Machine Learning: Past, Present and Future

Akshita Tyagi[1], Swetta Kukreja[2], Meghna Manoj Nair[1], Amit Kumar Tyagi[1]
School of Computer Science and Engineering,Vellore Institute of Technology
Chennai, 600127, Tamil Nādu, India
Terna Engineering College, University of Mumbai
akshitatyagi222@gmail.com, swettakukreja@gmail.com, mnairmeghna@gmail.com, amitkrtyagi025@gmail.com

Abstract:

Recent well-publicized triumphs in Machine Learning have solved issues that were previously believed to take decades to solve, rekindling interest in the fields of Artificial Intelligence (AI) And Machine Learning(ML). For instance, deep learning's recent success and quick commercialization have propelled technological advancements in a variety of industries, such as computer vision, speech recognition, gaming, and machine translation. These accomplishments offer fresh chances to advance Nondestructive Assessment (NDE) methods. The foundational ideas of MLare reviewed, along with the relationship they have to statistics. Then, we go over previous ML for NDE applications and techniques. We also discuss the problems that this discipline is still trying to overcome, namely the lack of trustworthy training data. We then go into recent ML for NDE research that aims to address these issues. Finally, we discuss how current ML developments like deep learning and transfer learning have the potential to fundamentally alter how we develop future NDE solutions.

Keywords: Artificial Intelligence, Internet of Things; Healthcare; Smart Cities; Smart Grid; Supply Chain Management; Machine Learning

DOI Number:10.14704/nq.2022.20.8.NQ44468 **NeuroQuantology 2022; 20(8): 4333-4357**

1. Introduction

Recent high-profile successes in several sectors that have made media headlines and gained widespread recognition are the result of significant discoveries in artificial intelligence (AI) and machine learning (ML). For instance, improvements in computer vision techniques have produced techniques for driverless vehicles, facial identification, and medical imaging. Technologies for mobile assistants and home automation have been made possible by advancements in voice recognition. The success of the computer player in video games for amusement and Go is a result of advancements in AI for gaming. Auto-translation features for websites and video conferences have been altered by new Machine Learning technologies. There is a lot of interest in Machine Learning and how AI can affect other industries as a result of these almost daily triumphs in the press. Governments, business, and academia are currently investing strategically in AI and ML. The necessity for quantitative NDE and the (renewed) interest in AI/ML raise the concerns of what role AI and ML have played in NDE and what ML and NDE could entail in the future. In a recent article [1], the use of ML for NDE is highlighted, along with applications for automatic interpretation of NDE signals. This work, which is divided into various sections, expands on the issue of how ML and NDE interact. First, a summary of Machine Learning ideas is given. An explanation of previous and contemporary uses of ML in NDE follows here.

The following are the types of models of Machine Learning,

a) **Supervised learning:** It is a process in which input and desired output is provided to the machine. Input and output data are labelled for classification purpose. Target variable is predicted from the given set of predictors. The supervised learning algorithms include, Random Forest (RF), k Nearest Neighbour (kNN), Decision Tree, Regression, Logistic Regression, boosting algorithm.
b) **Unsupervised learning**: This type of learning is used to draw inferences that come from datasets, which consist of input data without having labelled responses as if no supervisors to guide. The unsupervised learning algorithms include, Apriori algorithm, K-means, Adaptive Resonance Theory, Self-Organising Map (SOM Model).
c) **Reinforcement learning**: Reinforcement learning is same as supervised learning that trains algorithms using a reward and punishment technique. A reinforcement learning algorithm learns machine by interacting with its environment. Supervisor will give reward at the time of correct action and punishment at every wrong action. Here, machine is trained and therefore it will take its own decision. Reinforcement learning learns from the previous experiences. To make accurate decision, it will capture the best possible knowledge. Markov Decision Process is an example of Reinforcement learning

Organization of the work:

Section 2 discusses about the brief overview of Machine Learning and its types and forms. Following this, in section 3, the paper discusses about the how Machine Learning has evolved and transformed over the years and made automation an integral part of human lives. Section 4 throws light on the important Machine Learning applications and use cases in the past and present and also elucidates the possible inventions that'd significantly benefit from Machine Learning techniques in the future. While Section 5 is all about a thorough literature review on the efforts and contributions to Machine Learning over the years through supervised and unsupervised methods, section 6 covers the issues and contradictions in applications that incorporate Machine Learning. Section 7 focuses on the challenges and hurdles that one may encounter while deriving the use cases of Machine Learning applications and section 8 emphasizes on the future advancements with Machine Learning. Section 9 highlights the future scope and possible opportunities to work with Machine Learning and section 10 concludes the paper.

2. Machine Learning: A Brief Overview

"Automated methods for data analysis" or "methods for automatically detecting patterns and forecasting future data" are some of the definitions given for Machine Learning [1, 2]. The term "Machine Learning" refers to a technique that uses a computer to make decisions and generate output based on prior knowledge (i.e., data). The approaches and algorithms are often designed such that when more data is collected, they become more effective at doing a certain task. For instance, a Machine Learning model ought to generate more precise forecasts as theincreasing volume of training data (assume the training data well represents the desired scenario). Numerous statistical models also exhibit these qualities. Additionally, a lot of Machine Learning techniques have their roots in statistics. One theory holds that Machine Learning is just statistics repackaged with an emphasis on prediction. Additionally, Machine Learning places a lot greater emphasis on actual performance than it does on theory and justifications. In conversations on Machine Learning, artificial intelligence is frequently brought up. Fig. 1 shows how closely related the sciences of statistics, Machine Learning, and artificial intelligence are. Systems that are implemented on computers and carry out any work in a way that is viewed as intelligent fall under the category of artificial intelligence. The decision-making tools in AI systems heavily rely on Machine Learning techniques, therefore ML is a subset of AI.

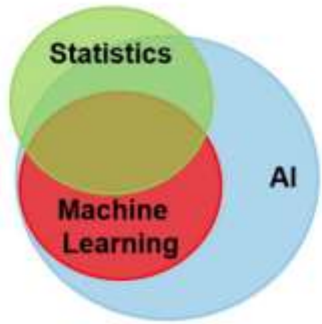

Fig. 1. Statistics, Machine Learning, and AI Machine Learning (ML).

Not that ML is a subset of artificial intelligence (AI), and both areas heavily rely on statistics. As a result, there is substantial overlap between these three subjects. Fig. 2 shown below shows the workflow followed when it comes to the process of Machine Learning.

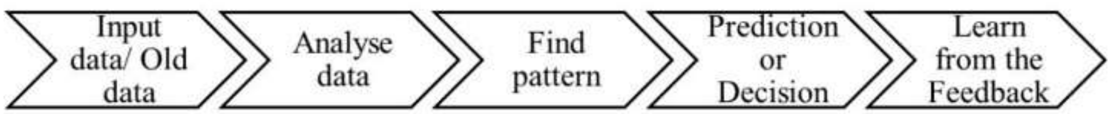

Fig. 2. Process of Machine Learning

2.1 Types of Machine Learning

Depending on the particular job of interest, Machine Learning is separated into groups. supervised learning and unsupervised learning are the two main kinds. In supervised learning, a collection of training data with inputs, xn, and outputs, yn, is given. The objective is to train a model to predict the matching outputs, ym, for some fresh set of inputs, xm. Finding patterns or groupings within the data that are helpful for its comprehension or explanation is the aim of unsupervised learning, which uses data that just comprises the inputs, xi. The inputs are frequently referred to as "features" in Machine Learning terminology.

2.1.1 Supervised Learning

Depending on the kind of input the model will predict supervised learning is frequently classified into several categories. This process is known as classification if the model is attempting to predict a categorical output, such as membership in a certain class [2]. Fig. 3(a), which depicts the data as various points in the input space and the related outputs as classes for the colors of the points ('red' and 'purple'), exemplifies this. Based on existing training data, a Machine Learning classifier will learn the line that most effectively divides the two groups. The color of any new input or feature is then determined by which side of the line the point is located. Regression is a different kind of supervised learning when the model aims to predict a continuous output, such as a real (or occasionally complex) number. Fig. 3(b), which plots the training data as black dots with outputs vs. inputs, serves as an illustration of this. A regression model built using training data was used to predict the dark, red line that runs across the spots. A confidence interval for those forecasts is shown by the lighter pink bars encircling the data. Regression models come in a wide variety, including support vector regression, neural networks, kriging, linear models, and more [1, 2]. Regression occurs often in high-dimensional, hard to visualize spaces in practice.

