Received 16 November 2022
Revised 29 June 2023
Accepted 3 October 2023

Machine learning for sustainable development: leveraging technology for a greener future

Muneza Kagzi
T A Pai Management Institute, Manipal Academy of Higher Education, Manipal, India

Sayantan Khanra
School of Business Management, Narsee Monjee Institute of Management Studies University, Mumbai, India, and

Sanjoy Kumar Paul
UTS Business School, University of Technology Sydney Business School, Sydney, Australia

Abstract

Purpose – From a technological determinist perspective, machine learning (ML) may significantly contribute towards sustainable development. The purpose of this study is to synthesize prior literature on the role of ML in promoting sustainability and to encourage future inquiries.

Design/methodology/approach – This study conducts a systematic review of 110 papers that demonstrate the utilization of ML in the context of sustainable development.

Findings – ML techniques may play a vital role in enabling sustainable development by leveraging data to uncover patterns and facilitate the prediction of various variables, thereby aiding in decision-making processes. Through the synthesis of findings from prior research, it is evident that ML may help in achieving many of the United Nations' sustainable development goals.

Originality/value – This study represents one of the initial investigations that conducted a comprehensive examination of the literature concerning ML's contribution to sustainability. The analysis revealed that the research domain is still in its early stages, indicating a need for further exploration.

Keywords Artificial intelligence, Clean energy, Industry innovation, Learning algorithms, Responsible consumption, Smart cities

Paper type Literature review

1. Introduction

As global agreement to take proactive measures to save the planet is increasing (Pérez-Ortiz *et al.*, 2014), policymakers are soliciting proactive solutions for sustainable development (Zhao *et al.*, 2020). Sustainable development is used as a strategic approach to achieve economic growth while optimizing immediate and long-term impacts on people and planet (Elkington, 1998). In this context, the effective utilization of machine learning (ML) techniques has emerged as a potentially vital tool (Abdella *et al.*, 2020). ML techniques leverage the analysis of extensive training data sets to identify patterns and generate predictive estimations (El-Alfy and Mohammed, 2020; Jing *et al.*, 2022). By leveraging statistical methods, these techniques enable prediction and classification based on key insights extracted from input data (Fahdi *et al.*, 2021).

Journal of Systems and Information Technology
Vol. 25 No. 4, 2023
pp. 440-479

1328-7265
DOI 10.1108/JSIT-11-2022-0266

The literature on the applications of ML to achieve sustainability is continuously progressing. Noteworthy applications of ML are reported regarding many dimensions of sustainability, including the supply chain (De Brito *et al.*, 2008; Hall *et al.*, 2012; Sharma *et al.*, 2020; Wu and Pagell, 2011), cleaner energy (Sohani *et al.*, 2022a), waste management (Li *et al.*, 2022b), among others. Technological determinism argues that advancements in technology drive socioeconomic transformations (Smith and Marx, 1994; Veblen, 1978). Thus, it is important to explore the role of ML to achieve sustainability. However, the extant literature observes a paucity of a comprehensive synthesis encompassing the various ways in which ML contributes to sustainable development. The objective of this study is to advance the extant literature by addressing this research gap.

The United Nations has outlined a set of 17 sustainable development goals (SDGs), which serve as a comprehensive guide for attaining a sustainable future by 2030 [Department of Economic and Social Affairs (DESA), 2022]. These SDGs hold significant importance owing to their broad scope and dedicated efforts to tackle global challenges (DESA, 2022). In the quest to synthesize existing literature on the role of ML in promoting sustainability, we aim to address the following research questions (RQs):

RQ1. What is the present state of literature on ML for sustainability?

RQ2. How is ML addressing different SDGs?

RQ3. Which are the avenues for future research on ML for sustainability?

The article is structured as follows: Section 2 provides an overview of the relationship between sustainable development and ML. Section 3 outlines the methodology used to identify relevant studies and presents a bibliometric analysis of 110 selected studies. Section 4 summarizes the contributions of ML to 13 SDGs. Section 5 discusses various ML algorithms examined in the reviewed studies. Section 6 highlights potential areas for future research. Finally, Section 7 concludes the article.

2. Conceptual background

2.1 Sustainable development

Elkington (1998) introduced the triple bottom-line framework as a means of assessing corporate performance, highlighting the importance of economic, environmental and social criteria. Businesses often undertake socially responsible activities that align with their core values (Mansouri and Momtaz, 2022; Robert *et al.*, 2005; Svanberg *et al.*, 2022). An increasingly common approach to this is the implementation of green initiatives, which involve the development of recyclable products, reduction of environmental impacts in manufacturing facilities, and pursuit of green certification for overall operations (Gimenez *et al.*, 2012; Jacobs *et al.*, 2010; Walk *et al.*, 2023). Such initiatives not only mitigate manufacturing costs by using recycled materials and reducing energy consumption but also offer financial benefits through improved environmental performance (Ambec and Lanoie, 2008). In the context of logistics operations, it is advisable for companies to incorporate environmental considerations into their decision-making processes (Agrawal *et al.*, 2021; Arunmozhi *et al.*, 2022; Bhattacharya *et al.*, 2014; Campbell, 2007; Seuring, 2013). Collaborative efforts with network partners can lead to optimal solutions, achieved through coordinated business forecasting, resource sharing in transportation and warehousing and efficient supply route planning (De Brito *et al.*, 2008). Nevertheless, firms often harbour doubts about the financial returns associated with the initial investment in environmentally conscious initiatives (Gimenez *et al.*, 2012).

2.2 Machine learning

With a wide range of algorithms at its disposal, ML offers robust solutions across various scientific research domains (Kondapaka *et al.*, 2023). The application of ML contributes to enhanced accuracy and efficiency in data analysis, empowering organizations and policymakers to extract valuable insights pertaining to sustainability challenges (Svanberg *et al.*, 2022). Through the processing of voluminous data, ML algorithms discern patterns, trends and correlations that may elude human observers, thereby facilitating well-informed decision-making (Kazemeini and Swei, 2023; Lu *et al.*, 2022). ML further aids in predictive modelling, enabling the identification of potential environmental risks and resource constraints and promoting proactive strategies for mitigation (Carrera *et al.*, 2022; Li *et al.*, 2023).

ML optimizes resource allocation and utilization by identifying inefficiencies and recommending improvements, thereby minimizing waste and fostering resource conservation (Shahidzadeh *et al.*, 2022; Abbate *et al.*, 2023). In sectors such as energy (Lohrmann *et al.*, 2022), transportation (Liu *et al.*, 2021a, 2021b; Ma *et al.*, 2022) and agriculture (Krause and Bokinala, 2023; Prioux *et al.*, 2023), ML optimizes processes (Azadi *et al.*, 2023) and promotes sustainable practices (Golafshani *et al.*, 2023). Furthermore, ML-powered technologies facilitate the development of smart cities (De Las Heras *et al.*, 2020) and enhance the quality of life for residents (Cai *et al.*, 2022; Khanna and Khanra, 2023). By harnessing the capabilities of ML, companies may address business problems while generating positive environmental and social outcomes (Dong *et al.*, 2023; Wang *et al.*, 2022a, 2022b, 2022c). An overview of previous similar review studies on the application of ML in sustainability is presented in Table 1. In Table 1, it can be seen that the existing literature review studies have discussed the one particular SGD, whereas the present study looks at the multiple SGDs to capture the wider aspect of sustainability.

3. Research methodology

The present study adopted a standard protocol to conduct a systematic literature review (Khanra *et al.*, 2020b). The protocol involves three sequential phases to plan, perform and present the review, as presented in Figure 1.

3.1 Planning the review

The present research's intellectual interest is to summarize ML's roles in sustainability. Therefore, two sets of keywords related to ML and sustainability are required to specify the conceptual boundaries of the intellectual interest (Tandon *et al.*, 2020). The following keywords are identified from a preliminary exploration of the databases:

- *Set 1*: ML, deep learning (DL), learning algorithms, learning systems and predictive modelling.
- *Set 2*: sustainability, sustainable development, circular economy and carbon footprint.

Scopus is recognized as the most extensive abstract and citation database in management (Khanra *et al.*, 2020a; Tandon *et al.*, 2020). Hence, the extant literature is identified from the Scopus database. The date of exploration is 31st May 2023.

3.2 Performing the review

The search was conducted using a pairwise query approach, with each query consisting of a single keyword from two distinct sets. The resulting initial search queries generated a

Sr. no.	Citation	SGD
1	D'Amato, D., Droste, N., Allen, B., Kettunen, M., Lähtinen, K., Korhonen, J., [. . .] and Toppinen, A. (2017). Green, circular, bio economy: A comparative analysis of sustainability avenues. *Journal of Cleaner Production*, 168, 716-734.	SGD 12: Responsible consumption and production
2	Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V. and Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers and Operations Research*, 119, 104,926.	SGD 12: Responsible consumption and production
3	Agrawal, R., Wankhede, V. A., Kumar, A., Luthra, S., Majumdar, A. and Kazancoglu, Y. (2021). An exploratory state of the art review of artificial intelligence applications in circular economy using structural topic modelling. *Operations Management Research*, 1–18	SGD-9- Industry-Innovation and Infrastructure
4	Lee, C. and Lim, C. (2021). From technological development to social advance: A review of Industry 4.0 through machine learning. *Technological Forecasting and Social Change*, 167, 120,653	SGD-9- Industry-Innovation and Infrastructure
5	Grybauskas, A., Stefanini, A. and Ghobakhloo, M. (2022). Social sustainability in the age of digitalization: A systematic literature review on the social implications of Industry 4.0. *Technology in Society*, 101,997	SGD-9- Industry-Innovation and Infrastructure
6	Li, J., Suvarna, M., Li, L., Pan, L., Pérez-Ramírez, J., Ok, Y. S. and Wang, X. (2022). A review of computational modeling techniques for wet waste valorization: Research trends and future perspectives. *Journal of Cleaner Production*, 133,025	SGD-11- Sustainable cities and communities
7	Rieder, E., Schmuck, M. and Tugui, A. (2022). A Scientific Perspective on Using Artificial Intelligence in Sustainable Urban Development. *Big Data and Cognitive Computing*, 7(1), 3.	SGD-11- Sustainable cities and communities
8	Sohani, A., Sayyaadi, H., Cornaro, C., Shahverdian, M. H., Pierro, M., Moser, D., [. . .] and Li, L. K. (2022). Using machine learning in photovoltaics to create smarter and cleaner energy generation systems: A comprehensive review. *Journal of Cleaner Production*, 132,701	SGD-9- Industry-Innovation and Infrastructure
9	Wang, J. and Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, 129, 103,925	SGD-11- Sustainable cities and communities
10	Abbate, S., Centobelli, P. and Cerchione, R. (2023). The digital and sustainable transition of the agri-food sector. *Technological Forecasting and Social Change*, 187, 122,222	SGD-9- Industry-Innovation and infrastructure
11	Mansoursamaei, M., Moradi, M., González-Ramírez, R. G. and Lalla-Ruiz, E. (2023). Machine Learning for Promoting Environmental Sustainability in Ports. *Journal of Advanced Transportation*, 2023	SGD-9- Industry-Innovation and Infrastructure

Table 1.
Prior literature review studies

corpus of 394 research articles in the field of management and related disciplines published in English-language journals. To ensure methodological rigour, this study excluded documents published in books, conference proceedings and trade journals. Subsequently, a meticulous evaluation of the abstracts led to the exclusion of 113 irrelevant articles, leaving 281 articles for full-text examination. Ultimately, a set of 110 articles was identified as pertinent to the intellectual pursuits of this investigation. Details about these articles are available from Table 2.

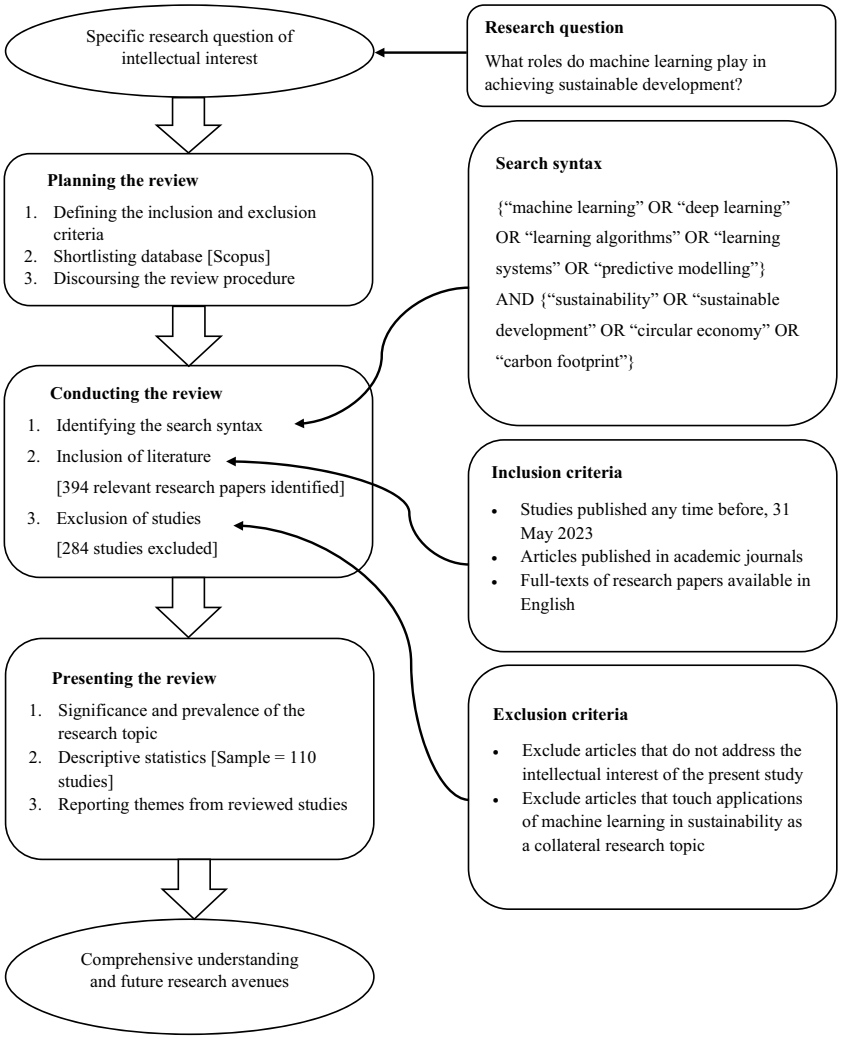


Figure 1.
Literature review approach followed in this study

Source: Created by authors

3.3 Bibliometric analysis

A total of 219 authors affiliated with 207 organizations located in 41 countries contributed to the sample of 110 articles. Year-wise distribution of the articles in our sample reveals that the topic under investigation is relatively new in the literature, with the oldest study being published in 2014. However, since then, there has been a noticeable upward trend in the number of papers published on the subject (see Figure 2).

