recent successes in scm are related to SL, this technique went through a rapid growth since 2010 and became one of the most applied techniques. Forecasting demand and inventory levels, predictive maintenance and smart logistics are being applied in a broad scene. This recent application is caused by different reasons. First, through the rapid growth of data production, a lot of data is available about the customers, suppliers and sometimes even the machines. Analyzing this data can help organizations in being more productive, it can help in customer engagement and in many more domains. Therefore, the data growth is the most important reason to explore the possibilities of ML in this project.

Having huge amounts of data does not directly imply to use ML. Other techniques like data analytics or business intelligence can also be explored. However, ML can have an added value and can be more useful in the RM and supplier selection processes. Data analytics is probably more than 25 years old, visualize the queries of your data to create dashboards that describe behavior from the past. In this way certain trends from the past can be explained, companies can learn from this data and improve their business strategies (Gandomi & Haider, 2015). What ML tries to do, is figuring out how, from these large datasets, predictions can be made on a new data that can support future decisions. Doing this with high accuracy and confidence it can bring huge business value (Gomez-Uribe & Hunt, 2016). The classification of different customers can be an example of SL in SCM. Some customers can be too expensive to have. If a customer returns a high percentage of their products, and the company invests a lot of money in personal advertisement. This can lead to an unprofitable client and therefore strategy must be changed to serve this customer.

The classification of input data can be of importance in the process of RM and supplier selection. Classifying an object from an image can be of help in the process of RM. In the process of returns a classification model can be helpful to smoothen and fasten the process. When a product returns in the warehouse, a SL algorithm can classify a product through image recognition. This can lead to a certain rack in the warehouse where it must be transported. It can even go a step deeper when the connection is made with inventory level, in this case the inventory level is up to date directly and the company can sell the product faster. However, dealing with damaged returns, the SL algorithm could also be classifying if the product is damaged. In this case there must be a link between the customer service and the returns. The reason behind the return should be known before the product enters the warehouse. This indicates the importance of the communication between the different departments. These are all helpful and they contribute in making the SC more efficient. Improvement of the KPIs will directly indicate the performance progression of the processes.

5.1.1 Supervised Learning - Potential in relation to Return Management & Supplier Selection

In the Return Process of companies there can be a lot of KPIs in place to measure the performance of the return flows. An example of a KPI can be the number of returns, measured monthly. This number can be hard to predict especially when there is little information available. Based on customer data, product data and sales data predictions can be made about the returns. In this way business will have better insights in their costs of returns. Products are being returned due to several reasons. Sometimes because of the damage on a product or sometimes due to wrong deliveries. In case of a wrong delivery, the product itself is still intact and main priority for the business is to have this product in stock again, as soon as possible. Managing the process of return delay distribution can save costs and boost sales. The faster the product is returned in the inventory, the faster it can be sold and shipped again (Toktay et al, 2003). When a regular customer, places an order of 5 trousers every three months, fits the trousers at home and returns four of the five trousers again because the customer only needs one trouser. Seeing this happening on regular basis, an algorithm can make predictions for this client and detect when it is going to happen again. In this case, personal advertisement can be in place to prevent the returns from happening. An example could be to give the customer an extra 10% discount when he does not return the ordered goods. Or suggest a better size table to give more precise advice to the customer.

Previous described examples are all in the forward process. This includes the process of buying and tries to prevent the returns from happening. SL can also contribute in the return process itself and can create an impact on different KPIs. The total time of the product processing from customer back to the state of origin. Make sure the goods are stored back in the right place and inventory is updated, this can improve revenue significantly. The different possibilities of SL in this process will need to be explored to accomplish KPI improvement.

Quoting Li & Olorunniwo (2008) "The ability to collaborate with various players in the reverse chain is as important as in the forward supply chain. In fact, what makes a forward supply chain successful is the visibility of products in motion as well as collaboration and trust amongst the various entities in the chain. This is also true for the reverse chain, especially since the RL process is also heavily demand driven – that is, the downstream customers make the final decision in orders and returns." This citation is matching perfectly with the inclusion of Supplier Selection in this project. The importance of communication and collaboration between the different parties in the SC is huge. Especially in the return process where there can be a lot more complexity and uncertainty.

In this process of supplier selection, there are many influenceable variables involved that makes it a complex decision-making process. Models have been developed to forecast the optimal relationship with multiple suppliers. ML algorithms can deal with multiple parameters that are set to grade the performance of suppliers. Making the link towards RM, there are products that are being received as damaged through the customers. There could be scenarios that this product was already damaged when it came from the supplier. There is a need for connecting this data and see how this can be analyzed and future damaged products can be prevented. This can have influence on the relation with the supplier.

5.2 Unsupervised Learning - Application

To create a full image of the main three ML techniques in the field, USL is included as second ML technique that will be explored for this project. Due to the different possibilities of USL, it can be helpful in improving the KPIs within this project. Through the clustering capabilities of USL, new insights can be created what will ask for new identified KPIs. New KPIs will give a more complementary view of the business performance. This creates an added value for the company and these KPIs can also be applied in the broader scene. Not only will they be applicable in this company, these KPIs can be measured market wide and therefore contribute in many different companies that are dealing with the same complex issues.

The return process of a company will be analyzed in this research. However, the communication and collaboration between the sending and returning departments is crucial to improve the way of working. USL can be less effective and applicable in the return process, however it can be of more contribution towards KPI improvement in the sending process. Change a certain process or handling in the process of sending can prevent certain issues from happening in the return flows. USL can therefore be more applicable when certain information from both flows are included and analyzed.

5.2.1 Unsupervised Learning - Potential in relation to Return Management & Supplier Selection

USL is a technique that can find structure in unlabeled data. The problems USL can solve are most of the time impossible for humans. Clustering can be used to compare the input data with before defined data clusters. Another application of USL in the return process can be anomaly detection, when the input data does not match the expected pattern, a further look in to the data is needed and it can lead to certain conclusions and insights in the process. When system failure occurs and a product is in the process for a very long lifespan this can raise notifications towards the responsible person so that they can react upon these matters in a suitable and efficient way.

Within Return Management USL can have much impact on the performance of the KPIs. Different processes within the operational and tactical level can be affected. The analysis of the processes will point out issues, possibilities and challenges that are open for the implementation of ML techniques. These issues will be tested to see if they can fulfill the requirements that are needed for the different ML techniques. One of the examples can be, clustering on customer level, detecting which customer groups are being evolving and what the return rates are per group. This can create new insights and can shape new KPIs that will indicate the performance of the company even better.

5.3 Reinforcement Learning - Application

This technique is chosen due to the several reasons. First, RfL gives insights and possibilities that are not reachable for SL and USL. RfL acts in a dynamic environment and adapts on continuously changing factors. It is interacting and reacting with the real world and this makes it, especially in manufacturing companies, very attractive to apply. Due to the continuous changes in the manufacture process of a company RfL can help in making decisions on human level, without the intervention of humans. Not only in Manufacturing companies, in the whole supply chain RfL can play an impactful role. The famous beer game, where a manufacturer, distributor, wholesaler and retailer are reenacted and the end goal of the game is to lower the cost as much as possible while there are incoming orders of different size every week. This game was played using a RfL algorithm and it showed very decent results (Oroojlooyjadid, 2017). The game is originally invented to show the bullwhip effect, however the RfL algorithm, after 45000 runs, reduced the bullwhip effect significantly. This shows that RfL can be very helpful in determining stock levels, reducing costs and predicting demands. Multi-layer collaboration and communication within the supply chain can be beneficial for all and will help the overall supply chain in performing the optimal. Therefore, RfL can be even of more use when there is collaboration between the different levels in the supply chain. In this research the Suppliers will be analyzed to see what kind of link or collaboration is possible between the supplier and retailer. RfL turns out to be beneficial in many ways, and it is still a technique that is only implemented in very few real-world examples. It will be hard for RfL to be implemented and used, this will ask for a lot of knowledge and investment costs. However, exploring the different possibilities can contribute in research for the future.

5.3.1 Reinforcement Learning - Potential in relation to Return Management & Supplier Selection

RfL is applied in many cases within the field of robotics. Robot arms in manufacturing automotive factories that reduce the fault rate, the factories of Tesla automotive are an example of these giant robotic arms (Tesla, 2017). Other activities in inventory management like, grasp and place items from racks into boxes. In the SC domain RfL can improve and bring a lot of new opportunities to work more efficient and reliable. In the process of return management there are a lot of activities that can be replaced by robotics. Certain decision processes can be supported by RfL algorithms. Streamlining and automating the reverse flows and striving for zero human intervention can be a goal to obtain cost and fault reduction. Based on the KPIs from the different analyzed processes the suitable applications of RfL in this project will be further explored.

6 Impact of Machine Learning on Return process

Most companies operating in the e-commerce market are gathering and using huge amounts of data. They analyze browsing behavior of customers to suggest products that they are likely to buy (Veitch, 2016). Data gathered by Facebook can be used to personalize the advertisements and can help the websites in approaching a specific customer with a targeted product (Sennaar, 2018). Where customers have zero knowledge about the usage of their own data, e-commerce parties are making use of it. The upcoming GDPR regulations will give the customers more control over their own data. Still, the e-commerce players will use the data on great scale to personalize search results and advertisements. Shaping a personalized website for every specific customer is the goal of many web shops to create the best customer experience (Zalando, 2017). This can only be done by gathering huge amounts of data. The e-commerce market becomes more and more data driven, this enables multiple ways of applying data analysis and ML techniques. The customer data can be valuable for the return process. The more data is available from the customer, better predictions can be made. In this study there is no data analysis and ML algorithm performed. Companies were not willing to give up their data and secondly, due to the different focus and lack of time in this study no implementation is performed. However, through the interviews conducted and the exploration of current literature, the opportunities of ML in return management are explored. The following sub question will be answered in this paragraph:

Sub-Question 4: What are the possible applications of Supervised Learning, Unsupervised Learning or/and Reinforcement Learning, on the bottlenecks in the Return process of an e-commerce company?

Sub question 4 tries to go a step deeper into the application of ML in the field of product returns. Based on the bottlenecks discovered in paragraph 3.1.2, the possible applications of ML will be explored. The bottlenecks are both from literature and case study and are divided into the forward and return process. The forward process can have a huge impact on the return volume and by making the right decisions in the forward process, returns can be prevented. E.g. complete product descriptions and meaningful customer reviews (Minnema, 2017). In this chapter both forward as return processes will be described separately to give the best overview of the applicable ML techniques. First the bottlenecks and application of ML in the Return process will be described. Secondly, the bottlenecks in the forward process will be presented. In chapter 7, a further specification on KPI level will be made to show the impact of ML as tangible and measurable performance.

Table 2 indicates the four different previously identified bottlenecks. These four bottlenecks will be analyzed to explore the possible applications of ML. Due to the problems that are arising with the bottlenecks, they demand for improvements. Therefore, the ML applicability is explored on these bottlenecks to see how these bottlenecks can be improved. Machine Learning is not just randomly applied to see what kind of effect it can bring. It is specifically focused on the concerns and bottlenecks around product returns. This gives opportunities for improvements in performance, efficiency and costs reduction.

Bottlenecks	Content
Customer contact	Contact moments of customers with the customer service in the return process
Cost and time in return process	Find the right balance between speed and control in the return process.
Agreements with suppliers	Suppliers need to be involved to manage returns efficiently and resilient through the whole SC.
Forward process	Prevention and reduction of returns must be created in the forward process of the SC.

Table 2, Bottlenecks in Forward and Return process.

6.1 Return Process

In this paragraph the return process will be further analyzed and the impact of SL, USL and RfL will be explored. This process starts after the customer decided to return the product. The first two bottlenecks will be discussed in this paragraph. 'Customer contact' and 'Cost and time in return process' will be described in 6.1.1 and 6.1.2. These are both part of the bottlenecks in the return process. The suggested improvements and applications of ML are based on current literature.

6.1.1 Customer contact in return process

In the past several years, customer demands grew intensively. Customers are high demanding in terms of satisfaction, speed and knowledge. This means that customers are demanding more information about their shipments, faster refunds and higher customer satisfaction. In the return process this seems to work the same way. Customers are continuously asking questions about their status of return. Online customer services have their hands full on questions regarding customer returns. The case study example approved these contact moments and described it at as 'labor intensive' and 'costly'. Customers are frequently asking questions like; What is the location of my package? Or when can I expect my return? The contact with the customer is of huge impact on the customer satisfaction level. Previous studies approved that satisfied customers in the return process are more likely to order again at the same web shop (Minnema, 2017). Trying to lower the labor intensity in this process while keeping the customer satisfaction high, will be challenging.

In the current return process, customers often call to the customer service to request information about their return. Customer data can be used to reduce these contact moments. When a customer purchased a product and requests for a return, regardless the reason behind the return. The customer will go to the website and search for a telephone number or search for the return page where the customer can print a ticket. When looking for these return related questions and the customer is logged in, the website can directly provide clarifying information about the previously purchased products. It can first ask for which product the customer is looking to be returned. Than ask what the reason is for the return. Depending on the reason of return, the website, can provide useful information towards the customer. If the customer indicates that it is not satisfied with the product, the website can suggest other products in the same product domain.

