Machine Learning Applications for Sustainable Manufacturing: A Bibliometric-based Review for Future Research

Purpose: The role of data analytics is significantly important in manufacturing industries as it holds the key to address sustainability challenges and handle the large amount of data generated from different types of manufacturing operations. The present study, therefore, aims to conduct a systematic and bibliometric-based review in the applications of machine learning (ML) techniques for sustainable manufacturing (SM).

Design/Methodology/Approach: In the present study, we use a bibliometric review approach that is focused on the statistical analysis of published scientific documents with an unbiased objective of the current status and future research potential of ML applications in sustainable manufacturing.

Findings: The present study highlights how manufacturing industries can benefit from ML techniques when applied to address SM issues. Based on the findings, a ML-SM framework is proposed. The framework will be helpful to researchers, policymakers and practitioners to provide guidelines on the successful management of SM practices.

Originality: A comprehensive and bibliometric review of opportunities for ML techniques in SM with a framework is still limited in the available literature. This study addresses the bibliometric analysis of ML applications in SM, which further adds to the originality.

Keywords: Sustainable manufacturing; Data analytics; machine learning; manufacturing systems; Industry 4.0; bibliometric review.

1. Introduction

The role of manufacturing industries in both developing and developed nations is critical. Industries are now working to achieve the 2030 sustainability goals. Therefore, there is a need for a strategic shift from the current manufacturing paradigm to manufacturing sustainability (Jamwal et al., 2021b; Machado et al., 2020; Malek & Desai, 2020; Sharma et al., 2020a). To maintain sustainability in the manufacturing sector, different types of knowledge-based systems have been developed in the last few years, i.e. intelligent scheduling, expert systems and fuzzy controllers (Machado et al., 2020). ML techniques can help industries to automate time-consuming processes of knowledge acquisition, which is an essential requirement in the development of knowledge-based system architecture for SM. The application of ML algorithms in manufacturing has increased in the last 15 years (Cavalcante et al., 2019). The increase in the data generated in the manufacturing industry is one of the major issues at present. This generated data is in a different format to that which is usually generated from the production line, machine tool parameters and environmental data (Garetti & Taisch, 2012). At present, industries are focusing on SM processes in their practices. These processes focus on the manufacturing of products and goods based on economically sound processes that reduce the negative impacts on the environment (Haapala et al., 2013). The implementation of SM practices in industries enhances employee, customer, community and product safety (Despeisse et al., 2012).

Recent studies on SM with ML algorithms are covering areas in SM that include production planning, total energy consumption, job shop scheduling, product design and sustainable machining (Tayal et al., 2020). The inclusion of emerging technologies such as IoT, artificial intelligence, data analytics, digital delivery services are influencing SM practices in the Industry 4.0 era (Jamwal et al., 2021a; Machado et al., 2020). Some studies have reported that the use of these advanced technologies in sustainable manufacturing results in minimizing total energy consumptions, reduced labours inputs, and better condition-based maintenance prediction (Wu et al., 2017). As of now, industries are focusing on the implementation of Industry 4.0 practices in their business practices. In this context, the main challenge is the generation of a large amount of data. The generated data that may be useful for future decision-making can be organized and understood through the use of data analytics tools (Anbesh Jamwal et al., 2020; Kumar et al., 2018).

Manufacturers are now applying AI and machine learning to improve efficiency, employee safety and enhance product quality. In the manufacturing industries, the ongoing maintenance of production lines and machinery results in major expenses, which have also a major impact on the bottom line of any asset reliant production operation (Sharp et al., 2018). To overcome this issue, industries are focusing on predictive maintenance. Advanced AI algorithms can be used for handling cognitive tasks and take real-time decisions in real-time environments (Giovannini et al., 2012). Both ML and deep learning algorithms are the subsets of AI. These are widely used in sustainable manufacturing in production planning and job shop scheduling problems (W. Cui et al., 2019).

Previously published studies show that there are opportunities for ML techniques to be applied in sustainable manufacturing, i.e. sustainable planning scheduling and predictive maintenance (Li & Lu, 2017). But a comprehensive review and bibliometric analysis which reports the opportunities of ML techniques in sustainable manufacturing are still missing in the available literature. In the past, studies reviews have been done on machine learning for traditional manufacturing systems but there is no review work on the applications of ML algorithms in sustainable manufacturing. In fact, in March 2019 Hendrik Fink, Head of Sustainability services at the PricewaterhouseCoopers, stated "*If we properly incorporate artificial intelligence, we can achieve a revolution with regards to sustainability. AI will be the driving force of the fourth industrial revolution"* (Majorel Deutschland GmbH Artificial Intelligence and Sustainability,2020).

The subfields of AI technologies such as machine learning and data mining have also become an important research area in manufacturing engineering. According to the above-discussed consideration in the present study, a systematic review and bibliometric analysis are done for articles on sustainable manufacturing in which ML techniques are used. This will help manufacturers to understand how ML techniques support manufacturing processes in their optimization, which will result in achieving sustainability in the manufacturing practices. The systematic review and bibliometric analysis are conducted with the following research questions in mind:

RQ1: What is the research publication trend with its citation structure in this research area?

RQ2: What work has been done in this area and what are future research domains?

The remaining of the paper is structured as follows; Section 2 presents a discussion on manufacturing and sustainability; Section 3 introduces the challenges to sustainable manufacturing; Section 4 discusses the application of ML in manufacturing; Section 5 presents machine learning algorithms; Section 6 reviews the methodology, systematic literature review methodology and bibliometric analysis. Section 7 of the paper presents its findings and based on them, a framework is proposed. Section 8 discusses the implication of the study. Finally, Section 9 presents the contributions, conclusion and limitations of the present study with some future scopes.

2. Concepts used in the study

In the present section, we discuss the role of sustainability in manufacturing, challenges to sustainable manufacturing and opportunities for ML techniques in manufacturing. We also discuss the different ML algorithms.

2.1 Manufacturing and Sustainability

Manufacturing is important for societies as it supports the GDP of nations and hence has a major contribution to the world's economy. The definition of manufacturing with regards to industrial operation can be defined as:

"activities related to the mobilization of resources (banking, finance, engineering, recruiting and training), the conception and innovation of products and processes (R&D, design, engineering), the actual organization and management of production (consulting, information processing, accounting, legal services), production itself (quality control, maintenance, logistics), the promotion and distribution of products (transportation, commercial intermediation, marketing, advertising)" (Martinelli, 1991)

SM helps in the development and growth of the country, which is addressed in the Brundtland report termed as *"Sustainable development"*, in which the development of the present time should not affect the needs of the future generations (Seliger et al., 2008). The U.S Department of Commerce defined SM as (ITA, 2012) :

"*The creation of manufactured products that use processes that minimize negative environmental impacts, conserve energy and natural resources, are safe for employees, communities, and consumers and are economically sound*."

Various studies have reported the importance of social and economic factors in manufacturing in various industrial sectors. If we consider the third pillar of sustainability, i.e. environmental, we cannot neglect that manufacturing is having a significant impact on the environment from the processes in which energy consumption and raw material used are considered the first important factor. In the present time, manufacturing is considered as the source for all goods which can be used in transportation, production, living and health purposes and completely depends on modern technologies (Ahuja et al., 2019). Currently, manufacturers are redefining and innovating manufacturing processes to achieve sustainability (Stock & Seliger, 2016). Sustainability in manufacturing is a critical issue at present and will be a critical issue for future industries and generations as natural resources are not infinite and cannot complete future generation demands (Joung et al., 2013).

2.2. Challenges to Sustainable Manufacturing

Overpopulation and rapid industrialization in the world are consuming natural resources rapidly. Current manufacturing systems are facing many challenges due to the lack of sufficient natural resources which have been discussed in Table 1.

Table 1: Main challenges to sustainable manufacturing

S. No	Challenges	Authors
1	Higher competitiveness	(Bhanot et al., 2017; Jamwal et al., 2021b; Yadav et al., 2020)
2	Manufacturing systems	(Machado et al., 2020; Manupati et al., 2018; Yadav et al., 2020)
3	Sustainable planning and scheduling	(Carmita Camposeco-Negrete, 2015; Manupati et al., 2018; Sealy et al., 2016)
4	Product recovery	(Kuik et al., 2016; Meng et al., 2020)
5	New employee abilities	(Bhanot et al., 2017; Malek & Desai, 2020; Yadav et al., 2020)
6	New policies, regulations and requirements	(Haapala et al., 2013; Jamwal et al., 2021a; Sharma et al., 2020a; Yadav et al., 2020)
7	Smart disposal	(Machado et al., 2020; R. Sharma et al., 2020)
8	Lack of knowledge about advanced technologies e.g. Machine learning and Artificial intelligence	(Jamwal et al., 2021a; Sharp et al., 2018; Yadav et al., 2020)
9	Closed-loop supply chains	(Abdirad & Krishnan, 2020; Bag et al., 2018; Malek & Desai, 2020; Yadav et al., 2020)
10	Smart manufacturing	(Sharp et al., 2018; Yadav et al., 2020)
11	Smart distribution	(Weiwei Cui et al., 2019; Dev et al., 2020; Sharp et al., 2018)
12	Maintenance	(Kumar et al., 2018; Lieber et al., 2013; Seliger et al., 2008; Wang et al., 2015)
13	Sustainable design	(Jamwal et al., 2021a; Machado et al., 2020; Malek & Desai, 2020; Sharma et al., 2020a)

2.3 ML applications in manufacturing

The application of ML techniques in manufacturing has gained popularity in the last two decades (Rohit Sharma et al., 2020). Today in manufacturing industries, ML tools are being used in various areas, e.g. troubleshooting, control and optimization purposes (Syafrudin et al., 2018). The scope of ML techniques in the manufacturing industry is shown in Figure 1.

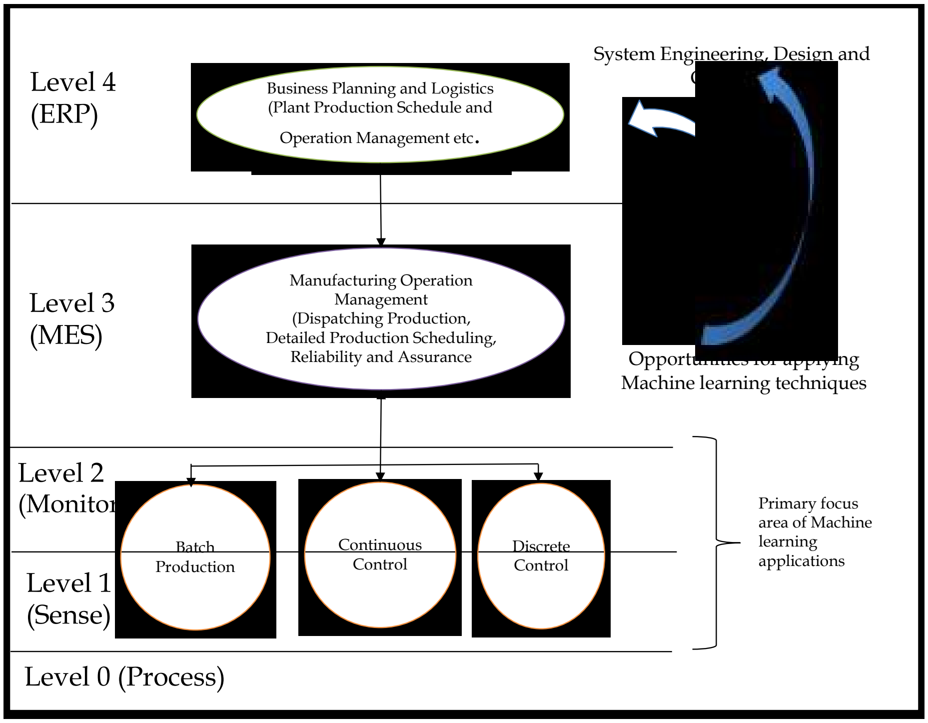

Fig.1: Opportunities for machining learning techniques in manufacturing industries

ML techniques are part of artificial intelligence and they can learn and adapt to new changes in systems (Loyer et al., 2016). Therefore, ML techniques provide a strong argument as to why their application is necessary at this present time in the manufacturing sector (Priore et al., 2001). Learning from the environment and changing it automatically according to its needs is the major strength of ML techniques. Generally, ML techniques are designed in such a way so that they can derive knowledge from data sets.