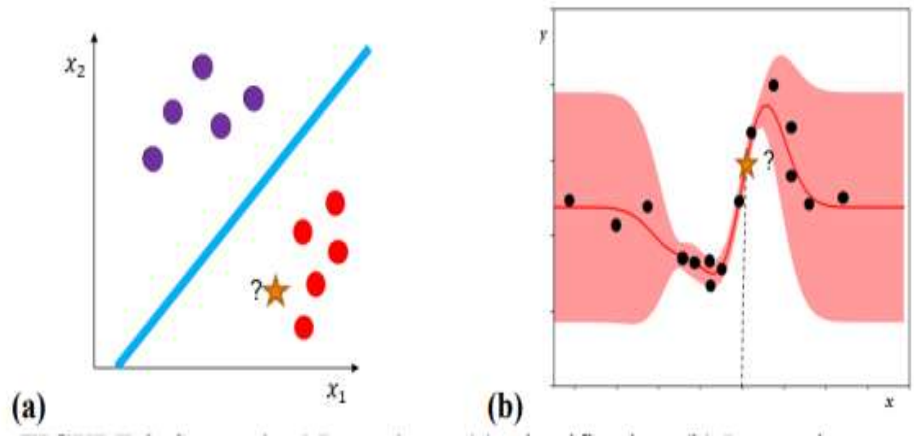

Fig. 3. Classification and regression under supervision

Based on classification and regression, there are a number of supervised learning algorithms integrated with the process of Machine Learning. One of the outrageously used techniques is that of Neural Networks which practically trainsthe models on data by imitating the logical connectivity of human brain nerves. It's basically a mapping function which keeps adjusting the learned concepts on the basis of the loss function and gradient descent. One of the other techniques is the Naïve Bayes algorithm. It's a classic and typical classification strategy that relies on the concept of class conditional independence which extends from the Bayes theorem which implies that the occurrence of an event does not necessarily impact that of another event. Multinomial Bayes, Bernoulli Naïve Bayes, and Gaussian Naïve Bayes are the types of Naïve Bayes classifiers. Next up is linear regression which is a technique used to detect the relation between variables in which one of them is dependent while the others are independent. Logistic regression, on the other hand, is similar to linear regression, except for the fact that it is used when the outcomes are binary in nature. One of the other popular supervised Machine Learning algorithms is support vector machine which serves dual purposes of solving classification as well as regression related issues by constructing a hyperplane where the two data points are distanced at its maximum. K Nearest Neighbors is one of the other non-parametric algorithms which distinguishes between data points on the basis of the proximity and closeness to the other data points and focuses on ensuring that data points that are similar to each other are placed closely. It usually makes use of the Euclidean Distance algorithm for distance calculation. Last, but not the least, is the random forest algorithm which is one of the most flexible algorithms useful for classification as well as regression which refer to the usage of uncorrelated decision trees and reduced variance. One of the other algorithms is the decision tree algorithm which depends on information gain and entropy for predicting results based on its training. C4.5 is one of the major decision tree techniques works well with discrete and continuous data and can also deal with the problem of incomplete information.

2.1.2 Unsupervised Learning

Unsupervised learning uses algorithms to analyze data without labels (which are required for classification) and without producing any results (needed for regression). Finding and describing structure in the data is the intended outcome instead. Only a handful of the numerous methods are covered here and are shown in Fig. 4. It shows examples of each algorithm type using the same dataset in feature space. Data is divided into groups using clustering algorithms, and each group is given a class label. In order to measure the uncertainty in the data, density estimation techniques estimate the joint probability density function for the data. Finding transformations of the data that can project them into a lower-dimensional, more effective representation of the data is the goal of dimensionality reduction techniques.

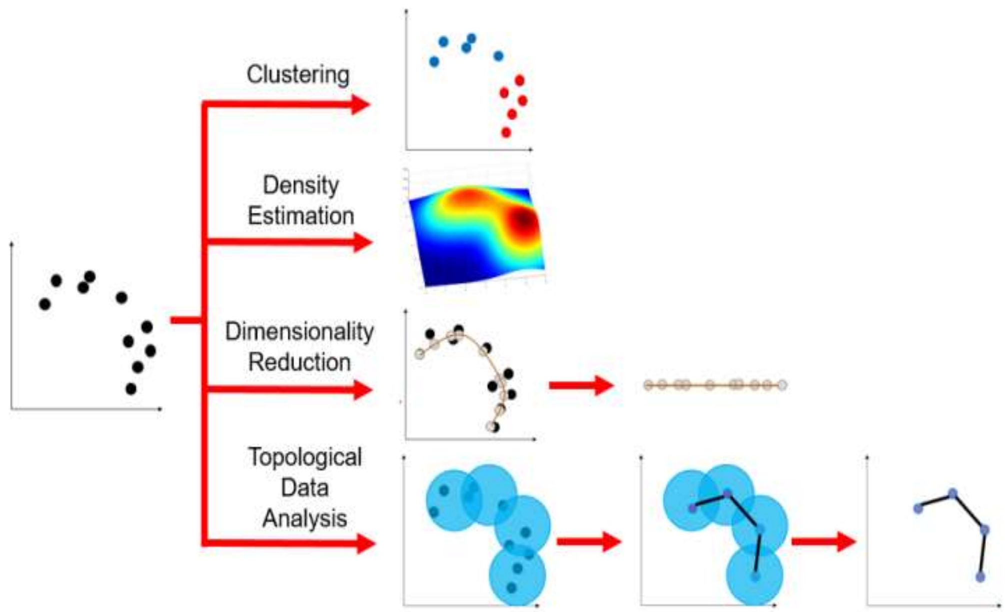

Fig. 4. Clustering, density estimation, dimensionality reduction, and topological data analysis