Another useful application of Machine Learning is the ability to provide useful information based on the input data of the customer. When the customer indicates that the product is damaged or not working, provide a short instruction video of how the product needs to be installed and set up. When the customer indicates that the product was damaged on arrival, the algorithm can try to guide the client towards a physical shop, in this case the customer can directly receive a new product. These will help the customers of the e-commerce players in the return process and will lower the contact moments with customer service. From the input of the interview with the case study, more than 30% of the employees of the company are dealing with returns on daily basis. Lowering this percentage could decrease operating and processing costs for the e-commerce player.

Another beneficial scenario for a ML technique can be providing continuous update to a specific group of targeted customers. This is a group of customers that previously returned products and reached out to the customer services in this process. Studies proved that customers that have returned in the past are more likely to return products in the future, compared to non-frequent returns (Petersen & Kumar, 2009; Minnema et al, 2016). As called by Shah et al (2012) these 'habitual returners' are far less profitable than other customers. In the long term these buyers can have seriously negative effects on the selling organizations (Shah et al, 2014). By using SL and USL, these groups of customers can be clustered and classified, when they return a new purchased product, the algorithm can detect these customers and they will continuously receive updates about their return. In this way, the updates can reduce customer contacts of the customer service. These updates can be sent to the customer by email or phone and they are informed about the status of their return. It can also be linked with the refund process, this means when the refund has taken place, the customer receives a message and knows that the refund has taken place. This technique is only necessary in specific groups, customers that are very impatient will probably be more targeted by this technique. These customers have higher demands than other customers, this makes them vulnerable and more sensitive to interact with.

The contact with customer service is crucial for a company and can have impact on the customer satisfaction. Implementing new technologies in this field can help in serving the customers. It can decrease the labor intensiveness at the customer service due to automatization and standardization. These customer service employees can now focus on other important tasks, where human decisions and judgements are needed. For the analyzed company of the case study, this can be a major improvement and reduction on labor intensity. However, they have to deal with an important trade-off, retain customer satisfaction high while lowering costs and labor intensity in the return process. ML can help in this process, it will lower costs and labor intensity due to automation, classification and customer clustering. This will enable the employees to focus more on customer satisfaction and be more customer orientated.

Using new technologies can also create negative side effects. Probably certain groups of customers are not yet ready for the change in customer service desks. People that prefer to speak with a person instead of automated suggestions. However, another group of customers probably prefers to deal with the interactive ML tool to receive instant responses. An e-commerce player can try this out and only aim for clients under the age of 30. These customers are probably more likely to use the website for returns and can have a more open attitude towards automation. If the test is a success and customer service contact moments decrease, this technique can be rolled out over a bigger group of customers.

Concluding this paragraph, this study tries to make clear that not only the positive effects are considered. Both advantages and dis-advantages are given and this is important when it comes to implementations for such techniques. ML algorithms are not humans, they cannot sense the emotions of a person. They will respond in the same way to every customer, where employees can interact with a customer and can detect and respond to the mood of the customer. A person can calm down a customer and help them with further elaborating questions. While ML algorithms are strictly bounded to their data and they cannot expand on other concerning issues of the customer. Nevertheless, applications of ML can help in the improvement of the return process at e-commerce players. They can contribute in responsiveness to customers and can decrease the labor intensity of the customer service.

6.1.2 Costs and time in the return process.

'Time is money' is a famous old saying in The Netherlands. The faster you take care of things, the cheaper it is. For the return process this seems a valid saying. Having a short processing time, have the returned items back in stock fast and refund the customer quickly will benefit all parties involved. However, it is not as simple as it seems. Being fast in the process can have negative influences on other parts in the returns. It can influence the accuracy and reliability of working. Being in control is important to guarantee the customer service and have full knowledge about the returns. In this paragraph, the bottleneck concerning the costs and time in the return process will be analyzed to explore the potential applicability of ML. Based on this identified bottleneck, two different processes are described. First, the incoming returned products in the warehouse. Possible applications of ML will be explored in this part of the return process. Second, the total time and process of returns will be analyzed. The output of the case study indicated some differences in the return process, relating to days of the week or product ranges. Is ML able to lower the differentiation in the process of returns?

1st Entrance of the goods in the warehouse:

The analyzed case study operates with one single warehouse where all returning products arrive from the customer. To tackle the labor intensity of the return process in the warehouse, a robot or machine based on ML technique can be implemented to perform simple human tasks. A machine can first unbox the product and then will process and classify the product. It will be using image recognition to classify the product and directly send it towards the next destination in the process. By scanning the product, the machine uses Supervised Learning techniques to link the input to a pre-defined output. When there is a match, the machine can come with an answer. It can also scan the return order that was on the box and match the return information with the package. If this information matches, the algorithm decides what the next step is. For example, the product is returned due to damage on arrival. The algorithm classifies the product as damaged, it could send it directly through towards the repair center without any human interaction. This can decrease the labor intensity in the return process while the customer does not notice anything. Optimal sorting is an example of this technology. In the union industry, optical sorting is done to reduce labor intensiveness. The machine can process 24 tons per hour and only needs 5 people to manage the process. Previous process without the classification machine, needed 16 people for the same amount of unions per hour. The machine enabled a decrease of 11 people per hour in labor intensity. The sorting machine can be differently adjusted, dependent on factors like; quality, roughness and size of the unions (JDC, 2018). It can assure a higher quality and higher supply reliability, while labor intensiveness decreased.

This is a perfect example of the application of ML in a sorting process of a company. Applying such a machine in the return process will have no influence on customer satisfaction, and the machine can easily be linked to other automatization processes. It can for example, automatically update the status of the product, this can create a message to update the customer about their return. The sorting machine in the union industry can be taken as an example to further explore the possible applications of ML in the return process.

2nd The totality of the return process

The second application of ML is more focused on totality of the return process. Where Griffis et al, (2012) already proved that a higher refund speed will improve the relationship with the client. The total spending can increase by 158% - 457% if the e-commerce players offer returns for free (Bower & Maxham, 2012). Reducing the processing time of the total return process can cut costs and reduce complexity. However, this is hard to manage while both speed as control needs to be secured. Data analyses can filter out outliers that are influencing the return process. For instance, all packages returned on Friday have a longer processing time than packages arrived on Tuesdays. This can be through simple reasons due to the weekend period. However, other underlying reasons can be detected that can suggests new strategies or changes in the process.

Striving for a faster return process can be beneficial for both the selling party as the customers. It will lead to a shorter refund time for customers and the selling organization has their products faster in stock again. To accomplish a faster return process there can be removed intermediate steps in the process. This can already start at the customer, if they received a damaged product, they can directly send it back to the original seller or to the repair center, instead of first sending it towards the selling party. This cuts out a step in the return process, however it will demand intense cooperation between the parties to keep serving the customers at a high level. Different information systems must be linked to keep the data controlled and manageable. ML techniques can be used to analyze the data of these processes and propose the best strategy for future returns. The USL techniques can cluster based on product range, and can offer different strategies per product. This means, products from different ranges and types must be handled differently. Perhaps, cheap products can be handled differently from expensive products. Give priority to more expensive products, refund money fast, this will create higher customer satisfaction. Or it can identify differentiations in the process of returns between the weekdays and propose a deviant strategy based on the day of the week. In this way they ensure the quality and time of the process of returns and prevent longer processing times. Products that are returned on Fridays will need other handling than products returned on Wednesdays.

6.2 Forward Process

As this study showed, the return process is not a stand-alone department that can completely be isolated from the rest of the organization. Striving for high performance in returns can be influenced by many other departments and factors. Communication between the return employees and other employees in the company can contribute in improving the returns. This paragraph is based on the two bottlenecks of 'Forward Process' and 'Supplier Selection' both identified in the case study analysis. The focus on preventing the returns is important to lower the costs and increasing the profit margin. Three different possible improvements in the Forward process will be discussed to see what the impact of ML can be. First, the introduction of a 'keep reward' strategy will be introduced. Can returns be prevented by rewarding customers, only when they do not return the purchased product. Second, how can agreements with suppliers contribute in the improvement of the returns. And third, the clustering of customer groups and behavior to prevent returns. Several studies proved that the return rates are strongly influenceable by adjustments in the purchase process of the customer (Minnema, 2017; Bower and Maxham, 2012; Petersen and Kumar, 2009). These adjustments must be further explored to see the potential effect on product returns.

6.2.1 Reward and keep

Preventing returns can be done in several ways. Provide incentives for the customer to keep the product is important. This keep reward system is research by Gelbrich et al (2017), they found a positive effect between rewarding system and the keep intention of the customer. When customers receive rewards for keeping the product, they are less likely to return the product. Especially with frequently buying customers this can be an effective strategy, these customers will come back multiple times and rewarding them will have a bigger impact on sales and profit margins.

This rewarding method must be specified on customer and product type. Certain products have very low return rates, therefore it is not necessary to implement the reward strategy. Also, certain customers never return products because they know what to buy. These customers do not need any personal reward system to prevent returns. Customer clustering can be used to group the frequent returns and reward them if they do not return a product. This can encourage them to keep a product and do not return products in the future. The personalization and specification can benefit the company, the algorithm only approach the customers that are likely to return their product.

6.2.2 Supplier Relationship Agreements

This study tried to link the Return process with the supplier to build a strong relation and serve the customer as it best. Generous refund policies from the suppliers can incentivize the e-commerce player to offer complete and full refunds to the customers. This will boost sales and will be beneficial for the whole supply chain (Minnema, 2017). The contracts with suppliers are crucial and the selection of suppliers is a complex and important process. There have been studies that tried to do fuzzy approaches on multi criteria selection on suppliers (Kahraman et al, 2003; Kuo et al, 2010). Scoring the suppliers on as many relevant criteria as possible can improve the relations and give more certainty for the e-commerce players. In the latest literature the focus shifted to more sustainable and environmental friendly criteria, these aspects were not considered in the past when it was about supplier relationship. However, for the e-commerce players the inclusion of Return agreements and policies will be more and more important. By having well covered agreements with the suppliers about returning products they can have better guarantee and refund policies towards their customers.

ML can be used through using historical data of a product sales. When there will be a new product launch of a brand, historical data can be used to predict the returns for this launch. Take for example the apple iPhone, they launch a new iPhone every several years. All data of returns of the previously launched iPhones can be used to make a precise prediction about the numbers of return for the new upcoming iPhone. When the e-commerce player knows this information, it can use it in their negotiation on the returns. They can make a cost based analyses on their previous iPhones and trough the predictions of the algorithm they can be very precise on the next iPhone. By using these prediction, they can cover all costs of their returns and can provide high service level to the customers. With the use of the previous data they can easily see what the main reasons were for the returns. If the main reason is a damaged product, they can use this in the negotiating with their suppliers and can demand lower damaged rates. Following the output of the case study, this point can be crucial for the profitability of a product. When the agreements with the suppliers does cover all costs related to returns, the e-commerce players can help and serve the customer fully. When there is no budget to handle the returns, probably the e-commerce player will try to handle the returns as cheap as possible to avoid big losses.

6.2.3 Customer groups and clustering

One of the strongest possibilities of ML lies in its predictive ability. ML algorithms can make precise estimation of future behavior based on historical data. The detection of certain specific groups can benefit an e-commerce player. For instance, studies showed that first time buyers have higher uncertainty and therefore return more often compared to frequent buyers (Minnema, 2017). This can variate between the products. The return sensitive people could be more supported in the purchase process and can be helped in several manners to prevent the return. The browsing behavior of customers do also have effect on return behavior. This is researched by Minnema et al (2017) and displayed in figure 23.

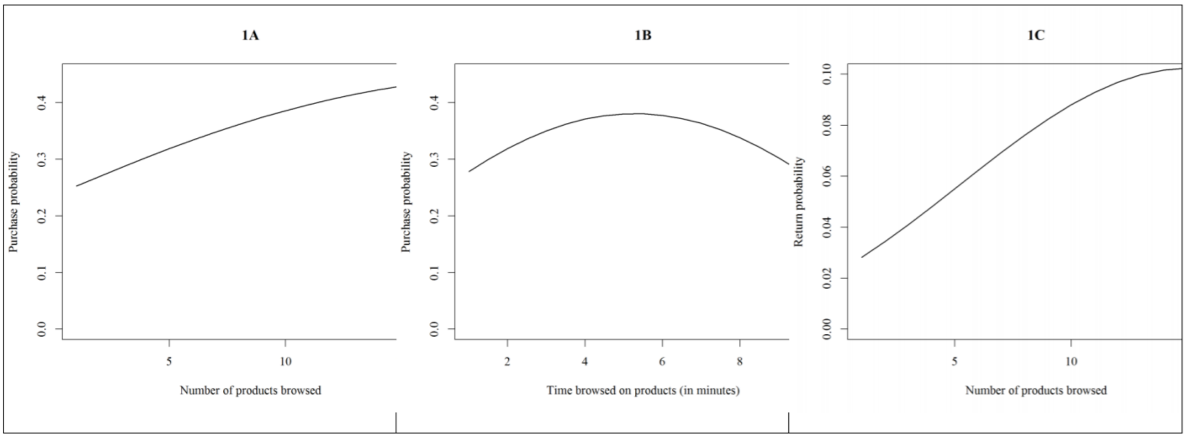


Figure 23, Influence of browsing behavior on Return and purchase behavior (Minnema et al, 2017)

Graph 1C shows the return probability on the Y-axes and the number of products browsed on the X-axes. The more products browsed by a customer, the higher the return rate. Based on this insight an e-commerce player can target these group of customers. Detecting this behavior is relatively simple, however reacting immediately to lower the return and help the customer in purchasing the product is much more difficult. The algorithm must detect this behavior and provide the right information. This can be in various ways, by providing short reviews of the browsed products or video clips that explain the different products. All these handlings will try to support the customer in making the right decision what will eventually lower the return rates.

