

Review

Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions

Mohammad Mustafa Taye 

Data Science and Artificial Intelligence, Philadelphia University, Amman 19392, Jordan; mtaye@philadelphia.edu.jo

Abstract: In recent years, deep learning (DL) has been the most popular computational approach in the field of machine learning (ML), achieving exceptional results on a variety of complex cognitive tasks, matching or even surpassing human performance. Deep learning technology, which grew out of artificial neural networks (ANN), has become a big deal in computing because it can learn from data. The ability to learn enormous volumes of data is one of the benefits of deep learning. In the past few years, the field of deep learning has grown quickly, and it has been used successfully in a wide range of traditional fields. In numerous disciplines, including cybersecurity, natural language processing, bioinformatics, robotics and control, and medical information processing, deep learning has outperformed well-known machine learning approaches. In order to provide a more ideal starting point from which to create a comprehensive understanding of deep learning, also, this article aims to provide a more detailed overview of the most significant facets of deep learning, including the most current developments in the field. Moreover, this paper discusses the significance of deep learning and the various deep learning techniques and networks. Additionally, it provides an overview of real-world application areas where deep learning techniques can be utilised. We conclude by identifying possible characteristics for future generations of deep learning modelling and providing research suggestions. On the same hand, this article intends to provide a comprehensive overview of deep learning modelling that can serve as a resource for academics and industry people alike. Lastly, we provide additional issues and recommended solutions to assist researchers in comprehending the existing research gaps. Various approaches, deep learning architectures, strategies, and applications are discussed in this work.

Keywords: machine learning (ML); deep learning (DL); recurrent neural network (RNN); convolutional neural networks (CNN) artificial intelligence (AI)



Citation: Taye, M.M. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers* **2023**, *12*, 91. <https://doi.org/10.3390/computers12050091>

Academic Editors: M. Ali Akber Dewan and Aditya Kumar Sahu

Received: 11 February 2023
Revised: 19 April 2023
Accepted: 22 April 2023
Published: 25 April 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Machine learning is used to make computers execute activities that humans can perform more effectively [1]. Using computer algorithms, machine learning enables the machine to access data automatically and with enhanced experience as it learns. It has simplified life and become an indispensable instrument in several industries, such as agriculture [2], banking [3], optimisation [4], robotics [5], structural health monitoring [6], and so on. It may be utilised in cameras for object detection, picture, colour, and pattern recognition, data collection, data sorting, and audio-to-text translation [7].

Deep learning [8] is one of the machine learning methods that dominate in various application areas. Machine learning functions similarly to a newborn infant. There are billions of linked neurons in the brain, which are engaged when a message is sent to the brain. When a baby is shown a vehicle, for instance, a specific set of neurons are active. When the infant is shown another automobile of a different model, the same set of neurons plus some extra neurons may be triggered. Thus, humans are trained and educated during childhood, and during this process, their neurons and the pathways linking them are modified.

If artificial intelligence is like a brain, then machine learning is the process through which AI gains new cognitive abilities, and deep learning is the most effective self-training system now available. Machine learning is the study of making computers learn and improve in ways that mimic or outperform human learning ability. Developers train models to predict the intended result from a set of inputs. This understanding takes the place of the computer programme in the past. The entire discipline of artificial intelligence known as machine learning is founded on this principle of learning by example, of which deep learning is a subset. Instead of providing a computer with a long list of rules to follow in order to solve a problem.

This is also how machine learning works, where the computer is trained using several instances from the training data sets, neural networks are trained, and their routes adjusted accordingly. The machine receives fresh inputs and generates outputs. Real-world applications of this technology include spam filters in Gmail, Yahoo, and the true caller app, which filters spam emails; Amazon's Alexa; and the recommended videos that appear on our YouTube homepage based on the kind of videos we watched before. Tesla, Apple, and Nissan are among the businesses developing autonomous technology based on deep learning. Deep learning is one of the machine learning approaches [9].

Due to the introduction of many efficient learning algorithms and network architectures [10] in the late 1980s, neural networks became an important topic in the fields of machine learning (ML) and artificial intelligence (AI). Multilayer perceptron networks trained with "Backpropagation" type algorithms, self-organising maps, and radial basis function networks were examples of such novel techniques [11–13]. Despite the fact that neural networks are successfully utilised in a variety of applications, interest in exploring this issue has declined with time.

In 2006, Hinton et al. [8] presented "Deep Learning" (DL), which was based on the notion of an artificial neural network (ANN). After that, deep learning became a significant topic, resulting in a renaissance in neural network research, hence the term "new generation neural networks." This is because, when trained correctly, deep neural networks have shown to be very good at a wide range of classification and regression problems [10]. Because of its ability to learn from the provided data, DL technology is currently one of the hottest issues within the fields of machine learning, artificial intelligence, data science, and analytics. In terms of its working domain, DL is regarded as a subset of ML and AI; hence, DL may be viewed as an AI function that mimics the processing of data by the human brain.

Deep learning algorithms profit from increased data production, better processing power currently available, and the growth of artificial intelligence (AI) as a service. Even with highly varied, unstructured, and connected data collection, deep learning enables machines to solve complicated issues. Deep learning algorithms perform better the more they learn [14–17].

This study's main objective is to draw attention to the most important DL elements so that researchers and students may quickly and easily get a thorough grasp of DL from a single review piece. Additionally, it let people understand more about current developments in the area, which will enhance DL research. To provide more precise opportunities to the field, researchers would be allowed to select the best route of study to pursue.

This type of learning is at the heart of the fourth industrial revolution (Industry 4.0).

The general contribution of this study is summarised as follows:

To investigate several well-known ML and DL methods and provide a taxonomy reflecting the differences between deep learning problems and their applications.

The main focus of the review that follows is deep learning, including its fundamental ideas and both historical and current applications in various domains.

This article focuses on deep learning workflow and modelling, that is, the ability of DL techniques to learn.

This article helps developers and academics gain a broader understanding of DL methodologies; I have summarised numerous potential real-world application areas of DL.

This article gives developers and academics an overview of the future directions of deep learning focused on medical applications.

An overview of the development of machine learning is provided in Section 2. A full view of machine learning, learning methodologies, including supervised learning, unsupervised learning, and hybrid learning, advantages and disadvantages of deep learning, a timeline of DL, DL workflow, and the various machine learning types and algorithms are depicted in Section 2. A summary of deep learning applications is provided in Section 3. Section 4 provides an overview of the future directions of deep learning.

2. Machine Learning and Deep Learning

In machine learning, a computer programme is given a set of tasks to complete, and it is said that the machine has learned from its experience if its measured performance in these tasks improves over time as it obtains more and more practice completing them. This means that the machine is making judgments and forecasts based on historical data. Consider computer software that learns to diagnose cancer based on a patient's medical records. When it analyses medical investigation data from a larger population of patients, its performance will increase through the accumulation of knowledge [18].

The number of cancer cases correctly predicted and detected, as verified by a seasoned oncologist, will serve as the performance metric. Machine learning is used in many different areas, including but not limited to robotics, virtual personal assistants (such as Google), video games, pattern recognition, natural language processing, data mining, traffic prediction, online transportation networks (such as Uber's surge pricing estimates), product recommendations, stock market forecasts, medical diagnoses, fraud predictions, agricultural advice, and search engine result refinement (such as Google's search engine) [19].

In artificial intelligence (AI), machine learning is the ability to automatically adapt with little to no human intervention, and deep learning is a subset of machine learning that uses neural networks to simulate the human brain's learning procedure. There is a wide gap between these two ideas. Although it needs more data to train on, deep learning can adapt to new circumstances and correct for its own faults.

Contrarily, machine learning permits training on smaller datasets, but it requires more human intervention to learn and correct its errors. Machine learning relies on human intervention to categorise data and highlight attributes. In contrast, a deep learning system aims to acquire these qualities without any human input. In the simplest of explanations, machine learning functions like an obedient robot. Patterns in the data are analysed in order to make predictions. If you can imagine a robot that learns on its own, that is what deep learning is like. It can learn more intricate patterns and generate independent predictions.

Deep neural networks constitute a subfield of machine learning. It is a model of a network consisting of neurons with multiple parameters and layers between input and output. DL uses neural network topologies as its basis. Consequently, they are known as deep neural networks [20]. DL provides autonomous learning of characteristics and their hierarchical representation at multiple levels. In contrast to conventional machine learning approaches, this robustness is a result of deep learning's powerful process; in brief, deep learning's whole architecture is used for feature extraction and modification. The early layers do rudimentary processing of incoming data or learn simple features, and the output is sent to the upper layers, which are responsible for learning complicated features. Hence, deep learning is suited for handling larger data sets and greater complexity.

2.1. The Key Distinctions between Deep Learning and Machine Learning

Machine learning learns to map the input to output given a specific world representation (features) that have been manually designed for each task.

Deep learning is a type of machine learning that seeks to represent the world as a nested hierarchy of concepts that are automatically detected by the architecture of deep learning.

The paradigm of ML and DL is the development of data-driven algorithms. Either structured or unstructured data are utilised to collect and derive the necessary task-related information.