This study conducts bibliographic coupling and co-authorship analysis using VOSviewer, which is a reliable software to analyse bibliometric data (Cobo *et al.*, 2011). Also, VOSviewer is an efficient software for analyzing co-occurrences of keywords from bibliometric data sets

Table 2.
List of studies considered for the literature review

SGD	Authors (year)	Title	Journal
Zero hunger (SDG 2)	Abdella *et al.* (2020)	Sustainability assessment and modeling based on supervised machine learning techniques: the case for food consumption	*Journal of Cleaner Production*
	Bedi (2022)	Transfer learning augmented enhanced memory network models for reference evapotranspiration estimation	*Knowledge-Based Systems*
	Jato-Espino and Mayor-Vitoria (2023)	A statistical and machine learning methodology to model rural depopulation risk and explore its attenuation through agricultural land use management	*Applied Geography*
	Benites-Lazaro *et al.* (2018)	Sustainability and governance of sugarcane ethanol companies in Brazil: topic modeling analysis of CSR reporting	*Journal of Cleaner Production*
	Taghizadeh-Mehrjardi *et al.* (2020)	Land suitability assessment and agricultural production sustainability using machine learning models	*Agronomy*
Good health and well-being (SDG 3)	Zolbanin *et al.* (2022)	Data analytics for the sustainable use of resources in hospitals: predicting the length of stay for patients with chronic diseases	*Information and Management*
	Balfaqih *et al.* (2022)	An Accident Detection and Classification System Using Internet of Things and Machine Learning towards Smart City	*Sustainability*
	Lan *et al.* (2021)	Constructing urban sprawl measurement system of the Yangtze River economic belt zone for healthier lives and social changes in sustainable cities	*Technological Forecasting and Social Change*
	Al-Jamimi and Saleh (2019).	Transparent predictive modelling of catalytic hydrodesulfurization using an interval Type-2 fuzzy logic	*Journal of Cleaner Production*
Clean water and sanitation (SDG 6)	Dong *et al.* (2023)	Combined water quality forecasting system based on multi-objective optimization and improved data decomposition integration strategy	*Journal of Forecasting*
	Priyadarshini *et al.* (2022a)	Water pollution reduction for sustainable urban development using machine learning techniques	*Cities*
	Rahman *et al.* (2021)	Development of flood hazard map and emergency relief operation system using hydrodynamic modeling and machine learning algorithm	*Journal of Cleaner Production*
	Baroi *et al.* (2020)	Sustainability assessment of phosphorus in the waste management system of Bangladesh using substance flow analysis	*Journal of Cleaner Production*

(*continued*)

Table 2.

SGD	Authors (year)	Title	Journal
Affordable and clean energy (SDG 7)	Kosovic *et al.* (2020)	Using Artificial Intelligence on environmental data from Internet of Things for estimating solar radiation: comprehensive analysis	*Journal of Cleaner Production*
	Priyadarshini *et al.* (2022b)	A machine-learning ensemble model for predicting energy consumption in smart homes	*Internet of Things*
	Tsoka *et al.* (2022)	Explainable artificial intelligence for building energy performance certificate labelling classification	*Journal of Cleaner Production*
	Sohani *et al.* (2022a, 2022b)	Thermo-electro-environmental analysis of a photovoltaic solar panel using machine learning and real-time data for smart and sustainable energy generation	*Journal of Cleaner Production*
	Yin and Zhou (2022)	Performance evaluation of China's photovoltaic poverty alleviation project using machine learning and satellite images	*Utilities Policy*
	Coro and Trumpy (2020)	Predicting geographical suitability of geothermal power plants	*Journal of Cleaner Production*
	Redchuk *et al.* (2023)	Adoption Case of IIoT and Machine Learning to Improve Energy Consumption at a Process Manufacturing Firm, under Industry 5.0 Model	*Big Data and Cognitive Computing*
	Jamwal *et al.* (2022)	Deep learning for manufacturing sustainability: models, applications in Industry 4.0 and implications	*International Journal of Information Management Data Insights*
	Gao *et al.* (2023)	A digital twin-based approach for optimizing operation energy consumption at automated container terminals	*Journal of Cleaner Production*
	Leong *et al.* (2020)	Enhancing the adaptability: lean and green strategy towards the Industry Revolution 4.0	*Journal of Cleaner Production*
	Lohrmann *et al.* (2022)	Global and regional models for identification of cooling technology in thermal power generation for water demand estimations in water-energy nexus studies	*Journal of Cleaner Production*
	Zafar *et al.* (2022)	Towards green energy for sustainable development: machine learning based MPPT approach for thermoelectric generator	*Journal of Cleaner Production*
	Bakay and Ağbulut (2021)	Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms	*Journal of Cleaner Production*

(*continued*)

SGD	Authors (year)	Title	Journal
	Bas *et al.* (2021)	Classification of potential electric vehicle purchasers: a machine learning approach	*Technological Forecasting and Social Change*
	Aaldering and Song (2019)	Tracing the technological development trajectory in post-lithium-ion battery technologies: a patent-based approach	*Journal of Cleaner Production*
	Wang and Zhao (2021)	The impact of the global stock and energy market on EU ETS: a structural equation modelling approach	*Journal of Cleaner Production*
Decent work and economic growth (SDG 8)	Kim *et al.* (2021)	A value of civic voices for smart city: a big data analysis of civic queries posed by Seoul citizens	*Cities*
	De Las Heras *et al.* (2020)	Machine Learning Technologies for Sustainability in Smart Cities in the Post-COVID Era"	*Sustainability*
	Mele and Magazzino (2020)	A machine learning analysis of the relationship among iron and steel industries, air pollution and economic growth in China	*Journal of Cleaner Production*
	Saura *et al.* (2022)	Exploring the challenges of remote work on Twitter users' sentiments: from digital technology development to a post-pandemic era	*Journal of Business Research*
Industry, innovation and infrastructure (SDG 9)	Ashraf *et al.* (2022)	Developing a sustainable concrete incorporating bentonite clay and silica fume: mechanical and durability performance	*Journal of Cleaner Production*
	Huo *et al.* (2022)	Development of machine learning models for the prediction of the compressive strength of calcium-based geopolymers	*Journal of Cleaner Production*
	Golafshani *et al.* (2023)	Modeling the chloride migration of recycled aggregate concrete using ensemble learners for sustainable building construction	*Journal of Cleaner Production*
	Munir *et al.* (2022)	Development of novel design strength model for sustainable concrete columns: a new machine learning-based approach	*Journal of Cleaner Production*
	Naseri *et al.* (2020)	Designing sustainable concrete mixture by developing a new machine learning technique	*Journal of Cleaner Production*
	Sadrossadat *et al.* (2023)	Innovative AI-based multi-objective mixture design optimisation of CPB considering properties of tailings and cement	*International Journal of Mining, Reclamation and Environment*
	Adedoyin *et al.* (2022)	A simulation experiment on ICT and patent intensity in South Africa: an application of the novel dynamic ARDL machine learning model	*Technological Forecasting and Social Change*

(*continued*)

Table 2.

Table 2.

SGD	Authors (year)	Title	Journal
	Jana *et al.* (2021)	Determinants of electronic waste generation in Bitcoin network: evidence from the machine learning approach	*Technological Forecasting and Social Change*
	Naveed *et al.* (2017)	Co-evolution between streaming and live music leads a way to the sustainable growth of music industry–Lessons from the US experiences	*Technology in Society*
Reduced inequalities (SDG 10)	Hidalgo *et al.* (2020)	The digital divide in light of sustainable development: an approach through advanced machine learning techniques	*Technological Forecasting and Social Change*
	Lakner *et al.* (2022)	How much does reducing inequality matter for global poverty?	*The Journal of Economic Inequality*
	Ribeiro (2021)	Sulfur dioxide emissions in Portugal: prediction, estimation and air quality regulation using machine learning	*Journal of Cleaner Production*
	Budak and Sarvari (2021)	Profit margin prediction in sustainable road freight transportation using machine learning	*Journal of Cleaner Production*
	Li *et al.* (2022a, 2022b, 2022c, 2022d)	A review of computational modeling techniques for wet waste valorization: research trends and future perspectives	*Journal of Cleaner Production*
Sustainable cities and communities (SDG 11)	Li *et al.* (2019)	An agent-based procedure with an embedded agent learning model for residential land growth simulation: the case study of Nanjing, China	*Cities*
	Zhao *et al.* (2020)	China's population spatialization based on three machine learning models	*Journal of Cleaner Production*
	Wang *et al.* (2022a)	Spatiotemporal evolution of urban-agricultural-ecological space in China and its driving mechanism	*Journal of Cleaner Production*
	Liu *et al.* (2021a, 2021b)	Factors of ecosystem service values in a fast-developing region in China: insights from the joint impacts of human activities and natural conditions	*Journal of Cleaner Production*
	Kim *et al.* (2021)	A value of civic voices for smart city: a big data analysis of civic queries posed by Seoul citizens	*Cities*
	Lai *et al.* (2023)	An accurate and adaptable deep learning-based solution to floating litter cleaning up and its effectiveness on environmental recovery	*Journal of Cleaner Production*

(*continued*)

Table 2.

SGD	Authors (year)	Title	Journal
	Mansoursamaei *et al.* (2023)	Machine Learning for Promoting Environmental Sustainability in Ports	*Journal of Advanced Transportation*
Responsible consumption and production (SDG 12)	Gunawan *et al.* (2020)	Institutional pressures, environmental management strategy, and organizational performance: the role of environmental management accounting	*Business Strategy and the Environment*
	D'Amato *et al.* (2017)	Green, circular, bio economy: a comparative analysis of sustainability avenues	*Journal of Cleaner Production*
	Teng *et al.* (2023)	Machine-learned digital phase switch for sustainable chemical production"	*Journal of Cleaner Production*
	Nilashi *et al.* (2019a)	Preference learning for eco-friendly hotels recommendation: a multi-criteria collaborative filtering approach	*Journal of Cleaner Production*
	Jafari *et al.* (2022)	Producer-consumer sustainability continuum: mutual understanding to implement extended producer responsibility	*Journal of Cleaner Production*
	Arranz *et al.* (2022)	Institutional pressures as drivers of circular economy in firms: a machine learning approach	*Journal of Cleaner Production*
	Behera *et al.* (2023)	Creation of sustainable growth with explainable artificial intelligence: an empirical insight from consumer-packaged goods retailers	*Journal of Cleaner Production*
	Sharma *et al.* (2020)	A systematic literature review on machine learning applications for sustainable agriculture supply chain performance"	*Computers and Operations Research*
	Wu *et al.* (2020)	Partner selection in sustainable supply chains: a fuzzy ensemble learning model	*Journal of Cleaner Production*
	Miguéis *et al.* (2022)	Reducing fresh fish waste while ensuring availability: demand forecast using censored data and machine learning	*Journal of Cleaner Production*
	Ayeleru *et al.* (2021)	Forecasting municipal solid waste quantity using artificial neural network and supported vector machine techniques: a case study of Johannesburg, South Africa. *Journal of Cleaner Production*	*Journal of Cleaner Production*
	Gue *et al.* (2022)	Predicting waste management system performance from city and country attributes, *Journal of Cleaner Production*	*Journal of Cleaner Production* *Tourism Management*

(*continued*)

Table 2.

SGD	Authors (year)	Title	Journal
	Li *et al.* (2022a, 2022b, 2022c, 2022d)	Climate adaptation planning for cultural heritages in coastal tourism destinations: a multi-objective optimization approach	
Climate action (SDG 13)	Kumar *et al.* (2021)	A novel framework for risk assessment and resilience of critical infrastructure towards climate change	*Technological Forecasting and Social Change*
	Jin *et al.* (2023)	Empirical evidence of urban climate adaptation alignment with sustainable development: application of LDA	*Cities*
	Chiu *et al.* (2023)	Applying unsupervised learning method to develop a regional risk model based on TCFD: a case study in the United States	*Journal of Cleaner Production*
	Zhu *et al.* (2022)	China's population spatialization based on three machine learning models	*Journal of Cleaner Production*
	Li *et al.* (2022a, 2022b, 2022c, 2022d)	An accurate and adaptable deep learning-based solution to floating litter cleaning up and its effectiveness on environmental recovery	*Journal of Cleaner Production*
Life on land (SDG 15)	Bao and Xie (2022)	Artificial intelligence in animal farming: a systematic literature review	*Journal of Cleaner Production*
	Shafi *et al.* (2023)	Leveraging machine learning and remote sensing to monitor long-term spatial-temporal wetland changes: towards a national RAMSAR inventory in Pakistan	*Applied Geography*
Peace, justice and strong institutions (SDG 16)	Wang *et al.* (2020)	Detecting corporate misconduct through random forest in China's construction industry	*Journal of Cleaner Production*
	Jabeur *et al.* (2021)	CatBoost model and artificial intelligence techniques for corporate failure prediction	*Technological Forecasting and Social Change*
	Wilson and Van Der Velden (2022)	Sustainable AI: an integrated model to guide public sector decision-making	*Technology in Society*
	Nilashi *et al.* (2019b)	Measuring sustainability through ecological sustainability and human sustainability: a machine learning approach	*Journal of Cleaner Production*
	Pérez-Ortiz *et al.* (2014)	Classification of EU countries' progress towards sustainable development based on ordinal regression techniques	*Knowledge-Based Systems*

(Fahimnia *et al.*, 2015). Bibliographic coupling analysis (Cobo *et al.*, 2011) was used to identify the key journals publishing in the domain of ML and sustainability.

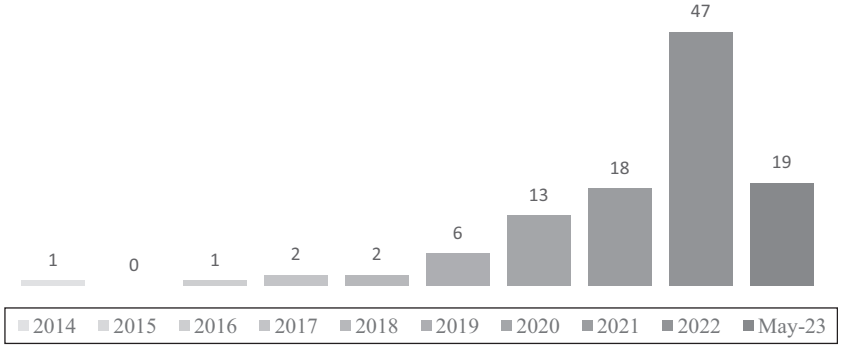
3.3.1 Co-authorship analysis. The tendency within a network of co-authors to refer to certain publications may influence literature on a research topic. The co-authorship network from our sample reveals two prominent networks of collaborations, as presented in Figure 3. Nilashi M. plays a pivotal role in these co-authorship networks.

3.3.2 Bibliographic coupling. Two publications referring to a document are coupled in this technique, as high instances of shared references indicate common intellectual capital of the coupled documents (Cobo *et al.*, 2011). Tables 3 and 4, respectively, acknowledge the influential journals and countries contributing important articles to the sample for this study. Figures 4 and 5 present bibliographic links among journals and countries, respectively.

