symmetry

MDPI

Review

Artificial Intelligence Methodologies for Data Management

Joel Serey [1], Luis Quezada [1], Miguel Alfaro [1], Guillermo Fuertes [1,2,*], Manuel Vargas [1], Rodrigo Ternero [1,3], Jorge Sabattin [4], Claudia Duran [5] and Sebastian Gutierrez [6,7]

1 Industrial Engineering Department, University of Santiago de Chile, Avenida Ecuador 3769, Santiago 9170124, Chile; joel.serey@usach.cl (J.S.); luis.quezada@usach.cl (L.Q.); miguel.alfaro@usach.cl (M.A.); manuel.vargasg@usach.cl (M.V.); rodrigo.ternero@usach.cl (R.T.)
2 Facultad de Ingeniería, Ciencia y Tecnología, Universidad Bernardo O'Higgins, Avenida Viel 1497, Ruta 5 Sur, Santiago 8370993, Chile
3 Escuela de Construcción, Universidad de las Américas, Santiago 7500975, Chile
4 Facultad de Ingeniería, Universidad Andres Bello, Antonio Varas 880, Santiago 7500971, Chile; jorge.sabattin@unab.cl
5 Departamento de Industria, Facultad de Ingeniería, Universidad Tecnológica Metropolitana, Santiago 7800002, Chile; c.durans@utem.cl
6 Facultad de Economía, Gobierno y Comunicaciones, Universidad Central de Chile, Santiago 8330507, Chile; sebastian.gutierrez@ucentral.cl
7 Facultad de Ciencias, Universidad Mayor, Chile, Santiago 7500628, Chile
* Correspondence: guillermo.fuertes@usach.cl

Abstract: This study analyses the main challenges, trends, technological approaches, and artificial intelligence methods developed by new researchers and professionals in the field of machine learning, with an emphasis on the most outstanding and relevant works to date. This literature review evaluates the main methodological contributions of artificial intelligence through machine learning. The methodology used to study the documents was content analysis; the basic terminology of the study corresponds to machine learning, artificial intelligence, and big data between the years 2017 and 2021. For this study, we selected 181 references, of which 120 are part of the literature review. The conceptual framework includes 12 categories, four groups, and eight subgroups. The study of data management using AI methodologies presents symmetry in the four machine learning groups: supervised learning, unsupervised learning, semi-supervised learning, and reinforced learning. Furthermore, the artificial intelligence methods with more symmetry in all groups are artificial neural networks, Support Vector Machines, K-means, and Bayesian Methods. Finally, five research avenues are presented to improve the prediction of machine learning.

Keywords: machine learning; data management; artificial intelligence; big data

check for updates

Citation: Serey, J.; Quezada, L.; Alfaro, M.; Fuertes, G.; Vargas, M.; Ternero, R.; Sabattin, J.; Duran, C.; Gutierrez, S. Artificial Intelligence Methodologies for Data Management. *Symmetry* **2021**, *13*, 2040. https://doi.org/10.3390/sym13112040

Academic Editor: Jeng-Shyang Pan

Received: 21 August 2021
Accepted: 14 October 2021
Published: 29 October 2021

1. Introduction

Information asymmetry in business based on data management can be reduced by using machine learning (ML) techniques, allowing free competition between market agents. Information asymmetry in data management comes from two sources: (i) patterns of public information not observed by some actors in the negotiation and (ii) actions carried out by an economic actor that are difficult to read by the rest of the market. The concept of ML has contributed to the new industrial revolution (industry 4.0) in particular, with the massive use of big data (BD) and cloud techniques. Obtaining information through real-time data processing is a strategy that offers competitive advantages for decision-making, regardless of the size and commercial sector of the organization.

The machine learns by improving its calculation results without human intervention. The machine needs three fundamental elements for this learning: process data, communication with the cloud and BD, and calculation models. These three elements require technological advances in data science, ML, and artificial intelligence (AI).

At the beginning of the 19th century, mainly in the mail order industry, the practice of collecting, preserving, analyzing, and using data was initiated [1]. The increasing use of email and the web, together with the recording of interactions (speech-text analyzer software and the flow of clickstreams from websites, among others), led to an explosion in data volumes.

For the purpose of this article, a definition adjusted of AI to our general line of thought is the one proposed by Haenlein and Kaplan [2]. These authors defined AI as a system's ability to interpret external data correctly, learn from said data, and use the knowledge to achieve specific goals and tasks through flexible adaptation [3]. In recent years, the use of AI in the service industry has been rapidly gaining ground. According to Xu et al. [4], AI in the context of customer service is defined as a technology-enabled system that evaluates service scenarios in real-time, using data collected from digital and/or physical sources to provide recommendations, alternatives, and personalized solutions to customer queries or problems. For example, AI technology can personalize services and product processing, processing past customer purchases and records. According to Dwivedi et al. [5], AI includes different branches used to obtain data some examples are: expert systems, ML, pattern recognition (PR), fuzzy logic, evolutionary computing, deep learning (DL), probability theory, discriminant analysis, support systems, learning systems, decision trees, and DL with convolutional neural networks (CNN). These approaches have become very popular in recent years for being powerful visual models that automatically produce hierarchies of characteristics [6], commonly trained by supervised learning. These types of networks are feedforward-type models [7]. The difference between CNN and artificial neural networks (ANN) lies in the hidden layer. The hidden layer of a CNN model is generally made up of three layers, namely the convolutional layer, the subsampling layer (clustering layer), and the fully connected layer [8]. CNN has made impressive achievements in many areas, including but not limited to computer vision and natural language processing (NLP) [9].

Interest in AI has grown in different areas of engineering, achieving significant and hopeful methodological advances. Engineering has witnessed the growth of different AI methods in its various areas. One of these methods, and the focus of this review, is ML-supported AI methods for obtaining and managing data developed during the last five years. The scope of the review is to summarize the theoretical background of the methods, provide a critical analysis on their use, and summarize and discuss the latest research on the methodological approaches in the area.

The use of digitization to obtain and manage information was the subject of some previous studies. Stone et al. [10] explained the evolution of ecosystems and the platforms used to obtain customer information and identify the management, research, and teaching implications of this evolution. The model considers the calculation of the customer's functional life value. Zheng et al. [11] surveyed AI-based intelligent visualization tools to extract information from BD. Lin et al. [12] studied the satisfaction of the virtual customer-seller link in the context of conflicting recommendations. Likewise, Refs. [13,14] proposed the use of textual characteristics to analyze the information on the news for sensitive stock market prediction [15].

Zhong et al. [16] presented an AI-based BD analysis for RFID logistics data by defining different behaviors of smart manufacturing objects. Hollebeek et al. [17] explored the use of robotic process automation (RPA), ML, and DL applications. Olshannikova et al. [18] discussed how the capabilities of augmented reality and virtual reality could be applied to the field of BD visualization. Brill et al. [19] conducted a qualitative empirical study to determine the level of satisfaction with digital assistants (Apple's Siri, Amazon's Alexa, Google's Google Assistant). Based on RPA, ML, and DL, Sampson [20] proposed a strategic framework to face the increasing effects of automation in highly skilled professional services jobs. Pantano and Pizzi [21] investigated the technological advance of chatbots, indicating the real areas of development; providing a complete understanding of the actual progress. Xiao and Kumar [22] carried out a conceptual framework that includes antecedents and consequences of the adoption and integration of robotics by companies in their customer service, technology marketing, and information technology operations.

Recently, Hoyer et al. [23] proposed a new framework to understand the role of new technologies powered by AI (internet of things (IoT), augmented reality, virtual reality, mixed reality (MR), virtual assistants, chatbots, robots, blockchain, and 3D printing) in the customer/buyer process. In addition, Duan et al. [24] and Liebowitz [25] analyzed the advancement of AI technology and its capacity to process BD for decision-making. Duan et al. propose twelve research proposals in AI information systems. Kokina and Davenport [26] discussed cognitive AI capabilities in auditing processes, with four large accounting firms launching numerous projects. While Singh et al. [27] developed a conceptual framework of companies' capabilities to operate with "one voice" to offer fluid, harmonious, and reliable interactions through various interfaces. Authors such as Kreutzer and Sirrenberg [28] evaluated the capacity of AI systems for: prediction and profiling of potential customers, conversational commerce, sentiment analysis, and the creation and distribution of content. Furthermore, Heller et al. [29] proposed an integrated framework to automate services based on augmented reality.

The articles of the debate highlight methodological advances aimed at developing AI applications mainly in the service industry (obtaining and managing data to aid decision-making). Methods, such as RPA, PR, and ML, have seen remarkable developments and increased use in database development and optimization in recent years. Recently, to address the limitations of RPA, authors, such as Berruti et al. [30], have proposed intelligent process automation, which refers to the combination of AI, ML, and cognitive automation.

The popularity of ML is by and large due to the availability of powerful new computing tools and hardware and the increasing ease of generating and having access to large datasets, but adoption has been slow. Taking all web search categories into account, a google trends analysis [31] (Figure 1) on the popularity of ML, BD, and AI interestingly reveals an increase arithmetic mean in ML since 2016, reaching a peak between the years 2018 to 2020. Meanwhile, in AI, the behavior is stable without presenting high peaks from 2011 to 2016. However, the popularity of AI increased from 2016 to 2018, which corresponds to the positive result of the ML. Finally, the behavior of BD has remained stable from 2014 to 2021, presenting some popularity peaks every year. Together, this information highlights a positive correlation between ML, BD, and AI. All the above is evidenced in the growing behavior of the number of articles per year. In this study, the comparison of AI/ML/BD trends was only illustrative to show recent growth of AI and ML compared to BD.

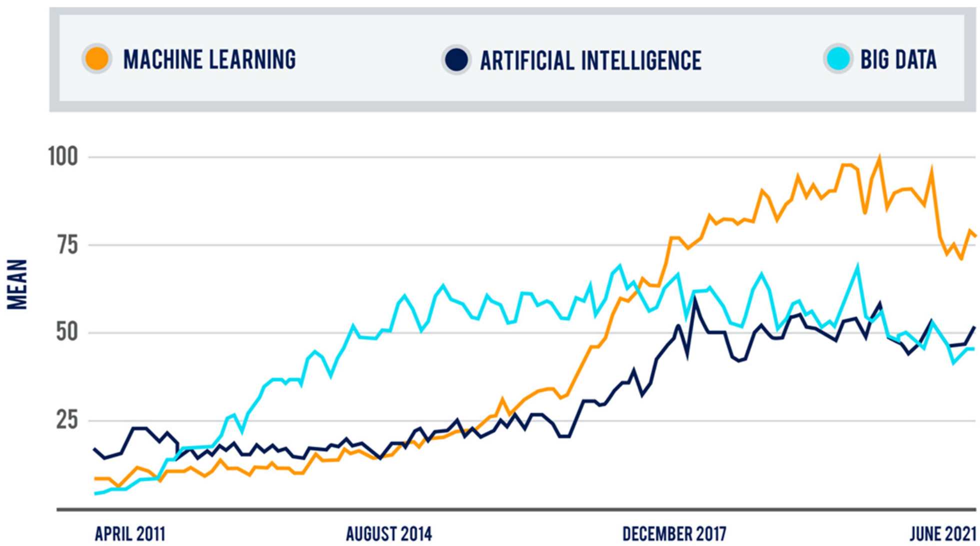

Figure 1. ML vs. AI/BD domain popularity trend.

The review article is structured as follows. Section 1 presents a general introduction to the subject of AI and its importance for obtaining and managing data. Within the subject, we present the opinions of different authors, where some advances in AI applications are discussed. Next, we correlate the popularity of the ML domain, and the main limitations and contributions of the study are revealed. Section 2 describes the methodology used. Section 3 presents a conceptual framework for classifying studies and provides a literature review of the latest AI methodologies used in the ML domain, where the differences between these techniques are detailed. Additionally, a descriptive analysis of the studies is carried out. Section 4 presents the discussions on the study thematic. Finally, the conclusions are provided in Section 5.

Limitations and Contributions

This review article presents a broad perspective of research efforts on using emerging ML-supported AI methods for data collection and management. Research question: What are the main methodological proposals for data management that contribute to the development of the ML domain? This study is limited to the literature regarding AI/ML applications in relation to engineering disciplines. For each ML group, the review of each article focuses on the domain addressed by the study, the subgroup and type of research to which the study belongs, the research results, and the AI method used.

The contributions of this review article are: (1) studying and summarizing the AI categories used to obtain customer information; (2) defining and analyzing the main groups and subgroups that make up the ML theme; (3) identifying study types, future directions, and emerging approaches that use AI supported by ML methodologies for data management; (4) identifying the main AI methodological approaches used in the last five years to obtain and manage data; and (5) highlighting the main AI categories, areas of knowledge, research results, and current limitations/challenges of AI methods with ML.

The ML domain is constantly growing, and it is not possible to cover all of the algorithms in a single article. The multidisciplinary nature of ML was the most challenging difficulty to overcome in this review. However, the contributions of co-authors allowed the search to be limited to widely-used AI methodologies.

2. Methodology

This work corresponds to an extensive review of the literature published of the recent advances in methodological proposals that contribute to developing the AI concept supported by ML technology. The methodology used for this literature review was content analysis; a valid technique for the study of scientific documents [32], used to: identify, classify, and analyze services in smart cities [33], study advances in nanotechnology applied to innovative packaging [34], propose a conceptual framework for strategic management [35], and analyze reverse-logistics models aimed at solid waste management [36–38].