Deep learning models have an advantage over traditional machine learning models due to the increased number of learning layers and a higher level of abstraction. One more reason for this advantage is that there is direct data-driven learning for all model components. Due to the algorithm, they rely on, traditional machine learning models encounter limitations as data size and demand for adequate insights from the data increase. The expansion of data has spurred the development of more sophisticated, rapid, and accurate learning algorithms. To maintain a competitive advantage, every organisation will employ a model that produces the most accurate predictions.

As a simple comparison between deep learning and traditional machine learning techniques, demonstrating how DL modelling can improve performance as data volume increases, see Table 1.

Table 1. Comparison between deep learning and traditional machine learning.

	Machine Learning	Deep Learning
Human Intervention	To achieve outcomes, machine learning requires more continuous human engagement.	Deep learning is more difficult to implement initially but requires little intervention afterward.
Hardware	Machine learning programmes are typically less complicated than deep learning algorithms and may frequently be executed on standard computers.	Deep learning systems necessitate significantly more robust hardware and resources. The increasing demand for power has increased the utilisation of graphics processing units. GPUs are advantageous due to their high bandwidth memory and thread parallelism's ability to conceal memory transfer latency (delays) (the ability of many operations to run efficiently at the same time.)
Time	Machine learning systems can be installed and used quickly, but their results may not be as good as they could be.	Deep learning systems take more time to set up, but they can give results right away (though the quality is likely to get better as more data becomes available).
Approach	Typically, machine learning requires organised data and uses conventional techniques such as linear regression.	Deep learning utilises neural networks and is designed to handle massive volumes of unstructured data.
Applications	Email, bank, and doctor's office all currently utilise machine learning.	Deep learning technology enables more complicated and autonomous programmes, such as self-driving automobiles and surgical robots.
Usage	There are numerous applications for machine learning, including regression analysis, classification, and clustering.	Deep learning is typically employed for complicated tasks such as picture and speech recognition, natural language processing, and autonomous systems.
Data	In general, machine learning algorithms use less data than deep learning algorithms, although data quality is more crucial.	Deep learning algorithms require enormous amounts of data to train neural networks but may learn and improve autonomously as additional data are processed.

DL is quite data-hungry given that it includes representation learning [21]. A well-behaved performance model for DL requires an enormous quantity of data, i.e., as the data accumulates, a more well-behaved performance model may be created (Figure 1). In the majority of instances, the given data are sufficient to build a reliable performance model. Unfortunately, there is occasionally an absence of data for the direct use of DL [22,23].

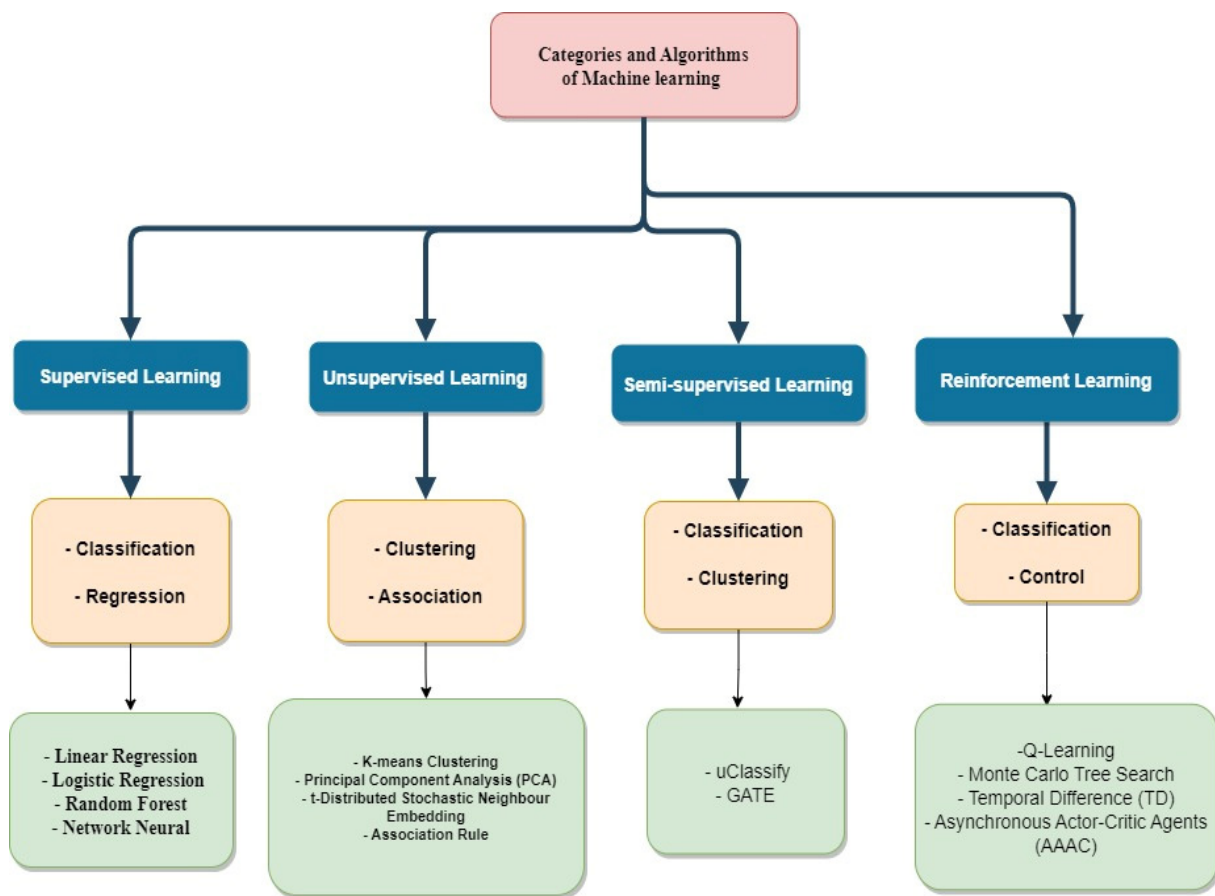


Figure 1. Different machine learning categories and algorithms [14].

DL does not require human-designed rules to function; rather, it leverages a massive amount of data to map the provided input to particular labels. DL is created utilising multiple layers of algorithms (artificial neural networks, or ANNs), each of which delivers a unique interpretation of the data provided to it [24,25]. Conventional machine learning (ML) techniques involve multiple sequential steps to accomplish the classification task, including pre-processing, feature extraction, intelligent feature selection, learning, and classification. In addition, feature selection has a significant impact on the performance of machine learning algorithms. Biased feature selection may lead to inaccuracy in class distinction. In contrast, DL can automate the learning of feature sets for several tasks, unlike standard ML algorithms [25,26]. DL enables learning and classification to be accomplished simultaneously.

Due to the complex multi-layer structure, a deep learning system needs a large dataset to smooth out the noise and generate accurate interpretations. Deep learning requires far more data than traditional machine learning algorithms. Machine learning may be utilised with as few as 1000 data points, but deep learning often only needs millions of data points [25,27]. Table 2 shows the advantages and disadvantages of DL.

2.2. Different Machine Learning Categories

The various machine learning types and algorithms are depicted in Figure 1 and will be discussed in detail below.

Classical machine learning is sometimes classified according to how an algorithm learns to make more precise predictions. There are four fundamental learning methodologies: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [28,29]. As seen in Figure 1, the type of method data scientists select depends on the sort of data they wish to forecast.

Table 2. Advantages and disadvantages of deep learning.

Advantages of Deep Learning	Disadvantages of Deep Learning
The potential to generate novel features from the limited existing training data.	There is less room for improvement in the training process because the entire training process depends on the constant flow of data.
Can produce results for tasks that are dependable and actionable by using unsupervised learning approaches.	With more datasets available, computational training becomes substantially more expensive.
It cuts down on the amount of time needed for feature engineering, one of the activities involved in learning how to use machine learning.	Transparency in fault revision is lacking. There are no intermediary stages to support a particular fault's claims. A whole algorithm is updated to address the problem.
Continuous training has made its architecture change-adaptive and capable of solving a variety of issues	For training the data sets, you need pricey resources, fast processors, and potent GPUs.

Supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, and deep learning are the five primary machine learning approaches.

2.2.1. Supervised Learning

The learning algorithm for this kind of machine learning is trained using labelled data. The data are referred to be labelled because it consists of pairs, the desired output that can be defined as a supervisory signal, and input that can be expressed as a vector [30,31].

Supervised learning occurs when the correct result is known beforehand [31]. Over time, the learning algorithm refines its predictions of this output in an effort to narrow the gap between its predictions and the actual output.

When the output is discrete, the supervised learning algorithm may create a classifier function, and when the output is continuous, a regression function is generated from the training data [30]. The learned function accurately predicts the output corresponding to any given input by making plausible generalisations of the patterns and features from the training data to fresh input data.

The two primary subcategories of supervised learning are regression algorithms (continuous output) and classification algorithms (discrete output). Regression algorithms look for the optimum function to match the training dataset's points. The three primary categories of regression algorithms are linear regression, multiple linear regression, and polynomial regression [32]. Assigning each input to the appropriate class allows classification algorithms to determine which class best fits the supplied data. The output of the predictive function in this instance is discrete and its value belongs to one of the possible classes [30].