3.3.3 Co-word analysis technique. Keywords used by the authors for articles under review provide a snapshot of the literature on a research topic (Fahimnia *et al.*, 2015). The authors of articles in the sample of this study provide 402 keywords. Figure 6 presents the network diagrams of author keywords that co-occurred more than once.

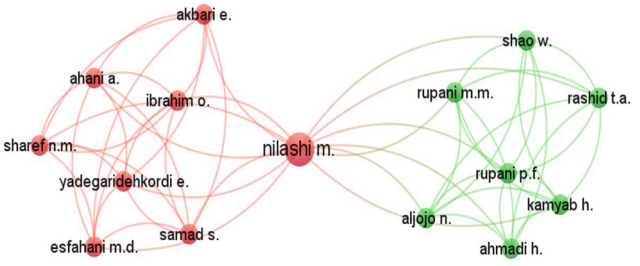
4. Thematic areas

The systematic literature review elucidated that ML can substantially contribute to various SDGs, as subsequently discussed.



Source: Created by authors

Figure 2. Year-wise distribution of publications in the study sample



Source: Created by authors

Figure 3. Network of authors from co-authorship analysis technique

4.1 Zero hunger (sustainable development goal 2)

Prior studies have explored the potential ML applications in evaluating the impact of sustainability initiatives on food production and consumption (Abdella *et al.*, 2020; Taghizadeh-Mehrjardi *et al.*, 2020). Taghizadeh-Mehrjardi *et al.* (2020) suggested using ML techniques to assess soil quality and ensure sustainable crop production. Abdella *et al.* (2020) recommended the use of K-means clustering and logistic regression to investigate the impact of food consumption in the USA.

Bedi (2022) introduced a novel approach, referred to as the transfer learning augmented enhanced memory network, to enhance the accuracy of estimating reference evapotranspiration, which is a crucial parameter in hydrological and agricultural studies as it quantifies the amount of water that evaporates from a reference crop surface. El Hathat *et al.* (2023) conducted a study on greenhouse gas emissions in tomato production in Morocco. They identified primary contributors to emissions and shared valuable insights into the environmental impact of tomato production, informing sustainable agricultural practices (El Hathat *et al.*, 2023).

Jato-Espino and Mayor-Vitoria (2023) developed a statistical and ML methodology to model the risk of rural depopulation and investigate its mitigation through agricultural land use management. The expansion of the Brazilian sugarcane industry has been associated with deforestation in the Amazon rainforest (Benites-Lazaro *et al.*, 2018). To gauge the extent to which sustainability is prioritized by sugarcane companies, Benites-Lazaro *et al.* (2018) applied the Latent Dirichlet allocation (LDA) algorithm, identifying 36 themes encompassing agro-environmental issues, biofuels and flex-fuel, among other topics.

Table 3.
Top ten journals from bibliographic coupling analysis

Journal	Total link strength
Journal of Cleaner Production	16
Cities	6
Information and Management	6
International Journal of Information Management Data Insights	6
Utilities Policy	5
Journal of Enterprise Information Management	3
Knowledge Based System	3
Big Data and Cognitive Computing	2
Journal of Business Research	2
Technology in Society	2

Table 4.
Top ten countries from bibliographic coupling analysis

Country	Total link strength
China	1,813
United Kingdom	1,718
Australia	1,164
Iran	1,163
USA	1,145
Hong Kong	887
India	838
Malaysia	797
Singapore	747
Italy	670

Source: Created by authors

Figure 4.
Connections from bibliographic coupling of journals

Implementing standardized procedures and codes of conduct guided by this algorithm could assist the industry in promoting sustainability (Benites-Lazaro *et al.*, 2018).

4.2 Good health and well-being (sustainable development goal 3)

The role of health-care intrastate is of paramount importance in facilitating sustainable growth (Zolbanin *et al.*, 2022). Zolbanin *et al.* (2022) focus on the application of data analytics to enhance resource utilization and promote sustainability within the health-care sector.

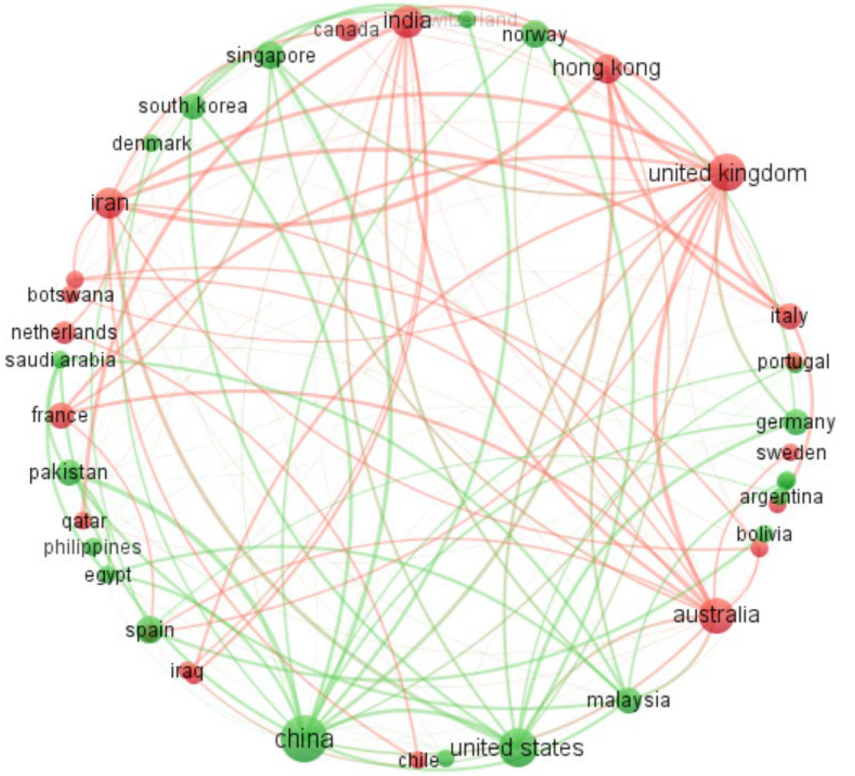


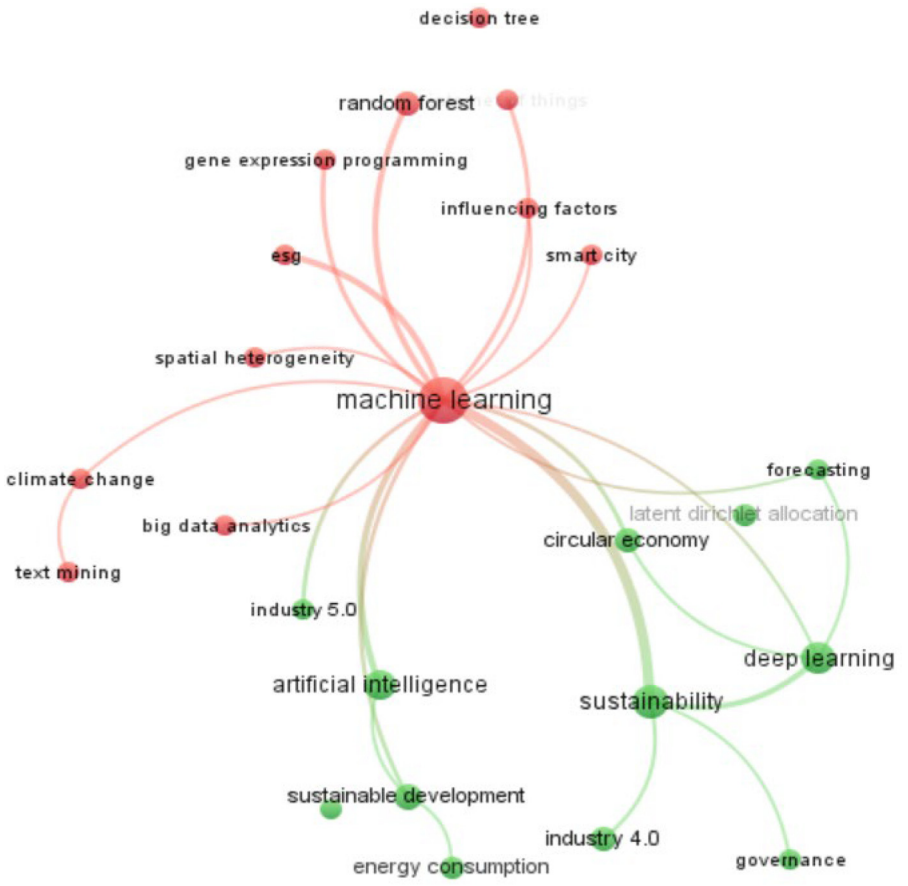
Figure 5.
Connections from bibliographic coupling of journals

Source: Created by authors

Their primary objective revolves around predicting the duration of hospital stays for patients with chronic illnesses, thereby enabling hospitals to optimize resource allocation and improve the quality of patient care.

An increasing number of daily traffic accidents annually leads to a significant loss of lives and injuries. Delayed arrival of emergency medical assistance contributes to the majority of fatalities and injuries in these accidents (Balfaqih *et al.*, 2022). Timely medical treatment plays a crucial role in improving survival chances and reducing the risk of disabilities. To address this issue, Balfaqih *et al.* (2022) developed an accident detection and classification system using the Internet of Things framework. Balfaqih *et al.* (2022) reported that the Gaussian mixture model and classification and regression trees exhibited superior precision and recall in accident classification.

Lan *et al.* (2021) propose an algorithm that combines DL and neural networks to aid government authorities in integrating diverse sustainable development initiatives. This algorithm effectively explores practices related to environmental conservation, healthy lifestyles and sustainable urban development, ultimately leading to the establishment of a robust regional economy (Lan *et al.*, 2021; Rieder *et al.*, 2022; Yin and Zhou, 2022; Warsame *et al.*, 2023). The Interval Type-2 fuzzy logic is found effective in predictive tasks, estimating sulphur concentration in fuels, which poses risks to public health and the environment (Al-Jamimi and Saleh, 2019).



Source: Created by authors

Figure 6.
Co-occurrences of author keywords

4.3 Clean water and sanitation (sustainable development goal 6)

Dong *et al.* (2023) introduce a combined water quality forecasting system that uses multi-objective optimization and an improved data decomposition integration strategy. By considering multiple objectives and integrating diverse data sources, their approach enhances the accuracy and reliability of water quality predictions. Priyadarshini *et al.* (2022a) explored water pollution reduction for sustainable urban development using ML techniques. Rahman *et al.* (2021) demonstrate the creation of a flood hazard map and an emergency relief operation system through the utilization of hydrodynamic modelling and ML algorithms. The investigation targets the resolution of flood-related difficulties and endeavours to enhance strategies for flood management.

Phosphorus plays a pivotal role in ensuring global food and water security, necessitating policymakers to formulate efficacious strategies for regulating phosphorus movement within the waste sector (Baroi *et al.*, 2020). Baroi *et al.* (2020) proposed an ML approach to overcome limited data availability in assessing long-term phosphorus quantities.

4.4 Affordable and clean energy (sustainable development goal 7)

Solar radiation measurements play a crucial role in enhancing energy efficiency within intelligent buildings and facilitating solar energy generation (Kosovic *et al.*, 2020). The high cost of purchasing and maintaining expensive sensors often limits data availability regarding solar radiation (Kosovic *et al.*, 2020). To address this limitation, Kosovic *et al.* (2020) propose a technique to estimate solar radiation using the data obtained from meteorological sensor stations. Priyadarshini *et al.* (2022b) developed an ensemble model based on ML techniques to forecast energy consumption in smart homes. Tsoka *et al.* (2022) concentrated on the utilization of artificial neural network (ANN) for classifying building energy performance certificates, which offer valuable insights into the energy efficiency of buildings and aid in decision-making processes for energy optimization and sustainability. Sohani *et al.* (2022a, 2022b) introduced an ANN-based framework that integrates ML techniques with solar panels to enhance energy generation efficiency and sustainability. Yin and Zhou (2022) evaluate China's photovoltaic poverty alleviation project.

Geothermal energy also serves as a sustainable power source for domestic and industrial applications, but its effectiveness relies on various complex environmental factors, posing challenges for conventional assessment methods (Coro and Trumpy, 2020). A maximum entropy-based algorithm using environmental data is found suitable to assess geothermal site suitability and optimize plant operations, aiding companies and regulators in conducting feasibility studies for potential geothermal power plant locations (Coro and Trumpy, 2020).

Redchuk *et al.* (2023) conducted a case study on the adoption of the industrial Internet of Things and ML for improving energy consumption in a process manufacturing firm operating under the Industry 5.0 paradigm. The Industry Revolution 4.0 and 5.0 emphasize intelligent production and real-time monitoring to optimize manufacturing performance by identifying the most effective pathway for process improvement (Jamwal *et al.*, 2022; Redchuk *et al.*, 2023). Gao *et al.* (2023) proposed a digital twin-based approach to optimize operational energy consumption in automated container terminals. Additionally, the implementation of lean and green approaches in a combined heat and power plant resulted in a significant performance improvement of 18.25% compared to previous levels (Leong *et al.*, 2020).

Prior studies investigated water demand for cooling technology in thermal power generation using ML (Lohrmann *et al.*, 2022; Zafar *et al.*, 2022). The researchers construct global and regional models using support vector machine (SVM)s, K-nearest neighbours (KNN) and random forest algorithms to gauge water demand, taking into account diverse factors like cooling technology, climate conditions and power plant attributes (Lohrmann *et al.*, 2022; Zafar *et al.*, 2022). Bakay and Ağbulut (2021) focus on the prediction of greenhouse gas emissions in Turkey using ML techniques. Their study compares the performance of DL, SVM and ANN algorithms based on electricity production data.

Bas *et al.* (2021) use a random forest algorithm to examine the classification of prospective buyers of electric vehicles. This research aids in the advancement of targeted marketing strategies and policy interventions aimed at promoting the adoption of electric vehicles. Aaldering and Song (2019) used three classifier algorithms, namely, SVM, decision tree classifier and Gaussian Naive Bayes, to investigate the trajectory of technological development in post-lithium-ion battery from patent data analysis.

The improvement of the carbon trading mechanism is crucial in promoting the energy market (Wang and Zhao, 2021). Wang and Zhao (2021) have presented a Bayesian Network from their research to effectively forecast the performance of the carbon futures market.

4.5 Decent work and economic growth (sustainable development goal 8)

Urbanization serves as a catalyst for economic growth; however, its intricate nature profoundly impacts both humanity and the environment (Kim *et al.*, 2021). Consequently, countries worldwide are striving to develop smart cities that foster sustainable living through the utilization of advanced technologies (De Las Heras *et al.*, 2020). Mele and Magazzino (2020) proposed a ML model incorporating long short-term memory aimed at harmonizing economic growth and environmental sustainability within the Chinese steel industry.