This review identifies and analyzes research that proposes new AI methods emerging as reliable and efficient tools in data management. The development of the proposed methodology provides technical background on the indicated methods and knowledge on using these algorithms for data management problems. AI methodological developments for BD processing as a solution for data management were used by Allam and Dhunny [39] to propose a framework that regulates and formulates BD processing policies through AI and ML aimed at the smart city concept.

Using the same methodology, Henrique et al. [40] analyzed different ML methods and techniques to predict financial market values, resulting in a bibliographic review of the most critical studies on this topic. Likewise, van Klompenburg et al. [41] used it to extract and synthesize ML algorithms used in predictive studies of agricultural crop yield.

This study is divided into categories, groups, and subgroups. The categories are represented by the 12 proposed emerging AI technologies. The groups constitute the four AI techniques represented in the ML domain. Each group contains the methodological contributions (subgroups) that illustrate some of the most outstanding algorithms used in the ML, identifying their degree of development through the investigations.

In this work, a systematic review of the scientific literature published between the years 2017 and 2021 has been carried out. For its preparation, the guidelines of the PRISMA statement [42,43] have been followed. Figure 2 summarizes the proposed PRISMA methodology. The systematic search was carried out with the Google Scholar search engine using the WOS and Scopus digital platforms, mainly databases, such as Springer Link, EmeraldInsight, Science Direct, Wiley Online Library, Taylor & Francis Group, and IEEE Xplore Digital Library. The keywords used were machine learning, data management, BD, and artificial intelligence.

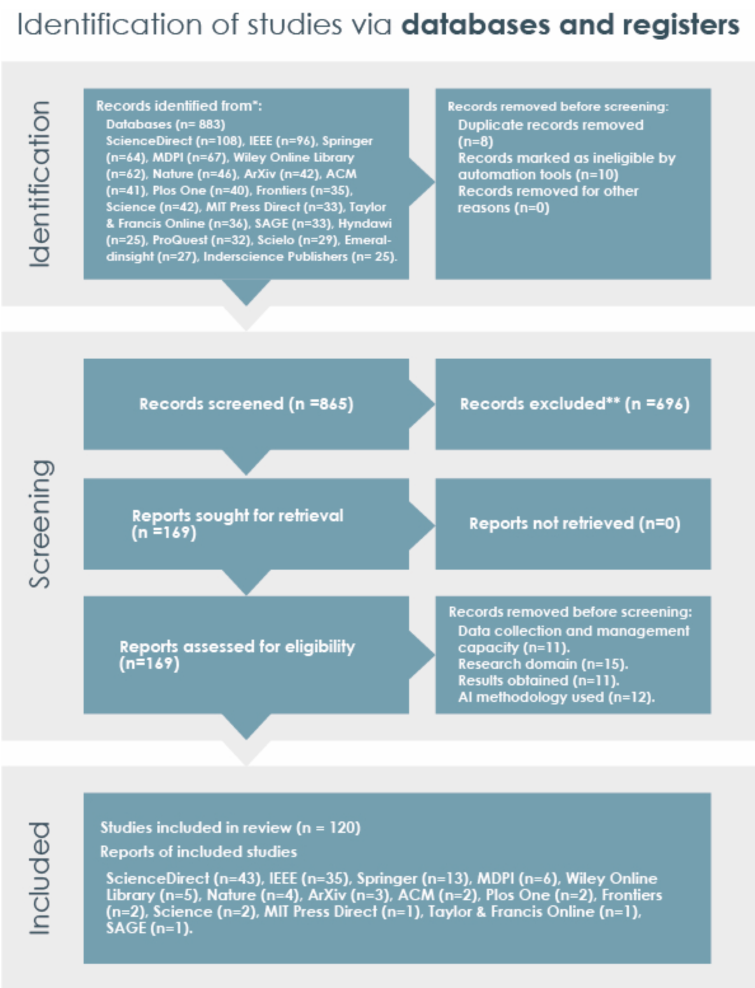

Figure 2. PRISMA flow diagram on three levels.

In addition, to choose an article, two types of quality measures were mainly considered: journal impact factor (JIF) and journal citation indicator (JCI), while not being excluding factors (Table S1). Initially, we reviewed about 4000 publications in scientific journals, identifying 883 articles for the first step. For the second step, the studies were selected by reviewing the most relevant titles. Subsequently, in the third step, the summaries were read. After the final choice (reading the abstracts), in the fourth and last step, we read the complete publication. Then, the articles were reviewed in terms of the inclusion criteria: (1) scientific studies that are part of the WOS and SCOPUS digital platforms; (2) which propose methodological solutions for data collection and management; (3) that in the context of methodological advance, the conceptual bias is studied; (4) that develop comparisons between the solution methods and results obtained; and (5) that have been published between the years 2017 and 2021. The exclusion criteria were: (1) data collection and management capacity, (2) research domain, (3) results obtained, and (4) AI methodology used.

Categorical Classification of the Emerging AI Technologies, ML Groups and Subgroups

The 120 investigations classified in the literature review contribute to the development of the main AI technologies. To facilitate the analysis of the ML domain, based on the 12 technologies proposed by Purcell and Curram [44], we sought to establish a conceptual framework, which is summarized in Figure 3.

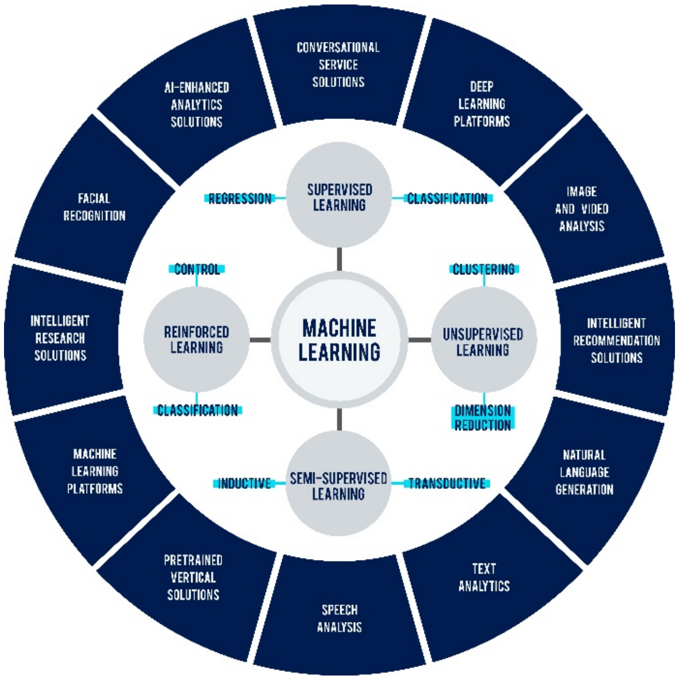

Figure 3. Overview of ML techniques/emerging AI technologies.

This study proposes to adapt and classify the AI technologies proposed by [44] into 12 AI categories used to obtain and manage data (Table 1); with four being the mature categories offering commercial value and impact on customer perception (AI-enhanced analytics solutions, DL platform, natural language generation (NLG), and speech analytics) [45]. These technologies can be used at different levels to provide optimal solutions to specific problems.

Table 1. AI categories used for data collection and management.

AI Categories	Description
AI-enhanced analytics solutions	These solutions are a new generation of business intelligence. It relies on NLP and NLG, as well as information retrieval, to respond to user queries. Many of these solutions use ML to personalize the user experience (improvements in service delivery), automatically revealing alerts based on preferences learned by the user. Today these solutions still need to learn from human analysts.
Conversational service solutions	They act as virtual agents, using NLP and ML to understand and address individual customer service problems. These solutions provide a conversation interface that is generally text-based but can also include voice or image, allowing users to participate through natural language. A virtual chat agent can quickly answer routine questions providing the requested information.
DL platform	A branch of ML, which provides access to DL algorithms (interconnected neural networks). These platforms are used for image and video recognition and auditory analysis. Each algorithm in the hierarchy applies a non-linear transformation to its input and uses what it learns to create a statistical model as the output. The number of layers and the iterations continue until the output has reached an acceptable level of precision.
Facial recognition	A type of application software that allows people to be identified by analyzing the biometric characteristics of faces (facial patterns). Examples: unlocking of electronic devices, identification of faces in social networks-marketing, virtual payments, security and online education, among others. Similar to image and video analysis, IR has issues with privacy.

Table 1. *Cont.*

AI Categories	Description
Image and video analysis	Understands tools and technology to analyze images and videos in order to understand and interpret objects and the characteristics of objects within them. Although many of these tools are pre-trained, adapting your own needs will have to be refined (transfer training). Currently, these technologies require a considerable repository of relevant and correctly classified images.
Intelligent recommendation solutions	Smart recommendation tools leverage AI to provide users with information search results close to their needs. To do this, these new engines continually learn (1) from the individual behavior and conversational interactions, (2) use DL to classify images, identifying interests and suggesting products, and (3) they use NLP to show users' needs and wants.
Intelligent research solutions	They help examine large amounts of structured and unstructured and internal and external data by leveraging NLP, ML, and, in some cases, NLG to generate information that can be used by product developers, sales teams, marketing specialists, scientific research teams, among others.
Machine learning platforms (ML platforms)	These platforms provide users with tools to build, implement, and monitor ML algorithms. Some platforms offer pre-built algorithms and interactive workflows, while others require a greater understanding of development and coding (regressions, decision trees, Bayesian models, unsupervised grouping methods, and statistical models, among others).
NLG	Includes tools and technology that use advanced models to produce high-quality texts in natural language, generally from a corpus of answers or made up of defined textual components. Currently, this technology provides value in areas such as the production of media content (academic evaluation reports, articles for online newspapers or medical, engineering, financial reports, among others).
Pre-trained vertical solutions	Solutions trained in a specific vertical data corpus with functionality adapted to each sector. Examples: agriculture—collection management, phytosanitary control, machinery and equipment control; financial services—detection of transactions and fraudulent accounts; real estate—synchronization of real estate with different portals, optimization of information for monitoring and commercial actions; investment advisers—client portfolio management; medicine—diagnosis of diseases, optimal treatment methods; journalism-writing articles, among others.
Speech analytics	Also called audio mining, it includes tools that understand and interpret the spoken word. This technology is made up of three parts: acoustic speech recognition; speech to text transcription and text analysis. It is an example of a technology that makes unstructured data ready for analysis.
Text analytics	Text analysis allows you to identify hidden information patterns and structures in the data and gain insight from the document collection for decision-making. Text analytics converts unstructured text data to analyzable structured data. Among the tasks carried out by the text analysis are: descriptive statistical analysis, entity extraction, concept extraction and self-tracking, cross-relevance analysis, sentiment analysis, and automatic categorization.

Table 2 defines the ML groups and subgroups used for classifying the studies.

Table 2. General description of the ML domain for the literature review.

ML Groups	Concept
Supervised learning (SL)	The algorithms work with labeled data, trying to find a model/function that, given the input variables, assigns the appropriate output label. The algorithm is trained with a history of real data and thus learns to assign the appropriate output label to a new value, predicting the output value. If the objective of the model is to forecast continuous variables, it is classified as a regression. However, if the goal is to predict discrete variables, it is known as classification.
Unsupervised learning (UL)	Unlike SL, UL is only given the characteristics without providing the algorithm with any labels. Its function is clustering; therefore, the algorithm should catalog by similarity and create groups. The system itself forms the groups from the input patterns.
Semi-supervised learning (SSL)	Aims to produce better classifiers, combining SL and UL techniques to learn from both labeled and unlabeled data. To later retrain the model, SSL methods use the untagged data to modify or reform the assumptions obtained only from the tagged data.

Table 2. *Cont.*

ML Groups	Concept
Reinforced learning (RL)	RL tries to get an agent or intelligent machine to learn to decide through their own experience. In other words, depending on the environment (real or virtual), in a given situation, execute the best possible action through an interactive trial and error process, depending on the observed state (knowledge of the environment). As a result, the agent gets the best possible reward.
ML Subgroups	**Concept**
Regression	Y function attempts to predict the estimated value of a response variable based on one or more independent variables of interest; that is, it predicts the Y value (dependent variable), given values of the X variable. Models can be linear, exponential, or logarithmic. Forecast of continuous variables.
Classification	Used when the expected result is a discrete label. Different performance metrics are used to evaluate the classification models (accuracy, precision, sensitivity, specificity, and F1 score). Depending on the target classes, the model prediction can be binary or multivariate.
Clustering	Consists of grouping a series of vectors according to a criterion in groups or clusters. Usually, the criterion is similarity (grouping similar vectors into groups). These models predict which is the best grouping of data.
Dimension reduction	The process of reducing the number of random variables involved (dimension reduction). These algorithms map a dataset to subspaces derived from the original space of less dimension, which describes the data at a lower cost.
Inductive	Methods that build a classifier that can generate predictions for any object in the input space.
Transductive	Methods that are limited during the training phase in obtaining label predictions for unlabeled data points. These methods are based on graphics.
Control	Flexible and adaptable methods that can be introduced into a control system to analyze differential equations (components of the control loop).

3. Literature Review

The literature review corresponds to the period 2017–2021. Different authors have provided models to the SL, UL, SSL, and RL groups during this period. These authors propose AI methodologies to solve ML problems in data collection and management. For each group, the maximum limit was 30 articles. A total of 120 studies were selected and are classified in Table 3.

Table 3. Literature review—AI methodological proposals supported by ML for data collection and management.