Regression is used to solve regression issues [33], while SVMs are used to classify [34] Random forest is used to solve classification and regression issues [30].

When data are tagged and the classifier is used for classification or numerical prediction, supervised learning is employed. LeCun et al. [25] offered a concise but comprehensive overview of the supervised learning technique and the formation of deep architectures. Deng and Yu [14] described a variety of deep networks for supervised and hybrid learning, such as the deep stacking network (DSN) and its derivatives. Schmidhuber [20] discussed all neural networks, from the first neural networks to the most recent convolutional neural networks (CNN), recurrent neural networks (RNN), and long short-term memory (LSTM) and their advancements [35].

- Usage:

In industries such as sales, commerce, and the stock market, machine learning algorithms are frequently employed to anticipate prices. These are the sectors that rely heavily on projections for the future, and by employing supervised machine learning algorithms, more accurate forecasts may be created. Supervised algorithms are used by sales platforms such as Highspot and Seismic.

2.2.2. Unsupervised Learning

This approach trains the learning algorithm using an input dataset devoid of any labelled outputs, in contrast to supervised learning. For each input item, there is no right or wrong output, and unlike supervised learning, there is no human involvement to correct or adapt. Unsupervised learning is hence more arbitrary than supervised learning [10].

Unsupervised learning's primary objective is to obtain a deeper understanding of the data by recognising its basic structure or pattern of distribution. The algorithm tries to represent a specific detected input pattern while reflecting it on the general structure of input patterns as it learns on its own. As a result, the various inputs are grouped depending on the features that were taken from each input item [10]. Unsupervised learning is used to solve association and clustering issues.

Unsupervised learning is used to extract features from unlabelled data and categorise or label them when the input data are not labelled. LeCun et al. [25] projected the future of unsupervised deep learning. Schmidhuber [20] also outlined neural networks for unsupervised learning. Deng and Yu [9,29] provided an overview of deep learning architectures. Autoencoders are neural networks (NN) in which outputs are inputs. AE begins with the original input, encodes it into a compressed form, and then decodes it to recreate the original input (Wang).

In a deep AE, lower hidden layers are used for encoding while higher ones are utilised for decoding, and error back-propagation is employed for training [14,36].

- Usage:

Many unsupervised algorithms are currently being used in the digital advertising and marketing space. They are used to analyse the available customer-centric data and adapt services to individual customers. Further, it might help identify potential customers. Case in point: Salesforce, which is ideal for such purposes.

2.2.3. Semi-Supervised Learning

This technique uses a huge quantity of input data, some of which are labelled, while the rest are not, and lies between supervised and unsupervised learning. This branch of machine learning deals with several real-world learning issues. Because it uses a huge amount of unlabelled input and a very small amount of labelled data, semi-supervised learning requires less human interaction. Since fewer labelled datasets are more difficult to get, more costly, and perhaps need access to domain specialists, using them is more enticing. On the other hand, unlabelled datasets are less expensive and simpler to retrieve [37].

Both supervised and unsupervised training approaches can be used to teach the learning algorithm in semi-supervised learning. By employing unsupervised learning methods, the input dataset's latent patterns and structures may be exposed.

By contrast, supervised learning methods may be applied to unlabelled data to provide predictions based on best guesses, which can subsequently be applied to new data sets. Thus, we may say that unlabelled data are used to re-rank or re-evaluate a hypothesis or forecast established using labelled data [37].

In order to make use of the unlabelled training data, all semi-supervised learning methods rely on either the smoothness assumption, the cluster assumption, or the manifold assumption [37].

Hybrid learning architectures combine supervised (or "discriminative") and unsupervised (or "generative") learning components. By fusing together distinct architectures, a hybrid deep neural network may be constructed. Using action bank features to recognise human activities with them is anticipated to produce considerably better outcomes [37,38].

- Usage:

Semi-supervised machine learning is widely employed in the healthcare industry. It is utilised in the identification and analysis of speech, as well as the categorisation and management of digital content. There are several places where it may be used, including the regulatory sector. This technology allows for more accurate voice and picture analysis.

2.2.4. Reinforcement

Learning via interaction with the environment of the problem is called reinforcement learning. Instead of being explicitly instructed on what to do, a reinforcement learning agent learns via its own activities [39]. It chooses a current course of action based on previous encounters (exploitation) and fresh options (exploration). Thus, it might be characterised as a learning process based on trial and error. The reinforcement learning agent receives a signal in the form of a monetary reward value that indicates whether or not an action was successful. The agent aspires to develop the ability to choose options that maximise the worth of the monetary reward [40]. Actions may have an impact on future circumstances and reward values in addition to the current situation and current reward value.

Learning agents often have objectives established, and they can sense the state of the environment they are in to some extent. As a result, they can act to change the environment's state and get closer to the goals they have been given. Based on how each approach learns, reinforcement learning and supervised learning differ from one another.

The supervised learning approach uses case studies offered by an outside supervisor to teach. In contrast, reinforcement learning gains information through direct interactions with the issue environment [41].

The algorithm is learned by reinforcement learning, which utilises a system of rewards and punishments. In this, an agent or algorithm pulls up information from its environment. An agent receives a reward for appropriate behaviour and punishment for inappropriate behaviour. The agent in a self-driving car, for instance, would be rewarded for arriving at the location safely but punished for veering off the road. Similar to this, a chess-playing machine may incorporate a reward state of winning and a punishment state of being checkmated. The agent makes an attempt to maximise the reward and lower the penalty. In reinforcement learning, the algorithm is not told how to learn, thus it must figure out how to do so on its own [41].

For the next step generated by the learning model, reinforcement learning utilises a reward and punishment mechanism. Often employed for games and robotics, it handles decision-making issues [39]. Schmidhuber [20] outlined the advancements of deep learning in reinforcement learning (RL) as well as the applications of deep feedforward neural network (FNN) and recurrent neural network (RNN) for RL. The author of [39] reviewed deep reinforcement learning (DRL), its architectures such as deep Q-network (DQN), and its applications in a variety of disciplines. Mnih et al. [42] presented a DRL framework for DNN optimisation utilising asynchronous gradient descent. Van Hasselt et al. [43] suggested a deep neural network-based DRL architecture (DNN).

- Usage

It is advisable to employ reinforcement learning techniques when there is little or inconsistent information available. The gambling sector is where it is primarily used. The system can adapt to inconsistent player behaviour and modify the games with the help of machine learning algorithms. This method is used in the game creation of the well-known video game series Grand Theft Auto.

In self-driving automobiles, the approach is also being used. It can recognise streets, make turns, and choose which direction to turn. When the AI program AlphaGo defeated the human champion in the board game "Go", the technology garnered media attention [44]. Another such use is natural language processing.

It is evident that machine learning is advancing into practically every sphere of human activity and assisting in the resolution of several issues. We rely on technology heavily nowadays for chores that are part of our daily lives, whether it be social media, a meal delivery app, or an online taxi service [45].

Deep neural networks (DNN) have achieved tremendous success in supervised learning (SL). In addition, deep learning (DL) models are tremendously effective in unsupervised, hybrid, and reinforcement learning [25].

2.3. What Role Does Deep Learning Play in AI? (Evolution of Deep Learning)

The majority of signal processing methods up until recently relied on the use of shallow designs. Support vector machines (SVMs) [34], linear or nonlinear dynamical systems, and Gaussian mixture models (GMMs) are an example of these designs [46]. The human voice, language recognition, and visual sceneries are a few examples of real-world issues that call for a more sophisticated and layered architecture to be able to extract such complex information [46].

Artificial network research served as the initial inspiration for the notion of deep learning. Deep neural networks, which are an excellent illustration of models with a deep architecture, are the term “feed-forward neural networks” is frequently used. One of the most often utilised techniques for figuring out these networks’ parameters was backpropagation (BP) [18].

Henry J. Kelley presented the first-ever iteration of the continuous backpropagation model. Although his model is based on control theory [13,34], it paves the way for future advancement and will be used in ANN [11].

However, learning networks struggled when BP was used alone. They have a substantial number of hidden layers [47]. local optimum’s recurrent appearance in non-convex deep networks’ goal functions is the primary source of information. The challenge of optimisation with the deep models was actually mitigated when a method for unsupervised learning was presented.

The first CNN [48,49] architecture that can identify visual patterns such as handwritten letters, was presented by Kunihiko Fukushima. Geoffrey Hinton, Rumelhart, and Williams successfully used backpropagation in the neural network in 1986. Using backpropagation, Yann LeCun [25,47] taught a CNN to detect handwritten numbers. In a paper they released in 2006, the authors of [50] introduced DBN. DBN is a collection of restricted Boltzmann machines (RBMs). A greedy learning method that optimises DBN weights with time complexity proportional to the size and depth of the networks lies at the heart of the DBN [35].