Saura *et al.* (2022) use sentiment analysis techniques to examine the sentiments of Twitter users concerning remote work, offering valuable insights into the intricate dynamics of remote work arrangements. Their findings inform organizational and policymaking efforts to tackle remote work challenges, enhance work-life equilibrium and advance well-being in the aftermath of the pandemic.

4.6 Industry, innovation and infrastructure (sustainable development goal 9)

The construction industry's reliance on concrete has led to significant environmental threats, including energy consumption and depletion of natural resources. Several studies have explored the integration of ML techniques to optimize concrete mixtures, enhance mechanical properties and reduce environmental impacts (Iftikhar *et al.*, 2022; Hull and Rothenberg, 2008). For example, Ashraf *et al.* (2022) investigated the incorporation of bentonite clay and silica fume in the development of a sustainable concrete mixture. Huo *et al.* (2022) used ML algorithms to determine the compressive strength of calcium-based geopolymers. Golafshani *et al.* (2023) emphasized the importance of sustainable construction practices using recycled aggregate concrete and used ensemble learning techniques to model chloride migration. Munir *et al.* (2022) proposed a novel ML-based design strength model for sustainable concrete columns, surpassing conventional design methods in terms of accuracy and efficiency. To achieve eco-friendly concrete production, it is crucial to balance economic factors and environmental impacts. Naseri *et al.* (2020) introduced six ML algorithms, including ANN, SVM, genetic algorithm, regression-based algorithm, soccer league competition algorithm and water cycle algorithm (WCA), to predict the compressive strength of eco-friendly concrete. Among these models, the WCA exhibited superior accuracy in terms of mean absolute error and coefficient of determination. Sadrossadat *et al.* (2023) proposed a multi-objective mixture design in optimizing cemented paste backfill.

Adedoyin *et al.* (2022) conducted a simulation experiment in South Africa to investigate the correlation between information and communication technology and patent intensity. Using a dynamic autoregressive distributed lag model (ARDL), the study reveals that the adoption and effective use of information and communication technology and patent foster enhanced innovation and increased patent activity (Adedoyin *et al.*, 2022). On the flip side, technologies may come with environmental consequences, such as e-waste generation from cryptocurrencies network (Jana *et al.*, 2021). Employing random forest and support vector regression, Jana *et al.* (2021) identify the significant determinants of e-waste production.

The rise of digital music streaming services is attributed to the decline in revenue of the recorded music industry in the USA (Dangelico, 2016; Naveed *et al.*, 2017). To address the threat to artists and other stakeholders reliant on the music industry, Naveed *et al.* (2017) propose the implementation of an ML algorithm that promotes the co-evolution of live streaming and the recorded music industry. By transforming the live music industry into a "live-concert-streaming music industry", stakeholders can engage in participatory creativity (Naveed *et al.*, 2017).

4.7 Reduced inequalities (sustainable development goal 10)

It is crucial to ensure economic and social equality to establish a fair and just society (Hidalgo *et al.*, 2020). Economic inequality can lead to adverse outcomes such as poverty, unemployment and limited access to fundamental necessities including health care and education (Hidalgo *et al.*, 2020). Various studies have examined the effects of reducing inequality on global poverty (Lakner *et al.*, 2022), the prediction of sulphur dioxide emissions and regulation of air quality (Ribeiro, 2021), the prediction of profit margins in sustainable road freight transportation (Budak and Sarvari, 2021), and computational modelling techniques for wet waste valorization (Li *et al.*, 2022a, 2022b, 2022c, 2022d). Achieving a balance between social, economic and environmental sustainability necessitates the development of inclusive, safe and resilient cities where the actions of one community do not adversely affect others (Li *et al.*, 2019). Li *et al.* (2019) propose the utilization of ML approaches, specifically agent-based modelling (ABM) and embedded geo-simulation, to foster sustainable cities and communities.

4.8 Sustainable cities and communities (sustainable development goal 11)

Rapid urbanization presents a complex and multifaceted challenge, requiring sustainable development strategies for smart cities (Li *et al.*, 2019). To achieve this, a crucial aspect is understanding population specialization (Zhao *et al.*, 2020). Zhao *et al.* (2020) proposed that convolutional neural network, deep neural network and random forest models may analyse population specialization. By comparing the outputs of these algorithms with actual Chinese population data sets, the study found the convolutional neural network algorithm to be the most accurate (Zhao *et al.*, 2020).

Wang and Biljecki (2022) reported the global spatiotemporal estimation of daily high-resolution surface carbon monoxide concentrations using the deep forest model (Wang *et al.*, 2022a). Such quantitative approaches to support spatial planning, industry selection and ecological regulation require an accurate assessment of ecosystem service value of combined natural, land-use and socioeconomic factors (Liu *et al.*, 2021b). Liu *et al.* (2021a, 2021b) proposed a conditional inference tree to explore the collective impact of seven factors and various ecological combinations to forecast ecosystem service values (Liu *et al.*, 2021a, 2021b).

ML techniques can be leveraged by authorities to address urban challenges and enhance the quality of life for city residents by improving government performance (Kim *et al.*, 2021). Kim *et al.* (2021) developed a dynamic topic model to identify relationships between keywords in civic complaints and provide relevant suggestions. The effectiveness of the model was successfully demonstrated using a large data set from Seoul, South Korea (Kim *et al.*, 2021). Lai *et al.* (2023) focus on the inference of household size distribution and its correlation with the built environment, leveraging extensive mobile phone data for their analysis.

Mansoursamaei *et al.* (2023) highlight the growing significance of sustainability in the maritime sector and underscores the necessity for inventive approaches to address environmental impacts of port activities. The study investigates the utilization of ANNs and SVMs to scrutinize extensive data sets and extract valuable insights regarding energy consumption, emissions, waste management and other ecological factors within ports. Mansoursamaei *et al.* (2023) recommend predictive models and decision support tools to optimize port operations, minimize environmental hazards and bolster sustainability.

4.9 Responsible consumption and production (sustainable development goal 12)

The adoption of ML techniques in various domains has shown promise in addressing sustainability challenges and promoting responsible consumption and production

(Gunawan *et al.*, 2020; D'Amato *et al.*, 2017; Lu *et al.*, 2022; McMeekin and Southerton, 2012). Teng *et al.* (2023) use neural architecture search (NAS) to develop a system that dynamically optimizes energy consumption and minimizes environmental impact in chemical processes through phase switching. Nilashi *et al.* (2019a) propose a model for recommending green hotels based on consumer feedback, using ML techniques such as self-organizing maps and adaptive neuro-fuzzy inference systems.

Jafari *et al.* (2022) examine the implementation of extended producer responsibility, highlighting the significance of mutual understanding between producers and consumers in achieving sustainability goals. Arranz *et al.* (2022) and Prioux *et al.* (2023) investigate institutional pressures driving circular economy in firms. Behera *et al.* (2023) explore the adoption of explainable artificial intelligence (XAI) in consumer-packaged goods retail, emphasizing the importance of transparency and accountability in decision-making for sustainable growth.

Sharma *et al.* (2020) emphasize the need for sustainable supply chains and expand decision-making criteria beyond economic factors to include social and environmental aspects. Wu *et al.* (2020) developed a multi-partner classification model based on fuzzy set theory and ensemble learning to categorize potential supply partners and recommend appropriate strategies for each category. In the study by Miguéis *et al.* (2022), the authors tackle the task of minimizing waste in the fresh fish industry while ensuring an adequate supply for consumers. Their focus lies in developing an accurate demand forecasting model using ML models and censored data.

Insufficient estimation of waste generated within municipalities often leads to a lack of appropriate solid waste management infrastructure (Ayeleru *et al.*, 2021). Ayeleru *et al.* (2021) propose a solution by creating a model that uses ANNs and SVMs to forecast the amount of solid waste generated. By effectively aiding municipal authorities in establishing the necessary infrastructure for solid waste management, this model promotes sustainable development in Johannesburg, South Africa. Further, Gue *et al.* (2022) predict waste management system performance based on city and country attributes. Li *et al.* (2022a, 2022b, 2022c, 2022d) focused on sustainable decision-making for contaminated site risk management and reviewed computational modelling techniques for wet waste valorization.

4.10 Climate action (sustainable development goal 13)

Kumar *et al.* (2021) propose a novel framework aimed at evaluating and fortifying the resilience of critical infrastructure in the context of climate change. The authors emphasize the necessity of addressing the vulnerabilities of infrastructure systems, such as transportation, energy and water, to effectively cope with the repercussions of climate change, including extreme weather events and rising sea levels. To support their claims, Jin *et al.* (2023) present empirical evidence demonstrating the alignment of urban climate adaptation with sustainable development, using LDA as their methodology. Similarly, Chiu *et al.* (2023) concentrate on the application of unsupervised learning techniques to establish a regional risk model based on the Task Force on Climate-related Financial Disclosures framework. By using these unsupervised learning methods (Khodabandehlou and Rahman, 2017), the authors analyse data and identify patterns and relationships pertaining to climate risks at the regional level. This approach fosters a more comprehensive understanding of the challenges associated with climate change, enabling informed decision-making and effective risk mitigation strategies to promote sustainable development. In a related study, Zhu *et al.* (2022) devised an anticrime information support system in the city of Chicago using the K-means algorithm. In the domain of climate adaptation planning in coastal tourism destinations, Li *et al.* (2022a, 2022b, 2022c, 2022d) adopt a multi-objective

optimization approach. Their methodology aims to optimize climate adaptation strategies while considering various objectives.

4.11 Life on land (sustainable development goal 15)

The effectiveness of sustainable animal farming can be enhanced through the implementation of solutions addressing animal health and welfare (Bao and Xie, 2022). In this regard, SVMs, random forest and clustering techniques have shown promise in addressing animal health and welfare-related issues (Bao and Xie, 2022). By using ML applications, which rely on animal activities as vital indicators of health, the promotion of sustainable and environmentally friendly animal farm production becomes feasible (Bao and Xie, 2022; Darnall *et al.*, 2010).

Shafi *et al.* (2023) underscore the significance of using ML and remote sensing techniques to monitor the spatiotemporal changes occurring in wetlands in Pakistan over an extended period. Their primary objective revolves around the creation of a national inventory of wetlands, aiming to preserve and protect these critical ecosystems.

4.12 Peace, justice and strong institutions (sustainable development goal 16)

An essential element of strong institutions is the promotion of ethical practices within organizations (Wang *et al.*, 2020). To this end, Wang *et al.* (2020) have introduced the application of variable importance analysis using the random forest model as a means to forecast corporate misconduct. By analyzing 953 observations from 93 Chinese construction companies between 2000 and 2018, the researchers identified 11 significant variables pertaining to corporate governance. This application of ML demonstrates its potential in assisting corporations to mitigate the risks associated with illegal and unethical activities (Wang *et al.*, 2020). In the domain of financial risk management and business sustainability, Jabeur *et al.* (2021) explore the use of ML, specifically the CatBoost model, to predict corporate failure. Jabeur *et al.* (2021) emphasize the importance of early identification and prediction of corporate failure to mitigate financial risks and ensure the continuity of business operations.

Recognizing the increasing influence of ML in policy and decision-making processes, Wilson and Van Der Velden (2022) propose an innovative model that seeks to enhance sustainable decision-making within the public sector. The authors emphasize the need for sustainability considerations to be integrated into ML systems, as they play a pivotal role in shaping policy outcomes. The model incorporates multiple dimensions of sustainability, encompassing economic, social and environmental aspects, to guide decision-making processes effectively (Wilson and Van Der Velden, 2022).

Assessing and managing the performance of sustainability initiatives in a country is a critical requirement in ensuring their effectiveness (Nilashi *et al.*, 2019b). To address this need, Nilashi *et al.* (2019b) used fuzzy clustering and supervised ML techniques to measure sustainability initiatives in 128 countries. However, indicators for measuring and monitoring sustainable development at the macro level often demonstrate limited predictive power (Pérez-Ortiz *et al.*, 2014). Thus, Pérez-Ortiz *et al.* (2014) proposed an ML technique that combines cluster analysis and ordinal classification to rank countries based on their sustainability initiatives so that policymakers can effectively monitor the impact of policies on sustainable development.

4.13 Partnerships for the goals

Green crowdfunding projects play a crucial role in financing SDGs. However, the identification of these projects depends on either subjective evaluation or existing

classifications provided by crowdfunding platforms, both of which have limitations (Butticè *et al.*, 2019). To address these issues, Butticè *et al.* (2019) proposed the use of ML algorithms to develop a content-specific classifier for green crowdfunding campaigns. ML algorithms, such as random forest, automatically extract informative keywords to differentiate between green and non-green campaigns (Butticè *et al.*, 2019).

Cacciarelli and Boresta (2022) use ML to predict donor responses in non-profit direct marketing campaigns, identifying key drivers behind donation decisions. Through the application of decision trees and random forest algorithms, Cacciarelli and Boresta (2022) develop predictive models that can accurately estimate the likelihood of donor response.

5. Discussion

This study reviewed a diverse range of ML models from prior literature that contribute to accurate predictions, insights and planning for sustainable development. Such models include applications of supervised learning algorithms, unsupervised learning algorithms, DL techniques and scenario planning approaches, among others.

Supervised learning involves constructing models using labelled data, wherein each data point is associated with a specific target or label to establish a relationship or mapping between the input features and the desired output labels (Abdella *et al.*, 2020). The studies under review extensively used supervised learning approaches to tackle a range of predictive tasks using various algorithms, including autoregressive distributed lag (Adedoyin *et al.*, 2022), maximum entropy (Coro and Trumpy, 2020), Bayesian network (Wang and Zhao, 2021) and Gaussian Naive Bayes model (Aaldering and Song, 2019). The ARDL is a regression model that analyses the relationship between a dependent variable and its lagged values as well as the lagged values of other independent variables (Adedoyin *et al.*, 2022). The maximum entropy-based algorithm is a multinomial logistic regression that models the probability distribution of a categorical dependent variable using the principle of maximum entropy (Coro and Trumpy, 2020). Bayesian Network algorithm uses graphical models that represent probabilistic relationships among variables to learn patterns from the data (Wang and Zhao, 2021). Gaussian Naive Bayes model, which is commonly used for text classification tasks, involves a probabilistic classifier based on Bayes' theorem assuming mutual exclusivity of the features (Aaldering and Song, 2019).