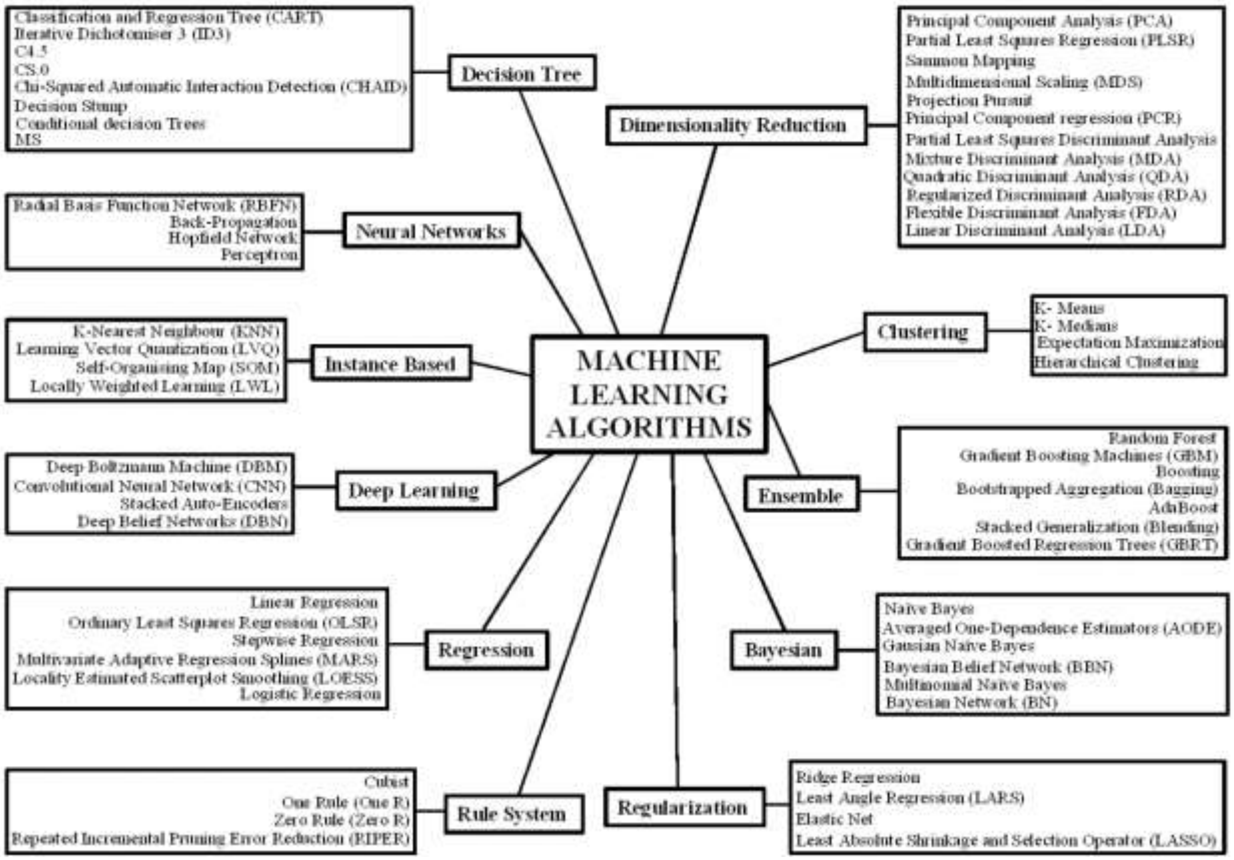

Fig. 5. Algorithms of Machine Learning Techniques

As observed in the previous section, classification and regression are the two important techniques in supervised Machine Learning. Similarly, in the case of unsupervised learning, clustering and association are the main phenomena. Clustering is all about grouping and gathering different objects into segments based on the similarity criteria with the aim of finding common grounds between the elements of a cluster. Association, on the contrary, is mainly for determining the relationship between variables in massive data collections and is strongly utilized in the marketing sector. Fig. 5 displays the various types of Machine Learning algorithms and the different versions available in each of them. Let's consider the Machine Learning algorithms mentioned in the above diagram one by one. The very first and most commonly used algorithm is that of Decision Trees which is a supervised learning technique that is often utilized for classification and regression related applications using tree-based structure at the grass root level. The decision node and leaf nodes together help in bringing out the final outcome based on the conditional factors. One commonly used decision tree algorithm is the CART algorithm. Coming to the next algorithm in the diagram, it is neural networks. This algorithm imitates the architecture of neural networks of human brains in terns of learning and training the model followed by accurately making predictions. The input layer, hidden layers, and the output layer are the layers which constitute the Machine Learning model. The training of the model enhances the strength and bond between the neurons thus leading to accurate predictions. Back propagation is one of the majorly used neural network algorithm which helps in efficient calculation of derivates and mathematical problems quickly. Next up is the instance-based algorithm in Machine Learning and systems which make use of this algorithm induce the technique of learning and training by heart which is further generalized to expand and familiarize itself with new instances. K Nearest Neighbor (K-NN) algorithm is an example of the instance-based technique which categorizes fata point based on certain similarity criteria. Moving on to Deep Learning techniques. Deep Learning is a subset of Machine Learning which focuses on the utilization of artificial neural networks. The infamous Convolutional Neural Network (CNN) algorithm in deep learning is extremely useful and extensively put to use in the classification, categorization, and differentiation of images for a variety of use cases. Regression and clustering are some of the other techniques mentioned in the diagram which have already been addressed in various parts of this paper. One of the other algorithms mentioned in the above diagram is that of dimensionality reduction which involves the process of cutting down on the number of variables randomly taken into consideration with the help of principal variables which can be mainly divided as feature selection/extraction. Principal Component Analysis (PCA) is an example for the dimensionality reduction algorithm which is basically a statistical computation that converts the observations of extracted features into linear features which may not be necessarily correlated. Then comes the ensemble technique which ultimately is a combination of a number of Machine Learning models with the main aim of improving the obtained results. Random forest algorithm is an ensemble of many decision trees and thus classifies input on various subsets by considering the mean outcome for increased accuracy. This gives a basic idea of the plethora of algorithms available in Machine Learning which are used in various relevant fields and aspects of automation and model implementation.

3. Evolution of Machine Learning

Artificial Intelligence (AI) applications such as Machine Learning have some remarkable skills. Software that can learn unsupervised can be created using a Machine Learning algorithm. Through the addition of historical data, the system appears to be able to "learn" and improve its predictive abilities without being explicitly programmed. Although they are frequently used synonymously, artificial intelligence and Machine Learning are not the same. One of the most active fields and a means of achieving AI is Machine Learning. There are a few reasons, including the ones listed below but not exclusively, for why Machine Learning is so good today.

- The huge data explosion
- The need for new customers and revenue sources in this industry is dwindling.
- Machine Learning algorithms have improved.
- creation of a super-efficient device with a large capacity and quick processing speed
- Ability to store

Today's machines are learning and carrying out tasks that, in the past, could only be completed by people, such as making better judgement calls, making decisions, playing games, etc. Machines can now analyze patterns, read them, and store knowledge for later use, making this conceivable. The main challenge nowadays is to locate resources that are qualified enough to differentiate their learning from university and PhD textbooks in actual business settings rather than only debating others on social media.

Business teams, managers, and executives should all have a working understanding of Machine Learning as part of the organization's culture. There must be ongoing programs and road shows for them if we want to attain this as a culture.

3.1 The journey of automation through Machine Learning

Between the data's actual source and the learning platform, which may be on the cloud, different algorithms and services are managed in a Machine Learning environment or setup. Modern Machine Learning always uses centralized infrastructure to complete its tasks. Even said, there are success stories where it also functions remarkably well on a distributed architecture. These techniques might not be the most effective, but for the time being, they are effective.

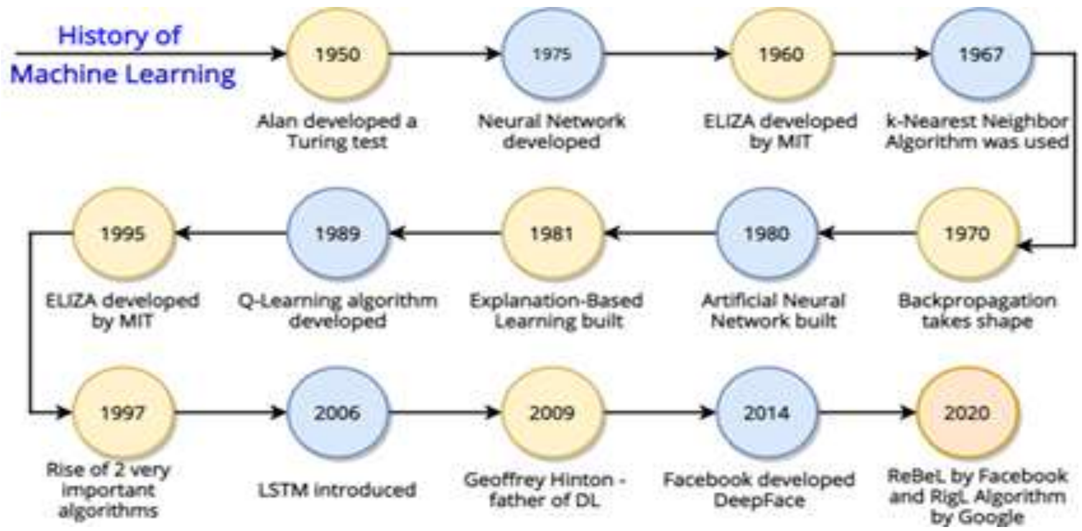

Fig. 6. Evolution of Machine Learning Over the Years

Turing machines were the beginning of the progression that led to today's extremely intelligent robots. How far Machine Learning has advanced from its inception is tough to gauge and quantify, yet the results are still obvious and observable. Despite the fact that research in the topic has been ongoing for at least 50 years, the phrase "Machine Learning" has recently gained a lot of popularity with developers and business owners alike. Fig. 6 describes the evolutionary timeline of Machine Learning over the years.

3.2 Machine Learning in Today's Era

The significance and importance of technology in almost everything we do in our lives has been a huge factor of boost for Machine Learning and Artificial Intelligence. The main benefit derived from the usage of Machine Learning is its ability to ease human chores and maintain a sense of efficiency, automation, and reliability. The first traces of Machine Learning dates back down to 1950 when Alan Turing proposed the Turing Test which turned out to be the classic test for detecting the intelligence of machines. From then on, there has been a stark increase in the development of algorithms, programs and software that revolve around Machine Learning which spanned across a large area of use cases. Tic Tac Toes game programs, checkers program etc. were some of the popular games during the 1960s. By 1970, there was an observed phenomena of AI Winter wherein the British government cut off funding for research in Artificial Intelligence and Machine Learning among the universities. By 1990's, IBM came up with Deep Blue which was a chess playing computer program. By 2010, Machine Learning was flourishing worldwide and was extensively being used by companies like Netflix, Kinect, Kaggle, and for ImageNet classification. A few years down the lane, some remarkable inventions such as Facebook's DeepFace (for efficient identification of humans from photos) and Google's Sibyl (for predicting user recommendations) have been curated. Today, we are at a stage where Machine Learning can be seen in healthcare, transport, business, education, and almost every other aspect of life. Fig. 7 shown below summarizes the history and progress of deep learning throughout the years.

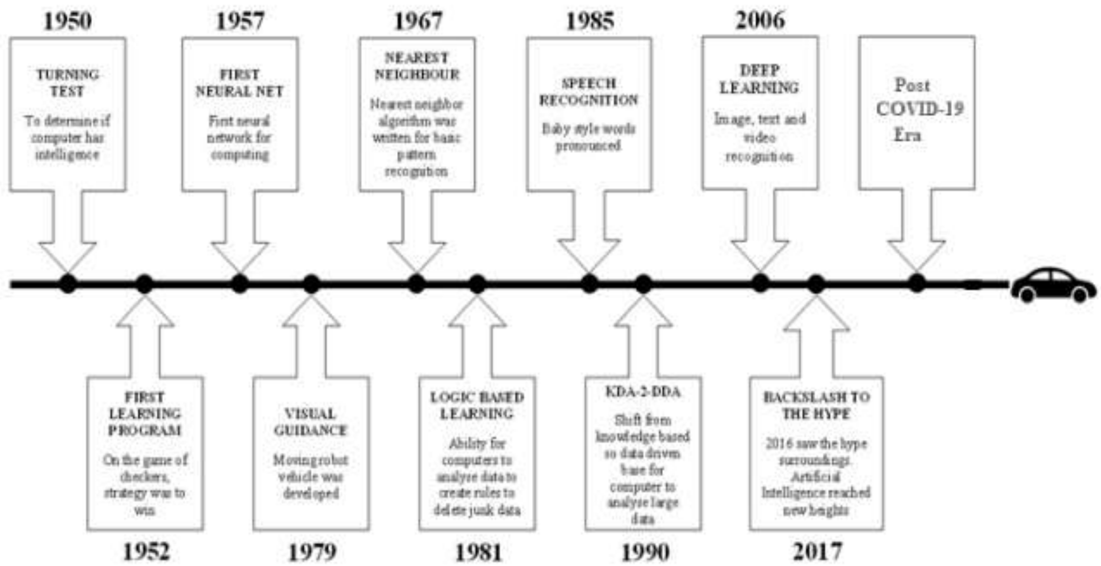

Fig. 7. Progress of Deep Learning

3.3 Machine Learning in the Future

With Machine Learning and Artificial Intelligence blooming in almost every other industry these days, it is highly predictable that the future applications and implementations in the field of technology would heavily rely on Machine Learning techniques, strategies, and algorithms. The fact that it has potential for competitive advantages and benefits makes it highly economical for incorporation into upcoming startups and business plans with the aim of getting more leverage. The use of Machine Learning in various different fields including those of medicine, education, transportation, military, etc. would only further expand in the future as people are able to derive large benefits from the use of Machine Learning to expedite their needs and demands. In the transportation sector, the very famous Tesla owned by Elon Musk has made use of extensive Machine Learning based algorithms for object detection to implement self-driving cars. A few years down the lane, self-driving cars would be the new normal with additional features and automation techniques. The effect of Machine Learning in education would be intensified considering the globalization aspect wherein revolutionary thinkers wish to provide global classrooms for students across the world, for automating scoring and reporting test scores of students, and for aiding the students who are physically challenged. One of the other major fields in which Machine Learning is likely to flourish further is that of cybersecurity with the aim of preventing cyberbullying, cyber fraud, online cat fishing, and other cyber crimes in general. The military and defense sector is one of the other areas which can easily derive the benefits of Machine Learning. Be it for complementing soldiers in terms of decision making, developing better plans and attack strategies, or for monitoring and surveillance, Machine Learning is bound to pave the way.Health care and medical sector, home, workplace, unmanned aerial vehicles, etc. are some of the major fields which are likely to further invest in Machine Learning in the future.