A DNN’s modelling capabilities have been demonstrated to be greatly enhanced by adding hidden layers with a large number of neurons, which leads to the creation of several configurations that are nearly ideal [20,21]. The resultant DNN is still capable of functioning pretty well even in the event where parameter learning was trapped within a local optimum since the likelihood of having subpar local optimum decreases as the number of neurons utilised increases. Deep neural networks would, however, demand a lot of computing power during the training phase. Deep neural networks have just recently begun to receive major attention from academics because of the difficulty in obtaining enormous computing capabilities in the past.

Stanford scientist FeiFei Li established ImageNet [49] in 2009. There are 14 million accurately annotated photos in ImageNet. In 2012, Alex Krizhevsky created AlexNet, a CNN model that is GPU-implemented and had an accuracy of 84 percent in the ImageNet image classification competition [35,49]. In comparison to other gains, it achieved the best accuracy. Then, Ian Goodfellow created GAN [36].

GAN provides a new avenue for the use of DL in the industries of fashion, art, and science since it can synthesise data that are comparable to that found in the real world. The documentation listed above is summarised in Figure 2.

2.4. How Deep Learning Works? DL Workflow

The application of neural network-based DL technology is widespread across many industries and areas of study, including healthcare, sentiment analysis, NLP, the id of images, BI, cybersecurity, and many more [44,51,52].

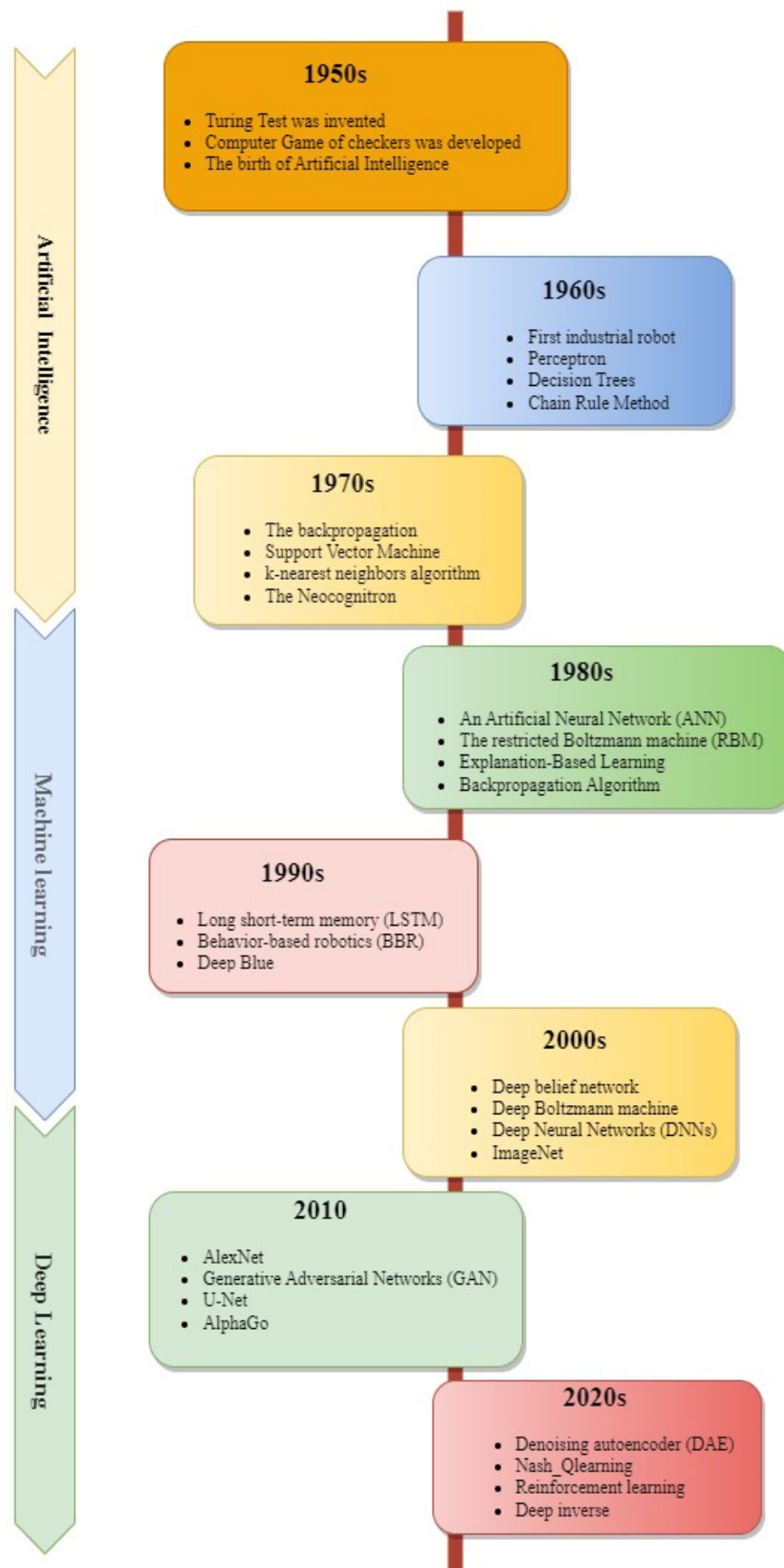


Figure 2. Deep learning timeline.

The dynamic nature and variety of real-world circumstances and data make it difficult to design an acceptable deep learning model, even if DL models are successfully implemented in the numerous application categories described above. There is also the belief that DL models are inherently mysterious and hence hinder the development of the field of deep learning [10].

Technology derived from ANN, which is called DL technology, is essential to accomplishing this goal. Many small, linked processing units (neurons) comprise the backbone of a typical neural network; these neurons are responsible for producing a sequence of real-valued activations that together provide the desired result [20]. The mathematical model of an artificial neuron (or processing element) is depicted in Figure 3 in a simplified schematic form. Signal output (s) is highlighted together with its input (X_i), weight (w), bias (b), summation function (Σ), activation function (f), and associated input (X_i) (y).

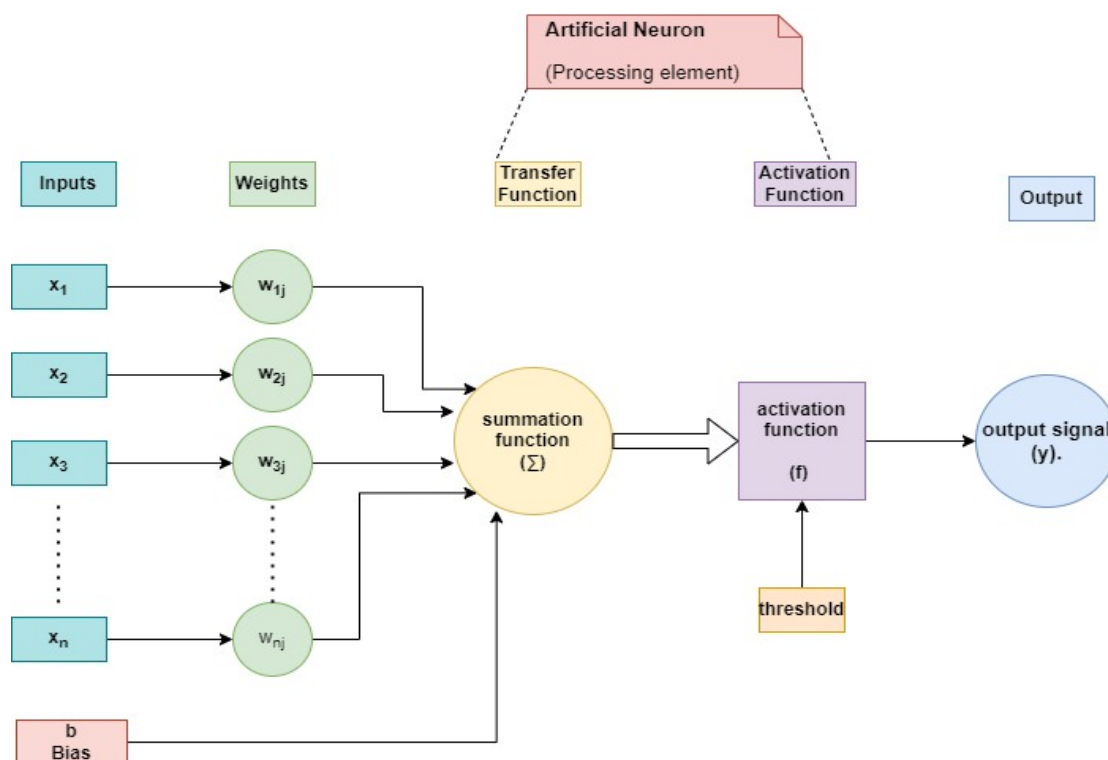


Figure 3. The mathematical model of an artificial neuron.

DL technology has the potential to transform the world as we know it, particularly in terms of a potent computational engine and its ability to support technology-driven automation smart and intelligent systems [28].