Groups	Subgroups	2017–2021	AI Methodology Used
1. Supervised learning	Regression	2	ANN
		1	Statistical algorithm
		1	Logistic regression
		1	Regression tree
	Classification	3	Support vector machines (SVM)
		1	Bayesian methods
		2	Random Forests
		1	Reduced complexity algorithm
		1	Decision tree
		13	ANN
		1	Manifold learning
2. Unsupervised learning	Clustering	3	Hierarchical clustering
		6	K-Means
		1	Proven predictive coding
		11	ANN
		1	Hidden Markov model
		1	Graphic schema theory

Table 3. *Cont.*

Groups	Subgroups	2017–2021	AI Methodology Used
2. Unsupervised learning	Dimension reduction	1 2 1 1	Geographically weighted regression ANN Bayesian methods Principal component analysis
3. Semi-supervised learning	Inductive	1 1 1 1 2 8 1 1 3 1	Virtual adversarial training Multiple learning Bayesian methods K nearest neighbor Graphic Schema Theory ANN Gaussian mixture model Fuzzy C-Means SVM MixMatch algorithm
	Transductive	1 3 1 1 1	SOGSFE algorithm ANN Bayesian methods Gaussian mixture model K nearest neighbor
4. Reinforced learning	Control	2 9 1 1 3	Markov's decision processes Learning the time difference ANN Dynamic scheduling Deep reinforcement learning Ad-Hoc Techniques Direct policy search
	Classification	6 1	ANN Multiple learning

3.1. *Supervised Learning Models for Data Collection and Management*

Table 4 classified and listed the group 1 SL model (Table 3) literature investigations into 12 AI categories (mentioned in Table 1). SL can be used in regression problems and classification problems. In regression problems, the outputs are continuous, while in a classification problem, the outputs are categorical.

Table 4. Classification of SL models for data collection and management.

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
AI-enhanced analytic solutions	Supervised learning under distributed features	Classification/Research article [46]	Convergence in new solutions	Reduced complexity algorithm
	Grouping of patents	Classification/Case study [47]	Method comparison	SVM
Conversational service solutions	-	-	-	-
DL platforms	Lung carcinoma detection	Classification/Research article [48]	Improved prediction rate and time	ANN
	Characterization of images and videos	Classification/Survey [49]	Self-supervised approaches	ANN
	Self-supervision of medical images	Classification/Research article [50]	Improved semantic performance	ANN

Table 4. *Cont.*

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
DL platforms	Deep convolutional network training	Classification/Research article [51]	Improved error rate	ANN
Facial Recognition	-	-	-	-
Analysis of videos and images	Structural damage detection	Classification/Research article [52]	Improved prediction rate and time	ANN
	Featured Object Detection	Classification/Research article [53]	Unified optimization framework	ANN
	Structural magnetic resonance	Classification/Research article [54]	Quality in automatic scans	Random forest
	Large-scale image search	Classification/Research article [55]	Improved prediction rate and time	ANN
	Plant segmentation	Classification/Research article [56]	Improved prediction rate and time	ANN
Smart recommendation solutions	Analysis of mobility patterns	Classification/Research article [57]	Improved prediction rate and time	Bayesian methods
Smart research solutions	Quantum classifier	Classification/Research article [58]	Improved prediction rate and time	SVM
	Opinion spammers	Regression/Research article [59]	Algorithm performance	Logistic regression
	Comparison of classification methods	Classification/Survey [60]	Algorithm performance	Literature review
	Bibliometric research	Classification/Research article [61]	Algorithm performance	ANN
ML platforms	SCADA-based intrusion detection	Both/Survey [62]	Method comparison	Systematic review
	Navigation system	Regression/Research article [63]	Integration of the BADGR algorithm	ANN
	Accumulated local effects	Regression/Research article [64]	Framework for adjusting predictors	ANN
	Detection of pipe damage	Regression/Research article [65]	Improved prediction rate and time	Statistical algorithm
	Unsafe behavior in construction	Classification/Research article [66]	Improved prediction rate and time	Decision tree
NLG	Training of coding models	Regression/Research article [67]	Representation of sentences	Regression tree
Pre-trained vertical solutions	Connection between ANNs	Both/Review article [68]	Method comparison	ANN
	Training of spiking networks	Classification/Research article [69]	Improved error rate	ANN
	PR in biology and medicine	Classification/Research article [70]	Method comparison	SVM
	Prediction of mental disorders	Both/Review article [71]	Method comparison	Literature review
	Training of a multilayer spiking network	Classification/Research article [72]	Improved error rate	ANN
Speech analytics	Supervised voice separation	Both/Review article [73]	Method comparison	ANN

Table 4. *Cont.*

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
Text analysis	Contrastive learning	Classification/Review article [74]	Method comparison	Literature index
	Fake news detection	Classification/Research article [14]	Method comparison	Random forest

In SL, the correct input/output pairs are available, and the goal is to correctly map them from the input space to the output space. Table 4 shows that 70% of studies correspond to the classification subgroup, 17% to the regression subgroup, and the remaining 13% to studies where both subgroups are addressed. This indicates a greater interest in the development of categorical methodologies for solving classification problems. In the literature review in Table 4, 22 research articles, 4 literature review articles, 3 surveys and, one case study were found. Thus, the literature on SL presents a significant progress regarding the collection and organization of knowledge, reflected in the solid understanding of the approaches and algorithms developed.

Table 4 shows that the most popular AI methodology for SL is ANN (50% of publications). Subsequently, multidimensional AI methodologies were found, classified as SVM (10% of publications), followed by non-parametric-highly flexible methodologies, such as decision trees (7%). Another recursive partition method that involves predictions based on a collection of individual decision trees is random forests (7%).

The literature reviewed for each SL subgroup has expanded in volume and scope and now encompasses a broad algorithmic spectrum. The main AI methodologies used included comparisons with models of (1) regression: assembly methods, regression analysis, learning metrics, regression tree, non-linear regression, Bayesian model, among others; and (2) classification: K-nearest neighbor, Bayesian belief networks, principal component analysis, linear discriminant analysis, assembly methods, learning metrics, collaborative filtering, etc.

Three AI categories feature the most significant breakthroughs—image and video analytics, ML platforms, and pre-trained vertical solutions, all with five publications. However, we did not find studies for the categories AI-conversational service solutions and facial recognition.

3.2. *Unsupervised Learning Models for Data Collection and Management*

Table 5 detailed the methodological contributions found in the literature review of the group 2 UL model (Table 3) into 12 AI categories (mentioned in Table 1). UL is used to build clustering or dimension reduction models based on the input data without the corresponding output labels [75].

Table 5. UL model classification for data collection and management.

	Domain	Subgroup/Study Type	Results	AI Methodology Used
AI-enhanced analytic solutions	Personal thermal comfort	Clustering/Research article [76]	Improved prediction rate and time	Hidden Markov model
	Modeling human activities	Dimension reduction/Research article [77]	Improved performance	Bayesian methods
Conversational service solutions	Speech features in nature	Clustering/Research article [78]	Improved performance	Proven predictive coding
	Bot and human classification	Clustering/Research article [79]	Improved prediction rate and time	K-means

Table 5. *Cont.*

	Domain	Subgroup/Study Type	Results	AI Methodology Used
Conversational service solutions	Semantic representation of audios	Clustering/Research article [80]	Improved performance	ANN
	Gesture distribution scenarios	Dimension reduction/Research article [81]	Improved prediction rate and time	Geographically weighted regression
DL platforms	Contrast by magnetization transfer	Clustering/Research article [82]	Improved performance	ANN
	Inkjet printing	Clustering/Research article [83]	Behavior prediction	ANN
	Overview -networks domain	Both/Survey [84]	Method comparison	Literature review
	Optimization of wireless systems	Clustering/Research article [85]	Optimization policies	ANN
	Reconfigurable smart surface	Clustering/Research article [86]	Improved performance	ANN
Facial recognition	-	-	-	-
Image and video analysis	Visual tracking	Clustering/Research article [87]	Improved performance	ANN
	Electrical failure diagnosis	Clustering/Research article [75]	Improved performance	K-means
	Extraction of primary objects	Clustering/Research article [88]	Improved prediction rate and time	ANN
	Image fusion	Clustering/Research article [89]	Displayable HDR images	ANN
Smart recommendation solutions.	Searching for new phenomena in data	Clustering/Research article [90]	Statistical test method	Hierarchical clustering
	Detection—false financial positives	Clustering/Research article [91]	Improved prediction rate and time	K-means
Smart research solutions	Prediction of software defects	Both/Review article [92]	Method comparison	Systematic review
	Simultaneous localization and mapping	Dimension reduction/Research article [93]	Improved prediction rate and time	ANN
	Cluster quality	Clustering/Research article [94]	Improved performance	K-means
ML platforms	Evaluation of synaptic dynamics	Clustering/Research article [95]	Improved performance	ANN
	Un-manned surface vehicle	Clustering/Research article [96]	Multiple Tasks Assignment	K-means
	Data—IoT framework	Clustering/Research article [97]	Method comparison	Hierarchical clustering
	Forest fires.	Dimension reduction/Research article [98]	Resource allocation	Principal component analysis
NLG	Automatic speech emotion recognition	Clustering/Research article [99]	Method comparison	ANN

Table 5. *Cont.*

	Domain	Subgroup/Study Type	Results	AI Methodology Used
Pre-trained vertical solutions	Fault detection	Clustering/Research article [100]	Improved performance	ANN
	Clustering	Clustering/Research article [101]	Method comparison	K-means
Speech analytics	Identification of patterns in calls	Clustering/Research article [102]	Reliable mobile network	ANN
Text analysis	Semantic analysis between sentences	Clustering/Research article [103]	Improved performance	Hierarchical clustering
	Patent classification	Clustering/Research article [104]	Improved prediction rate and time	Graphic schema theory

The output data is not available in UL, and the goal is to find patterns in the input data. Table 5 shows that 80% of the studies correspond to the clustering subgroup, 13% to the reduction dimension subgroup, and the remaining 7% of studies addressed both subgroups. This indicates a greater interest in developing exploratory methodologies (structural description of data) for solving clustering problems. According to the analysis of the studies, 28 research articles were found, one literature review article, and one survey; no findings were presented for case studies. Thus, UL takes advantage of large amounts of unlabeled data. Currently, there are significant developments in mathematical modeling, reflected in the number of research articles.

According to Table 5, the most widely used AI solution methodology is the ANN (43% of publications). AI methodologies, such as k-means (20% of publications), aimed to determine the number, quality, and cohesion of the groupings in a data set. Statistical methodologies, such as Markov's (7% of publications), relate observable events and hidden events.

Different AI methodologies in each subgroup include comparisons with models of (1) clustering: partitional clustering, spectral clustering, single-link, bi-clustering, multinominal regression, leader clustering algorithm, gaussian mixture model, non-linear regression, clustering feature tree, literal fuzzy c-means, among others; and (2) dimension reduction: singular value decomposition, independent component analysis, locally linear embedded, spectral embedding, isomap embedding, factor analysis, multidimensional scaling, t-distributed stochastic neighbor embedding, non-negative matrix factorization etc.

The AI category that has made the greatest advancements is DL platforms, with five publications. We did not find studies for the category AI-facial recognition.

3.3. Semi-supervised Learning Models for Data Collection and Management

The SSL approach builds inductive or Transductive models based on the original tagged data and the untagged data with new tags [105]. Table 6 presented an updated methodological description of the group 3 SSL model (Table 3) into 12 AI categories (mentioned in Table 1).

Table 6. Classification of SSL models for obtaining and managing data.

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
AI-enhanced analytic solutions	-	-	-	-
Conversational service solutions	-	-	-	-

Table 6. *Cont.*

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
DL platforms	Training model	Inductive/Research article [106]	Improved error rate	ANN
	Hyperspectral image classification	Inductive/Research article [107]	Improved prediction rate and time	ANN
	Modeling of energy consumption	Inductive/Research article [108]	Improved prediction rate and time	ANN
	Seismic impedance inversion	Transductive/Research article [109]	Improved performance	ANN
	Broad learning system	Transductive/Research article [110]	Method comparison	ANN
Facial recognition	Character extraction	Transductive/Research article [111]	Improved performance	SOGSFE algorithm
Image and video analysis	Federated learning	Inductive/Research article [112]	Information gathering	ANN
	Label propagation and graphic fusion	Inductive/Research article [113]	Improved performance	Graphic schema theory
	Engine failure diagnosis	Inductive/Research article [114]	Improved prediction rate and time	Graphic schema theory
	Image classification model	Inductive/Research article [115]	Improved error rate	MixMatch algorithm
	Diagnosis—computed tomography	Inductive/Research article [116]	Improved prediction rate and time	ANN
	Diagnosis of mechanical failures	Inductive/Research article [117]	Improved performance	Multiple learning
Smart recommendation solutions	Facade defects	Inductive/Research article [118]	Improved prediction rate and time	ANN
	Educational data mining	Inductive/Research article [119]	Method comparison	K nearest neighbor
	Label prediction	Transductive/Research article [120]	Method comparison	ANN
Smart research solutions	Graphics-based learning with security awareness	Inductive/Research article [121]	Adaptive selection of graphics	SVM
	Semi-supervised learning	Both/Survey [105]	Method comparison	Literature survey
	Adaptive graph	Transductive/Research article [122]	Method comparison	K nearest neighbor
	Data security	Both/Review article [123]	Research advances	Literature index
	Graph-based semi-supervised learning	Both/Review article [124]	Method comparison	Literature review
ML platforms	Discriminative and generative models	Inductive/Research article [125]	Method comparison	Bayesian methods
	ANN regularization	Inductive/Research article [126]	Improved performance	Virtual adversarial training
	Cybersecurity in IoT	Inductive/Research article [127]	Improved prediction rate and time	Fuzzy c-means

Table 6. *Cont.*

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
ML platforms	Large data and zero noise graph limits	Transductive/Research article [128]	Improved performance	Bayesian methods
	Cyber threats	Inductive/Research article [129]	Classifier optimization	ANN
NLG	Big social data analysis	Inductive/Research article [130]	Recognition of emotions	SVM
Pre-trained vertical solutions	Manifold adversarial training	Inductive/Research article [131]	Improved performance	Gaussian mixture model
	Hybrid S3L method	Transductive/Research article [132]	Improved performance	Gaussian mixture model
	Broad Learning System	Inductive/Research article [133]	Improved prediction rate and time	ANN
	Diagnosis of mechanical failures	Inductive/Research article [134]	Improved prediction rate and time	SVM
Speech analytics	-	-	-	-
Text analysis	-	-	-	-

SSL takes advantage of large amounts of labeled and untagged data. Conceptually it is situated between SL and UL. Table 6 shows that 67% of studies correspond to the inductive subgroup, 23% to the Transductive subgroup, and the remaining 10% to studies that address both subgroups. The preceding indicates more interest in developing methodologies that optimize predictive models for solving classification problems. In the literature review, we found 26 research articles, 2 literature review articles, one survey, and one case study. In recent years, research in this area has followed the general trends observed in machine learning, with much attention directed to models based on ANNs and generative learning.