The manufacturing requirements and theoretical ability of ML techniques to meet those requirements are discussed in Table 2.

Table 2: Manufacturing requirement and ability of ML techniques to fulfil requirements

Manufacturing Requirement	ML ability to meet the manufacturing requirement
Ability to handle large data sets and high dimensional problems	ML algorithms such as SVM (support vector machine) are capable to handle large data sets with high dimensionality.
Ability to adapt to new environments due to technological change with reasonable cost and effort	As ML techniques are part of Artificial intelligence, they can learn and adapt to changes in the environment in the manufacturing industry due to new technological adoption.
Ability to minimize the complex nature of results	ML techniques are capable to derive patterns from large data sets and can also help to derive the future behaviour of a system.
Ability to work with the available machine data or manufacturing data	ML techniques are designed in such a way so that they can derive knowledge from existing data when it is

	analysed. This analysed data can be used to make future predictions, e.g. maintenance predictions.
Ability to identify the intra and interrelationship between processes	ML techniques can contribute to generating new information and knowledge by identifying patterns from existing data sets.

However, there is a need to discuss and analyse the existing limitations and strengths of ML techniques. Also, there is a need to answer questions such as *how can industries benefit from the adoption of ML techniques and how can these techniques be used to handle qualitative information?* ML techniques are a promising choice to improve quality control and optimization in different types of manufacturing systems. These techniques can be used in complex manufacturing environments where problem detection is difficult. In the past few years, ML in manufacturing has grown as an independent research domain on which researchers are focusing on. In this section, we have attempted to provide a generalized view of ML techniques in manufacturing. However, every ML technique has its limitations and advantages, which need to be considered while using the ML technique for manufacturing applications. ML techniques help to improve cycle time and resource utilization in complex manufacturing problems.

2.4 ML algorithms

ML is defined as "the scientific study of algorithms and computational models on computers using experience for progressively improving the performance on a specific task or make accurate forecasts" (PRIORE et al., 2001; Wuest et al., 2016). In this definition, the term "experience" is defined as the historical data which can be used to build a prediction model. Common terminologies used in ML are discussed in Table 3. Structures of ML techniques and their algorithms are illustrated in Figure 2.

Table 3: Common terminologies used in ML

S.NO	Terminology	Definition
1	Example	Instances of data used for learning purposes.
2	Feature	Feature is defined as an individual property which can be measured or characteristic of phenomena being observed. Feature can also be defined as a set of vectors or attributes which are associated with an example.
3	Hyperparameters	Hyperparameters can be defined as parameters whose values can be used to control the learning process.
4	Hypothesis set	Hypothesis set may be defined as the set of a function that can be used to map a set of features with a set of labels.
5	Label	Labels are the final output in machine learning.
6	Loss function	Loss function is used to measure the difference between a true label and a predicted label.
7	Test sample	Test samples are used to evaluate the performance of a learning algorithm.
8	Training sample	Training samples are used to train the learning algorithms.
9	Validation sample	Validation samples are used to tune the parameters for a learning sample.

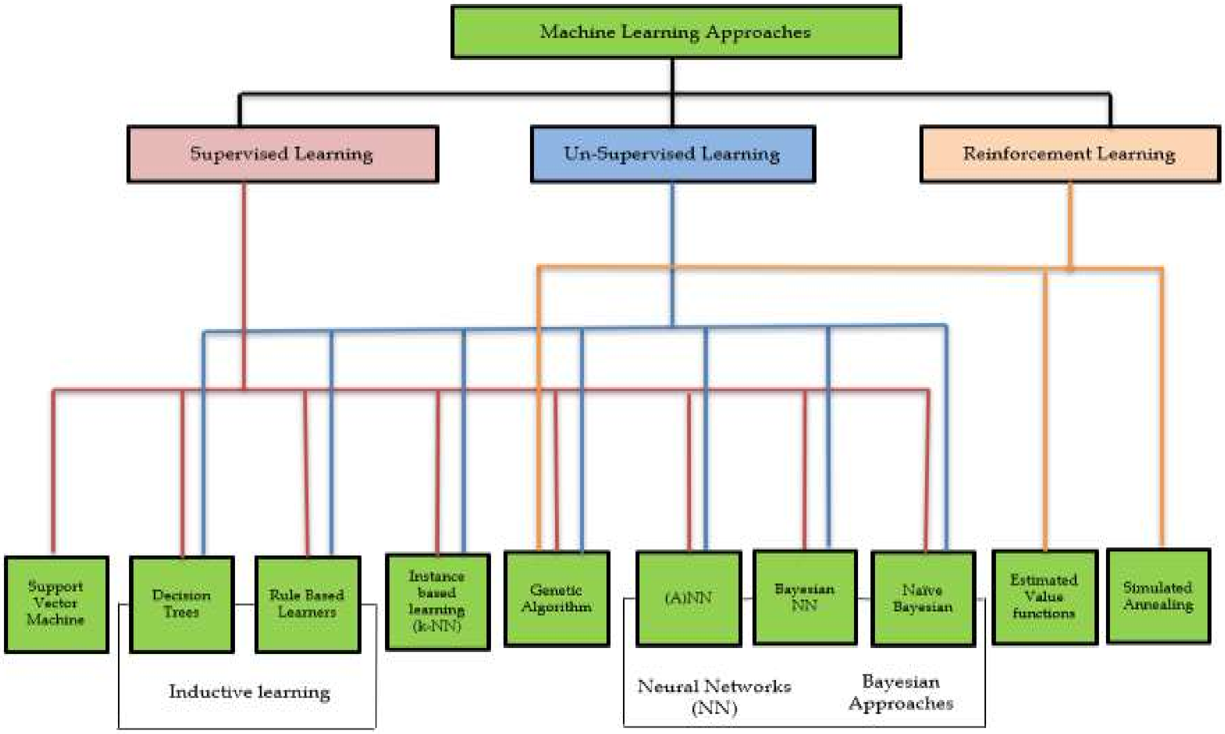

Fig.2: Structures of ML techniques and algorithms

A general configuration for ML algorithms is presented in Figure 3. In both labelled and unlabelled training data is extracted as input to ML systems from different sources. Table 4 shows the various machine learning algorithms with their application area in manufacturing.

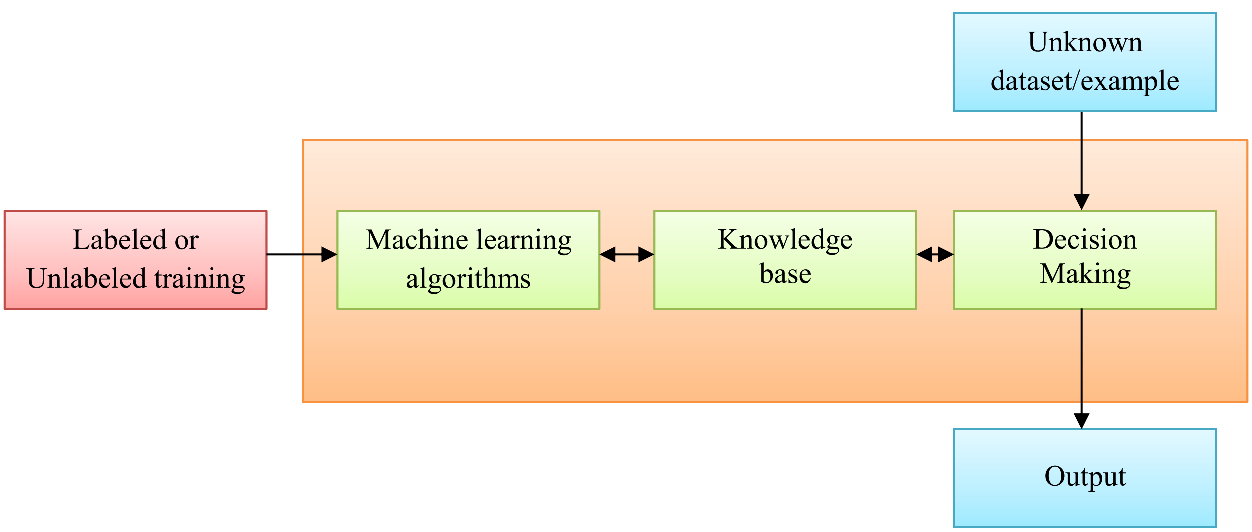

Fig. 3: Configuration of the ML system [Adapted from (Du & Sun, 2006)]

Table 4: Machine learning algorithms with their application area in manufacturing

ML Algorithm	Description	Application area in Manufacturing	Strengths	Weakness
Support Vector Machine (Alpaydin, 2020)	Support Vector Machine is a boundary detection algorithm that is used to define or identify the different data points in multi-dimensional boundaries	Process planning and scheduling, Quality assessment, Tool wear prediction, Forecasting, Control chart pattern recognition	(a) Memory efficient algorithm (b) More effective in the case of high dimensional spaces and when number of cases are more than number of samples.	(a) Not suitable for larger data sets and data with higher noise. (b) Underperform if no. of features at each data point exceeds the no. of samples.
Decision Tree (Kodratoff, 2014)	Decision tree algorithm is used to classify the larger datasets into smaller datasets where each small dataset contains the responses (either "YES or "NO")	Technology Selection, Yield enhancement, Life cycle engineering, Green Design, Quality management	(a) Requires less data preparation and effort during the pre-processing phase. (b) Model is easy to explain for technical teams or stakeholders. (c) Normalization of data is not required in the decision tree. (d) Scaling of data is not necessary.	(a) A small change in data is leading to a larger change in the structure of a decision tree. (b) Calculations are more complex sometimes as compared to other algorithms. (c) Inadequate for predicting continuous values and regression.
Regression analysis (Kubat, 2017)	Regression analysis is a predictive model which is used to express the relationship between the various inputs parameters and output parameters in the form of equations. The different types of regression models are: polynomial regression, logistic regression and linear regression	Forecasting, Machining parameters optimization, Supplier performance evaluation, Performance measurement	(a) Implementation is simple as compared to other algorithms. (b) Overfitting can be reduced by regularization.	(a) Prone to underfitting and not suitable to complex relationships. (b) Sensitive to outliers.