The foundation of deep learning is artificial neurons that mimic the human brain's neurons. A perceptron, also called an artificial neuron, mimics the behaviour of a real neuron by receiving information from a collection of inputs, each of which is given a certain amount of weight. The neuron uses these weighted inputs to compute a function and provide an output. N inputs are sent to the neuron (one for each feature). Then, it adds up the inputs, performs some kind of operation on them (the activation), and produces the output (see Figure 3). The significance of an input is measured by its weight. The neural network will place more importance on inputs that carry more weight. The output of the neural network can be fine-tuned for each individual perceptron by adjusting its bias parameter. It ensures the best feasible model-data fit. An activation function is a transformation between inputs and results. Applying a threshold results in an output. Linear, identity, unit, binary step, sigmoid, logistic, tanh, ReLU, and SoftMax are some examples of activation functions. Since a single neuron is unable to process many inputs,

multiple neurons are employed in order to reach a conclusion. As can be seen in Figure 4, a neural network is made up of perceptrons that are coupled in different ways and run on distinct activation functions. Any neural network with more than two layers is considered a deep learning model. In data processing, “hidden layers” refer to the intermediate levels between input and output. These layers have the potential to enhance precision. Although neural networks can resemble the brain, their processing power is nowhere near that of the human brain. Remember too that neural networks benefit from huge datasets to train from. As one of the most rapidly expanding subfields in computational science, deep learning makes use of large multi-layered networks to model top-level patterns in data.

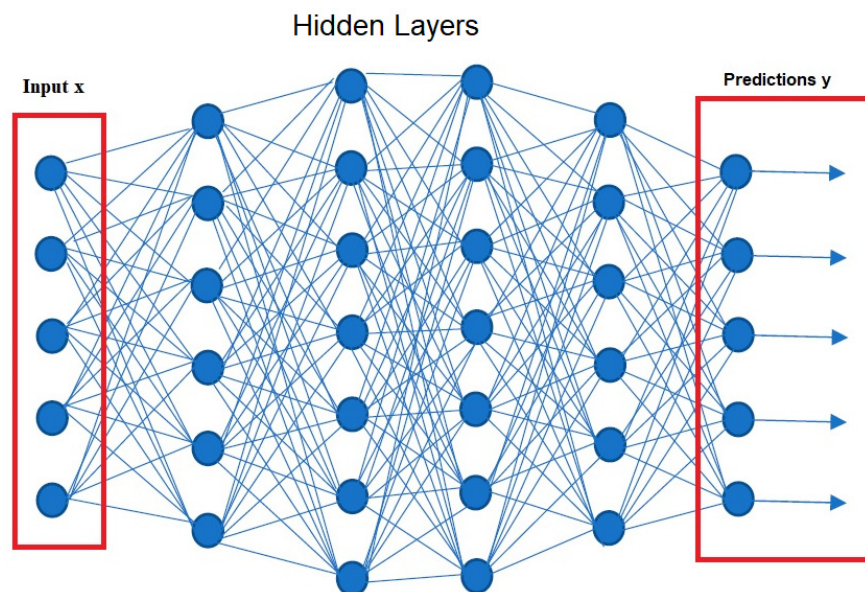


Figure 4. Representation of a neural network [20].

The network’s weights are iteratively tweaked to reduce the loss function, or the gap between the observed and ideal output veci. Internal hidden layers that are neither inputs nor outputs become prominent after weight modification. Interactions between these levels guarantee task domain regularities [18,47].

The most frequent activation function, the logic sigmoid function, is used to summarise the backpropagation procedure.

$$\sigma(\alpha) = \frac{1}{1 + e^{-\alpha}} \quad (1)$$

where α denotes activation and is a linear combination of the inputs $x = \{x_1, \dots, x_p\}$ and $w = \{w_1, \dots, w_p\}$ are the weights.

$$\hat{y} = \delta \left(w_0 + \sum_{i=1}^p w_i x_i \right) \quad (2)$$

The extra weight w_0 , often known as the bias, is included even though it has nothing to do with the input. The common activation function, the logic sigmoid function, provides the actual outputs \hat{y} .

Convolutional Neural Networks CNN

CNN is the most prominent and widely used algorithm in the field of DL [49,53,54]. The main advantage of CNN over its predecessors is that it automatically picks out important parts without any help from a person. CNNs have been utilised widely in a variety of fields, such as computer vision, voice processing, face recognition, etc. Similar to a normal

neural network, the structure of CNNs is influenced by neurons in human and animal brains. CNN simulates the complicated sequence of cells that make up the visual cortex of a cat's brain. Goodfellow et al. [36] identified three significant advantages of CNN: comparable representations, sparse interactions, and parameter sharing. In contrast to typical fully connected (FC) networks, CNN employs shared weights and local connections to make full use of 2D input data structures such as picture signals.

This method uses an extremely small number of parameters, which makes training the network easier and speeds it up. This is the same as the cells of the visual cortex. Interestingly, only tiny parts of a scene are perceived by these cells as opposed to the entire picture (i.e., these cells spatially extract the available local correlation in the input, similar to local filters over the input).

A popular version of CNN is similar to the multi-layer perceptron (MLP) [55] in that it has many convolution layers followed by subsampling (pooling) levels and FC layers as the last layers. Figure 5 is an example of the CNN architecture for image classification. The input x of each layer in a CNN model is structured in three dimensions: height, width, and depth, or $m \times m \times r$, where the height (m) equals the width (m). The term depth is also known as the channel number. In an RGB image, for instance, the depth (r) is equal to three.

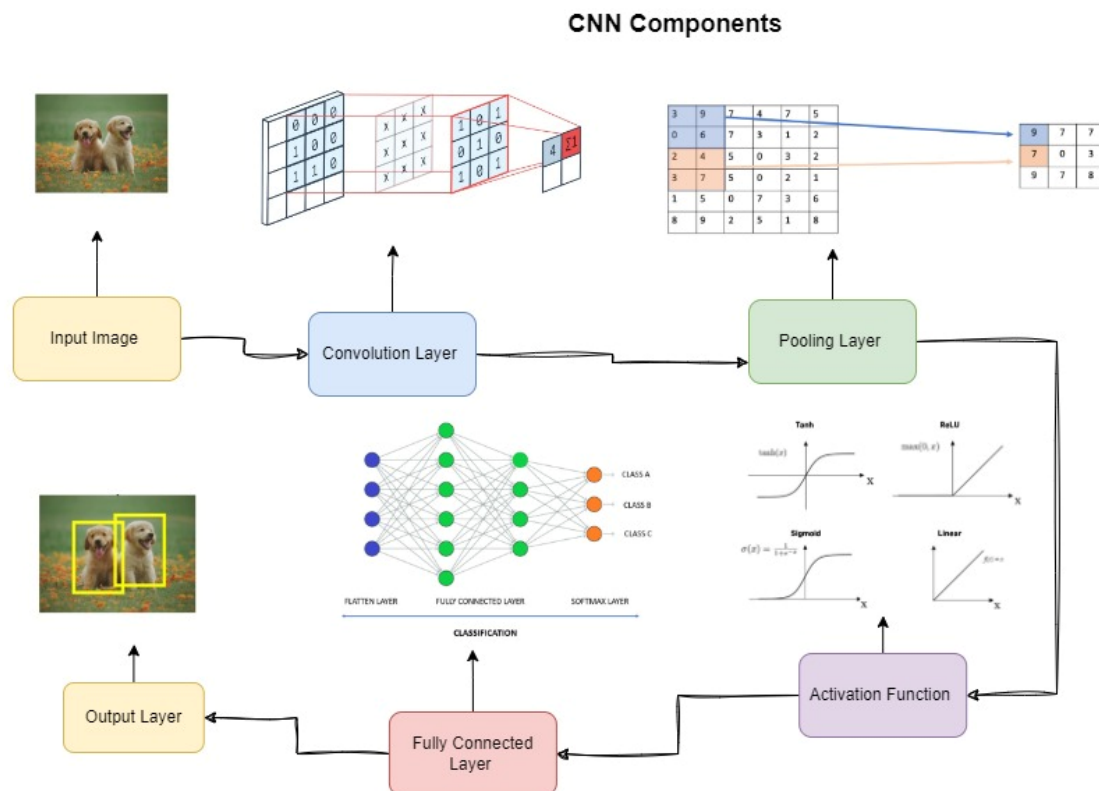


Figure 5. The CNN components.

Multiple kernels (filters) accessible in each convolutional layer are designated by k and have three dimensions ($n \times n \times q$), comparable to the input picture; however, n must be less than m and q must be equal to or less than r . In addition, the kernels serve as the foundation for the local connections, which share comparable characteristics (bias b^k and weight W^k) for producing k feature maps h^k with a size of $(m - n - 1)$ and are convolved with input, as described before. Similar to NLP, the convolution layer generates a dot product between its input and the weights as shown in Equation (1), but its inputs are less

than the original picture size. Then, by adding nonlinearity or an activation function to the output of the convolution layer, we obtain the following:

$$h^k = f(W^k \times x + b^k) \quad (3)$$

Then, each feature map in the subsampling layers is downsampled. This results in a decrease in network parameters, which speeds up the training process and facilitates the resolution of the overfitting problem. The pooling function (such as maximum or average) is applied to a neighbouring region of size $p \times p$, where p is the kernel size, for all feature maps. The FC layers then receive the mid- and low-level data and generate the high-level abstraction, which corresponds to the final stage layers of a normal neural network. The classification scores are produced by the last layer (e.g., support vector machines (SVMs) or SoftMax) [34]. Each score reflects the likelihood of a specific class for a particular event.