According to Table 6, the most widely used AI solution methodology was ANN (36% of publications). Next, was AI methodologies that analyze networks, known as graph theory (7% of publications). Probabilistic graphical methodologies provide simple ways to visualize the structure and properties of a probability model, such as Bayesian methods (7% of publications). Another method found commonly used for clustering data was the Gaussian mixture model (7%) and the k-nearest neighbor classification algorithm (7%).

The AI category with the greatest advancements was image and video analysis, with six publications. We found no publications of AI categories—AI-enhanced analytics solutions, conversational service solutions, speech analytics, and text analytics. We found only one recent survey that collects and organizes this knowledge, which may hamper the ability of researchers and engineers to use SSL. The literature on this subject has expanded in volume and scope and now encompasses a broad spectrum of theories, algorithms, and applications.

3.4. Reinforced Learning Models for Data Collection and Management

An RL algorithm aims to maximize cumulative rewards by learning strategies through interaction with the environment. Table 7 classified and listed the group 4 RL model (Table 3) literature investigations into 12 AI categories (mentioned in Table 1).

RL is a framework for decision-making problems where the agent interacts through trial and error with its environment to discover optimal behavior. Table 7 shows that 57% of studies correspond to the control subgroup, 24% to the classification subgroup, and 18% to studies that address both subgroups. The above indicates a greater interest in methodological developments where computers make decisions on complex and stochastic systems to solve control problems. In the content analysis, we found 23 research articles, 4 literature review articles, and 3 surveys; no findings were presented for case studies.

According to the literature, RL is the most popular technique for artificial agents to learn optimal strategy closely through experience. Such techniques are validated with the algorithmic developments found in the reviewed studies. Different studies analyzed models and solved RL problems using Markov's decision process theory, Monte Carlo, and dynamic programming. RL is a potent engineering tool for modeling dynamic behaviors and achieving goals based on rewards and penalties.

Table 7. Classification of RL models for data collection and management.

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
AI-enhanced analytic solutions	Multi-agent	Control/Research article [135]	Improved performance	ANN
Conversational service solutions	-	-	-	-
DL platforms	Gold immunochromatographic strip	Classification/Research article [136]	Improved performance	ANN
	Intelligent discharge system	Control/Research article [137]	Improved performance	Markov decision processes
	Collision avoidance in pedestrian-rich environments	Control/Research article [138]	Improved performance	ANN
	Autonomous driving	Both/Survey article [139]	Method comparison	Literature survey
	Drone navigation	Control/Research article [140]	Improved prediction rate and time	Direct policy search
	Airfoil shape optimization	Control/Research article [141]	Improved performance	ANN
	Traffic control	Control/Research article [142]	Improved prediction rate and time	Deep reinforcement learning
Facial recognition	Person re-identification	Classification/Research article [143]	Improved prediction rate and time	ANN
Image and video analysis	Robust hierarchical tracker	Classification/Research article [144]	Improved performance	ANN
	Head shake in panoramic video	Control/Research article [145]	Improved performance	ANN
Smart recommendation solutions	Financial forecast	Classification/Research article [146]	Improved prediction rate and time	ANN
	AlphaGo Zero program	Control/Research article [147]	Improved performance	ANN
	Long-term recommendation system	Classification/Research article [148]	Improved prediction rate and time	ANN
Smart research solutions	Learning techniques in economics and finance	Both/Review article [149]	Method comparison	Literature review
	Control of a Quadrotor	Control/Research article [150]	Improved performance	Direct policy search
	Building control	Control/Review article [151]	Identification—trends, progress and gaps	Literature review
	Adaptive controller	Control/Research article [152]	Improved performance	ANN

Table 7. *Cont.*

AI Categories	Domain	Subgroup/Study Type	Results	AI Methodology Used
Smart research solutions	Novo molecular design—ReLeaSE	Classification/Research article [153]	Chemical library	ANN
	Learning techniques in energy systems	Both/Review article [154]	Method comparison	Literature review
	Adaptive optimal control scheme	Control/Research article [155]	Improved performance	ANN
	Approaches in social robotics	Both/Survey article [156]	Method comparison	Literature survey
ML platforms	Network slicing	Control/Research article [157]	Improved performance	ANN
	Inverse reinforcement learning	Both/Survey article [158]	Identification—challenges, methods and advances	Literature survey
	Adaptive production control system	Control/Research article [159]	Improved performance	Markov decision processes
NLG	-	-	-	-
Pre-trained vertical solutions	Dynamic multi-objective optimization problems	Classification/Research article [160]	Improved prediction rate and time	Multiple learning
	Adaptive optimal control	Control/Research article [161]	Improved performance	ANN
	Biological data extraction	Both/Review article [162]	Method comparison	Literature review
	Estimation error	Control/Research article [163]	Improved performance	Direct policy search
	Optimal consensus control	Control/Research article [164]	Improved performance	Dynamic scheduling
Speech analytics	-	-	-	-
Text analysis	-	-	-	-

Table 7 shows that the most popular AI methodology for SL is ANN (50% of publications). Next was control methodologies that assign probabilities, such as the direct search for policies (10%). Followed by methodologies to obtain optimal policies, such as Markov's decision processes (7%).

The two AI categories with the biggest breakthroughs are smart research solutions (8 publications) and DL platforms (7 publications). We found no studies for the categories AI-conversational service solutions, NLG, speech analytics, and text analysis.

3.5. Descriptive Analysis of the Studies

Figure 4, using the VOS viewer software, summarizes the main keywords found in the literature review.

According to Figure 4, the number of papers for each main keyword per year is the following: 2017—machine learning (23 publications); 2018—supervised learning (10 publications); unsupervised learning (20 publications); 2019—semi-supervised learning (21 publications), deep learning (14 publications); 2020—reinforcement learning (13 publications); and 2021—deep reinforcement learning (9 publications).

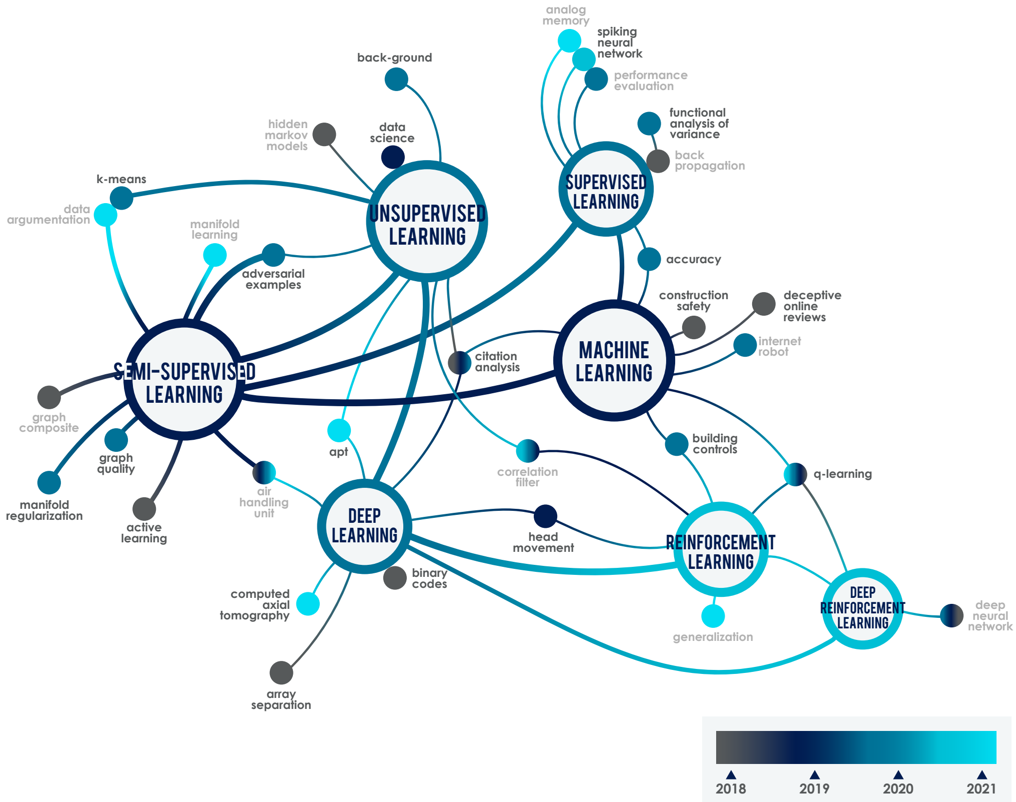

Figure 4. Main keywords found.

The distribution of publications by study type: the type of study most evoked were research articles (with 99 investigations), followed by literature reviews (with 11), surveys (with 8), and case studies (with 2) (Figure 5).

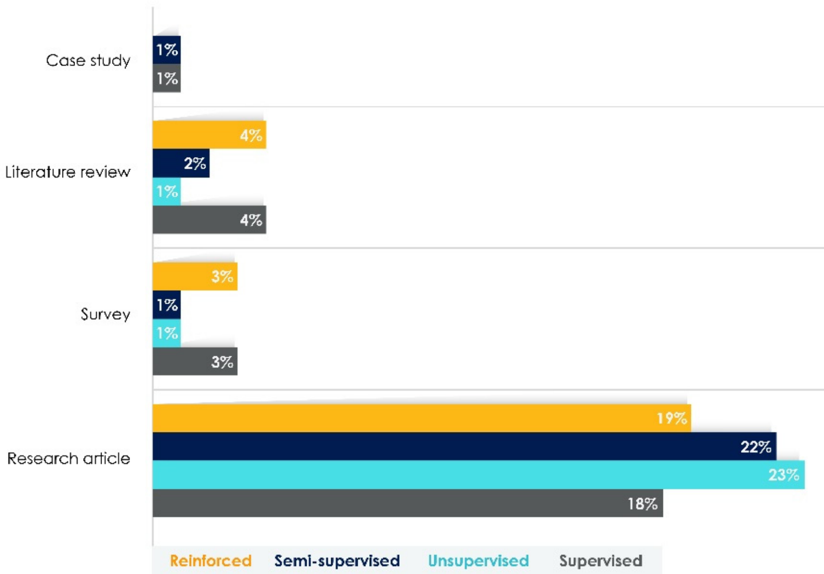

Figure 5. Distribution of publications by study type.

The number of publications per subgroup: eight subgroups with 120 articles were used in the literature review. The classification subgroup made the most significant contribution, adding the groups SL and RL would be 28 publications, followed by the clustering (24 publications), the inductive subgroup (20 publications), the control subgroup (17 publications), transductive (7 publications), regression (5 publications), and dimension reduction (4 publications) (Figure 6).

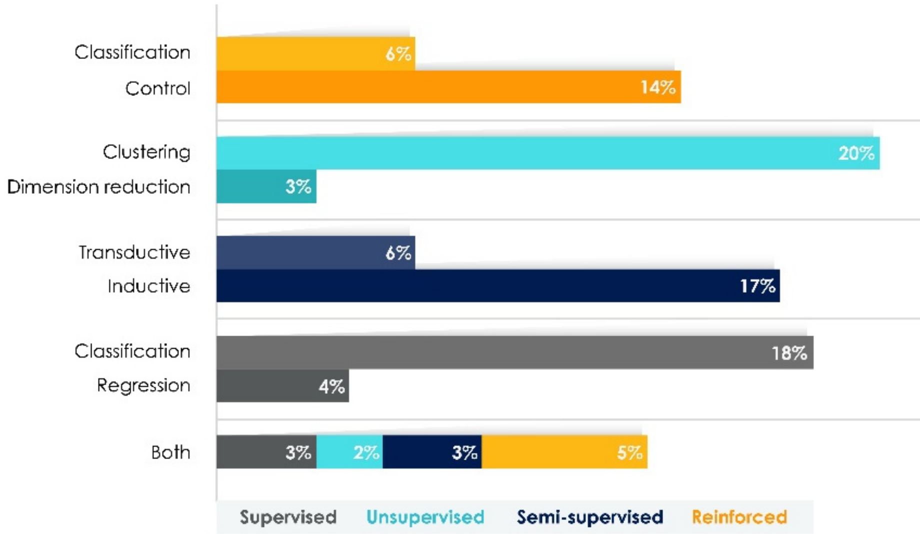

Figure 6. Distribution of studies by subgroups.