Another application of USL is the detection of outliers, anomaly detection. This can be used to detect fraudulent behavior among customers. Customers that order cheap products because they receive a certain gift additional on their order, are keeping the gift and returning the product. In this way they try to receive free products in a fraudulent way. Or they order cheap products in the hope they do not have to return the product. E-commerce players have often more costs on the whole return process and decided to let the customer keep the product. This is behavior can be detected as outliers if it is compared with normal return behavior. These customers are causing problems at the e-commerce players and can be tackled by using ML techniques. Now this behavior is detected after an x amount of returns, probably an amount above 30 before anyone notice. However, ML techniques can identify this behavior earlier and prevent future fraudulent returns (Appriss, 2017).

As noticed in this paragraph, there are several possible ways to lower the return rate of customers with the help of ML. And unlike Amazon, this study does not want to ban customers but tries to suggest other possibilities to interact with customers. Customer bans can be prevented through good agreements with suppliers, interactive contact with the customer and extensive data usage to create a future proof strategy on returns.

7 Performance in Return Process Improved

The continuous improvement on performance is necessarily for the e-commerce businesses to keep up with their competitors. Key Performance indicators are receiving more attention to keep track of their business performance. These KPIs can give a representable image of the behavior of a company, acting upon the KPIs can help a company in detecting problems and react upon them. When the performance on a certain KPI is low, there can be changes suggested in strategy or operations to improve performance. Examples of important KPIs in the e-commerce markets are; net promotor score, profit, conversion rate, gross margin, website tracking and customer lifetime value.

Both in literature and business there is no extensive knowledge that specifies on KPI measurement in the return process. The complexity in the returns is related to the vertical integration, multiple levels of the supply chain are involved in the process of returns. The return process is not limited to one level in the supply chain, returns overarch more levels and stages in the SC. To establish an efficient supply chain, sharing of information, resources and costs are required between the different stakeholders (Huang et al, 2010). These multi-layer relationships make it hard to measure the performance and steer upon the KPIs. Based on extensive literature review and practical examples from the case study, chapter 3 showed the identified KPIs in the return process. These KPIs will be used as a representation for companies participating in the electronics e-commerce market. This chapter will indicate the impact and applicability of ML on the KPIs in the return process. The following sub-question was created;

What can be the applicability of Supervised Learning, Unsupervised Learning or Reinforcement Learning on the defined Key Performance Indicators of the Return Process?

a. What is the applicability of the different ML techniques on the KPIs in the return process?

b. What can be the potential impact of the different ML techniques on the performance in the return department of an e-commerce player?

This chapter will explore the applicability of ML in the field of product returns. The impact on KPI level will be presented to see the tangible output and potential gains on business performance in the e-commerce market.

7.1 Applicability of Machine Learning on Key Performance Indicators

In this paragraph, the identified KPIs from paragraph 3.3 will be evaluated to indicate the applicability of Supervised Learning, Unsupervised Learning and Reinforcement Learning. Table 3, presents a detailed overview of the 4 main KPIs, the matrix indicates the applicability of the ML techniques in Return Management. To indicate the applicability of the ML techniques on the KPIs, a scoring range is defined from strong negative to strong positive applicability. Through the analysis of the return process and corresponding KPIs, the applicability of the ML techniques can be weighted within this range. The results are presented in table 3, the table shows the degree of applicability of ML on the KPIs in the return process. For instance, Looking at KPI 2, SL and USL can be strongly applicably to decrease the customer contact moments in the return process. This indicates a strong positive applicability of SL and USL on one of the bottlenecks discovered in the case study. It can eventually lead to a decrease in costs and increase in responsiveness of the customer service. The basis of these results lies in the analysis of the case study, where bottlenecks in the return process were explored. To measure improvements on these bottlenecks, KPIs were identified that can measure the performance of the return process of an e-commerce player. Now, the ML techniques are linked to the KPIs of the case study to provide tangible insights that can help in applying ML to improve the performance in the return process.

Range	Degree of Applicability
++	Strong Positive Applicability
+	Positive Applicability
=	No Impact
-	Negative Applicability
--	Strong negative Applicability

Table 3, Legend

The two plusses at top in table 2, indicate a strong positive applicability of one of the described ML techniques. The two minuses at the bottom of table 4, indicate a strong negative applicability. This can indicate a negative influence of the ML technique on the KPIs, this can arise when the Machine Learning technique is not performing well due to the lack of data. The difference between 'Strong Positive Applicability' or 'Positive Applicability' will mainly lay in the data volume that is available for the Machine Learning algorithms. The more data is available, the better the ML algorithms will perform (Ng, 2018). The amount of data will enable the ML algorithm to perform well and create a visible impact on the business performance. When there is less data available, the applicability will be less. The effectiveness of the Machine Learning techniques will be coherent with many other requirements. There are people needed with the right knowledge and a company needs to have the right data architecture in place to take advantage of ML. However, this study is not focused on an implementation or execution of a ML algorithm. It provides insights and new knowledge on the applicability and the possible impact ML can have on business process in the complex dynamic environment of product returns. While interpreting these results, it must be considered that there is very limited research and business practices available in this field. This makes it very innovative and attractable for further research, however it also brings some uncertainty, this will be further explained in the discussion.

KPI 1: Return Volume	SL	USL	RfL	Possible Impact
- **Fraudulent Returns**	=	+/++	=	Minimize fraudulent returns.
- **Number of returns**	+	++	=	Lower total numbers of returns. Specified per customer and product
- **Labor intensity**	=	++	=	Less workers needed in return Department and customer service.
KPI 2: Process Quality	**SL**	**USL**	**RfL**	**Possible Impact**
- **Customer Satisfaction level in Return process**	++/--	++/--	=	Customer responsiveness will go up. Customer satisfaction can differ in certain customer groups. Can have positive and/or negative influence.
- **Come back rate of customer**	++	+	=	Will grow (Meaning: Customers will come back to same web shop)
- **Contact Moments with customer in return process**	++	++	=	Decrease of contact moments with human employee.
- **Labor intensity**	++	++	=	Decrease of manpower (involved in return process)
KPI 3: Process Costs	**SL**	**USL**	**RfL**	**Possible Impact**
- **Costs of Return**	++	=	+	Cost can decrease per returned product.
- **Margin of returned product**	++	=	=	Margin is retained in more cases of returned products.
- **Indirect Costs**	+	=	=	Indirect cost can be reduced, back office is less involved.
- **Total time of return**	+	+	=	Process can be accelerated through automation.
KPI 4: Supplier Agreements	**SL**	**USL**	**RfL**	**Possible Impact**
- **Refund Policy**	+	=	=	Better covering refunds from suppliers-side. Can result give higher customer service to end customer
- **Quality of Contract**	+	+	=	Contract can be more complete in order to serve complete supply chain better.

Table 4, Applicability of ML techniques on KPIs in return process

7.1.1 KPI 1: Return Volume

The first KPI, Return Volume, is influenced by many underlying factors and parameters. The return reduction is the main goal of this KPI, and many other smaller, more specific KPIs contribute in return reduction. The current industry estimates a 9% return rate on electronical products. However, this variates up to 50% in different sectors like online apparel businesses (Minnema, 2017). Reducing this return rate can significantly reduce the costs of the total organization. An example with external hard disks at Coolblue showed a return reduction of 50%. Customers did not know how to use the hard disk and this resulted in a high return rate. Including a link to a short instruction video helped customers in installing the hard disk properly. This reduced the return rate of this product by 50% (Keswiel, 2018). Simple adjustments can result in return reduction, SL and USL can help to optimize these adjustments and create insights on the causes of returns.

USL can be applicable to prevent and detect fraudulent returns. Through the cluster capabilities of USL, the fraudulent returners can be identified and separated from the other clients. These fraudulent returns can be clustered by identifying their buying behavior. If this can be detected and labelled as a fraudulent returner. The website can offer them products with discounts which are not possible to return. Or the website can offer these customers products that have full refund coverage agreements with their suppliers. In this way, the e-commerce player secures itself and avoids high return costs. USL techniques will base their decision on all available information. Where at Amazon, persons were banned unfairly, an USL algorithm can prevent these unfairly bans. It will make the decision based on more data and complete information. By applying USL on the customer groups, outliers can be detected and the fraudulent returns can be warned and eventually be banned or penalized.

Reducing the rates of returns will have huge impacts on net profits and margins. If the return rate can be reduced by 50% through Machine Learning, this means a saving of 158.080 € per day, according to the numbers of the case study. To achieve this reduction, the integration of the Machine Learning techniques should work seamlessly and they need to have enough and trustful data. Further explanation on the business effect will be explained in 7.2. SL can be applied to make analyses of input data and match this with the output data. For instance, helping customers when they have troubles in choosing the right product. Based on their browse behavior, other products or short movie clips can be suggested to help them in choosing the right product.

USL will have other possible applications, as discussed in chapter 6, rewarding systems can be applied on the future behavior of clients. USL can help in predicting the future purchase behavior of customers, based on these analyses, the selling parties can reward the customer. Through the customer clusters, and the return behavior of the customers, every group of customers will be rewarded and targeted differently. A group with a higher return rate and high purchase rate could be approached differently than a group with low returns and low frequency of purchases.

7.1.2 KPI 2: Process Quality

The second KPI, Process Quality, is a broad KPI and can be measured on different levels in the SC. Due to the great diversity of this KPI, different applicability's are possible for the ML techniques. The process quality differs between companies, some will probably be easier satisfied with their quality than others. For example, the customer journey and satisfaction levels are different between Coolblue and Ali-express. Both companies have different focus and this translates into different standards on quality level. There are no ISO standards found for the process quality in returns. These ISO standards require a certain threshold for the quality in a company. Setting up ISO standards for the return process can contribute in the improvement of quality.

In this KPI, the customer satisfaction is a very sensitive factor. This KPI can be influenced both in a positive as negative way by ML techniques. The quality in the return process must remain high while innovations and automations are possible ways to cut on costs. It is important for the e-commerce players to find the right balance between the use of Machine Learning and remaining customer satisfaction high. With ML, the speed of responsiveness can be improved to instant replies when customers ask status updates of their returns. However, aspects as sensitivity and sympathy can be better performed by humans.

Where the first KPI, Return Volume, focusses on the prevention of return, this KPI specifies on the management and performance of the return process itself. How to perceive the best performance in the return process? In the return process, many contrasting interests are opposing each other. Where the customers want high service, quality and speed. The sellers want fast, reliable and cheap processes to be more efficient and less costly. These incentives can be conflicting and the adaptiveness and flexibility from the retailer side will be crucial to maintain their customers. In this KPI, the come-back rate of the customers is of huge importance, it indicates indirectly the attractiveness and quality of the e-commerce player. Only customers who are satisfied about the return process will come back and purchase again. Another way of retaining a high come back rate of customers is to offer them high service in the return process.

There are several possibilities of ML to improve the customer satisfaction in the return process. SL can use the input data; a customer that returned previously and was impatient in the process of returning. This means that such a customer contacted the customer service multiple times in the return process. Now, in case of a new return, provide continuous updates about their return can satisfy them and prevent customer contact moments with the service desk. SL can be applicable to reduce contact moments, this results in a reduction on labor intensity at the service desk. On the other hand, SL can also decrease customer satisfaction. Customers can react negatively on automation at the customer service. Therefore, the use of SL must be explored further to see the influence on customer satisfaction and to make sure that it will cause no harm on the relationship with the customers.

It is beneficial for both customer as retailer to reduce the processing time of a returned product. The customer will receive their refund faster and the retailer can gain advantage from speed due to lower costs and indirect higher customer satisfaction. When the total time of the return process is analyzed, a company can see how long the return process takes for a certain product (type) and see if it differentiates between days, weeks or certain periods in the year. If this indicates that a certain product type, returned on Fridays, always takes on average 2 days longer to be handled, a retailer can act upon this data and handle these products differently. The usage of SL and or USL can be applicable in this case to make analysis and make predictions for certain product groups. It can automatedly activate a special return program for sensitive products that are returned on Fridays. This raises awareness at the employees and makes sure they process the product in a short time period. Another possible option is to prioritize customers in the return process. Customers that are likely to repurchase can be prioritized in the return process and will receive their money faster.

7.1.3 KPI 3: Process Costs

The total costs of the return process must be spread over the different levels of the supply chain. However, in most cases the e-commerce player is responsible for all costs involved in the process. And when the product is returned, they have a damaged or opened product which cannot be sold as new anymore. To maintain a profitable margin on the products, the total process of returns must be lowered as much as possible. Beginning at the customer, products will be shipped back to the warehouse, repair center or other destinations. These return flows have some simple traceable costs such as; transportation costs and handling costs. Nonetheless, the return process also has indirect costs from back office and management. These indirect costs must be considered to have a full overview of the costs of returns. Several ML techniques can be applicable to reduce the costs in returns. The application of Supervised Learning can reduce costs by automating a part of the return process. The implementation of a machine that classifies returned products can lower the labor costs related to returns. However, it will require a change in process and demands for a big investment. To implement a machine and make all processes work, there is need for specific knowledge, data and money. It will require training and continuous updates about the products that are new in the assortment. This automation can reduce labor intensity and can create a seamless integration with information systems. This enables automatic updates to customers when the product arrived at the machine in the warehouse, this update can be given without human intervention.