2.5. Deep Learning Approaches

Deep neural networks are effective in hybrid learning as well as supervised, unsupervised, and reinforcement learning [56]. Table 3. Show the advantage and limitations of the most well-known deep learning approaches.

Table 3. Deep learning approaches.

	Advantages	Limitation
CNN	Highly effective for visual recognition	The quantity and ability of the training data have a significant impact on CNN performance.
	After learning a segment inside a certain area of an image, CNN can recognise that segment anywhere else in the picture.	extremely sensitive to noise
RNN	An RNN uses the same parameters throughout each phase, unlike a conventional neural network. This significantly lowers the number of parameters we need to memorise.	It is challenging for RNNs to monitor long-term dependence. This is particularly true when there are too many words between the noun and the verb in extended phrases and paragraphs.
	For unlabelled photos, RNNs may be used in conjunction with CNNs to produce precise descriptions.	RNNs can't be combined to create highly complex models. The reason for this is that the gradient decays over several layers as a result of the activation function employed in RNN models.
Generative Adversarial Networks (GANs)	GANs enable effective semi-supervised classifier training.	The effectiveness of the generator and discriminator is essential to GAN's success. Even if one of them fails, the entire system collapses.
	The produced data are practically indistinguishable from the original data due to the model's increased accuracy.	The discriminator and generator are distinct systems that were trained using various loss functions. It might thus take a long time to train the entire system.
Autoencoders	A model is produced that is mostly dependent on data rather than predetermined filters.	Sometimes training demands a lot of time.
	Very little complexity makes them simpler to train.	The information that emerges from the model may be hazy and confusing if the training data are not indicative of the testing data.
ResNets	In some situations, ResNets are more accurate and need fewer weights than LSTMs and RNNs.	If a ResNet has too many levels, faults may be difficult to see and difficult to transmit back fast and accurately. However, if the layers are too thin, the learning may not be as effective.
	A network may be built by adding tens of thousands of residual layers, which can then be trained.	

2.6. DL Properties and Dependencies

Typically, a DL model follows the same processing phases as machine learning models. Figure 6 depicts a deep learning workflow for solving real-world issues. This workflow consists of four processing steps: data comprehension and preprocessing, DL model construction and training, and validation and interpretation.

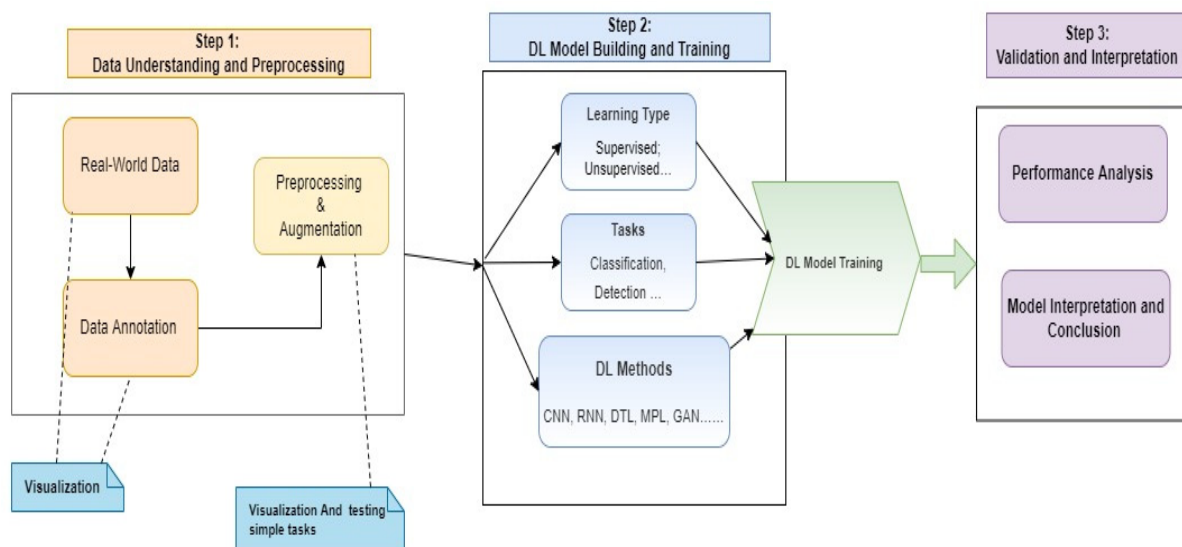


Figure 6. A typical DL workflow to solve real-world problems.

2.6.1. Understanding Different Types of Data

In order to create a data-driven intelligent system in a certain application area, as DL models learn from data, it is crucial to have a thorough understanding of and representation of the data [57]. Data in the actual world can take on a variety of shapes, however, for deep learning modelling, it is often possible to describe it as follows:

- **Sequential Data:**

Any type of data where the order matters, or a collection of sequences, is referred to as sequential data. While developing the model, it must explicitly take into consideration the input data's sequential character. Sequential data includes but is not limited to, text streams, audio snippets, and video clips.

- **2D data or an image:**

A matrix, or rectangular array of numbers, symbols, or expressions arranged in rows and columns in a 2D array of integers, is the building block of a digital image. The four fundamental elements or properties of a digital image are matrices, pixels, voxels, and bit depth.

- **Tables of Data:**

The main components of a tabular dataset are rows and columns. As a result, tabular datasets have data that is organised into columns just like a database table.

Each field (column) must be given a name, and each field may only hold data of the designated kind.

Overall, the data are organised logically and methodically into rows and columns based on the characteristics or aspects of the data. We can create intelligent systems that are driven by data by using deep learning models, which can learn effectively from tabular data.

Deep learning applications in the real world frequently use the kind of data mentioned above. So, various kinds of DL techniques perform differently depending on the nature and qualities of the data. Yet depending on the application, traditional machine learning

algorithms, especially those based on logic rules or trees, perform much better in many real-world application areas [14,56].

2.6.2. The Dependencies and Properties of DL

DL modelling usually employs the same processing stages as machine learning modelling. In Figure 6, we depict a deep learning workflow that consists of three stages: understanding and preparing the data, creating and training the DL model, and finally validating and interpreting the results. In contrast to ML modelling, feature extraction in the DL model is handled automatically. Machine learning methods such as k-nearest neighbours, decision trees, random forests, naive Bayes, linear regression, association rules, and k-means clustering are widely used in a wide variety of fields of application. The DL model incorporates various types of neural networks, such as convolutional networks, recurrent networks, autoencoders, deep belief networks, and many more.

Hardware requirements and large computational operations are needed by the DL algorithms when a model is being trained using a lot of data. The GPU is mostly used to optimise the processes effectively since the more the computations, the greater the benefit of a GPU over a CPU. GPU hardware is therefore required for deep learning training to function effectively. As a result, DL relies more on powerful computers with GPUs than traditional machine learning techniques [58].

2.6.3. Process for Feature Engineering

Using domain expertise, feature engineering is the process of removing features (characteristics, qualities, and attributes) from unstructured data. The attempt to directly extract high-level properties from data distinguishes DL from other machine-learning techniques in key ways. As a result, DL reduces the time and effort needed to build a feature extractor for each issue.

2.6.4. Model Training and Execution Time

Time for model training and execution: Due to the enormous number of parameters in the DL algorithm, training a deep learning algorithm typically takes a long time; as a result, model training takes longer. For instance, training with DL models can take longer than a week, whereas training with ML algorithms just takes a few seconds to a few hours [16,18]. Compared to some machine learning techniques, deep learning algorithms run extremely quickly during testing.

2.6.5. Black-Box Perception and Interpretability

Understanding and Interpretability of Black Boxes: When contrasting DL and ML, interpretability is a crucial consideration. Understanding a deep learning result, or “black box”, is challenging. The machine learning algorithms, in particular rule-based machine learning approaches [16,30]. There are a plethora of deep learning (DL) libraries and tools [59] that provide these fundamental utilities, as well as numerous pre-trained models and other crucial features for DL model construction and development. Among these are PyTorch [23] (which has a lightning-level API) and TensorFlow [60,61] (which also offers Keras as a high-level API).

Figure 6 shows a typical DL workflow to solve real-world problems, which consists of three sequential stages (i) data understanding and preprocessing (ii) DL model building and training (iii) validation and interpretation.

3. Some Deep Learning Applications

Over the past few years, deep learning has been applied successfully to address a wide range of problems in a variety of application fields.

Robots, enterprises, cybersecurity, virtual assistants, image recognition, healthcare, and many more are examples of these. They also involve sentiment analysis and natural language processing [14].

In Figure 7, we have listed a few possible real-world application domains for deep learning.