4. Machine Learning Applications in Past, Present and Future

The Internet of Everything (IoE) includes a number of internet- and network-based concepts, including the Internet of Things (IoT), Internet of People (IoP), and Industrial Internet.A universal connection network of machines and smart gadgets capable of networking with one another, business processes, people, and the social environment is what is meant by the IoE, a current smart technology paradigm.

4.1 Investigating Health Research Trends Using Bibliometric Networks

We referred to the Web of Science (WoS) database [45] to discover the research trends in health considering IoT and Machine Learning topics since the year 2000. In Fig. 8, the research fields are categorized by different colors, and the size of the spheres shows the volme of papers published in each area. For example, the keywords such as "security", "healthcare system", "communication technology", "IoT technology" and "Radio-Frequency Identification (RFID)" are in a green color. In this context, most of the research has focused on security.

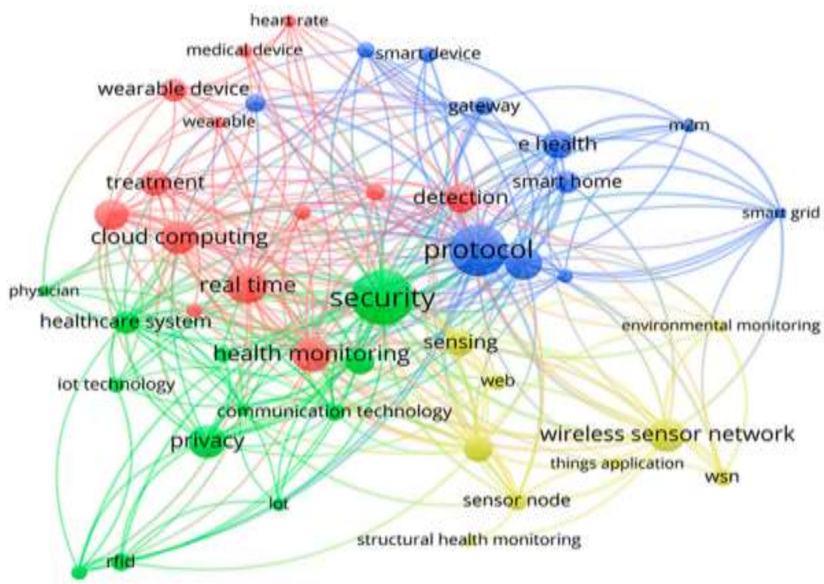

Fig. 8. The bibliometric networks in health and internet of things (IoT) themes from 2000 to 2021.

Infection, genomics, critical care, bioinformatics, and pathogens are only a few examples of the health themes with which artificial intelligence, big data, expert systems, and ontology have a strong interaction, as shown in Fig. 9. K closest neighbors, artificial neural network, and support vector machine techniques are represented in this diagram by the letters KNN, ANN, and SVM.

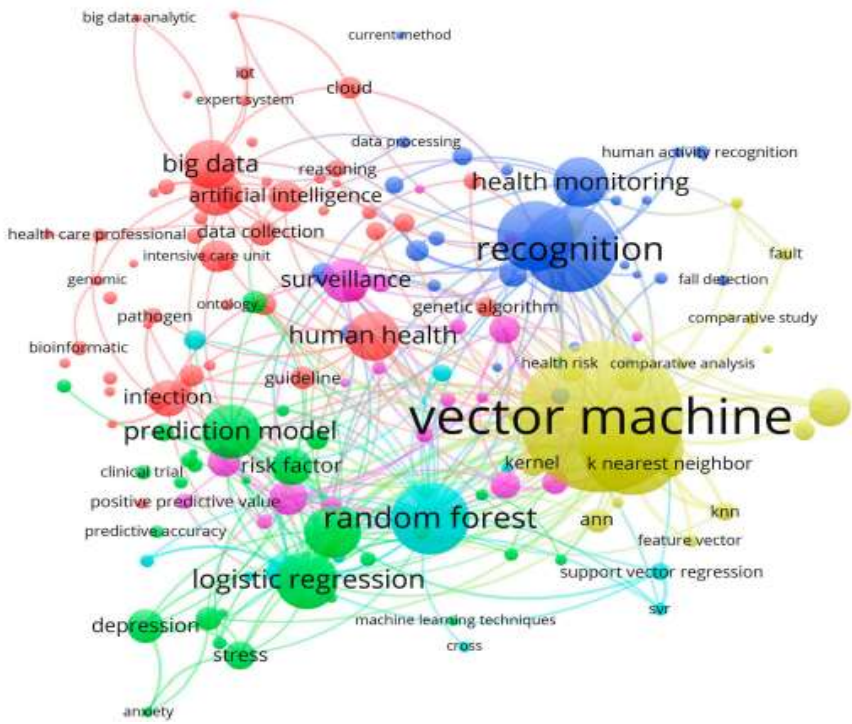

Fig. 9. The bibliometric networks in subjects related to Machine Learning and healthcare sector

The logistic regression method has, however, primarily been used for research on mental health issues like depression, stress, and anxiety. We selected the VOSviewer software to analyze the data since researchers have extensively used it to examine literature reviews in a variety of fields[46].

4.2 Applications of Machine Learning in Smart Grid

Machine Learning is now a widely used new technology in both business and non-commercial applications. Machine Learning has advanced particularly due to the rapidly expanding applications of deep learning. Deep learning has consequently become a popular research topic at academic institutions and tech firms. A multi-layer neural network model is typically used in deep learning to extract high-level features from a collection of low-level abstractions in order to locate scattered data features and solve challenging Machine Learning issues. Some of the early adopters of Machine Learning technologies include the smart grid in general and clean technology.

The forecast of the power consumer from the retailer's perspective is one of the additional uses of Machine Learning in the smart electricity environment. Machine Learning is used to predict the behavior of a day-ahead (DA) retail electrical energy market with price-based Demand Response (DR) using a hierarchical multi-agent framework. A retailer agent purchases energy from the DA wholesale market and sells it to consumers at the top of the hierarchical structure. The retailing agent's objective is to maximize profit by determining the ideal retail prices.

4.3 Deep Reinforcement Learning Applications for Smart Cities

The demand side's ability to modify consumption in response to signals from the electrical market is one of the key factors boosting the penetration level of green energy resources. Buildings used for domestic and commercial purposes are the main consumers of power. A larger penetration of green energy could be achieved by enhancing energy efficiency and using novel energy management strategies, signaling a change towards a sustainable energy environment. Understanding individual consumption patterns based on information from the massive data collected from Advanced Metering Infrastructure (AMI) as well as new data sources from smart home IoT is the important element of the practical demand side[47].

4.4 Applications in Supply Chain Management

Organizations can't function independently, and they rely on the skills and resources that suppliers, customers, and partners have to provide. Supply chain management (SCM) has received attention since 1980, when organizations began to see the advantages of fostering cooperative connections within their own walls through SCM [48]. Managing supply chains sustainably is a growing problem for businesses of all sizes and in a variety of sectors. The creation and marketing of sustainable goods can supplement this more reactive strategy of responding to external pressure from governments, consumers, non-governmental organizations (NGOs), and the media.308 publications on the subjects of "IoT" and "supply chain" were published between 1980 and 2018, according to our search of the Web of Science database. Note that in recent, China has made the largest contribution, contributing 45 percent of the knowledge domain of ML, while the USA came in second with 20 percent.

5. Literature Survey on Past and Present Machine Learning Efforts

The interpretation of NDE signals for detection and characterization has been the subject of decades of study in statistical and Machine Learning methods. This section makes no attempt to provide an exhaustive summary of all pertinent works in order to represent the general contours of commonly used Machine Learning techniques in NDE research.

5.1 Supervised Methods

Models are developed using labelled data using the Machine Learning technique known as supervised learning. Models must identify the mapping function in supervised learning in order to link the input variable (X) with the output variable (Y).

