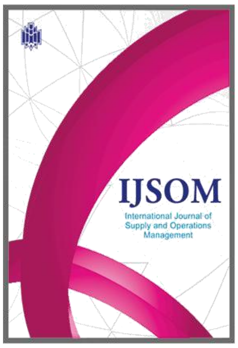
International Journal of Supply and Operations Management

IJSOM

November 2022, Volume 9, Issue 4, pp. 398-416
ISSN-Print: 2383-1359
ISSN-Online: 2383-2525
www.ijsom.com



Machine Learning in Supply Chain Management: A Systematic Literature Review

Mahraz Mohamed-Iliasse [a,*], Benabbou Loubna [b] and Berrado Abdelaziz [a]

[a] *Research team AMIPS, The Mohammadia School of Engineers, Mohammed V University in Rabat, Morocco*
[b] *Department of management sciences, University of Quebec at Rimouski, Levis, QC, Canada*

Abstract

The supply chain ecosystem is currently benefiting from a great dynamic resulting from the digitalization of organizations and trades. For all the stakeholders in the area, this is a real breakthrough, and machine learning is at the core of this revolution. It has profoundly revolutionized companies in many aspects including the evolution of communication methods, the automation of many processes, the growing importance of information systems, etc. With shrinking margins and more demanding customers, supply chain management in increasingly becoming a source of competitive advantage. Its management and optimization requires a factual to Supply Chain decision making at strategique, tactical and operational levels. In this context and data rich environment, machine learning approaches and techniques find numerous useful applications for supply chain decision making. Today, companies have no choice but to apply Machine Learning solutions in almost every part of their processes. This fact seems even clearer in markets where competition is fierce. While Machine Learning does not redefine the enterprise, it is certainly a powerful asset for both marketing and process optimization purposes. It is so ingrained in the strategies of companies that now most of them rely heavily on it for all processes from creation, to product quality control, to public relations. In recent years, a series of practical applications of machine learning (ML) for supply chain decisions have been introduced. By interconnecting the ML methods applied for SC decision making, this paper identifies current SC applications and indicates potential research gaps. In this paper, we examine the general usability of machine learning techniques to assist with supply chain decisions. The main objective of this research is therefore to study how machine learning techniques can be integrated into the range of tools available to supply chain decision makers in order to take advantage of the increasing volume of data generated in the supply chain.

Keywords: Supply chain; Supply network; Expert system; Machine Learning; Digital Transformation; Supply Chain Analytics.

1. **Introduction**

In this new environment, the digital transformation of the supply chain is needed today more than ever. The Digital Transformation has emerged as an important preoccupation and a key strategic matter for all kinds of organizations (Arribas, V., & Alfaro, J. A., 2018). Today, we are in a new era of digital transformation of the supply chain (SC), which is more than ever one of the main pillars of any industrial or commercial activity. An efficient and agile supply chain represents a major competitive advantage for companies.

The gradual transition to supply chain 4.0 allows for a more flexible and autonomous organization for companies seeking productivity. They are investing in new tools to empower and automate their production resources. Big data, connected objects and artificial intelligence are invading all levels of the supply chain.

The digital transformation of a value chain can only be as strong as its weakest link. It is therefore important that organizations apply Industry 4.0 principles to all levels of their production chain, including inventory management and, by extension, purchasing. It is therefore time for them to rethink their supply chain 4.0.

Corresponding author email address: m.mahraz@hotmail.fr
DO: 10.22034/ijsom.2021.109189.2279

The transition to digital offers company's new opportunities in terms of inventory management, production cost reduction, time savings, and more. Through an optimal alliance of Industry 4.0 solutions (connected objects, IoT, field data collection and analysis, etc.), organizations have to place the customer at the heart of the supply chain by adapting production to their needs. In other words, adopting a new strategy to digitalize the supply chain 4.0 is now a necessary step to understand the emerging technological landscape, to take advantage of its benefits, and thus to strengthen its position against the competition.

Digital technologies offer different solutions that meet the supply chain challenges of companies and market expectations (Kache, F., 2017). Numerous significant digitalization trends could be applied in a supply chain to significantly improve the future, and "Machine learning" (ML) is one of them. The ML has emerged from research labs to become ubiquitous in our private lives, and has increasingly taken over the enterprise. Like the agricultural and industrial revolutions that preceded it, this new revolution is redefining many aspects of modern life around the world.

Discovering new patterns in supply chain data has the potential to revolutionize any business. Machine learning algorithms find these new patterns in the data it has access to on a daily basis, without the need for manual intervention or the definition of a taxonomy to guide the analysis. The algorithms iteratively query the data with many patterns, using constraint-based modeling, to find the core set of factors with the highest predictive accuracy. Key factors that affect inventory levels, supplier quality, demand forecasting, procurement, cash ordering, production planning, transportation management and more are known and understood, for the first time. New knowledge and insights from machine learning are thus revolutionizing logistics management.

This work aims to provide an in-depth overview of possible applications in different parts of the supply chain, and thus analyse the challenges and benefits by integrating machine learning along the supply chain. It intend to provide a research agenda on Machine Learning in the supply chain. For that, we tracked the recent academic and trade literature to find research in this field.

The paper is based on a Systematic literature review by searching the "Google Scholar", "Scopus" and "Web of Science" databases under the search terms "Supply Chain" and "Machine Learning" to provide the reader with comprehensive and useful information on the areas of application and influence on the SC. In light of recent developments in machine learning, the results were restricted to publications from 2010 to 2020. In all, 40 relevant articles were identified.

2. Related methods

2.1. The Supply chain and their components

The term "supply chain" is well documented in the academic literature and is generally referred to as the alignment of companies that deliver products or services to the consumer market (Lambert, D. M. et al., 1998). Many researchers have studied the supply chain, but they do not observe it from a single angle (Jones & Riley (1985), Ayers,J. B.(2001), Mentzer, J.(2001), Chopra et al. (2007), Feniès (2006)). Each one proposes a definition according to the discipline from which it comes and the objectives which direct its analysis. Some definitions take a "product" viewpoint and others take a "business" or "process" viewpoint. One of the first definitions was proposed by (Jones & Riley, 1985) which defined the SC as being the planning and the piloting of the whole of the material flow since the supplier until the final customer while passing by the producer and the distributor. On their side, (Chopra & Meindl, 2001) define it as the whole of the activities impacting directly or indirectly the realization of the customer order. The supply chain does not only include producers and suppliers but also transporters, warehouses, retailers and the customers themselves.

In general, The Supply Chain is the process that is generated when a customer places an order until the product or service is delivered and paid for. Therefore, the Supply Chain includes the planning, execution and control of all activities related to the flow of materials and information, the purchase of raw materials, the intermediate transformation of the product and its delivery to the final customer. It is the set of interdependent companies (considered as the different links of the chain) coordinating in the realization of activities (supplies, production and distribution) to ensure the circulation of products or services from their conception to their end of life (after-sales service and withdrawal logistics). Customers' needs are changing; they expect more and more a service including a particular mode of delivery, replenishment, delay, reliability, security of supply, data transfer, after-sales service. This leads all the stakeholders to integrate more and more directly into the final consumer's sales act, as far as packaging, replenishment and forecasting modes are concerned, through data capture and direct transfer techniques. Nowadays, Supply Chains have become very complex due to internationalization, the increase in the types of flows and the evolution of global consumption patterns.

In principle, the objectives of Supply chain are: to anticipate and foresee customer requests; to process orders; to ensure the supply and storage of products and raw materials; to plan orders and deliveries. It include several component including:

- *Demand* Forecasting & Planning,
- Production Planning,
- Inventory management and Transportation management.

Based on these principles, it aims to minimize the costs of these components, namely storage, transportation and product availability. With good practice, quality supply chain can achieve significant results, such as "zero stock", just-in-time delivery and, above all, the absence of global stock shortages.

The supply chain can therefore be seen as a network of organizations, which is committed to efficiently coordinating the flow of all services at the lowest cost. All this is optimized with the use of information technologies, as well as by the adequacy of the flow of information between the different networks of organizations, which are located along the distribution chain of the product.

If companies are to survive in the coming years, they will need to manage their supply chain much more effectively than it is now. Many factors drive companies to change, among the main factors: - Continuous customer demand for better service, choicer and lower prices; more and more competitors; technologies continues to evolve what is available.

2.2. The transition to a supply chain 4.0

Having defined the concept of the supply chain, interest in recent years has focused on a phenomenon known as "supply chain 4.0". This can be explained as "the reorganization of supply chains through the use of Industry 4.0 technologies".

These technologies have emerged in the 21st century and are mainly implemented in high-income countries (Budak, A., et al., 2017). Given this concept, it is necessary to define Industry 4.0, from which the term "Supply Chain 4.0" derives. According to Lasi et al. (2014) the term was coined in Germany to refer to an impending fourth industrial revolution. This is characterized by a paradigm shift in manufacturing production, based on advanced digitalization of factories and wireless technologies in the context of 'smart objects'. According to the authors, the future vision of industry consists of modular and efficient production systems, which allow for the manufacture of goods in individual batches, but retain the economic conditions of mass production. This transformation is still in development, it is broadly defined as it lacks the historical perspective, and the literature has not yet managed to document it comprehensively. There is also a lack of econometric indicators that clearly show the impact of this fourth revolution. Indeed, while the common goal of industrial revolutions throughout history has been the acceleration of productivity, there is still no evidence of a fourth revolution on a global scale, at least not in a comprehensive way. However, it is possible that some industries are developing their own revolutions, and that this is happening in different parts of the world.