In this field, Reinforcement Learning could be applicable to smoothen and automate processes in the warehouse. When the product needs to be stocked again, a potential application of RfL can be in putting the products back in stock. However, the application of RfL will be made not only in the return process. This will have an impact on the whole warehouse and therefore it is hard to see if the RfL agents can be applicable in the return process alone. If RfL is applicable for an organization, they probably will not only do it in the return department. It is a technique that will be implemented throughout the whole warehouse and therefore a focus just on returns is too narrow.

7.1.4 KPI 4: Supplier Agreements

It is vital for e-commerce players to have complete agreements and contracts with their suppliers. This takes out uncertainties in their relations and help both parties in solving issues that arise along the way. To collaborate in the return process and solve return related issues together, contracts are needed that cover most handling costs and refunds. By analyzing the data and specify on supplier level, a return rate per supplier can be identified. This data can be used to predict return rates of future product launches from suppliers. It will support the e-commerce players in their negotiation rounds with the suppliers. Using historical data for future negotiations do not necessarily have to include Machine Learning. These are rather simple calculations that can be seen through the historical data. However, including external factors such as; time of the year, weather or geographical location can be of an impact. ML algorithms can enable insights that cannot be achieved by other techniques and can be more precise on the prediction for future products.

Better relations between the stakeholders in the supply chain can eventually benefit all involved parties. This can lead to higher coverage of the refunds and better contracts. Where now ML is already applied in supplier selection to score them on supplier assessments, audits and credit scoring (Jenks, 2017). Including a return parameter in this metrics, there can be success from the first point of collaboration. The algorithm can support in choosing the right supplier, if for a certain product the return rate is high, independently the supplier. The parameter of return policy must be weighted heavier to find the supplier that covers the returns best. In this way, the return department is fully covered and can this can be translated towards a higher customer service.

To summarize this first paragraph, figure 24 is mapped to give an overview of the applicability of the different ML techniques on the identified KPIs.

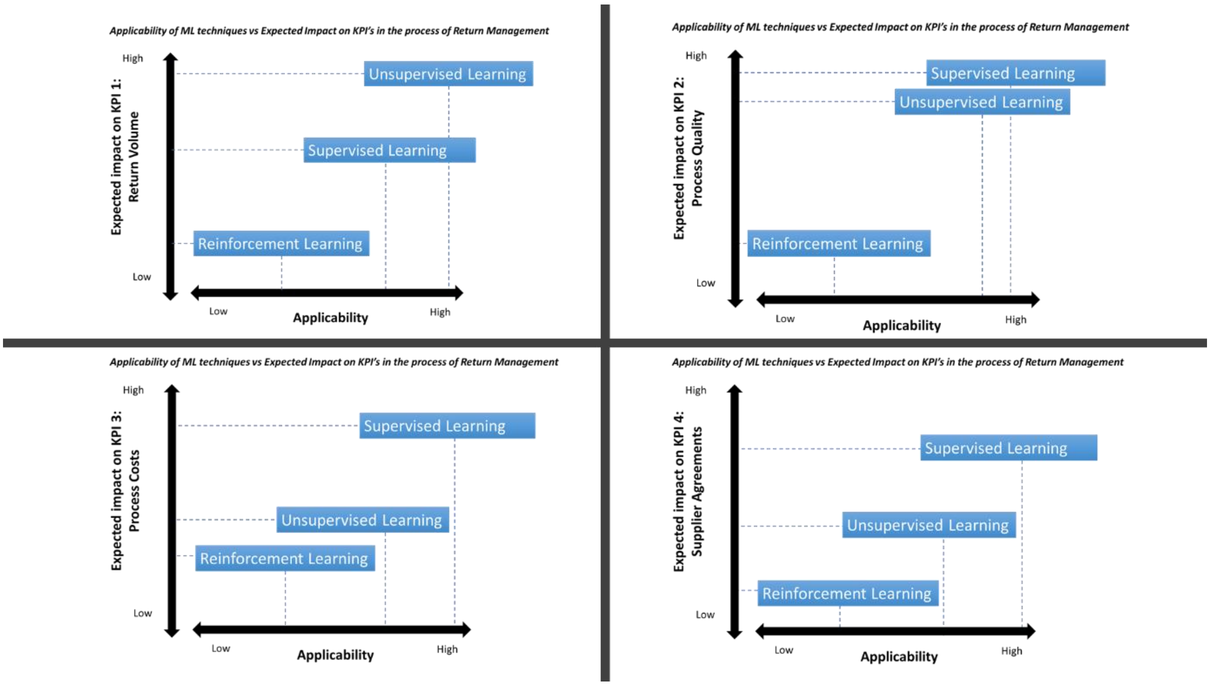


Figure 24, Applicability vs Expected impact on KPIs in Return process

In general, USL and SL will be most applicable in return management and can have the most impact on the KPI performance. Dealing with returns and the prevention of returns is a complex process that is in need for new developments and improvements. ML can be one of those new technologies that can improve this process and help e-commerce players with their costly returns. The applicability of the ML techniques is based on several aspects. First, based on the bottlenecks and processes behind the KPIs. These processes must be suitable to apply ML, otherwise there is now added value in the use of ML. Second, the available data will play a crucial role, huge amount of data is required to apply ML and receive useful insights.

7.2 Potential Gains and Impact of Machine Learning in Return Process

In this paragraph the potential gains and impact of Machine Learning on product returns will be presented. The effect and impact on the bottlenecks of the return process will be described and the potential improvements are made tangible. Figure 25 demonstrates the possible intrusions and interventions that can help in optimizing the return process. These are related back to the improvements on KPI measurement as depicted in paragraph 7.1. On the right-hand side of the figure, the potential impact is indicated. In process 1, Machine Learning can help in reducing the return rate with extensive analysis of purchasing and browsing behavior of the customers. Providing elaborated advice to the customers can provide support in their decision making. In figure 25, the return process has already started and the customer returns the item via one of the channels of the e-commerce player. However, in the forward process even more return reduction can be achieved through Machine Learning models. Process 2, which indicates the incoming returned products at the warehouse, can be improved by a text and image recognition. This will decrease labor intensity and can fasten the handling process of returns. The third possible impact of Machine Learning is related to the customer service. Customer contact moments can be decreased with the help of automation, which give customers continuous updates. This reduction of customer contacts can be throughout the whole return process.

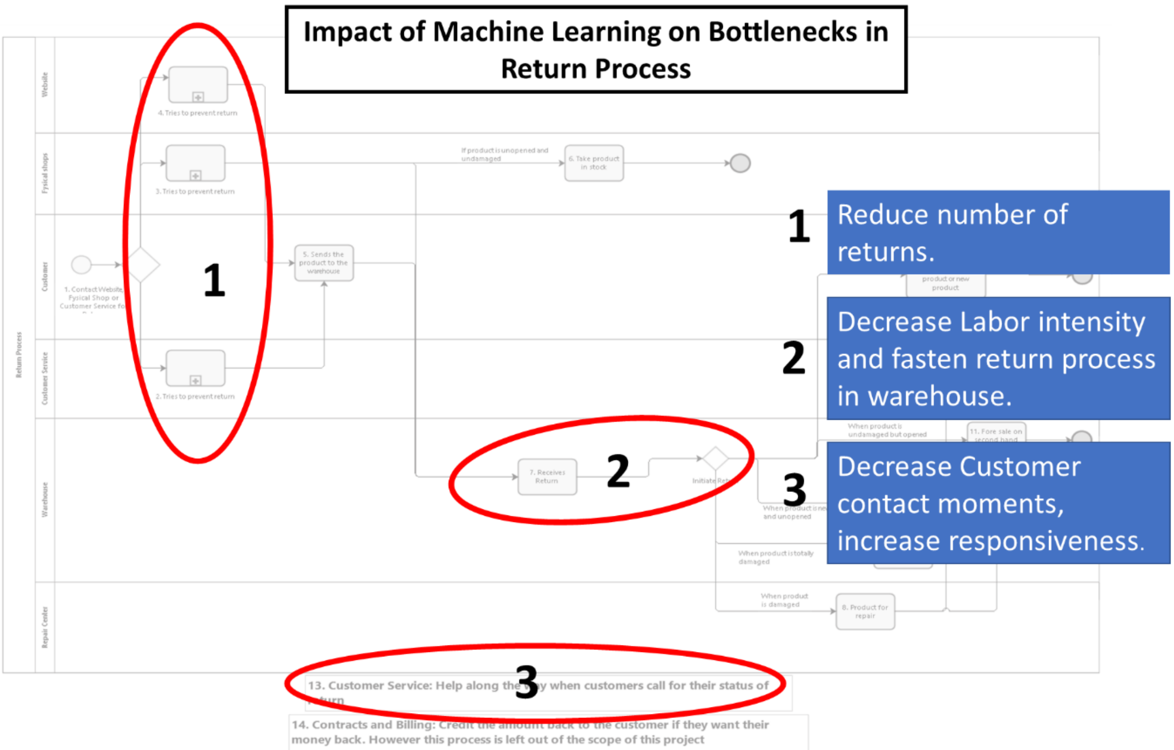
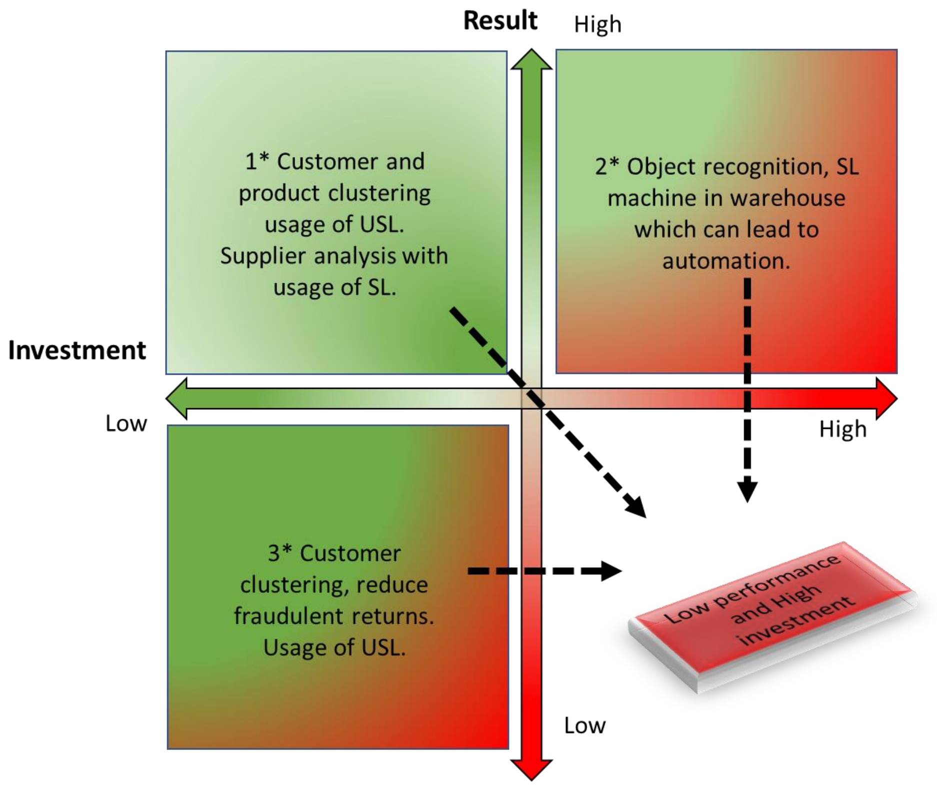


Figure 25, Potential Impact of Machine Learning in Return Process

A translation of these application into concrete results is needed to expand the insights of this research. This will show the impact on business performance and costs related to the return process. Figure 26 is presented which shows the level of investment on the horizontal-axis and the potential results on the vertical-axis. The three different blocks all indicate possible ML applications that can be applied to impact the performance of the return process. The lower right corner of figure 26, indicates an undesirable scenario. High implementation costs and low results are the consequence of bad implementation and lack of knowledge and structure. This scenario must be avoided to avert failure of implementations. The three desirable situations will be described to give concrete insights on the effect of ML on the return process.



26, Investment on ML techniques vs potential impact

7.2.1 Low investment and potential high result

In the first scenario, customer and product clustering can be used to reduce product returns in the e-commerce market. In the electronics e-commerce market, return rates are 9%, which results in approximately 1000 returns per day. According to the sales numbers found of the case study. Table 1 is presented again to indicate the numbers of return.