5.1.1 Support Vector Machines

Although they may be expanded for regression, support vector machines (SVMs) are mostly utilized for classification. Finding a hyperplane that divides the classes is the core premise of SVMs. Fig. 10 shows how to do this. The best separator from the several alternatives is identified and solved for [2] as having the biggest margin between classes. This criterion can then be relaxed to allow for a little amount of training data inaccuracy. When the classes cannot be linearly separated, it is possible to translate the data onto a higher-dimensional space using a kernel technique [1, 2] so that the data can be linearly divided (at least approximately).

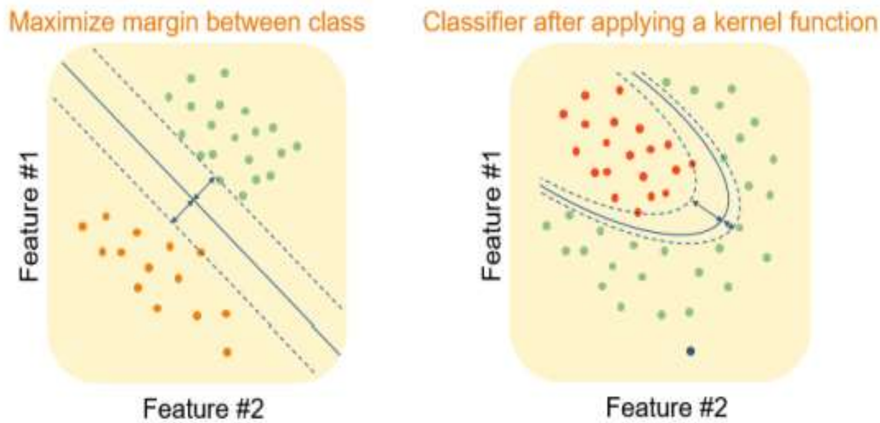

Fig. 10. Support Vector Machines

In SVM, the biggest margin across classes is predicted for a linear discriminator. Data that cannot be separated linearly can still be mapped using kernel functions.

Eddycurrent (the loops of electric current which is generated by changing magnetic fields within a conductor) [7], sonic IR [8], and ultrasound are a few instances of how SVMs have been used to NDE issues in the literature. For instance, in [9], damage was localized using simulated B-scans to build an SVM regression model. As shown in Fig. 11, the SVM operated as an inverse model, using a fresh B-scan as input and returning the position of the second concealed delamination as an output.

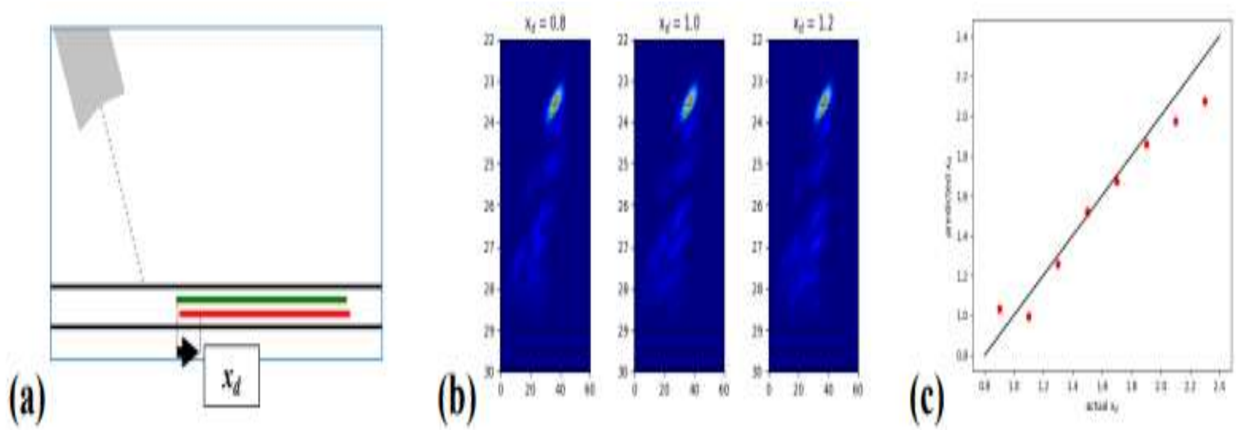

Fig. 11. Application of SVMs to UT B-scan Inversion - The B-scans in (b) were created by simulating oblique angle incidence UT in (a) (b). An SVM-based inverse model that predicts the site of the concealed delamination was trained using the simulations. The projected vs. actual position (in mm) in the graph displays the accuracy (c).

5.1.2 Neural Networks

An input layer and an output layer are joined by one or more hidden layers to form a neural network [1,2]. We refer to a neural network as a deep neural network when there are several hidden layers. A single value is multiplied by a learning weight for each edge in the network. The total of the weighted values refers to the congruence of edges at a node. Each node often additionally includes a learn bias word associated with it. An affine transformation of the inputs is made up of the weights and biases. A nonlinear function known as the activation function is then applied to the affine transform result. The node's output is formed by the activation's output, which is subsequently used as an input by the neural network's next layer. Each of these elements is shown in Fig. 12. The parameters of the neural network that need to be solved are the weights and biases for each node. There is a lengthy history of the use of neural networks in NDE, including the categorization of UT signals for the identification of fractures [10-12], corrosion [24], weld flaws [25], and other defects [24-28]. Additionally, the characterization of material characteristics [27], the localization of flaws, and the characterization of damage [29] have all been accomplished using neural networks. The fact that neural networks approximate generic functions is one of its key advantages.Several layers make up a neural network, with the input layer and the output layer being connected by one or more hidden layers.

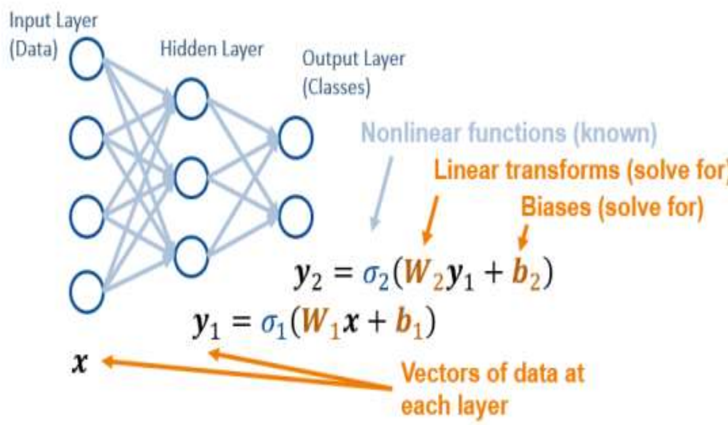

Fig. 12. Layers used in a Neural Network

5.2 Unsupervised Methods

Another Machine Learning technique is unsupervised learning, which infers patterns from unlabeled input data. Finding structure and patterns in the incoming data is the aim of unsupervised learning. There is no requirement for monitoring during unsupervised learning. Instead, it uses the data to discover patterns on its own.

5.2.1 Clustering

For the purpose of grouping a set of data into clusters, clustering is an unsupervised learning job. K-means clustering [2] is the most used clustering algorithm. Hierarchical clustering [1], spectral clustering [1], and the EM method [2] are further alternatives. The k-means algorithm's functionality for a problem with two inputs for two clusters is demonstrated in the example in Fig. 13. Two means are started randomly, starting with the unlabeled data. Then, depending on the separation between each mean and each point, each point is allocated to one of the k distinct clusters. Then, using each of the points in that cluster as input, a new set of k distinct means is computed.

Fig. 13 K-Means Clustering

Note that Clustering - Clustering is carried out by estimating new means for each of the newly labeled clusters up until convergence. Labels are iteratively assigned to each data point based on the nearest mean.

5.2.2 Matrix decomposition methods

To identify efficient representations of data, matrix decomposition techniques are utilized, including dictionary learning, independent component analysis, and singular value decomposition. These algorithms frequently pick up on data changes that are applied in later analysis, including better visualization or categorization. A 2D matrix containing the data is frequently used as the input for these transformations, which then divide the data into a matrix of components (also known as atoms or bases) and a matrix of coefficients. Fig. 14 displays an illustration of one of these matrices. Different assumptions are made by each decomposition method. For instance, orthogonal component vectors and orthogonal coefficient vectors are produced using singular value decomposition. A components matrix that is statistically independent is produced via independent component analysis. A sparse matrix of coefficients is created as a result of dictionary learning. Matrix decomposition approaches have been widely used in guided wave structural health monitoring to separate damage events and their changes over long time periods (i.e., spanning 10s to 1000s to 100,000s of guided wave data) [3-12]. Therefore, even in the presence of noise of noise and other confusing variables, such as temperature fluctuations, these matrix decomposition techniques may be utilized to acquire important information from datasets.

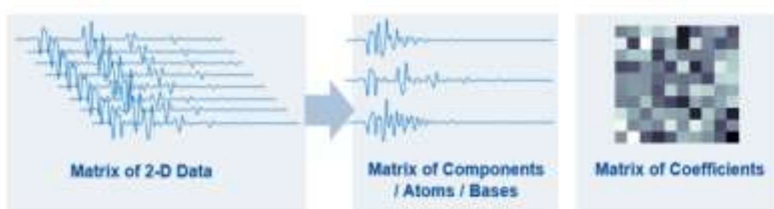

Fig. 14. Methods for Matrix Decomposition –

5.2.3 Reinforcement/ Feedback based Learning

Reinforcement learning which is often denoted as a feedback-based Machine Learning technique is one of the other forms of Machine Learning algorithms in which an agent is trained so as to behave in certain environments and scenarios on the bases of actions and corresponding results. This is a technique different from supervised learning which makes use of labelled data for training related activities because of which the training is restricted to experiences only. In simpler terms, reinforcement learning focuses on interacting with a particular environment and learns to act and behave within it. There are three main approaches when it comes to reinforcement learning – Value Based, Policy Based, and Model Based. The value-based approach is all about finding the optimal value function with the maximum value of a state. The policy-based approach is to configure the optimal policy with the maximum future rewards in the absence of any specific value function. And the model-based approach involves a virtual model that is created for the environment and doesn't revolve around any specific solution or algorithmic technique.