Artificial Neural Networks (ANN) (Lison, 2015)	ANN is the mathematical and computational model which is inspired by the biological nervous systems.	Demand Forecasting, Capacity utilization, Quality improvement, Machining parameters optimization	(a) Can be used for large volume of data and complex relationships (b) No restriction on input variables (how the variables are distributed).	(a) More data is required as compared to other algorithms. (b) Model development time is more. (c) Computationally expensive.
Clustering (Kodratoff, 2014)	The clustering algorithms are used to divide data into the k clusters, i.e. k means clustering	Manufacturing system design, Engineering design, Quality assurance, Production planning and process planning	(a) Quite effective as an unsupervised learning method with larger data sets (b) Implementation process is easy and no distributional assumption about the data	(a) Not suitable for complex geometric shapes (b) Does not consider the data points that are far away from each other even if it belongs to the same cluster
Bayesian Networks (Parsons, 2005)	Bayesian network algorithm is used the Bayes theorem to predict the output class. In this prior probability and conditional probability is used.	Monitoring and Diagnosis, Performance evaluation, Fault Diagnosis, Resilence modelling, Product quality management	(a) Suitable to determine the causal relationship between random variables (b) Both direct pieces of evidence and indirect pieces of evidence can be incorporated in a single analysis	(a)Not suitable for high dimensional data (b) Computation process is expensive (c) Sometimes models are complex
Ensemble Learning (Kodratoff, 2014)	Ensemble learning algorithms are used to solve the computational intelligence problems by the multiple models, which can also help to improve the function approximation, prediction and classification of data.	Material removal rate predictions, System designing, System diagnosis, Supplier selection	(a) Improves the ML results by combining the multiple models that allow formulating a better prediction model as compared to a single model. (b) Noise level is less	(a) Difficult to interpret (b) High computation cost
Deep Learning (Lison, 2015)	Deep learning is a subfield of ML which is concerned with ANN-based algorithms. Deep learning can be classified into three categories: (1) Deep recurrent neural networks, (2) Deep	Distortion prediction in additive manufacturing, Quality analysis, Fault Diagnosis	(a) The architecture is flexible and capable to adapt to new changes in the system. (b) Parallel computational processes can be done	(a) Requires large volume of data for more accuracy and high computation cost (b) Requires higher skill level

	belief networks, and (3) stacked autoencoder.		(c) Same neural network approach can be used for different data types	
Genetic Algorithms (Kodratoff, 2014; Parsons, 2005)	Genetic algorithms are known as search heuristics that are inspired by Charles Darwin's theory. Genetic algorithms are used to solve both unconstrained or constrained optimization problems in different sectors	Planning and Scheduling, Parameter optimization, Supply chain management	(a) Solution search from the population of points rather than a single point. (b) Suitable for multi-objective optimization and noisy environments	(a) Computationally expensive and time-consuming (b) As it requires less information about the problem so designing an objective function is difficult sometimes.
Instance-based learning (Kubat, 2017)	Instance-based learning algorithms are also known as memory-based learning algorithms. They are used to compare the new problem instances with the different instances that occur in training and that are already stored in the memory.	Reconfigurable manufacturing system design, Simulation and animations	(a) Rather than estimating the entire instance set its focus on the local approximation, which can be made to the target function (b) It can adapt new changes in data easily	(a) Higher computation cost (b) A large amount of memory is required to store the data as this is a memory-based learning algorithm.

3. Review Methodology

A large amount of literature is available in scientific databases that are not fully accessible to researchers, policymakers and practitioners. The process of extracting useful information from these databases is time-consuming for academics and policymakers (Antony et al., 2020; Jamwal et al., 2021b). Thus, to answer the research questions raised in the present study, we have adopted a systematic literature review and bibliometric analysis. A systematic literature review study focuses on the identification, selection and assessment of research to answer the research questions formulated to achieve the research objectives. Bibliometric analysis is a statistical analysis of research articles. It is an effective way to measure the influence of research articles in a particular research area over time (Baker et al., 2020). In the next sections, we present a systematic literature review and bibliometric analysis.

3.1 Systematic literature review

The present study adopted a three-phase systematic literature review technique, as suggested by (Tranfield et al., 2003), in which the first phase is "*planning the review*", the second phase is "*conducting the review*" and the third phase is "*review findings based on review*".

3.1.1 Planning the review

The first objective of this study was to review the applications of ML algorithms in sustainable manufacturing. For the specification of conceptual boundaries, the Scopus database is utilized with the keywords covering the ML techniques discussed in Figure 4. All the matched review papers and technical papers were refined and reviewed with alignment to objectives and checked whether sustainable manufacturing is included in the title, abstract or keywords of papers.

3.1.2 Conducting the review

The search for related keywords was performed on the Scopus database. Scopus is the largest database, it includes documents from reputed publishers like Elsevier, Springer, Taylor and Francis, Inderscience, IGI Global, Wiley, IEEE, IOP science and Emerald (Jamwal et al., 2021c). In this study, the search was limited to articles, conference papers and books. Conference reviews, book reviews, undefined and editorial research documents were excluded from the study to maintain the quality of the study. As a first step, the search keywords used were “Sustainable manufacturing” AND “Machine learning”. From this search, 104 articles were found, including conference reviews and book reviews. To maintain the quality of the study, a second search was conducted with more specific keywords. The keywords used in the second search are shown in Figure 4.

With these keywords, a total of 96 articles were found and finalized. The articles included journal papers, conference papers and book chapters. Furthermore, following the inclusion criteria, the resulting articles were finalized. The inclusion criteria were based on the fact that (1) journal or conference papers had to be from a peer-reviewed source, (2) these had to contain any of the algorithms, (3) ML was applied to at least one of the research areas of sustainable manufacturing, (4) the articles had to be in English language only, and (5) all the finalized articles had to pass the interrater-reliability test. It was ensured that all articles contained at least one of the keywords in each category. The articles were finalized by decoding them independently by three authors (having research background) and the coding of all three authors was compared with each other to understand the different scores, which helped to assess the inter-rater-reliability test. It was found that all the articles had zero differences in scores, ensuring the quality of the review. These criteria were applied, resulting in 96 articles. The selection process for articles is shown in Figure 4.

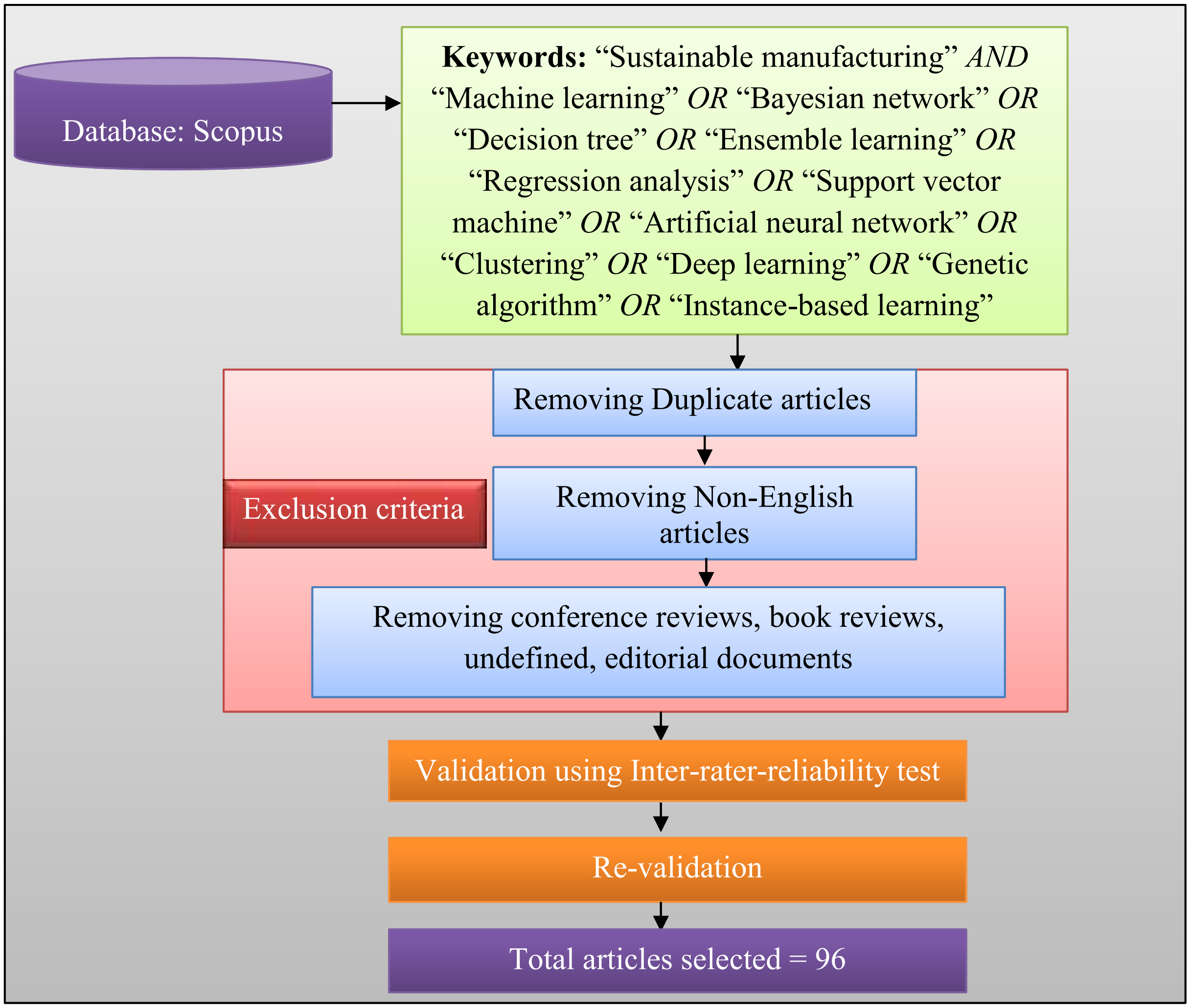

Fig.4: Article selection process

3.2 Bibliometric Analysis

In past studies, authors have used different software, with various advantages and limitations, for bibliometric analysis (Baker et al., 2020). In the present study, we divided the bibliometric analysis into four major categories, i.e. (1) year-wise citation structure of articles; (2) top authors, institutes and countries working in this area; (3) top journals and keywords in this area; (4) network analysis; and (5) cluster analysis. A summary of the bibliometric analysis is shown in Table 5.

Table 5: Bibliometric analysis summary

Description	Results
Source (Journal, Conference, Books)	50
Total Research Documents	96
Authors Keywords	305
Time Period	2009-2020
Total Authors	305
Total Funders	44
Total Citations	1541
Total Cited Articles	81
Total Subject Areas	9
Total Countries	31

3.2.1 Year Wise publication trend

The year-wise publication trend is shown in Figure 5. The first article reporting the use of ML with a focus on sustainable manufacturing was published in 2009. A total of 96 research articles have been published in the last 11 years. A sharp increase in publications in ML in sustainable manufacturing can be seen after 2015.

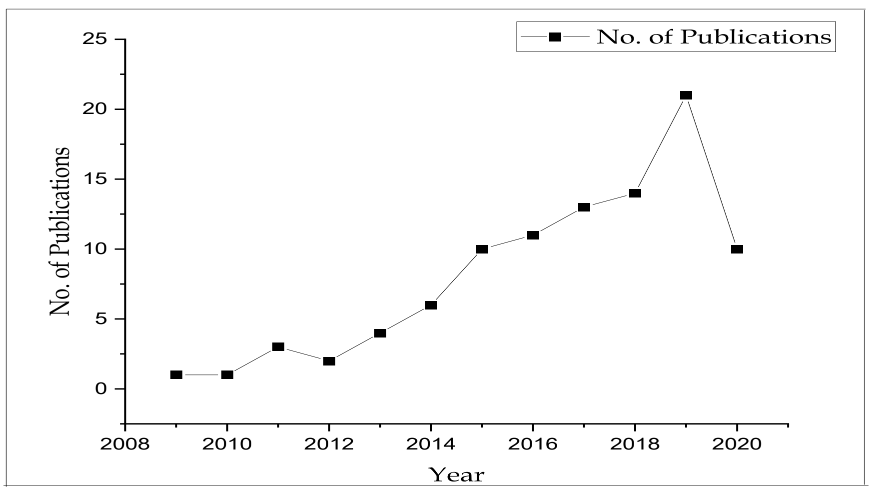

Figure 5: Year-wise publication trend

3.2.2. Year-wise citation Structure

The citation structure of articles on the application of ML in sustainable manufacturing is shown from the period 2009-2020 in Table 6. A total of 96 articles were published in this area. Also, it was found that the articles published in 2014 and 2015 were more productive as they had an average citation per article of 59.83 and 49.9 respectively. The average citation per article and average citation per cited document was calculated as:

$$\frac{C}{P} = \frac{TC}{TP} \tag{1}$$

$$\frac{C}{CP} = \frac{TC}{NCP} \tag{2}$$

Table 6: Year-wise citation structure of articles

						Publications with Citations ≥				
Year	**TP**	**TC**	**NCP**	**C/P**	**C/CP**	**250**	**100**	**50**	**25**	**1**
2009	1	3	1	3	3	0	0	0	0	3
2010	1	3	1	3	3	0	0	0	0	3
2011	3	9	2	3	4.5	0	0	0	0	2
2012	2	5	2	2.5	2.5	0	0	0	1	2
2013	4	66	3	16.5	22	0	0	0	1	2
2014	6	359	6	59.8	59.8	0	1	2	2	1
2015	10	499	9	49.9	55.4	0	1	3	5	0
2016	11	250	11	22.7	22.7	0	0	1	4	6
2017	13	156	13	12	12	0	0	0	2	11
2018	14	96	11	6.8	8.7	0	0	0	1	10
2019	21	90	19	4.2	4.7	0	0	0	0	19
2020	10	5	3	0.5	1.6	0	0	0	0	3

"TP: Total Publications, NCP: No. of cited publications, TC: Total citations, C/P: average citations per publication, C/CP: average citations per cited publications, Publications with Citations ≥: Publications with at least 1, 25, 50,100 and 250 citations."