Sale of goods	Average amount per order	Number of orders	Percentage of Returns	Total returns	Returns per day	Total value of returns per day
1,1 billion €	320 €	3.437.500	9%	309.375	988	316.160 €

Table 5, Number of orders and returns case study

Target habitual returners with products in sale which cannot be returned can lead to a reduction of returns. The same for anonymous customers that are purchasing for the first time, if their behavior indicates a high probability for a return. This can be due to long browsing and comparing many products, this customer can also be directed to products that cannot be returned. Another option is to offer these customers videos where experts give advice about the products, this can help them in making a better choice. Additional on this, customers that have similar alternative products in their shopping cart are likely to return one of those. Probably they order the goods to compare and want to have a hands-on experience. The website can suggest them to try and feel the products in the physical shops or provide a video where an expert compares the products. Another way to prevent return with the use of current data can be done by product selection. For instance, a specific smartphone always returns after 7 months due to failure. This failure occurs when customers do not execute the required update. In the future, these customers can be contacted and reminded to do the update. This will prevent the returns of this smartphone. Being proactively involved with the customer that buys a specific product can help in preventing return and can avoid complaints from the customers. Another example can be with washing machines, when an e-commerce player experiences a high return rate of a certain washing machine after three years. When the problem is determined, it can be prevented by future customers. ML can help in detecting the customers with the affected product and can automatically send a message to warn the customer. Probably the washing machines needs cleaning or maintenance to prevent failure.

Approaching these customer clusters differently can help in reducing the return rate. If this results in lowering the return rate from 9% to 5% it can save money and labor. Returns per day can drop with almost 450, this means around 140.000 € in product value. This only indicates the value of the product, if the transportation costs and labor costs are included this will result in much more savings for the e-commerce player. The estimation of a 4% reduction is based potential of the machine learning applications. Using ML algorithms on their currently available data, will not require any big investments and will create huge impact in returns. The return volumes can be decreased through customer and product clustering. Based on customer behavior and product types, the clusters can be created to impact return rates.

Additionally, the existing data can be used to have facts based negotiations with the suppliers. The use of historical data from the supplier can help in predict future returns. Having these numbers at the negotiations rounds with the supplier can give the e-commerce player an advantage. Specification on reasons of return will provide insights of the quality of the product. Is the product returned due to damage or was the customer not satisfied about the product? All these indicators can help the e-commerce player in their collaboration with suppliers. Now they can ensure to have full cost covering on their returns from that supplier. The expected costs concerning returns were determined in advance and both e-commerce player as supplier knows what percentage of products will return. Forging the suppliers to help in covering the costs of returns, will make them more aware of their own product quality. They will strive for high quality products with low return rates. This will benefit the whole supply chain, they will try to guarantee their quality of products and make sure they minimize the amount of returns. The collaboration between e-commerce player and supplier results in a resilient and robust return process and can cover all costs for the e-commerce player. If the remaining 5% of returns is covered by the suppliers, this can save the e-commerce player up to 175.000 € per day.

In any scenario of the three desirable situations in figure 26, bad ML implementations and management can cause a translocation of the block towards the lower right corner. This includes high investment and low potential result. The dotted black arrows indicate these translocations. This can be caused due to the continuous costs for implementing and maintenance of a ML implementation project. It is no exception for IT projects to go over budget and time limit. Therefore, there is need for an efficient strategy and an experienced and skilled team. First, smaller test pilots must be created to only focus on a small group of clients or products. This can show the results of the ML techniques, if the test is a success, it can be rolled out wider to impact the whole return process. Bad implementation cannot only impact the organization itself, it can also affect customers and they can be lost due to bad service or other failures. Communicating with the customers in a more automated way, can lower the customer satisfaction. These issues indicate the importance of the implementation and a good approach to make ML work in the Return process is required.

7.2.2 High investment and potential high results

High investment and high results, the upper right corner of the figure, new investments are required to make the ML techniques work. In this situation, an intrusive design is needed in the return process and this requires relatively big investments. A possible solution in the return process, that will require a relatively big investment, is the implementation of a text and image recognition machine that handles all incoming returned products in the warehouse. As presented in the sorting machine in the union industry, such classification machines can drastically lower labor intensity. Based on the number of employees working at the incoming product returns in a warehouse, a business case can be created for an e-commerce player to explore their potential gain and calculate the cost savings. Such machines can easily take over human labor and can even lead to higher quality of the process (JDC, 2018).

In the current situation, an average of 1000 products enter the warehouse per day. These returns are now manually handled and require intensive labor from employees. A returned product must be initiated, linked to the right customer and directed to the next location. These products will return on various times and will never come all in once. For this reason, it is hard to schedule the right amount of people for these return handling. Making use of a text and image recognition machine, the complexity and intensity of the labor can be reduced. The machine can do the handling and send the product towards the next destination. Handling time of a return is now estimated on 5 minutes per product. This includes unpacking, initiating the product, linking the product to the customer and determine the next destination. For 1000 products this means 5000 minutes of work. Which approximately will take 10 persons with an 8-hour workday, if this can be decreased by 80% due to the handling capabilities of the machine, the handling process will be much cheaper. Based on a 15 € hourly tariff, approximately 1000€ per day can be saved. The machine will be less error sensitive compared to human workers, it will improve over time and will grow performance due to the increase of data. Another issue in the warehouse is related to the retardation of product returns. Currently, when a huge amount of product returns arrive at the same time and employees are busy with other activities, they will leave the returns and focus on other tasks. This can delay the process of return and results in a slower refund to the customer. Implementing a text and image recognition machine can prevent such congestions through the reduction of dependency on human labor.

Such a machine will change the way of working in the warehouse. When this machine performs well, it can be an example for other areas in the warehouse and it can open new ways of working. This investment is probably quite high, however can create huge advantages and cost savings in business processes. When the classification is successful, other parts of the warehouse can benefit. It can impact indirectly other processes and will be able to reduce labor intensity. To successfully implement Machine Learning, the current process must be analyzed and all related costs and people must be indicated. The company can propose a desirable situation with design requirements. Based on these requirements, the ML applications can be tested to see if they can create business performance improvements.

7.2.3 Low investment and potential low result

In the lower left corner, low investment and low results, the use of current data can lead to relatively low results. This is relatively a small percentage of the total returns so the impact and result of this improvement on KPIs will be relatively smaller compared to the first two solutions. Possible applications of the USL and SL algorithms can be detected along the way, when the data is extensively used. Therefore, an actual data simulation and implementation of ML will be needed as next steps for this research. This will present the possible impact and improvement on KPI level and can open new ways for dealing with the 'monster of returns'. The problems with returns are by no means solved with Machine Learning, step by step businesses and literature will need to keep pushing in this area to reduce returns and deal with them properly.

In electronics, the fraudulent return rate is lower compared to the apparel industry. In the apparel industry, 8% of the total returns in this sector is related to fraudulent returns (Appris, 2017). A return can be fraudulent in different ways. For instance, people order many cheap products and request to return all of them after a couple of days. These people hope that the e-commerce player does not desire the product anymore because of the high costs of the return process. In the apparel industry, fraudulent returns return their piece of clothing after they worn it. They do not want to buy an expensive garment and return it after their night out. This can also happen in the electronics markets, customers that order a lens for their DSLR camera. Go on holidays and use the lens for one week, when their holiday is over, they return the lens again. European laws and regulations make it very hard for the e-commerce players to refuse or decline such returns. By extensive use of Machine Learning algorithms these fraudulent returns can be detected and prevent. The clustering and classification of such 'fraudulent returners' can be done based on their browsing and purchase behavior. Probably type of product will play a key role. Or geographical location can be important, where certain areas of a country have richer or poorer people. All these factors must be indicated to come to a rightful conclusion. The annual revenue of Zalando in 2017 contained 4,4 billion euros. 8% of fraudulent returns means a value of 320 million euros. Reducing the fraudulent return rate in the online apparel industry can create a bigger impact if it is compared to the online electronics industry. Therefore, these ML applications will need to be further specified on market level or even company level, to indicate the specific impact on business performance.

8 Conclusion and Discussion

The issues regarding Return Management in the e-commerce market created the basis for this research. This study tried to explore the possible applications of ML in the complex environment of product returns in the Dutch e-commerce market. The involvement of different actors and conflicting arguments between stakeholders makes the returns a very interesting field to research. Continuous research and improvements will be needed to make sure that companies can deal efficiently with their product returns. This research has led to an answer on the following research question:

Research Question:

"Concerning the high product returns in the Dutch e-commerce market; What can be the applicability and the potential impact of Machine Learning on the Key Performance Indicators measured in the process of Return Management?"

This research question was supported by several sub-questions that all contributed in finding the possible applications of ML in Return Management. To conclude, this study indicated that Supervised Learning and Unsupervised Learning are most applicable in the process of Return Management. Reinforcement Learning is less applicable, due to the complexity and scale of possible implementation. Both SL and USL, will have potential beneficial impacts for organizations that deal with product returns in the e-commerce market. Extensive use of current data with help of Machine Learning can lead to a decrease in costs. Labor intensity in different departments of an e-commerce player can be reduced, both at customer service level as operational levels in the warehouse. It can increase the responsiveness of the e-commerce player, which can lead to higher customer satisfaction. Eventually product returns can also be prevented, by clustering certain customer groups and provided them with matching reward systems. Furthermore, the relationship with the supplier plays a key role in the process of returns. Machine Learning can help in predicting future product returns, based on historical data of a certain supplier. This can support e-commerce players in having the right agreements with their suppliers, which will give them the opportunity to offer high customer loyalty in the process of returns. Serving the customer at a high level in the return process will be beneficial for the whole supply chain.

Mainly two solution streams were identified that can be applicable for companies to further explore. These solution streams are presented in figure 27. The first possible application of ML is the extensive usage of current and existing data. This can be used in the forward process to lower the return rate and specify on product and customer level. This will create better insights into the reasons of returns, customers preferences and returning behavior. This can be used to predict returns in future situations and help in reducing and preventing returns. This first possible application of ML will require relative small investments and is less intrusive to implement, it will not affect the business organization structure and processes.

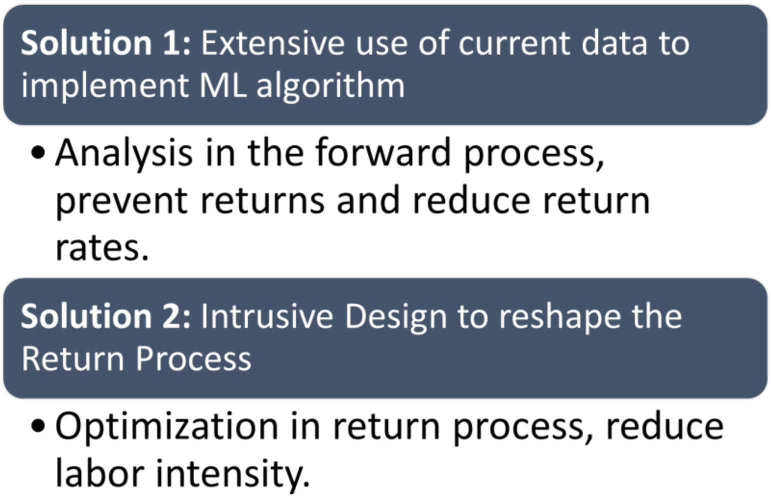


Figure 27, Solution Design of ML applicability in Return Process

The second possible application of ML is the intrusive design in the return process. The reshaping of the total process can optimize the return process and can contribute in reducing labor intensity and increasing communication. It can help in making the process faster and there will be more automation due to the ML algorithm in the machine that handles all incoming returns at the warehouse. Such machines can be very efficient and increase quality and reduce labor intensiveness as we earlier saw in the union production line (JDC, 2018). This possible application of ML will ask for a bigger investment and will have a heavier impact on the business organization. Where unions did not change in 20 years, e-commerce players have fast changing products in their assortment. This makes it harder for ML to classify all new products, is the machine able to process these new products without any major improvements? These are implications that must be considered when describing the potential results of ML on KPIs in the return process. However, this study did not implement ML algorithms and therefore cannot say much about the potential impact it will create. This research tried to clear some of the questions and issues around product returns in the e-commerce market. Now, future research will need to implement a Machine Learning algorithm. It will need to test if the data and processes are suitable and have potential to improve. The e-commerce market will keep on growing in the coming years and so will product returns. The data driven attitude of these e-commerce players can support in the battle against returns. This will be necessary to survive in the market and keep up with the competitors. The exploration of disruptive technologies is important to create an innovative character and focus on continuous improvement of performance in the field of product returns.

8.1 Implications for Business, Society and Science

This research is partly based on a case study that was conducted with an e-commerce player operating in the Dutch electronics market. This research gives new insights for the business on how to deal with their high product returns. These e-commerce players are having hard times dealing with the high returns and therefore new technologies and applications can improve business performance. They can use the outcome of this research to further develop their data structures and data gathering to enable the strength of data analysis and ML algorithms. It can change their way of communicating with clients in the return process and can help in fastening the process of returns. Probably these huge e-commerce players are already using their data efficient and maybe already use some ML algorithm in parts of their businesses. This study showed them to expand this data opportunities into the field of Return Management and explore the ML possibilities in this field.

Returns became part of the shopping behavior of society. Returning products became normal and customers expect e-commerce players to accept the returns free of charge. The ease of return will not change, it is simply too valuable for the e-commerce player and they use it as a competitive factor. Customers can be approached in differently than before, their online behavior will play a more important role. Return behavior of the customers will be analyzed more extensive and their return behavior will shape the relationship with the e-commerce player. The data of the public will become more and more important and e-commerce players will need this data to exploit ML. Regulations as GDPR, must ensure that data of the customer is secure and used by the rightful parties. These changes for both business and society can influence the success factor of the implementation of ML algorithms.

