Matti Vahermo

Exploring Machine Learning to Improve Procurement and Purchasing Processes

School of Technology and Innovation
Master of Science in Economics and
Business Administration
Information systems

Vaasa 2023

Vaasan Yliopisto
School of Technology and Innovation

Tekijä:	Matti Vahermo
Tutkielman nimi:	Exploring Machine Learning to Improve Procurement and Purchasing Processes
Tutkinto:	Kauppatieteiden maisteri
Oppiaine:	Information systems
Työn ohjaaja:	Tomi Pasanen
Valmistumisvuosi:	2023 **Sivumäärä:** 85

TIIVISTELMÄ:

Koneoppimisella tarkoitetaan tekoälyn osa-aluetta, joka mahdollistaa järjestelmien suorituskyvyn parantamisen oppimalla datasta ilman, että sitä on tarkoituksenmukaisesti ohjelmoitu kyseisestä tehtävää varten. Oppiminen tapahtuu kouluttamalla algoritmeja tunnistamaan suurista tietomääristä korrelaatioita ja malleja, joiden perusteella pystytään luoman ennusteita sekä tekemään johtopäätöksiä. Koneoppiminen on kasvattanut viime vuosina suosiotaan erilaisten kaupallisten sovelluskohteiden muodossa, eikä myöskään hankinta- ja ostoprosessit ole tässä asissa poikkeus. Jatkuvasti parantuva tietokoneiden laskentakyky ja tiedonhallinta ovat mahdollistaneet entistä kehittyneempiä koneoppimista hyödyntäviä sovelluksia, mikä on myös laajentanut koneoppimisen ympärillä tapahtuvaa tutkimustyötä. Kuitenkin kun tarkastellaan tutkimuksia, joissa käsitellään hankinta - ja ostoprosessien kehittämistä koneoppimisen avulla on julkaisumäärä rajallista, erityisesti sellaisten tutkimusten osalta, joissa on hyödynnetty kokemusperäistä, haastatteluista saatua tietoa.

Täten tämän työn tavoitteena oli tarjota ajankohtainen katsaus koneoppimista hyödyntävien sovellusten tarjoamista mahdollisuuksista ja potentiaalisista haasteista niitä hankinta ja osto prosseihin käyttöönotettaessa. Lisäksi suoritettiin haastatteluita, joiden tavoitteena oli saada selville syyt, jotka ovat esteinä tehokkaan hankinnan- ja ostoprosessien tapahtumiselle sekä mihin työtehtäviin haastateltavat toivoisivat erityisesti apua tietojärjestelmien kautta ja voisiko koneoppiminen tarjota apua koettuihin ongelmiin. Työn tutkimusmetodit olivat teoreettisia sekä empiirisiä. Teoreettinen osio koostui koneoppimisen sekä hankinnta- ja ostoprosessien eri osa-aluiden esittelystä, joiden lähdemateriaalina hyödynnettin saatavilla olevia akateemisia sekä koneoppimisen ja hankinnan ja oston alojen julkaisuja. Lähdemateriaali pyrittiin pitämään mahdollisimman ajankohtaisena. Haastattelut suoritettiin yhden yrityksen kanssa, johon otti osaa kahdeksan henkilöä.

Tutkimuksen tulosten perusteella koneoppimisen sovellukset, jotka auttavat hinnoittelussa, kustannusten analysoinnissa sekä materiaalitarpeiden ennustamisessa nähtiin erityisen hyödyllisinä. Haasteina koettin ongelmat, jotka johtuivat erityisesti käytetyn datan heikosta laadusta, datan suuresta määrästä sekä datan jäljitettävyydestä. Haastatteluiden perusteella haasteina koettiin prosessien vähäinen automatisointi, datan luotettavuussongelmat, materiaalitarpeiden ennustaminen sekä materiaalien hinnoitteluun ja analysointiin liittyvät haasteet. Tuloksista voitiin tulkita, että hankinnan ja oston prosesseja voidaan kehittää koneoppimisen avulla, ja empiirinen tutkimusosio myös tukee tätä johtopäätöstä.

Avainsanat: Koneoppiminen; hankinta; ostaminen; algoritmi; koneoppimismallit.

UNIVERSITY OF VAASA
School of Technology and Innovation

Author:	Matti Vahermo
Title of the Thesis:	Exploring Machine Learning to Improve Procurement and Purchasing Processes
Degree:	Master of Science in Economics and Business Administration
Discipline:	Information systems
Supervisor:	Tomi Pasanen
Year:	2023 **Pages:** 85

ABSTRACT:

Machine learning is an area of artificial intelligence that enables systems to improve their performance by learning from data without being purposefully programmed for the task. Learning occurs by training algorithms to identify correlations and patterns in large amounts of data, which can be then utilized to make predictions and conclusions. Machine learning has grown in popularity in recent years for a variety of commercial applications, and the purchasing process is no exception. The continuously improving computing power and data management capabilities of computers have enabled more sophisticated machine learning applications, which has also expanded the research work around machine learning. However, when looking at studies on the development of purchasing and purchasing processes using machine learning applications, the number of publications is limited, especially for studies that have used empirical data from interviews.

Thus, the aim of this research was to provide a current overview of the opportunities and potential challenges of implementing machine learning applications in procurement and purchasing processes. In addition, interviews were conducted with the aim of finding out the reasons that are preventing efficient procurement and purchasing processes, the work tasks that interviewees would most like to see assist from information systems, and whether machine learning could offer help with perceived problems. The research methods used in this thesis were both theoretical and empirical. The theoretical part consisted of an introduction to different aspects of machine learning and procurement and purchasing processes, using available academic and industry material as sources. The aim was to keep the source material as up to date as possible. Interviews were conducted with one company, involving eight participants in total.

Based on the results of the research, machine learning applications that provide assist with pricing, cost analysis and forecasting of material requirements were seen especially useful. Challenges were perceived, in particular due to the poor quality of the used data, the large amount of data and the traceability of the data. Based on the interviews the low level of automatization of processes, data reliability problems, forecasting material requirements, and pricing and analysing of materials were seen as challenges. The results suggested that machine learning can be used to improve purchasing and procurement processes, and the empirical research supports this conclusion.

KEYWORDS: Machine learning; Purchasing; Procurement; Algorithm; Machine learning models.

Contents

Figures

Tables

1 Introduction

This chapter will be introducing the study by first discussing the background and context of the topic, then moving on to discussing the research problem, followed by research aims, objectives and questions, the significance of the research, and lastly, the introduction of the structural outline of this study.

The core idea of machine learning is to create computer applications that will learn and improve their performance through experience. Over the recent two decades machine learning has spread widely across commercial use in different fields and industries such as manufacturing, logistics, health care, finance, and marketing. The main enablers for this development have been new learning algorithms and progress in the theory itself. In addition, the cost of computing has decreased lately and the rapidly increased data which is accessible online have been as well important factors. Therefore, machine learning has been rising its popularity among professionals who are trying to receive valuable predictions, insights, and support for decision-making from large quantities of data. In the area of artificial intelligence, machine learning plays a significant role and has provided applications for instance to speech recognition, processing of natural language, and robotic control. (Jordan & Mitchell, 2015.)

Procuring and purchasing the required goods for manufacturing or services are one the most crucial aspects of the supply chain process and have a significant impact on how business can meet their customer commitments. During the past decades, these fields have also gone through major digitalization, and this has affected the processes and the way of working in many organizations. Digitalization has turned from an option to a necessity in recent years as it has increased and tightened the competition between different companies (Viale & Zouari, 2020). The recent rise of machine learning technologies and the digitalization of purchasing and procurement have created an interest in solutions by companies and software providers. For instance, one of the leading Enterprise Resource Planning (ERP) software providers SAP has already released its solutions for purchasing and procurement processes that utilize machine learning techniques (SAP Products & Services, 2018a). Besides the benefits of digitalization, purchasing and procurement operations face problems such as new technology adaption, data quality, pricing, and spend analytics, as well as supplier relationship management. In general machine learning has been mentioned to be one of the technologies that will drive supply chain organizations towards more agile, cost effective, flexible, and dependable. (Ageron et al., 2020.)

As machine learning solutions are rapidly developing and purchasing and procurement processes facing further digitalization it can be seen that there is a lack of research to find current challenges of purchasing and procurement process and how to counter these with machine learning solutions, especially in a way that includes surveys or interviews conducted with actual professionals of purchasing and procurement. Even though every purchasing and procurement organization has its individual ways of working and processes, there is a lack of research that would provide insights about their current challenges. The current research is insufficient in terms of collecting valuable feedback from actual professionals who are working in the fields of purchasing and procurement. Without receiving firsthand information about what is hindering the purchasing and procurement of being efficient it is difficult to conclude if machine learning offers solutions to help with the issues. Furthermore, businesses may find themselves in a situation where they believe machine learning will assist them in solving problems, but the problems may be created by issues that are not optimal for machine learning solutions to tackle, and so resources may be allocated inappropriately. As a result, the purpose of this study is to address that gap and provide insights that firms can use to determine whether machine learning technologies can fit into their procurement and purchasing processes.

For the above-mentioned reasons, this study will aim to identify and discuss the potential of implementing machine learning techniques in purchasing and procurement organizations. Therefore, the objectives of this study are to first identify the current advantages and challenges of implementation of machine learning techniques into purchasing and procurement process. Then identify the issues which are preventing effective purchasing and procurement processes to occur and lastly compare if these issues could be solved or reduced with the help of machine learning techniques. Based on the revealed objectives this study contains three research questions:

1. *What are the current advantages of implementing machine learning techniques in purchasing and procurement processes?*

2. *What are the current challenges of implementing machine learning techniques in purchasing and procurement processes?*

3. *Which issues prevent effective purchasing and procurement and how machine learning techniques could possibly help with the found issues?*

According to Jordan and Mitchell (2015), machine learning is still a relatively new subject for research, and its applications in procurement are constantly evolving. Therefore, this study will contribute by providing knowledge on the present status of machine learning solutions that have been or will be deployed in procurement and purchasing. In addition, this study will provide information about possible machine learning use cases as well as general information about the technology to industry professionals. As a result, the information offered by this study could be valuable for businesses determining whether to invest in machine learning solutions. Finally, this study contains interviews with procurement and purchasing professionals, illustrating the types of difficulties these organizations face and how machine learning can help with them.

The next chapter discusses past relevant publications which have been conducted within the context of machine learning and its connection to both purchasing and procurement. In chapter three, machine learning as a technology is introduced starting from its different definitions, followed by its different types and algorithms. In chapter four, the core process of purchasing and procurement are presented, and their fundamental differences are explained. In chapter five, the research methodology used in this study is introduced and different choices of how this research was conducted are explained.

In chapter six, results based on the interviews are unveiled. In chapter seven, the results are discussed and interpreted. In chapter eight, the study is concluded with a brief discussion of what has been achieved and what are the recommendations for future studies.

2 Literature review

The purpose of this chapter is to introduce existing research on the selected topic. This will provide an outlook on where the research currently stands. The first part of this chapter focuses on the existing and predicted use cases of machine learning techniques in the area of purchasing and procurement. The second part highlights the state of research regarding the current challenges in purchasing and procurement. Lastly, a brief conclusion is provided.

Procurement processes are seen as one of the fields which can benefit from machine learning applications. Yamusa et al. (2020) proposed a framework for the development of machine learning models for procurement processes and explained how machine learning involvement in procurement is reducing human error and increasing transparency in the processes. Rejeb et al. (2018) discuss how machine learning is having a significant impact on procurement processes in the field of supply chain management alongside with artificial intelligence and robotics. Learning algorithms provide benefits by processing efficiently unstructured documents, such as contracts, bill of materials, and specification drawings, and recognizing if errors occur, for instance with payment terms. Correspondingly McCrea (2019) argues that procurement processes include various repetitive and manually performed tasks that could benefit from machine learning solutions. Kiefer et al. (2019) conducted research with the REIF group to compare procurement forecasting with a multilayer perceptron (MLF) machine learning algorithm. Results indicated that the algorithm was able to predict similar trends or occasionally more accurate predictions of future demands, and stated that artificial intelligence, such as machine learning has a justified role in forecasting processes. Bajari et al. (2008) found that machine learning algorithms provide significant accuracy in demand prediction in comparison to traditional regression and logic models. Also, Carbonneau et al. (2008) found out through their study that machine learning techniques were able to provide more precise demand forecasts compared to more traditional techniques such as trends and moving average. In turn, Priore et al. (2019) discovered that an inductive machine learning algorithm was able to correct the inventory replenishment rule an average of 88% of the time in a wholesale simulation model.

In recent years, research on supplier selection and how it might be upgraded using new technology has been popular in the procurement industry. Especially, the usage of support vector machine algorithms and neural network models has been gaining interest (Bousqaoui et al., 2017; Kıran et al., 2020). Lau et al. (2001) developed an intelligent decision system for business partner selection by utilizing neural network and argue that quick adaptation ability for most developed and state-of-the-art technologies is an important aspect of building stronger collaboration with business partners, both for suppliers and customers. Mori et al. (2012) utilized a support vector machine–algorithm to find potential business partners and the results were showing positive signs of real-life usage. Aksoy & Öztürk (2011) propose a neural network-based supplier selection and performance evaluation approach, that can provide help to Just in time (JIT) manufacturers. Bodendorf et al. (2021) state that cost estimation is an area where machine learning has the potential to reduce workload significantly. Enríquez García & Roque Huerta (2020) conducted research on machine learning and deep learning applications and discovered that there is a potential for cost reduction and for various industrial and business sectors to gain a competitive advantage.

Borges et al. (2009) tell that for the successful deployment of information system tools, it is important to understand where in the processes they can be used strategically in order to provide value for the organization. Allal-Chérif et al. (2021) state that although purchasing organizations together with suppliers and partners are creating great quantities of data the potential of it is very rarely fully utilized. One of the key reasons for poor analytical utilization of both internal and external data is the lack of required skills, as well as the use of inappropriate tools. (Mikalef et al., 2018). Also, opposition against change is mentioned as an obstacle in digitalization and artificial intelligence-related projects within procurement organizations (Allal-Chérif et al., 2021). Tanner et al. (2008) explain that the high cost of new information technology solutions is a challenge in procurement and therefore slowdowns the development of the process in organizations. In addition, Tanner et al. state that expectations of new information technology implementations in procurement are rarely met and do not correspond to the initial plans. Although concurrently, Dryden Group (2022) explain that outdated technology and ineffective process are mentioned to be one of the challenges of purchasing and procurement. In addition, data inaccuracy plays an influential role in the challenges of purchasing and procurement as inaccurate data will also lead to inaccurate decisions which could potentially lead to financial losses.

Unstable market situations and events that have caused major disturbances in supply chains such as the COVID-19 pandemic are creating pressure on the material availability and costs of raw materials. In these scenarios, suppliers are often moving the raw materials costs into the product prices which causes challenges in procurement organizations. To handle the situation of increased costs, procurement organizations are forced to optimize their process in order to purchase and spend more effectively than before (Sheinfield & Forman, 2022). Correspondingly, ProcurePort (2022) sees managing inflation and its impacts on price levels as one of the key challenges of maintaining competitiveness.

As a conclusion from the literature review, we can notice that machine learning has already several use cases regarding purchasing and procurement and the scene is likely to develop even further and the mentioned potential use cases might turn into reality. Current challenges of purchasing and procurement appear to be mostly around data utilization and its quality, outdated information systems, and unstable market situation which tend to lead to pricing issues. The lack of research where actual professionals from purchasing or procurement are interviewed was distinct and this strengthens the necessity of this study. Dubois & Araujo (2007) recommend combining theoretical and practical methods in purchasing and supply management research, which this study has done as well.