Moreover, the move towards the fourth industrial revolution is likely to take different forms in different sectors: while productivity gains are always expected from an industrial revolution, the fall in employment has not been the consequence in all cases. Industrial revolutions have also created new jobs through the use of new technologies (e.g. with the third industrial revolution came the internet, creating previously non-existent jobs such as software developer) and even new industries (The automotive industry). Many "old" technologies that are still widely used today are proving their stability and efficiency every day. Assessing each technology separately is complicated because their impact often depends on the application of other complementary technologies. This makes their impact almost zero, but very important when several technologies are implemented together.

As it is very difficult to anticipate all the possible uses and complementarities, in this paper we will focus on the use of machine learning in the new supply chain based on successful implementations on the different components of the supply chain.

2.3. Overview of Machine learning

Machine learning is a subset of artificial intelligence that allows an algorithm, software, or system to learn and adapt without being specifically programmed to do so. ML typically uses data or observations to form a computer model in which different patterns of data (combined with actual and predicted outcomes) are analyzed and used to improve the operation of the technology. Machine learning (ML) models, based on algorithms, are excellent at analyzing trends, spotting anomalies, and gaining predictive insights into massive data sets. These powerful features make it an ideal solution for addressing some of the key challenges in the supply chain industry. We can recognize three main categories of algorithms, which differ in the data they use in the learning or training phase, and the type of result that they provide. These families are characterized by different learning paradigms:

-Supervised learning in the context of artificial intelligence (AI) and machine learning, is a system that provides both input data and expected output data. The input and output data are labeled for classification to establish a learning base for further processing of the data. Supervised machine learning systems feed the learning algorithms with known quantities that will support future decisions (Kuo, R. J., & Li, P. S., 2016). The data used for supervised learning is a set of examples comprising pairs of input subjects and expected outputs (also called supervisory signals) (Makkar, S.,Devi,G. N. R., & Solanki, V. K., 2019). These models are more likely to make decisions that humans can relate to because they rely on human input. But with a retrieval-based approach, supervised learning systems have difficulty processing new information.

-Unsupervised Learning Unlike supervised learning, the unsupervised context is the one where the algorithm must operate from unannotated examples. It must automatically generate the categories to associate with the data submitted to it in order to recognize that a cat is a cat, a car is a car, as animals and humans are capable of doing (Makkar, S.,Devi,G.

N. R., & Solanki, V. K., 2019). The most common unsupervised learning problem is segmentation (or clustering) where we try to separate data into groups (category, class, cluster...): grouping images of cars, cats, etc. A lot of hope is put on anomaly detection for predictive maintenance, cybersecurity, but also early detection of diseases, etc. In general, the algorithm seeks to maximize the homogeneity of the data within the data groups and to form groups that are as distinct as possible: depending on the context, one chooses to use this or that algorithm to classify the data, for example, according to their density or their density gradient. In the case of anomaly detection, it is rather the extreme or atypical character of the values or of a pattern in the data that is sought. The underlying metric plays a key role in determining what is the norm and what deviates from it (Zhou,L., et al., 2017).

-Reinforcement learning Rather than telling a computer precisely how to solve a problem, Machine Learning teaches it to learn to solve a problem on its own. This field of study includes dozens of algorithms. They are also called trainable systems because these algorithms are able to make mathematical rules emerge from data by training themselves on the basis of examples, and then to apply these rules to new data by constantly improving with experience. Among the most common algorithms, we find SVM (Support Vector Machine), boosting, random forests, neural networks, Bayesian networks, etc. They operate in various contexts. They operate in various contexts: supervised, semi-supervised or unsupervised, in sequential or batch mode, by reinforcement, etc. They are "input-output" systems with an input (image, sound, text) and an output (such as the category of the object in the image, the word spoken, the subject of the text). All the tasks requiring to enter data and to classify them can thus be automated: it allows to equip computers or machines with systems of perception of their environment like vision, recognition of objects (faces, diagrams, natural languages, writing, syntactic forms... On the Internet, it allows to filter undesirable contents (spam), to order answers to a search, to make recommendations or to select interesting information for each user (search engines); to conceive systems of assistance to diagnoses, medical in particular, game programs, brain-machine interfaces, systems of detection of frauds to the credit card, financial analysis, classification of the sequences of DNA.

Due to the large number of machine learning techniques, we have tried to group them by family in order to allow a better analysis of the results. These families of techniques have been built with the help of other articles (Seif, G., 2018) (Makkar, S.,Devi,G. N. R., & Solanki, V. K., 2019). Table 2 gives an overview of the different techniques used in machine learning and their family. Table 1 Different techniques used in machine learning depending on their family

<table>
<tr><td rowspan="2">Supervised Learning (SL)</td><td>Classification</td><td>-K-Nearest Neighbor(KNN)
-Discriminant Analysis (DA)
-Naïve Bayes (NB)
-Logistic Regression (LR)
-Decision Trees (DT) (CART, ID3, C4.5...)
-Ensemble Learning techniques (Boosting, Bagging, Staking, Random Forest..)
- Support Vector Machines (SVM)
- Deep Learning/Neural Networks (NN): MCP, ADALINE, MLP, CNN, RNN, LSTM…</td></tr>
<tr><td>Regression</td><td>-Linear Regression (LR’)
-SVR,GPR
-Decision Trees (DT) (CART, ID3, C4.5...)
-Ensemble Learning techniques (Boosting, Bagging, Staking, Random Forest..)
- Deep Learning/Neural Networks (NN): MCP, ADALINE, MLP, CNN, RNN, LSTM…</td></tr>
<tr><td rowspan="2">Unsupervised Learning (UL)</td><td>Clustering</td><td>-Hierarchical methods: Agglomerative methods (Linkage methods, Median, Centroid) and divisive methods (Graph bicoloring)
-K-Means, Fuzzy C-Means, K-Medoids
-Neural Networks (Self Organizing Maps)
-Gaussian Mixture</td></tr>
<tr><td>Association Rule Minining</td><td>-Apriori
-FP-Growth</td></tr>
<tr><td>Reinforcement Learning (RL)</td><td>RL</td><td>-Value –Iteration
-Q-Learning</td></tr>
</table>

3. Material and Methods

To conduct this review, we used the principles of systematic reviews to obtain reproducibility and high quality results (Okoli & Schabram, 2010; Barbosa-Povoa et al, 2017). It is an academic method that uses a more formal and clearly established approach to reviewing the literature in a particular field. It is a "research search" that involves analyzing the theory, methodologies, and findings of already available qualitative research and then summarizing this evidence to generate a new understanding of the field. It is guided by a formulated question, identifies relevant studies, assesses their quality and summarizes the evidence using a specific methodology.

1- Propose a taxonomy for the use of machine learning in the supply chain ;
2- Organize the main concepts related to this field;
3- Present the main models of machine learning in supply chain;
4- Identify key challenges and future issues related to supply chain.

The protocol used for this research is that of Okoli & Schabram (2010) (Okoli, Chitu, and Kira Schabram, 2010) , through its stringent and clear approach with strict rules, it allowed us to collect all relevant and valuable evidence on machine learning in the supply chain

3.1. Planning : Research Identification

As a first step for our study, we specified a problem and reformulated the objectives in the form of clear, structured and unambiguous questions to guide this study. The RQ1 lists the major channels for targeting machine learning in the supply chain. The RQ2 is about identifying and linking the methods and the models currently in place. The RQ3 identifies the most commonly found terms, for the creation of a standardization and presentation of the proposed taxonomy. The RQ4 discusses the applications, taking into account the cases and data collected.

Table 2. Research questions

Identifier	Issue
RQ 1	What are the primary channels are target of research of Machine learning in the supply chain?
RQ 2	Is it feasible to build a taxonomy for the use of machine learning in the supply chain?
RQ 3	Which parts of the supply chain are most targeted by the application of machine learning and the most common Machine learning methods in supply chain?
RQ 4	Which approach is more valuable for the supply chain, Machine Learning or conventional approach?

3.2. Selection: search strategy

The second step was to identify the relevant literature resources and databases, as well as the search terms used to conduct our study. The Literature search has focused on scientific articles published in English between 2010 and 2020, and is investigated in the major online databases of peer-reviewed literature: "Google Scholar", "Scopus" and "Web of Science". The digital databases for the study were identified from prior papers and in the literature in the same or similar field of study. Scopus, Web of Science, and Google Scholar libraries were traditionally used by the authors, and all studies were selected from these databases. Using the digital databases, we conducted our research, but using different methods with a specific string configuration. This allowed us to generate a larger volume of results, even if the return of some publications are not focused on our issue. In the process of selecting the initial results, we did not impose any kind of restrictions.

The search method used in our study is called a Boolean search (AND/OR). We tried to combine different terms using the OR and AND logic operator. A set of relevant keywords were selected through previous articles and paper in the same field or with similar scope. The different compositions of terms were searched in keywords, abstracts, and title.