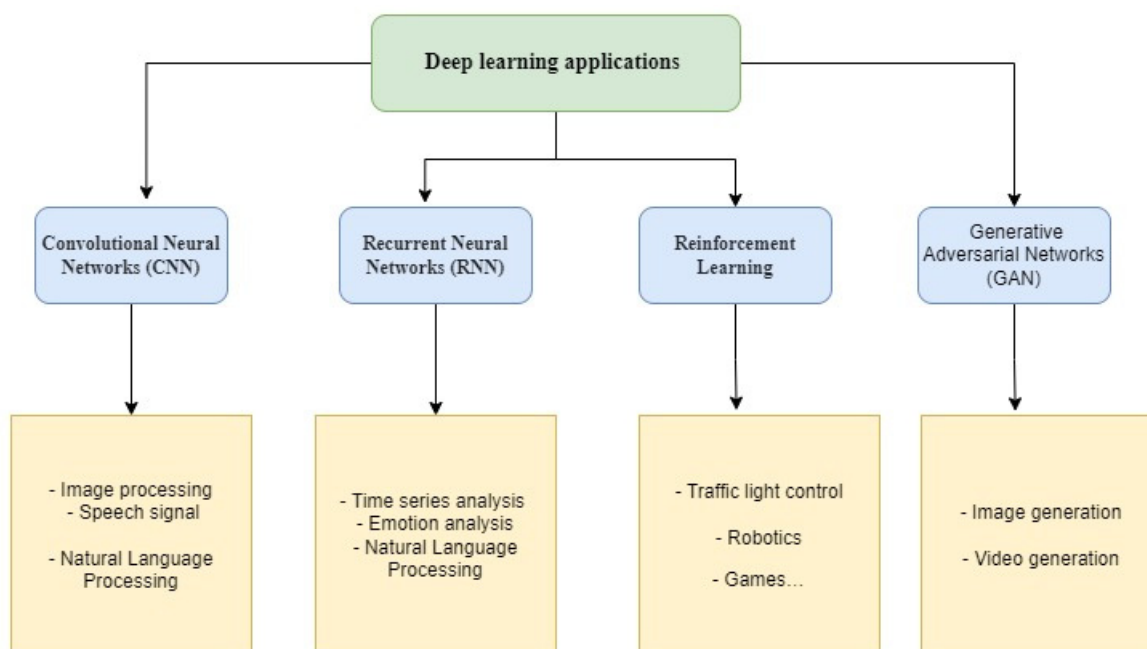


Figure 7. Several potential real-world application areas of deep learning.

Diverse deep learning techniques are applied in several application areas.

Overall, Figure 7 makes it quite evident that deep learning modelling has a bright future and a variety of possible applications.

Deep learning can be viewed as a technique for enhancing outcomes and streamlining processing times in many computing processes. Deep learning techniques have been used for handwriting generation [11] and image caption production [62] in the field of natural language processing. The following uses fall under the categories of biometrics, medicine, and pure digital image processing.

3.1. Recognition of Objects in Images

Before deep learning emerged as a new method of investigation, a number of applications had been implemented based on the principle of pattern recognition via layer processing.

By merging Bayesian belief propagation with particle filtering, an intriguing example was developed in 2003. The main theory behind this application is that since a human can recognise a person's face by only glancing at a half-cropped photograph of their face [14], a computer could be able to rebuild the image of a face from a cropped one.

By combining the greed algorithm with a hierarchy, a program was developed in 2006 that could read and interpret handwritten numbers [62]. Deep learning has been employed as the principal approach for digital image processing in recent research. For instance, adopting a convolutional neural network (CNN) instead of standard iris sensors can be more successful for iris identification.

Finally, facial recognition is a noteworthy use of deep learning in digital image processing. Deep learning facial recognition models are proprietary to Google, Facebook, and Microsoft [63]. Recent advancements have transformed facial image-based identification into automatic recognition by using age and gender as the basic parameters. For instance, Sighthound Inc. tested a deep convolutional neural network system that can identify emotions in addition to age and gender. Furthermore, by using a deep multi-task learning architecture, a reliable system was created to precisely identify a person's age and gender from a single photograph [62].

3.2. Biometrics

Automatic voice recognition was tested in 2009 to lower the phone error rate (PER) [64] using two different deep belief network topologies. In 2012, the CNN method was used using a Hybrid Neural Network-Hidden Markov Model (NN-HMM) [65]. This resulted in an effective rate of return of 20.07 percent. In compared to a prior three-layer neural network baseline approach, the PER obtained is better [66]. Smartphones and their camera resolution have been tried and tested for iris recognition.

When it comes to safety, especially for the purpose of access control, deep learning is used in conjunction with biometric features. The creation and fine-tuning of the Facial Sentinel face recognition devices were sped up with the use of DL. This firm estimate that in nine months, its products' one-to-many identification capabilities may be increased. This engine development may have taken five times as long without DL's implementation. It sped up manufacturing and market entry for the new machinery. Deep learning algorithms improve as they get more experience. The next few years should be remarkable as technology continues to advance [66].

3.3. Natural Language Processing

Deep learning is used in many fields of natural language, including voice translation, machine translation, computer semantic comprehension, etc.

In reality, deep learning is only successful in two areas: image processing and natural language processing.

In 2012, Schwenk et al. [67] suggested a deep neural network-based phrase-based statistical machine translation system (DNN).

It was capable of learning accurate translation probabilities for unseen phrases that were not included in the training set. Dong et al. [9] introduced a new adaptive multi-compositionality (AdaMC) layer in the recursive neural network in 2014.

This model included many composition functions, which were selected adaptively based on the input features. In 2014, Tang et al. [68] presented a DNN for sentiment analysis on Twitter data. Google introduced deep learning-based Word Lens recognition technology in 2015, which was utilised for real-time phone translation and video translation.

Not only could this technology read the words in real time, but it could also translate them into the desired target language. In addition, translation work might be performed over the phone without the need for networking. The existing technology might be used to translate more than twenty languages visually. In addition, Google offered an automatic mail reply function in Gmail that extracted and semantically analysed email text using a deep learning model. Ultimately, a response is generated based on the semantic analysis. This method is fundamentally distinct from conventional email auto-response capabilities [69].

ChatGPT is an NLP (natural language processing) technique that enables machines to comprehend and respond to human language. It is a form of artificial intelligence (AI) system developed by the US company OpenAI that uses deep learning algorithms to generate human-like dialogues in natural language [70]. It is a form of text-based communication between machines and people. It uses a lot of data to build a deep learning model that can understand and respond to real language. This model is based on the Transformer model. The model was trained using Internet-based text databases. This contained an astounding 570 GB of material acquired from online books, web texts, and articles. To be even more precise, the algorithm was given 300 billion words. The technology is built primarily to comprehend and generate natural language. This distinguishes it from the majority of other chatbots, which are often built to offer predetermined responses to specific inputs or execute specified tasks. It distinguishes itself from all other chatbots on the market due to its capacity to remember and answer differently to previous questions. The responses are also less "conversational" than those of a traditional chatbot. Although the online interface is in English, ChatGPT understands and answers in French, German, Russian, and Japanese, among others [71].

3.4. Recommender Systems (RS)

Recommender Systems (RS) are an efficient tool to deal with the “overload” of information provided by users. The three most common recommendation approaches are content-based (CB), collaborative filtering (CF), and hybrid recommendation methods. Recently, DL techniques have been applied in the field of RS, facilitating the overcoming of obstacles of conventional models and achieving high recommendation quality. Several applications of deep learning models in the RS framework are described in the following, classified based on the DL model used in the RS. AEs can be used to build an RS either by learning a low-dimensional representation of features or by directly providing the missing entries of the rating matrix in the construction layer.

Sedhain et al. [72] proposed an AE-based collaborative filtering RS, referred to as AutoRec, where the original partial observed vectors have been replaced by integer ratings. Experimental results imply that AutoRec outperforms biased matrix factorisation, RBM-based CF [73], and local low-rank matrix factorisation (LLORMA) with respect to accuracy on MovieLens and Netflix 62 data sets. Wu et al. and Wang et al. have proposed a method which combines stacked denoising AEs and probabilistic matrix factorisation for improving the performance of rating prediction. RBMs have also been exploited in the context of RS, with Salakhutdinov et al. developing a model which combines RBM with singular value decomposition (SVD) to improve performance.

He et al. [74] have proposed a framework named neural network collaborative filtering (NCF). The MLP is utilised to replace MF and learn the user–item interaction function. Tay et al. have proposed a new neural architecture for recommendation with reviews, which utilises an MLP for multi-pointer learning schemes. The proposed method performs significantly better than state-of-the-art methods, achieving relative improvements of up to 71% compared.

3.5. Mobile

Smartphones and wearables with sensors are revolutionising a number of mobile applications, including health monitoring [75]. As the distinction between consumer health wearables and medical devices becomes more blurred, it is now conceivable for a single wearable device to monitor a variety of medical risk factors. These devices might potentially provide patients with direct access to personal analytics that contribute to their health, facilitate preventative care, and assist in the management of chronic sickness [75]. Deep learning is seen as a crucial component for understanding this new form of data. Nevertheless, only a few recent publications in the healthcare sensing arena employed deep models, mostly due to hardware restrictions. In actuality, executing an efficient and dependable deep architecture on a mobile device to analyse noisy and complicated sensor data is a tough effort that is likely to deplete device resources [76]. Numerous research examined methods for overcoming these hardware restrictions. As an illustration, Lane and Georgiev [77] suggested a low-power deep neural network inference engine that used both the central processing unit (CPU) and the digital signal processor (DSP) of the mobile device without causing a significant hardware overload. In addition, they presented DeepX, a software accelerator capable of reducing the device resources required by deep learning, which operates as a significant barrier to mobile adoption at present. This architecture enables the effective execution of large-scale deep learning on mobile devices and outperformed cloud-based offloading alternatives [78,79].