3 Machine learning

This chapter provides information about different aspects of machine learning. First, it discusses machine learning as a definition. Secondly, the most recognized types of machine learning are introduced followed by a brief introduction to the core ideas of some algorithms. This study consciously does not display formulas of algorithms but instead illustrates with the use of figures models' ideas which should be comprehended effortlessly even without prior knowledge about machine learning.

3.1 Machine learning definition

Machine learning, like many other technologies, has various definitions based on its use cases; thus, it is useful to introduce some of them in order to provide a solid impression that helps in understanding the concept behind machine learning.

Hurwitz & Kirsch (2018) describe machine learning as a technique that utilizes algorithms that are designed to learn from the given data set, that then turns it into representative form and generates predictions for the user. Mitchell (1997) defines machine learning as computer algorithms that are able to improve on their own automatically based on experience and by the utilization of the given data set. Correspondingly Bourke (2020) states, fairly distinctly, that machine learning is about turning data into numbers and finding patterns in those numbers which then provide hopefully useful information for the end user. Turning data into numbers refers in this case to the machine learning algorithms which are using numbers in order to function and provide outputs. According to Samuel (1967), machine learning is a field of study that gives computers the ability to learn without being explicitly programmed to do so. Hao (2018a) states that machine learning algorithms search through data for patterns using statistics and the dataset is usually massive by size and can include for instance words, images, and numbers. This definition also brings up the fundamental origins of machine learning which are combination of mathematics and statistics. In a conclusion machine learning enables computer systems to accomplish tasks without being precisely programmed to do so while continuously learning through the received input data and also in some cases making independent decisions based on the results.

3.2 Machine learning and artificial intelligence

Often machine learning is seen as a subset of artificial intelligence as Figure 1 shows. Deep learning is correspondingly seen as a subset of machine learning and it will be not discussed in close detail in this study. Gavrilova (2020) states that artificial intelligence is a science that focuses on studying different ways to create intelligent computer programs that solve problems by using their creativity. This ability has been previously considered a privilege only for humans. As Opperman (2022) underlines, artificial intelligence programs are able to sense, reason, act and adapt based on the data. Thus, the goal of artificial intelligence is to mimic human intelligence, whereas machine learning focuses solely on training programs to learn from data. Furthermore, machine learning is currently the most popular method of utilizing artificial intelligence, and it is responsible for the majority of its achievements in terms of useful applications. (Hao, 2018b.)

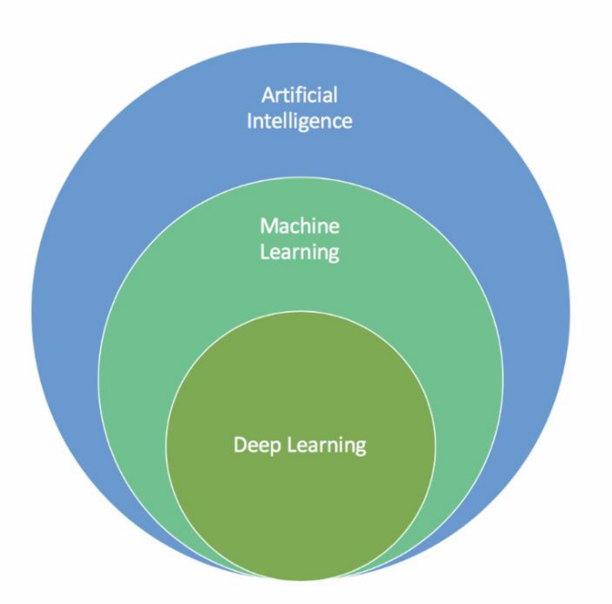

Figure 1. Artificial Intelligence, Machine Learning, Deep Learning (Wu, 2019).

Nevertheless, some questions have been raised about the relationship between artificial intelligence and machine learning. According to Schmelzer (2021), some machine learning programs are closer to automation and should not be considered to perform anything intelligently demanding. In addition, in Walch's & Schemlzer's interview (2018) Perez-Breva argues that many of the current machine learning applications are in fact statistical exercises and analytical applications that use mathematical algorithms without any demanding aspect related to artificial intelligence. Therefore, as a reader and researcher, it is critical to be wary of how the terms artificial intelligence and machine learning are combined, as they may be used as a buzzword to draw attention to their marketed applications.

3.3 Machine learning types

This chapter goes through four main types of machine learning which are currently supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. There is no official determination of the above-mentioned types, though the following three ones are usually the most mentioned and used as the main types in the machine learning literature (Bourke, 2020). In addition, examples of relevant real-life use cases are introduced that occur in different types of learning. Later in chapter seven machine learning algorithms are introduced and their functions and example use cases are generally explained.

3.3.1 Supervised learning

In recent years supervised learning has been the most researched area of machine learning. Supervised learning requires labelled datasets which are then implemented into the trained algorithms which are most often classifying or predicting outcomes from the data. As the user has the knowledge of the data and the labels the selected machine learning algorithm is attempting to learn the relationships between the data and its labels. This process is happening during the training phase where the user is “supervising” in order to ensure that the program is behaving as intended. (Bourke, 2020.) This supervising is possible because the user already has knowledge about the data and can identify whether the outcome of the model is matching with the details of the data labels thus the user is able to judge whether the result is coherent based on the data that has been input in (Hurwitz & Kirsch, 2018). Similarly, if any input values are missing from the model, it is practically impossible to conclude and use the results for anything relevant. Often probabilities for inputs are not defined, which is acceptable if all inputs are available for the model. In contrast, in unsupervised learning, which will be discussed in greater detail in the following chapter, latent variables are assumed to be the bases for all observations, implying that observations are at the end of the chain of causation. (Oladipupo, 2010.) Figure 2 shows the main difference between supervised learning and unsupervised learning models. Supervised learning inputs are known for the user but in unsupervised learning the input variables are latent, referring to random variables which the user cannot observe before or during the testing phase.

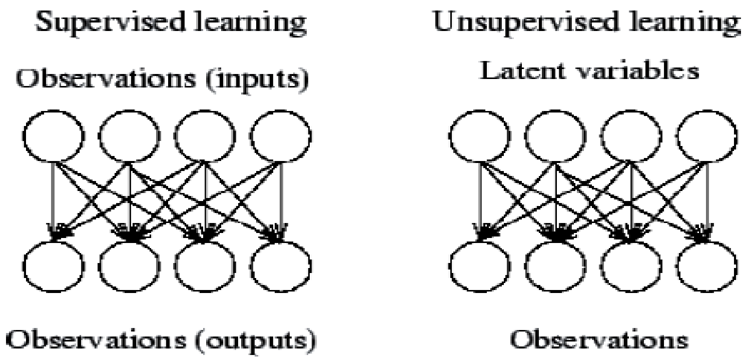

Figure 2. Examples of Supervised and Unsupervised Learning (Ayodele, 2010).

Regression is used when the data that is getting inputted to the model is continuous. Regression provides knowledge about the correlation that exists between two or more variables. For example, in weather forecasting historical weather conditions and patterns are used together with the current weather conditions, thus providing a prediction of what future weather conditions might end up occurring. (Hurwitz & Kirsch, 2018). Correspondingly when data is originally from a limited set classification method is used. Instead of regression, the model is trying to learn the relationship between the data and its labels. A categorization challenge might be one in which the user has pictures of dogs and cats, and the label indicates which photo contains which animal. And therefore, during the training process, the model is trained to understand which photo is a cat and which is a dog based on the label on the picture, and thus in the future, it can detect the correct animal based on the training session. (Bourke, 2020.)

3.3.2 Unsupervised learning

Unsupervised approaches, as opposed to supervised learning, are best suited when the tasks include a large amount of data but no labeling. It is implemented into an algorithm that seeks to classify the data in order to gain a knowledge of the patterns and connections within it. Thus, in contrast to supervised learning, there are no correct solutions, and no prior teaching is required (Brownlee, 2016). Unsupervised learning is an excellent way for organizations to understand patterns in huge amounts of unlabeled data that would otherwise be exceedingly difficult to analyse without the assistance of machine learning techniques. As a result, when compared to manually conducted procedures, unsupervised learning gives a more efficient way to examine data. (Hurwitz & Kirsch, 2018; Delua, 2021). For instance, various social media companies are using unsupervised learning methods as their platforms consists of innumerable amount of unlabeled data. Thus, unsupervised algorithms try to understand clusters and patterns within the data and are hopefully able to classify the data successfully. Facebook, Instagram, and Twitter are examples of applications that are using unsupervised learning algorithms to provide accurate content and other recommendations based on their users' prior actions and given data on their platforms

Unsupervised learning methods can be divided into the following two main categories: clustering and association. The clustering approach is used to mine enormous amounts of unlabeled and raw data, which are then classified into groups based on patterns and structural differences, if they exist. Figure 3 illustrates how clustering separates unstructured data into certain groups based on their similar characteristics.

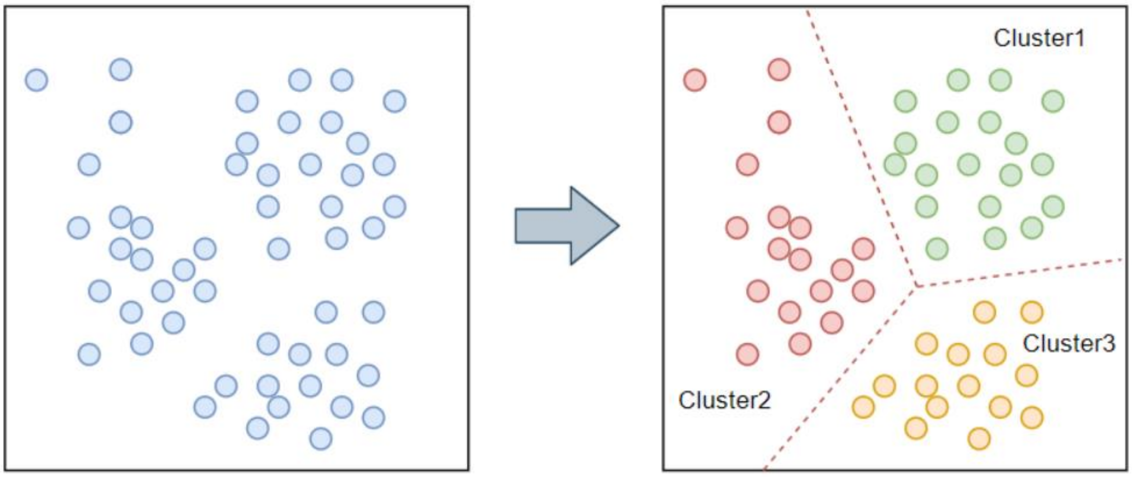

Figure 3. Clustering (Chen, 2021).

Association method is designed to search for interesting relationships which occur between different variables from a massive data set. In practice, the association algorithm is counting frequencies of different occurrences and it is trying to find if any underlying correlations exist between them. Simply the unsupervised model is trying to find out if A happens is B happening too. (Brownlee, 2016a.) Figure 4 illustrates the core rule about the association method and how it shares the same concept as IF - Else clause in coding. For instance, retail stores are trying to find relationships about which products are more likely to be bought together, and how this finding can be utilized in marketing and product placement inside the store. (DeepAI, 2019.)

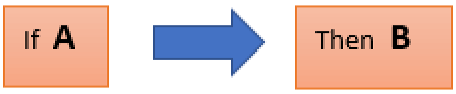

Figure 4. Association rule learning (JavaTpoint, n.d.-a).

3.3.3 Semi-supervised learning

Sometimes problems are between supervised and unsupervised learning which occur when there is a small amount of labelled data and a large amount of unlabelled. Often in real – life problems belong to this category. Zhu & Goldberg (2009) explain that semi-supervised learning can utilize available unlabelled data to enhance supervised learning in its tasks in cases where the available labelled data is expensive which makes it a useful addition to machine learning styles. Thus, the goal of semi-supervised learning is to understand how learning behaviour changes when combining both labelled and unlabelled data (Zhu & Goldberg, 2009). Brownlee (2016a) tells that often real-world problems require semi–supervised learning as they are implemented for making best guesses and predictions. For example, suppose a dataset contains photos, but only some of them are labeled, such as automobile, bike, or bus, and the model is attempting to determine which category each unlabeled photo belongs to.

3.3.4 Reinforcement learning

Reinforcement learning is mentioned to be most similar to how humans learn as it interacts with its environment and receives either positive or negative feedback and learns through these experiences (Coursera, 2022). Sutton & Barto (2014) explain how this kind of learning model is not told what to do but instead the goal is to try out all actions and find out which of them leads to the highest reward. Thus, the model is called behavioral machine learning because all its results are outcomes of “reinforcement” processes where the model has developed itself the best possible solution for the problem by trial and error. Therefore, reinforcement training differs from other learning styles by focusing on maximization of the rewards. In addition, reinforcement learning is based on making decisions from the data sequentially as the next input varies based on what has been the output from the previous input. (Mummert et al., 2022.) Figure 5 demonstrates the idea behind reinforcement learning by showing how the model learns through punishment or reward as a result of environmental changes that are either advantageous to specific behaviors or actions, or unfavorable.

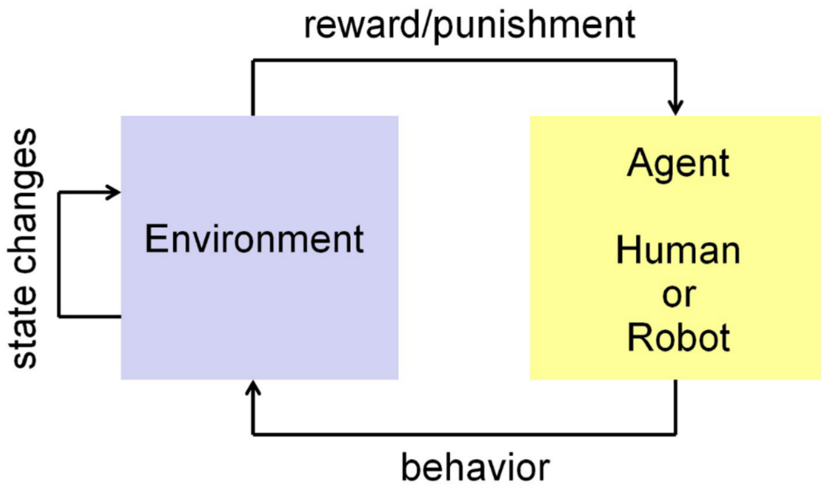

Figure 5. Universal model of reinforcement learning (R. Liu et al., 2021).

Hurwitz & Kirsch (2018) explain that robotics is one of the most popular applications where reinforcement learning is implemented as trial and error is an effective way for the robot to learn from its mistakes and correct its behavior. In addition to robotics, reinforcement training is implemented to train artificial intelligence systems that are used in video games (Berner et al., 2019). According to Pan et al. (2017) one conceivable application of reinforcement learning is self-driving automobiles, because the models used would be trained virtually in a confined environment before being implemented on real vehicles. It is currently unknown what kind of learning models Tesla is utilizing in their self-driving car applications, but given that they now have that many cars in their fleet and that number is growing, it is thought that they may be able to apply reinforcement training as well (Bouchard, 2019).

3.4 Algorithms used in machine learning

This chapter introduces some commonly known algorithms that are used in machine learning-based applications. In addition, this chapter aims to provide clear information to assimilate the main idea behind each of the algorithms which are being introduced. The focus is to illustrate how the data of these algorithms are presented.