The search string used is: *“Machine” AND “ Learning” AND “Supply” AND ”Chain” OR “ Network”; “AI” AND “Supply” AND “Chain” OR “ Network”; “Expert” AND “System” AND “Supply” AND “Chain”.*

Table 3. Databases and search terms

Databases	Search terms/Keywords
Google Scholar	Machine + Learning + Supply + chain
Scopus	AI + Supply + chain
Science Direct	Expert + System + Supply+ chain
	Machine+ Learning + Supply + network
	AI + Supply + network

Table 4. Systematic review process

Institute for Scientific information – Google Scholar, Scopus, Web of Science	
Criteria	Filters
Restriction	Topic (Title, abstract, Keywords)
Documents type	Articles and conference proceedings
Language	English
Years	2010-2020

3.3. Selection*:* Article filtering phases and Quality appraisal

The application of the search string in the databases allowed us to Google scholar, Scopus, WOS with date filter 2010-2020 generate a large number of papers. We started the articles filtering phase, based on the criteria described on the table 5. To be included in the sample, the paper had to meet the objectives and requirements of our study as presented in the first section, namely the project had to deal with application cases and not just theoretical ones. It also had to deal with machine learning in the supply chain. The impact, whether positive or negative, was not considered as an inclusion criterion. The search resulted in 145 publications that were filtered to remove duplication. They were then filtered first on the abstract reading, and finally through a full reading when necessary based on the inclusion criteria listed below. No articles were removed due to the quality of the search method. Table 6 shows the results of the screening, with 40 papers being included in the search.

Table 5. The criteria of quality appraisal

Criteria	Description
Criteria 1	The publication year of the papers must be 2010 or above
Criteria 2	Remove editorial, Keynote, opinion, tutorial, workshop, summary report, poster or paper such as unpublished articles, master's theses, and books.
Criteria 3	Remove studies that their full text is not available and are not in English
Criteria 4	Remove the studies that are not related to Machine learning and supply chain, and to the research questions.

All the criteria were applied using OR logical operator between them.

Table 6. The results of the selection of papers

	#Initial	#Final
Machine + Learning + Supply + chain	75	26
AI + Supply + chain	20	2
Expert + System + Supply+ chain	12	5
Machine+ Learning + Supply + network	33	7
AI + Supply + network	5	0
Total	145	40

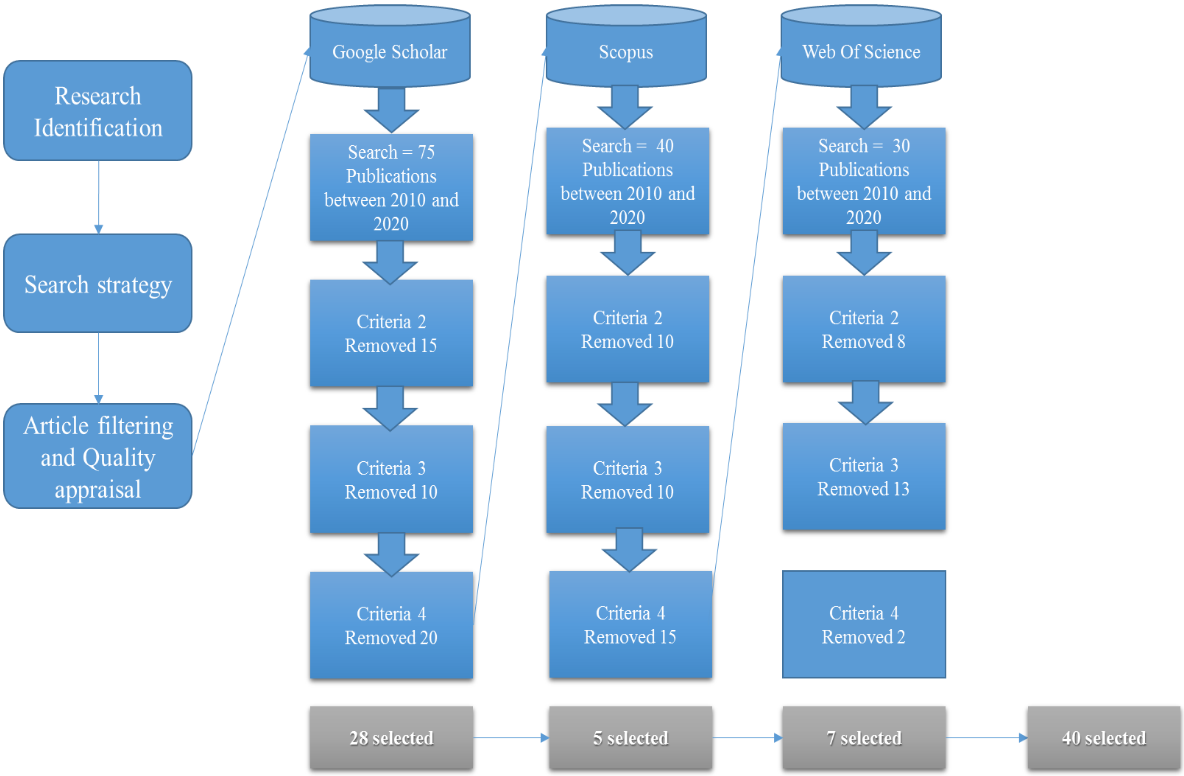


Figure 1. Article filtering phases

Table 7. Selected Articles by type and year

Type	Publication Year	Author	Publication Title
Conference Paper	2009	Tian, Jingwen; Gao, Meijuan; Zhou, Shiru	2009 International Conference on Computational Intelligence and Security
Journal Article	2010	Tamagawa, Dai; Taniguchi, Eiichi; Yamada, Tadashi	Procedia - Social and Behavioral Sciences
Journal Article	2011	Fasli, Maria; Kovalchuk, Yevgeniya	Information Sciences
Journal Article	2011	Huang, Jia-Yen; Tsai, Po-Chien	Journal of the Chinese Institute of Industrial Engineers
Journal Article	2011	Yu, Min-Chun	Expert Systems with Applications
Journal Article	2012	Ghorbani, Mazaher; Arabzad, S. Mohammad; Bahrami, Mahdi	Procedia - Social and Behavioral Sciences
Journal Article	2012	Beutel, Anna-Lena; Minner, Stefan	International Journal of Production Economics
Journal Article	2013	Fan, Xuemei; Zhang, Shujun; Wang, Longzhao; Yang, Yinsheng; Hapeshi, Kevin	Journal of Bionic Engineering
Journal Article	2013	Kartal, Hasan Basri; Cebi, Ferhan	International Journal of Machine Learning and Computing
Journal Article	2014	Gupta, Rajan; Pathak, Chaitanya	Procedia Computer Science
Journal Article	2014	Jaipuria, Sanjita; Mahapatra, S.S.	Expert Systems with Applications
Journal Article	2014	Ciupan, Emilia	Procedia Engineering

Table 7. Continued

Type	**Publication Year**	**Author**	**Publication Title**
Journal Article	2015	Mokhtarinejad, Maede; Ahmadi, Abbas; Karimi, Behrooz; Rahmati, Seyed Habib A.	Applied Soft Computing
Journal Article	2015	Mortazavi, Ahmad; Arshadi Khamseh, Alireza; Azimi, Parham	Engineering Applications of Artificial Intelligence
Journal Article	2015	Rana, Rupal; Oliveira, Fernando S.	Expert Systems with Applications
Journal Article	2016	Kocamaz, Uğur Erkin; Taşkın, Harun; Uyaroğlu, Yılmaz; Göksu, Alper	Computers & Industrial Engineering
Journal Article	2016	Kartal, Hasan; Oztekin, Asil; Gunasekaran, Angappa; Cebi, Ferhan	Computers & Industrial Engineering
Journal Article	2017	Fan, Zhi-Ping; Che, Yu-Jie; Chen, Zhen-Yu	Journal of Business Research
Journal Article	2017	Lolli, Francesco; Ishizaka, Alessio; Gamberini, Rita; Balugani, Elia; Rimini, Bianca	Procedia Manufacturing
Journal Article	2017	Vahdani, Behnam; Mousavi, S. Meysam; Tavakkoli-Moghaddam, Reza; Hashemi, Hassan	International Journal of Computational Intelligence Systems
Journal Article	2017	Budak, Aysenur; Ustundag, Alp; Guloglu, Bulent	Transportation Research Part A: Policy and Practice
Journal Article	2017	Qiu, Xueheng; Ren, Ye; Suganthan, Ponnuthurai Nagaratnam; Amaratunga, Gehan A.J.	Applied Soft Computing
Journal Article	2018	Blackhurst, Jennifer; Rungtusanatham, M. Johnny; Scheibe, Kevin; Ambulkar, Saurabh	Journal of Purchasing and Supply Management
Journal Article	2018	Kara, Ahmet; Dogan, Ibrahim	Expert Systems with Applications
Journal Article	2018	Villegas, Marco A.; Pedregal, Diego J.; Trapero, Juan R.	Computers & Industrial Engineering
Journal Article	2018	N Polyzotis, S Roy, SE Whang, M Zinkevich	ACM SIGMOD Record
Type	Publication Year	Author	Publication Title
Conference Paper	2018	P Stanula, A Ziegenbein, J Metternich	Procedia CIRP
Journal Article	2019	Q Min, Y Lu, Z Liu, C Su, B Wang	International Journal of Information Management
Conference Paper	2019	A Mayr, D Kißkalt, M Meiners, B Lutz, F Schäfer…	Procedia CIRP
Conference Paper	2019	H Tercan, A Guajardo, T Meisen	2019 IEEE 17th International Conference on Industrial Informatics
Journal Article	2019	J Krauß, M Frye, GTD Beck, RH Schmitt	Machine Learning for Cyber Physical Systems
Conference Paper	2019	T Sobottka, F Kamhuber, M Faezirad, W Sihn	Procedia Manufacturing
Journal Article	2020	Martínez, Andrés; Schmuck, Claudia; Pereverzyev, Sergiy; Pirker, Clemens; Haltmeier, Markus	European Journal of Operational Research
Journal Article	2020	R Cioffi, M Travaglioni, G Piscitelli, A Petrillo…	Sustainability

Table 7. Continued

Type	Publication Year	Author	Publication Title
Journal Article	2020	Z Kang, C Catal, B Tekinerdogan	Computers & Industrial Engineering
Journal Article	2020	K Guo, M Yang, H Zhu	Neural Computing and Applications
Journal Article	2020	S Thiede, A Turetskyy, T Loellhoeffel, A Kwade, S Kara…	CIRP Annals
Journal Article	2020	A Garre, MC Ruiz, E Hontoria	Operations Research Perspectives
Journal Article	2020	Y Li, S Carabelli, E Fadda, D Manerba, R Tadei…	The International Journal of Advanced Manufacturing Technology

4. Results and discussion

The results and discussions presented in this section are on the questions developed earlier.