3.2.3 Author's citation structure

Table 7 shows the authors' citation structure of published articles on ML in sustainable manufacturing. It can be seen that most authors working in this area are from China and the United Kingdom. Most authors from China working in this area are affiliated with "Huazhong University of Science and Technology". Among the top 15 authors in the Scopus database, 11 authors had more than 2 publications in this area. Haibo Dong, Ying Liu, Petrovic Sanja, Lohse Niels are working collaboratively with each other on several research articles in different sub research areas. Also, it was found that the majority of top authors that were working in the area are from industrial engineering and operation research backgrounds.

3.2.4 Country wise citation structure

The use of ML techniques in sustainable manufacturing has attracted the attention of a large number of researchers, as indicated by the contributions of authors in sustainable manufacturing from 31 countries. Table 8 shows the top publishing countries with their citation structure. In the top 15 countries, China, the United States and the United Kingdom are the most active. China has published 37 articles in this area. The citation structure of other countries is shown the Table 8.

Table 7: Citation structure of top authors

									Publications with Citations≥				
Author	**Affiliation**	**Country**	**TP**	**TC**	**NCP**	**C/P**	**C/CP**	**h-index**	**250**	**100**	**50**	**25**	**1**
Haibo Dong	University of Nottingham Ningbo	China	3	249	3	83	83	3	0	1	1	1	0
John Donovan	Institute of Technology Sligo	Ireland	3	2	2	0.6	1	1	0	0	0	0	2
Liang Gao	Huazhong University of Science and Technology	China	3	46	2	15.3	23	2	0	0	0	1	1
Weidong Li	Coventry University	United Kingdom	3	100	3	33.3	33.3	2	0	0	1	1	1
Xiaoxia Li	Huazhong Agricultural University	China	3	9	2	3	4.5	1	0	0	0	0	2
Xinyu Li	Huazhing University of Science and Technology	China	3	24	2	8	12	2	0	0	0	0	2
Ying Liu	University of Glasgow	UK	3	249	3	83	83	3	0	1	1	1	0
Lohse Niels	Loughborough University	United Kingdom	3	249	3	83	83	3	0	1	1	1	0
Petrovic Sanja	Nottingham University Business School	United Kingdom	3	249	3	83	83	3	0	1	1	1	0
Tormey David	Institute of Technology Sligo	Ireland	3	2	2	0.6	1	1	0	0	0	0	2
Peng Wang	University of Kentucky	US	3	44	2	14.6	22	2	0	0	0	1	1
Raunak Bhinge	Infinite Uptime	United States	2	28	2	14	14	1	0	0	0	1	1
Xianto Cai	Wuhan University	China	2	29	2	14.5	14.5	1	0	0	0	1	1
Jaime Andres Camelio	Virginia Polytechnic Institute and State University	United States	2	44	2	22	22	2	0	0	0	1	1
Carmita Camposeco-Negrete	Tecnologico de Monterrey	Mexico	2	89	2	44.5	44.5	2	0	0	1	0	1

Table 8: Citation structure of top countries

								Publications with Citations≥				
Country	**TP**	**TC**	**NCP**	**C/P**	**C/CP**	**h-index**	**Category**	**250**	**100**	**50**	**25**	**1**
China	37	751	33	20.2	22.7	17	Developing	0	1	2	9	21
United States	21	446	18	21.2	24.7	8	Developed	0	1	2	5	10
United Kingdom	13	502	12	38.6	41.8	10	Developed	0	1	2	5	4
India	10	153	6	15.3	25.5	4	Developing	0	0	0	1	5
Singapore	5	92	5	18.4	18.4	4	Developed	0	0	0	2	3
Germany	4	4	1	1	4	1	Developed	0	0	0		1
South Korea	4	51	4	12.7	12.7	2	Developed	0	0	0	1	3
Hong Kong	3	11	2	3.6	5.5	2	Developed	0	0	0	0	2
Ireland	3	2	2	0.6	1	1	Developed	0	0	0	0	2
Italy	3	111	3	37	37	2	Developed	0	1	0	0	2
Japan	3	90	3	30	30	3	Developed	0	0	1	1	1
Mexico	3	107	3	35.6	35.6	3	Developing	0	0	1	0	2
Australia	2	29	2	14.5	14.5	2	Developed	0	0	0	1	1
Brazil	2	7	2	3.5	3.5	1	Developing	0	0	0	0	2
Malaysia	2	15	2	7.5	7.5	1	Developing	0	0	0	0	2

3.2.5 Source wise citation structure

Based on our dataset, a total of 50 sources were found for 96 articles in 9 different subject areas. The leading journals are the Journal of Cleaner Production and Procedia Manufacturing, see Table 9. Further, many of the top journals are indexed in both Scopus and Web of Science databases and have high impact score and impact factor. This means that this research area has received the attention of a large number of top journals in the engineering domain.

3.2.6 Institute wise citation structure

The data for the Institute wise citation structure was extracted from the Scopus database. In total, 134 institutes are working in this area. The citation structure for the top 15 Institutes working in this area is shown in Table 10. It was found that “Huazhong University of Science and Technology” from China has most of the publication in this area. Liang Gao is associated with this institute and has worked in the optimization area for machining processes. The Coventry University of United Kingdom is in the second position with a total of 4 publications. Weidong Li is affiliated with this institution and having to work in the sustainable planning and scheduling area.

Table 9: Citation structure of top sources

								Publications with Citations ≥				
S. No.	**Source**	**TP**	**NCP**	**TC**	**C/P**	**C/CP**	**h-index**	**≥40**	**≥30**	**≥20**	**≥10**	**≥1**
1	Journal of Cleaner Production	13	13	562	43.2	43.2	10	4	2	4		3
2	Procedia Manufacturing	7	5	21	3	4.2	3	0	0	0	1	4
3	Sustainability	7	6	46	6.5	7.6	5	0	0	0	1	5
4	International Journal of Advanced Manufacturing Technology	6	6	102	17	17	5	0	1	2	1	2
5	Procedia CIRP	6	5	86	14.3	17.2	3	1	0	0	0	4
6	International Journal of Production Research	5	5	198	39.6	39.6	4	2	0	0	1	2
7	Journal of Manufacturing Systems	3	3	123	41	41	3	2	1	0	0	0
8	International Journal of Product Lifecycle Management	2	1	3	1.5	3	1	0	0	0	0	1
9	Journal of Manufacturing Processes	2	2	32	16	16	2	0	0	1	0	1
10	Journal of Manufacturing Science and Engineering Transaction of The ASME's	2	2	33	16.5	16.5	2	0	1	0	0	1

Table 10: Citation structure of top institutes

								Publications with Citations≥				
Institute	**TP**	**TC**	**NCP**	**C/P**	**C/CP**	**h-index**	**Country**	**250**	**100**	**50**	**25**	**1**
Huazhong University of Science and Technology	5	55	4	11	13.7	4	China	0	0	0	1	3
Coventry University	4	148	4	37	37	4	United Kingdom	0	0	1	2	1
Cranfield University	3	93	3	31	31	3	United Kingdom	0	0	0	2	1
Tecnologico de Monterrey	3	107	3	35.6	35.6	3	Mexico	0	0	1	0	2
National University of Singapore	3	51	3	17	17	3	Singapore	0	0	0	1	2

Harbin Institute of Technology	3	41	3	13.6	13.6	2	China	0	0	0	1	1
Nanjing University of Aeronautics and Astronautics	3	63	3	21	21	3	China	0	0	0	1	2
University of Connecticut	3	71	3	23.6	23.6	3	United States	0	0	0	2	1
National Institute of Standards and Technology	3	99	3	33	33	2	United States	0	0	1	1	1
Institute of Technology Sligo	3	2	2	0.6	1	1	Ireland	0	0	0	0	2
Wuhan University	3	37	3	12.3	12.3	2	China	0	0	0	1	2
Shandong University	3	42	3	14	14	3	China	0	0	0	1	2
Huazhong Agriculture University	3	37	3	12.3	12.3	2	China	0	0	0	1	2
Loughborough University	2	35	2	17.5	17.5	2	United Kingdom	0	0	0	1	1
Ministry of Education China	2	8	1	4	8	1	China	0	0	0	0	1

3.2.7 Top keywords

Keywords analyses identify the most frequently used keywords by authors in their research articles. For the ML applications in sustainable manufacturing articles, authors have used different keywords, see Table 11. Also, the keywords network analysis of frequently used keywords is shown in Figure 6. It was found that sustainable manufacturing is the most used keyword with the maximum no. of occurrences, i.e. 44. Energy consumption and energy efficiency are also the most used keywords.

Table 11: Top keywords in ML and sustainable manufacturing

Keyword	Occurrence
Sustainable Manufacturing	44
Energy consumption	11
Energy efficiency	7
Machine learning	11
Multi-objective optimization	8
Scheduling	4
Genetic algorithm	5
Process planning	4
Sustainability	5
Energy modelling	3
Turning	4
Industry 4.0	3
Optimization	4

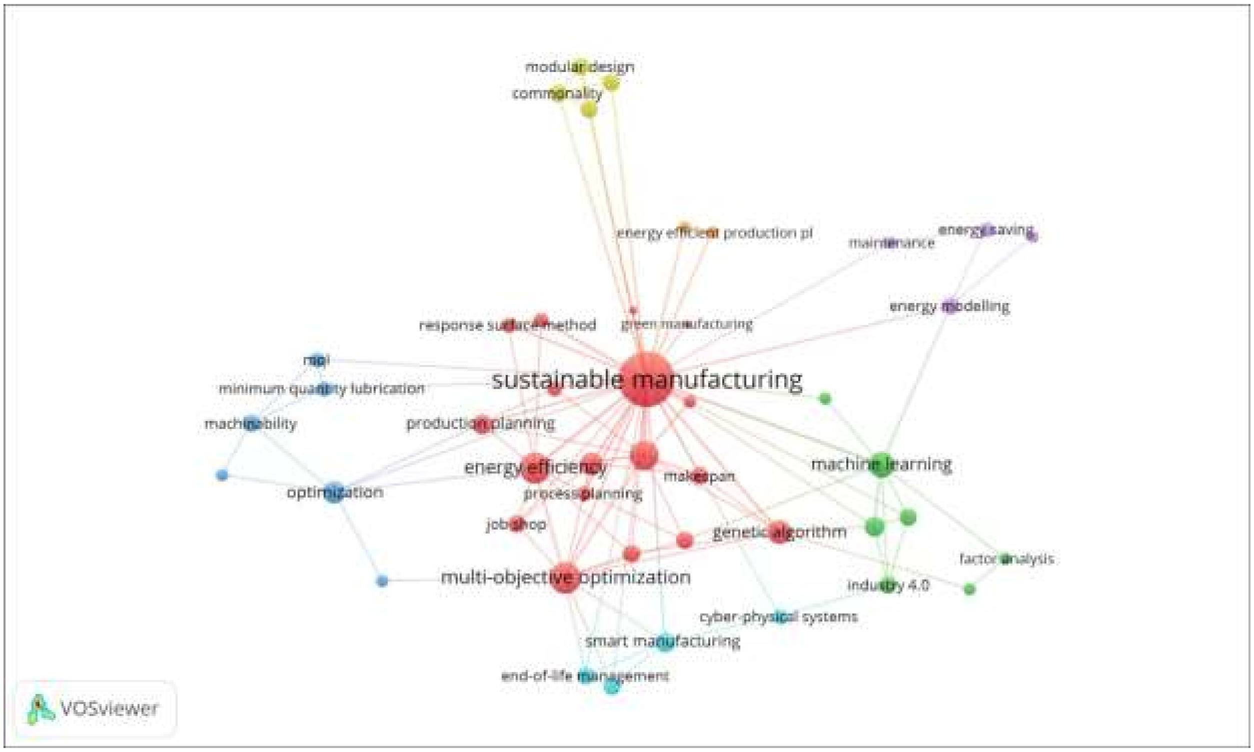

Figure 6: A network analysis of keywords

3.2.8 Network analysis of publications

The network analysis of publications of ML application in sustainable manufacturing was carried out with the support of the VOSviewer software. The dataset used in this study was extracted from the Scopus database and saved in the CSV format, which contained the bibliographic data for the study.