On scientific level, this research expanded the knowledge of Machine Learning applications in the field of Return Management. This research indicated the applicability of ML techniques in return management and concluded that both SL as USL will be most suitable. There was little, to almost no knowledge available of linking product returns with ML. This research tried to explore the possibility of linking these two concepts to create a better, less costly and more efficient way of dealing with product returns in the e-commerce market. Where lately the research into return management has increased, there is still much left to explore. The case study example gave the possibility to grasp the essence from the business side and indicated how they are dealing with the returns. Combining this with the literature knowledge and the performance frameworks, this research gave new insights on the impact on the KPIs in the return management process.

8.2 Limitations of research

This research is based on literature and a case study in the Dutch e-commerce market. While this case study gives a good representation of the Dutch e-commerce market, it is not completely representable. Every company will have different return processes and will set different standards on their quality. This research would have higher representativeness if more e-commerce players were interviewed. However, due to time constraints and the focus of this study it was not relevant to include more case studies. The case study was conducted with a successful company that already had a strong focus on returns and are managing returns quite successful. A case study with a company that is less successful in terms of Return Management could have created other insights. The fact that Amazon is destroying their returns indicates huge problems and bad management in their return department (NOS, 2018). Interviewing Amazon would probably have led to different outcomes of this study. It would also create a broader or different set of KPIs that cover the performance in Return Management. The case study was conducted with an e-commerce player in the electronical market, every market will have different return process and return rates. Therefore, this research is only representative to the comparable companies.

The environmental aspect of product returns is left out of the scope of this project. More and more businesses are focusing on their green footprints and trying to improve on these aspects. Closed loop supply chains are currently extensively researched in literature and this in combination with return management does give very interesting material to research. However, the bottlenecks and KPIs related to environmental and sustainable issues are left out of the scope of this project. In this research 4 main KPIs were identified and chosen to measure the performance in the return management. A more extensive search on KPIs and a bigger set of KPIs could have created a broader overview of the performance in product returns. However, the researcher believes that the essence of return management is cropped into the 4 KPIs presented. The focus of this research was to explore the applicability of ML in product returns, and therefore the KPIs from literature and case study analysis were considered as complete and sufficient.

This research only focused on three Machine Learning techniques. Where these techniques are currently most used and explored in literature, there are many other possible ML techniques that could be applicable in Return Management. However, due to limited time, limited knowledge and personal interests these three (SL, USL and RfL) were chosen to represent ML in this research. The researcher of this study thought that these three techniques could be most potential and applicable in the product return market. Other innovative techniques could be explored in combination with Return Management. The use of Blockchain can enable customer to customer shipments without the intervention of the e-commerce players. This will enable product returns from a customer to another customer. This will safe time, money and resources. Further research is needed to explore the ability of Blockchain to create a more robust and efficient Return Management network.

This research did not implement or test any ML algorithm. Therefore, the real impacts or effects that ML techniques will create cannot be describe in this research. This research mentioned the possible applicability of ML techniques and did not try to show the results of an implementation. This can be considered as a next step in the exploration of ML in return management.

There were no ethical or data privacy issues explored that will be playing important roles in the applicability of Machine Learning. The influence of GDPR regulations will affect the data gathering of the e-commerce players. These are factors that will need to be addressed in future research to see what their effect and influence will be. The company willingness for example will play an important role when ML is explored. The disruptions in process and changes in way of working cannot be underestimated and companies will need to prepare for these changes. Another requirement is related to the data, data must be in the right order and condition to be used properly and efficient. This includes data cleaning and management, however due to the explorative approach of this study the data requirements are left out of scope.

8.3 Reflection on research

The case study approach was chosen to have an example from the business side and have real numbers, data and information about returns in the current market. While this was used as basis, frameworks from both Bernon et al. (2011) and Shaik and Abdul-Kader (2012) were used to support the identified KPIs for this research. These papers were one of the few that researched into the field of performance measurements in reverse logistics related topics. This limited basis of research gives a lot of opportunities to explore in the future, where on the other hand it gives uncertainty about measuring the right KPIs.

The process of returns was mapped based on the BPMN approach. This is now a static model and was not tested or evaluated. This model can be simulated to see what the effect can be of several scenarios. These scenarios can help in implementing new strategies and can show the effect of several intrusions in the design of the return process. Now the BPMN was only used to identify bottlenecks and give a schematic overview of the process, it however has far more opportunities to help in researching a specific process. This research tried to identify a knowledge problem. Where there was no current knowledge of the application of ML in the field of product returns. And secondly, this research tried to explore a design problem on how to apply ML in the field of returns. Where the first problem was clearly identified and this study shaped the link between ML and product returns. The second, design related problem, was harder to solve and still leaves openings to further explore. Possible ways of implementation and application of ML algorithms are suggested, however design requirements and details are not presented. Data requirements, organizational structures and many more issues will play a role in making ML algorithms a success in return management. This research made assumptions about the available customer data and this will need to be further explored future research.

This study mapped an existing process of a case study, this was used to go into depth on the ML techniques. A more comparative study that analyses multiple e-commerce players in the Dutch market could be an option to validate the results better. The single case study in this research makes the quality and results of this research somewhat biased and a multiple case study analysis could help in identifying a broader supported performance measurement structure.

8.4 Recommendation for Future Research

In this section, possible opportunities for future research will be suggested. Based on this research and the identified gaps, there is research needed to show the impact of ML algorithms in the return management process. Future research can model and test algorithms to see what they potentially can bring in the field of returns. What will be data requirements relating to the return data to make ML work? Can every e-commerce player use Machine Learning to reduce their returns or will it not profitable enough for certain companies to invest in Machine Learning? Comparable questions must be further researched to explore the effect and potential success of implementing Machine Learning in the return process. Actual implementations and test pilots or Proof of concepts will show the impact of Machine Learning. Company requirements must be aligned and determined to create a selective group of possible companies.

This study tried to clear the bush and explore the potential applications of Machine Learning. Now, future research will need to go into detail about data requirements and structures. For instance, the effect of so called ‘drop-offs’ must be explored. These are customers that first compare all their products online. After they compared everything, they take some time to think about their purchase. Than later they go to the website, possibly on another device, and order directly one of the products he compared before. The algorithm will not detect this customer as a ‘high probability returner’, the effect of such ‘drop-offs’ must be researched to see the applicability of Machine Learning. Other interesting issues for future research are related to the fast change in product assortment and the frequency of purchasing. Continuously, new products are launched on the market, customers will develop other behavior on the new products. Can Machine Learning algorithms deal with these fast-changing assortments in products? And do customers buy frequently enough to identify them as ‘habitual returners’?

Not only algorithms have to be tested, the organization structure and the company willingness is of importance when new techniques are implemented. Therefore, future research is needed on more ethical issues and the willingness of all stakeholders involved. Not only the company employees, also customers and suppliers will have to deal with the new technologies. The field of Reinforcement Learning is still quite unexplored however it is a very promising technique that can have potential impact in the supply chain. Future research can contribute to see what the specific applications are for RFL in the field of returns.

E-commerce players are not the only companies that suffer under returns. Transporters are influenced as well by the massive increase in returns. DHL expected a bold decrease in their expected profits over the year 2018. More than 1 billion euro less in profits is expected for 2018 (Gurp, 2018). The moved an increase number of package however are experiencing huge problems with the return flows. They have no idea how to be profitable in their return process. These results show the impact of product returns throughout the different parties in the supply chain. Return volumes are harder to predict and transporters do not know how many trucks and drivers are needed. New possibilities of transportation must be research in the return sector. Predicting the amount of returns will play a key role for these transporters to know how many trucks they need for the transportation.

Apart from the machine learning opportunities in product returns, return management itself still has a lot to improve. Not only machine learning algorithms can contribute in the field of product returns. Other methods or techniques can contribute in lifting the return process to a higher level. Defining more extensive KPIs to better measure the performance of the return process. Explore the possibilities of outsourcing vs inhouse dealing with returns, these are opportunities to improve the performance in product returns in another way. Another option for future research is the effect of physical shops on the product returns. Can it be beneficial for web shops to open physical shops that will impact their return department? Comparative analysis can be conducted to see what can be most beneficial for companies and the impact of returns can be measured on the business performance.

There will come up some moral and ethical issues when applying ML algorithms in the returns. Based on customer specific data, probably one customer will be offered a discount on their next purchase where another customer, that never returns, will not receive a discount. Is this morally acceptable, prelude one customer before the other? Studies will have to conduct research in this dilemma to see what the possibilities are of these rewarding programs. The new regulations from EU policies will influences the exploration and possibilities of ML. GDPR regulations will affect the market, research is needed to see what the effect can be and see if this significantly change the effectiveness of ML in e-commerce.

This study scoped out the environmental effects of the returns and the institutional externalities involved. These two subjects will play important roles in the future and possible solutions and new designs must be explored to see where the product returns are shifting. EU legislation is probably lacking behind due to the fast changes and movements in the e-commerce markets and there can be need for change in laws and regulations. What can be the role of an external regulator that performs checks on product returns and makes sure that both customer and retailer are treated well. Both knowledge as design research is needed in the domain of product returns. Wherever it is in combination with Machine Learning techniques or other disruptive technologies. The field of product returns in the e-commerce market will remain complex and the dynamic environment gives enough incentives to perform explorative and research studies in the future.

9 Literature

Akkermans, H.A., and Van Wassenhove, L.N. 2013. “Searching for the Grey Swans: The Next 50 Years of Production Research.” International Journal of Production Research 51 (23–24):6746–55.

Appriss, (2017) Optimize retail returns. Retrieved from: https://apprissretail.com/solutions/verify/

Aramyan, L., Lansink, A., Vorst, J. and Kooten, O. (2007), “Performance measurement in agri-food supply chains: a case study”, Supply Chain Management, Vol. 12 No. 4, pp. 304-15.

Awaysheh, A., & Klassen, R. D. (2010). The impact of supply chain structure on the use of supplier socially responsible practices. International Journal of Operations & Production Management, 30(12), 1246-1268.

Bernon, M, Silvia Rossi, John Cullen, (2011) "Retail reverse logistics: a call and grounding framework for research", International Journal of Physical Distribution & Logistics Management, Vol. 41 Issue: 5, pp.484-510:

Beske-Janssen, P., Johnson, M. P., & Schaltegger, S. (2015). 20 years of performance measurement in sustainable supply chain management–what has been achieved?. Supply chain management: An international Journal, 20(6), 664-680.

Blut, M., Frennea, C. M., Mittal, V., & Mothersbaugh, D. L. (2015). How procedural, financial and relational switching costs affect customer satisfaction, repurchase intentions, and repurchase behavior: A meta-analysis. International Journal of Research in Marketing, 32(2), 226-229.

Branch, S. T. (2017). A new enhanced support vector model based on general variable neighborhood search algorithm for supplier performance evaluation: A case study.

Brignall, 2016 Banned by Amazone by returning faulty goods. Retrieved from: https://www.theguardian.com/money/2016/mar/18/banned-by-amazon-returning-faulty-goods-blocked-credit-balance

Brooks, R. A. (1991). Intelligence without representation. Artificial intelligence, 47(1-3), 139-159.

Chae, B., Olson, D., and Sheu, C. 2014. “The Impact of Supply Chain Analytics on Operational Performance: A ResourceBased View.” International Journal of Production Research 52(16):4695–710.

Choy, K. L., Lee, W. B., & Lo, V. (2003). Design of an intelligent supplier relationship management system: a hybrid case based neural network approach. Expert Systems with Applications, 24(2), 225-237.

Coase, R. (1998). The new institutional economics. The American Economic Review, 88(2), 72-74.

Cohen, P. R., & Feigenbaum, E. A. (Eds.). (2014). The handbook of artificial intelligence (Vol. 3). Butterworth-Heinemann.

CoolBlue. (2018) Jaarverslag 2017. Retrieved from: https://nieuws.coolblue.nl/jaarboek-2017/

Dallemule, L., & Davenport, T. H. (2017). What’s Your Data Strategy?. Harvard Business Review, 95(3), 112-121.

Davidson, S., De Filippi, P., & Potts, J. (2016a). Disrupting Governance: The New Institutional Economics of Distributed Ledger Technology. Available at SSRN 2811995

Dennis. S. (2018) The ticking Time Bomb of E-commerce Returns. Forbes. https://www.forbes.com/sites/stevendennis/2018/02/14/the-ticking-time-bomb-of-e-commerce-returns/#19267144c7f1

Ding, Y., Xu, H., & Tan, B. C. (2016, June). Predicting Product Return rate with" Tweets". In PACIS (p. 345).

Ellis. S. (2016). IDC: Launch of Watson Supply Chain looks to address gaps in data analytics and resources. Marketplace. Retrieved from: https://www.ibm.com/blogs/watson-customer-engagement/2016/10/28/idc-launch-watson-supply-chain-looks-address-gaps-data-analytics-resources/

Fallahpour, A., Wong, K. Y., Olugu, E. U., & Musa, S. N. (2017). A Predictive Integrated Genetic-Based Model for Supplier Evaluation and Selection. International Journal of Fuzzy Systems, 19(4), 1041-1057.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144.