RQ1: What are the primary channels are target of research of Machine learning in the supply chain?

In order to answer this question, we have tried to present a detailed analysis, first of the distribution of papers by publisher. Figure 2 allows us to highlight Elsevier as the main publisher with the most meaningful number of publications. The analysis of the occurrence of papers by year of publication and type of publication has allowed us to notice a significant increase over time, especially in the publications on scientific journals.

In figure 2, we can see the growth of the publications, emphasizing from 2018, which shows the permanent interest and the impact, positive or negative, of machine learning in the supply chain.

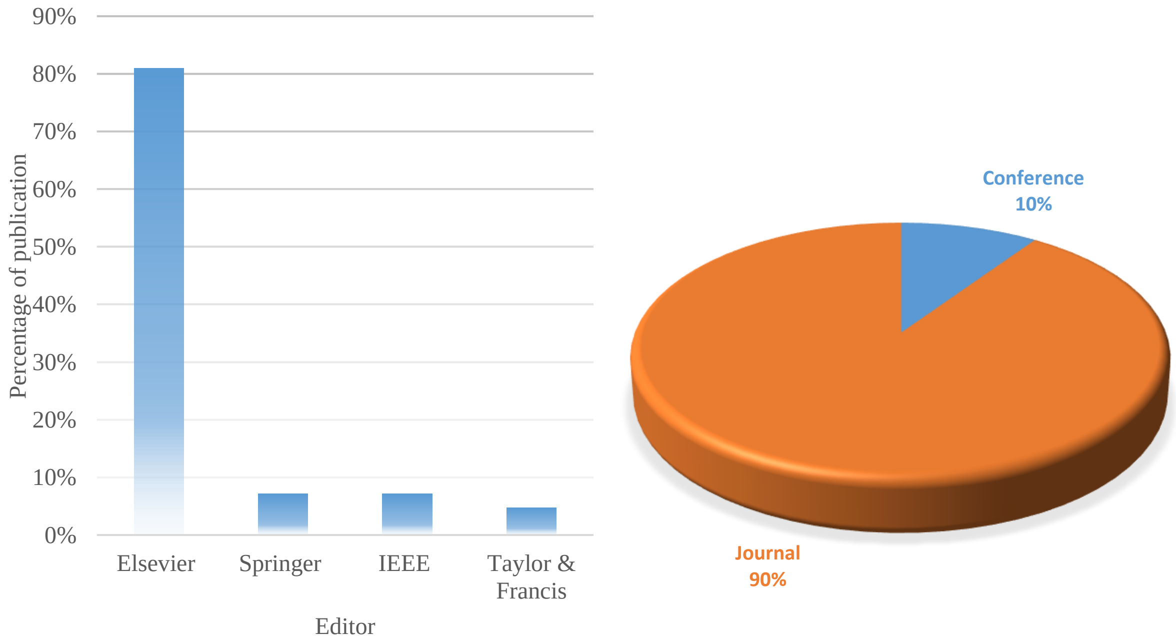


Figure 2. The distribution of papers by type and Editors

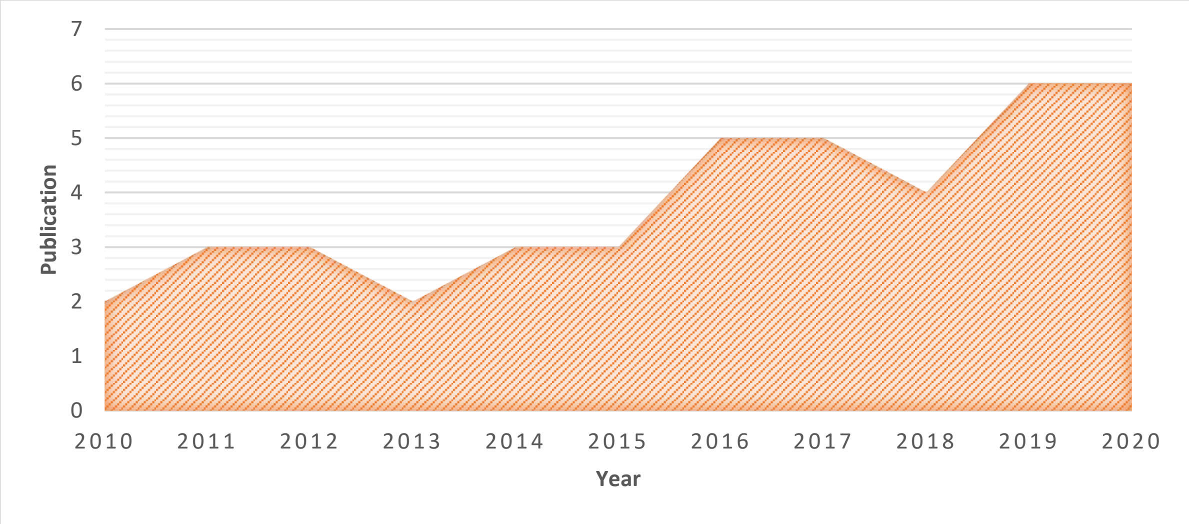


Figure 3. Distribution and tendency of publications by year

RQ2 : Is it feasible to build a taxonomy for the use of machine learning in the supply chain?

The first step in identifying our taxonomy was to generate a mapping and clustering using VOSviewer which is a program that allows to create maps of publications, authors, co-citation network, journals from the co-citation network, or to create maps of keywords from the co-occurrence network. VOSviewer is developed by the University of Leiden. It allows to create maps from network data using two methods, mapping and clustering. First, we had to generate a file containing the essential information to create the maps: authors, terms and citations. For this, we used the Google Scholar database. This first analysis allowed us to identify the links between words using a natural language algorithm. In order to avoid the reproduction of synonymous words on the map, we applied filters for the same terms (Table 4.). Subsequently, in order to identify the different terms that are hierarchically related in direct relation to machine learning and supply chain, we generated the relationship map and the heat map of VOSviewer.