The number of articles per year: according to the analysis in Figure 1, the growing popularity of ML, AI, and BD is evident. This confirms the increasing interest that researchers have been giving to ML in the last five years. In this research, 35 articles belonging to the literature review were published in 2020 (Figure 7).

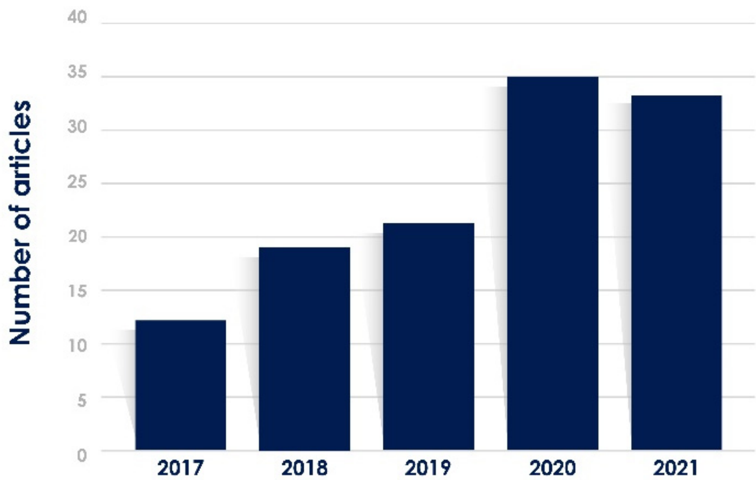

Figure 7. Number of studies per year.

Distribution of AI categories: according to the literature review, five AI categories make the most significant methodological contributions. As shown in Figure 8, the most important is the DL platform (with 21 studies), followed by intelligent research solutions (with 20), image and video analytics—ML platforms (with 17), and pre-trained vertical solutions (with 16).

Distribution of the areas of knowledge: the analysis of the 120 studies that make up this literature review shows that the area of computer engineering and systems presents the most widely used methodological developments (with 62 investigations). This is followed by telecommunications (with 14), infrastructure (with 12), transportation (with 10), health (with 10), the financial area (with 6), marketing and news (with 4), and agriculture (with 2).

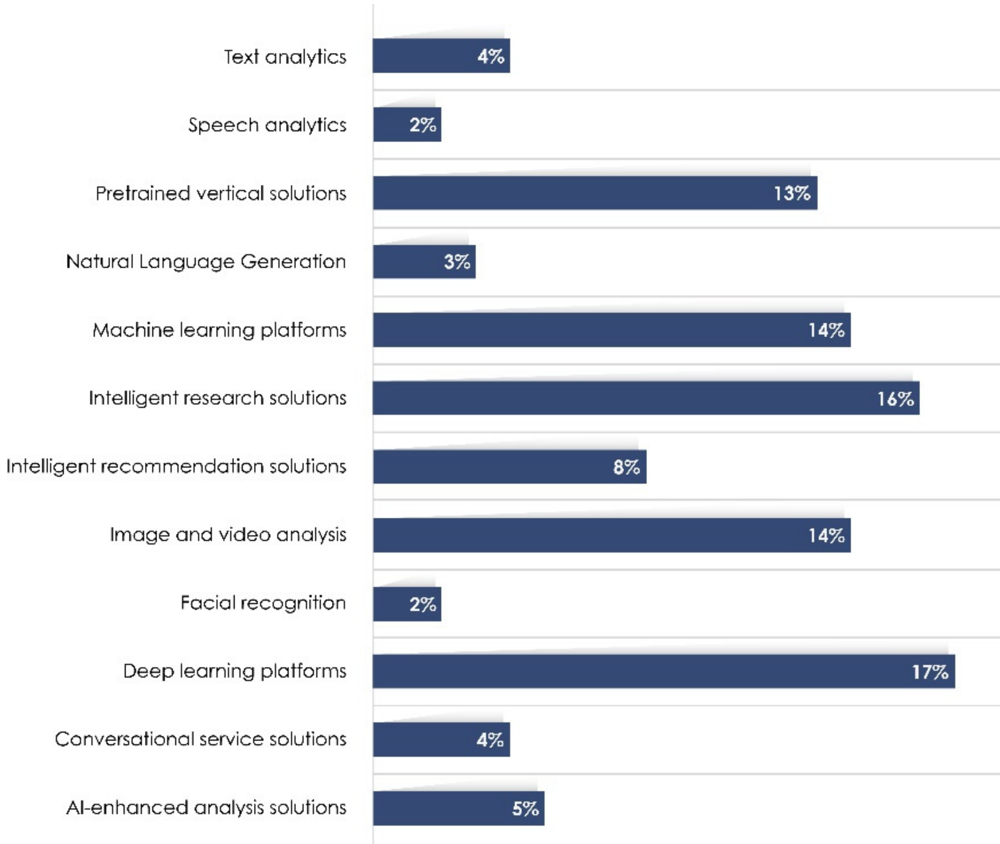

Figure 8. Distribution of AI categories.

The journals with the most publications are Neurocomputing, IEEE Access, IEEE Transactions on Pattern Analysis and Machine Intelligence, and IEEE Transactions on Neural Networks and Learning Systems (Table S1). Regarding the origin of the research by country, China proposes the greatest ML methodological developments (39), followed by the USA (32) (Figure S1). We reviewed 883 publications under the ML, SL, UL, SSL, and RL criteria and/or obtaining and managing information. Finally, 120 publications were selected in the fourth step; the largest number of contributions were made by Science Direct databases, IEEE, and Springer (Figure S2). In total, 113 articles and 7 conference articles made the most important contributions.

4. Discussion

The most widely used ML technique corresponds to SL; even so, today's BD requires UL and RL learning paradigms. However, the accuracy of UL and RL techniques is accompanied by high computational costs.

The literature review suggests using different metrics to evaluate the performance and efficiency of ML models. We found different metrics to evaluate the performance and efficiency of AI methodologies; area under the curve (AUC), Nash–Sutcliffe coefficient radius (NS), relative percentage difference (RPD), precision, accuracy, median absolute error (MedAE), recall, normalized mean squared error (NMSE), root mean square prediction error (RMSEP), mean squared prediction error (MSPE), correlation coefficient (R), and specificity. The accuracy metric was the most used by the classification subgroup, followed by mean absolute percentage error (MAPE), mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), F-score, and normalized root mean square deviation or error (NRMSE) for the regression subgroup. Future works are necessary to obtain precision levels close to 100%.

The analysis of the data partition formats suggests that the most typical partition ratio is (80:20) training/testing. Additionally, other studies adopted the training/validation/test data partition (80:10:10). We found no studies that stated a general rule for adopting data partitioning. The literature suggests that partition formats (80/20) following the Pareto rule provide optimal divisions for AI and ML data analysis.

Based on this study, we observed possible research avenues to improve ML predictions (1) integration of two or more AI methodologies; (2) integration of new AI methodologies with soft computing or other conventional methods; (3) use of data decomposition techniques to improve data set quality; (4) use of a set of methods to generalize models and reduce uncertainty; and (5) use of complementary algorithms to improve the quality of new AI methodological proposals.

The deep RL-DL/RL combination promises to revolutionize the future of AI in areas such as automatic driving, NLP, robots, among others. The findings of the review mainly suggest the use of two types of RL models: when the environment and state are known, they use model-based RL solutions (e.g., AlphaZero); when the environment and the state are partially known, they use the model-free RL, whose algorithms are mainly Q-learning (value-based) and gradient policy algorithms (probability-based).

The ML architecture through the IoT concept analyses and interprets complex and large volumes of data, particularly CNN. The ANN analysis of learning rates found that most studies use fixed rates. However, some studies suggest the use of adjustable rates using special algorithms. Regarding the activation function, it was observed that the linear function was the most used. Additionally, within the findings in ANN, some information processing architectures were found for signals supported on graphs, known as graph neural networks (GNN). Most of these methodological proposals are supported in deep learning and mainly propose GNN architectures based on CNN, recurrent ANN, and deep autoencoders.

The next great challenge lies in the superposition of the four ML groups: the ability to select the most appropriate AI method. This involves anticipating various scenarios (selection of parameters) and dealing with different levels of uncertainty (missing or incomplete data, computational capacity, classification precision, among others).

Despite multiple advancements of the 12 AI categories used for data collection and management, there are still multiple problems, challenges, methodologies, and future trends that AI/ML must overcome. While some UL techniques remove unnecessary data, there is still a need for massive processing power capable of analyzing all scenarios. NLG processing is a long way from being a natural and accurate translation. Jargon, accents, and understanding the language remain big challenges for ML because although image classification is a settled issue, the machine does not really understand the meaning of the image. For now, we continue to classify everything without defining intermediate states, despite the constant developments in fuzzy or soft systems.

The lack of video training is a sensitive topic for ML. Video data sets are much richer in content than still images; therefore, ML needs deeper systems capable of learning and responding efficiently with little input data. This challenge requires solving storage capacity (memory capacity to store past events) through technologies such as a collective memory network between all artificial thinking entities and differentiable neural computers, added to a modular system that integrates different algorithms. The ML reasoning ability is associated with the future development of a model of ideas; this model should serve as an interface, helping to interpret ML's own language.

Currently, changes in the importance and frequency of participating in online activities before and during COVID-19 created new challenges for ML. According to Mouratidis and Papagiannakis [165], during the pandemic, there were substantial increases in the importance of teleworking (31% increase), teleconferencing (34% increase), e-learning (34% increase), and telehealth (21% increase), among others. To reduce the effect of the pandemic on the education sector, most educational institutions were forced to teach online classes. As an academic tool (Zoom Microsoft Teams, Moodle, Google Classroom, virtual reality applications, etc.), the web provides a global open platform for storing data and presenting it in text, graphic, audio, and video formats, and in communication tools for synchronous and asynchronous communication [166]. AI-supported e-learning (AIeL) refers to the use of AI techniques in e-learning (the use of computer and network technologies for learning or training) [167]. Through web platforms, AIeL proposes ML approaches to: identify the learning style and personalize learning experiences [168,169], personalized hybrid recommender for the adjustment or association of content to students [170,171], DL algorithms for monitoring student emotions in real-time [172,173], a multi-agent system to improve the Moodle platform in intelligent tutoring systems [174], a cyber threat detection model in e-learning systems [175], and fuzzy ANN for learning English [176]. Finally, different authors have evaluated the impact of AIeL during the COVID 19 pandemic [177–181].

5. Conclusions

This review article presents the importance of continuous AI methodological developments for ML applications during 2017–2021. A total of 181 studies were used, of which 120 are part of the literary analysis. The literature indicated that, among the numerous methods, ML has been increasingly adopted and used to develop emerging AI technologies. In general, ML areas are closely related, as they fundamentally overlap in scope.

The most used tools to evaluate the performance of AI methods were accuracy, RMSE, R, specificity, MAPE, MPAE, followed by MSE, in addition to generalizability, robustness, calculation cost, and speed. The most commonly used ML algorithm was ANN, followed by SVM, k-means, and Bayesian methods. Some studies adopted hybrid methodologies to harness the power of different techniques to compensate for the weakness of specific techniques. The knowledge areas evaluated make software and systems engineering the next generation approaches to perform data collection and management, with fault diagnosis being the application area with the greatest solution proposals, followed by robotics, autonomous computing, and driving. This review showed that classification tasks are the most frequently used methods by so-called intelligent systems, either by statistical algorithms or AI. The literature also suggests that the AI-based DL platforms require less information, which improves complicated decision problems, making it the alternative solution with the greatest AI methodological proposals.

Based on the methodology proposed in this study, mature AI technologies, such as speech analytics, facial recognition, NLG, and conversational service solutions, had slower methodological developments, showing researchers are interested in less developed AI categories.

Future work should amplify the discussions in the proposed study areas. For example, one of the concepts that needs to be expanded is the emerging PR AI method; the objective of this study would be to know the impact that PR advances generate in different areas of engineering, from a perspective framed in generative models versus discriminative models.

It is essential to broaden the literature review, also focusing on the emerging DL AI method, for example, to analyze the advances that CNN applications have had in different areas of engineering. Likewise, analyzing the methodological advances of DL architectures, recurrent neural networks, automatic encoders, deep belief networks, among others, is important. Finally, the design of an adequate methodology that integrates multiple emerging AI technologies to facilitate data collection and management is envisaged.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/sym13112040/s1, Table S1: Number of publications by journal, Figure S1: Number of publications by country where the study is carried out (2017–2011), Figure S2: Number of articles identified.