3.2.9 Citation analysis

The citation analysis of documents shows the patterns, frequency and graph of citations in a research document. It also measures the degree of connectivity between various research publications. The citation analysis of documents conducted determined the top-cited articles in this area. In the top-cited articles, (C. Liu et al., 2014) had the highest number of citations (i.e. 162), followed by (May et al., 2015) with 105 citations. These two articles are from the sustainable planning and scheduling research area, which indicates that ML techniques have extensive research opportunities in this field of manufacturing. Total citation per year (TCP) measures the impact of research articles per year. It can be observed that the top 6 papers in this area had more than 10 average citations per year. The most cited articles in ML applications for SM is shown in Figure 7 and the global citation network is shown in Figure 8.

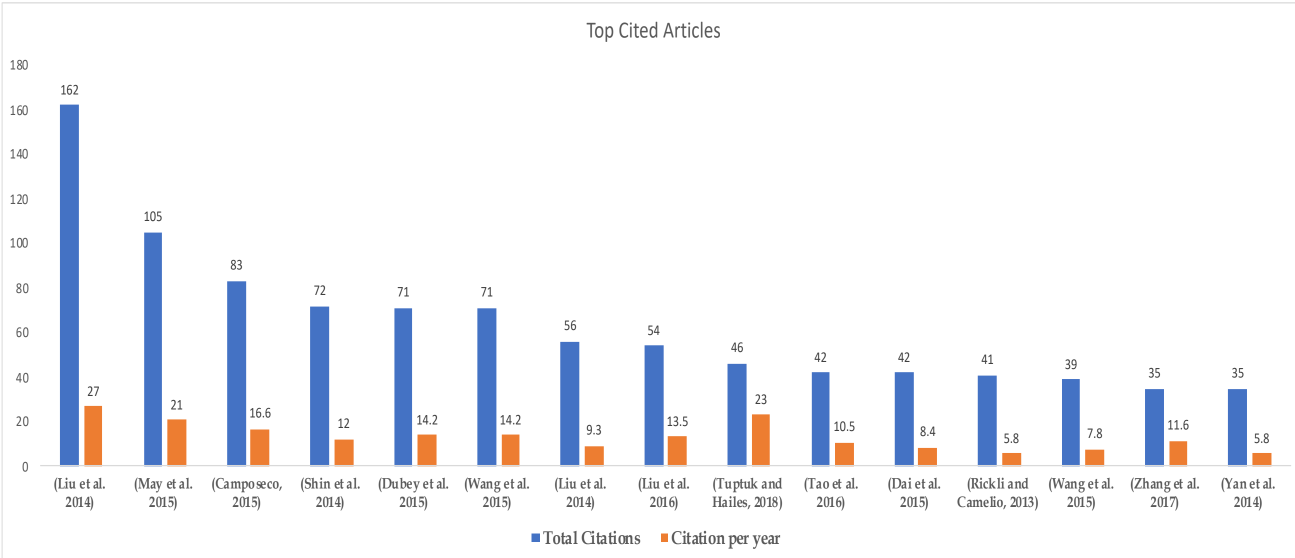

Figure 7: Top cited articles

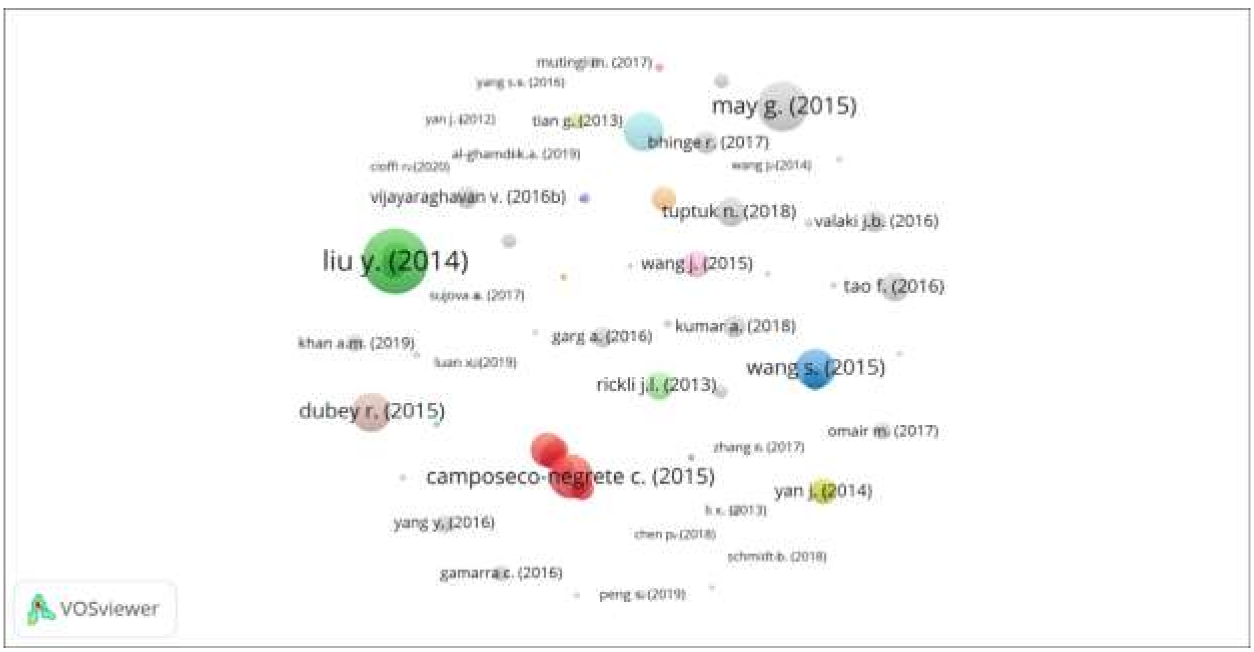

Figure 8: Global citation network

3.2.10 Co-Citation Analysis

Co-citation analysis involves the tracking pairs of research articles that are cited together in the source publications. Co-citation of two research documents occurs when both of the research documents are cited in the third research document (Baker et al., 2020). Generally, a co-citation network can be obtained using the formula:

$$B_{Cocit} = Article \times cited\ reference\ article \qquad (3)$$

Here:

B_{cocit} represents a symmetrical matrix

The leading papers of each cluster obtained from the cluster analysis is shown in Table 12.

3.2.11 Cluster analysis

In bibliometric studies, cluster analysis helps to study the bibliographic networks. It helps to provide a deeper insight into the structure of the network. There are many clustering techniques such as hierarchical agglomerative clustering and density clustering etc.

In this section, we analyze and present the findings of the systematic literature review in all the sub research areas of sustainable manufacturing. These represent the use and applications of ML techniques in the different sub research areas of sustainable manufacturing. The use of ML approaches in each research area serves the specific applications of improving sustainable manufacturing practices in a particular research area.

Table 12: Clusters analysis

Cluster 1	Total Link Strength
(C. Liu et al., 2014)	1
(Carmita Camposeco-Negrete, 2015)	2
(Sealy et al., 2016)	1
(Zhang et al., 2017)	1
(Deng et al., 2017)	1
(Preez & Oosthuizen, 2019)	1
(C Camposeco-Negrete, 2020)	2
Cluster 2	
(Wang et al., 2015)	2
(Huang et al., 2017)	2
(Li & Lu, 2017)	2
Cluster 3	
(C. Liu et al., 2014)	2
(Dai et al., 2014)	1
(Y. Liu et al., 2015)	1
Cluster 4	
(Dubey et al., 2015)	1
(Khatri et al., 2019)	1
Cluster 5	
(Kuik et al., 2016)	1
(Meng et al., 2020)	1
Cluster 6	
(Bhinge et al., 2017)	1
(Ferguson et al., 2018)	1

In the cluster analysis, six major clusters with their total link strength were identified, see Table 12.

In Cluster 1, a total of seven papers were identified. (Liu et al., 2014) stated that industries should focus on minimizing their carbon economy and carbon dioxide emissions generated through their manufacturing processes without affecting their economic factors. A multi-objective optimization model was developed with NSGA-II (Non-dominated sorting genetic algorithm) to minimize the total completion time and CO_2 emissions. The proposed model was validated with the help of industrial cases and it was found that the proposed algorithm had a better performance than the previously proposed algorithms. (Carmita Camposeco-Negrete, 2015) optimized the cutting parameters of AISI 6061 T6 aluminium in a turning operation and concluded that advanced optimization techniques play a vital role in minimizing energy consumption. (Sealy et al., 2016) investigated the effect of total energy consumption in a milling process. It was found that the role of total energy consumption is important to balance the economic and environmental dimensions of sustainability. In this study, regression models were proposed to predict the specific energy consumption. It was also found that milling consumes more energy than the down milling process. (Zhang et al., 2017) investigated the process parameters effect on the optimization of a multi-pass dry drilling process. The main objective of the study was to maximize efficiency while minimizing CO_2 emissions and total energy consumption. A multi-objective-based optimization model was proposed and solved by a genetic algorithm approach. It was found that the balance of these three objectives can help manufacturing companies achieve sustainability in their manufacturing processes. (Deng et al., 2017) investigated the effect of process parameters in a cutting process. The main objective of the study was to minimize the total energy consumption during the cutting process. A multi-objective optimization model based on GA (genetic algorithm) was proposed to minimize the total energy consumption as well as total processing time in the cutting process. (Preez & Oosthuizen, 2019) discussed the opportunities for ML approaches in the manufacturing sector. It was found that the adoption of ML approaches in manufacturing processes, e.g. cutting processes, act as an enabler for smart-sustainable manufacturing in industries. (C Camposeco-Negrete, 2020) investigated the effect of sustainable machining on AISI 1045 steel and concluded that the development of ML approaches has enhanced the research area of sustainable manufacturing. It was concluded that sustainable machining processes can enhance the quality of the product as well they minimize the negative environmental impact.

In Cluster 2, a total of three papers were found in which (Wang et al., 2015) proposed a sustainable process planning optimization model for the milling process with a genetic algorithm approach with the objectives e.g. surface quality and energy efficiency. A genetic algorithm was used to establish the non-linear relationship between measured datasets and process parameters. Further, the proposed model was validated through an industrial case study in the European machining industry. (Li & Lu, 2017) developed an energy consumption model based on artificial neural network (ANN) and response surface methodology (RSM). In their study, a two-level optimization model was proposed to demonstrate the relationship between machining parameters and energy consumption. (Huang et al., 2017) proposed a mathematical model based on the NSGA-II algorithm for minimizing carbon emission and machining cost. Also, the TOPSIS methodology was adopted to determine the optimal solution from a Pareto front.