6. Metrics used in Machine Learning

There are a number of metrics which come handy and useful when it comes to evaluation of models and other neural networks developed using Machine Learning. Using these evaluation metrics, it is possible to identify the status of efficiency of the model/application and make necessary improvements to optimize it to the maximum. Metrics themselves can be classified into different types. One of the first types is the threshold-based discriminator metrics which is majorly used in binary classification related problems. There are four possible outcomes here – false positive, true positive, false negative, true negative.

Table 1. Threshold Metrics for Classification Evaluations

Metrics	Formula	Evaluation Focus
Accuracy (acc)	$\frac{tp + tn}{tp + fp + tn + fn}$	In general, the accuracy metric measures the ratio of correct predictions over the total number of instances evaluated.
Error Rate (err)	$\frac{fp + fn}{tp + fp + tn + fn}$	Misclassification error measures the ratio of incorrect predictions over the total number of instances evaluated.
Sensitivity (sn)	$\frac{tp}{tp + fn}$	This metric is used to measure the fraction of positive patterns that are correctly classified
Specificity (sp)	$\frac{tn}{tn + fp}$	This metric is used to measure the fraction of negative patterns that are correctly classified.
Precision (p)	$\frac{tp}{tp + fp}$	Precision is used to measure the positive patterns that are correctly predicted from the total predicted patterns in a positive class.
Recall (r)	$\frac{tp}{tp + tn}$	Recall is used to measure the fraction of positive patterns that are correctly classified
F-Measure (FM)	$\frac{2 * p * r}{p + r}$	This metric represents the harmonic mean between recall and precision values
Geometric-mean (GM)	$\sqrt{tp * tn}$	This metric is used to maximize the *tp* rate and *tn* rate, and simultaneously keeping both rates relatively balanced
Averaged Accuracy	$\frac{\sum_{i=1}^{l} \frac{tp_i + tn_i}{tp_i + fn_i + fp_i +}}{l}$	The average effectiveness of all classes
Averaged Error Rate	$\frac{\sum_{i=1}^{l} \frac{fp_i + fn_i}{tp_i + fn_i + fp_i +}}{l}$	The average error rate of all classes
Averaged Precision	$\frac{\sum_{i=1}^{l} \frac{tp_i}{tp_i + fp_i}}{l}$	The average of per-class precision
Averaged Recall	$\frac{\sum_{i=1}^{l} \frac{tp_i}{tp_i + fn_i}}{l}$	The average of per-class recall
Averaged F-Measure	$\frac{2 * p_M * r_M}{p_M + r_M}$	The average of per-class F-measure

Note: - each class of data; tp_i - true positive for C_i; fp_i - false positive for C_i; fn_i – false negative for C_i; tn_i - true negative for C_i; and M macro-averaging.

This metric is easy to comprehend, analyze and evaluate due to its simplicity and ability to be applied to multi-class problems as well. The next evaluation metric is Mean Square Error (MSE) which is best known for obtaining the difference between the predicted and the desired solution which is squared to avoid the possibility of nulling off of values. The lower the value of MSE, greater is the precision and accuracy of the model/algorithm. Then comes the Area under the ROC curve metric which is one of the popular ranking metrics which represents the total inclusive ranking of performance of a certain classifier. Apart from the above-mentioned metrics, there also exists hybrid metrics and certain other metrics like graphic-based metrics. These metrics are specifically significant to flamboyantly evaluate and analyze the results. As far as metrics are concerned, there are quite a few factors to be considered. The major factors include the possibility of issues in the case of multiclass algorithms, reduced complexity and cost of computation, uniqueness and discriminability, informativeness, favorability towards the minority class, etc. Here, Table 1 includes all Threshold Metrics for Classification Evaluations.

7. Issues In Machine Learning Based Applications

Even if Machine Learning has great potential and is currently helping businesses all over the world, there are still problems and difficulties in the industry. Machine Learning, for instance, excels at identifying patterns but struggles when it comes to generalizing information. There is also the issue of "algorithm fatigue" for users.Some Machine Learning-related problems have important repercussions that are already manifesting themselves. The "black box problem" - the absence of interpretability and explain ability — is one. Even its creators are unable to fully comprehend how Machine Learning models generate their own actions and judgments. This makes it challenging to correct mistakes and guarantee that the data a model disseminates is accurate and impartial. People noted, for instance, that Apple's credit card algorithm offered women much lesser credit limits than men, but the corporation was unable to explain why and had no idea how to correct the problem. This relates to the biggest problem affecting the area right now: algorithmic bias and biased data. Algorithmic bias elucidates the possibility of repeatable errors and issues that often lead to unfair and biased outputs like favoring one arbitrary group of users over that of others. Biased data is what often leads to inaccuracy in predictions due to distortions of perceptions in decision making and data gathering.

Despite the fact that this problem is widespread and well-known, there is opposition to taking the considerable action that many experts in the field believe is required. TimnitGebru and Margaret Mitchell, the co-leaders of Google's ethical AI team, were sacked by the business in what thousands of workers referred to as a "retaliatory termination" when Gebru refused to retract research on the dangers of using huge language models. Researchers, decision-makers, and activists were polled on their concerns about the future of AI, and the majority expressed concern that, by 2030, it will still be largely focused on maximizing profits and societal control at the expense of ethics. Across the nation, legislation pertaining to AI is being discussed and implemented, particularly considering its immediate and evident detrimental uses, like as facial recognition for law enforcement.

7.1 Healthcare Problems and Challenges

The healthcare budget is under stress due to the ageing population, so innovative solutions must be put forth to make up for the limited healthcare resources. The Internet of Things is a promising solution to address this issue. IoT can be applied to the healthcare industry as a facilitator for monitoring, diagnostics, and even the potential for online telesurgery. This study discusses the difficulties in implementing smart and linked devices for healthcare applications.

One of the biggest issues would be connected to the security and identification of the nodes given the widespread use of IoT devices in the healthcare industry. This is required to recognize the information that has been received and allocate it to the appropriate node. The security of the nodes is very important since hostile activity can compromise the system. Use of trusted execution environments may be one remedy for this issue.

Telecommunication management is a different issue that needs to be handled, especially in the health sector. The availability of patient monitoring systems and device downtime could both be greatly decreased by choosing the best possible combination of telecommunication technologies [44]. The emergence of location technology presents another difficulty for the IoT in healthcare. Real-time location solutions are necessary due to the increased IoT device penetration in the healthcare sector. With the aid of this technology, the healthcare system can be reconfigured based on the availability of distributed resources and the secure tracking of the treatment process. One of the most useful methods in thisarea might be a synthesis of the Internet of Things (IoT) and Global Positioning System (GPS). To improve location accuracy, local positioning systems must be added.

7.2 Issues towards other Sectors

Just like how every coin has two sides, Machine Learning can also cause problems and issues in a few sectors of its application as well. When it comes to the marketing sector, the fact that data inaccessibility and security can cost a lot is a potential problem as Machine Learning models require ample amounts of data for training. Similarly, the affordability and time-consuming application development process can often be cumbersome for business at the initial stages of their growth. Another reason as to why Machine Learning can be a possible issue is because application developers and businesses need to understand and comprehend the nuance of where exactly can they derive the benefits of Machine Learning in their application. Problems like lack of quality and abundance of data, lack of self-resources, inadequate infrastructure, efficient implementation techniques, overfitting and underfitting of data, and investing in monitoring and maintain the models and systems can be tedious tasks for those involved in using Machine Learning based models for their use cases and applications. Additional issues like irrelevancy of features and data and algorithmic biases can also contribute to the deficiency of accuracy in the model. And the above-mentioned problems can occur with respect to any of the sectors involved in Machine Learning applications – education, transportation, health, business, finance, etc.

8. Challenges towards Machine Learning ApplicationsImplementation

The practical applications for Machine Learning will soon grow significantly if present trends in algorithm development and hardware performance continue. Deep-learning algorithms are already used to identify objects and language in photos, as well as emotions from face photographs and other things. Think about these potential future scenarios:

- Self-driving cars that can not only recognize objects, such as pedestrians, bicycles, and animals, but predict what these objects are likely to do
- Systems that can "watch," score, and report on sporting events
- Systems that can recognize, and neutralize, hacking attempts and fraudulent activity, even for approaches and vulnerabilities that were not previously known
- Robots that can detect our emotional state and tailor their interactions with us.