Author Contributions: Conceptualization, J.S. (Joel Serey) and L.Q.; methodology, G.F. and J.S. (Joel Serey); software, S.G. and C.D.; validation, M.V., J.S. (Jorge Sabattin) and M.A.; formal analysis, R.T.; investigation, G.F. and J.S. (Joel Serey); resources, L.Q.; data curation, M.A. and M.V.; writing—original draft preparation, C.D.; writing—review and editing, S.G.; visualization, R.T.; supervision, J.S. (Joel Serey); project administration, G.F. and L.Q.; funding acquisition, J.S. (Jorge Sabattin) and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by DICYT (Scientific and Technological Research Bureau) of the University of Santiago of Chile (USACH) and the Department of Industrial Engineering. This work was supported in part by Fondecyt (Chile) Grant No. 11200993 (M.V.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research has been supported by DICYT (Scientific and Technological Research Bureau) of the University of Santiago of Chile (USACH) and the Department of Industrial Engineering.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Abbreviations

The main abbreviations of this work are:

R	Correlation coefficient
AI	Artificial intelligence
BD	Big data
ML	Machine learning
SL	Supervised learning
UL	Unsupervised learning
DL	Deep learning
RL	Reinforced learning
NS	Nash–Sutcliffe coefficient radius
PR	Pattern recognition
ANN	Artificial neural network
GNN	Graph neural networks
CNN	Convolutional neural networks
RPA	Robotic process automation
IoT	Internet of things
NLP	Natural language processing
NLG	Natural language generation
SSL	Semi-supervised learning
AUC	Area under the curve
RPD	Relative percentage difference
MSE	Mean squared error
MAE	Mean absolute error
AIeL	AI-supported e-learning
NMSE	Normalized mean squared error
MSPE	Mean squared prediction error
MAPE	Mean absolute percentage error
RMSE	Root mean squared error
RMSEP	Root mean square prediction error
NRMSE	Normalized root mean square deviation or error
MedAE	Median absolute error

References

1. Stone, M.; Woodcock, N.; Wilson, M. Managing the change from marketing planning to customer relationship Managment. *Long Range Plan.* **1996**, *29*, 675–683. [CrossRef]
2. Haenlein, M.; Kaplan, A. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *Calif. Manag. Rev.* **2019**, *61*, 5–14. [CrossRef]
3. Kaplan, A.; Haenlein, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Bus. Horiz.* **2019**, *62*, 15–25. [CrossRef]
4. Xu, Y.; Shieh, C.H.; van Esch, P.; Ling, I.L. AI customer service: Task complexity, problem-solving ability, and usage intention. *Australas. Mark. J.* **2020**, *28*, 189–199. [CrossRef]
5. Dwivedi, Y.K.; Hughes, L.; Ismagilova, E.; Aarts, G.; Coombs, C.; Crick, T.; Duan, Y.; Dwivedi, R.; Edwards, J.; Eirug, A.; et al. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *Int. J. Inf. Manag.* **2019**, *57*, 101994. [CrossRef]
6. Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [CrossRef]
7. Yu, Y.; Wang, C.; Gu, X.; Li, J. A novel deep learning-based method for damage identification of smart building structures. *Struct. Health Monit.* **2018**, *18*, 143–163. [CrossRef]
8. Yu, H.; Yang, L.T.; Zhang, Q.; Armstrong, D.; Deen, M.J. Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Neurocomputing* **2021**, *444*, 92–110. [CrossRef]
9. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–21. [CrossRef]
10. Stone, M.; Aravopoulou, E.; Gerardi, G.; Todeva, E.; Weinzierl, L.; Laughlin, P.; Stott, R. How platforms are transforming customer information Managment. *Bottom Line* **2017**, *30*, 216–235. [CrossRef]
11. Zheng, Y.; Wu, W.; Chen, Y.; Qu, H.; Ni, L.M. Visual analytics in urban computing: An overview. *IEEE Trans. Big Data* **2016**, *2*, 276–296. [CrossRef]
12. Lin, Y.-T.; Doong, H.-S.; Eisingerich, A.B. Avatar design of virtual salespeople: Mitigation of recommendation conflicts. *J. Serv. Res.* **2021**, *24*, 141–159. [CrossRef]
13. Fuertes, G.; Alfaro, M.; Vargas, M.; Espinoza, A.; Galvez, D.; Sepalveda-Rojas, J.P. Measure of Semantic Likeness among Business Process Activities in a Telecommunication Company. *IEEE Access* **2020**, *8*, 32332–32340. [CrossRef]
14. Reis, J.C.S.; Correia, A.; Murai, F.; Veloso, A.; Benevenuto, F.; Cambria, E. Supervised learning for fake news detection. *IEEE Intell. Syst.* **2019**, *34*, 76–81. [CrossRef]
15. Usmani, S.; Shamsi, J.A. News sensitive stock market prediction: Literature review and suggestions. *PeerJ Comput. Sci.* **2021**, *7*, e490. [CrossRef] [PubMed]
16. Zhong, R.Y.; Xu, C.; Chen, C.; Huang, G.Q. Big data analytics for physical internet-based intelligent manufacturing shop floors. *Int. J. Prod. Res.* **2017**, *55*, 2610–2621. [CrossRef]
17. Hollebeek, L.D.; Sprott, D.E.; Brady, M.K. Rise of the machines? Customer engagement in automated service interactions. *J. Serv. Res.* **2021**, *24*, 3–8. [CrossRef]
18. Olshannikova, E.; Ometov, A.; Koucheryavy, Y.; Olsson, T. Visualizing big data with augmented and virtual reality: Challenges and research agenda. *J. Big Data* **2015**, *2*, 1–27. [CrossRef]
19. Brill, T.M.; Munoz, L.; Miller, R.J. Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *J. Mark. Manag.* **2019**, *35*, 1401–1436. [CrossRef]
20. Sampson, S.E. A strategic framework for task automation in professional services. *J. Serv. Res.* **2021**, *24*, 122–140. [CrossRef]
21. Pantano, E.; Pizzi, G. Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis. *J. Retail. Consum. Serv.* **2020**, *55*, 102096. [CrossRef]
22. Xiao, L.; Kumar, V. Robotics for customer service: A useful complement or an ultimate substitute? *J. Serv. Res.* **2021**, *24*, 9–29. [CrossRef]
23. Hoyer, W.D.; Kroschke, M.; Schmitt, B.; Kraume, K.; Shankar, V. Transforming the customer experience through new technologies. *J. Interact. Mark.* **2020**, *51*, 57–71. [CrossRef]
24. Duan, Y.; Edwards, J.S.; Dwivedi, Y.K. Artificial intelligence for decision making in the era of big data—Evolution, challenges and research agenda. *Int. J. Inf. Manag.* **2019**, *48*, 63–71. [CrossRef]
25. Liebowitz, J. *Data Analytics and AI*; Taylor & Francis: New York, NY, USA, 2020.
26. Kokina, J.; Davenport, T.H. The emergence of artificial intelligence: How automation is changing auditing. *J. Emerg. Technol. Account.* **2017**, *14*, 115–122. [CrossRef]
27. Singh, J.; Nambisan, S.; Bridge, R.G.; Brock, J.K.-U. One-voice strategy for customer engagement. *J. Serv. Res.* **2021**, *24*, 42–65. [CrossRef]
28. Kreutzer, R.T.; Sirrenberg, M. Fields of application of artificial intelligence—Customer service, marketing and sales. In *Understanding Artificial Intelligence*; Springer: Cham, Switzerland, 2020; pp. 105–154.
29. Heller, J.; Chylinski, M.; de Ruyter, K.; Keeling, D.I.; Hilken, T.; Mahr, D. Tangible service automation: Decomposing the technology-enabled engagement process (TEEP) for augmented reality. *J. Serv. Res.* **2021**, *24*, 84–103. [CrossRef]

30. Berruti, F.; Nixon, G.; Taglioni, G.; Whiteman, R. Intelligent Process Automation: The Engine at the Core of the Next-Generation Operating Model. Available online: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/intelligent-process-automation-the-engine-at-the-core-of-the-next-generation-operating-model# (accessed on 12 March 2021).
31. Google Trends Analysis Artificial Intelligence-Big Data-Machine Learning. Available online: https://trends.google.es/trends/explore?date=2011-03-132021-03-13&q=Machinelearning,Artificialintelligence,bigdata (accessed on 13 March 2021).
32. Lopez, F. El análisis de contenido como método de investigación. *Rev. Educ.* **2002**, *4*, 167–180.
33. Serey, J.; Quezada, L.; Alfaro, M.; Fuertes, G.; Ternero, R.; Gatica, G.; Gutierrez, S.; Vargas, M. Methodological proposals for the development of services in a smart city: A literature review. *Sustainability* **2020**, *12*, 10249. [CrossRef]
34. Fuertes, G.; Soto, I.; Carrasco, R.; Vargas, M.; Sabattin, J.; Lagos, C. Intelligent packaging Systems: Sensors and Nanosensors to Monitor Food Quality and Safety. *J. Sens.* **2016**, *2016*, 1–8. [CrossRef]
35. Fuertes, G.; Alfaro, M.; Vargas, M.; Gutierrez, S.; Ternero, R.; Sabattin, J. Conceptual framework for the strategic Managment: A literature review—Descriptive. *J. Eng.* **2020**, *2020*, 1–21. [CrossRef]
36. Banguera, L.; Sepulveda, J.M.; Fuertes, G.; Carrasco, R.; Vargas, M. Reverse and inverse logistic models for solid waste Managment. *S. Afr. J. Ind. Eng.* **2017**, *28*, 120–132. [CrossRef]
37. Vargas, M.; Alfaro, M.; Karstegl, N.; Fuertes, G.; Gracia, M.D.; Mar-Ortiz, J.; Sabattin, J.; Duran, C.; Leal, N. Reverse Logistics for Solid Waste from the Construction Industry. *Adv. Civ. Eng.* **2021**, *2021*, 1–11. [CrossRef]
38. Valenzuela, J.; Alfaro, M.; Fuertes, G.; Vargas, M.; Sáez-Navarrete, C. Reverse logistics models for the collection of plastic waste: A literature review. *Waste Manag. Res.* **2021**, *39*, 1–19. [CrossRef]
39. Allam, Z.; Dhunny, Z.A. On big data, artificial intelligence and smart cities. *Cities* **2019**, *89*, 80–91. [CrossRef]
40. Henrique, B.M.; Sobreiro, V.A.; Kimura, H. Literature review: Machine learning techniques applied to financial market prediction. *Expert Syst. Appl.* **2019**, *124*, 226–251. [CrossRef]
41. van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [CrossRef]
42. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef]
43. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*, n160. [CrossRef]
44. Purcell, B.; Curram, R. *TechRadar^{TM}: Artificial Intelligence Technologies and Solutions*; Q1 2017; Forrester: Cambridge, MA, USA, 2017. Available online: https://www.forrester.com/report/the-top-emerging-technologies-in-artificial-intelligence/RES137806 (accessed on 13 October 2021).
45. Purcell, B. *The Top. Emerging Technologies in Artificial Intelligence*; Forrester: Cambridge, MA, USA, 2017. Available online: https://www.forrester.com/report/TechRadar-Artificial-Intelligence-Technologies-Q1-2017/RES129161 (accessed on 13 October 2021).
46. Ying, B.; Yuan, K.; Sayed, A.H. Supervised learning under distributed featuresss. *IEEE Trans. Signal. Process.* **2018**, *67*, 977–992. [CrossRef]
47. Kyebambe, M.N.; Cheng, G.; Huang, Y.; He, C.; Zhang, Z. Forecasting emerging technologies: A supervised learning approach through patent analysis. *Technol. Forecast. Soc. Chang.* **2017**, *125*, 236–244. [CrossRef]
48. Kanavati, F.; Toyokawa, G.; Momosaki, S.; Rambeau, M.; Kozuma, Y.; Shoji, F.; Yamazaki, K.; Takeo, S.; Iizuka, O.; Tsuneki, M. Weakly-supervised learning for lung carcinoma classification using deep learning. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef] [PubMed]
49. Jing, L.; Tian, Y. Self-supervised visual feature learning with deep neural networks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 1–22. [CrossRef] [PubMed]
50. Chen, L.; Bentley, P.; Mori, K.; Misawa, K.; Fujiwara, M.; Rueckert, D. Self-supervised learning for medical image analysis using image context restoration. *Med. Image Anal.* **2019**, *58*, 101539. [CrossRef]
51. Mostafa, H.; Ramesh, V.; Cauwenberghs, G. Deep supervised learning using local errors. *Front. Neurosci.* **2018**, *12*, 1–16. [CrossRef]
52. Sarmadi, H.; Entezami, A. Application of supervised learning to validation of damage detection. *Arch. Appl. Mech.* **2021**, *91*, 393–410. [CrossRef]
53. Luo, A.; Li, X.; Yang, F.; Jiao, Z.; Cheng, H. Webly-supervised learning for salient object detection. *Pattern Recognit.* **2020**, *103*, 107308. [CrossRef]
54. Klapwijk, E.T.; van de Kamp, F.; van der Meulen, M.; Peters, S.; Wierenga, L.M. Qoala-T: A supervised-learning tool for quality control of FreeSurfer segmented MRI data. *Neuroimage* **2019**, *189*, 116–129. [CrossRef]
55. Yang, H.F.; Lin, K.; Chen, C.S. Supervised learning of semantics-preserving hash via deep convolutional neural networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 437–451. [CrossRef]
56. Adams, J.; Qiu, Y.; Xu, Y.; Schnable, J.C. Plant segmentation by supervised machine learning methods. *Plant. Phenome J.* **2020**, *3*, e20001. [CrossRef]
57. Zhu, Z.; Sun, L.; Chen, X.; Yang, H. Integrating probabilistic tensor factorization with bayesian supervised learning for dynamic ridesharing pattern analysis. *Transp. Res. Part. C Emerg. Technol.* **2021**, *124*, 102916. [CrossRef]
58. Havlíček, V.; Córcoles, A.D.; Temme, K.; Harrow, A.W.; Kandala, A.; Chow, J.M.; Gambetta, J.M. Supervised learning with quantum-enhanced feature spaces. *Nature* **2019**, *567*, 209–212. [CrossRef]