The studies under review also deployed supervised algorithms like KNN model (Lohrmann *et al.*, 2022), SVM (Naseri *et al.*, 2020), decision trees (Balfaqih *et al.*, 2022), random forest (Bao and Xie, 2022; Jana *et al.*, 2021; Zhao *et al.*, 2020), gradient boosting (Golafshani *et al.*, 2023) and the CatBoost model (Jabeur *et al.*, 2021). KNN is a non-parametric algorithm that classifies new data points based on the majority class of their nearby neighbours in the training set (Lohrmann *et al.*, 2022). SVM algorithm that uses a hyperplane to separate different classes in the feature space to find the optimal hyperplane with the largest margin between classes (Bakay and Ağbulut, 2021). Decision trees partitions the feature space based on a series of binary decisions for both regression and classification tasks (Cacciarelli and Boresta, 2022). Random Forest is an ensemble method that combines multiple decision trees to make predictions introducing randomness in the tree-building process to reduce overfitting and improve generalization (Bas *et al.*, 2021; Butticè *et al.*, 2019; Wang *et al.*, 2020). Gradient boosting is another ensemble method that combines multiple weak predictive models – often decision trees – to create a strong predictive model (Golafshani *et al.*, 2023). CatBoost, an abbreviation for categorical boosting model, is a gradient boosting algorithm that is designed to handle categorical features effectively (Jabeur *et al.*, 2021).

Prior research on ML for sustainability uses several unsupervised ML algorithms. Unsupervised learning involves training a model on unlabelled data, particularly when there is no prior knowledge about the data and the underlying structure of the data is undiscovered (Chiu *et al.*, 2023). Prominent unsupervised learning algorithms include K-means clustering (Zhu *et al.*, 2022), hierarchical clustering (Bao and Xie, 2022) and Gaussian Mixture Models (Balfaqih *et al.*, 2022). Clustering techniques group similar data points together based on their inherent attributes (Abdella *et al.*, 2020; Chiu *et al.*, 2023). Further, dynamic topic modelling is an ML approach used to analyse the evolution of topics from a collection of documents over time (Kim *et al.*, 2021). Dynamic topic modelling is an extension of traditional topic modelling algorithms, such as LDA, that account for the temporal aspect of textual data (Benites-Lazaro *et al.*, 2018).

Several studies discussed the deployment of DL that focuses on training ANNs with multiple layers to learn and extract hierarchical representations from complex data (Bakay and Ağbulut, 2021; Zhao *et al.*, 2020). ANN is a method based on the functioning of the human brain to identify patterns and associations from large data for prediction and classification tasks (Ayeleru *et al.*, 2021; Sohani *et al.*, 2022a, 2022b). LSTM helps computers understand and remember things over time like human brains do (Mele and Magazzino, 2020). NAS aims to systematically search and evaluate various architectures to identify the most effective network design that can achieve superior performance on a specific task (Teng *et al.*, 2023). DL models excel at tasks such as image and speech recognition, natural language processing and generative modelling (Lan *et al.*, 2021).

Optimization of the performance is crucial in training ML models and finding optimal solutions (Dong *et al.*, 2023; Li *et al.*, 2022a, 2022b, 2022c, 2022d). Naseri *et al.* (2020) explored the best set of parameters for their objective function with a genetic algorithm (GA) and WCA. GA is a search and optimization technique that involves creating a population of candidate solutions, evaluating their fitness, applying genetic operators such as mutation and crossover to generate new solutions and iteratively evolving the population to find the best solution to a given problem (Naseri *et al.*, 2020). The WCA is a population-based optimization algorithm that simulates the natural water cycle stages such as precipitation, evaporation and infiltration, to optimize a given problem (Naseri *et al.*, 2020). Transfer learning can help improve the performance of models, particularly when the new data set is small or when data is limited (Bedi, 2022). This ML technique applies knowledge gained from solving one problem to a different but related problem, wherein a pre-trained model is used as a starting point and then fine-tuned on the new task (Bedi, 2022).

The studies under review further presented scenario planning with ABM (Li *et al.*, 2019), soccer league competition algorithm (Naseri *et al.*, 2020) and planning land use scenarios (PLUS) (Li *et al.*, 2022a, 2022b, 2022c, 2022d). PLUS involves simulation of different land use patterns and their potential impacts on urban development to help urban planners and policymakers make informed decisions regarding land use, zoning regulations and sustainable development (Li *et al.*, 2022a, 2022b, 2022c, 2022d). An ABM simulates the behaviour and interactions of autonomous agents to study complex systems, particularly in the context of social sciences (Li *et al.*, 2019). A soccer league competition algorithm is traditionally used to simulate and optimize the scheduling of a tournament and predict outcomes from historical data (Naseri *et al.*, 2020).

The need for transparency and interpretability in ML models is increasingly recognized to avoid treating artificial intelligence as black boxes (Behera *et al.*, 2023). Research in the field of XAI is growing to develop methods and techniques that enhance the transparency, interpretability and inclusiveness of ML models (Tsoka *et al.*, 2022). Some approaches involve the use of interpretable models, such as decision trees or linear regression, which

inherently possess transparent decision-making processes (Behera *et al.*, 2023; Tsoka *et al.*, 2022). Other approaches focus on post hoc explanations, which generate explanations by analyzing the model's internal representations or the significance of input features after making predictions (Behera *et al.*, 2023; Tsoka *et al.*, 2022). The brief description of the ML techniques is mentioned in Appendix.

6. Future research scopes

A critical review of future research scopes in the context of ML applications for sustainability is summarized in Table 5 and discussed as follows:

6.1 Model extension

Future research in the field of ML applications in sustainability can build upon previous studies to advance existing models. For instance, Mele and Magazzino (2020) developed a model that traces the path of emissions, and future researchers can enhance it by incorporating appropriate vectors and considering different emission patterns. Regarding waste management performance prediction, Gue *et al.* (2022) propose the application of ML techniques, such as regression models, time series analysis and ensemble methods. These approaches can assist policymakers and stakeholders in developing efficient waste management strategies, resource allocation and infrastructure planning. In multiple-criteria decision-making related to production cost and environmental concerns, Garg and Lam (2016) suggest including additional factors, and future studies can explore this approach further. In the context of supply chain partner selection, Wu *et al.* (2020) propose using the interval linguistic model, trapezoidal fuzzy method and dynamic evaluation techniques. Future researchers can leverage these methods in a rapidly changing environment to make more suitable partner selections.

Li *et al.* (2019) present an ABM, and its predictive power can be improved by incorporating agents' learning processes, experiences exchange and a combination of different learning algorithms. For responsible hospitality and tourism services, Nilashi *et al.* (2019a) suggest considering energy conservation, eco-friendly temperature control and water-saving washrooms in recommendation algorithms for hotel selection. In solid waste management, Ayeleru *et al.* (2021) propose a model that future studies can extend by incorporating policy implications and disposal parameters. Jin *et al.* (2021) designed a DL-based policy guarantee mechanism with multiple applications. Future research is necessary to explore the use of suitable algorithms to expand the functionalities of public services (Khanra, Joseph, and Ruparel, 2019). Zhu *et al.* (2022) present an anticrime information support system using K-means in Chicago. Future research can expand the system's capabilities by incorporating additional data sources like social media, surveillance footage and emergency service data. This can improve crime prediction, identify hotspots, and enhance urban safety and security (Khanna and Khanra, 2023). Similarly, Kim *et al.* (2021) propose enhancing their model for civic queries in smart cities by addressing complex semantic relationships.

6.2 Model establishment

Future research offers several avenues to advance the literature on ML applications in sustainability. Leong *et al.* (2020) suggest implementing enhanced adaptive lean and green models in multiple industries transitioning to Industry 4.0 (Grybauskas *et al.*, 2022; Lee and Lim, 2021). Validating Naseri *et al.* (2020) algorithm across multiple case studies can aid in developing eco-friendly commercial machines. Arnold's (2018) model, aimed at overcoming human biases, may benefit from additional training data and augmented techniques.

Theme	Research questions	References
Model extension	• What are the additional factors (production cost and environmental concerns) in multiple-criteria decision-making impacting manufacturing practices?	Garg and Lam (2016)
	• How can other methodologies be used to select a supply partner selection?	Wu *et al.* (2020)
	• How may considering factors such as energy conservation, eco-friendly temperature control and water-saving washrooms in the recommendation algorithm help responsible consumers of hospitality and tourism services choose appropriate hotels?	Nilashi *et al.* (2019a)
	• How can the solid waste management model be extended by including policy implications and disposal parameters?	Ayeleru *et al.* (2021)
	• How to integrate additional ML application model-regression models, time series analysis and ensemble methods, to improve waste management prediction?	Gue *et al.* (2022)
	• How to enhance urban safety by incorporating additional data sources, such as social media, CCTV footage and emergency service data, to improve crime prediction, hotspots identification and resource allocation.	Zhu *et al.* (2022)
Model establishment	• How can human biases be overcome with additional training data and augmented techniques?	Arnold (2018)
	• How to commercialize the eco-friendly concrete machine?	Naseri *et al.* (2020)
	• How can machine learning be applied in designing a sustainable port with a special focus on carbon emission?	Fahdi *et al.* (2021)
	• How to validate the existing land suitability assessment framework developed by testing the framework in cases of multiple crops	Taghizadeh-Mehrjardi *et al.* (2020)
	• How to refine and optimizing the multi-objective optimization and data decomposition integration strategy to improve the accuracy and reliability of water quality predictions?	Dong *et al.* (2023)
	• How to integrate ML methodology with other technologies such as – sensor technologies, real-time monitoring and advanced data analytics to develop comprehensive and adaptive systems for water pollution management?	Priyadarshini *et al.* (2022a, 2022b)
Scope extension	• How to advance the existing machine learning models to detect financial misconduct in other industries?	Wang *et al.* (2020)

Table 5.
Future research questions

(*continued*)

Theme	Research questions	References
	• How to assess the impact of the sustainable sugarcane industry on related sectors such as energy, food and land?	Benites-Lazaro *et al.* (2018)
	• How to see the recent impact of the model offered to predict population spatialization data collected from broader demographic groups?	Zhao *et al.* (2020)
	• How can machine learning algorithms be used for classification in emerging technologies, such as Internet of Things, nanotechnologies and self-driving vehicles, which may contribute to sustainability but are not comprehensively covered by standard classifications yet?	Butticè *et al.* (2019)
	• How to enhance the predictive model by additionally, incorporating ecological indicators to evaluate the ecological services provided by wetlands?	Shafi *et al.* (2023)
	• How to scale and generalize the models by applying to different geographical regions and considering a wider range of climate risk factors?	Chiu *et al.* (2023)
Potential application of ChatGPT promoting sustainability	• How to use ChatGPT to spread the education in emerging and developing countries to improve the standard of living of people?	Firat (2023)
	• How can ChatGPT be used to forecast the environmental impact in different countries?	Jungwirth and Haluza (2023)

Table 5.

Exploring the internal mechanisms that cause variations in the performance of DL and neural network models can provide a profound understanding of sustainable urban development (Lan *et al.*, 2021). Combining clustering and supervised learning techniques could lead to more effective sustainability assessment compared to the approach proposed by Nilashi *et al.* (2019b).

Investigating the impact of ML applications on carbon emission reduction in ports is another area for future study (Fahdi *et al.*, 2021). The land suitability assessment framework developed by Taghizadeh-Mehrjardi *et al.* (2020) can be validated through testing in cases involving multiple crops. Dong *et al.* (2023) propose a combined approach for water quality forecasting, which can be refined and optimized by improving the multi-objective optimization and data decomposition integration strategy. Integrating real-time data sources, advanced ML techniques and sensor technologies can enhance water pollution management systems in urban areas, enabling early warning systems, decision support tools and targeted interventions (Priyadarshini *et al.*, 2022a, 2022b).

6.3 Scope extension

Future research has ample opportunities for extending the scope of ML models developed in previous studies to advance sustainable applications. For instance, Wang *et al.* (2020)

presented a random forest algorithm that can be tested across multiple countries and industries to predict financial misconduct, thereby promoting ethical practices in the corporate sector. Benites-Lazaro *et al.* (2018) emphasized the need to investigate the impact of the sugarcane industry on related sectors like energy, food and land for ethanol production. Abdella *et al.* (2020) conducted a study on the sustainability impacts of food consumption, and future researchers can expand upon it by comparing country-specific impacts. Considering potential barriers and implementation plans, Liu *et al.* (2021a, 2021b) stressed the importance of region-specific ML applications due to issues with data standardization. Revisiting the findings of Pérez-Ortiz *et al.* (2014) with contemporary data can capture the recent impact. The model proposed by Zhao *et al.* (2020) for population spatialization can be tested with recent data collected from broader demographic groups.

Chiu *et al.* (2023) demonstrated the application of unsupervised learning methods in developing a regional risk model for climate-related risks. Further research can explore the scalability and generalizability of such models by applying them to different geographical regions and considering a wider range of climate risk factors. Additionally, investigating the integration of supervised and unsupervised learning approaches can provide a more comprehensive understanding of climate-related risks and support effective risk management strategies. Shafi *et al.* (2023) showcased the use of ML and remote sensing techniques for monitoring wetland changes. Future research can expand this approach to other regions and countries, develop standardized methodologies for wetland monitoring and assessment, and incorporate ecological indicators to enhance the understanding of wetland dynamics and support conservation efforts. Naveed *et al.* (2017) proposed a blueprint for the future of the music industry, which can also be implemented in different cultural contexts and other industries, such as print and digital media, facing similar challenges.

6.4 Potential application of large language models

Language models such as generative pre-trained transformers (GPT) are fast spreading the use of DL techniques to generate coherent and contextually relevant text. These models may complement human intelligence with support for evaluation, interpretation, communication and decision-making. For instance, ChatGPT has been used for automating material supply in construction projects, reducing material overruns and human effort (Ashiq *et al.*, 2023). Language models can also enhance education by providing customized interactive assistance, promoting independence and autonomy (Firat, 2023). A newer age GPT has shown promise in various fields such as education, health and communication, enabling human-like interactions (Jungwirth and Haluza, 2023). ChatGPT may be used in climate change research, enabling forecasting and generating scenarios of climate change and its environmental impacts (Biswas, 2023; Steward, 2012). Further exploration on possible contributions of large language models towards sustainability is warranted.

7. Conclusion

This study presents a systematic literature review that examines the role of ML in promoting sustainability. *RQ1* regarding the current state of the literature is addressed with bibliometric analysis. *RQ2* involves a critical analysis of how ML contributes to different SDGs. The 110 studies under review address a total of 13 out of 17 SDGs, excluding SDG 1, SDG 4, SDG 5 and SDG 14. *RQ3* solicits exploration of potential avenues for future research on ML for sustainability. Future research scopes highlight model extension, model establishment, scope extension and the application of large language models. Future research may also focus on using ML to address SDGs related to poverty alleviation, quality

education, gender equality and life below water. The findings of this study have implications for business managers in terms of operational planning and resource allocation, as well as for government authorities involved in designing and implementing sustainable development policies.

References

Aaldering, L.J. and Song, C.H. (2019), "Tracing the technological development trajectory in post-lithium-ion battery technologies: a patent-based approach", *Journal of Cleaner Production*, Vol. 241, p. 118343.

Abbate, S., Centobelli, P. and Cerchione, R. (2023), "The digital and sustainable transition of the agri-food sector", *Technological Forecasting and Social Change*, Vol. 187, p. 122222.