3.4.1 Gradient descent

Gradient descent is an optimization algorithm that is used to discover the minimum from a convex function. In machine learning, gradient descent algorithm is utilized to train other machine learning models as it is attempting to minimize differences between the actual and anticipated results. Hence, gradient descent is commonly implemented in machine learning because it can be used to optimize the parameters of other machine learning algorithms, such as linear regression and logistical regression, which are discussed in greater detail later. (Brownlee, 2016; DataCamp Team, 2022.)

As gradient descent plays a big role in machine learning in general it is important to understand its core idea. To comprehend what gradient descent does DataCamp Team (2022) explain the following, “We imagine a person who has lost his way in the mountains. Basically, it will be a matter of finding the way back by first looking for the direction with a steep downward slope. After having followed this direction for a certain distance, this method must be repeated until a valley is reached (the lowest value). In machine learning, the gradient descent consists of repeating this method in a loop until finding a minimum for the cost function.” As a result, depending on the size of the dataset, gradient descent frequently necessitates a considerable number of calculation iterations.

Figure 6 illustrates how the gradient descent algorithm is visualized. Learning rate is a value and determinates how big or small steps the algorithm takes to reach its goal, the minimum. A low learning rate means that the prediction is more accurate but takes more time and computing power to process. Visa versa high learning rate is faster and more efficient but not as accurate and can lead to a situation where the actual minimum is missed. Usually learning rates are quite low values such as 0,1 and 0,001. Random initial value is the value where functions start calculation towards the minimum.

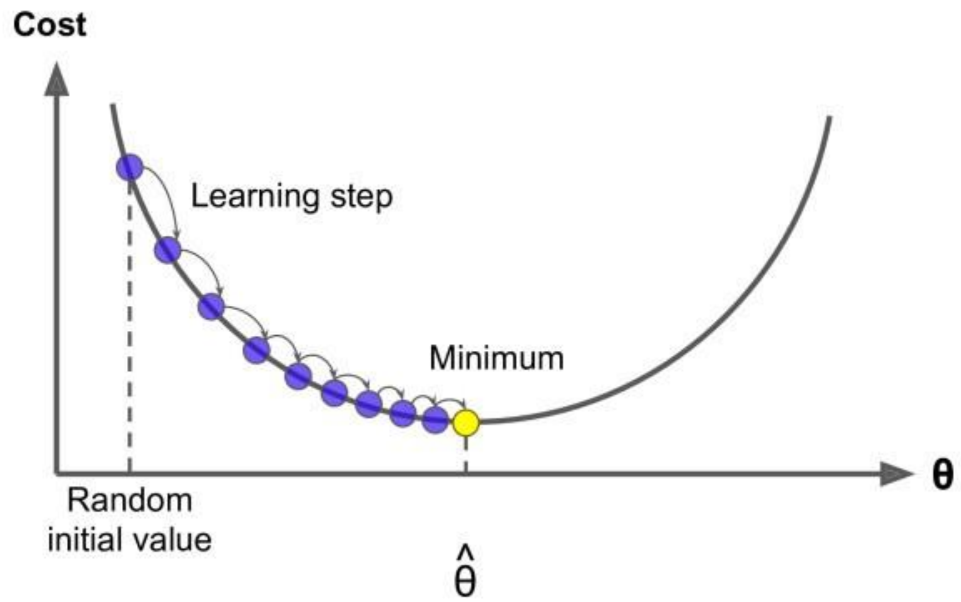

Figure 6. Concept of Gradient Descent in Machine Learning (Gudimalla, 2021).

Gradient descent algorithms can be divided into three types: Stochastic Gradient Descent, Batch Gradient Descent, and Mini Batch Gradient Descent. These styles vary based on how much they include training data when taking one learning step. Style decisions are therefore based on the size of the dataset and how efficient the processing is required to be. (Ray, 2019.) As seen in this chapter, instead that gradient descent algorithm is used for a particular problem, it lays a foundation to optimize other machine learning models by reducing their errors.

3.4.2 Linear regression

Linear regression is a simple machine learning algorithm that is especially utilized for forecasting and prediction analyses. It is mentioned to be one of the most popular and easiest machine learning algorithms to use and understand (Savcı, 2022). Sunil (2017) explains that linear regression is able to show a relationship between independent and dependent variables. While the independent variable is able to measure continuous or categorical values the dependent variable has to have continuous and real values in order to function. Real values can be for instance material price, age, sales values, or cost. Linear regression is under supervised learning approach as it is using labelled datasets to learn. Hurwitz & Kirsch (2018) state that with regressions analysis it is important to know the context related to data as otherwise predictions based on the result may be inaccurate. Linear Regression can be separated into two types of models: simple and multivariate linear regression (Maulud & Abdulazeez, 2020).

Simple regression is a model that illustrates the relationship between a single dependent variable and an independent variable. Ray (2019) describes Linear regression as a suitable fit for analysis if it is known that the connection between the covariates and the response variable is linear. However, due to its simplicity, simple regression is not an effective fit for nonlinear and complex issues. As a result, more sophisticated algorithms may be needed to solve real-life problems because multiple independent variables are frequently present.

Contrary to single regression, multivariate linear regression has one dependent variable but multiple independent variables. Thus, it is an extension of single regression as it takes more independent variables into account (Maulud & Abdulazeez, 2020). For instance, when predicting a value of a house it is useful to consider also other independent variables such as location, nearby services, and past prices of similar houses in the neighborhood. Thus, multivariate regression has the potential to provide more realistic predictions compared to single regression. However, Ray (2019) reminds us that more variables don't automatically lead to more accurate predictions. Adding more variables increase the probability of multicollinearity, meaning that independent variables have a high correlation with each other which causes redundant information. Therefore, independent variables should correlate with dependent variable but not between each other. Figure 6 expresses the linear relationship between dependent variables and independent variables and them having a positive linear relationship. Therefore, as a result, the value of Y (dependent variable) grows together with the value of X (independent variables). Line of regression presents a straight line that has the best fine meaning by minimizing its distance from the data points. In addition, the below figure shows how results of linear regression are visualized.

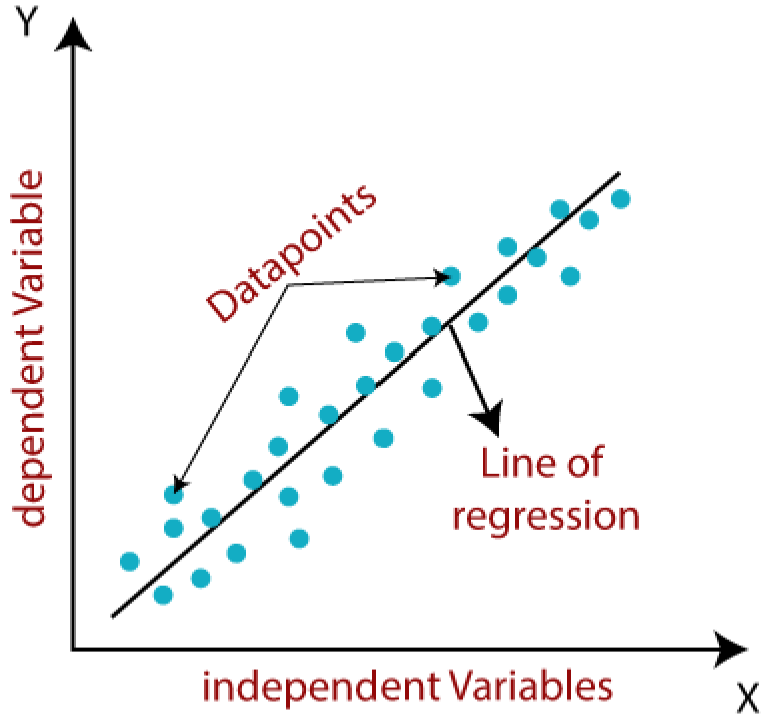

Figure 6. Linear Regression in Machine Learning (Javatpoint, 2021).

3.4.3 Logistic regression

Logistics regression is a popular machine learning algorithm and is implemented when it is required to classify given data set into two or more classes. Thus, it is utilized specifically for classification problems, and it is part of supervised learning techniques.

Unlike linear regression which is used to output continuous values such as a price of a stock, logistic regression is creating probabilities that are discrete and limited to certain values. This difference is due to sigmoid function that enables to generate probabilities between 0 and 1 from any real numbers and these probabilities are then utilized to classify data. Probabilities with two classes are called binary classifications and above two are multiclass classifications. Therefore, the predicted outputs have to be either discrete or categorical values such as 0 or 1, Yes or No, or with multiclass classes words such as “car”, “plane” or “bike”. (Brownlee, 2016.)

Figure 7 illustrates how logistics regression is presented after the data has passed through the algorithm. It creates a S-shaped curve that presents probabilistic values as 0 and 1. Threshold value is separating data into either category 0 or 1.

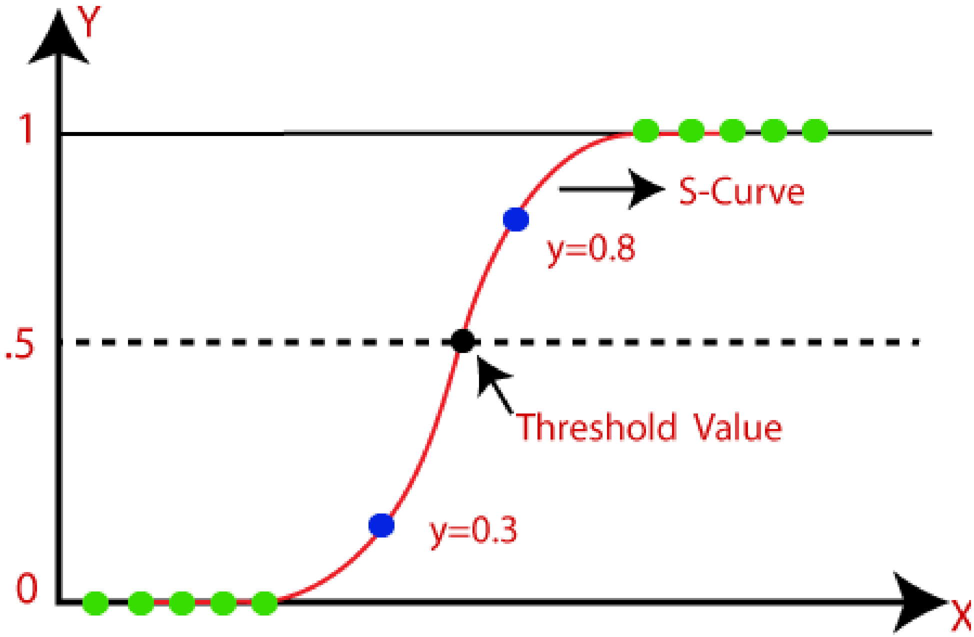

Figure 7. Logistic regression (JavaTpoint, n.d.-b).

Logistic regression is utilized mainly implemented to solve large-scale problems for example to predict whether an e-mail is spam or not or in the financial sector if a bank transfer is fraudulent or not. In addition, sales, healthcare, and marketing are sectors that have seen applications of logistics regression. In a summary, logistic regression is implemented when it is required to know the likelihood of something happening or not. (Ray, 2019; Chandrasekaran, 2021.)

3.4.4 Decision tree

A decision tree is a supervised machine learning algorithm, and it can be implemented for both classification and regression problems. Decision tree models where the target variable can have discrete values are called classification trees. Correspondingly, when a target variable can take the form of a continuous value such as real numbers these are called regression trees. Decision trees are one of the most popular and practical methods of supervised learning (Mitchell, 1997).

Decision tree's main purpose is to create a tree-like model that is capable of predicting the value of a variable. The nodes of the tree illustrate different kind of features the dataset consist of, decision rules are represented as branches and lastly, leaf nodes represent the outcome of the decision (Mitchell, 1997). Thus, decision tree's structure can be seen as similar to flowcharts which are using if else-based decision-making logic. Figure 8 is representing the general idea decision tree where the outlook of the weather presents the root node. This then expands to branches that include different weather conditions and their nodes and based on the decision rules a decision can be made whether to play tennis or not.

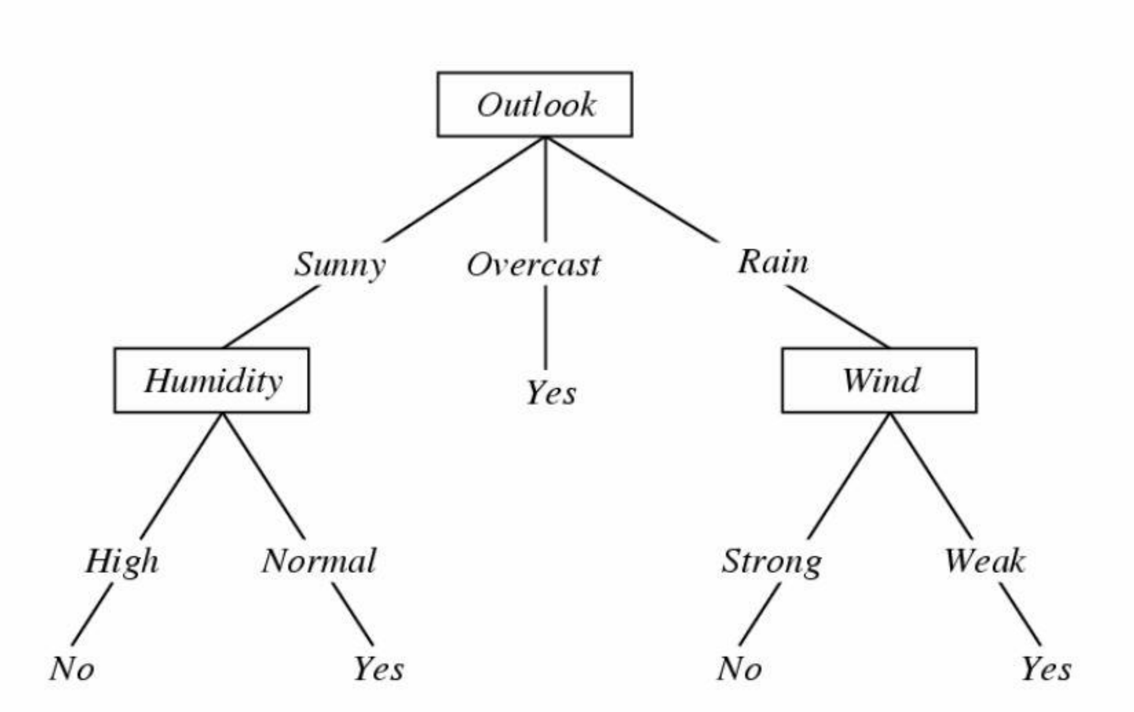

Figure 8. Decision Tree for PlayTennis (Mitchell, 2005).

Applications of decision trees are known to be used in text classification and extraction, for example, in libraries where they can classify papers into various genres based on the types of words and symbols they contain. Decision tree applications can also be used by medical experts to diagnose diseases. Decision trees could be used in stock markets to forecast changes in stock worth. (Navada et al., 2011.)

3.4.5 Random forest

Random forest is a supervised machine learning algorithm that can be used for both classification and regression problems. Due to its versatility in use cases, it is also one of the most widely used machine learning algorithms. Y. Liu et al. (2012) explain that random forest is a combination of multiple decision tree classifiers. In contrast to the previously presented decision tree algorithm, random forests view trees as weak learners whose efforts are combined to perform better than a single strong learner. Therefore, the reasoning behind the random forest tree algorithm is as follows: each tree selects the most well-liked class, after which all votes are added together to determine the outcome, which then serves as a shared prediction of the model. In addition, this method also helps in overfitting issues as it reduces the variance of the model. (CFI Education Inc, 2021; Y. Liu et al., 2012.)