In Cluster 3, (C. Liu et al., 2014) investigated sustainable planning and scheduling problems in job shop scheduling considering the total energy consumption and total weighted tardiness as objectives. The optimal solution for the problem was investigated with NSGA-II algorithm and a case study on a 10×10 job shop system was carried out to explore the effectiveness of the proposed model. (Dai et al., 2014) reported that sustainable planning and scheduling play an important role in the implementation of sustainability practices in planning and scheduling. In their study, an energy-aware based integrated process plan was proposed for job shops. The objectives makespan and energy consumption were considered and Modified-GA (genetic algorithm) approach was used to find the optimal solution. (Y. Liu et al., 2015) investigated the planning and scheduling problem to minimize total electricity consumption, cost of electricity and total weighted tardiness. A case study was carried out in the manufacturing industry of China to solve the Rolling Blackout policy problem. The optimal solution was investigated with NSGA-II algorithm on a modified job shop problem.

In Cluster 4, only two papers on empirical investigations were found. (Dubey et al., 2016) proposed a world-class sustainable manufacturing framework for manufacturing industries. The data in this study was collected through the “Dillman’s total design method”, which helped to improve the quality of responses. Also, regression analysis and CFA were performed for hypothesis testing. (Khatri et al., 2019) proposed a framework by considering success factors for sustainable manufacturing. In their study, a regression model and ANN were developed to find the suitability of the importance of sustainable manufacturing variables in manufacturing industries. The study revealed that government support is a must for the successful implementation of sustainable manufacturing practices. It also determined that the government should provide financial assistance to manufacturing firms for the implementation of sustainability.

In Cluster 5, only two papers related to product recovery decision were found. (Kuik et al., 2016) suggested that recovery operations in the implementation of sustainability practices are more complicated than the traditional manufacturing systems due to uncertainty issues. In their study, an integrated model was proposed to find the optimal recovery plan by considering quality, waste and lead time as constraints. The optimal solution was found with a genetic algorithm. (Meng et al., 2020) proposed a framework for the implementation of end of life management based on condition monitoring. A multi-objective optimization model was formulated to find the best roadmap and meet business goals. In their study, the NSGA-II algorithm was used to find the optimal solution. Numerical experiments were also conducted for the validation of the proposed model.

Finally, in Cluster 6, two papers were found for energy prediction models for machine tools. Machine tools are important in manufacturing processes as they contribute to energy-efficient process planning. (Bhinge et al., 2017) developed a model for energy prediction in manufacturing industries for machining tools. In their study, Gaussian process regression, which is a non-parametric ML approach, was used for optimization. (Ferguson et al., 2018) suggested that at present, it is possible to use real-time ML approaches to monitor the current state of machines. ML techniques are capable to handle large datasets that are complex in nature. In their study, the condition of the milling tool was predicted with the data processing pipeline. The novel Kernel function ML approach was used for prediction. Further, Table 13 shows the different studies with their statistical results in SM with ML techniques.

Table 13: Studies on different areas of SM with machine learning algorithms

Author	Focused area	Observed Feature	Model/Algorithm	Results
(C. Liu et al., 2014)	Operation decision making for Single machine system	Carbon emission, total time completion	NSGA-II	There is a significant trade-off between total CO_2

				emission and total completion time
(Carmita Camposeco-Negrete, 2015)	Optimization of cutting parameters	Surface roughness and specific energy	ANN	Specific energy reduced to 14.41% and surface roughness reduced to 360.47%
(Sealy et al., 2016)	Energy consumption modelling	Material removal rate, Specific energy consumption	Regression	Specific energy reduced to 11.3%
(Zhang et al., 2017)	Process parameters optimization	Energy, Carbon emission and efficiency	GA	Energy consumption reduced by 0.8%
(Deng et al., 2017)	Process parameters optimization	Specific energy consumption	GA	Specific energy consumption reduced by 32.07%
(Li & Lu, 2017)	Energy consumption	Material removal rate and Energy consumption	ANN	Predicted data are 3.08%, 0.42% and 1.17%
(Y. Liu et al., 2015)	Manufacturing systems	Environmental impacts, and delivery rate	GA	Delivery rate is improved by 60%
(Dai et al., 2014)	Planning and scheduling	Energy-aware planning and scheduling	GA	Total energy consumption is reduced by 2.5%
(Kuik et al., 2016)	Product recovery value	Recovery decision making	GA	Recovery value is increased from $18.91to $19.21
(Meng et al., 2020)	Condition monitoring	End of life management using condition monitoring	GA	Bearing cases can be reused in the industry when age is<4000 hours
(Bhinge et al., 2017)	Energy prediction modelling	Energy prediction	Regression	Energy prediction functions can be model using the 10% of data points

These six clusters show how ML techniques have been used in sustainable manufacturing in different sub research areas.

3.3 Benefits of Machine learning adoption in Sustainable manufacturing

From the above discussion, it is found that there is a growing importance of digitalization and innovation in products and processes. Consequently, it means that the adoption of ML techniques in the manufacturing sector is also an emerging issue. This suggests that ML techniques represent an opportunity for manufacturing companies to handle the large amount of data generated from different production practices. Now, due to strict government policies and regulations related to the environment, many countries are adopting sustainable manufacturing practices *(Machado et al., 2020; Malek & Desai, 2020)*. Both ML and AI-based technologies are also important for the development of Industry 4.0 practices with regards to sustainability.

From the sustainability perspective, the analysis undertaken in the present study highlighted that the adoption of ML approaches has the potential to bring new improvements, e.g. better resource utilization, tool life prediction and quality management, in manufacturing industries. There is a growing interest in the applications of sustainable development and green manufacturing. In this line, ML techniques can play an important role in business practices to increase sustainability through the better utilization of energy and resources, intelligent decision-making, environmental footprint monitoring and better life prediction.

Furthermore, it was also found that ML techniques in sustainable manufacturing offer a wide range of opportunities for sustainable development, which include supply chain management, condition monitoring and predictive maintenance. The different technologies with their dimensions and benefits are shown in Figure 9.

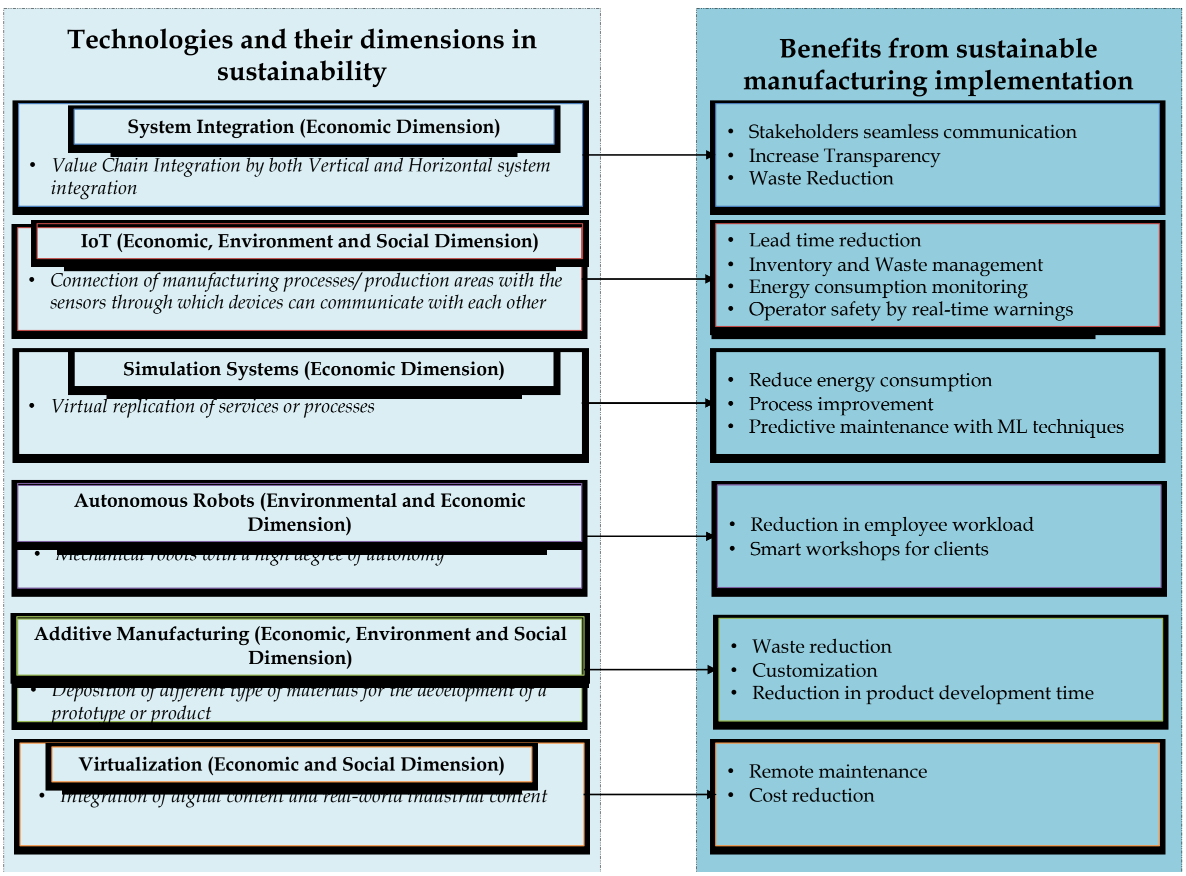

Figure 9: Key research areas for future ML research in SM

4. Proposed Framework and implications

In this section, we have proposed a ML-based SM framework. In the framework, three main components are considered i.e. different phases of SM, opportunities of ML techniques and benefits from ML in all three dimensions of sustainability.

4.1 ML-SM framework

The review findings of the present study reveal that ML-driven technologies play an important role in SM as evidence shows that they improve the overall efficiency of the manufacturing industries. We have used the findings of literature to develop a ML-based SM application framework for manufacturing industries that can be used by the practitioners in future. The proposed framework is shown in Figure 10.

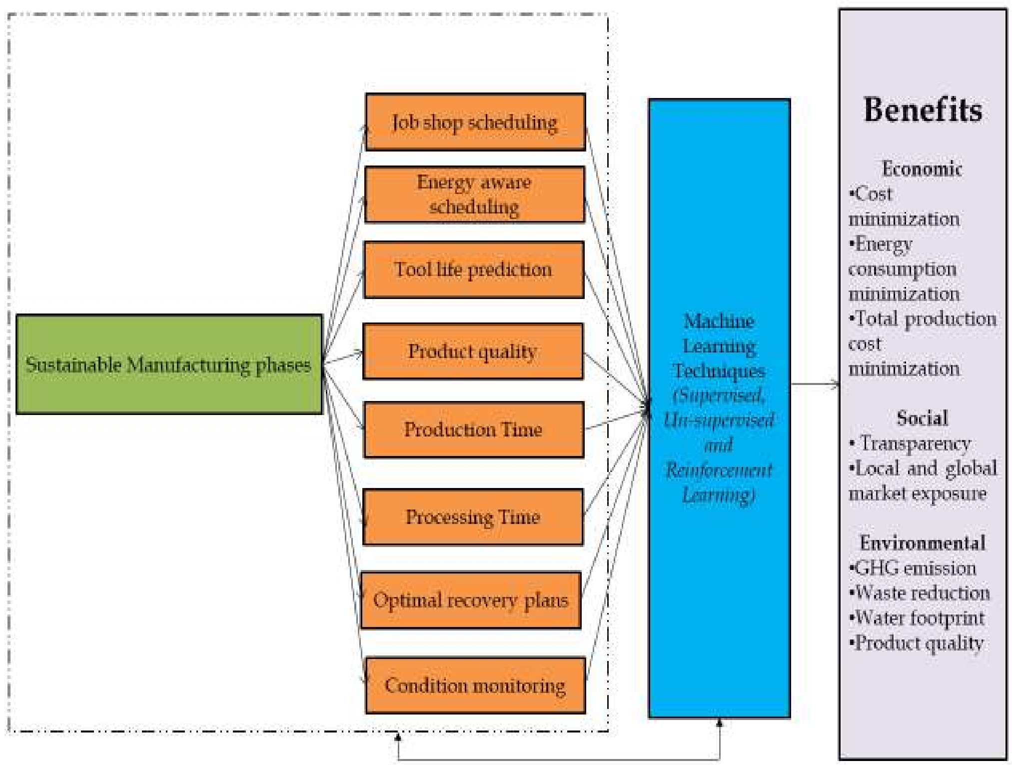

Fig. 10: ML-based SM framework

4.1.1 Sustainable manufacturing phases

In the proposed ML-SM performance framework the first phase includes the different phases of SM i.e. pre-production or planning, processing phase, production phase and product recovery phase. In these phases of SM, a lot of data is generated which is of large volume and complex in nature. ML techniques are suitable for handling such type of complex data. The present study reveals that ML techniques improve the performance of SM system e.g. condition monitoring, tool life predictions, sustainable planning and scheduling, machining parameters optimization, shop floor optimization, product recovery plans, simulation systems, system integrations, supply chain, IoT integration and additive manufacturing. The data generated from the pre-production phase can be used for decision making in the production phase.