Abdella, G.M., Kucukvar, M., Onat, N.C., Al-Yafay, H.M. and Bulak, M.E. (2020), "Sustainability assessment and modeling based on supervised machine learning techniques: the case for food consumption", *Journal of Cleaner Production*, Vol. 251, p. 119661.

Adedoyin, F.F., Mavengere, N. and Mutanga, A. (2022), "A simulation experiment on ICT and patent intensity in South Africa: an application of the novel dynamic ARDL machine learning model", *Technological Forecasting and Social Change*, Vol. 185, p. 122044.

Agrawal, R., Wankhede, V.A., Kumar, A., Luthra, S., Majumdar, A. and Kazancoglu, Y. (2021), "An exploratory state-of-the-art review of artificial intelligence applications in circular economy using structural topic modelling", *Operations Management Research*, Vol. 15 Nos 3/4, pp. 1-18.

Al-Jamimi, H.A. and Saleh, T.A. (2019), "Transparent predictive modelling of catalytic hydrodesulfurization using an interval type-2 fuzzy logic", *Journal of Cleaner Production*, Vol. 231, pp. 1079-1088.

Ambec, S. and Lanoie, P. (2008), "Does it pay to be green? A systematic overview", *The Academy of Management Perspectives*, Vol. 22 No. 4, pp. 45-62.

Arnold, M.G. (2018), "Combining conscious and unconscious knowledge within human-machine-interfaces to foster sustainability with decision-making concerning production processes", *Journal of Cleaner Production*, Vol. 179, pp. 581-592.

Arranz, C.F., Sena, V. and Kwong, C. (2022), "Institutional pressures as drivers of circular economy in firms: a machine learning approach", *Journal of Cleaner Production*, Vol. 355, p. 131738.

Arunmozhi, M., Venkatesh, V.G., Arisian, S., Shi, Y. and Sreedharan, V.R. (2022), "Application of blockchain and smart contracts in autonomous vehicle supply chains: an experimental design", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 165, p. 102864.

Ashiq, A., Hassan, U. and Zubair, M.O. (2023), "Utilizing chat GPT for automation of material supply in construction projects using programming and primavera P6 scheduling. Sustainable structures and materials", *An International Journal*, Vol. 6 No. 1, pp. 49-53.

Ashraf, M., Iqbal, M.F., Rauf, M., Ashraf, M.U., Ulhaq, A., Muhammad, H. and Liu, Q.F. (2022), "Developing a sustainable concrete incorporating bentonite clay and silica fume: mechanical and durability performance", *Journal of Cleaner Production*, Vol. 337, p. 130315.

Ayeleru, O.O., Fajimi, L.I., Oboirien, B.O. and Olubambi, P.A. (2021), "Forecasting municipal solid waste quantity using artificial neural network and supported vector machine techniques: a case study of Johannesburg", South Africa", *Journal of Cleaner Production*, Vol. 289, p. 125671.

Azadi, M., Yousefi, S., Saen, R.F., Shabanpour, H. and Jabeen, F. (2023), "Forecasting sustainability of healthcare supply chains using deep learning and network data envelopment analysis", *Journal of Business Research*, Vol. 154, p. 113357.

Bakay, M.S. and Ağbulut, Ü. (2021), "Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms", *Journal of Cleaner Production*, Vol. 285, p. 125324.

Balfaqih, M., Alharbi, S.A., Alzain, M., Alqurashi, F. and Almilad, S. (2022), "An accident detection and classification system using internet of things and machine learning towards smart city", *Sustainability*, Vol. 14 No. 1, p. 210.

Bao, J. and Xie, Q. (2022), "Artificial intelligence in animal farming: a systematic literature review", *Journal of Cleaner Production*, Vol. 331, p. 129956.

Baroi, A.R., Chowdhury, R.B., Roy, B.B. and Sujauddin, M. (2020), "Sustainability assessment of phosphorus in the waste management system of Bangladesh using substance flow analysis", *Journal of Cleaner Production*, Vol. 273, p. 122865.

Bas, J., Cirillo, C. and Cherchi, E. (2021), "Classification of potential electric vehicle purchasers: a machine learning approach", *Technological Forecasting and Social Change*, Vol. 168, p. 120759.

Bedi, J. (2022), "Transfer learning augmented enhanced memory network models for reference evapotranspiration estimation", *Knowledge-Based Systems*, Vol. 237, p. 107717.

Behera, R.K., Bala, P.K. and Rana, N.P. (2023), "Creation of sustainable growth with explainable artificial intelligence: an empirical insight from consumer packaged goods retailers", *Journal of Cleaner Production*, Vol. 399, p. 136605.

Benites-Lazaro, L.L., Giatti, L. and Giarolla, A. (2018), "Sustainability and governance of sugarcane ethanol companies in Brazil: topic modeling analysis of CSR reporting", *Journal of Cleaner Production*, Vol. 197, pp. 583-591.

Bhattacharya, A., Mohapatra, P., Kumar, V., Dey, P.K., Brady, M., Tiwari, M.K. and Nudurupati, S.S. (2014), "Green supply chain performance measurement using fuzzy ANP-based balanced scorecard: a collaborative decision-making approach", *Production Planning and Control*, Vol. 25 No. 8, pp. 698-714.

Biswas, S. (2023), "ChatGPT and the future of medical writing", *Radiology*, Vol. 307 No. 2, p. e223312.

Budak, A. and Sarvari, P.A. (2021), "Profit margin prediction in sustainable road freight transportation using machine learning", *Journal of Cleaner Production*, Vol. 314, p. 127990.

Butticè, V., Colombo, M.G., Fumagalli, E. and Orsenigo, C. (2019), "Green oriented crowdfunding campaigns: their characteristics and diffusion in different institutional settings", *Technological Forecasting and Social Change*, Vol. 141, pp. 85-97.

Cacciarelli, D. and Boresta, M. (2022), "What drives a donor? A machine learning-based approach for predicting responses of nonprofit direct marketing campaigns", *Journal of Philanthropy and Marketing*, Vol. 27 No. 2, p. e1724.

Cai, Z., Demuzere, M., Tang, Y. and Wan, Y. (2022), "The characteristic and transformation of 3D urban morphology in three Chinese mega-cities", *Cities*, Vol. 131, p. 103988.

Campbell, J.L. (2007), "Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility", *Academy of Management Review*, Vol. 32 No. 3, pp. 946-967.

Carrera, B., Piñol, V.L., Mata, J.B. and Kim, K. (2022), "A machine learning based classification models for plastic recycling using different wavelength range spectrums", *Journal of Cleaner Production*, Vol. 374, p. 133883.

Chiu, M.C., Wei, C.J., Wang, Y.C. and Kao, M.C. (2023), "Applying unsupervised learning method to develop a regional risk model based on TCFD: a case study in the United States", *Journal of Cleaner Production*, Vol. 400, p. 136669.

Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E. and Herrera, F. (2011), "Science mapping software tools: review, analysis, and cooperative study among tools", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 7, pp. 1382-1402.

Coro, G. and Trumpy, E. (2020), "Predicting geographical suitability of geothermal power plants", *Journal of Cleaner Production*, Vol. 267, p. 121874.

D'Amato, D., Droste, N., Allen, B., Kettunen, M., Lähtinen, K., Korhonen, J., Leskinen, P., Matthies, B.D. and Toppinen, A. (2017), "Green, circular, bio economy: a comparative analysis of sustainability avenues", *Journal of Cleaner Production*, Vol. 168, pp. 716-734.

Dangelico, R.M. (2016), "Green product innovation: where we are and where we are going", *Business Strategy and the Environment*, Vol. 25 No. 8, pp. 560-576.

Darnall, N., Henriques, I. and Sadorsky, P. (2010), "Adopting proactive environmental strategy: the influence of stakeholders and firm size", *Journal of Management Studies*, Vol. 47 No. 6, pp. 1072-1094.

De Brito, M.P., Carbone, V. and Blanquart, C.M. (2008), "Towards a sustainable fashion retail supply chain in Europe: organisation and performance", *International Journal of Production Economics*, Vol. 114 No. 2, pp. 534-553.

De Las Heras, A., Luque-Sendra, A. and Zamora-Polo, F. (2020), "Machine learning technologies for sustainability in smart cities in the Post-COVID era", *Sustainability*, Vol. 12 No. 22, p. 9320.

Department of Economic and Social Affairs (DESA) (2022), "Do you know all 17 SDGs? United Nations", available at: https://sdgs.un.org/goals

Dong, Y., Wang, J., Niu, X. and Zeng, B. (2023), "Combined water quality forecasting system based on multiobjective optimization and improved data decomposition integration strategy", *Journal of Forecasting*, Vol. 42 No. 2, pp. 260-287.

El Hathat, Z., Sreedharan, V.R., Venkatesh, V.G., Zouadi, T., Arunmozhi, M. and Shi, Y. (2023), "Modelling and analyzing the GHG emissions in the VUCA world: evidence from tomato production in Morocco", *Journal of Cleaner Production*, Vol. 382, p. 134862.

El-Alfy, E.S.M. and Mohammed, S.A. (2020), "A review of machine learning for big data analytics: bibliometric approach", *Technology Analysis and Strategic Management*, Vol. 32 No. 8, pp. 984-1005.

Elkington, J. (1998), "Partnerships from cannibals with forks: the triple bottom line of 21st-century business", *Environmental Quality Management*, Vol. 8 No. 1, pp. 37-51.

Fahdi, S., Elkhechafi, M. and Hachimi, H. (2021), "Machine learning for cleaner production in port of Casablanca", *Journal of Cleaner Production*, Vol. 294, p. 126269.

Fahimnia, B., Sarkis, J. and Davarzani, H. (2015), "Green supply chain management: a review and bibliometric analysis", *International Journal of Production Economics*, Vol. 162, pp. 101-114.

Firat, M. (2023), "How chat GPT can transform autodidactic experiences and open education", Department of Distance Education, Open Education Faculty, Anadolu Unive.

Gao, Y., Chang, D. and Chen, C.H. (2023), "A digital twin-based approach for optimizing operation energy consumption at automated container terminals", *Journal of Cleaner Production*, Vol. 385, p. 135782.

Garg, A. and Lam, J.S.L. (2016), "Modeling multiple-response environmental and manufacturing characteristics of EDM process", *Journal of Cleaner Production*, Vol. 137, pp. 1588-1601.

Gimenez, C., Sierra, V. and Rodon, J. (2012), "Sustainable operations: their impact on the triple bottom line", *International Journal of Production Economics*, Vol. 140 No. 1, pp. 149-159.

Golafshani, E.M., Kashani, A., Behnood, A. and Kim, T. (2023), "Modeling the chloride migration of recycled aggregate concrete using ensemble learners for sustainable building construction", *Journal of Cleaner Production*, Vol. 407, p. 136968.

Grybauskas, A., Stefanini, A. and Ghobakhloo, M. (2022), "Social sustainability in the age of digitalization: a systematic literature review on the social implications of industry 4.0", *Technology in Society*, Vol. 70, p. 101997.

Gue, I.H.V., Lopez, N.S.A., Chiu, A.S., Ubando, A.T. and Tan, R.R. (2022), "Predicting waste management system performance from city and country attributes", *Journal of Cleaner Production*, Vol. 366, p. 132951.

Gunawan, J., Permatasari, P. and Tilt, C. (2020), "Sustainable development goal disclosures: do they support responsible consumption and production?", *Journal of Cleaner Production*, Vol. 246, p. 118989.

Hall, J., Matos, S. and Silvestre, B. (2012), "Understanding why firms should invest in sustainable supply chains: a complexity approach", *International Journal of Production Research*, Vol. 50 No. 5, pp. 1332-1348.

Hidalgo, A., Gabaly, S., Morales-Alonso, G. and Urueña, A. (2020), "The digital divide in light of sustainable development: an approach through advanced machine learning techniques", *Technological Forecasting and Social Change*, Vol. 150, p. 119754.

Hull, C.E. and Rothenberg, S. (2008), "Firm performance: the interactions of corporate social performance with innovation and industry differentiation", *Strategic Management Journal*, Vol. 29 No. 7, pp. 781-789.

Huo, W., Zhu, Z., Sun, H., Ma, B. and Yang, L. (2022), "Development of machine learning models for the prediction of the compressive strength of calcium-based geopolymers", *Journal of Cleaner Production*, Vol. 380, p. 135159.

Iftikhar, B., Alih, S.C., Vafaei, M., Elkotb, M.A., Shutaywi, M., Javed, M.F., … and Aslam, F. (2022), "Predictive modeling of compressive strength of sustainable rice husk ash concrete: ensemble learner optimization and comparison", *Journal of Cleaner Production*, Vol. 348, p. 131285.

Jabeur, S.B., Gharib, C., Mefteh-Wali, S. and Arfi, W.B. (2021), "CatBoost model and artificial intelligence techniques for corporate failure prediction", *Technological Forecasting and Social Change*, Vol. 166, p. 120658.

Jacobs, B.W., Singhal, V.R. and Subramanian, R. (2010), "An empirical investigation of environmental performance and the market value of the firm", *Journal of Operations Management*, Vol. 28 No. 5, pp. 430-441.

Jafari, S.Q., Shokouhyar, S. and Shokoohyar, S. (2022), "Producer-consumer sustainability continuum: mutual understanding to implement extended producer responsibility", *Journal of Cleaner Production*, Vol. 374, p. 133880.

Jamwal, A., Agrawal, R. and Sharma, M. (2022), "Deep learning for manufacturing sustainability: models, applications in industry 4.0 and implications", *International Journal of Information Management Data Insights*, Vol. 2 No. 2, p. 100107.

Jana, R.K., Ghosh, I., Das, D. and Dutta, A. (2021), "Determinants of electronic waste generation in bitcoin network: evidence from the machine learning approach", *Technological Forecasting and Social Change*, Vol. 173, p. 121101.

Jato-Espino, D. and Mayor-Vitoria, F. (2023), "A statistical and machine learning methodology to model rural depopulation risk and explore its attenuation through agricultural land use management", *Applied Geography*, Vol. 152, p. 102870.

Jin, S., Stokes, G. and Hamilton, C. (2023), "Empirical evidence of urban climate adaptation alignment with sustainable development: application of LDA", *Cities*, Vol. 136, p. 104254.

Jin, X., Zuo, X., Dong, X., Dong, Y. and Ding, H. (2021), "Analysis of the policy guarantee mechanism of rural infrastructure based on deep learning", *Technological Forecasting and Social Change*, Vol. 166, p. 120605.

Jing, Z., Luo, Y., Li, X. and Xu, X. (2022), "A multi-dimensional city data embedding model for improving predictive analytics and urban operations", *Industrial Management and Data Systems*, Vol. 122 No. 10.

Jungwirth, D. and Haluza, D. (2023), "Feasibility study on utilization of the artificial intelligence GPT-3 in public health".