Random forest gives more accurate predictions than a single decision tree in general because prediction performance is connected with the number of trees in the model and it is resistant to overfitting, which means it works well with new data. Overfitting difficulties occur when an algorithm performs well on the data used for training but produces erroneous predictions when new data is used. The disadvantage of its enchased prediction skills is that it is slower in processing and has higher computational needs than the decision tree method. (Hoffman, 2020); (IBM Cloud Education, 2021.) Figure 9 illustrates the random forest algorithm, and the above-explained logic can be seen visualized.

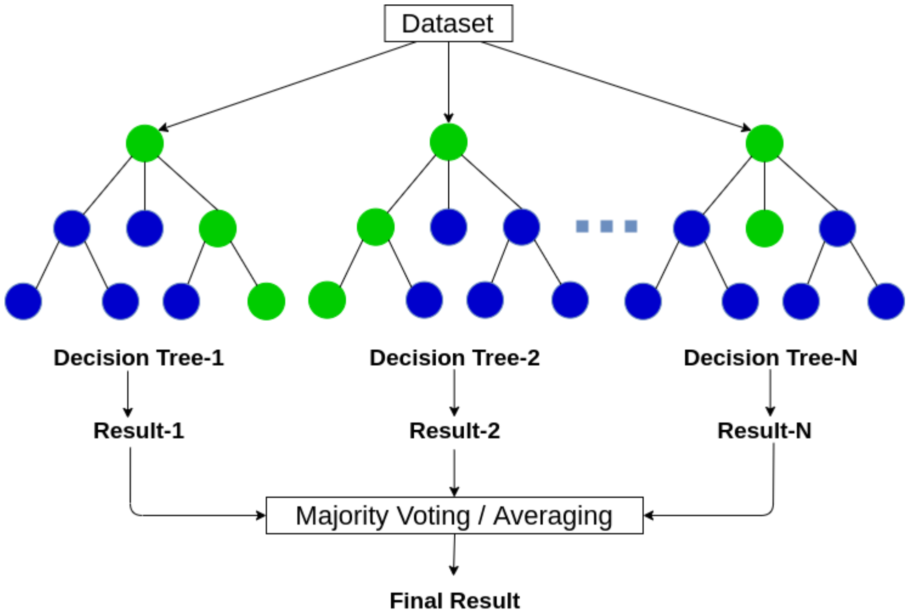

Figure 9. Random Forest (Sharma, 2020).

3.4.6 K Means clustering

K means clustering is one of the most studied algorithms of clustering and in general it is mentioned to provide accurate predictions. This algorithm belongs to unsupervised learning style as the target variable is unknown (Nazeer & Sebastian, 2009). K means clustering groups up unlabelled data into different groups based on their similarities. These groups are clusters, and the grouping process is called clustering. Letter K in the algorithm determines how many clusters the algorithm has to create when it's processing the data. Figure 10 shows how the unlabelled data is separated into three different groups. It also shows centroids which present the center of each cluster, and they illustrate the mean of data points inside the cluster. In this manner, the algorithm process until each data point is nearer to its own cluster's centroid than to other cluster's centroids (Lopez Yse, n.d.). Ayodele (2010) explains that as the locations of centroids will affect results, they should be decided wisely for instance as far as possible from each other. Google Developers (2022) mention how simple implementation, scaling ability to large data sets, guaranteed convergence, and adaptation to new examples are one of the key advantages of K means.

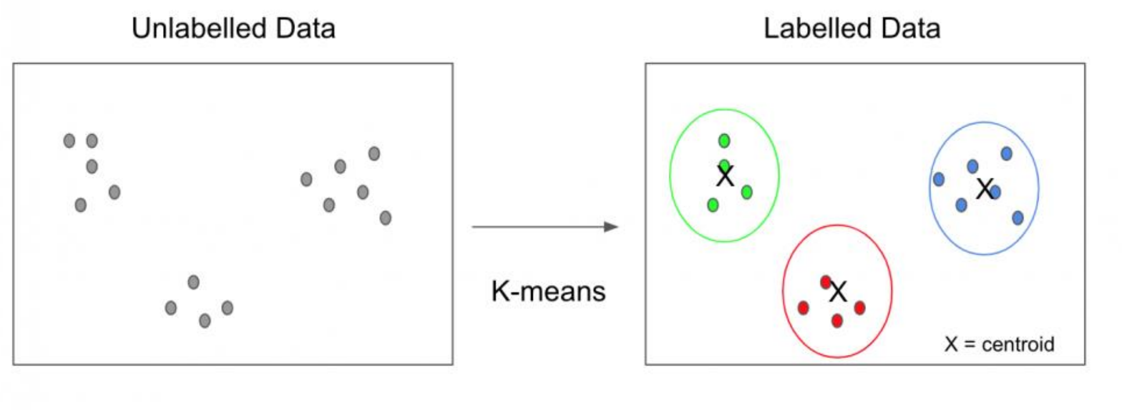

Figure 10. K Means Clustering (Team Data Science, 2019).

Lopez Yse, (n.d.) mentions that K means clustering algorithm is utilized for example in inventory categorization, spotting anomalies, and behavioural segmentation. Ranjan (2021) adds that segmentation of customers and optimization of delivery routes, and document classification are examples where K means clustering is implemented.

3.4.7 Support vector machine

Support vector machine algorithms were first introduced in the 1990s and since that they have been very popular in machine learning and found to be highly efficient, especially for pattern recognition. Pattern recognition is a process where the goal is to classify data on prior knowledge that has been already received or to extract statistical information from the data. (Karamizadeh et al., 2014). Support vector machine techniques have primarily been employed to handle supervised learning issues, but they have also been successfully used for unsupervised learning in recent years (Zhao et al., 2009). Support vector machines may address classification and regression issues, however classification appears to be the most common application case. Common use cases include face, speech, and picture recognition, text categorization, fraudulent credit card recognition, and handwritten text recognition. (Noble, 2006.)

The target of the support vector machine algorithm is creating a hyperplane that separates data into two distinct classes. The separation is done by finding the most optimal hyperplane which is the one that has the greatest distance between data points for both classes. The distance is called margin in this model. As the algorithm is in most cases adjusted to maximize the distance it ensures that the classification is done accurately. (Karamizadeh et al., 2014.) As mentioned, the support vector machine is able to separate data points only to two classes in its common simple form which is called binary classification. However, further complex models have been created for the classification of multiple classes but this work focuses explaining the binary classification with support vector algorithm (Brownlee, 2016).

Figure 11 illustrates how the most optimal hyperplane separates the data into two classes. It also shows support vectors which are data points that are positioned closer to the hyperplane, and they are utilized to maximize the margin between them and the hyperplane. An optimal hyperplane is not selected automatically but during the process, there is a selection of hyperplanes that are then trying to solve the optimization problem but in the end, there can be only one that maximizes it most accurately and it will be then selected. (Karamizadeh et al., 2014.)

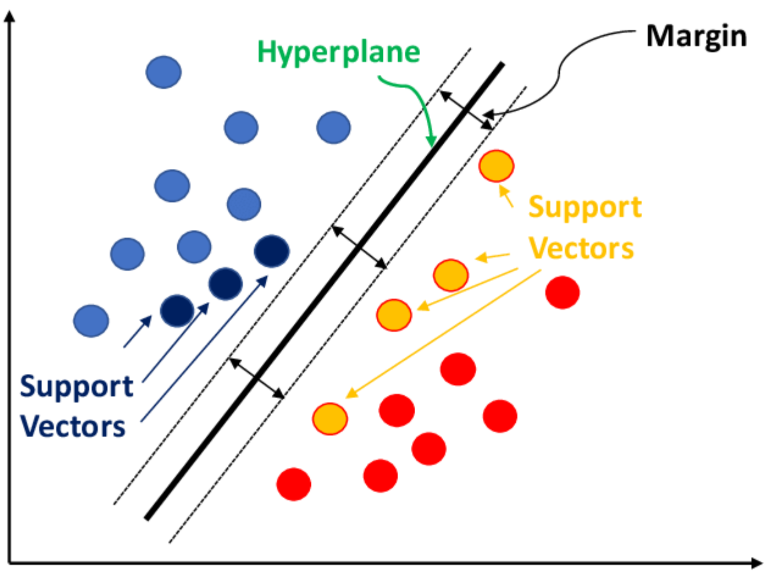

Figure 11. Support Vector Machine Visualization (Manjrekar & Dudukovic, 2019).

4 Procurement and Purchasing

This chapter aims to introduce what kind of processes are included in purchasing and procurement and how they differ from each other. First, the terminology is discussed. Secondly, the main processes of purchasing are introduced followed by procurement. Lastly, the relationship between procurement and purchasing process is introduced and how they could function in a real-life scenario.

4.1 Terminology

When searching for information about procurement and purchasing researchers might find a situation where different definitions and terms are provided about similar processes. As there are no official or strictly used definitions it might be challenging to choose what kind of terms should be used and why. Hence, this chapter introduces some of the commonly known definitions.

Procurement is also seen defined as strategic procurement and purchasing as operational procurement (Bartolini; 2014; Taulia, 2022). Logistiikan Maailma (2023) separates these two as strategic procurement and operational purchasing. Also, the term strategic purchasing has been mentioned. Strategic purchasers and operational purchasers are common titles when companies are looking for employees to fill procurement and purchasing positions. Furthermore, Young (2022) explains that procurement includes a strategic process with a proactive approach about finding suppliers and sourcing materials, whereas purchasing includes transactional process and has a reactive approach to purchasing materials when requirements arise. On the other hand, some are seeing purchasing as a subset of procurement (Reich, 2020). For this study, I have chosen to follow similar terminology as Young.

Depending on the organization the responsibilities and roles might differentiate. For instance, in some organizations, demand forecasting might be a task for procurement but in some, it falls under purchasers’ responsibility. Also, in some companies, a purchaser might be responsible for both procurement and purchasing processes, especially if the company is a smaller one and there is no requirement or resources to share the roles. Therefore, it is critical to remember that the procurement and purchasing processes outlined in this study are more like broad guidelines for how the process might work in most businesses.

4.2 Purchasing

As already mentioned above purchasing as a process is focusing generally more on transactional tasks to acquire the required goods or services which the company is needing to meet its customer commitments. Purchasing process is focusing on activities such as converting purchase requisitions into purchase orders, expediting orders, ensuring that orders are received, and invoices have been received and payments fulfilled accordingly. Khan et al. (2019) describe the core functions of purchasing as follows : to purchase the correct material, component, or service in the correct quantity at the correct place and time. Correspondingly, Codeless Platforms (2019) explains how purchasing focuses on transactional activities and meeting short-term goals such as quantity, cost, and timing. Therefore, purchasing can be seen as an operational day-to-day process to secure the arisen demands. Codeless Platforms suggest that the following tasks are common for purchasing professionals :

- To receive a purchase requisition
- To evaluate received RFQs
- To create and distribute purchase orders
- To receive products or services
- To assure quality of received product or service
- To arrange payment to suppliers

In addition, for purchasers it is important to monitor that the order has been left from the suppliers’ premises as it has been agreed between the purchaser and the supplier. Often purchasers receive tracking information when the order is getting delivered which they can follow to stay updated about orders' estimated time of arrival. This is especially important when the ordered goods are needed urgently or ordered by using just in time method and thus there is no room for delays (Banton, 2022). If a problem emerges, operational purchasers are responsible for initiating an expediting process, and if the problem necessitates it, often procurement will also assist in the process because they are responsible for the supplier.

4.3 Procurement

Procurement, on the other hand, is seen often focusing on a more strategical process of sourcing products or services. It is responsible for instance finding suitable suppliers who would be able to fulfill demands based on products' specific requirements, creating functional supplier relationships, analyze cost and suppliers' KPI results. Procurement also focuses into more longer-term goals and uses generally a more proactive and strategic approach compared to purchasing. Reich (2020) explains that procurement is a strategic process to source material or service and lists the followings tasks that are commonly associated with procurement :

- To plan a corporate strategy for buying
- To identify short-term and long-term goals
- To assist with budgeting and forecasting research
- To source and evaluate potential suppliers
- To negotiate contracts with suppliers
- To administrate the purchasing and payment process
- To conduct spending analysis and reporting

In addition to the items listed above, maintaining and developing relationships with suppliers is critical in procurement because it allows for the establishment of long-term business relationships between the procurement organization and the supplier, lowering the risk of material shortages and other potential issues. Fostering this relationship provides a situation that benefits both the procurement organization and the supplier.

Furthermore, ensuring that the quality of the materials is in line with the internal requirements is an important aspect of procuring because it adds no value to create relationships with suppliers if they are unable to meet the quality standards to begin with. Therefore, relevant supplier research is a challenge for procurement professionals. (SourceDogg, 2021.)

4.4 Relationship between procurement and purchasing

Magal & Word (2011) explain that the purchasing process starts as the company's warehouse recognizes that there is a requirement to purchase materials, for instance as the levels of inventory have declined under a certain limit. Purchasing requisition is created by the warehouse which is then sent and received at the purchasing department. The purchasing department is responsible for identifying the most suitable vendor for the material at that moment and converting the requisition into a purchase order. After the purchase order is created and sent to the desired vendor the purchasing organization is ensuring that the purchase order is confirmed within the required schedule, quantity, and agreed price level. Before the ordered goods are received purchaser monitors that the order is dispatched on the agreed delivery date. Warehouse then receives the materials, and the vendor sends an invoice. The accounting department is responsible for the invoice processing and therefore sends payment to the vendor. After the payment is successfully sent to the vendor the procurement process is completed for that single requirement. Before all above-mentioned has occurred, procurement department have identified need for the material in the company's processes sourced suppliers for the material, negotiated contract including prices and other important aspects such as material lead times and contract penalty clauses.

Therefore, in order that the processes work smoothly both procurement and purchasing organizations must work in cooperation. From table 1 which illustrates key differences between procurement and purchasing activities, we can see how procurement focuses on processes that happen before the actual purchase requisition and purchase order are created. We can consider procurement to be the entity that plays a critical role in ensuring that purchasing is possible and that it occurs with as few issues as possible on the part of suppliers. Therefore, as procurement is responsible for creating contracts and sourcing the materials which then are then used by the purchasers to create orders, it is important that the information is correct for example in contracts, and that suppliers also know what is required from them regarding aspects such as quality and delivery times.

Table 1. Procurement and purchasing differences. Adapted from (Arbuzova, 2021).

Procurement	Purchasing
A complex process associated with sourcing and obtaining goods and services	A set of activities directly related to acquiring goods and services
Includes multiple stages and has numerous people involved	Fewer steps and fewer employees involved
Stages can be tailored depending on the company and vendor	Stages are usually standard among all businesses
Takes place before, during, and after the purchase	The actual process of purchasing
Aimed at recognizing and satisfying the company's internal needs	Aimed at ordering and paying for goods
Proactive approach	Reactive approach
Involves need identification, supplier management, contract management, and purchasing	Comprises placing an order, expediting, receipt of goods in the stock, and making payments
Has long-term goals like gaining competitive advantage for the business	Has short-term goals like obtaining the necessary goods at the right time
Focused on creating and maintaining strong supplier relationships	Focused on making transactions
Places more importance on the item's value than its cost	More focused on the item's price

5 Machine learning within purchasing and procurement

This chapter focuses first on introducing how machine learning techniques are currently or have the potential to be utilized within purchasing and procurement based on publications both from academics and industry professionals. The potential benefits of machine learning implementations in purchasing and procurement are first discussed. Second, the potential challenges of these machine learning implementations are discussed.