The ways that "intelligent" machines will help humanity in the future are just beginning to be explored. Although computers may never be able to think like humans can, they may be ideally suited to activities like these that call for a lot more perseverance, speed, or patience than the majority of people can manage. Be prepared for some fascinating advancements in AI.Many of the applications in the banking sector make use of Machine Learning and artificial intelligence, especially for fault/fraud detection in credit cards. However, the process if very often induced by redundancy and can make it tough for the systems to spot and detect frauds without massive amounts of data. One of the other areas which highly exploits the techniques of Machine Learning and artificial intelligence is that of recommendation systems or proposal engines. While a few systems can be highly depended upon, not all recommendation engines serve the purpose with precision. One of the biggest challenges is the lack of sufficient experts, professionals and talents in the field of Machine Learning who are capable enough to spearhead the innovation process. Similarly, problems like making wrong and uncalculated assumptions, lack of suitable infrastructural resources, obsoleteness of algorithms with the growth in dataset, high complexity, concepts drift, heavy maintenance, and high likelihood of errors are some of the major challenges to be overcome as we move towards the full-fledged use of Machine Learning in the future. The major hinderances and troubles faced in Machine Learning implementations are with respect to volume, variety and velocity which have been detailed in Table 1. given below.

Table 1. Challenges faced in Machine Learning [45]

Characteristic Feature of Machine Learning	**Challenges Faced**
With Regards to Volume	
Performance in terms of Processing Power	Space and time complexity increases
Imbalance of Class	Sampling strategies may lead to data corruption
Variance and Biasness	Issue of overfitting and failure in case of generalization prediction with increase in data size
Non-Linearity	This in itself is a challenge
With Regards to Variety	
Localization of Data	Storing all required data locally can never be guaranteed completely
Heterogenous Data	Homogenous distribution pertaining to various data sources is highly unlikely
Noise in Data	Increasing cost for cleaning, maintaining, and managing missing data and outliers
With Regards to Velocity	
Streaming and Availability of data	Data being streamed is not exposed to the train, test and predict framework leading to susceptible model changes
Independent and Uniformly distributed Random Variables	There's no guarantee for completely independent variables for the entire population of data
Drift of Concept	Initially accurate models may drift leading to worse outputs

9. The Future with Machine Learning

We quickly go through some of the existing issues with using Machine Learning to NDE applications in this part. We next go over some recent advances in Machine Learning that are expected to help resolve these issues and further the study of NDE.

9.1 Challenges with Traditional Machine Learning

Applying Machine Learning to actual issues presents a number of difficulties. In the realm of NDE, certain of these are particularly important. The feature problem, the black box problem, and the data problem are three especially important difficulties that we discuss in this article.The chosen data features, which are often extracted from the raw data in a procedure known as "feature engineering" or "feature extraction," are frequently highly dependent on the accuracy of Machine Learning models. Two examples are shown in Fig. 15 to demonstrate the feature problem. The features that are not related to the output and for which it would be challenging to create an appropriate classifier are shown in the image on the left. This is due to the features' minimal physical connection to the relevant problem, such as damage detection. The figure on the right solves this issue by utilizing two far more sensible characteristics. We see that the two classes can be distinguished linearly, which enables us to create a suitable classifier.

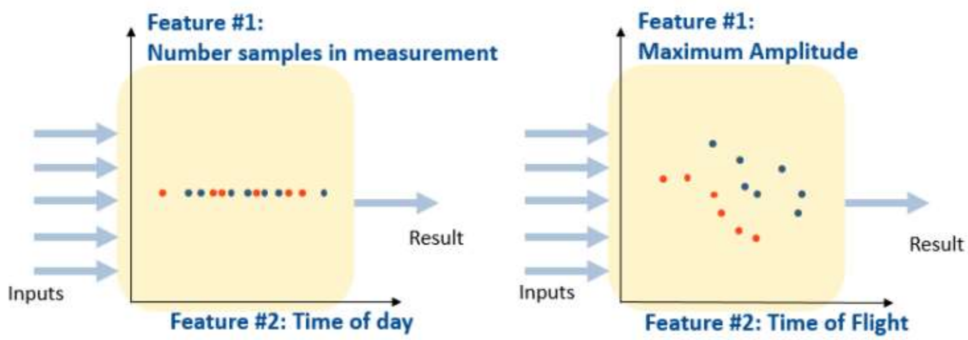

Fig. 15. The Feature Problem

Note that it is possible that features have no connection to the output. Some attributes could serve as a more accurate representation of the data and serve categorization purposes better.

Ultrasonic testing, eddy current, sonic infrared, thermography, and other measures with a very large number of dimensions are often used to solve NDE problems in the form of time series or pictures. For the purpose of creating meaningful models, many of the existing techniques entail extracting a maximum amplitude, duration of flight, spectral data, or other properties. Although they have in the past, these are not the best solutions for all issues. Data is frequently impacted by many degradation processes in complicated and varied ways. The common data characteristics that result from these variations are not always the same.Deep learning has recently been utilized in many domains to overcome this problem (cascading several Machine Learning models connected in layers, with the results of the first model feeding as inputs into the next model) [30-33]. A deep neural network, which connects several hidden layers between the input and output layers, is a popular method of deep learning. Deep learning techniques can successfully extract more optimum features from raw data in an early Machine Learning layer if there are enough layers and data [30]. The categorization or regression is then carried out by later layers. This can therefore solve the feature issue.

The Black Box Challenge

Machine Learning models are sometimes referred to as "black boxes" and are challenging to understand. When a signal is classed for NDE, this means that Machine Learning models will not usually offer much extra information to describe what aspect of the signal or characteristics was utilized to reach that determination. The behavior of most Machine Learning models can vary greatly depending on the data at hand because they are not based on any physical laws.

As a result, the black box dilemma has effects on how Machine Learning in NDE is validated.Machine Learning based on physics has been highlighted in recent artificial intelligence research [34–39]. By limiting the optimizations required to develop each model, these techniques often integrate physical information. These techniques are therefore easier to understand and comprehend. These techniques also require less data than a conventional Machine Learning approach. But keep in mind that these initiatives are still rather new.

The Data Challenge

A lot of training data is frequently needed for statistical and Machine Learning models, especially deep models. Overfitting can occur when a deep neural network, for example, is trained with sparse data. The model will therefore very precisely recreate the training data or data that are extremely similar to the training data. For fresh, more generic data, the model won't be able to generate reliable predictions. Because we lack the enormous databases that are available for other Machine Learning applications, such as image and audio processing, this is a typical difficulty in NDE and other engineering fields. Due to their experimental difficulty and the fact that many NDE datasets are not directly comparable, we frequently lack these databases.

Transfer learning [40-43] is a recent development in Machine Learning that can help with this problem. Transfer learning enables us to transfer our knowledge from one learning assignment to another. As a result, we may frequently train models that are considerably simpler and do a certain goal. The links between these smaller models may then be discovered by combining them into a larger model. Each issue employs simpler models as a result, requiring less information overall [43]. Additionally, it is not necessary for the data used to train each job to be directly comparable.

10. Opportunity with Machine Learning

The concept of "deep learning" is at the core of a lot of contemporary Machine Learning research. This method divides the "neurons" of a neural network into three layers: an input level, an output level, and one or more intermediate "hidden" levels. The intricacy of the potential connections between input and output increases with the number of layers and "neurons" in each level. The computer can reach many conclusions by attempting various pathways. The computer may be taught to draw the appropriate conclusions from inputs that are similar to, but not precisely like, ones it has already experienced by "training" it is using a collection of known inputs and outputs. This is notably helpful in image processing, where the computer can be taught to tell photos of dogs apart from images without them by being "shown" thousands of distinct photographs of dogs.The ways that "intelligent" machines will help humanity in the future are just beginning to be explored [49, 50]. Although computers may never be able to think like humans can, they may be ideally suited to activities like these that call for a lot more perseverance, speed, or patience than the majority of people can manage. Be prepared for some fascinating advancements in AI. Few of other popular opportunities in near future towards ML will be:

- Quantum Machine Learning
- ML based IoT based Communication
- Blockchain based ML

- ML based cloud/ fog/ edge computing

11. Conclusion and Future Work

Nondestructive assessment and related subjects stand to benefit significantly from research on Machine Learning, which is a potent tool in its own right. As said, there are a number of issues (including the feature problem, the data challenge, and the black box challenge) that have limited the effectiveness of earlier Machine Learning initiatives. This study outlines these difficulties, discusses how they connect to earlier Machine Learning initiatives in NDT, and offers a perspective on how to address them (via deep learning, transfer learning, and Machine Learning that is inspired by physics).

We urge the NDE community to use Machine Learning techniques more frequently and actively. We do advise against viewing Machine Learning as a tool, but rather as a field of study. Machine Learning for designed systems remains a serious issue in environments with limited data availability and high levels of variability. We require researchers who are knowledgeable about physics and engineering issues to take the initiative in advancing the Machine Learning discipline in order to meet this challenge. Then we can develop durable Machine Learning solutions for NDE issues.