4.1.2 ML algorithms

The ML algorithm is the second main component of the framework which include different types of ML techniques. As discussed in our study ML techniques can handle large data sets which are even complex. But every ML technique has its own advantages and limitations. The proposed framework in our study suggests that ML techniques should have a feedback loop capable of extracting the information from data generated from the machines. This data may be in a complex form which can be extracted by the use of an appropriate ML technique. The best outcome of this phase improves the performance of manufacturing industries by balancing the sustainability dimensions.

4.1.3 Performance of sustainable manufacturing system

The findings of the present study reveal that manufacturing data extracted from the machines using the ML algorithms can be used to develop a better and efficient SM system. The adoption of ML approaches not only helps to achieve economic benefits but also helps to improve the environmental and social performance of the industry. We have found from the cluster analysis that ML techniques help to optimize the various objectives in planning and scheduling problems, developing energy-aware oriented models and machining parameter optimization thus improving the sustainability of manufacturing industries. The detailed analysis of the various studies in sub-research areas is presented in the cluster analysis section. The discussions help us to include the SM system performance as the third main component of the proposed framework. This proposed framework addresses the three main dimensions (economic, social and environmental) of SM.

5. Implications of the study

A large number of experts and authors in the field of SM support the view that there is a dearth of thorough systematic literature review-based study on ML applications in SM. This has motivated us to conduct a systematic literature review and bibliometric analysis while also keeping the sustainability theme in our mind. Below are some of the implications for researchers, practitioners and policymakers to drive future research agenda.

5.1 Theoretical contributions

This article discusses the opportunities for ML applications in SM. Moreover, this article shows the key contributions in this area, what topics have been explored in context to SM, what are the strengths and weakness of existing literature available on ML applications in SM. The identification of themes and clusters i.e. ML in sustainable planning and scheduling, condition monitoring, optimal recovery plans, quality management is the uniqueness of this article. In the present study, we have identified six clusters based on the research areas. These clusters have been discussed in Table 12. Thus our study provides a comprehensive understanding of the ML application in a SM context.

5.2 Managerial Contributions

This review article delivers some meaningful insights which will be equally helpful for both the practitioners and policymakers as it helps them to understand the importance of ML applications in the manufacturing industries to achieve sustainability. The development in Industry 4.0 and sustainability concerns are decreasing the manual jobs of skilled people due to advanced technologies such as cyber-physical systems, internet of things, additive manufacturing and smart manufacturing. In this study, we have discussed different research opportunities in manufacturing with ML techniques which will be helpful for industries to create more jobs in Industry 4.0 also. Secondly, managers and practitioners must assess the themes discussed in this study. The findings will help practitioners to understand how ML techniques can be applied to achieve sustainability in manufacturing practices. The question that arises from the present study is whether the adoption of ML techniques in industries provides uniform benefits to all industry sectors or countries. There is a need to conduct short surveys to investigate whether this technological advancement in the industry is contributing to the betterment of industry or leading towards the destruction of manufacturing activities. This issue can be addressed by managers or policymakers in future.

5.3 Implications for researchers

Based on the bibliometric analysis and systematic literature review we present the following broad research areas for researchers which needs further investigations:

- The present study discusses how ML techniques have changed SM practices in the manufacturing industries. This study is mainly focused on the different ML applications in the SM and less attention has been paid to the data capturing from the various machines, i.e., how data is captured, synthesized and used for optimization purposes in manufacturing.
- The finding of our study indicates that applications of ML techniques have changed the SM scenario in the industries. We also found that these techniques are helping the industries in the adoption of Industry 4.0 practices. It would be more interesting to know the impact of ML techniques in Industry 4.0 with SM perspective which can be addressed in future studies. This will also help to provide the specific guidelines that how ML techniques can be helpful for SM in Industry 4.0 for manufacturing systems optimization, condition monitoring and recovery plans.
- In future, more studies can be conducted on ML application in SM which will help in optimal utilization of energy resources in machining processes. As machining in manufacturing is a major area where more amount of energy is consumed.
- There is also a need for empirical studies focusing on the identification of key drivers and barriers for the implementation of ML approaches in SM. Additionally, the interrelationship between these drivers and barriers can be identified with various techniques such as ISM modelling. The identification of drivers and barriers will help in expediting the ML applications in SM.
- Future studies can further focus on understanding how the adoption of ML techniques with SM practices can transform the existing production systems to production systems in Industry 4.0.
- Finally, future studies can also explore the performance of different ML techniques and provide a comparative assessment in a single problem. For example: comparing the job scheduling problem with NSGA-II algorithm with some other ML technique.

6. Conclusion

The present study focused on the systematic and bibliometric review of ML applications in SM which is an emerging research area. The literature review available in any research area tends to be broad and complete coverage of all published research documents is very challenging and almost impossible. Therefore, in the present study, a bibliometric study of ML applications in SM was conducted using the Scopus database and analyzing 96 articles between the period 2009-2020. In the bibliometric analysis top journals, top countries, top authors, top institutes and publication trend in last 11 years is analyzed together with the citation structure. Our findings show that the Journal of Cleaner Production and Procedia Manufacturing are the leading sources that publish the articles in ML applications for SM. The main contribution of this study is the development of a ML-based SM framework which will be beneficial for the industries to achieve sustainability in their practices. The study reveals that the number of ML techniques that have been introduced in the last few years of them has their advantages and limitations. The adoption of ML techniques helps to reduce the cycle time and handle complex data generated from the manufacturing processes which can be further used for monitoring and prediction purposes. As with any research, every research has certain limitations. This article has also certain limitations. We have limited our study to only the SCOPUS database and in the future other databases can be included. Other tools like Gephi and BibExcel can also be used in future research studies. Although, our research team have identified six major clusters in the different research areas where researchers have applied ML techniques. The proposed framework in this study is based on the literature finding which has not been empirically investigated. In future studies, this proposed framework can be validated by conducting survey-based studies. Moreover, future studies can also explore the applications of ML techniques in sustainable manufacturing in different industry sectors and different regions of the world which will help us to gain a holistic understanding of the potential of ML techniques.

References

Abdirad, M., & Krishnan, K. (2020). Industry 4.0 in Logistics and Supply Chain Management: A Systematic Literature Review. *EMJ - Engineering Management Journal*, *00*(00), 1–15. https://doi.org/10.1080/10429247.2020.1783935

Ahuja, J., Panda, T. K., Luthra, S., Kumar, A., Choudhary, S., & Garza-Reyes, J. A. (2019). Do human critical success factors matter in adoption of sustainable manufacturing practices? An influential mapping analysis of multi-company perspective. *Journal of Cleaner Production*, *239*. https://doi.org/10.1016/j.jclepro.2019.117981

Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

Antony, J., Psomas, E., Garza-Reyes, J. A., & Hines, P. (2020). Practical implications and future research agenda of lean manufacturing: A systematic literature review. *Production Planning {\&} Control*, 1–37.

Bag, S., Telukdarie, A., Pretorius, J. H. C., & Gupta, S. (2018). Industry 4.0 and supply chain sustainability: Framework and future research directions. *Benchmarking*. https://doi.org/10.1108/BIJ-03-2018-0056

Baker, H. K., Kumar, S., & Pandey, N. (2020). Correction to: Thirty years of small business economics: A bibliometric overview. *Small Business Economics*, *56*(1), 519–520. https://doi.org/10.1007/s11187-020-00368-2

Bhanot, N., Rao, P. V., & Deshmukh, S. G. (2017). An integrated approach for analysing the enablers and barriers of sustainable manufacturing. *Journal of Cleaner Production*, *142*, 4412–4439. https://doi.org/10.1016/j.jclepro.2016.11.123

Bhinge, R., Park, J., Law, K. H., Dornfeld, D. A., Helu, M., & Rachuri, S. (2017). Towards a generalized energy prediction model for machine tools. *Journal of Manufacturing Science and Engineering*, *139*(4), 041013. PubMed. https://doi.org/10.1115/1.4034933

Camposeco-Negrete, C. (2020). Optimization of printing parameters in fused deposition modeling for improving part quality and process sustainability. *International Journal of Advanced Manufacturing Technology*, *108*(7–8), 2131–2147. https://doi.org/10.1007/s00170-020-05555-9

Camposeco-Negrete, Carmita. (2015). Optimization of cutting parameters using Response Surface Method for minimizing energy consumption and maximizing cutting quality in turning of AISI 6061 T6 aluminum. *Journal of Cleaner Production*, *91*, 109–117. https://doi.org/10.1016/j.jclepro.2014.12.017

Cavalcante, I. M., Frazzon, E. M., Forcellini, F. A., & Ivanov, D. (2019). A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing. *International Journal of Information Management*, *49*, 86–97. https://doi.org/10.1016/j.ijinfomgt.2019.03.004

Cui, W., Li, L., & Lu, Z. (2019). Energy-efficient scheduling for sustainable manufacturing systems with renewable energy resources. *Naval Research Logistics*, *66*(2), 154–173. https://doi.org/10.1002/nav.21830

Cui, Weiwei, Sun, H., & Xia, B. (2019). Integrating production scheduling, maintenance planning and energy controlling for the sustainable manufacturing systems under TOU tariff. *Journal of the Operational Research Society*, *71*(11), 1760–1779. https://doi.org/10.1080/01605682.2019.1630327

Dai, M., Tang, D., Xu, Y., & Li, W. (2014). Energy-aware integrated process planning and scheduling for job shops. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *229*(1_suppl), 13–26. https://doi.org/10.1177/0954405414553069

Deng, Z., Zhang, H., Fu, Y., Wan, L., & Liu, W. (2017). Optimization of process parameters for minimum energy consumption based on cutting specific energy consumption. *Journal of Cleaner Production*, *166*, 1407–1414. https://doi.org/10.1016/j.jclepro.2017.08.022

Despeisse, M., Mbaye, F., Ball, P. D., & Levers, A. (2012). The emergence of sustainable manufacturing practices. *Production Planning and Control*, *23*(5), 354–376. https://doi.org/10.1080/09537287.2011.555425

Dev, N. K., Shankar, R., & Qaiser, F. H. (2020). Industry 4.0 and circular economy: Operational excellence for sustainable reverse supply chain performance. *Resources, Conservation and Recycling*, *153*. https://doi.org/10.1016/j.resconrec.2019.104583

Du, C.-J., & Sun, D.-W. (2006). Learning techniques used in computer vision for food quality evaluation: A review. *Journal of Food Engineering*, *72*(1), 39–55. https://doi.org/10.1016/j.jfoodeng.2004.11.017