5.1 Advantages of machine learning applications

As usual in process development key targets are saving time and cost and increasing efficiency and machine learning solutions for procurement and purchasing are not an exception in that regard (EYVO, 2021). According to Deloitte Switzerland's (2019) survey, 51% of procurement leaders responded that they have implemented advanced analytics technologies such as machine learning, natural language processing, and cognitive analytics, and 93% are satisfied with the benefits they are receiving. Furthermore, according to Schmitz (2017) machine learning, especially combined with other artificial intelligence features can be seen providing value for procurement processes. Schmitz tells how machine learning applications can learn price developments from the past, develop forecasts, and provide suggestion to find the supplier for the material with the most efficient price level. For instance, this can be advantageous when purchasing raw materials where price fluctuations can be dramatic even in a short period of time. Also, Schmitz explain how machine learning based automated tax calculations are reducing the risk of paying incorrect amount of taxes. Incoming invoices will be audited based on specifics regulations and irregularities are detected before they would cause potentially any monetary harm for the company. In general, machine learning is mentioned to help with different kind of data maintenance problems (Hemant, 2021). Likewise, Hemant adds how machine learning should be considered as one of the options when selecting technologies for procurement processes while companies are searching solutions for their current business issues and challenges. Below main advantages of machine learning implementations for purchasing and procurement are introduced.

5.1.1 Forecasting

Taranenko (2021) states that in the area of demand forecasting machine learning techniques have the possibility to increase data processing speed, produce more accurate predictions, analyze larger amounts of data, identify hidden patterns from data, and have more adaptability to demand changes. McKinsey's (2017) report indicates that machine learning-based techniques are able to reduce cost of sales by 65% due to increased accuracy of forecasting. More precise forecasts also improve supplier relationships as they would be receiving more stable forecasts without surprise purchase orders that were not forecasted before. Samir (2014) explains that manual forecasting requires a significant amount of time, and still, the forecasting might end up not providing accurate enough results. Thus, Samir argues that with help of machine learning that is able to take into account more variables and process more data in a shorter amount of time, forecasting results would end more accurate than before. In addition, this would help persons who are creating forecasts to focus on other daily tasks which cannot be automated as easily.

Inventory value is often one of the key indicators that are carefully followed, especially by operative purchasers and their management. Samir explains that machine learning can optimize safety stock levels in cases where they are kept purposely overly high due to mistrust of suppliers to meet required demands. Also, Samir mentions that manual forecasting requires a great deal of time and yet fails to be accurate enough. Gezgin et al. (2017) argue that machine learning applications can provide advice for instance about material planning and how it should be changed. Gezgin et al. also mention that in a company where artificial intelligence–supported application was implemented, the company was able to reduce its inventory by 20%. Responsibility of who creates the forecasts is depending based on organizational structures as it could be for example a purchaser or a dedicated planner but nevertheless, procurement is always benefitting from accurate forecasting.

5.1.2 Spend analytics

According to Biedron (2019), Unnecessary transaction expenses, regulatory noncompliance, and overall process inefficiencies consume 3 to 4% of a company's external costs. Thus, regardless of the size of the organization this spending should be avoided or very least minimalized, and machine learning techniques are seen as one of the ways to assist with it. Dilmegani (2020) mentions that proactive spending analytics can spot opportunities for cost savings and manage possible risks and issues. For procurement departments, it is essential to receive accurate spending data to optimize and development processes even further and machine learning algorithms are mentioned to be able to provide this information by classifying spending data into different categories.

Usually, items without material numbers are challenging to categorize without using a significant amount of manual work to it. For instance, products might have only descriptions that provide challenges to the categorization process. However, Determine (2018) states that machine learning algorithms can help with classification issues by automatically categorizing spending types by detecting similarities in their context. Another example of how spend analytics including machine learning techniques provides value is supplier name normalization. Organizations may have different abbreviations and spellings of the same supplier in their separate systems, such as in accounts payable and purchasing. With the help of machine learning this data is collected and normalized based on the rules for instance by removing words and letters. (Issar, 2020.)

5.1.3 Monitoring transits

Purchasers are frequently asked, either internally or externally, what the estimated arrival time of purchased items in transit is. This information is critical for internal manufacturing scheduling, as well as for customers who are waiting for ordered items. According to McCrea (2019), predictive analytics for materials in transit can assist procurement because purchasers are responsible for ensuring that orders arrive on time and in the correct location. As a result, it is critical to monitor the status of orders in transit and to respond as quickly as possible if delays arise. Furthermore, SAP has previously released a machine learning-based solution for estimating the arrival times of in-transit stocks (SAP Products & Services, 2018b).

5.1.4 Sourcing

For a procurement organization, searching for and locating the right supplier based on specific needs can be a time-consuming procedure with no guarantee of satisfaction. As a result, sourcing typically consumes a large number of working hours that could be given to other tasks if it could be done more automatically and efficiently. Sengewald & Lackes (2022) explain how machine learning with the help of the Bayesian method can reduce both purchasing and procurement process costs. In this use case, the Bayesian method is learning stochastically optimal sourcing strategy directly from quotation data. Nagar et al. (2021) argue that machine learning can help in the process of determining the correct prices that are economically viable for the company to produce. Because sending quotations to several vendors and comparing pricing is part of the sourcing process, procurement organizations would benefit from having this price estimation as a guideline and assistance.

McCrea (2019) has mentioned sourcing to be one of the potential uses of machine learning applications as it would reduce the requirement for manually looking for materials with the best price, delivery time, and quality combination. For instance, the system would automatically send a request for a quotation to pre-defined selection suppliers when a source of supply for the material is not available. Correspondingly, Hemant (2021) explains that machine learning could help sourcing by identifying which suppliers are most suitable for specific projects.

5.1.5 Supplier relationship management

Vollmer (2018) states that after the supplier is decided machine learning helps monitor actions of the suppliers to find behaviors and patterns which might indicate that the supplier is not able to meet the current business or regulatory requirements. With the help of machine learning this process is more efficient than a human employee would be able to do and with the help of proved facts, it may be easier to bring up these issues to the supplier to discuss. Also, Nagar et al. (2021) point out that machine learning's pattern recognition capabilities provide fast information if the quality level by the supplier is not meeting criteria which could be for instance unstable delivery times or an increasing amount of quality defects on ordered materials.

Schmitz (2017) adds that machine learning would allow procurement experts to have an intelligent assistant to whom they could ask queries. For example, intelligence assistance would help how to form contracts. Machine learning could also provide intelligent assistance for the procurement employee during the process of contract selection with a supplier as machine learning applications would analyze contract texts based on suppliers' and purchasers' responsibilities and payment agreements. Machine learning applications would utilize in the above-mentioned process contract library that stores different kinds of agreement examples (Schmitz, 2017).

5.2 Challenges of machine learning applications

Typically, when implementing new technological applications for enterprise environment challenges arise. This has been noticed in machine learning applications as well. Some of the major issues encountered when implementing machine learning applications are addressed here. Examples of how these challenges may affect procurement purchasing groups are also offered.

5.2.1 Data quality

Döbel et al. (2018) explain that in machine learning challenges are different compared to more traditionally programmed applications. For instance, the learning model must be kept generic enough to function effectively in cases where it receives input that it did not observe and interpret during the training phase. As a result, the machine learning model's affectability is determined by the amount of high-quality training data it received throughout the training phase. By receiving too many incorrect data examples, the model will not be able to produce worthwhile outputs. As machine learning models are using already existing data and providing predictions based on that used data must be as accurate as possible to reality. (Alpaydin, 2010.) For instance, in procurement, if the material's stock and price values are not up to date the result that the model is providing will be certainly incorrect and could lead users to do decisions that might have an adverse impact on the procurement and purchasing organization. Therefore, beneficial usage of machine learning models could be challenging especially because inaccurate data has been mentioned to be one of the common issues of the procurement industry (Steers, 2021). In fact, this problem is not only limited to machine learning because in general one of the issues of procurement and new IT solutions implementations have been poor "master data" (Chopra, 2019).

Simpson Rochwerger (2021) argues that in academic publications machine learning models are not often used to solve problems but instead to prove the functionality of the introduced model. In real-life cases, both high-quality and accurate data are challenging to collect compared scientific research scenarios where data is easily available and to apply into the model. Thus, he states that data is clearly more important than the model that is being used. In order to use the model in operational use the user must have confidence in the results, and this trust can be achieved by implementing high-quality data into the model (Lee & Shin, 2020). Also based on research by Forrester (2022), 38% of responders reported that it is challenging to translate academic machine learning models into actual products that can be deployed into processes. As a result, it can be difficult to identify useful machine learning models for procurement and purchasing solely based on how they performed in scientific research, in a different because it is difficult to validate the data authenticity that has been used. To counter possible issues caused by inaccurate or otherwise unsuitable data, Lee & Shin (2020) suggest that companies should set up a process to control data quality that includes steps such as developing quality metrics, new data collection, and data quality evaluation.

5.2.2 Data traceability

In addition to data quality issues, Döbel et al. (2018) tells that traceability of the model can be also a problem. For example, decision tree models' output rationale can be tracked relatively easily, however deep neural network models' behavior can be hard to comprehend, and hence the reasoning behind the model may be challenging to explain to related stakeholders. Also, Forrester (2022) tell how data traceability and explainability are viewed as one of the most difficult challenges of integrating machine learning algorithms into business operations. Thus, the demand for in-house experts who can program or at the very least able to comprehend the functionality behind the models is rising and is predicted to increase during upcoming years. Therefore, also procurement organizations are required to support the learning of their employees who are responsible for developing processes. (Lee & Shin, 2020.) Furthermore, Lee & Shin remind us that for managers it can be challenging to understand and keep up with machine learning solutions and their possibilities and challenges as they are still new techniques to most of the businesses. Thus, underlying that the importance of knowledge sharing about machine learning across the business organizations.

5.2.3 Data volume

The fact that organizations are receiving and maintaining increasing amounts of data is creating also challenges for machine learning applications. Monteiro et al. (2021) state that in addition to the pure volume, the data is also becoming more diverse, meaning that it’s in different formats and from various sources. Organizations are forced to make decisions in a fast phase and rapidly which requires machine learning algorithms to learn efficiently from different data sources and large quantities of data. This phenomenon is causing the risk of outdated model because the data used to train the model is no longer relevant. Thus, in order to ensure that the models' decision-making performance is efficient, models should be able to adapt dynamically to the data changes nearly instantaneously. (Monteiro et al., 2021.) Therefore, also in procurement and purchasing it is important to ensure that the models are up to date and ensure data collected from different sources are functioning efficiently in the model and is that the data still relevant. For instance, purchasing and procurement data include price values, date information, and letter digits and these might change frequently, and thus implying that the previously listed concerns are equally valid inside procurement and purchasing organizations.

Procurement organizations generate large amounts of data daily from their processes for instance about spending and suppliers’ key performance indicators. This large amount of data is often referred to as “big data” and companies are attempting to benefit from it through analyzation (Little, 2014). Therefore, it is natural that machine learning is a viable option for these analyzations’ tasks. However, L’Heureux et al. (2017) raise a concern about how large amounts of data are bringing challenges for using machine learning techniques. As the size of the data increases, it influences also computational complexity. An increasingly growing set of data may raise expenses and require more time-consuming processing. In addition to data size challenges, Barrad & Valverde (2019) state that the availability of data is one of the main limitations of implementing machine learning and data analytic technologies specially when not all organizations collect enough information about their procurement processes.

5.2.4 Ethics & compliance

In recent years there has been an increased amount of attention pointed toward ethical and compliance guidelines as cases of misuse have come into light (Bailey & Shantz, 2018). Thus also, ethical questions have risen often when discussing machine learning implementations. For instance, procurement organizations are handling personal information such as names, email addresses, and telephone numbers. Also, in contractual matters, compliance laws and norms must be strictly observed, and in the event of wrongdoing, the use of machine learning is not considered as a valid justification for remission. (Lee & Shin, 2020.) These mentioned concerns have to be taken into an account for instance when vendor management applications are implemented which are using machine learning features as they are include personal information, especially when this information is external, and the supplier and other stakeholders have to have trust that the information is correctly maintained and securely stored in systems used by procurement and purchasing organizations.

5.2.5 Investing in emerging technologies

Justifying investments in emerging technologies has also viewed as troublesome, even though procurement managers are often interested in artificial intelligence - technologies such as machine learning (Deloitte, 2019). Similarly, Baier et al. (2019) highlight how machine learning specialists regard it as difficult to convince their consumers to pay for their services, which are typically new functionality to current software and services. Khuan (2019) discusses how, in procurement, it is difficult to find the financial rationale for investments in emerging technologies, and hence, in some circumstances, it has proven more efficient to hire new employees to complete existing work tasks. Furthermore, events such as global turmoil and trade wars are mentioned as reasons for hesitation in investing in emerging technologies. Previous findings align with Deloitte (2017) that states how 79% of procurement leaders are focusing on cost reduction. Hence if the potential benefits of machine learning techniques are uncertain it seems to be challenging to justify investments as the focus is mainly on cost-cutting rather than finding new technological solutions to invest in.

Currently, machine learning implementation projects do have a higher failure rate compared to more traditional technologies. Also, Lee & Shin (2020) remind us that machine learning models are not solutions for all business problems and require careful planning and research to determine whether implementing machine learning techniques is financially viable. This should be kept in mind, especially during unpredictable periods like the COVID-19 outbreak, which has interrupted procurement and purchase processes, making it difficult to validate new, uncertain investments (Deloitte, 2022).

6 Research methodology

According to Kothari (2004), research methodology is a systematic method to solve a research problems. It includes different steps that are supposed to be adopted by researchers and pushes researchers to evaluate which methods or techniques are relevant for the research that is being conducted. This chapter introduces the methodology of this research, data collection, analysis, limitations and a brief summary.

6.1 Research strategy

This research consists of theoretical and empirical parts. The theoretical part includes secondary research data from literature, journals, articles, educational YouTube videos, etc. The empirical section consists of interviews which were conducted as part of the research to provide practical experience about the topic and in addition, this enables me to utilize primary sources.

A semi-structured interview approach was chosen for this study as it provides space and flexibility for interviewees to respond and discuss topics (Kallio et al., 2016). In addition, a semi-structured provides more reciprocity between the participant and the interviewer compared to a structured interview approach (Galletta & Cross, 2013). The interview structure for this study followed predetermined questions in their content, and the number of questions was likewise same on each interview. Mixed methods of research were used with a high focus on qualitative methods due to the nature of interviews as only one question from the interview required a numerical answer. Before sending the interview attendance proposals one test interview was done with an operational purchaser from the same organization where the actual interviews were later also conducted.

6.2 Data collection

The company where interviews were conducted is an energy solution provider with industrial manufacturing as its primary activity and it is based in Finland and has a strong emphasis on exporting its products worldwide. The research was not conducted by request of the company instead I contacted the company personally and requested permission to carry out interviews and use them as a resource for this research. After permission was given, I contacted the potential interview participants. It was also agreed that the name of the company will stay unknown.