Gao, J., & Jamidar, R. (2014). Machine learning applications for data center optimization. Google White Paper.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS quarterly, 27(1), 51-90.

Gelbrich, K., Gäthke, J., & Hübner, A. (2017). Rewarding customers who keep a product: How reinforcement affects customers' product return decision in online retailing. Psychology & Marketing, 34(9), 853-867.

George, M., Kumar, V., Grewal, D., (2013). Maximizing Profits for a Multi-Category Catalog

Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning. Machine learning, 3(2), 95-99.

Gomez-Uribe, C. A., & Hunt, N. (2016). The netflix recommender system: Algorithms, business value, and innovation. ACM Transactions on Management Information Systems (TMIS), 6(4), 13.

Govindan, K., Soleimani, H., & Kannan, D. (2015). Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future. European Journal of Operational Research, 240(3), 603-626.

Gunasekaran, A., Patel, C., & McGaughey, R. E. (2004). A framework for supply chain performance measurement. International journal of production economics, 87(3), 333-347.

Gurp Van. T., (2018) Winstwaarschuwing DHL: ebit komt bijna 25% lager uit. Retrieved from: https://www.nieuwsbladtransport.nl/de-markt/2018/06/11/winstwaarschuwing-dhl-ebit-komt-bijna-25-lager-uit/?utm_source=ntemail&utm_medium=nieuwsbrief&utm_campaign=ntemail

Ha, S. H., & Krishnan, R. (2008). A hybrid approach to supplier selection for the maintenance of a competitive supply chain. Expert systems with applications, 34(2), 1303-1311.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning. In The elements of statistical learning (pp. 485-585). Springer, New York, NY.

He, Y. (2015). Acquisition pricing and remanufacturing decisions in a closed-loop supply chain. International Journal of Production Economics, 163, 48-60.

Heiman, A., McWilliams, B., & Zilberman, D. (2001). Demonstrations and money-back guarantees: Market mechanisms to reduce uncertainty. Journal of Business Research, 54(1), 71-84.

Heydari, J., Choi, T. M., & Radkhah, S. (2017). Pareto improving supply chain coordination under a money-back guarantee service program. Service Science, 9(2), 91-105.

Hielscher, H., Goebel J., & Brück M., (2018) Amazon vernichtet massenhaft Neuware. Retrieved from: https://www.wiwo.de/unternehmen/handel/das-ist-ein-riesengrosser-skandal-amazon-vernichtet-massenhaft-neuware/22658544.html

Huang, R. H., Yang, C. L., Wuang, M. S., & Tsui, C. S. (2010, June). Constructing a performance evaluation model for reverse logistics—Cases of recycled tire traders. In Management of Innovation and Technology (ICMIT), 2010 IEEE International Conference on (pp. 606-611). IEEE.

Huscroft J. R. Benjamin T. Hazen Dianne J. Hall Joe B. Hanna, (2013),"Task-technology fit for reverse logistics performance", The International Journal of Logistics Management, Vol. 24 Iss 2 pp. 230 - 246

Ivan, M. (2018) Types of Machine Learning. https://en.proft.me/2015/12/24/types-machine-learning-algorithms/

Jarvis. J. (2017) Predictions for the state of AI and Robotics by 2025. Retrieved from: http://www.pewinternet.org/2014/08/06/predictions-for-the-state-of-ai-and-robotics-in-2025/

Jayaraman, V., Ross, A. D., & Agarwal, A. (2008). Role of information technology and collaboration in reverse logistics supply chains. International Journal of Logistics: Research and Applications, 11(6), 409-425.

JDC, (2018) Optically sorting onions with a click of the mouse. Retrieved from: http://www.jdcgrading.com/projects/optically-sorting-onions-with-a-click-of-the-mouse/

Jeng, S. P. (2017). Increasing customer purchase intention through product return policies: The pivotal impacts of retailer brand familiarity and product categories. Journal of Retailing and Consumer Services, 39, 182-189.

Jenks, M. (2017) Applications of artificial intelligence for your supply chain. Retrieved from: https://medium.com/@KodiakRating/6-applications-of-artificial-intelligence-for-your-supply-chain-b82e1e7400c8

Kahraman, C., & Oztaysi, B. (2014). Supply Chain Management Under Fuzziness. Springer.

Kahraman, C., Cebeci, U., & Ulukan, Z. (2003). Multi-criteria supplier selection using fuzzy AHP. Logistics information management, 16(6), 382-394.

Keswiel, M., (2018) Zo doorbrak Coolblue de magische 1 miljard omzetgrens. Retrieved from: https://www.sprout.nl/artikel/retail/zo-doorbrak-coolblue-de-magische-1-miljard-omzetgrens-6-redenen

Ko, M., Tiwari, A., & Mehnen, J. (2010). A review of soft computing applications in supply chain management. Applied Soft Computing, 10(3), 661-674.

Kobayashi, T., Tamaki, M., & Komoda, N. (2003). Business process integration as a solution to the implementation of supply chain management systems. Information & Management, 40(8), 769-780.

Krishna, G. J., & Ravi, V. (2016). Evolutionary computing applied to customer relationship management: A survey. Engineering Applications of Artificial Intelligence, 56, 30-59.

Kumar, S., & Putnam, V. (2008). Cradle to cradle: Reverse logistics strategies and opportunities across three industry sectors. International Journal of Production Economics, 115(2), 305-315.

Kuo, R. J., Wang, Y. C., & Tien, F. C. (2010). Integration of artificial neural network and MADA methods for green supplier selection. Journal of cleaner production, 18(12), 1161-1170.

Laan. M. (2015) Gepersonilaseerde webshops voor meer conversie. Retrieved from: https://www.guapa.nl/gepersonaliseerde-webshop-voor-meer-conversie/

Lambert, D.M. and Pohlen, T.L. (2001), "Supply chain metrics", International Journal of Logistics Management, Vol. 12 No. 1, pp. 1-19.

Lambert, S., Riopel, D., & Abdul-Kader, W. (2011). A reverse logistics decisions conceptual framework. Computers & Industrial Engineering, 61(3), 561-581.

Leancor, (2018) end to end supply chain improvement. Retrieved from: https://leancor.com/end-to-end-supply-chain-advancement/

Li, G., & Li, W. (2015). The Analysis of Return Reverse Logistics Management Strategy Based on B2C Electronic Commerce.

Lima-Junior, F. R., & Carpinetti, L. C. R. (2016). Combining SCOR® model and fuzzy TOPSIS for supplier evaluation and management. International Journal of Production Economics, 174, 128-141.

Lycett, M. 2013. "Datafication: Making Sense of (Big) Data in a Complex World." European Journal of Information Systems 22(4):381–86.

Macua, S, V. (2015) https://www.quora.com/What-are-the-disadvantages-of-reinforcement-learning-And-what-would-be-a-good-alternative-to-use-cases-where-its-used

Malhotra, R. (2015). A systematic review of machine learning techniques for software fault prediction. Applied Soft Computing, 27, 504-518.

McFarland. M., (2016) Amazon only needs a minute of human labor to ship your next package. Retrieved from: http://money.cnn.com/2016/10/06/technology/amazon-warehouse-robots/index.html

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. Journal of Business logistics, 22(2), 1-25.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning.

Moghaddam, K. S. (2015). Fuzzy multi-objective model for supplier selection and order allocation in reverse logistics systems under supply and demand uncertainty. Expert Systems with Applications, 42(15-16), 6237-6254.

Ng, A. (2017, February). Artificial Intelligence is the New Electricity. In presentation at the Stanford MSx Future Forum.

Ngai, E. W. T., Peng, S., Alexander, P., & Moon, K. K. (2014). Decision support and intelligent systems in the textile and apparel supply chain: An academic review of research articles. Expert Systems with Applications, 41(1), 81-91.

Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications, 36(2), 2592-2602.

North, D. C. (2003). The new institutional economics and third world development. In The new institutional economics and third world development (pp. 31-40). Routledge.

Olugu, E. U., & Wong, K. Y. (2012). An expert fuzzy rule-based system for closed-loop supply chain performance assessment in the automotive industry. Expert Systems with Applications, 39(1), 375-384.

Oroojlooyjadid, A., Nazari, M., Snyder, L., & Takáč, M. (2017). A Deep Q-Network for the Beer Game with Partial Information. arXiv preprint arXiv:1708.05924.

Petersen, J. A., & Kumar, V. (2009). Are product returns a necessary evil? Antecedents and consequences. Journal of Marketing, 73(3), 35-51.

Petersen, J. A., & Kumar, V. (2015). Perceived risk, product returns, and optimal resource allocation: evidence from a field experiment. Journal of Marketing Research, 52(2), 268-285.

Prakash, O., & Pandey, V. (2014) Bullwhip effect decreased by AI techniques. Journal of Production Research & Management. 4(2), 1-13.

Pur, S., Stahl, E., Wittman, M., Wittman, G., and Weinfurter, S. (2013) Retourenmanagement im online handel. Das beste daraus machen, Regensburg: ibi research an der Universität Regensburg GmbH.

Quotenet,. (2017) Miljoenen bestelde pakjes kunnen niet voorkomen dat Wehkamp nog altijd verlies lijdt. Retrieved from: http://www.quotenet.nl/Nieuws/Miljoenen-bestelde-pakjes-kunnen-niet-voorkomen-dat-Wehkamp-nog-altijd-verlies-lijdt-201302

Rajesh, R. (2016). Forecasting supply chain resilience performance using grey prediction. Electronic Commerce Research and Applications, 20, 42-58.

Rajesh, R., & Ravi, V. (2015). Supplier selection in resilient supply chains: a grey relational analysis approach. Journal of Cleaner Production, 86, 343-359.

Raman, V., & Reddy, S. (2017). Predicting Delivery Time of Components in a Supply Chain.

Ravi, V., Shankar, R., & Tiwari, M. K. (2005). Analyzing alternatives in reverse logistics for end-of-life computers: ANP and balanced scorecard approach. Computers & industrial engineering, 48(2), 327-356.

René B.M. de Koster Marisa P. de Brito, Masja A. van de Vendel, (2002) "Return handling: an exploratory study with nine retailer warehouses", International Journal of Retail & Distribution Management, Vol. 30 Issue: 8, pp.407-421,

Reverse Logistics Executive Council (2007), glossary, available at: www.rlec.org/glossary.html

Rigby, D. K., (2014) Online shopping isn’t as profitable as you think. Retrieved from: https://hbr.org/2014/08/online-shopping-isnt-as-profitable-as-you-think

Rogers, D. S., Lambert, D. M., Croxton, K. L., & García-Dastugue, S. J. (2002). The returns management process. The International Journal of Logistics Management, 13(2), 1-18.

Ronald S. Tibben-Lembke, Dale S. Rogers, (2002) "Differences between forward and reverse logistics in a retail environment", Supply Chain Management: An International Journal, Vol. 7 Issue: 5, pp.271-282

Roßmann, B., Canzaniello, A., von der Gracht, H., & Hartmann, E. (2017). The future and social impact of Big Data Analytics in Supply Chain Management: Results from a Delphi study. Technological Forecasting and Social Change.

Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited,.

Safdar. K., Stevens. L. (2018) Banned from Amazon, the Shoppers who make too many returns. Retrieved from: https://www.wsj.com/articles/banned-from-amazon-the-shoppers-who-make-too-many-returns-1526981401

Safeir. (2017) Retrieved from: https://medium.com/ai-society/my-first-experience-with-deep-reinforcement-learning-1743594f0361

Salden, J. (2016) De Zalando Zeepbel. Retrieved from: https://eenvandaag.atavist.com/webwinkels

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. Journal of Business Logistics, 36(1), 120-132.

Sennaar, K., (2018) Artificial Intelligence in eCommerce – Comparing the Top 5 Largest Firms. Retrieved from: https://www.techemergence.com/artificial-intelligence-in-ecommerce-amazon-alibaba-jd-com/

Shaharudin, M. R., Govindan, K., Zailani, S., Tan, K. C., & Iranmanesh, M. (2017). Product return management: Linking product returns, closed-loop supply chain activities and the effectiveness of the reverse supply chains. Journal of cleaner production, 149, 1144-1156.

Shaik, M., & Abdul-Kader, W. (2012). Performance measurement of reverse logistics enterprise: a comprehensive and integrated approach. Measuring Business Excellence, 16(2), 23-34.

Silver, B. (2011). BPMN Method and Style, with BPMN Implementer's Guide: A structured approach for business process modeling and implementation using BPMN 2.0 (p. 23). Aptos: Cody-Cassidy Press.

Solarz, T., (2016) De groei van internet retailer Bol.com is spectaculair maar kloppen de cijfers wel? Retrieved from: http://tadeksolarz.blogspot.co.uk/2016/06/de-groei-van-bolcom-is-spectaculair.html

Spector, L. (2006). Evolution of artificial intelligence. Artificial Intelligence, 170(18), 1251-1253.

Statista.com. (2017). B2B e-commerce volume in the United States from 2006 to 2015 (in billion U.S. dollars) https://www.statista.com/statistics/239366/us-b2b-e-commerce-volume-since-2006/

Sutton, R. S., & Barto, A. G. (1998). Introduction to reinforcement learning (Vol. 135). Cambridge: MIT press.

Suwelack, T., Hogreve, J., & Hoyer, W. D. (2011). Understanding money-back guarantees: Cognitive, affective, and behavioral outcomes. Journal of Retailing, 87(4), 462-478.