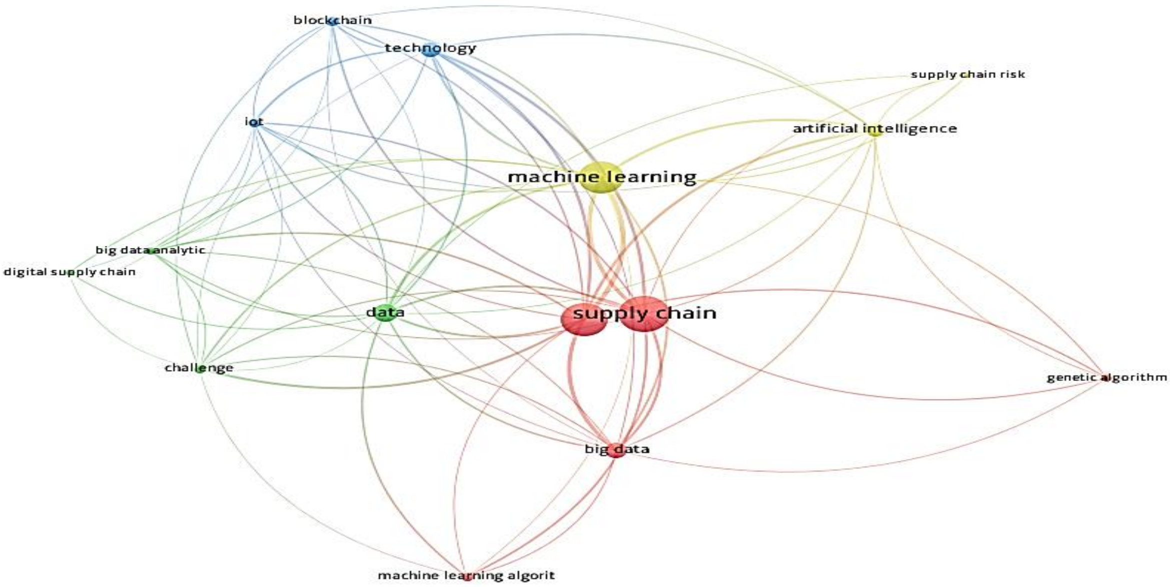


Figure 4. Relationship maps

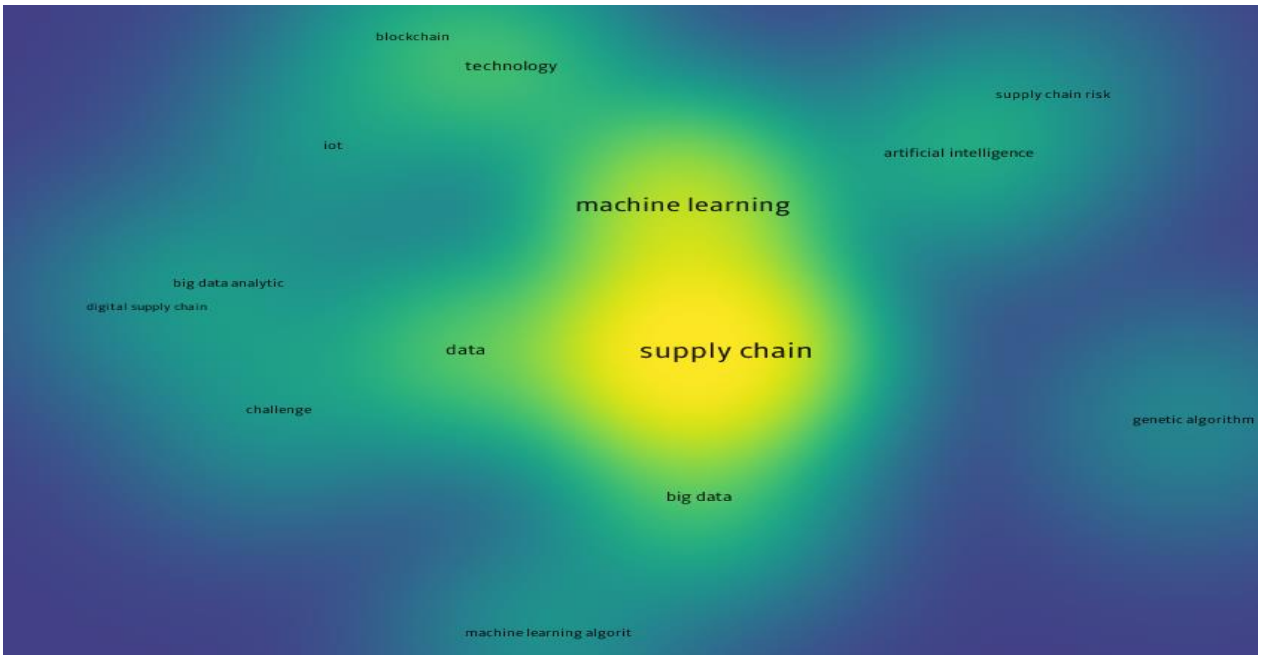


Figure 5. Heat map

The maps in Figure 4. refer to a relationship and 5 to the heat map. This allowed us to highlight the main terms found for machine learning with a focus on the supply chain which we have reproduced in table 8. In order to have a clearer view of the literature on which to build our taxonomy, we performed the mapping to key terms, and the results for each cluster related are presented in Table 10.

Table 8. The structure of the co-occurrence terms

Label	**Replace by**
Machine Learning technique	Machine Learning Algorithm
Internet of things	Iot
AI	Artificial Intelligence
Supply chain Network	Supply chain

Table 9. Key terms identified

Key Terms
Digital Supply chain
Supply chain risk
Big Data / Big Data analytic
Artificial Intelligence
Genetic Algorithm / Machine Learning Algorithm
Supply chain Network

Table 10. The main terms and related clusters

The main terms of the main clusters	**Clusters**
Supply chain	Big data, Genetic algorithm, artificial intelligence, supply chain risk, Iot, blockchain, Machine learning algorithm, Big data analytic, Digital Supply chain
Machine Learning	Digital Supply chain, Big data analytic, Iot, Blockchain, Technology, Big data, Genetic algorithm, artificial intelligence, supply chain risk, Supply chain
Technology	Iot, Blockchain, Big Data analytic, Artificial Intelligence, Machine Learning

RQ3: Which parts of the supply chain are most targeted by the application of machine learning and the most common Machine learning methods in supply chain?

The two most important characteristics of Machine Learning are its predictive capacity and its ability to incorporate many different sources of information. Indeed, the decision making in the Supply Chain is usually driven by sales forecasts. The method of elaboration depends on the level of the planning. Nevertheless, the main aspect of ML is that it is able to predict, and therefore make, the most accurate decisions in the future based on past data. Machine Learning allows for the development of models from high volumes of both internal and external data. After training and testing these models, they will be capable of predicting and adjusting to the evolution of the system. Furthermore, decision makers can use Machine Learning to find new data variables that are relevant to their business processes. From this perspective, Machine Learning improves the understanding of problems and enriches all the information used in decision making.

In order to analyze the different areas of use of machine learning in the supply chain, we linked each selected article to an application area, thus measuring its percentage of use depending. Figure 6. Presents the main areas of application of Machine Learning to decision making in the Supply Chain domain. Until recently, greater profitability in logistics was achieved by increasing volume and taking advantage of economies of scale to reduce costs. Now, that's not enough: it' s about making better decisions faster. The integration of Machine Learning into enterprise technologies has become a race to the bottom for many companies. Our literature review showed that 60% of articles focus on demand forecasting and planning, as well as production planning, and are benefiting from these advances and making huge leaps forward in performance.. Applied to this part of the supply chain, Machine Learning allows to project and analyze very finely the demand of its customers to anticipate and plan future demands. Also, 40% of the papers showed that companies are already exploiting the potential of data by integrating ML tools into inventory planning and management, and transportation. Through this question, we aimed also to concretely illustrate the different areas of use of machine learning in the supply chain. In order to answer this question; we have linked each selected article that are presented on the table 6 to an application area, thus measuring its percentage of use. Using the components of the supply chain that reflect the 3 levels (strategic, tactical and operational) and the main families of machine learning algorithms (supervised or unsupervised, classification and regression), we were able to give a detailed analysis of the literature.

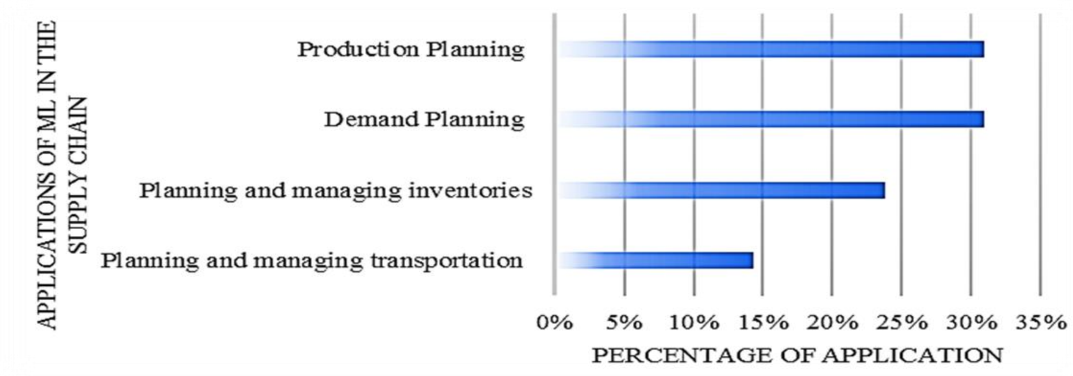


Figure 6. Applications of ML in the Supply chain by percentage of use

Table 11. ML applications in the supply chain by authors

Component of the supply chain	Identifiers
Demand Forecasting & Planning	Blackhurst, Jennifer, et al, 2018 ; Fan, X., et al.,2013 ; Fan, Zhi-Ping, et al, 2017 ; Fasli, Maria, Yevgeniya Kovalchuk. 2011 ; Gupta, Rajan et al, 2014 ; Jaipuria, Sanjita et al, 2014 ; Kocamaz, Uğur Erkin, et al., 2016 ; Kuo, R. J et al, 2016 ; Martínez, Andrés, et al, 2018 ; Qiu,X., et al.,2017 ; Rana, Rupal et al, 2015 ; Vahdani,B., et al.,2017 ; Villegas, Marco A. et al, 2018
Inventory Management & Planning	Beutel et al, 2012 ; Ciupan, Emilia, 2014 ; Huang, Jia-Yen et al, 2011 ; Kara, Ahmet et al, 2018 ; Kartal et al, 2013 ; Kartal, Hasan, et al, 2016 ; Lolli, Francesco, et al, 2017 ; Mortazavi et al,2015 ; Sui, Zheng. et al, 2010; Yu, Min-Chun, 2011
Production Management & Planning	Polyzotis et al, 2018 ; Stanula et al, 2018 ; Min, et al, 2019 ; Mayr, et al, 2019 ; Tercan et al, 2019 ; Krauß et al, 2019 ; Sobottka et al, 2019 ; Martínez et al, 2020 ; Kang et al, 2020 ; Guo et al, 2020 ; Thiede et al, 2020 ; Garre et al,2020 ; Li et al, 2020 ;
Transportation Management & Planning	Budak, Aysenur et al, 2017 ; Ghorbani, Mazaher et al, 2012 ; Tibor, Žácik et al, 2018 ; Mokhtarinejad et al. (2015) ; Dai, Tamagawa et al., 2010 ; Jingwen, Tian et al., 2010

- **Production Management & Planning**

Production lines are no longer linear sequences of operations, and they have become progressively more difficult to optimize. Complexities are numerous and streamlining operations is more difficult than ever. Moreover, that does not even include information constraints, where knowledge and data are fragmented across organizational units (Yu, M. C., 2011). Currently, through machine learning algorithms that are able to explore new ways of thinking allow from past data to build a global vision in real time, creating and adapting models to the reality of production lines.