Data for the research was collected by requesting interviews from ten different purchasing or procurement professionals. The invitation for the interview was sent by email and two weeks of time were given to respond to inform if he or she was interested in attending. In total eight persons answered. In addition, an in-person interview possibility was offered if the participant accepted that, which was then accepted by two participants. The remaining six interviews were answered in writing without an additional discussion about the questions. It was necessary to provide this option for the sake of flexibility, as interviewees' busy schedules could have made it difficult to find time for a meeting, even though the interaction between interviewer and interviewee is limited compared to in-person interviews. We can therefore say that the data was gathered in the form of surveys, backed by two interviews, given that some participants responded to it as an interview while the majority responded on it as a survey. However, during this study, I am using the term interview as it is also written on the question templates which were sent to participants and as interviews were initially the important part of this study.

To select potential participants for interviews non – probability sampling was utilized as all the participants were from the same company. In particular, convenience sampling was used because potential participants were selected based on the fact that they work on purchasing or procurement in one company. Therefore, the population was limited to one company and two fields of profession.

6.3 Data analysis

Received answers from interviews were analyzed by using Microsoft Excel spreadsheet – program. First, the written answers were transferred from the interview templates to Excel where they were categorized and labelled. Then occurring themes and words were highlighted from the answers. Also, simple calculations were done about responders' experience years and counting how many times different topics were mentioned among the answers. Due to the simple structure of received data from interviews, I did not see the necessity to use any other software than the earlier mentioned ones.

6.4 Limitations of research methodology

The fact that the interviews were collected from a single company reduced the representation of different viewpoints as the process and practices vary between different companies. More companies and organizations would have brought most likely different types of answers but on the other hand, this would have taken more time. Also, answers are leaning more toward experiences and thoughts from purchasing professionals as the response rate was higher among them. Time was a significant limitation as it had to be set aside for preparing and sending the interviews, conducting interviews, and analyzing the results. Thus, a compromise between sample size and time resources was made.

The option of not conducting face-to-face interviews raises the possibility that respondents misunderstood questions or were unsure what questions meant. Therefore, creating a risk of receiving not relevant answer which could have been avoided by having a face-to-face meeting but in my point of view the flexibility in answering was justified due to responders' busy schedules or in general responders might feel more comfortable answering in writing and without intercourse with the researcher.

6.5 Summary of research methodology

This chapter introduced how this research was designed and which reasons affected how the research was conducted. Secondary data came from various academic and industry-related sources and primary data through semi-structured interviews. Personally, an important aspect was the capability to collect primary research data in the form of interviews. Interview questions were collected to a Microsoft Word template and analyzed using Microsoft Excel. The primary limitations of the research methodology were the small sample size and the fact that interviews were only done in two distinct organizations within the same company. The main reason for the mentioned limits was a lack of time resources, as a larger number of interviews would not have been possible with the current resources and schedule.

7 Results

This chapter covers the findings from the research interviews, which attempted to answer the third research question concerning the issue which prevent effective purchase and procurement and how machine learning solutions could potentially help to overcome them.

7.1 Introduction

The purpose of this part of the research was to discuss the current challenges and disadvantages of the implementation of machine learning techniques in purchasing and procurement processes. Furthermore, interviews were conducted in a company to collect firsthand feedback on the current state of purchasing and procurement tasks and processes. Later chapters discuss and interpret the results in relation to the current state of machine learning implementation challenges and opportunities.

The goal of the interviews was to identify the currently time-consuming and challenging tasks in procurement and purchasing. In addition, responders were asked to tell which job task they would like to receive more support from information technology. All questions can be seen at the end of this research paper in the Appendices – section. As the questions were similar for both purchasing and procurement but answers between them were following mostly different themes, therefore the results have been also divided into purchasing and procurement answers. Both results chapters start by providing information about the sample size and experience level of the responders. Afterwards, the following sub-chapters focus on the main topics and themes which came up during the interviews. These chapters are structured by using answers from interview questions 5, 6, and 7.

7.2 Purchasing

Overall, five purchasing professionals took part in the interviews and all of them were working as operative purchasers, despite one having a different job title but the job description and tasks were similar to an operative purchaser. Responders' experience of purchasing was the following: Two had less than a year, one had four years, one had six and a half and one had seven years.

7.2.1 Forecasting

Four out of five responders mentioned demand forecasting, being one of the main job tasks that require lots of manual work and is time-consuming. One respondent stated that the sheer volume of material numbers is a major reason why the process takes so long, while another explained that the materials are complex, and that forecasting must be done individually based on the specific requirements of each project and their characteristics. In addition, trust in the production program's correctness was mentioned two times as an issue and its providing challenges to forecasting and ordering.

"*If it would be possible to somehow automatization manual forecast process. […] there are quite many variables what to think about (types, Suppliers, etc.*)"

7.2.2 Order follow-up

Order follow-up was mentioned among three responders as one of the key tasks which require constantly time. The order follow-up process includes such tasks as requesting earlier deliveries, postponements, cancellations and in general taking care that the purchaser has the latest knowledge about the status of each purchase order. These purchase order rescheduling requirements may be caused by internal such as production or external such as supplier related reasons. Three respondents specifically wished for increased support for the order follow-up process. For example, one responder wished below about the automatization of the order follow-up process:

“Automatic reminder about unconfirmed orders, automatic reminder if some order is confirmed drastically different than requested or lines are somehow noticed immediately from the list. “

Another respondent hoped for a corresponding improvement in order follow-up and assistance in receiving an accurate picture of orders' current status:

“Automatic tracking […] For example, which orders the vendor should prioritize, blocked invoices […] & reliability of deliveries.”

7.2.3 Work task notifications

Work task notifications are tasks in ERP – software such as SAP that are pointed to a person to carry out a requested activity. These activities can be for example checking the current stock level of a certain material if a recently received batch has quality issues and how it could potentially risk material availability for production (SAP, 2015). Four responders brought up how much time is used to deal with different kinds of tasks which are caused by quality notifications. Quality notifications are created internally when a material does not pass the required quality standards and therefore a chain of tasks is spread among different stakeholders to solve the issue. It is mentioned that often in these cases responders are spending time investigating how to solve the issue. This might require contacting different stakeholders internally and externally and this process requires time and knowledge to know who to contact. For instance, sometimes it is unclear who to contact in these cases and what is expected from responders. When asked about particularly difficult tasks one of the responders mentioned below:

” Some quality notifications – close communication internally and externally about what is needed and how to solve the issue, what parts are needed etc.

7.3 Procurement

In total there were three responders from procurement and all their work titles were strategic purchaser. Experience in years among the responders was three, 11, and 14. Due to misalignment the experience also includes experience years from purchasing if the person had it before procurement.

7.3.1 Material prices and pricing management

From all procurement responders’ different kinds of material price-related issues were risen during the interviews. For instance, price negotiations are very challenging as they require time, preparation, and resources. The high inflation rate and unstable market situation were mentioned as one of the key reasons for these challenges. Also, price differences in invoices and purchase orders and updating them were mentioned as tasks that consume time and resources but do not provide extra value to the work itself. Therefore, automation for the price updating process was wished as a topic for further development. Also, analytic tools to maintain and follow price and spending trends were wished to be implemented to ease the current workload. One responder commented below about challenges with price negotiations:

“Price negotiation is very challenging at the moment due to high inflation rate and unstable market”

7.3.2 Supplier relationship management

Supplier relationship management is the evaluation process of suppliers' strengths, performance, and capabilities in relation to the company's overall strategy, as well as the deciding of what activities to involve in with different suppliers, and the planning and execution of all interactions with suppliers in a consistent way, in order to maximize the value during the relationship life cycle (Hughes, 2009). Supplier relation management was mentioned by two responders as a process that requires a lot of time and effort. A third respondent also stated, without mentioning the phrase supplier relationship management, that they may become overly attached to their suppliers, which can lead to situations where ending a connection with a supplier is difficult and risky. This was mentioned to be due to the limited number of suppliers who are able to produce complex materials which are required for the company's processes.

7.3.3 Uncertainty of the future

During the interviews, the uncertainty of the future was mentioned by two responders specifically. Both prediction of future material volumes and price changes are seen currently as difficult due to unstable market situation and high inflation rate. In general, the uncertainty of future events compared to more stable times in the past is causing challenges to the procurement process. In the company where the interviews were conducted procurement professionals are responsible of longer-term volume planning while purchasing professionals have more responsibility within material-level forecasting. Therefore, as demands are closely based on the current global market situation longer-term volume forecasting is causing challenges to procurement organizations.

7.4 Summary of results

In the area of purchasing demand forecasting, order follow-up, and task notifications were mentioned as work assignments that cause most issues and prevent efficient and effortless purchasing to be possible. Especially demand forecasting and order follow-up were pointed out as areas where is a clear need for support from information technology solutions. Both heavy manual workload and uncertainty of the forecasts were the main issues that caused limited efficiency of forecasting. In general, lack of automatization and uncertainty with data are factors that consume vast amounts of time and effort from purchasers.

In procurement different kinds of topics around material pricing were mentioned by all responders as challenging tasks. In addition, in general, the prediction of what might happen in the future, and how future events will affect procurement was mentioned as a great concern due unstable market situation. Finally, respondents mentioned supplier management as a task that demands a significant amount of time and effort to be effective.

Purchasing and procurement professionals were struggling with quite identical issues in the case of forecasting demands and uncertainty of the future as these both are changing due to external factors such current market situation and global crises. Procurement professionals in the company where interviews were conducted focused on long-term volume planning, whereas purchasers dealt with short-term demand planning and forecasting.

8 Discussion

This chapter will discuss about the results of the study. First, research aims and questions of the study are reminded. Then key findings of this study are expressed by first discussing the findings from the literature and then from the interviews. Next results are interpreted. Then limitations and recommendations for future studies are presented.

8.1 Research aims and questions

This research aimed to explore machine learning techniques in purchasing and procurement by investigating the current advantages and challenges of their implementation. In addition, explore which issues are currently preventing effective purchasing and procurement processes. Respectively this study had three main aims and research questions. The first one was to discover what are the main advantages of implementing machine learning techniques into purchasing and procurement processes. Secondly, what are the key challenges of implementing machine learning techniques into purchasing and procurement processes. Thirdly, through interviews conducted with purchasing and procurement professionals, to learn what are the main challenges that prevent effective working and if machine learning techniques could provide help with these challenges.

8.2 Key findings

This chapter discusses the key findings of this research and they are separated into findings from literature and interviews. Later results are collectively interpreted.

8.2.1 Literature

From the literature different kinds of applications for pricing, spend analytics, and forecasting were mentioned to provide benefits for procurement process. Solutions to support sourcing are seen also as helpful. Furthermore, tracking of materials in transit is something that machine learning solutions are helping to support. Also, additional support for the sourcing process and supplier relationship management were mentioned as key advantages of machine learning implementations for purchasing and procurement. Additionally, the use of machine learning is mentioned to have compliance and ethical challenges to overcome as often large quantities of personal information is handled in companies.

In terms of challenges, data quality, volume and traceability are critical factors to consider. Moreover, ethical and compliance questions also provide issues in terms of correct implementation and the justification of investing resources in machine learning applications is currently seen as challenging. Justifying investment in machine learning applications is also considered difficult, even though managers appear to be interested in general, but the high failure rate of implementation projects and high prices are raising concerns about artificial intelligence and machine learning technology.

8.2.2 Interviews

Work tasks that require a notable amount of manual work and concentration or both were mentioned to be issues for purchasing process and therefore interrupt effective purchasing. For example, forecasting and order follow-up were mentioned to be one of these. Thus, in general lack of automatization and support information system applications were common themes. In addition, unclear work task assignments were also seen as an issue that extracts time from the daily schedule as the person must find out what needs to be done in to complete the task. Not being able to trust the data received by the system was also mentioned as an issue that disturbs daily work tasks.

In procurement challenges with material pricing and other issues related to prices were mentioned. Unstable market and inflation rates are affecting material prices and leading to evermore challenging price negotiation meetings with suppliers. Aside from meetings with suppliers, preparing for price discussions takes a significant amount of time and work. In general, supplier relationship management-related tasks require plenty of time. The lack of effective automatization solutions for price updating was also mentioned as an issue. Prediction of the future for example about long-term material volume requirements was brought up as a challenge for procurement professionals.

8.3 Interpretation of results

When comparing results from the interviews machine learning applications can potentially provide support for issues mentioned by responders. The need for less manual work on demand forecasting was mentioned by purchasing respondents. Also, the lack of trust for data received from the systems and the previously mentioned challenges of machine learning and its data quality issues must be pointed out here. Therefore, the data quality issues should be corrected before any implementation of machine learning solutions is done also in the company where the interviews were conducted or in any other company that faces similar kinds of issues with data. As previously stated, it is difficult to train machine learning algorithms to produce effective results if the training data used is of poor quality.

For procurement pricing estimation and spending analytics were mentioned as potential applications and correspondingly these same topics were mentioned in the interviews. This connection between practical issues and suggestions by the responders and existing literature indicates that there is a potential to implement machine learning applications to both purchasing and procurement processes. For instance, Samir (2014) mentioned safety stock level estimator is something that I can see useful in purchasing as it would release the burden of adjusting safety stock levels manually.

When comparing results and defined research questions both literature and interviews results provided useful content. Found advantages and disadvantages of machine learning implementation from the academic and industrial publications helped answer the research questions. Thus, the current situation of machine learning implementation in procurement and purchasing has been provided which was one main goal of this research. Also, the interview part and its results helped to answer the research question about issues that are preventing effective procurement and purchasing. Interviews answers provided essential empirical material for this research. Also, similarities were discovered when comparing interviewee responses and literature publications, as interviewers were pointing out concerns that were also highlighted in the literature as one of the areas where machine learning could bring advantages over more traditional technologies. This phenomenon of similarities raised the reliability of the results and collected literature content and helped to create a connection between the theoretical and empirical results of this research. The results of the literature follow similar issues as prior studies about machine learning and its applications in procurement and purchasing, which was expected given the number of sources on the subject. As a result, no significant differences were found when comparing the findings of this study to other publications.

A study by TealBook (2020) shows that one of the key issues of efficient procurement is poor data from suppliers as it influences negatively to decision-making processes, reaching deadlines, and weakens the ability to act in an agile way during supply chain turmoil. Poor supplier data was not mentioned directly during my interviews. This might be due to the reason that the questions did not mention directly if there are issues caused by suppliers. Global market disruptions and their effects on procurement were mentioned as one of the key issues during my interviews and this follows also other publications such as SupplyChain Management Review (2022). Also, these research interviews were done in the Autumn of 2022, which should be remembered given that the COVID-19 pandemic and global geopolitical situation had caused significant disruption in markets for several months at this point.

When comparing results between interviews and utilized resources it is important to keep in mind that each organization might have different information systems in use, for example, the ERP – software. Furthermore, even if identical ERP - software is used, the techniques on how different processes are carried out may differ greatly amongst firms. For example, notification processing issues raised during interviews may be handled differently in another organization and not cause difficulties to be mentioned. Thus, these factors must be taken into consideration when making wider conclusions based on this research's results and others which have been published around the same topic and with similar research approaches.

8.4 Limitations of the study

This study also has some potential limitations which are presented in this chapter. As most of the interviews were conducted without proper in-person interaction there is a lack of in-depth discussion about the topics and potentially missed opportunities to find key opinions and other insights. The decision not to mandate in-person interviews was made due to time resource limitations and to increase the likelihood of receiving answers. Nevertheless, all received answers provided useful insight into procurement and purchasing as was seen in the results.