Terry, L. (2014) Managing Retail Returns, the good the bad the ugly. http://www.inboundlogistics.com/cms/article/managing-retail-returns-the-good-the-bad-and-the-ugly/#sidebar1

Topno, S. (2014) A presentation on Supply chain management. Retrieved from: https://www.slideshare.net/topnovineet/a-presentaion-on-suppy-chain-management

Urbanke, P., Kranz, J., & Kolbe, L. (2015). Predicting Product Returns in E-Commerce: The Contribution of Mahalanobis Feature Extraction.

Veitch, M. (2016) Zalando relies on technology to aid fashionistas. Retrieved from: https://www.idgconnect.com/abstract/21150/zalando-relies-technology-aid-fashionistas

Viswarupan, N. (2017) K-means clustering. Retrieved from: https://towardsdatascience.com/k-means-data-clustering-bce3335d2203

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Production Economics, 176, 98-110.

Welch, C. (2018) Google just gave a stunning demo of Assistant making an actual phone call. The verge. Retrieved from: https://www.theverge.com/2018/5/8/17332070/google-assistant-makes-phone-call-demo-duplex-io-2018

Wellens, A,. (2017) Wehkamp woekert klanten stuk met enorme rente, maakt nog steeds verlies. Retrieved from: http://925.nl/archief/2017/08/01/wehkamp-woekert-klanten-stuk-met-enorme-rente-maakt-nog-steeds-verlies/

Williamson, O. E. (1975). Markets and hierarchies : analysis and antitrust implications : a study in the economics of internal organization. New York: The Free Press.

Williamson, O. E. (1979). Transaction-cost economics: the governance of contractual relations. The journal of Law and Economics, 22(2), 233-261.

Williamson, O. E. (1985). The economic intstitutions of capitalism: Simon and Schuster.

Williamson, O. E. (1998). Transaction cost economics: How it works; where it is headed. Economist, 146(1), 23-58.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

Wood, S. L. 2001. "Remote purchase environments: the influence of return policy leniency on two-stagedecision processes," Journal of Marketing Research (38:2), pp. 157–169

Wu. S. (2015) US B2B eCommerce Will Be Twice The Size Of B2C eCommerce By 2020. Retrieved from: https://www.forrester.com/report/US+B2B+eCommerce+Will+Be+Twice+The+Size+Of+B2C+eCommerce+By+2020/-/E-RES122366

Xing, B., Gao, W. J., Battle, K., Marwala, T., & Nelwamondo, F. V. (2010). Artificial intelligence in reverse supply chain management: the state of the art. arXiv preprint arXiv:1012.4046.

Zhu, J., Hans, M., & Fredershausen, A. C. (2016). What Drives Product Return Behavior in E-Commerce? The Moderating Effect of Cultural Values. In Academy of Management Proceedings (Vol. 2016, No. 1, p. 12657). Briarcliff Manor, NY 10510: Academy of Management.

10 Appendices

In the appendices all additional information of this thesis is attached. In appendix A, the methodology and tools used in this thesis is described in detail. In Appendix B, an extensive summary of the conducted interview is presented. This interview gives valuable information that is used in this research.

Appendix A Methodology and Tools

In this section the methodology that will be used in this research project is explained. Besides that, the tools are defined that can be of use in this research to analyze the gathered information and data. To answer the sub questions different methods or tools can be used to complement the project. This all will be displayed in the schematic overview of table 5.

Sub Question	Method	Tools
1. What are the current components of a return management process from an e-commerce player in the Dutch electronics market?	Interviews with responsible persons at company, case study, Business Process Modeling	To create the visualization, a business process tool will be used to sketch the process.
2. What is Supervised Learning, Unsupervised Learning and Reinforcement Learning?	Literature review and Desk research	
3. What are the possible applications in Return Management for Machine Learning	Interviews with responsible persons at company, case study, Business Process Modeling	
4. On which parts of the Return process can Supervised Learning, Unsupervised Learning or/and Reinforcement Learning be applied?	Desk research, Case study analysis.	

5. What can be the applicability of Supervised Learning, Unsupervised Learning or Reinforcement Learning on the defined Key Performance Indicators of the Return Process?	Case study analysis, Scenario analysis.	

Table 6, Methods & Tools used

In figure 28, the project planner for the master thesis is presented.

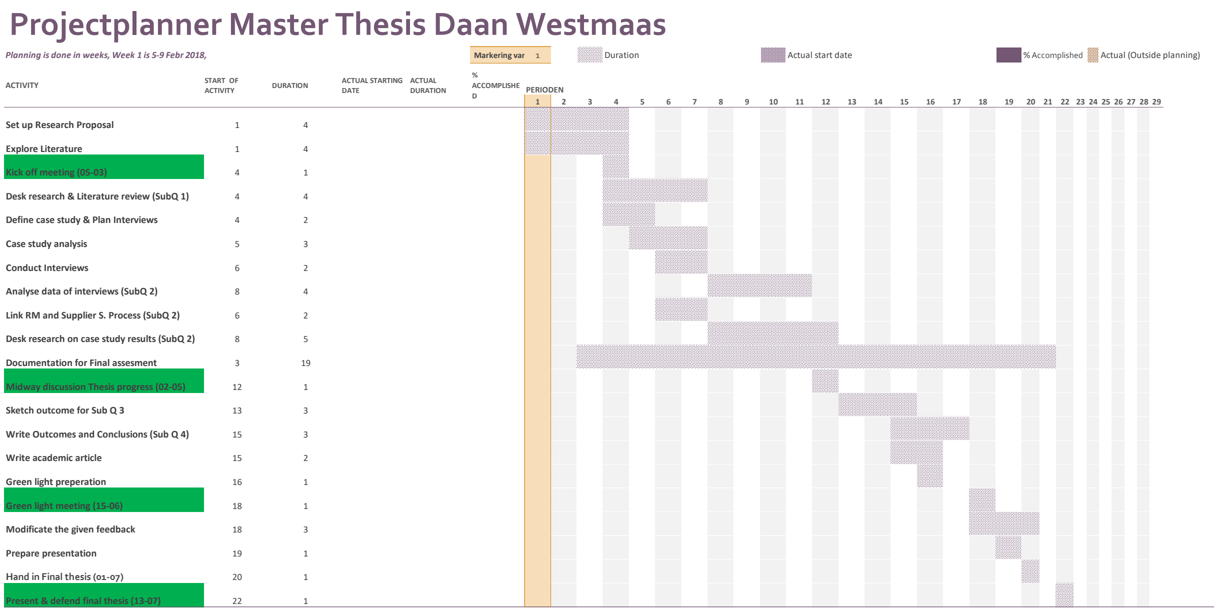


Figure 28, Project planner Master thesis

Appendix B Interview conducted (Not for public use)

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix C Background Information ML

SL application

Scientific papers and journals are describing the different possibilities and opportunities of SL. Recently, many scientists and companies are investing time and money in the exploration of SL techniques (Malhotra, 2015). Where the usage of SL in RM is rather unexplored, other fields in the SC are broadly discovered and can contribute in the outcome for this project. Due to the current extensive software tools and the knowledge available in literature, the translation from literature into tangible and applicable business cases is easier in SL compared to the other ML techniques. As described in literature, SL has a huge potential in many fields. Not only in SCM, also in healthcare, finance and entertainment, SL can have an impact and literature is exploring the possibilities. New methods and techniques in SL are being explored to further optimize and improve business processes, decision making and complex dynamic situations. The potential related to RM and supplier selection of SL will be discussed in the next paragraph.

RFL Meaning:

RFL is becoming an increasingly popular machine learning technique that is very suitable for addressing problems in a dynamic and adaptive environment (Hammond, 2017). In fact, it is a science of decision making, being rewarded for each decision the agent will be led to an optimal score. It is a way of unsupervised learning, no one tells the agent what the right action is. It receives rewards that can have different value. The agent interacts in a dynamic environment, it is not the same as a classical ML technique where a model learns on a static training set. In RFL the agent influence and affect the data that it sees. Therefore, the chronological order of time really matters. An example in the field of RFL is the management of an investment portfolio. Taking certain actions will be rewarded with money on the bank and the more money on the bank the better the agent is performing. Each reward is measured in dollars and will lead to an increase or decrease on the account balance. There is no control over the environment, however the agent influences the environment by the actions it takes. This makes it attractive to implement in real life situation, where the agent will be able to interact with the dynamic environment. Important to know that the objective of the agent is to maximize the expected reward over time, sometimes a step that leads to a less higher score can eventually lead to the highest score. Therefore, the overarching goal will be more important than the next step that the agent must take. In summary, the agent tries to find a policy that will lead him to the highest reward. (Macua, 2015) In RFL the so-called Deep Reinforcement Learning (DRL) models are using Neural Networks to optimize the behavior of their agent. One of the common known examples of this work is the algorithm that can play various Atari games on human level, with one single RFL algorithm. (Mnuh et al, 2013)

RFL Pitfalls:

Structure the problem in a curriculum. This seems natural when dealing with persons, so it is with Reinforcement Learning. Step by step learning will go much faster. Since you constrain the explore environment of the agent. It has less possible decisions to perceive the final state that you are after. The training that is conducted, should map the real world. It is a lot easier to build simulations that are not physically realizable. It does not need to have full fidelity, however it is important to have a solid representation of reality.

When training with Neural Networks in RL, the model must be very precise to come up with some meaningful results (Sefair, 2017). It can not be trained in a couple of hours, it will take long processing times to perform. Choosing the parameters, topology and activation functions can be very challenging and important. This can be the essence between a good performance and perform nothing at all. This complexity of settings and the long run times can be frustrating and will make it hard to apply RL.

RFL application

One example of the appliance of RfL in the supply chain is at Google. The number two of expenses at Google is electricity, second after employees' wages. Optimizing these electricity grids in their data centers can save them huge amounts of money. Reinforcement Learning techniques were used by Google to accomplish these power savings and make a more efficient use of their datacenters (Hammond, 2017). This can be applied in broader sense, manufacturing companies and warehouses are dealing with high electricity usage. When this can be brought down with RfL algorithms this can save money. To better manage the SC, new methodologies have been designed to improve results. Just in Time policies, lean production and ERP systems are integrated in the supply chains to have a better control and management of the activities. Additionally, on all these methods, RfL can help in improving the way of working and help in being lean and efficient when decisions are made (Gunasekaran et al, 2004).

Appendix D Background material from chapter 6

6.1.1

This demanding culture is partly created by the e-commerce parties themselves, where they advertise with high promising slogans that create intangible customers. For instance, Wehkamp is openly advertising about the shift of the fitting room from the physical shop towards every customers living room. When customers fit their clothes at home because they purchased it online, there is a chance of a misfit with unavoidable returns as an effect. New applications and technologies are created to meet the high demands of the customers. RFID chips are used to keep track of your order, these technologies anticipated on the changing demands in customer needs. The technologies are examples of new innovations that are continuously needed to keep serving the customers at high level.

In this study, there was no access retained to real customer data in the return process. However, to make a real-world representation a scenario will be sketched to show the applicability of Machine Learning. Current e-commerce companies are using all kind of data to analyze their customers. Facebook is used to do personalized advertisements and the data from the website is used to sketch customers profiles. In this study there is not complete knowledge of the available data in the process of return. However, when a customer returned a product in the past, this customer can be classified as a returner. Using data from previous purchases, social media data from Facebook or other platforms can link more personalized data to the customer. This can be geographical location, sex, estimation of age, product interest and so on. By combining this data there can be sketched a relatively representative image of the customer. This will help in serving the customer in both the return as well as in the forward process.

This way of working is a simple way of using Supervised Learning, the algorithm detects certain behavior from the customer and can classify this behavior. This interactive way of working can prevent customer service contact moments and will reduce their contact with the clients in the return process. It is hard to explicitly transfer these tasks into numbers. It can reduce the customer contact, however when the algorithm provides vague and incorrect information the customers will eventually call the customer service even more. Before the implementation of such a tool, e-commerce parties must first train and test the algorithm and make sure it is implemented correctly. Due to the learning behavior of the algorithm it will improve over time and can have the potential to serve even more customers.

6.2.3.

Several customers of Amazon had huge spending’s in their years as a customer. Through variety of reasons they got banned by Amazon, for instance a customer decided to buy some more cheap products on Amazon because they had a good warranty policy. Unfortunately, many orders were damaged on arrival and this customer send those products back. Due to the high number of orders and returns this customer was removed from Amazon. However, this customer did not try to trick Amazon and was simply forced by the bad quality of the products to send them back.

Appendix E Background information chapter 7

Bol.com, one of the biggest e-commerce players in The Netherlands had a total net profit of 2.4 billion euro over the year 2017. There are question marks on these numbers, probably they include the sending costs that the customer is paying. However, bol.com does not charge sending costs at orders above 20 euros because they offer the transport for free (Solarz, 2017). Having a 9% return rate (market average) means that a total product value of approximately 220 € millions is returned. However, this number can even be higher due to the higher return rate on expensive products (Minnema, 2017). Bol.com has a potential market of at least 220 € million to improve. This is a significant amount of money and improvements in this area can have impacts on business performance. The 220 € Million does not includes the costs of transport, personnel and processes. Including these costs probably the total scope of return goes over 300 € million.