59. Kumar, N.; Venugopal, D.; Qiu, L.; Kumar, S. Detecting review manipulation on online platforms with hierarchical supervised learning. *J. Manag. Inf. Syst.* **2018**, *35*, 350–380. [CrossRef]
60. Sen, P.C.; Hajra, M.; Ghosh, M. Supervised classification algorithms in machine learning: A survey and review. In *Advances in Intelligent Systems and Computing*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 937, pp. 99–111.
61. Song, M.; Kang, K.Y.; Timakum, T.; Zhang, X. Examining influential factors for acknowledgements classification using supervised learning. *PLoS ONE* **2020**, *15*, e0228928. [CrossRef]
62. Suaboot, J.; Fahad, A.; Tari, Z.; Grundy, J.; Mahmood, A.N.; Almalawi, A.; Zomaya, A.Y.; Drira, K. A taxonomy of supervised learning for IDSs in SCADA environments. *ACM Comput. Surv.* **2020**, *53*, 1–37. [CrossRef]
63. Kahn, G.; Abbeel, P.; Levine, S. BADGR: An autonomous self-supervised learning-based navigation system. *IEEE Robot. Autom. Lett.* **2021**, *6*, 1312–1319. [CrossRef]
64. Apley, D.W.; Zhu, J. Visualizing the effects of predictor variables in black box supervised learning models. *J. R. Stat. Soc. Ser. B* **2020**, *82*, 1059–1086. [CrossRef]
65. Sen, D.; Aghazadeh, A.; Mousavi, A.; Nagarajaiah, S.; Baraniuk, R.; Dabak, A. Data-driven semi-supervised and supervised learning algorithms for health monitoring of pipes. *Mech. Syst. Signal. Process.* **2019**, *131*, 524–537. [CrossRef]
66. Goh, Y.M.; Ubeynarayana, C.U.; Wong, K.L.X.; Guo, B.H.W. Factors influencing unsafe behaviors: A supervised learning approach. *Accid. Anal. Prev.* **2018**, *118*, 77–85. [CrossRef]
67. Conneau, A.; Kiela, D.; Schwenk, H.; Barrault, L.; Bordes, A. Supervised Learning of Universal Sentence Representations from Natural Language Inference Data. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark; 2017; pp. 670–680.
68. Wang, X.; Lin, X.; Dang, X. Supervised learning in spiking neural networks: A review of algorithms and evaluations. *Neural Netw.* **2020**, *125*, 258–280. [CrossRef]
69. Mostafa, H. Supervised learning based on temporal coding in spiking neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 3227–3235. [CrossRef]
70. Bzdok, D.; Krzywinski, M.; Altman, N. Machine learning: Supervised methods. *Nat. Methods* **2018**, *15*, 5. [CrossRef] [PubMed]
71. Jiang, T.; Gradus, J.L.; Rosellini, A.J. Supervised machine learning: A brief primer. *Behav. Ther.* **2020**, *51*, 675–687. [CrossRef]
72. Zenke, F.; Ganguli, S. SuperSpike: Supervised learning in multilayer spiking neural networks. *Neural Comput.* **2018**, *30*, 1514–1541. [CrossRef] [PubMed]
73. Wang, D.; Chen, J. Supervised speech separation based on deep learning: An overview. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2018**, *26*, 1702–1726. [CrossRef]
74. Jaiswal, A.; Babu, A.R.; Zadeh, M.Z.; Banerjee, D.; Makedon, F. A survey on contrastive Self-supervised learning. *Technologies* **2021**, *9*, 2. [CrossRef]
75. Swana, E.; Doorsamy, W. An unsupervised learning approach to condition assessment on a wound-rotor induction generator. *Energies* **2021**, *14*, 602. [CrossRef]
76. Ghahramani, A.; Castro, G.; Karvigh, S.A.; Becerik-Gerber, B. Towards unsupervised learning of thermal comfort using infrared thermography. *Appl. Energy* **2018**, *211*, 41–49. [CrossRef]
77. Wu, C.; Zhang, J.; Sener, O.; Selman, B.; Savarese, S.; Saxena, A. Watch-n-patch: Unsupervised learning of actions and relations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 467–481. [CrossRef]
78. Riviere, M.; Dupoux, E. Towards unsupervised learning of speech features in the wild. In Proceedings of the IEEE Spoken Language Technology Workshop, Shenzhen, China, 19–22 January 2021; pp. 156–163.
79. Rovetta, S.; Suchacka, G.; Masulli, F. Bot recognition in a web store: An approach based on unsupervised learning. *J. Netw. Comput. Appl.* **2020**, *157*, 102577. [CrossRef]
80. Jansen, A.; Plakal, M.; Pandya, R.; Ellis, D.P.W.; Hershey, S.; Liu, J.; Moore, R.C.; Saurous, R.A. Unsupervised learning of semantic audio representations. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, AB, Canada, 15–20 April 2018; pp. 126–130.
81. Jirak, D.; Biertimpel, D.; Kerzel, M.; Wermter, S. Solving visual object ambiguities when pointing: An unsupervised learning approach. *Neural Comput. Appl.* **2021**, *33*, 2297–2319. [CrossRef]
82. Kang, B.; Kim, B.; Schär, M.; Park, H.; Heo, H. Unsupervised learning for magnetization transfer contrast MR fingerprinting: Application to CEST and nuclear overhauser enhancement imaging. *Magn. Reson. Med.* **2021**, *85*, 2040–2054. [CrossRef]
83. Huang, J.; Segura, L.J.; Wang, T.; Zhao, G.; Sun, H.; Zhou, C. Unsupervised learning for the droplet evolution prediction and process dynamics understanding in inkjet printing. *Addit. Manuf.* **2020**, *35*, 101197. [CrossRef]
84. Usama, M.; Qadir, J.; Raza, A.; Arif, H.; Yau, K.L.A.; Elkhatib, Y.; Hussain, A.; Al-Fuqaha, A. Unsupervised machine learning for networking: Techniques, applications and research challenges. *IEEE Access* **2019**, *7*, 65579–65615. [CrossRef]
85. Liu, D.; Sun, C.; Yang, C.; Hanzo, L. Optimizing wireless systems using unsupervised and reinforced-unsupervised deep learning. *IEEE Netw.* **2020**, *34*, 270–277. [CrossRef]
86. Gao, J.; Zhong, C.; Chen, X.; Lin, H.; Zhang, Z. Unsupervised learning for passive beamforming. *IEEE Commun. Lett.* **2020**, *24*, 1052–1056. [CrossRef]
87. Wang, N.; Zhou, W.; Song, Y.; Ma, C.; Liu, W.; Li, H. Unsupervised deep representation learning for real-time tracking. *Int. J. Comput. Vis.* **2021**, *129*, 400–418. [CrossRef]

88. Zhao, D.; Ding, B.; Wu, Y.; Chen, L.; Zhou, H. Unsupervised learning from videos for object discovery in single images. *Symmetry* **2021**, *13*, 38. [CrossRef]
89. Qi, Y.; Zhou, S.; Zhang, Z.; Luo, S.; Lin, X.; Wang, L.; Qiang, B. Deep unsupervised learning based on color un-referenced loss functions for multi-exposure image fusion. *Inf. Fusion* **2021**, *66*, 18–39. [CrossRef]
90. De Simone, A.; Jacques, T. Guiding new physics searches with unsupervised learning. *Eur. Phys. J. C* **2019**, *79*, 289. [CrossRef]
91. López de Prado, M.; Lewis, M.J. Detection of false investment strategies using unsupervised learning methods. *Quant. Financ.* **2019**, *19*, 1555–1565. [CrossRef]
92. Li, N.; Shepperd, M.; Guo, Y. A systematic review of unsupervised learning techniques for software defect prediction. *Inf. Softw. Technol.* **2020**, *122*, 106287. [CrossRef]
93. Gao, X.; Zhang, T. Unsupervised learning to detect loops using deep neural networks for visual SLAM system. *Auton. Robot.* **2017**, *41*, 1–18. [CrossRef]
94. López-Rubio, E.; Palomo, E.J.; Ortega-Zamorano, F. Unsupervised learning by cluster quality optimization. *Inf. Sci.* **2018**, *436–437*, 31–55. [CrossRef]
95. Brivio, S.; Ly, D.R.B.; Vianello, E.; Spiga, S. Non-linear memristive synaptic dynamics for efficient unsupervised earning in spiking neural networks. *Front. Neurosci.* **2021**, *15*, 580909. [CrossRef] [PubMed]
96. Ma, S.; Guo, W.; Song, R.; Liu, Y. Unsupervised learning based coordinated multi-task allocation for unmanned surface vehicles. *Neurocomputing* **2021**, *420*, 227–245. [CrossRef]
97. Casolla, G.; Cuomo, S.; Di Cola, V.S.; Piccialli, F. Exploring Unsupervised Learning Techniques for the Internet of Things. *IEEE Trans. Ind. Inform.* **2019**, *16*, 2621–2628. [CrossRef]
98. Wei, Y.; Thompson, M.P.; Belval, E.J.; Calkin, D.E.; Bayham, J. Understand daily fire suppression resource ordering and assignment patterns by unsupervised learning. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 2. [CrossRef]
99. Eskimez, S.E.; Duan, Z.; Heinzelman, W. Unsupervised learning approach to feature analysis for automatic speech emotion recognition. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, AB, Canada, 15–20 April 2018; pp. 5099–5103.
100. Yan, K.; Huang, J.; Shen, W.; Ji, Z. Unsupervised learning for fault detection and diagnosis of air handling units. *Energy Build.* **2020**, *210*, 109689. [CrossRef]
101. Sinaga, K.P.; Yang, M.S. Unsupervised K-means clustering algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [CrossRef]
102. Bhowmik, A.; De, D. mTrust: Call behavioral trust predictive analytics using unsupervised learning in mobile cloud computing. *Wirel. Pers. Commun.* **2021**, *117*, 483–501. [CrossRef]
103. Pawar, A.; Mago, V. Challenging the boundaries of unsupervised learning for semantic similarity. *IEEE Access* **2019**, *7*, 16291–16308. [CrossRef]
104. Suominen, A.; Toivanen, H.; Seppänen, M. Firms' knowledge profiles: Mapping patent data with unsupervised learning. *Technol. Forecast. Soc. Chang.* **2017**, *115*, 131–142. [CrossRef]
105. Van Engelen, J.E.; Hoos, H.H. A survey on semi-supervised learning. *Mach. Learn.* **2020**, *109*, 373–440. [CrossRef]
106. Laine, S.; Aila, T. Temporal ensembling for semi-supervised learning. In Proceedings of the International Conference on Learning Representations, Toulon, France, 14–26 April 2017; pp. 1–13.
107. Wu, H.; Prasad, S. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. *IEEE Trans. Image Process.* **2017**, *27*, 1259–1270. [CrossRef]
108. Chen, C.; Liu, Y.; Kumar, M.; Qin, J.; Ren, Y. Energy consumption modelling using deep learning embedded semi-supervised learning. *Comput. Ind. Eng.* **2019**, *135*, 757–765. [CrossRef]
109. Wu, B.; Meng, D.; Zhao, H. Semi-supervised learning for seismic impedance inversion using generative adversarial networks. *Remote Sens.* **2021**, *13*, 909. [CrossRef]
110. Liu, Z.; Huang, S.; Jin, W.; Mu, Y. Broad learning system for semi-supervised learning. *Neurocomputing* **2021**, *444*, 38–47. [CrossRef]
111. Liu, Z.; Lai, Z.; Ou, W.; Zhang, K.; Zheng, R. Structured optimal graph based sparse feature extraction for semi-supervised learning. *Signal. Process.* **2020**, *170*, 107456. [CrossRef]
112. Yang, D.; Xu, Z.; Li, W.; Myronenko, A.; Roth, H.R.; Harmon, S.; Xu, S.; Turkbey, B.; Turkbey, E.; Wang, X.; et al. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. *Med. Image Anal.* **2021**, *70*, 101992. [CrossRef] [PubMed]
113. Bahrami, S.; Dornaika, F.; Bosaghzadeh, A. Joint auto-weighted graph fusion and scalable semi-supervised learning. *Inf. Fusion* **2021**, *66*, 213–228. [CrossRef]
114. Zaman, S.M.K.; Liang, X. An effective induction motor fault diagnosis approach using graph-based semi-supervised learning. *IEEE Access* **2021**, *9*, 7471–7482. [CrossRef]
115. Berthelot, D.; Carlini, N.; Goodfellow, I.; Papernot, N.; Oliver, A.; Raffel, C. MixMatch: A holistic approach to semi-supervised learning. In Proceedings of the Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 1–14.
116. Han, C.H.; Kim, M.; Kwak, J.T. Semi-supervised learning for an improved diagnosis of COVID-19 in CT images. *PLoS ONE* **2021**, *16*, e0249450. [CrossRef]
117. Yu, K.; Lin, T.R.; Ma, H.; Li, X.; Li, X. A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning. *Mech. Syst. Signal. Process.* **2021**, *146*, 107043. [CrossRef]