The fact that the sample size of responders was only eight, limits the reliability of the study in terms of creating wider conclusions based on the answers. In addition, all responders are from the same company which limits even further the ability to create wider conclusions of the current state of purchasing and procurement as work methods and used systems vary across companies and organizations. Thus, it must be kept in mind that the received answers come from a small sample size which was selected using non-probability sampling from one company. There is a certain level of "sample bias" as the participants were not selected completely randomly. Nonetheless, expanding the number of responders, organizations and even companies would have required distinctly more time and resources which in this case was not feasible. Regardless of above mention limitations regarding sample size, results represent one example of purchasing and procurement professionals and how they see their current work tasks and issues related to them. Therefore, I believe this study can also provide information to purchasing and procurement professionals from other companies as they reflect these findings on their own experiences.

The lack of previous research studies on the topic can be found also as a limitation for this research. It was difficult to compare especially the interview results to others that had similar structure and scope. In fact, studies that included surveys were mostly focusing on procurement. Therefore, this research lacks comparisons specifically in the area of purchasing. Yet, this fact makes it more crucial to learn about purchasing issues through this research. Here I have to also point out that the lack of personal experience in research is a limitation of this study. For instance, this was the first time for me to conduct several interviews in an academic manner.

8.5 Recommendations for future study

For future research around this topic, I would recommend increasing the sample size of interview or survey participants distinctly. An increased number of participants would provide more viewpoints of different individuals and thus include wider demographics for the research. In addition, to the wider demographics, I would also recommend including participants from two or more companies to have a perspective on how answers differ between companies. Including more companies in the research would also then provide more usable results in general if the answers are following the same kind of themes even if more companies and participants are included.

Also, I would recommend narrowing down the research scope only either to procurement or purchasing. During the research, I noticed that utilized sources were addressing more procurement related themes and it was challenging to provide an equal amount of content about purchasing. Although I would still argue that this research and conducted interviews provided a useful amount of information from both procurement and purchasing, especially when most interview participants were from purchasing.

9 Conclusion

This chapter will conclude the study by summarizing the key findings regarding the objectives and research questions, as well as their value and contribution. Additionally, it will discuss the study's limitations and propose opportunities for future research.

The purpose of this study was to determine the key benefits and obstacles of implementing machine learning techniques into purchasing and procurement processes. The other key goal was to determine what the main factors are that limit efficient purchase and procurement processes and whether machine learning approaches may help with these challenges. The results indicate that solutions for pricing, spend analytics, forecasting and sourcing and supplier relationship management are to provide the most value if implemented. The main challenges for implementations are data quality, data volume and its traceability. In addition, ethical and compliance issues are to be taken in consideration, and in general the high costs and fail rate of machine learning implementation projects.

The high quantity of manual work, absence of automated processes, difficulties in relying on incoming data, and unclear work assignments are important findings of the obstacles that purchasing professionals currently face. Pricing issues, such as price negotiations, are the most common procurement obstacles. In addition, an unpredictable market makes forecasting future volumes and pricing problematic. In general, supplier relationship management takes a long time, and there looked to be a lack of automation and analytics solutions for it.

This study provides knowledge about the current advantages and challenges of implementing machine learning techniques into purchasing and procurement process. In this respect, no new knowledge was offered, but rather an overview of existing knowledge in the domain, which can therefore be utilized as support if organizations seek to implement machine learning techniques into the purchasing and or procurement processes. Furthermore, conducted interviews provided valuable information what are the key challenges of purchasing and procurement in reality and how potentially machine learning could help with these. Similarities were found between the literature and interviews regarding what are causing issues and what kind of machine learning solutions would help with them.

There are other research publications about purchasing and procurement as well as studies about machine learning and its solutions. Thus, this study follows along with the existing theories and practices. Though this study provides knowledge on how machine learning could provide value in purchasing and procurement based on both interviews and industry and academic publications. Even if the machine learning advantages and challenges are widely published this study provides an example of what kind of challenges purchasing and procurement professionals are facing and how machine learning could help with these challenges. Therefore, I argue that results and their interpretations of this study can be used when purchasing or procurement organization is considering whether it would be beneficial to implement machine learning techniques into their process as they can compare this study's results to their own process. In addition, this study provided knowledge on how some of the machine learning algorithms function, and this information is shared in a way that no deep prior knowledge is required and should be suitable for any intelligence layman. As a result, this study could serve as an information package for purchasing and procurement managers who are thinking about the use of machine learning techniques.

The main limitations of this study were the low amount of interview participants and sampling issues as participants were not selected randomly. Also, bias from the researcher can be mentioned as he has work experience in the fields of purchasing and procurement and this might have affected on questions and structure of the research. Also, with more time there would have been an opportunity to add more participants or even more companies to the interview scope. Lack of personal researching experience is seen also as a limitation.

For future research around this same topic, I would recommend extending the number of participants and also paying attention to the sampling methods to avoid any bias regarding it. In addition, doing a similar kind of interview as this study in different kinds of industries and comparing how they differ from each other, and if there are certain characteristics why or why not machine learning would be beneficial to consider as a supporting technology. Therefore, for future research around this same topic I would recommend reserving a generous amount of time as enlarging the scope of interviews would require it intrinsically. Also, I would recommend focusing either on purchasing or procurement as it would provide more time to focus deeper on one theme and provide possibly further detailed results.

References

Ageron, B., Bentahar, O., & Gunasekaran, A. (2020). Digital supply chain: Challenges and future directions. *Supply Chain Forum: An International Journal*, *21*(3), 133–138. https://doi.org/10.1080/16258312.2020.1816361

Aksoy, A., & Öztürk, N. (2011). Supplier selection and performance evaluation in just-in-time production environments. *Expert Systems with Applications*, *38*(5), 6351–6359. https://doi.org/10.1016/j.eswa.2010.11.104

Allal-Chérif, O., Simón-Moya, V., & Ballester, A. C. C. (2021). Intelligent purchasing: How artificial intelligence can redefine the purchasing function. *Journal of Business Research*, *124*, 69–76. https://doi.org/10.1016/j.jbusres.2020.11.050

Alpaydin, E. (2010). *Introduction to machine learning* (2nd ed). MIT Press.

Arbuzova, L. (2021, August 26). *Procurement vs. Purchasing: Key Differences Explained*. Precoro Procurement and Spend Management Blog. https://precoro.com/blog/procurement-vs-purchasing/

Ayodele, T. (2010). *Types of Machine Learning Algorithms*. https://doi.org/10.5772/9385

Baier, L., Jöhren, F., & Seebacher, S. (2019). CHALLENGES IN THE DEPLOYMENT AND OPERATION OF MACHINE LEARNING IN PRACTICE. *Research Papers*. https://aisel.aisnet.org/ecis2019_rp/163

Bailey, C., & Shantz, A. (2018). Building an Ethically Strong Organization. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/creating-an-ethically-strong-organization/

Bajari, P., Chu, C., & Park, M. (2008). An Empirical Model of Subprime Mortgage Default From 2000 to 2007. *National Bureau of Economic Research, Inc, NBER Working Papers*.

Banton, C. (2022, July 3). *Just-in-Time (JIT): Definition, Example, and Pros & Cons*. Investopedia. https://www.investopedia.com/terms/j/jit.asp

Barrad, S., & Valverde, R. (2019). Architecture for the Payment of Suppliers in the Supply Chain Through Web Services. *International Journal of Organizational and Collective Intelligence*, *9*, 18–36. https://doi.org/10.4018/IJOCI.2019100102

Bartolini, A. (2014, May 29). *Skills for the Modern Procurement Pro – Operational Procurement « CPO RISING – THE SITE FOR CHIEF PROCUREMENT OFFICERS & LEADERS IN SUPPLY MANAGEMENT*. https://cporising.com/2014/05/29/skills-for-the-modern-procurement-pro-operational-procurement/

Berner, C., Brockman, G., Chan, B., Cheung, V., Dębiak, P., Dennison, C., Farhi, D., Fischer, Q., Hashme, S., Hesse, C., Józefowicz, R., Gray, S., Olsson, C., Pachocki, J., Petrov, M., Pinto, H., Raiman, J., Salimans, T., Schlatter, J., & Zhang, S. (2019). *Dota 2 with Large Scale Deep Reinforcement Learning*.

Biedron, R. (2019, September 12). *Machine Learning in Procurement*. Planergy Software. https://planergy.com/blog/machine-learning-in-procurement/

Bodendorf, F., Merkl, P., & Franke, J. (2021). Intelligent cost estimation by machine learning in supply management: A structured literature review. *Computers & Industrial Engineering*, *160*, 107601. https://doi.org/10.1016/j.cie.2021.107601

Borges, M., Hoppen, N., & Luce, F. B. (2009). Information technology impact on market orientation in e-business. *Journal of Business Research*, *62*(9), 883–890. https://doi.org/10.1016/j.jbusres.2008.10.010

Bouchard, Y. (2019, May 29). *Tesla And Real World Reinforcement Learning At Scale (NASDAQ:TSLA) | Seeking Alpha*. https://seekingalpha.com/article/4267031-tesla-and-real-world-reinforcement-learning-scale, https://seekingalpha.com/article/4267031-tesla-and-real-world-reinforcement-learning-scale

Bourke, D. (Director). (2020, July 12). *2020 Machine Learning Roadmap (95% valid for 2022)*. https://www.youtube.com/watch?v=pHiMN_gy9mk

Bousqaoui, H., Said, A., & Tikito, K. (2017). *Machine learning applications in supply chains: An emphasis on neural network applications*. 1–7. https://doi.org/10.1109/CloudTech.2017.8284722

Brownlee, J. (2016). *Master Machine Learning Algorithms*. *1.1*, 163.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research*, *184*, 1140–1154. https://doi.org/10.1016/j.ejor.2006.12.004

CFI Education Inc. (2021). *Bagging (Bootstrap Aggregation)*. Corporate Finance Institute. https://corporatefinanceinstitute.com/resources/knowledge/other/bagging-bootstrap-aggregation/

Chandrasekaran, M. (2021, November 8). *Logistic Regression for Machine Learning*. Capital One. https://www.capitalone.com/tech/machine-learning/what-is-logisitic-regression/

Chen, A. (2021). Introduction to Machine Learning: Is AutoML Replacing Data Scientists? *ECloudvalley | AWS Premier Consulting Partner*. https://www.ecloudvalley.com/mlintroduction/

Chopra, Dr. A. (2019). *AI in Supply & Procurement* (p. 316). https://doi.org/10.1109/AICAI.2019.8701357

Codeless Platforms. (2019, March 22). *What Is The Difference Between Procurement And Purchasing? | Codeless Platforms*. https://www.codelessplatforms.com/blog/what-is-the-difference-between-procurement-and-purchasing/

Coursera. (2022, September 14). *3 Types of Machine Learning You Should Know*. Coursera. https://www.coursera.org/articles/types-of-machine-learning

DataCamp Team. (2022, March). *Gradient Descent Tutorial*. https://www.datacamp.com/tutorial/tutorial-gradient-descent

DeepAI. (2019, May 17). *Association Learning*. DeepAI. https://deepai.org/machine-learning-glossary-and-terms/association-learning

Deloitte. (2017). *The Deloitte Global CPO Survey 2017 | Deloitte | Operations*. Deloitte United Kingdom. https://www2.deloitte.com/tw/en/pages/operations/articles/gx-cpo-survey.html

Deloitte. (2019). *Global Chief Procurement Officer Survey 2019*. Deloitte Switzerland. https://www2.deloitte.com/ch/en/pages/operations/articles/cpo-survey.html

Deloitte. (2022). *Shaping sourcing and procurement in a post-COVID environment*. Deloitte Switzerland. https://www2.deloitte.com/ch/en/pages/technology/articles/shaping-sourcing-and-procurement-in-a-post-covid-environment.html

Delua, J. (2021, March 12). *Supervised vs. Unsupervised Learning: What's the Difference?* https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning

Determine (Director). (2018, March 28). *Webinar: How Artificial Intelligence will change Procurement and Contracting*. https://www.youtube.com/watch?v=yl5og-nLR3Q

Dilmegani, C. (2020, December 3). *AI Procurement: Why it matters & applications / use cases 2022*. https://research.aimultiple.com/ai-procurement/

Döbel, I., Leis, M., Vogelsang, M. M., Neustroev, D., Petzka, H., Riemer, A., Rüping, S., Voss, A., Vegele, M., & Welz, J. (2018). *Maschinelles Lernen: Kompetenzen, Forschung, Anwendung*. Fraunhofer-Allianz Big Data
und Künstliche Intelligenz. https://www.bigdata-ai.fraunhofer.de/de/publikationen/ml-studie.html

Dryden Group. (2022). *Common Challenges in Procurement | Dryden Group*. https://drydengroup.com/common-challenges-in-procurement/

Dubois, A., & Araujo, L. (2007). Case research in purchasing and supply management: Opportunities and challenges. *Journal of Purchasing and Supply Management*, *13*(3), 170–181. https://doi.org/10.1016/j.pursup.2007.09.002

Enríquez García, H. C., & Roque Huerta, E. (2020). Machine Learning and Deep Learning Patentable Developments and Applications for Cost Reduction in Business and Industrial Organizations. *ANWESH: International Journal of Management & Information Technology*, *5*(2), Article 2.

EYVO. (2021). *Is There A Future For Machine Learning In Procurement? | eyvo*. https://www.e-procurement.com/machine-learning-and-procurement/

Forrester. (2022). *Operationalizing Machine Learning Achieves Key Business Outcomes*. https://ecm.capitalone.com/WCM/tech/forrester-pdfs/capitaloneforrestersnapshotmloctober2022.pdf

Galletta, A., & Cross, W. E. (2013). *Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication*. New York University Press. http://ebookcentral.proquest.com/lib/tritonia-ebooks/detail.action?docID=1187368

Gavrilova, Y. (2020, August 4). *AI vs. ML vs. DL: What's the Difference*. Serokell Software Development Company. https://serokell.io/blog/ai-ml-dl-difference

Gezgin, E., Huang, X., Samal, P., & Silva, I. (2017). Raising supply-chain performance to new levels. *Digital Transformation*.

Google Developers. (2022, July 18). *K-Means Advantages and Disadvantages | Machine Learning*. Google Developers. https://developers.google.com/machine-learning/clustering/algorithm/advantages-disadvantages

Gudimalla, V. (2021, August 26). Concept of Gradient Descent in Machine Learning. *Medium*. https://varshithagudimalla.medium.com/concept-of-gradient-descent-algorithm-in-machine-learning-44f587ac16ac

Hao, K. (2018a, October 11). *What is AI? We drew you a flowchart to work it out*. MIT Technology Review. https://www.technologyreview.com/2018/11/10/139137/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/

Hao, K. (2018b, October 17). *What is machine learning?* MIT Technology Review. https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/

Hemant, G. (2021, February 3). *Machine Learning Use Cases in Procurement | LinkedIn*. https://www.linkedin.com/pulse/machine-learning-use-cases-procurement-hemant-gupta/

Hoffman, K. (2020, November 29). Random Forest Hyperparameters Explained. *Medium*. https://ken-hoffman.medium.com/random-forest-hyperparameters-explained-8081a93ce23d

Hughes, J. (2009). *CIO Leadership Interview: Maximising the Value of Supplier Relationships*. https://info.vantagepartners.com/insights/cio-leadership-interview-maximising-the-value-of-supplier-relationships

Hurwitz, J., & Kirsch, D. (2018). *Machine Learning For Dummies®, IBM Limited Edition*. 75.