Kazemeini, A. and Swei, O. (2023), "Identifying environmentally sustainable pavement management strategies via deep reinforcement learning", *Journal of Cleaner Production*, Vol. 390, p. 136124.

Khanna, P. and Khanra, S. (2023), "Citizen empowerment through smart surveillance: evidence from Indian smart cities", *Digital Policy, Regulation and Governance*, Vol. 25 No. 4, pp. 385-401.

Khanra, S., Joseph, R.P. and Ruparel, N. (2019), "Dynamism of an e-government network in delivering public services", *Academy of Management Global Proceedings*, p. 376.

Khodabandehlou, S. and Rahman, M.Z. (2017), "Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior", *Journal of Systems and Information Technology*, Vol. 19 Nos 1/2, pp. 65-93.

Kondapaka, P., Khanra, S., Malik, A., Kagzi, M. and Hemachandran, K. (2023), "Finding a fit between CXO's experience and AI usage in CXO decision-making: evidence from knowledge-intensive professional service firms", *Journal of Service Theory and Practice*, Vol. 33 No. 2, pp. 280-308.

Kim, B., Yoo, M., Park, K.C., Lee, K.R. and Kim, J.H. (2021), "A value of civic voices for smart city: a big data analysis of civic queries posed by Seoul citizens", *Cities*, Vol. 108, p. 102941.

Kosovic, IN., Mastelic, T. and Ivankovic, D. (2020), "Using artificial intelligence on environmental data from Internet of things for estimating solar radiation: comprehensive analysis", *Journal of Cleaner Production*, Vol. 266, p. 121489.

Krause, P.J. and Bokinala, V. (2023), "A tutorial on data mining for Bayesian networks, with a specific focus on IoT for agriculture", *Internet of Things*, Vol. 22, p. 100738.

Kumar, N., Poonia, V., Gupta, B.B. and Goyal, M.K. (2021), "A novel framework for risk assessment and resilience of critical infrastructure towards climate change", *Technological Forecasting and Social Change*, Vol. 165, p. 120532.

Lai, J., Luo, T., Liu, X., Huang, L., Yu, Z. and Wang, Y. (2023), "Inferring household size distribution and its association with the built environment using massive mobile phone data", *Cities*, Vol. 136, p. 104253.

Lakner, C., Mahler, D.G., Negre, M. and Prydz, E.B. (2022), "How much does reducing inequality matter for global poverty?", *The Journal of Economic Inequality*, Vol. 20 No. 3, pp. 559-585.

Lan, H., Zheng, P. and Li, Z. (2021), "Constructing urban sprawl measurement system of the Yangtze river economic belt zone for healthier lives and social changes in sustainable cities", *Technological Forecasting and Social Change*, Vol. 165, p. 120569.

Lee, C. and Lim, C. (2021), "From technological development to social advance: a review of industry 4.0 through machine learning", *Technological Forecasting and Social Change*, Vol. 167, p. 120653.

Leong, W.D., Teng, S.Y., How, B.S., Ngan, S.L., Abd Rahman, A., Tan, C.P., Ponnambalam, S.G. and Lam, H.L. (2020), "Enhancing the adaptability: lean and green strategy towards the industry revolution 4.0", *Journal of Cleaner Production*, Vol. 273, p. 122870.

Li, X., Fu, J., Jiang, D., Lin, G. and Cao, C. (2022a), "Land use optimization in Ningbo city with a coupled GA and plus model", *Journal of Cleaner Production*, Vol. 375, p. 134004.

Li, J., Suvarna, M., Li, L., Pan, L., Pérez-Ramírez, J., Ok, Y.S. and Wang, X. (2022b), "A review of computational modeling techniques for wet waste valorization: research trends and future perspectives", *Journal of Cleaner Production*, Vol. 367, p. 133025.

Li, Q., Wang, Z., Li, G., Zhou, C., Chen, P. and Yang, C. (2023), "An accurate and adaptable deep learning-based solution to floating litter cleaning up and its effectiveness on environmental recovery", *Journal of Cleaner Production*, Vol. 388, p. 135816.

Li, P., Xiao, X. and Seekamp, E. (2022c), "Climate adaptation planning for cultural heritages in coastal tourism destinations: a multi-objective optimization approach", *Tourism Management*, Vol. 88, p. 104380.

Li, F., Xie, Z., Clarke, K.C., Li, M., Chen, H., Liang, J. and Chen, Z. (2019), "An agent-based procedure with an embedded agent learning model for residential land growth simulation: the case study of Nanjing", China", *Cities*, Vol. 88, pp. 155-165.

Li, X., Yi, S., Cundy, A.B. and Chen, W. (2022d), "Sustainable decision-making for contaminated site risk management: a decision tree model using machine learning algorithms", *Journal of Cleaner Production*, Vol. 371, p. 133612.

Liu, G., Ayello, F., Vera, J., Eckert, R. and Bhat, P. (2021a), "An exploration on the machine learning approaches to determine the erosion rates for liquid hydrocarbon transmission pipelines towards safer and cleaner transportations", *Journal of Cleaner Production*, Vol. 295, p. 126478.

Liu, Z., Wu, R., Chen, Y., Fang, C. and Wang, S. (2021b), "Factors of ecosystem service values in a fast-developing region in China: insights from the joint impacts of human activities and natural conditions", *Journal of Cleaner Production*, Vol. 297, p. 126588.

Lohrmann, A., Lohrmann, C. and Luukka, P. (2022), "Global and regional models for identification of cooling technology in thermal power generation for water demand estimations in water-energy nexus studies", *Journal of Cleaner Production*, Vol. 380, p. 134842.

Lu, J., Dai, E., Miao, Y. and Kusnierek, K. (2022), "Improving active canopy sensor-based in-season rice nitrogen status diagnosis and recommendation using multi-source data fusion with machine learning", *Journal of Cleaner Production*, Vol. 380, p. 134926.

McMeekin, A. and Southerton, D. (2012), "Sustainability transitions and final consumption: practices and socio-technical systems", *Technology Analysis and Strategic Management*, Vol. 24 No. 4, pp. 345-361.

Ma, Y., Miao, R., Chen, Z., Zhang, B. and Bao, L. (2022), "An interpretable analytic framework of the relationship between carsharing station development patterns and built environment for sustainable urban transportation", *Journal of Cleaner Production*, Vol. 377, p. 134445.

Mansouri, S. and Momtaz, P.P. (2022), "Financing sustainable entrepreneurship: ESG measurement, valuation, and performance", *Journal of Business Venturing*, Vol. 37 No. 6, p. 106258.

Mansoursamaei, M., Moradi, M., González-Ramírez, R.G. and Lalla-Ruiz, E. (2023), "Machine learning for promoting environmental sustainability in ports", *Journal of Advanced Transportation*, Vol. 2023.

Mele, M. and Magazzino, C. (2020), "A machine learning analysis of the relationship among iron and steel industries, air pollution, and economic growth in China", *Journal of Cleaner Production*, Vol. 277, p. 123293.

Miguéis, V.L., Pereira, A., Pereira, J. and Figueira, G. (2022), "Reducing fresh fish waste while ensuring availability: demand forecast using censored data and machine learning", *Journal of Cleaner Production*, Vol. 359, p. 131852.

Munir, M.J., Kazmi, S.M.S., Wu, Y.F., Lin, X. and Ahmad, M.R. (2022), "Development of novel design strength model for sustainable concrete columns: a new machine learning-based approach", *Journal of Cleaner Production*, Vol. 357, p. 131988.

Naseri, H., Jahanbakhsh, H., Hosseini, P. and Nejad, F.M. (2020), "Designing sustainable concrete mixture by developing a new machine learning technique", *Journal of Cleaner Production*, Vol. 258, p. 120578.

Naveed, K., Watanabe, C. and Neittaanmäki, P. (2017), "Co-evolution between streaming and live music leads a way to the sustainable growth of music industry–lessons from the US experiences", *Technology in Society*, Vol. 50, pp. 1-19.

Nilashi, M., Ahani, A., Esfahani, M.D., Yadegaridehkordi, E., Samad, S., Ibrahim, O., Sharef, N.M. and Akbari, E. (2019a), "Preference learning for eco-friendly hotels recommendation: a multi-criteria collaborative filtering approach", *Journal of Cleaner Production*, Vol. 215, pp. 767-783.

Nilashi, M., Rupani, P.F., Rupani, M.M., Kamyab, H., Shao, W., Ahmadi, H., Rashid, T.A. and Aljojo, N. (2019b), "Measuring sustainability through ecological sustainability and human sustainability: a machine learning approach", *Journal of Cleaner Production*, Vol. 240, p. 118162.

Pérez-Ortiz, M., de La PazMarín, M., Gutiérrez, P.A. and Hervás-Martínez, C. (2014), "Classification of EU countries' progress towards sustainable development based on ordinal regression techniques", *Knowledge-Based Systems*, Vol. 66, pp. 178-189.

Prioux, N., Ouaret, R., Hetreux, G. and Belaud, J.P. (2023), "Environmental assessment coupled with machine learning for circular economy", *Clean Technologies and Environmental Policy*, Vol. 25 No. 2, pp. 689-702.

Priyadarshini, I., Alkhayyat, A., Obaid, A.J. and Sharma, R. (2022a), "Water pollution reduction for sustainable urban development using machine learning techniques", *Cities*, Vol. 130, p. 103970.

Priyadarshini, I., Sahu, S., Kumar, R. and Taniar, D. (2022b), "A machine-learning ensemble model for predicting energy consumption in smart homes", *Internet of Things*, Vol. 20, p. 100636.

Rahman, M., Chen, N., Islam, M.M., Mahmud, G.I., Pourghasemi, H.R., Alam, M., Rahim, M.A., Baig, M. A., Bhattacharjee, A. and Dewan, A. (2021), "Development of flood hazard map and emergency

relief operation system using hydrodynamic modeling and machine learning algorithm", *Journal of Cleaner Production*, Vol. 311, p. 127594.

Redchuk, A., Walas Mateo, F., Pascal, G. and Tornillo, J.E. (2023), "Adoption case of IIoT and machine learning to improve energy consumption at a process manufacturing firm, under industry 5.0 model", *Big Data and Cognitive Computing*, Vol. 7 No. 1, p. 42.

Ribeiro, V.M. (2021), "Sulfur dioxide emissions in Portugal: prediction, estimation and air quality regulation using machine learning", *Journal of Cleaner Production*, Vol. 317, p. 128358.

Rieder, E., Schmuck, M. and Tugui, A. (2022), "A scientific perspective on using artificial intelligence in sustainable urban development", *Big Data and Cognitive Computing*, Vol. 7 No. 1, p. 3.

Robert, K.W., Parris, T.M. and Leiserowitz, A.A. (2005), "What is sustainable development? Goals, indicators, values, and practice", *Environment: Science and Policy for Sustainable Development*, Vol. 47 No. 3, pp. 8-21.

Sadrossadat, E., Basarir, H., Karrech, A. and Elchalakani, M. (2023), "Innovative AI-based multi-objective mixture design optimisation of CPB considering properties of tailings and cement", *International Journal of Mining, Reclamation and Environment*, Vol. 37 No. 2, pp. 1-17.

Saura, J.R., Ribeiro-Soriano, D. and Saldaña, P.Z. (2022), "Exploring the challenges of remote work on twitter users' sentiments: from digital technology development to a post-pandemic era", *Journal of Business Research*, Vol. 142, pp. 242-254.

Seuring, S. (2013), "A review of modeling approaches for sustainable supply chain management", *Decision Support Systems*, Vol. 54 No. 4, pp. 1513-1520.

Shafi, A., Chen, S., Waleed, M. and Sajjad, M. (2023), "Leveraging machine learning and remote sensing to monitor long-term spatial-temporal wetland changes: towards a national RAMSAR inventory in Pakistan", *Applied Geography*, Vol. 151, p. 102868.

Shahidzadeh, M.H., Shokouhyar, S., Javadi, F. and Shokoohyar, S. (2022), "Unscramble social media power for waste management: a multilayer deep learning approach", *Journal of Cleaner Production*, Vol. 377, p. 134350.

Sharma, R., Kamble, S.S., Gunasekaran, A., Kumar, V. and Kumar, A. (2020), "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance", *Computers and Operations Research*, Vol. 119, p. 104926.

Smith, M.R. and Marx, L. (1994), *Does Technology Drive History?: The Dilemma of Technological Determinism*, MIT Press.

Sohani, A., Sayyaadi, H., Cornaro, C., Shahverdian, M.H., Pierro, M., Moser, D., Karimi, N., Doranehgard, M.H. and Li, L.K. (2022a), "Using machine learning in photovoltaics to create smarter and cleaner energy generation systems: a comprehensive review", *Journal of Cleaner Production*, Vol. 364, p. 132701.

Sohani, A., Sayyaadi, H., Miremadi, S.R., Samiezadeh, S. and Doranehgard, M.H. (2022b), "Thermo-electro-environmental analysis of a photovoltaic solar panel using machine learning and real-time data for smart and sustainable energy generation", *Journal of Cleaner Production*, Vol. 353, p. 131611.

Steward, F. (2012), "Transformative innovation policy to meet the challenge of climate change: sociotechnical networks aligned with consumption and end-use as new transition arenas for a lowcarbon society or green economy", *Technology Analysis and Strategic Management*, Vol. 24 No. 4, pp. 331-343.

Svanberg, J., Ardeshiri, T., Samsten, I., Öhman, P., Rana, T. and Danielson, M. (2022), "Prediction of environmental controversies and development of a corporate environmental performance rating methodology", *Journal of Cleaner Production*, Vol. 344, p. 130979.

Taghizadeh-Mehrjardi, R., Nabiollahi, K., Rasoli, L., Kerry, R. and Scholten, T. (2020), "Land suitability assessment and agricultural production sustainability using machine learning models", *Agronomy*, Vol. 10 No. 4, p. 573.

Tandon, A., Dhir, A., Islam, N. and Mäntymäki, M. (2020), "Blockchain in healthcare: a systematic literature review, synthesizing framework and future research agenda", *Computers in Industry*, Vol. 122, p. 103290.

Teng, S.Y., Galvis, L., Blanco, C.M., Özkan, L., Barendse, R., Postma, G. and Jansen, J. (2023), "Machine-learned digital phase switch for sustainable chemical production", *Journal of Cleaner Production*, Vol. 382, p. 135168.

Tsoka, T., Ye, X., Chen, Y., Gong, D. and Xia, X. (2022), "Explainable artificial intelligence for building energy performance certificate labelling classification", *Journal of Cleaner Production*, Vol. 355, p. 131626.

Veblen, T. (1978), *The Theory of Business Enterprise*, Transaction Publishers.

Walk, J., Kühl, N., Saidani, M. and Schatte, J. (2023), "Artificial intelligence for sustainability: facilitating sustainable smart product-service systems with computer vision", *Journal of Cleaner Production*, Vol. 402, p. 136748.