118. Guo, J.; Wang, Q.; Li, Y. Semi-supervised learning based on convolutional neural network and uncertainty filter for façade defects classification. *Comput. Civ. Infrastruct. Eng.* **2021**, *36*, 302–317. [CrossRef]
119. Livieris, I.E.; Drakopoulou, K.; Tampakas, V.T.; Mikropoulos, T.A.; Pintelas, P. Predicting secondary school students' performance utilizing a semi-supervised learning approach. *J. Educ. Comput. Res.* **2019**, *57*, 448–470. [CrossRef]
120. Kejani, M.T.; Dornaika, F.; Talebi, H. Graph convolution networks with manifold regularization for semi-supervised learning. *Neural Netw.* **2020**, *127*, 160–167. [CrossRef] [PubMed]
121. Gan, H.; Li, Z.; Wu, W.; Luo, Z.; Huang, R. Safety-aware graph-based semi-supervised learning. *Expert Syst. Appl.* **2018**, *107*, 243–254. [CrossRef]
122. Yuan, Y.; Li, X.; Wang, Q.; Nie, F. A semi-supervised learning algorithm via adaptive laplacian graph. *Neurocomputing* **2021**, *426*, 162–173. [CrossRef]
123. Li, Y.F.; Liang, D.M. Safe semi-supervised learning: A brief introduction. *Front. Comput. Sci.* **2019**, *13*, 669–676. [CrossRef]
124. Chong, Y.; Ding, Y.; Yan, Q.; Pan, S. Graph-based semi-supervised learning: A review. *Neurocomputing* **2020**, *408*, 216–230. [CrossRef]
125. Gordon, J.; Hernández-Lobato, J.M. Combining deep generative and discriminative models for bayesian semi-supervised learning. *Pattern Recognit.* **2020**, *100*, 107156. [CrossRef]
126. Miyato, T.; Maeda, S.I.; Koyama, M.; Ishii, S. Virtual adversarial training: A regularization method for supervised and semi-supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 1979–1993. [CrossRef]
127. Rathore, S.; Park, J.H. Semi-supervised learning based distributed attack detection framework for IoT. *Appl. Soft Comput. J.* **2018**, *72*, 79–89. [CrossRef]
128. Dunlop, M.M.; Slepčev, D.; Stuart, A.M.; Thorpe, M. Large data and zero noise limits of graph-based semi-supervised learning algorithms. *Appl. Comput. Harmon. Anal.* **2020**, *49*, 655–697. [CrossRef]
129. Ashfaq, R.A.R.; Wang, X.Z.; Huang, J.Z.; Abbas, H.; He, Y.L. Fuzziness based semi-supervised learning approach for intrusion detection system. *Inf. Sci.* **2017**, *378*, 484–497. [CrossRef]
130. Hussain, A.; Cambria, E. Semi-supervised learning for big social data analysis. *Neurocomputing* **2018**, *275*, 1662–1673. [CrossRef]
131. Zhang, S.; Huang, K.; Zhu, J.; Liu, Y. Manifold adversarial training for supervised and semi-supervised learning. *Neural Netw.* **2021**, *140*, 282–293. [CrossRef]
132. Gan, H.; Guo, L.; Xia, S.; Wang, T. A hybrid safe semi-supervised learning method. *Expert Syst. Appl.* **2020**, *149*, 113295. [CrossRef]
133. Zhao, H.; Zheng, J.; Deng, W.; Song, Y. Semi-supervised broad learning system based on manifold regularization and broad network. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 983–994. [CrossRef]
134. Yan, K.; Zhong, C.; Ji, Z.; Huang, J. Semi-supervised learning for early detection and diagnosis of various air handling unit faults. *Energy Build.* **2018**, *181*, 75–83. [CrossRef]
135. Vinyals, O.; Babuschkin, I.; Czarnecki, W.M.; Mathieu, M.; Dudzik, A.; Chung, J.; Choi, D.H.; Powell, R.; Ewalds, T.; Georgiev, P.; et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **2019**, *575*, 350–354. [CrossRef] [PubMed]
136. Zeng, N.; Li, H.; Wang, Z.; Liu, W.; Liu, S.; Alsaadi, F.E.; Liu, X. Deep-reinforcement-learning-based images segmentation for quantitative analysis of gold immunochromatographic strip. *Neurocomputing* **2021**, *425*, 173–180. [CrossRef]
137. Ning, Z.; Dong, P.; Wang, X.; Rodrigues, J.J.P.C.; Xia, F. Deep reinforcement learning for vehicular edge computing: An intelligent offloading system. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–24. [CrossRef]
138. Everett, M.; Chen, Y.F.; How, J.P. Collision avoidance in pedestrian-rich environments with deep reinforcement learning. *IEEE Access* **2021**, *9*, 10357–10377. [CrossRef]
139. Kiran, B.R.; Sobh, I.; Talpaert, V.; Mannion, P.; Sallab, A.A.A.; Yogamani, S.; Perez, P. Deep reinforcement learning for autonomous driving: A survey. *IEEE Trans. Intell. Transp. Syst.* **2021**, 1–18. [CrossRef]
140. Hodge, V.J.; Hawkins, R.; Alexander, R. Deep reinforcement learning for drone navigation using sensor data. *Neural Comput. Appl.* **2021**, *33*, 2015–2033. [CrossRef]
141. Viquerat, J.; Rabault, J.; Kuhnle, A.; Ghraieb, H.; Larcher, A.; Hachem, E. Direct shape optimization through deep reinforcement learning. *J. Comput. Phys.* **2021**, *428*, 110080. [CrossRef]
142. Tan, T.; Bao, F.; Deng, Y.; Jin, A.; Dai, Q.; Wang, J. Cooperative deep reinforcement learning for large-scale traffic grid signal control. *IEEE Trans. Cybern.* **2019**, *50*, 2687–2700. [CrossRef]
143. Zhang, W.; He, X.; Lu, W.; Qiao, H.; Li, Y. Feature aggregation with reinforcement learning for video-based person re-identification. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3847–3852. [CrossRef]
144. Zhong, B.; Bai, B.; Li, J.; Zhang, Y.; Fu, Y. Hierarchical tracking by reinforcement learning-based searching and coarse-to-fine verifying. *IEEE Trans. Image Process.* **2018**, *28*, 2331–2341. [CrossRef]
145. Xu, M.; Song, Y.; Wang, J.; Qiao, M.; Huo, L.; Wang, Z. Predicting head movement in panoramic video: A deep reinforcement learning approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 2693–2708. [CrossRef]
146. Carta, S.; Corriga, A.; Ferreira, A.; Podda, A.S.; Recupero, D.R. A multi-layer and multi-ensemble stock trader using deep learning and deep reinforcement learning. *Appl. Intell.* **2021**, *51*, 889–905. [CrossRef]
147. Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science* **2018**, *362*, 1140–1144. [CrossRef]

148. Huang, L.; Fu, M.; Li, F.; Qu, H.; Liu, Y.; Chen, W. A deep reinforcement learning based long-term recommender system. *Knowl. Based Syst.* **2021**, *213*, 106706. [CrossRef]
149. Charpentier, A.; Élie, R.; Remlinger, C. Reinforcement learning in economics and finance. *Comput. Econ.* **2021**, 1–38. [CrossRef]
150. Hwangbo, J.; Sa, I.; Siegwart, R.; Hutter, M. Control of a quadrotor with reinforcement learning. *IEEE Robot. Autom. Lett.* **2017**, *2*, 2096–2103. [CrossRef]
151. Wang, Z.; Hong, T. Reinforcement learning for building controls: The opportunities and challenges. *Appl. Energy* **2020**, *269*, 115036. [CrossRef]
152. Bai, W.; Zhou, Q.; Li, T.; Li, H. Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation. *IEEE Trans. Cybern.* **2020**, *50*, 3433–3443. [CrossRef]
153. Popova, M.; Isayev, O.; Tropsha, A. Deep reinforcement learning for de novo drug design. *Sci. Adv.* **2018**, *4*, eaap7885. [CrossRef]
154. Perera, A.T.D.; Kamalaruban, P. Applications of reinforcement learning in energy systems. *Renew. Sustain. Energy Rev.* **2021**, *137*, 110618. [CrossRef]
155. Bai, W.; Li, T.; Tong, S. NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems. *IEEE Trans. Cybern.* **2020**, *50*, 4573–4584. [CrossRef] [PubMed]
156. Akalin, N.; Loutfi, A. Reinforcement learning approaches in social robotics. *Sensors* **2021**, *21*, 1292. [CrossRef] [PubMed]
157. Li, R.; Zhao, Z.; Sun, Q.; Chih-Lin, I.; Yang, C.; Chen, X.; Zhao, M.; Zhang, H. Deep reinforcement learning for resource Managment in network slicing. *IEEE Access* **2018**, *6*, 74429–74441. [CrossRef]
158. Arora, S.; Doshi, P. A survey of inverse reinforcement learning: Challenges, methods and progress. *Artif. Intell.* **2021**, *297*, 103500. [CrossRef]
159. Kuhnle, A.; Kaiser, J.P.; Theiß, F.; Stricker, N.; Lanza, G. Designing an adaptive production control system using reinforcement learning. *J. Intell. Manuf.* **2021**, *32*, 855–876. [CrossRef]
160. Zou, F.; Yen, G.G.; Tang, L.; Wang, C. A reinforcement learning approach for dynamic multi-objective optimization. *Inf. Sci.* **2021**, *546*, 815–834. [CrossRef]
161. Liu, Y.J.; Li, S.; Tong, S.; Chen, C.L.P. Adaptive reinforcement learning control based on neural approximation for nonlinear discrete-time systems with unknown nonaffine dead-zone input. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *30*, 295–305. [CrossRef]
162. Mahmud, M.; Kaiser, M.S.; Hussain, A.; Vassanelli, S. Applications of deep learning and reinforcement learning to biological data. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2063–2079. [CrossRef] [PubMed]
163. Wang, X.; Gu, Y.; Cheng, Y.; Liu, A.; Chen, C.L.P. Approximate policy-based accelerated deep reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 1820–1830. [CrossRef]
164. Zhang, H.; Jiang, H.; Luo, Y.; Xiao, G. Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method. *IEEE Trans. Ind. Electron.* **2016**, *64*, 4091–4100. [CrossRef]
165. Mouratidis, K.; Papagiannakis, A. COVID-19, internet, and mobility: The rise of telework, telehealth, e-learning, and e-shopping. *Sustain. Cities Soc.* **2021**, *74*, 103182. [CrossRef] [PubMed]
166. Aslam, S.M.; Jilani, A.K.; Sultana, J.; Almutairi, L. Feature evaluation of emerging e-learning systems using machine learning: An extensive survey. *IEEE Access* **2021**, *9*, 69573–69587. [CrossRef]
167. Tang, K.-Y.; Chang, C.-Y.; Hwang, G.-J. Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interact. Learn. Environ.* **2021**, 1–19. [CrossRef]
168. Rasheed, F.; Wahid, A. Learning style detection in e-learning systems using machine learning techniques. *Expert Syst. Appl.* **2021**, *174*, 114774. [CrossRef]
169. Semerci, Y.C.; Goularas, D. Evaluation of students' flow state in an e-learning environment through activity and performance using deep learning techniques. *J. Educ. Comput. Res.* **2020**, *59*, 960–987. [CrossRef]
170. Bhaskaran, S.; Marappan, R. Design and analysis of an efficient machine learning based hybrid recommendation system with enhanced density-based spatial clustering for digital e-learning applications. *Complex. Intell. Syst.* **2021**, 1–17. [CrossRef]
171. Bhaskaran, S.; Marappan, R.; Santhi, B. Design and analysis of a cluster-based intelligent hybrid recommendation system for e-learning applications. *Mathematics* **2021**, *9*, 197. [CrossRef]
172. Bhardwaj, P.; Gupta, P.K.; Panwar, H.; Siddiqui, M.K.; Morales-Menendez, R.; Bhaik, A. Application of deep learning on student engagement in e-learning environments. *Comput. Electr. Eng.* **2021**, *93*, 107277. [CrossRef]
173. Nandi, A.; Xhafa, F.; Subirats, L.; Fort, S. Real-time emotion classification using EEG data stream in e-learning contexts. *Sensors* **2021**, *21*, 1589. [CrossRef]
174. Vuković, I.; Kuk, K.; Čisar, P.; Banđur, M.; Banđur, Đ.; Milić, N.; Popović, B. Multi-agent system observer: Intelligent support for engaged e-learning. *Electronics* **2021**, *10*, 1370. [CrossRef]
175. Cvitić, I.; Peraković, D.; Periša, M.; Jurcut, A.D. Methodology for detecting cyber intrusions in e-learning systems during COVID-19 pandemic. *Mob. Netw. Appl.* **2021**, *2021*, 1–12. [CrossRef]
176. Dong, H.; Tsai, S.B. An empirical study on application of machine learning and neural network in english learning. *Math. Probl. Eng.* **2021**, *2021*, 8444858. [CrossRef]
177. Ho, I.M.K.; Cheong, K.Y.; Weldon, A. Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques. *PLoS ONE* **2021**, *16*, e0249423. [CrossRef] [PubMed]

178. Shim, T.E.; Lee, S.Y. College students' experience of emergency remote teaching due to COVID-19. *Child. Youth Serv. Rev.* **2020**, *119*, 1–7. [CrossRef]
179. Alqurshi, A. Investigating the impact of COVID-19 lockdown on pharmaceutical education in Saudi Arabia—A call for a remote teaching contingency strategy. *Saudi Pharm. J.* **2020**, *28*, 1075–1083. [CrossRef]
180. EDUCAUSE. EDUCAUSE DIY Survey Kit: Remote Work and Learning Experiences. Available online: https://er.educause.edu/blogs/2020/4/educause-diy-survey-kit-remote-work-and-learning-experiences/ (accessed on 20 September 2021).
181. Al-Maroof, R.S.; Alhumaid, K.; Akour, I.; Salloum, S. Factors that affect e-learning platforms after the spread of COVID-19: Post acceptance study. *Data* **2021**, *6*, 49. [CrossRef]