References

1. K. Murphy, Machine Learning: A Probabilistics Perspective, The MIT Press, Cambridge, Massachusetts, 2012.
2. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2ndEd., Springer ScienceBusiness Media, LLC, 2009.
3. C. Liu, J. Dobson, and P. Cawley, “Efficient generation of receiver operating characteristics for the evaluation of damage detection in practical structural health monitoring applications,” Proc. Math. Phys. Eng. Sci., vol. 473, no. 2199, p. 20160736, Mar. 2017.
4. J. Dobson and P. Cawley, “Independent Component Analysis for Improved Defect Detection in Guided Wave Monitoring,” Proc. IEEE, vol. 104, no. 8, pp. 1620–1631, Aug. 2016.
5. J. B. Harley, C. Liu, I. J. Oppenheim, and J. M. F. Moura, “Managing Complexity, Uncertainty, and Variability in Guided Wave Structural Health Monitoring,” SICE Journal of Control, Measurement, and System Integration, vol. 10, no. 5, pp. 325–336, 2017.
6. C. Liu, J. B. Harley, M. Berges, D. W. Greve, and I. J. Oppenheim, “Robust ultrasonic damage detection under complex environmental conditions using singular value decomposition,” Ultrasonics, vol. 58, pp. 75– 86, Apr. 2015.
7. C. Zang, M. I. Friswell, and M. Imregun, “Structural Damage Detection using Independent Component Analysis,” Structural Health Monitoring, vol. 3, no. 1, pp. 69–83, Mar. 2004.
8. V. Morovati and M. T. Kazemi, “Detection of sudden structural damage using blind source separation and time–frequency approaches,” Smart Mater. Struct., vol. 25, no. 5, p. 055008, Mar. 2016.
9. S. Marinetti, L. Finesso, and E. Marsilio, “Matrix factorization methods: Application to thermal NDT/E,” NDT E Int., vol. 39, no. 8, pp. 611–616, Dec. 2006.
10. B. Gao, W. L. Woo, and B. W.-K. Ling, “Machine Learning source separation using maximum a posteriori nonnegative matrix factorization,” IEEE Trans Cybern, vol. 44, no. 7, pp. 1169–1179, Jul. 2014.
11. G.-M. Zhang, D. M. Harvey, and D. R. Braden, “Advanced acoustic microimaging using sparse signal representation for the evaluation of microelectronic packages,” IEEE Trans. Adv. Packag., vol. 29, no. 2, pp. 271–283, May 2006.
12. M. Johnson, “Waveform based clustering and classification of AE transients in composite laminates using principal component analysis,” NDT E Int., vol. 35, no. 6, pp. 367–376, Sep. 2002.
13. M. Salucci, G. Oliveri, F. Viani, R. Miorelli, C. Reboud, P. Calmon, A. Massa, “A

learning-by-examples approach for nondestructive localization and characterization of defects through eddy current measurements", 2015 IEEE Int. Sym. Antennas Propag., 2015.
14. Z. Zeng, J. Zhou, N. Tao, L. Feng, C. Zhang, and X. Han, "Support vector machines-based defect recognition in SonicIR using 2D heat diffusion features," NDT & E International, vol. 47, 2012, pp. 116-123.
15. D. Sparkman, et.al, "A Direct Inverse Model for Ultrasonic Delmanination Characterization in PMCs", Review of Progress in Quantitative Nondestructive Evaluation, 2017.
16. B. Gao, W. L. Woo, and B. W.-K. Ling, "Machine Learning source separation using maximum a posteriori nonnegative matrix factorization," IEEE Trans Cybern, vol. 44, no. 7, pp. 1169–1179, Jul. 2014.
17. F. W. Margrave, K. Rigas, D. A. Bradley, and P. Barrowcliffe, "The use of neural networks in ultrasonic flaw detection," Measurement, vol. 25, no. 2, pp. 143–154, Mar. 1999.
18. S. Yuan, L. Wang, and G. Peng, "Neural network method based on a new damage signature for structural health monitoring," Thin-Walled Struct., vol. 43, no. 4, pp. 553–563, Apr. 2005.
19. S. W. Liu, J. H. Huang, J. C. Sung, and C. C. Lee, "Detection of cracks using neural networks and computational mechanics," Comput. Methods Appl. Mech. Eng., vol. 191, no. 25, pp. 2831–2845, Apr. 2002.
20. A. De Fenza, A. Sorrentino, and P. Vitiello, "Application of Artificial Neural Networks and Probability Ellipse methods for damage detection using Lamb waves," Compos. Struct., vol. 133, pp. 390–403, Dec. 2015.
21. C. Zang, M. I. Friswell, and M. Imregun, "Structural Damage Detection using Independent Component Analysis," Structural Health Monitoring, vol. 3, no. 1, pp. 69–83, Mar. 2004.
22. V. Morovati and M. T. Kazemi, "Detection of sudden structural damage using blind source separation and time–frequency approaches," Smart Mater. Struct., vol. 25, no. 5, p. 055008, Mar. 2016.
23. K. Virupakshappa and E. Oruklu, "Ultrasonic flaw detection using Support Vector Machine classification," 2015 IEEE IUS, Taipei, 2015, pp. 1-4.
24. M. Fahad, K. Kamal, T. Zafar, R. Qayyum, S. Tariq, and K. Khan, "Corrosion detection in industrial pipes using guided acoustics and radial basis function neural network," in 2017 International Conference on Robotics and Automation Sciences (ICRAS), 2017, pp. 129–133.
25. W. Hou, Y. Wei, J. Guo, Y. Jin, and C. Zhu, "Automatic Detection of Welding Defects using Deep Neural Network," J. Phys. Conf. Ser., vol. 933, no. 1, p. 012006, Jan. 2018.
26. J. Rafiee, F. Arvani, A. Harifi, and M. H. Sadeghi, "Intelligent condition monitoring of a gearbox using artificial neural network," Mech. Syst. Signal Process., vol. 21, no. 4, pp. 1746–1754, May 2007.
27. S. Sharma and A. Das, "Backcalculation of pavement layer moduli from falling weight deflectometer data using an artificial neural network," Can. J. Civ. Eng., vol. 35, no. 1, pp. 57–66, Jan. 2008.
28. S. Zhang, H. Wang, W. Wang, and S. Chen, "Damage Detection in Structures Using Artificial Neural Networks," in 2010 International Conference on Artificial Intelligence and Computational Intelligence, 2010, vol. 1, pp. 207–210.
29. A. Bernieri, L. Ferrigno, M. Laracca, and M. Molinara, "Crack Shape Reconstruction in Eddy Current Testing Using Machine Learning Systems for Regression," IEEE Trans. Instrum. Meas., vol. 57, no. 9, pp. 1958– 1968, 2008.
30. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
31. Y. Bengio, "Learning Deep Architectures for AI," Foundations and Trends® in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.

32. M. Abadi et al., “TensorFlow: A System for Large-Scale Machine Learning,” in OSDI, 2016, vol. 16, pp. 265–283.
33. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
34. A. Karpatne, W. Watkins, J. Read, and V. Kumar, “Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling,” arXiv [cs.LG], 31-Oct-2017.
35. R. Paolucci, F. Gatti, M. Infantino, C. Smerzini, A. GüneyÖzcebe, and M. Stupazzini, “Broadband Ground Motions from 3D Physics-Based Numerical Simulations Using Artificial Neural Networks,” Bull. Seismol. Soc. Am., vol. 108, no. 3A, pp. 1272–1286, 2018.
36. A. Karpatne et al., “Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data,” IEEE Trans. Knowl. Data Eng., vol. 29, no. 10, pp. 2318–2331, Oct. 2017.
37. D. Finol, Y. Lu, V. Mahadevan, and A. Srivastava, “Deep Convolutional Neural Networks for Eigenvalue Problems in Mechanics,” arXiv [physics.comp-ph], 17-Jan-2018.
38. M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations,” arXiv [cs.AI], 28-Nov-2017.
39. M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations,” arXiv [cs.AI], 28-Nov-2017.
40. Y. Bengio, “Deep learning of representations for unsupervised and transfer learning,” in Proceedings of ICML Workshop on Unsupervised and Transfer Learning, 2012, pp. 17–36.
41. S. Ghosh, “Transductive transfer learning using autoencoders,” Indian Statistical Institute, Kolkata, 2016.
42. H.-C. Shin et al., “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1285–1298, May 2016.
43. S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE Trans. Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, 2010.
44. Yin, Y.; Zeng, Y.; Chen, X.; Fan, Y. The internet of things in healthcare: An overview. J. Ind. Inf. Integr. 2016, 1, 3–13. [CrossRef].
45. Augenstein, Christoph, Norman Spangenberg and Bogdan Franczyk. “Applying Machine Learning to big data streams : An overview of challenges.” 2017 IEEE 4th International Conference on Soft Computing & Machine Intelligence (ISCMI) (2017): 25-29.
46. Journal Citation Report (JCR). Available online: https://clarivate.com/products/journal-citation-reports/ (accessed on 21 February 2018).
47. Van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010, 84, 523–538. [CrossRef] [PubMed].
48. Seunghyoung, R.; Jaekoo, N.; Hongseok, K. Deep Neural Network Based Demand Side Short Term Load Forecasting. In Proceedings of the 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), Sydney, Australia, 6–9 November 2016; pp. 308–313.
49. Giannakis, M.; Croom, S.R. Toward the Development of a Supply Chain Management Paradigm: A Conceptual Framework. J. Supply Chain Manag. 2004, 40, 27–37.
50. Mishra S., Tyagi A.K. (2022) The Role of Machine Learning Techniques in Internet of Things-Based Cloud Applications. In: Pal S., De D., Buyya R. (eds) Artificial

Intelligence-based Internet of Things Systems. Internet of Things (Technology, Communications and Computing). Springer, Cham. https://doi.org/10.1007/978-3-030-87059-1_4