The power of machine learning algorithms allows predicting what will happen as a result of the present actions, even determining optimal rules and suggesting the best course of action to achieve the objective that the company has defined beforehand. With ML techniques, we notice a transition from a partial understanding of production processes with standard models to a global view that measures future consequences, taking into account time and space constraints on production lines (Min, Q, Lu, Y, Liu, Z, Su, C, & Wang, B, 2019)

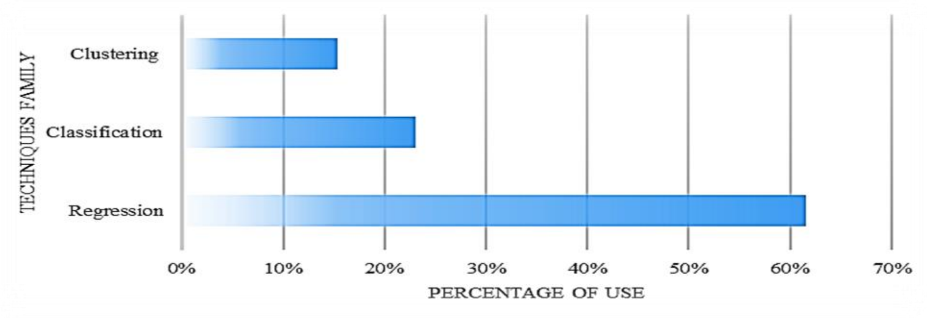


Figure 7. Percentage of application of techniques by techniques family in Production Management & Planning

Figure 7 demonstrates that regression is the main family technique, which represents about 63% of all the sample articles. Approximately 25% of the papers applied classification and 15% applied clustering.

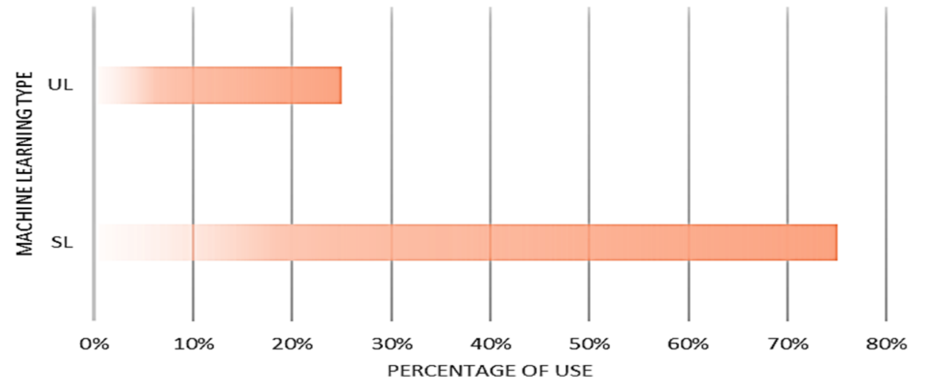


Figure 8. Percentage of application by ML type

According to Figure 8, supervised learning is the dominant type of machine learning applied to production lines. Some studies used both supervised and unsupervised learning methods.

- **Demand Forecasting & Planning**

In a highly competitive environment where innovation is permanent and customers are more volatile, demand is becoming more and more diversified, fragmented, complex and volatile, even as companies are faced with new economic, technical and regulatory pressures. Predicting and planning demand and understanding customer behavior allows for better planning of the company's flows. Predictive analysis becomes much more powerful thanks to Machine Learning (Gupta, R., & Pathak, C., 2014). Conventional demand forecasting tools are generally based on statistical analysis of historical sales data. This analysis, which is often aggregated and static, does not adapt to changing market conditions and consumer behaviors and requires a lot of human intervention. For supply chains, significant improvements in forecasting accuracy can translate into equally significant returns, enabling them to satisfy more customers, more quickly, and with less risk (Gupta, R., & Pathak, C., 2014).

The analysis of the results showed that regression methods are the most used in the forecasting and demand-planning part, even if some methods are more powerful than the conventional regression methods, this family includes some models that are of great interest considering their capacity to generate and their excellent representation of nonlinear relationships. In addition, they can be used effectively with unsupervised ML techniques and methods such as K-means.

The results also show that clustering techniques are also widely used. This is likely due to the fact that they can handle unlabeled and huge data in the forecasting domain. Figure 9 shows the percentage of technique application by family of techniques in demand planning.

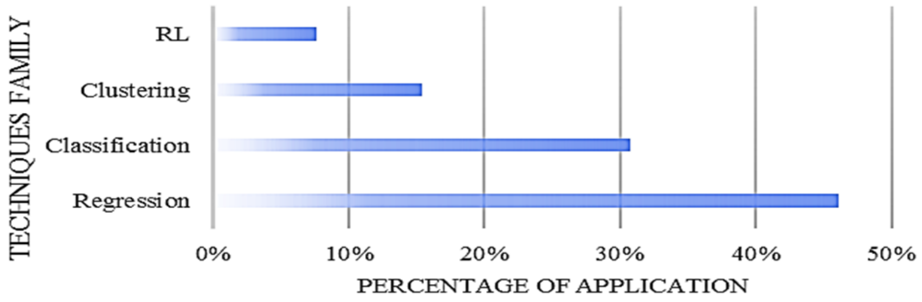


Figure 9. Percentage of application of techniques by techniques family in Demand Forecasting & Planning

As an article may use multiple ML techniques, it may refer to multiple learning types (according to the classification given in Table 2.) at the same time. The possible associations between the various learning types were considered. We have illustrated the results in Figure 9.

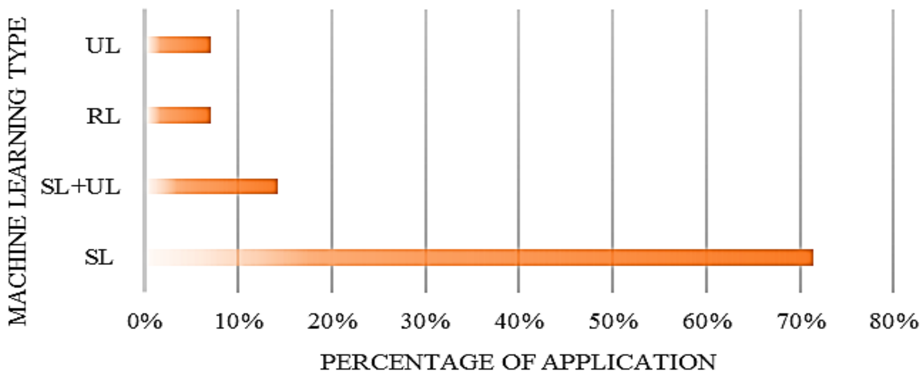


Figure 10. Percentage of application by ML type

- **Inventory planning & management**

The application of machine learning in a supply chain just for the purpose of automating and making reliable the forecast of demand for budgetary purposes is not sufficient because it is not the same exercise as a forecast intended to produce or supply goods. It is possible to forecast annual sales to within a few percent (Huang, J. Y., & Tsai, P. C, 2011) . It is much more difficult to forecast the weekly sales of thousands of goods individually.

Anticipatory inventory management is critical in very competitive sectors. It offers substantial cost savings in terms of logistics deployments and costs. It enables the better control of the inventory management chain (forecasting trend loss volumes, possible periods of turnover allowing to replay collections), based on inventory and sales data as well as other available external statistical data (such as economic data), or even the comparison of economic growth and decline of countries in the same situation as the company (Beutel, A. L., & Minner, S., 2012). In summary, this more in-depth analysis, based on current data but also on a prospective trend, enables the company to act on its inventories (of products, of raw materials) in an anticipatory manner to meet future trends that are not yet known.

The analysis of the literature has shown that classification methods are the most widely used in inventory planning and management. The aim of these techniques is to establish rules for classifying elements into clusters based on qualitative or quantitative variables that characterize these elements. The appearance of powerful algorithms for the classification of high-dimensional data, such as the support vector machine (SVM) or Naive-Bayes, has progressively transformed the field of conventional statistics. The latter was largely based on preprocessing by the human operator. Figure 11 showed the different types of learning used in the inventory planning & managing and Figure 12 present the percentage of application of ML type by percentage.

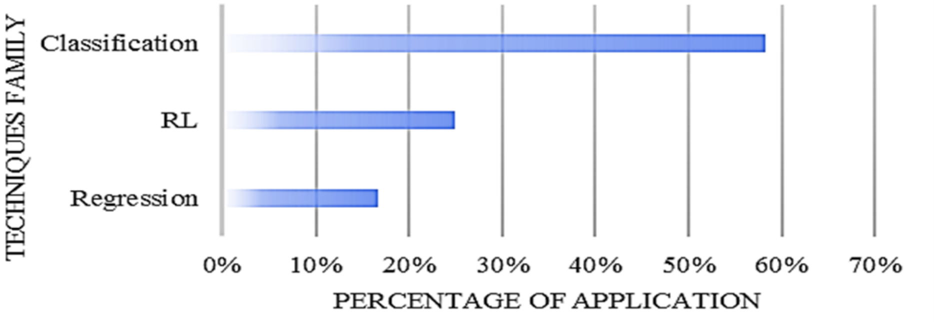


Figure 11. Percentage of application by techniques family in planning and managing inventories

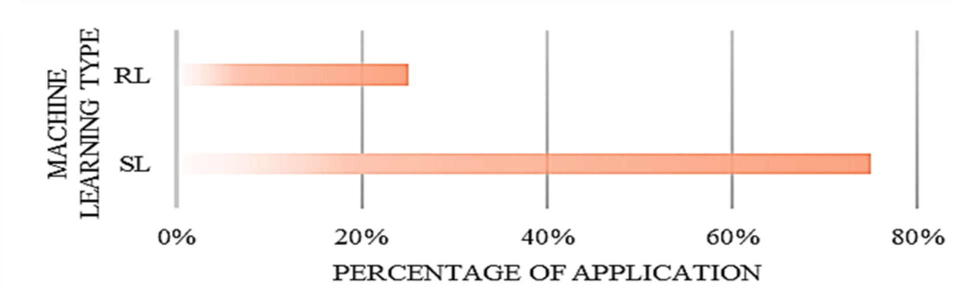


Figure 12. Percentage of application by ML type

According to Beutel et al, 2012 and Kara, Ahmet et al, 2018 the contribution of machine learning to inventory management can come as an advantageous substitute for businesses. This is because machine learning technology can identify patterns invisible to the human eye, such as complicated buying patterns that may fluctuate from year to year or season to season. Inventory is adjusted accordingly and the machine gains intelligence with each passing year.