Dubey, R., Gunasekaran, A., & Chakrabarty, A. (2015). World-class sustainable manufacturing: Framework and a performance measurement system. *International Journal of Production Research*, *53*(17), 5207–5223. https://doi.org/10.1080/00207543.2015.1012603

Dubey, R., Gunasekaran, A., Childe, S. J., Wamba, S. F., & Papadopoulos, T. (2016). The impact of big data on world-class sustainable manufacturing. *International Journal of Advanced Manufacturing Technology*, *84*(1–4), 631–645. https://doi.org/10.1007/s00170-015-7674-1

Ferguson, M., Bhinge, R., Park, J., Lee, Y. T., & Law, K. H. (2018). A Data Processing Pipeline for Prediction of Milling Machine Tool Condition from Raw Sensor Data. *Smart and Sustainable Manufacturing Systems*, *2*(1), 20180019. https://doi.org/10.1520/ssms20180019

Garetti, M., & Taisch, M. (2012). Sustainable manufacturing: Trends and research challenges. *Production Planning and Control*, *23*(2–3), 83–104. https://doi.org/10.1080/09537287.2011.591619

Giovannini, A., Aubry, A., Panetto, H., Dassisti, M., & El Haouzi, H. (2012). Ontology-based system for supporting manufacturing sustainability. *Annual Reviews in Control*, *36*(2), 309–317. https://doi.org/10.1016/j.arcontrol.2012.09.012

Haapala, K. R., Zhao, F., Camelio, J., Sutherland, J. W., Skerlos, S. J., Dornfeld, D. A., Jawahir, I. S., Clarens, A. F., & Rickli, J. L. (2013). A review of engineering research in sustainable manufacturing. *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, *135*(4). https://doi.org/10.1115/1.4024040

Huang, J., Jin, L., & Zhang, C. (2017). Mathematical Modeling and a Hybrid NSGA-II Algorithm for Process Planning Problem Considering Machining Cost and Carbon Emission. *Sustainability*, *9*(10), 1769. https://doi.org/10.3390/su9101769

ITA, U. (2012). How does commerce define sustainable manufacturing? *Online] Http://Trade. Gov/Competitiveness/Sustainablemanufacturing/How_doc_defines_SM. Asp*.

Jamwal, A., Agrawal, R., Sharma, M., & Kumar, V. (2021). Review on multi-criteria decision analysis in sustainable manufacturing decision making. *International Journal of Sustainable Engineering*. https://doi.org/10.1080/19397038.2020.1866708

Jamwal, Anbesh, Agrawal, R., Manupati, V. K., Sharma, M., Varela, L., & Machado, J. (2020). Development of cyber physical system based manufacturing system design for process optimization. *IOP Conference Series: Materials Science and Engineering*, *997*, 012048. https://doi.org/10.1088/1757-899x/997/1/012048

Jamwal, Anbesh, Agrawal, R., Sharma, M., Dangayach, G. S., & Gupta, S. (2021). Application of optimization techniques in metal cutting operations: A bibliometric analysis. *Materials Today: Proceedings*, *38*, 365–370. https://doi.org/10.1016/j.matpr.2020.07.425

Jamwal, Anbesh, Agrawal, R., Sharma, M., Kumar, V., & Kumar, S. (2021). Developing A sustainability framework for Industry 4.0. *Procedia CIRP*, *98*, 430–435.

Joung, C. B., Carrell, J., Sarkar, P., & Feng, S. C. (2013). Categorization of indicators for sustainable manufacturing. *Ecological Indicators*, *24*, 148–157. https://doi.org/10.1016/j.ecolind.2012.05.030

Khatri, A., Garg, D., & Dangayach, G. S. (2019). Critical success factors of sustainable manufacturing and procurement: An empirical study. *International Journal of Social Ecology and Sustainable Development*, *10*(3), 17–27. https://doi.org/10.4018/IJSESD.2019070102

Kodratoff, Y. (2014). *Introduction to machine learning*. Elsevier.

Kubat, M. (2017). *An introduction to machine learning*. Springer.

Kuik, S. S., Kaihara, T., & Fujii, N. (2016). Product Recovery Configuration Decisions for Achieving Sustainable Manufacturing. In T. R (Ed.), *Procedia CIRP* (Vol. 41, pp. 258–263). Elsevier B.V. https://doi.org/10.1016/j.procir.2016.01.195

Kumar, R., Singh, S. P., & Lamba, K. (2018). Sustainable robust layout using Big Data approach: A key towards industry 4.0. *Journal of Cleaner Production*, *204*, 643–659. https://doi.org/10.1016/j.jclepro.2018.08.327

Li, X., & Lu, J. (2017, April). *Machining process energy consumption modelling using response surface methodology and neural network*. 2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD). https://doi.org/10.1109/cscwd.2017.8066724

Lieber, D., Stolpe, M., Konrad, B., Deuse, J., & Morik, K. (2013). Quality Prediction in Interlinked Manufacturing Processes based on Supervised & Unsupervised Machine Learning. *Procedia CIRP*, *7*, 193–198. https://doi.org/10.1016/j.procir.2013.05.033

Lison, P. (2015). An introduction to machine learning. *Language Technology Group (LTG)*, *1*(35).

Liu, C., Yang, J., Lian, J., Li, W., Evans, S., & Yin, Y. (2014). Sustainable performance oriented operational decision-making of single machine systems with deterministic product arrival time. *Journal of Cleaner Production*, *85*, 318–330. https://doi.org/10.1016/j.jclepro.2014.07.025

Liu, Y., Dong, H., Lohse, N., & Petrovic, S. (2015). Reducing environmental impact of production during a Rolling Blackout policy – A multi-objective schedule optimisation approach. *Journal of Cleaner Production*, *102*, 418–427. https://doi.org/10.1016/j.jclepro.2015.04.038

Loyer, J.-L., Henriques, E., Fontul, M., & Wiseall, S. (2016). Comparison of Machine Learning methods applied to the estimation of manufacturing cost of jet engine components. *International Journal of Production Economics*, *178*, 109–119. https://doi.org/10.1016/j.ijpe.2016.05.006

Machado, C. G., Winroth, M. P., & Ribeiro da Silva, E. H. D. (2020). Sustainable manufacturing in Industry 4.0: An emerging research agenda. *International Journal of Production Research*, *58*(5), 1462–1484. https://doi.org/10.1080/00207543.2019.1652777

Majorel Deutschland GmbH. (n.d.). *Artificial Intelligence and Sustainability*. Retrieved January 8, 2020, from https://www.future- customer.com/artificial-intelligence-and-sustainability/

Malek, J., & Desai, T. N. (2020). A systematic literature review to map literature focus of sustainable manufacturing. *Journal of Cleaner Production*, *256*. https://doi.org/10.1016/j.jclepro.2020.120345

Manupati, V. K., Xavior, M. A., Chandra, A., & Ahsan, M. (2018). Workload assessment for a sustainable manufacturing paradigm using social network analysis method. In *Knowledge Computing and Its Applications: Knowledge Manipulation and Processing Techniques: Volume 1* (p. 108). Springer Singapore. https://doi.org/10.1007/978-981-10-6680-1_5

Martinelli, F. (1991). A demand-oriented approach to understanding producer services. *The Changing Geography of Advanced Producer Services*, 15–29.

May, G., Stahl, B., Taisch, M., & Prabhu, V. (2015). Multi-objective genetic algorithm for energy-efficient job shop scheduling. *International Journal of Production Research*, *53*(23), 7071–7089. https://doi.org/10.1080/00207543.2015.1005248

Meng, K., Qian, X., Lou, P., & Zhang, J. (2020). Smart recovery decision-making of used industrial equipment for sustainable manufacturing: Belt lifter case study. *Journal of Intelligent Manufacturing*, *31*(1), 183–197. https://doi.org/10.1007/s10845-018-1439-2

Parsons, S. (2005). Introduction to Machine Learning by Ethem Alpaydin, MIT Press, 0-262-01211-1, 400 pp., $50.00/£32.95. *The Knowledge Engineering Review*, *20*(4), 432–433. https://doi.org/10.1017/s0269888906220745

Preez, A. D., & Oosthuizen, G. A. (2019). Machine learning in cutting processes as enabler for smart sustainable manufacturing. In S. G. Badurdeen F. Kohl H. ,. Jawahir I. S. (Ed.), *Procedia Manufacturing* (Vol. 33, pp. 810–817). Elsevier B.V. https://doi.org/10.1016/j.promfg.2019.04.102

PRIORE, P., DE LA FUENTE, D., GOMEZ, A., & PUENTE, J. (2001). A review of machine learning in dynamic scheduling of flexible manufacturing systems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *15*(3), 251–263. https://doi.org/10.1017/s0890060401153059

Sealy, M. P., Liu, Z. Y., Zhang, D., Guo, Y. B., & Liu, Z. Q. (2016). Energy consumption and modeling in precision hard milling. *Journal of Cleaner Production*, *135*, 1591–1601. https://doi.org/10.1016/j.jclepro.2015.10.094

Seliger, G., Kim, H.-J., Kernbaum, S., & Zettl, M. (2008). Approaches to sustainable manufacturing. *International Journal of Sustainable Manufacturing*, *1*(1–2), 58–77. https://doi.org/10.1504/IJSM.2008.019227

Sharma, R., Jabbour, C. J. C., & Lopes de Sousa Jabbour, A. B. (2020). Sustainable manufacturing and industry 4.0: What we know and what we don't. *Journal of Enterprise Information Management*. https://doi.org/10.1108/JEIM-01-2020-0024

Sharma, Rohit, Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, *119*, 104926. https://doi.org/10.1016/j.cor.2020.104926

Sharp, M., Ak, R., & Hedberg, T., Jr. (2018). A Survey of the Advancing Use and Development of Machine Learning in Smart Manufacturing. *Journal of Manufacturing Systems*, *48 Pt C*(Pt C), 10.1016/j.jmsy.2018.02.004. PubMed. https://doi.org/10.1016/j.jmsy.2018.02.004

Stock, T., & Seliger, G. (2016). Opportunities of Sustainable Manufacturing in Industry 4.0. In M. J. Seliger G. Kohl H. (Ed.), *Procedia CIRP* (Vol. 40, pp. 536–541). Elsevier B.V. https://doi.org/10.1016/j.procir.2016.01.129

Syafrudin, M., Alfian, G., Fitriyani, N. L., & Rhee, J. (2018). Performance Analysis of IoT-Based Sensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing. *Sensors (Basel, Switzerland)*, *18*(9), 2946. PubMed. https://doi.org/10.3390/s18092946

Tayal, A., Solanki, A., & Singh, S. P. (2020). Integrated frame work for identifying sustainable manufacturing layouts based on big data, machine learning, meta-heuristic and data envelopment analysis. *Sustainable Cities and Society*, *62*. https://doi.org/10.1016/j.scs.2020.102383

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, *14*(3), 207–222.

Wang, L., Törngren, M., & Onori, M. (2015). Current status and advancement of cyber-physical systems in manufacturing. *Journal of Manufacturing Systems*, *37*, 517–527. https://doi.org/10.1016/j.jmsy.2015.04.008

Wu, D., Jennings, C., Terpenny, J., Gao, R. X., & Kumara, S. (2017). A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. *Journal of Manufacturing Science and Engineering*, *139*(7). https://doi.org/10.1115/1.4036350

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517

Yadav, G., Kumar, A., Luthra, S., Garza-Reyes, J. A., Kumar, V., & Batista, L. (2020). A framework to achieve sustainability in manufacturing organisations of developing economies using industry 4.0 technologies' enablers. *Computers in Industry*, *122*. https://doi.org/10.1016/j.compind.2020.103280

Zhang, H., Deng, Z., Fu, Y., Lv, L., & Yan, C. (2017). A process parameters optimization method of multi-pass dry milling for high efficiency, low energy and low carbon emissions. *Journal of Cleaner Production*, *148*, 174–184. https://doi.org/10.1016/j.jclepro.2017.01.077