IBM Cloud Education. (2021, March 6). *What is Overfitting?* https://www.ibm.com/cloud/learn/overfitting

Issar, R. (2020, October 9). Supplier Name Standardization using Unsupervised Learning. *Analytics Vidhya*. https://medium.com/analytics-vidhya/supplier-name-standardization-using-unsupervised-learning-adb27bed9e0d

JavaTpoint. (n.d.-a). *Association Rule Learning—Javatpoint*. Www.Javatpoint.Com. Retrieved 23 January 2023, from https://www.javatpoint.com/association-rule-learning

JavaTpoint. (n.d.-b). *Logistic Regression in Machine Learning—Javatpoint*. Retrieved 24 January 2023, from https://www.javatpoint.com/logistic-regression-in-machine-learning

Javatpoint. (2021). *Linear Regression in Machine learning—Javatpoint*. Www.Javatpoint.Com. https://www.javatpoint.com/linear-regression-in-machine-learning

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954–2965. https://doi.org/10.1111/jan.13031

Karamizadeh, S., Abdullah, S. M., Halimi, M., Shayan, J., & Rajabi, M. javad. (2014). Advantage and drawback of support vector machine functionality. *2014 International Conference on Computer, Communications, and Control Technology (I4CT)*, 63–65. https://doi.org/10.1109/I4CT.2014.6914146

Khan, S. A. R., Yu, Z., Khan, S. A. R., & Yu, Z. (2019). Introductory Chapter: Purchasing and Supply Management. In *Modern Perspectives in Business Applications*. IntechOpen. https://doi.org/10.5772/intechopen.85380

Khuan, L. S. (2019). *THE CHALLENGES OF EMERGING TECHNOLOGIES: THE EXPERIENCE OF PROCUREMENT PROFESSIONALS*. *9*. https://journal.oscm-forum.org/journal/proceeding/download_paper/20191207215718_OSCM_2019_paper_109.pdf

Kiefer, D., Ulmer, A., & van Dinther, C. (2019). *Application of Artificial Intelligence to optimize forecasting capability in procurement* (p. 80). https://doi.org/10.5281/zenodo.3539397

Kıran, M., Eşme, E., Torğul, B., & Paksoy, T. (2020). *Supplier Selection with Machine Learning Algorithms* (pp. 103–125). https://doi.org/10.1201/9780429327636-11

Kothari, C., R. (2004). *Research methodology: Methods and techniques (2nd revised edition)* (2nd ed., Vol. 2004). http://14.99.188.242:8080/jspui/bitstream/123456789/12457/1/Research%20Methodology%20-%20Methods%20and%20Techniques%202004.pdf

Lau, H. C. W., Lee, W. B., & Lau, P. K. H. (2001). Development of an intelligent decision support system for benchmarking assessment of business partners. *Benchmarking*, *8*(5), Article 5. http://dx.doi.org.proxy.uwasa.fi/10.1108/14635770110410389

Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, *63*(2), 157–170. https://doi.org/10.1016/j.bushor.2019.10.005

L'Heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. M. (2017). Machine Learning With Big Data: Challenges and Approaches. *IEEE Access*, *5*, 7776–7797. https://doi.org/10.1109/ACCESS.2017.2696365

Little, A. D. (2014). Big Data in Procurement. *Operations & Innovation Management Insight*. https://www.adlittle.com/sites/default/files/viewpoints/ADL_2014_Big_Data_in_Procurement.pdf

Liu, R., Nageotte, F., Zanne, P., De Mathelin, M., & Dresp, B. (2021). Deep reinforcement learning for the control of robotic manipulation: A focussed mini-review. *Robotics*, *10*(1), 22. https://doi.org/10.3390/robotics10010022

Liu, Y., Wang, Y., & Zhang, J. (2012). New Machine Learning Algorithm: Random Forest. In B. Liu, M. Ma, & J. Chang (Eds.), *Information Computing and Applications* (pp. 246–252). Springer. https://doi.org/10.1007/978-3-642-34062-8_32

Logistiikan Maailma. (2023). *Procurement and Purchases – Logistiikan Maailma*. https://www.logistiikanmaailma.fi/en/buyingselling/procurement-and-purchases/

Lopez Yse, D. (n.d.). *Introduction to K-Means Clustering*. Pinecone. Retrieved 25 January 2023, from https://www.pinecone.io/learn/k-means-clustering/

Magal, S., & Word, J. (2011). *Integrated Business Processes with ERP Systems*.

Manjrekar, O. N., & Dudukovic, M. P. (2019). Identification of flow regime in a bubble column reactor with a combination of optical probe data and machine learning technique. *Chemical Engineering Science: X*, *2*, 100023. https://doi.org/10.1016/j.cesx.2019.100023

Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, *1*(4), Article 4. https://doi.org/10.38094/jastt1457

McCrea, B. (2019, June 11). *What's the Role of Machine Learning in Procurement?* Source Today. https://www.sourcetoday.com/supply-chain/article/21867474/whats-the-role-of-machine-learning-in-procurement

McKinsey. (2017, April 21). *Smartening up with artificial intelligence | McKinsey*. https://www.mckinsey.com/industries/semiconductors/our-insights/smartening-up-with-artificial-intelligence

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and E-Business Management*, *16*(3), 547–578. https://doi.org/10.1007/s10257-017-0362-y

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.ISBN 978-0-07-042807-2.

Mitchell, T. M. (2005, September 13). *DSS News: Vol. 3, No. 23*. Machine Learning, Decision Trees, Overfitting. http://www.dssresources.com/newsletters/66.php

Monteiro, J. P., Ramos, D., Carneiro, D., Duarte, F., Fernandes, J. M., & Novais, P. (2021). Meta-learning and the new challenges of machine learning. *International Journal of Intelligent Systems*, *36*(11), 6240–6272. https://doi.org/10.1002/int.22549

Mori, J., Kajikawa, Y., Kashima, H., & Sakata, I. (2012). Machine learning approach for finding business partners and building reciprocal relationships. *Expert Systems* *with Applications*, *39*(12), Article 12. https://doi.org/10.1016/j.eswa.2012.01.202

Mummert, T., Subramanian, D., Vu, L., progress, N. P. P. 14 syyskuuta 2022 T. save your, & now, you must be logged in to I. D. L. in. (2022, September 15). *What is reinforcement learning?* IBM Developer. https://developer.ibm.com/what-is-automated-ai-for-decision-making

Nagar, D., Raghav, S., Bhardwaj, A., Kumar, R., Singh, P., & Sindhwani, R. (2021). Machine learning: Best way to sustain the supply chain in the era of industry 4.0. *Materials Today: Proceedings*, *47*. https://doi.org/10.1016/j.matpr.2021.01.267

Navada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011). Overview of use of decision tree algorithms in machine learning. *2011 IEEE Control and System Graduate Research Colloquium*, 37–42. https://doi.org/10.1109/ICSGRC.2011.5991826

Nazeer, K. A. A., & Sebastian, M. P. (2009). *Improving the Accuracy and Efficiency of the k-means Clustering Algorithm*.

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, *24*(12), Article 12. https://doi.org/10.1038/nbt1206-1565

Oladipupo, T. (2010). Types of Machine Learning Algorithms. In Y. Zhang (Ed.), *New Advances in Machine Learning*. InTech. https://doi.org/10.5772/9385

Opperman, V. (2022). *Artificial Intelligence*. Vic Opperman. https://www.vicopperman.com/artificial-intelligence/

Pan, X., You, Y., Wang, Z., & Lu, C. (2017). Virtual to Real Reinforcement Learning for Autonomous Driving. *ArXiv:1704.03952 [Cs]*. http://arxiv.org/abs/1704.03952

Priore, P., Ponte, B., Rosillo, R., & de la Fuente, D. (2019). Applying machine learning to the dynamic selection of replenishment policies in fast-changing supply chain environments. *International Journal of Production Research*, *57*(11), 3663–3677. https://doi.org/10.1080/00207543.2018.1552369

ProcurePort. (2022, May 25). *Top 6 Challenges in Procurement and How to Overcome Those*. ProcurePort Blog. https://blog.procureport.com/challenges-in-procurement/

Ranjan, R. (2021, July 19). *K-means Clustering and its applications | LinkedIn*. https://www.linkedin.com/pulse/k-means-clustering-its-applications-ritvik-ranjan/

Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35–39. https://doi.org/10.1109/COMITCon.2019.8862451

Reich, A. (2020, October 7). *Procurement vs Purchasing: Here's the Difference*. Order.Co. https://www.order.co/blog/procurement/procurement-vs-purchasing/

Rejeb, A., Keogh, J., & Edit, S. (2018). *Exploring New Technologies in Procurement*.

Samir, J. (2014, December 30). *Machine learning: A new tool for better forecasting*. https://www.supplychainquarterly.com/articles/933-machine-learning-a-new-tool-for-better-forecasting

Samuel, A. L. (1967). Some Studies in Machine Learning Using the Gameof Checkers. II-Recent Progress. *IBM Journal of Research and Development*, *11*(6), 601–617.

SAP. (2015). *Working with Tasks*. https://help.sap.com/doc/saphelp_nw73ehp1/7.31.19/en-us/47/54bc6492180924e10000000a42189c/frameset.htm

SAP Products & Services (Director). (2018a, October 31). *Intelligent ERP - SAP S/4HANA Invoice and Goods Receipt Reconciliation*. https://www.youtube.com/watch?v=FjA0KKJecpg

SAP Products & Services (Director). (2018b, December 4). *Intelligent ERP - SAP S/4HANA Predict Arrival of Stock in Transit*. https://www.youtube.com/watch?v=BQw8ADEbd48

Savcı, U. (2022, March 18). Machine Learning 101: Linear Regression. *Medium*. https://medium.com/@ugursavci/machine-learning-101-linear-regression-949b7faaebe0

Schmelzer, R. (2021, August 4). *Is Machine Learning Really AI?* Forbes. https://www.forbes.com/sites/cognitiveworld/2019/11/21/is-machine-learning-really-ai/

Schmitz, A. (2017, September 7). *Wie die Beschaffung von maschinellem Lernen profitiert*. SAP News Center. https://news.sap.com/germany/2017/09/ki-machine-learning-procurement/

Sengewald, J., & Lackes, R. (2022). *Prescriptive Analytics in Procurement: Reducing Process Costs*. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1059&context=wi2022

Sharma, A. (2020, May 11). Decision Tree vs. Random Forest—Which Algorithm Should you Use? *Analytics Vidhya*.

https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/

Sheinfield, R., & Forman, A. (2022, February 25). *How to Succeed in Procurement in 2022*. Bain & Company. https://www.bain.com/insights/how-to-succeed-in-procurement-in-2022/

Simpson Rochwerger, A. (2021, April 19). *The challenges of applied machine learning [Review of the book Real World AI: A Practical Guide for Responsible Machine Learnin, by Ben Dickson]*. TechTalks. https://bdtechtalks.com/2021/04/19/applied-machine-learning-challenges/

SourceDogg. (2021, September 10). The Importance of Effective Supplier Relationship Management. *SourceDogg*. https://www.sourcedogg.com/insight/the-importance-of-effective-supplier-relationship-management/

Steers, S. (2021, May 7). *Five common challenges faced by the procurement industry | Procurement Strategy*. Procurement. https://procurementmag.com/procurement-strategy/five-common-challenges-faced-procurement-industry

Sunil, R. (2017, September 8). Commonly Used Machine Learning Algorithms | Data Science. *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

SupplyChain Management Review. (2022, October 26). *Global survey shows procurement under pressure*.

https://www.scmr.com/article/global_survey_shows_procurement_under_pressure

Sutton, R. S., & Barto, A. G. (2014). *Reinforcement Learning: An Introduction*. 352.

Tanner, C., Wölfle, R., Schubert, P., & Quade, M. (2008). Current Trends and Challenges in Electronic Procurement: An Empirical Study. *Electronic Markets*, *18*(1), 6–18. https://doi.org/10.1080/10196780701797599

Taranenko, L. (2021, January 19). Machine learning in demand forecasting for retail. *Medium*. https://mobidev-biz.medium.com/machine-learning-in-demand-forecasting-for-retail-6ef35c502c93

Taulia. (2022). *What is strategic procurement?* Taulia. https://taulia.com/glossary/what-is-strategic-procurement/

TealBook. (2020, May 17). *Severe Business Consequences Stem from Poor Supplier Data, Survey Shows*. TealBook. https://www.tealbook.com/news/severe-business-consequences-stem-from-poor-supplier-data-survey-shows/

Team Data Science. (2019, December 28). A Layman's Guide to K Means Clustering—Machine Learning. *DATA SCIENCE*. https://datascience.eu/machine-learning/k-means-clustering-in-machine-learning/

Viale, L., & Zouari, D. (2020). Impact of digitalization on procurement: The case of robotic process automation. *Supply Chain Forum: An International Journal*, *21*(3), 185–195. https://doi.org/10.1080/16258312.2020.1776089

Vollmer, M. (2018, June 1). *Will Machine Learning Save Procurement Millions a Year?* Future of Sourcing. https://futureofsourcing.com/will-machine-learning-save-procurement-millions-a-year

Walch, K., & Schemlzer, R. (2018, March 28). AI Today Podcast #30: Interview with MIT Professor Luis Perez-Breva - Contrary Perspectives on AI and ML. *Cognilytica*. https://www.cognilytica.com/2018/03/28/ai-today-podcast-30-interview-with-mit-professor-luis-perez-breva-contrary-perspectives-on-ai-and-ml/

Wu, J. (2019, July 12). *AI, Machine Learning, Deep Learning Explained Simply*. Medium. https://towardsdatascience.com/ai-machine-learning-deep-learning-explained-simply-7b553da5b960

Yamusa, M., Ibrahim, Y., Ibrahim, A., Bala, K., Bello, A., Muhammad, A., Ahmadu, H., & Aminu, A. (2020, November 10). *A CONCEPTUAL FRAMEWORK FOR DEVELOPING MACHINE LEARNING MODELS FOR PROCUREMENT PROCESSES*.

Young, J. (2022, July 6). *What Is Procurement?* Investopedia. https://www.investopedia.com/terms/p/procurement.asp

Zhao, K., Tian, Y., & Deng, N. (2009). New Unsupervised Support Vector Machines. In Y. Shi, S. Wang, Y. Peng, J. Li, & Y. Zeng (Eds.), *Cutting-Edge Research Topics on Multiple Criteria Decision Making* (pp. 606–613). Springer. https://doi.org/10.1007/978-3-642-02298-2_89

Zhu, X., & Goldberg, A. B. (2009). Introduction to Semi-Supervised Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *3*(1), 1–130. https://doi.org/10.2200/S00196ED1V01Y200906AIM006

Appendices

Appendix 1. Interview questions

Thesis Interview Questions

1. **Do you work in purchasing or procurement?**

   ☐ Purchasing

   ☐ Procurement

2. **What is your job title?**

   Click or tap here to enter text.

3. **How many years of relevant work experience do you have either in purchasing or procurement?**

   Choose an item.

4. **What do you like most about your job?**

   Click or tap here to enter text.

5. **What are the three most time-consuming tasks in your job?**

   Click or tap here to enter text.

   Click or tap here to enter text.

   Click or tap here to enter text.

20.10.2022

Vaasan yliopisto

6. **What are the three most difficult tasks you encounter in your job? What is difficult about them, and why?**

   Click or tap here to enter text.

   Click or tap here to enter text.

   Click or tap here to enter text.

7. **If you could choose to have more support from information technology for a specific work task, which task would it be? (Supporting from Information technology could be by e.g., automatization, analytics, artificial intelligence)**

   Click or tap here to enter text.