- **Transportation planning & management**

The supply chain is in fact a chain like any other. Its robustness corresponds to that of its weakest link. A link that has not been given visibility and information in the right formats is likely to "break" and the whole chain will be affected. Transport management, considered as one of the most important functions in the supply chain, has a huge potential for savings, which can only be exploited, however, if the material requirement and transport processes are intelligently linked. Transport management is the most important task that can help a company to reduce costs and thus increase profitability. It is therefore necessary to identify the elements that make it up. This is where Machine learning comes in, as it can anticipate several things based on delivery times and points. By using optimization tools such as ours, adaptability is promoted. The company makes sure to maximize the filling capacities of the vehicles and to limit the unnecessary kilometers. All this has a positive impact on the production cost of the delivery. It can help to size fleets and vehicle fleets, promote better planning of drivers' working hours to optimize the entire process. Transport industry management is a complex process due to the huge volume of information that needs to be processed and the high quality of the data required.

However, it is possible today to predict trends using machine learning. These knowledge-based computing systems get to know the business and identify industry trends and consumer needs in an intelligent and efficient way that traditional analytics can only barely identify.

NNs is a powerful model for the transport industry because it can be used for a broad set of problems and does not require a deep knowledge of the processes involved in a given task. Its highlights also include the use of pattern recognition to link inputs to outputs. It can also process a large amount of data with good fit and performance when surrounded by noisy data. It saves time as it is a fast and efficient computational tool.

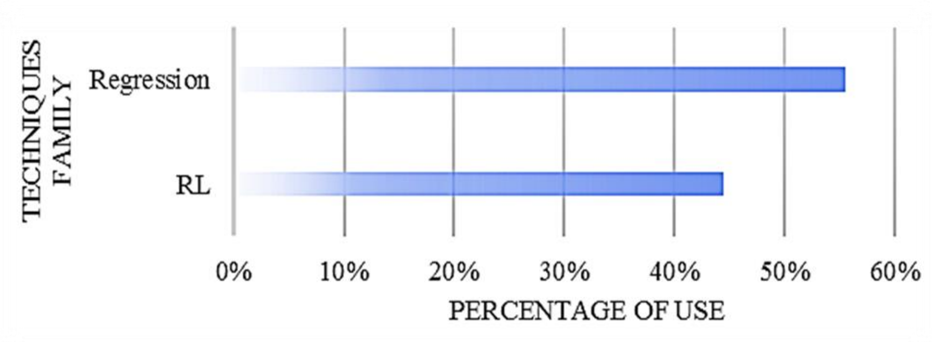


Figure 13. Percentage of application by techniques family in planning and managing Transportation

RQ 4: Which approach is more valuable for the supply chain, Machine Learning or conventional approach?

Machine learning is a very promising technology for supply chain optimization and also a major transformation lever. Indeed, the algorithms will allow to reach levels of accuracy that are currently impossible to imagine, which will allow gains and will give levers of improvement on the whole value chain.

Machine learning is concerned with the masses of data that can be used to highlight the elements on which it is possible to improve the CS. Based on algorithms, it is necessary for machine learning to have a sufficient quantity of data to analyze in order to translate it into useful indicators for decision-making. In the Supply Chain, the role of machine learning is therefore central and enables logistics to be significantly optimized, thereby improving profitability.

The selection of the approaches used should be based on the outcomes required and the completeness of the data, but it is also feasible to combine both approaches to achieve an accurate and comprehensive understanding of the relationships between the data and the observations, and to model these relationships while making the best predictions and getting the most out of your data.

The two main distinctions between a conventional supply chain decision-making approach and the utilization of Machine Learning are the ability of the latter to predict and its potential to integrate numerous sources of information. In fact, supply chain decision processes are normally guided by sales forecasts. The way in which they are developed depends on the different levels of planning. However, its main characteristic is to use historical data to make predictions and thus make the most pertinent decision in the future. Machine Learning offers the possibility of constructing patterns from vast quantities of external and internal data. They will then be in a position to expect system changes, and to adjust themselves accordingly. They will therefore be capable of improving sales forecasting processes, which feed all levels of decision making.

5. Conclusion

Improving supply chain efficiency plays a crucial role in any business. Operating their businesses in challenging profit margins, any type of process improvement can have a significant impact on the bottom line.

Innovative technologies such as machine learning make it easier to manage the challenges of volatility and accurately forecast demand in global supply chains. Studies predict that at least 50% of global companies in supply chain operations will use transformative technologies related to AI and ML by 2023. This speaks to the growing popularity of machine learning in the supply chain industry.

But, in order to take full advantage of machine learning, companies must plan for the future and begin investing in machine learning and related technologies today to benefit from increased profitability, efficiency, and resource availability in the supply chain industry.

References

Arribas, V., and Alfaro, J. A. (2018). 3D technology in fashion: from concept to consumer. *Journal of Fashion Marketing and Management: An International Journal*, Vol. 22(2), pp. 240-251.

Ayers, J. B. (2006). Handbook of supply chain management. Auerbach publications.

Becker, T., Illigen, C., McKelvey, B., Hülsmann, M., and Windt, K. (2016). Using an agent-based neural-network computational model to improve product routing in a logistics facility. *International Journal of Production Economics*, Vol. 174, pp. 156-167.

Beutel, A. L., and Minner, S. (2012). Safety stock planning under causal demand forecasting. *International Journal of Production Economics*, Vol. 140(2), pp. 637-645.

Blackhurst, J., Rungtusanatham, M. J., Scheibe, K., and Ambulkar, S. (2018). Supply chain vulnerability assessment: A network based visualization and clustering analysis approach. *Journal of Purchasing and Supply Management*, Vol. 24(1), pp. 21-30.

Budak, A., Ustundag, A., and Guloglu, B. (2017). A forecasting approach for truckload spot market pricing. *Transportation Research Part A: Policy and Practice*, Vol.97, pp. 55-68.

Cadavid, J. P. U., Lamouri, S., and Grabot, B., (2018). Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review, International Conference on Information Systems, Logistics and Supply Chain, July 2018, Lyon, France. Hal-01881362

Chen, I. J., and Paulraj, A. (2004). Towards a theory of supply chain management: the constructs and measurements. *Journal of operations management*, Vol. 22(2), pp. 119-150.

Chopra, S., and Meindl, P. (2007). Supply chain management. Strategy, planning & operation. In Das summa summarum des management (pp. 265-275). Gabler.

Ciupan, E. (2014). A study regarding the possibility of optimizing the supply batch using artificial neural networks. *Procedia Engineering*, Vol. 69, pp. 141-149.

Fan, X., Zhang, S., Wang, L., Yang, Y., and Hapeshi, K. (2013). An evaluation model of supply chain performances using 5DBSC and LMBP neural network algorithm. *Journal of Bionic Engineering*, Vol. 10(3), pp. 383-395.

Fan, Z. P., Che, Y. J., and Chen, Z. Y. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, Vol. 74, pp. 90-100.

Fasli, M., and Kovalchuk, Y. (2011). Learning approaches for developing successful seller strategies in dynamic supply chain management. *Information Sciences*, Vol. 181(16), pp. 3411-3426.

Ghorbani, M., Arabzad, S. M., and Bahrami, M. (2012). Applying a Neural Network algorithm to Distributor selection problem. *Procedia-Social and Behavioral Sciences*, Vol. 41, pp. 498-505.

Gupta, R., and Pathak, C. (2014). A machine learning framework for predicting purchase by online customers based on dynamic pricing. *Procedia Computer Science*, Vol. 36, pp. 599-605.

Huang, J. Y., and Tsai, P. C. (2011). Determination of order quantity for perishable products by using the support vector machine. *Journal of the Chinese Institute of Industrial Engineers*, Vol. 28(6), pp. 425-436.

J. Tian, M. GAO and S. Zhou (2009). The Research of Building Logistics Cost Forecast Based on Regression Support Vector Machine, 2009 International Conference on Computational Intelligence and Security, Beijing, , pp. 648-652.

Jaipuria, S., and Mahapatra, S. S. (2014). An improved demand forecasting method to reduce bullwhip effect in supply chains. *Expert Systems with Applications*, Vol. 41(5), pp. 2395-2408.

Jordan, M.I.,and Mitchell, T.M. (2015). Machine learning: Trends perspectives and prospects. Science, Vol.349, pp. 255–260.