Wang, J. and Biljecki, F. (2022), "Unsupervised machine learning in urban studies: a systematic review of applications", *Cities*, Vol. 129, p. 103925.

Wang, Z.J. and Zhao, L.T. (2021), "The impact of the global stock and energy market on EU ETS: a structural equation modelling approach", *Journal of Cleaner Production*, Vol. 289, p. 125140.

Wang, R., Asghari, V., Hsu, S.C., Lee, C.J. and Chen, J.H. (2020), "Detecting corporate misconduct through random Forest in china's construction industry", *Journal of Cleaner Production*, Vol. 268, p. 122266.

Wang, D., Fu, J., Xie, X., Ding, F. and Jiang, D. (2022a), "Spatiotemporal evolution of urban-agricultural-ecological space in China and its driving mechanism", *Journal of Cleaner Production*, Vol. 371, p. 133684.

Wang, Y., Yuan, Q., Li, T. and Zhu, L. (2022b), "Global spatiotemporal estimation of daily high-resolution surface carbon monoxide concentrations using deep Forest", *Journal of Cleaner Production*, Vol. 350, p. 131500.

Wang, S., Zhang, Z., Wang, Z. and Liu, G. (2022c), "Exploring the multidimensional factors and emergence mechanisms of industrial symbiotic relationships based on machine learning", *Journal of Cleaner Production*, Vol. 381, p. 135169.

Warsame, A.A., Abdi, A.H., Amir, A.Y. and Azman-Saini, W.N.W. (2023), "Towards sustainable environment in Somalia: the role of conflicts, urbanization, and globalization on environmental degradation and emissions", *Journal of Cleaner Production*, Vol. 406, p. 136856.

Wilson, C. and Van Der Velden, M. (2022), "Sustainable AI: an integrated model to guide public sector decision-making", *Technology in Society*, Vol. 68, p. 101926.

Wu, Z. and Pagell, M. (2011), "Balancing priorities: decision-making in sustainable supply chain management", *Journal of Operations Management*, Vol. 29 No. 6, pp. 577-590.

Wu, C., Lin, C., Barnes, D. and Zhang, Y. (2020), "Partner selection in sustainable supply chains: a fuzzy ensemble learning model", *Journal of Cleaner Production*, Vol. 275, p. 123165.

Yin, H. and Zhou, K. (2022), "Performance evaluation of China's photovoltaic poverty alleviation project using machine learning and satellite images", *Utilities Policy*, Vol. 76, p. 101378.

Zafar, M.H., Khan, N.M., Mansoor, M. and Khan, U.A. (2022), "Towards green energy for sustainable development: machine learning based MPPT approach for thermoelectric generator", *Journal of Cleaner Production*, Vol. 351, p. 131591.

Zhao, S., Liu, Y., Zhang, R. and Fu, B. (2020), "China's population spatialization based on three machine learning models", *Journal of Cleaner Production*, Vol. 256, p. 120644.

Zhu, Q., Zhang, F., Liu, S. and Li, Y. (2022), "An anticrime information support system design: application of K-means-VMD-BiGRU in the city of Chicago", *Information and Management*, Vol. 59 No. 5, p. 103247.

Zolbanin, H.M., Davazdahemami, B., Delen, D. and Zadeh, A.H. (2022), "Data analytics for the sustainable use of resources in hospitals: predicting the length of stay for patients with chronic diseases", *Information and Management*, Vol. 59 No. 5, p. 103282.

Further reading

Bruce, A., Adam, K.E., Buller, H., Chan, K.W. and Tait, J. (2021), "Creating an innovation ecosystem for rapid diagnostic tests for livestock to support sustainable antibiotic use", *Technology Analysis and Strategic Management*, Vol. 34 No. 11, pp. 1-14.

Caviggioli, F. and Ughetto, E. (2019), "A bibliometric analysis of the research dealing with the impact of additive manufacturing on industry, business and society", *International Journal of Production Economics*, Vol. 208, pp. 254-268.

Chabowski, B.R., Mena, J.A. and Gonzalez-Padron, T.L. (2011), "The structure of sustainability research in marketing, 1958–2008: a basis for future research opportunities", *Journal of the Academy of Marketing Science*, Vol. 39 No. 1, pp. 55-70.

Gunarathne, A.N., Lee, K.H. and Hitigala Kaluarachchilage, P.K. (2021), "Institutional pressures, environmental management strategy, and organizational performance: the role of environmental management accounting", *Business Strategy and the Environment*, Vol. 30 No. 2, pp. 825-839.

Khanra, S., Dhir, A., Kaur, P. and Mäntymäki, M. (2021a), "Bibliometric analysis and literature review of ecotourism: toward sustainable development", *Tourism Management Perspectives*, Vol. 37, p. 100777.

Khanra, S., Dhir, A. and Mäntymäki, M. (2020), "Big data analytics and enterprises: a bibliometric synthesis of the literature", *Enterprise Information Systems*, Vol. 14 No. 6, pp. 737-768.

Khanra, S., Dhir, A., Parida, V. and Kohtamäki, M. (2021b), "Servitization research: a review and bibliometric analysis of past achievements and future promises", *Journal of Business Research*, Vol. 131, pp. 151-166.

Khanra, S., Kaur, P., Joseph, R.P., Malik, A. and Dhir, A. (2022), "A resource-based view of green innovation as a strategic firm resource: present status and future directions", *Business Strategy and the Environment*, Vol. 31 No. 4, pp. 1395-1413.

Nanath, K., Kaitheri, S., Malik, S. and Mustafa, S. (2022), "Examination of fake news from a viral perspective: an interplay of emotions, resonance, and sentiments", *Journal of Systems and Information Technology*, Vol. 24 No. 2, pp. 131-155.

Sahoo, S., Kumar, S., Abedin, M.Z., Lim, W.M. and Jakhar, S.K. (2023), "Deep learning applications in manufacturing operations: a review of trends and ways forward", *Journal of Enterprise Information Management*, Vol. 36 No. 1, pp. 221-251.

Samuel, A.L. (1959), "Some studies in machine learning using the game of checkers", *IBM Journal of Research and Development*, Vol. 3 No. 3, pp. 210-229.

Saunila, M., Rantala, T., Ukko, J. and Havukainen, J. (2019), "Why invest in green technologies? Sustainability engagement among small businesses", *Technology Analysis and Strategic Management*, Vol. 31 No. 6, pp. 653-666.

Van Eck, N.J. and Waltman, L. (2014), "Visualizing bibliometric networks", *Measuring Scholarly Impact*, Springer, Cham, pp. 285-320.

Xu, X., Chen, X., Jia, F., Brown, S., Gong, Y. and Xu, Y. (2018), "Supply chain finance: a systematic literature review and bibliometric analysis", *International Journal of Production Economics*, Vol. 204, pp. 160-173.

Corresponding author

Muneza Kagzi can be contacted at: muneza.kagzi@manipal.edu

Appendix

Table A1.
Brief overview of machine learning algorithms

Sr. no.	Name of ML algorithm	Brief idea of ML algorithm	Reference
1	Agent-based model	ABM is a computational modelling technique that aims to simulate complex associations by representing individual agents and their interactions within a given environment. ABMs are useful to understand and study various aspects, such as social dynamics, ecological systems, economic interactions, traffic patterns and disease spread	Li *et al.* (2019)
2	Artificial neural network (ANN)	ANN is a method based on the functioning of the human brain. It aims to identify patterns and associations from large data for prediction and classification tasks. ANN models consist of interconnected artificial neurons that learn from training data to make predictions or decisions	Tsoka *et al.* (2022)
3	Autoregressive distributed lag model (ARDL)	ARDL model facilitates the analysis of long and short-run dynamics among variables to predict the relationship between the dependent variable and its lagged values. It is commonly used in econometrics and time series analysis	Adedoyin *et al.* (2022)
4	Bayesian network ML algorithm	A Bayesian Network is a machine learning method that uses probabilistic reasoning to model associations between variables. It represents the relationships between variables using a directed acyclic graph, enabling probabilistic inference and decision-making based on observed evidence	Wang and Zhao (2021)
5	CatBoost model	CatBoost refers to "Categorical Boosting" and is particularly useful when dealing with data sets that combine numerical and categorical observations. CatBoost offers several advanced features, including automatic handling of missing values, which simplifies the application of the algorithm	Jabeur *et al.* (2021)
6	Clustering	Clustering is a technique that groups similar data points together based on their inherent attributes. The goal of clustering is to identify patterns or structures within a data set without any prior knowledge of the class labels or target variables. Prominent clustering techniques include K-means, Gaussian mixture models and Hierarchical clustering	Bao and Xie (2022)
7	Decision trees	Decision trees follow a tree-like structure, with internal nodes representing features or attributes, branches representing decisions based on those features, and leaf nodes representing class labels or predicted values. Decision trees are commonly used for classification and regression tasks, providing an intuitive way to make decisions based on the data	Cacciarelli and Boresta (2022)
8	Deep learning (DL)	DL is a subset of machine learning that focuses on training artificial neural networks with multiple layers to learn and extract hierarchical representations from complex data. Deep learning models excel at tasks such as image and speech recognition, natural language processing and generative modelling	Lan *et al.* (2021)

(*continued*)

Table A1.

Sr. no.	Name of ML algorithm	Brief idea of ML algorithm	Reference
9	Dynamic topic modelling	Dynamic topic modeling is a machine learning approach used to analyse and model the evolution of topics from a collection of documents over time. It is an extension of traditional topic modelling algorithms, such as Latent Dirichlet allocation (LDA), that account for the temporal aspect of textual data	Kim *et al.* (2021)
10	Explainable artificial intelligence (XAI)	XAI aims to provide additional information to users, explaining the logic behind proposed solutions generated by AI models. It addresses the "black box" nature of some machine learning algorithms and allows users to understand and trust the decision-making process	Behera *et al.* (2023)
11	Gaussian Naive Bayes	Gaussian Naive Bayes is a machine learning algorithm based on Bayes' theorem. It is commonly used for classification tasks and assumes that the features of the data are independent of each other and follow a normal distribution. Gaussian Naive Bayes is efficient and performs well when the independence assumption holds	Aaldering and Song (2019)
12	Genetic algorithm (GA)	GA is a search and optimization algorithm inspired by the process of natural selection and genetics. It involves generating and evolving a population of potential solutions through successive iterations to find the best solution to a problem. Genetic algorithms are useful for solving complex optimization problems with multiple variables and constraints	Naseri *et al.* (2020)
13	Gradient boosting	Gradient boosting is an ensemble learning method that enhances predictive models by combining multiple weak models, typically decision trees, through an iterative process. It aims to minimize prediction errors by sequentially fitting new models to the residuals of the previous models. Gradient boosting is widely used for regression and classification tasks	Golafshani *et al.* (2023)
14	K-nearest neighbours (KNN)	KNN is a supervised machine learning algorithm used for classification and regression tasks. It operates by identifying the K nearest neighbours of a given data point within the feature space and assigns a class or predicts a value based on the neighbours' labels or values. KNN is simple yet effective, especially in situations where local patterns are important	Lohrmann *et al.* (2022)
15	Long short-term memory (LSTM)	LSTM is a special type of computer programme that helps computers understand and remember things over time, just like our brains do. It is often used in tasks such as speech recognition, language translation and predicting future events, where understanding sequences and capturing long-term dependencies is crucial	Mele and Magazzino (2020)

(*continued*)

Table A1.

Sr. no.	Name of ML algorithm	Brief idea of ML algorithm	Reference
16	Maximum entropy-based algorithm (MaxEnt)	MaxEnt is a machine learning algorithm used for classification and modelling problems. It aims to find the most probable distribution or model given the observed data, subject to a set of constraints. MaxEnt integrates the element of uncertainty and is particularly useful when there is limited knowledge or prior information about the problem	Coro and Trumpy (2020)
17	Neural architecture search (NAS)	NAS aims to systematically search and evaluate various architectures to identify the most effective network design that can achieve superior performance on a specific task. It automates the process of designing neural networks, saving time and effort in the manual trial-and-error process	Teng *et al.* (2023)
18	Planning land use scenarios (PLUS)	PLUS is a method used by city planners and urban developers to make decisions about how to use land in the best possible way. It helps them create plans for different locations and determine what should be constructed there, considering factors such as infrastructure, environmental impact and community needs	Li *et al.* (2022a, 2022b, 2022c, 2022d)
19	Random forest regression (RFR)	RFR is a powerful machine learning algorithm that assists in making predictions and understanding relationships between different variables, similar to solving puzzles. It combines many small decision trees to create a big and powerful model, providing accurate predictions for regression tasks	Wang *et al.* (2020)
20	Soccer league competition algorithm	The Soccer league competition algorithm is an interesting algorithm inspired by soccer games. It works by imitating how teams compete in a soccer league. Each team represents a potential solution to a problem, and the aim is to find the best solution through competition and evolution, similar to how soccer teams compete for the championship	Naseri *et al.* (2020)
21	Supervised machine learning	Supervised learning focuses on developing models on labelled data, where each data point is associated with a corresponding target or label. The goal is for the model to learn a mapping between the input features and the desired output labels. Prominent supervised machine learning algorithms include SVM, ANN, decision tree, random forest and logistic regression. Supervised learning is useful for tasks such as image classification, spam detection, sentiment analysis and more	Abdella *et al.* (2020)
22	Support vector machine (SVM)	SVM is a machine learning algorithm that aims to facilitate classification and regression tasks. It works by finding an optimal hyperplane that separates different classes or predicts continuous values based on the input features. SVMs are effective in handling high-dimensional data and are widely used in various applications, including image recognition, text classification and bioinformatics	Naseri *et al.* (2020)

(*continued*)

Table A1.

Sr. no.	Name of ML algorithm	Brief idea of ML algorithm	Reference
23	Transfer learning	Transfer learning is a machine learning technique that involves leveraging knowledge or models based on one task to improve performance on a different but related task. Instead of developing a model from scratch, transfer learning allows researchers to transfer the knowledge or representations learned from one task to another, saving time and computational resources	Bedi (2022)
24	Unsupervised machine learning	Unsupervised learning involves training a model on unlabelled data, where there are no predefined target labels. The aim is to explore inherent patterns, structures or associations within the data without any explicit guidance or supervision. Unsupervised learning algorithms aim to discover meaningful representations or groupings in the data based on its inherent structure. Prominent unsupervised machine learning algorithms include k-means, hierarchical clustering, principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE). Unsupervised learning is beneficial when there is no prior knowledge about the data, and discovering the underlying structure of the data is of interest	Chiu *et al.* (2023)
25	Water cycle algorithm (WCA)	WCA replicates the flow of natural water cycles observed in a natural setting to arrive at solutions for optimization problems. The algorithm mimics the process of evaporation, condensation and precipitation to find optimal solutions similar to how water cycles in nature. It is particularly useful for solving optimization problems with complex constraints and multiple objectives	Naseri *et al.* (2020)

Source: Created by authors