Kache, F. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations and Production Management*, Vol. 37(1), pp.10-36.

Kara, A., and Dogan, I. (2018). Reinforcement learning approaches for specifying ordering policies of perishable inventory systems. *Expert Systems with Applications*, Vol. 91, pp. 150-158.

Kartal, H. B., and Cebi, F. (2013). Support vector machines for multi-attribute ABC analysis. *International Journal of Machine Learning and Computing*, Vol. 3(1), pp. 154.

Kartal, H., Oztekin, A., Gunasekaran, A., and Cebi, F. (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. *Computers & Industrial Engineering*, Vol. 101, pp. 599-613.

Kocamaz, U. E., Taşkın, H., Uyaroğlu, Y., and Göksu, A. (2016). Control and synchronization of chaotic supply chains using intelligent approaches. *Computers & Industrial Engineering*, Vol. 102, pp. 476-487.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., and Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*, Vol.13, pp. 8-17.

Kuo, R. J., and Li, P. S. (2016). Taiwanese export trade forecasting using firefly algorithm based K-means algorithm and SVR with wavelet transform. *Computers & Industrial Engineering*, Vol.99, pp. 153-161.

Lambert, D. M., Stock, J. R., and Ellram, L. M. (1998). Fundamentals of logistics management. McGraw-Hill/Irwin.

Liu ,C. , Shu, T., Chen, S., Wang, S., Lai, KK. and Gan L.(2016). An improved grey neural network model for predicting transportation disruptions. *Expert Systems with Applications*, Vol. 45, pp. 331-340

Lolli, F., Ishizaka, A., Gamberini, R., Balugani, E., and Rimini, B. (2017). Decision trees for supervised multi-criteria inventory classification. *Procedia Manufacturing*, Vol.11, pp. 1871-1881.

Makkar, S.,Devi,G. N. R., and Solanki, V. K. (2019). Applications of Machine Learning Techniques in Supply Chain Optimization. *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management*, pp. 861–869.

Martínez, A., Schmuck, C., Pereverzyev Jr, S., Pirker, C., and Haltmeier, M. (2018). A machine learning framework for customer purchase prediction in the non-contractual setting. *European Journal of Operational Research*, Vol. 281(3), pp. 588-596.

Min, H., 2009. Artificial intelligence in supply chain management: theory and applications. *International Journal of Logistics Research and Applications*, Vol. 13(1), pp. 13–39.

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

Mercier, S., and Uysal, I. (2018). Neural network models for predicting perishable food temperatures along the supply chain. *Biosystems engineering*, Vol. 171, pp. 91-100.

Min, H., 2009. Artificial intelligence in supply chain management: theory and applications. *International Journal of Logistics Research and Applications*, Vol. 13(1), pp. 13–39.

Min, H., 2009. Artificial intelligence in supply chain management: theory and applications. *International Journal of Logistics Research and Applications*, Vol. 13(1), pp. 13–39.

Mokhtarinejad, M., Ahmadi, A., Karimi, B., and Rahmati, S. H. A. (2015). A novel learning based approach for a new integrated location-routing and scheduling problem within cross-docking considering direct shipment. *Applied Soft Computing*, Vol. 34, pp. 274-285.

Mortazavi, A., Khamseh, A. A., and Azimi, P. (2015). Designing of an intelligent self-adaptive model for supply chain ordering management system. *Engineering Applications of Artificial Intelligence*, Vol. 37, pp. 207-220.

Qiu, X., Ren, Y., Suganthan, P. N., and Amaratunga, G. A. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied Soft Computing*, Vol. 54, pp. 246-255.

Rana, R., and Oliveira, F. S. (2015). Dynamic pricing policies for interdependent perishable products or services using reinforcement learning. *Expert Systems with Applications*, Vol. 42(1), pp. 426-436.

Rana, R., and Oliveira, F. S. (2015). Dynamic pricing policies for interdependent perishable products or services using reinforcement learning. *Expert Systems with Applications*, Vol. 42(1), pp. 426-436.

Raschka, S.,and Mirjalili, V.:Python Machine Learning, 2nd Ed. Packt Publishing, Birmingham, UK, 2 edition, 2017.

Samuel, A.L. (1959). Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3, 210-229.

Samuel, A.L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, Vol. 3, pp. 210-229.

Seif, G., 2018. The 5 Clustering Algorithms Data Scientists Need to Know.

Sui, Z., Gosavi, A., and Lin, L. (2010). A reinforcement learning approach for inventory replenishment in vendor-managed inventory systems with consignment inventory. *Engineering Management Journal*, Vol. 22(4), pp. 44-53.

Tamagawa, D., Taniguchi, E., and Yamada, T. (2010). Evaluating city logistics measures using a multi-agent model. *Procedia-Social and Behavioral Sciences*, Vol. 2(3), pp. 6002-6012.

The MathWorks, I., 2016. Introduction to Machine Learning.

Vahdani, B., Mousavi, S. M., Tavakkoli-Moghaddam, R., and Hashemi, H. (2017). A new enhanced support vector model based on general variable neighborhood search algorithm for supplier performance evaluation: A case study. *International Journal of Computational Intelligence Systems*, Vol. 10(1), pp. 293-311.

Villegas, M. A., Pedregal, D. J., and Trapero, J. R. (2018). A support vector machine for model selection in demand forecasting applications. *Computers & Industrial Engineering*, Vol. 121, pp. 1-7.

Witten, H. I., Eibe, F. and Hall, A. M., 2011. Data Mining: Practical Machine Learning Tools and Techniques: Elsevier.

Yu, M. C. (2011). Multi-criteria ABC analysis using artificial-intelligence-based classification techniques. Expert Systems with Applications, Vol. 38(4), pp. 3416-3421.

Žácik, T., Mracka, I., Hajossy, R., and Hycko, M. (2018, May). Reinforcement Learning in Gas Transport Control. In PSIG Annual Meeting. Pipeline Simulation Interest Group.

Zhou,L.,Pan,S.,Wang,J.,and Vasilakos, A.V.(2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, Vol. 237, pp. 350–361,

Okoli, Chitu and Schabram, Kira, A Guide to Conducting a Systematic Literature Review of Information Systems Research (May 5, 2010). Available at SSRN: https://ssrn.com/abstract=1954824 or http://dx.doi.org/10.2139/ssrn.1954824

Polyzotis, N, Roy, S, Whang, SE, and Zinkevich, M (2018). Data lifecycle challenges in production machine learning: a survey. *ACM SIGMOD Record*, Vol. 47(2), pp. 17-28.

Min, Q, Lu, Y, Liu, Z, Su, C, and Wang, B (2019). Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management*, Vol. 49, pp. 502-519

Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., and De Felice, F. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability*, Vol. 12(2), pp.492.

Stanula, P, Ziegenbein, A, and Metternich, J (2018). Machine learning algorithms in production: A guideline for efficient data source selection. *Procedia CIRP*, Vol. 78, pp. 261-266.

Mayr, A., Kißkalt, D., Meiners, M., Lutz, B., Schäfer, F., Seidel, R. and Franke, J. (2019). Machine Learning in Production–Potentials, challenges and exemplary applications. *Procedia CIRP*, Vol. 86, pp. 49-54.

Tercan, H, Guajardo, A, and Meisen, T (2019). Industrial Transfer Learning: Boosting Machine Learning in Production. In 2019 IEEE 17th International Conference on Industrial Informatics (INDIN) (Vol. 1, pp. 274-279). IEEE.

Kang, Z, Catal, C, and Tekinerdogan, B (2020). Machine learning applications in production lines: A systematic literature review. *Computers & Industrial Engineering*, Vol. 149, 106773.

Krauß, J, Frye, M, Beck, GTD, and Schmitt, RH (2019). Selection and Application of Machine Learning-Algorithms in Production Quality. In *Machine learning for cyber physical systems* (pp. 46-57). Springer Vieweg, Berlin, Heidelberg

Krauß, J., Dorißen, J., Mende, H., Frye, M., and Schmitt, R. H. (2019). Machine learning and artificial intelligence in production: Application areas and publicly available data sets. In *Production at the leading edge of technology* (pp. 493-501). Springer Vieweg, Berlin, Heidelberg.

Sobottka, T, Kamhuber, F, Faezirad, M, and Sihn, W (2019). Potential for machine learning in optimized production planning with hybrid simulation. *Procedia Manufacturing*, Vol. 39, pp. 1844-1853.

Thiede, S., Turetskyy, A., Loellhoeffel, T., Kwade, A., Kara, S., and Herrmann, C. (2020). Machine learning approach for systematic analysis of energy efficiency potentials in manufacturing processes: A case of battery production. *CIRP Annals*, Vol. 69(1), pp. 21-24.

Guo, K, Yang, M, and Zhu, H (2020). Application research of improved genetic algorithm based on machine learning in production scheduling. *Neural Computing and Applications*, Vol. 32(7), pp. 1857-1868.

Garre, A, Ruiz, MC, and Hontoria, E (2020). Application of Machine Learning to support production planning of a food industry in the context of waste generation under uncertainty. *Operations Research Perspectives*, Vol. 7, 100147.

Li, Y., Carabelli, S., Fadda, E., Manerba, D., Tadei, R., & Terzo, O. (2020). Machine learning and optimization for production rescheduling in Industry 4.0. *The International Journal of Advanced Manufacturing Technology*, *110*(9), 2445-2463.