3.6. Clinical Imaging

After the success of computer vision, the first clinical uses of deep learning were in image processing, especially the analysis of brain MRI scans to predict Alzheimer’s disease and its variations [80,81]. In other medical domains, CNNs were used to infer a hierarchical representation of low-field knee MRI data in order to automatically segment cartilage and predict the likelihood of osteoarthritis [82]. While utilising 2D photos, our system yielded superior results compared to a state-of-the-art method employing manually

picked 3D multi-scale features. Deep learning was also used to segment multiple sclerosis lesions in multi-channel 3D MRI images [83] and to differentiate benign from malignant breast nodules in ultrasound pictures [83]. Gulshan et al. [84] employed CNNs to diagnose diabetic retinopathy in retinal fundus photos, achieving high sensitivity and specificity across around 10,000 test images in comparison to certified ophthalmologist comments. CNN also outperformed 21 board-certified dermatologists in classifying biopsy-proven clinical images of various types of skin cancer (keratinocyte carcinomas versus benign seborrheic keratoses and malignant melanomas versus benign nevi) over a large data set of 130,000 images (1942 biopsy-labelled test images) [85].

3.6.1. Medical Applications

Deep learning's predictive capability and its ability to automatically identify features make it a popular tool for disease detection. Using either frequency or species, the applications of deep learning in the medical industry are continually evolving. In 2014, Li et al. [81] introduced a CNN-based classification system for lung image patches. This model avoided overfitting by employing the dropout method and a single-volume structure. Li et al. [82] introduced a DNN-based framework to identify the Alzheimer's disease (AD) phases from MRI and PET scan data in 2015. Their SC-CNN algorithm outperformed the traditional method of feature categorisation. Google created a visual technology for the early detection of ocular disorders in 2016. They collaborated with the Moorfields Eye Hospital to develop early prevention strategies for conditions such as diabetic retinopathy and age-related macular degeneration. A month later, Google employed deep learning techniques to build a radiotherapy approach for head and neck cancer that effectively controlled the patient's radiotherapy time and minimised the patient's radiotherapy injury. Deep learning in the field of precision medical care will make more important contributions as deep learning technology gets better [83].

3.6.2. Tasks That Machine Learning Can Handle in Healthcare

Detecting and Diagnosing Critical Diseases:

Machine learning in healthcare is beneficial for identifying and diagnosing critical diseases such as cancer and genetic disorders. In addition, advances in image diagnostic tools will be included in the AI-driven diagnosis procedure.

Drug Discovery and Manufacturing:

In the early stages of drug development, machine learning in healthcare plays a crucial role. AI-based technology enables the discovery of alternate treatments for complex disorders. Using gadgets and biosensors with sophisticated health measuring capabilities, personalised medications and treatment alternatives will also be possible in the next years.

Keeping Health Records:

Machine learning has made it easier to maintain health records, saving time and money. In the next years, ML-based smart health records will also aid in the development of more precise diagnoses and recommend more effective therapeutic treatments.

Clinical Trials and Research:

Machine learning offers significant benefits in clinical trials and research since it enables researchers to concurrently access several data points. In addition, the system utilises real-time monitoring, data access for trial participants, and electronic recordkeeping to reduce data-based mistakes. In order to better the detection and diagnosis of life-threatening illnesses, researchers and medical professionals are currently soliciting vast quantities of data from the public with their permission.

The application of AI and ML in the clinical practice of healthcare poses a number of ethical considerations [86,87].

- **Knowledgeable Consent to Use**

This means figuring out how much of a patient's responsibility it is to teach them about the complexity of AI, such as the kind of data that can be collected and the possible limits of using AI.

- **Security and Openness**

Safety is one of the most serious problems associated with the use of AI in healthcare diagnosis and treatment. To minimise injury, professionals must assure the safety and dependability of these systems and be clear about them.

- **Fairness of algorithms and Biases**

A machine learning system is only as dependable and successful as its training (to interpret data and learn from it to perform a task with accuracy). So, the people who make AI should deal with this problem and get rid of biases at every level to make sure it does not hurt the effectiveness of healthcare solutions.

- **Data Privacy**

Patients must be provided with sufficient information on the collection and processing of their data in order to respect their basic privacy rights.

4. The Future Directions

To wrap up our study and show the future directions, an itemised analysis is next offered.

DL already has trouble concurrently modelling several complicated modalities of data. Multimodal DL is another popular strategy in the current advancements in DL.

Deep learning techniques have a significant influence on the model that emerges for a certain issue domain in terms of data availability and quality for training. Such data problems might result in subpar processing and false conclusions, which is a serious concern. Efficient methods for pre-processing data are needed so that designs can be made that fit the nature and characteristics of the data issue.

MLP, CNN, and RNN are the three primary types of deep networks for supervised learning, as well as their variations, which are extensively used in many different application sectors. However, a fresh contribution might be made by developing new methods or their variations that take into consideration model optimisation, accuracy, and application.

To forecast unknown data and train the models, DL needs huge datasets (labelled data are better). This problem becomes very challenging when real-time data processing is necessary or when there are few datasets available (such as in the case of healthcare data).

Over the past few years, DL and data augmentation have been looked at as possible solutions to this problem.

Many of the existing deep learning models use supervised learning, even though ML gradually moves to semi-supervised and unsupervised learning to handle real-world data without the requirement for manual human labelling.

Regarding the issue of a lack of training data, it is expected that various transfer learning techniques, such as training the DL model on large datasets of unlabelled images before using the knowledge to train the DL model on a limited number of labelled images for the same task, will be considered.

4.1. The Difficulties of Deep Learning

4.1.1. Absence of Originality in Model Structure

From 2006, when deep learning was re-recognised, the deep learning model has primarily been introduced as the aforementioned classical approaches. On the basis of more than a decade of evolution, these classical models saw the final introduction of deep learning models. Due to the stacking of models in the past, it is becoming increasingly difficult to improve data processing efficiency. Nonetheless, the depth of the advantages of learning technology is not yet fully implemented, as it is necessary to recognise that the development of a new depth of learning model, either the current depth of the learning model or other appropriate methods for effective integration, is required to resolve the issue.

4.1.2. Modernise Training Techniques

Supervised and unsupervised learning are the two training approaches used for the deep learning models currently in use. The use of supervised training methods, the

restricted Boltzmann machine, and the automatic encoder as the core model, as well as the use of a large variety of training methods as the primary pre-training. The method involves unsupervised learning. Moreover, they are used with supervised learning to refine learning training. Completely unsupervised instruction serves no meaningful use. Future research on the deep learning model will therefore focus on achieving unsupervised training in its entirety.

4.1.3. Decrease Training Duration

Nowadays, the majority of deep learning model identification is carried out in an ideal environment. Current technology is still incapable of achieving the required outcomes in reality's complex context. In addition, the deep learning model consists of either a single model or several models. As the complexity of the problem increases, so does the amount of data that must be handled, necessitating longer and longer training periods for the deep learning model.

Future studies in deep learning technology will investigate how to improve the accuracy and speed of data processing by modifying the deep learning model without hardware flexibility.

4.1.4. Online Education

Deep learning methods used today are mostly based on unsupervised pre-training and supervised fine-tuning.

Yet, online learning training requires global fine-tuning, resulting in a local minimum output.

The present training is therefore not suitable for the implementation of online learning. The enhancement of online learning capabilities based on a novel deep learning model must be addressed.

Despite the promising findings produced by utilising deep architectures, the clinical application of deep learning to health care still faces a number of unresolved obstacles. Specifically, I emphasise the following critical issues:

Data volume: Deep learning is a collection of computationally intensive models. Such examples are fully connected multi-layer neural networks in which a multitude of network parameters must be accurately evaluated. Massive amounts of data are needed to accomplish this objective. While there are no hard and fast rules about the minimum amount of training documents, a good rule of thumb is to have at least 10 times as many samples as network parameters. This is also one of the reasons why deep learning is so effective in sectors where vast quantities of data can be easily collected (e.g., computer vision, speech, natural language).

Temporality: The progression and evolution of diseases across time are nondeterministic. Nonetheless, many existing deep learning models, including those recently suggested in the medical area, assume static, vector-based inputs, which cannot naturally account for the time factor. Developing deep learning methods capable of handling temporal healthcare data is a crucial component that will necessitate the creation of unique solutions.

Interpretability: Despite the success of deep learning models in a variety of application domains, they are frequently considered black boxes. While this may not be a problem in other more deterministic domains, such as picture annotation (since the end user may objectively check the tags assigned to the photos), in health care, both the quantitative algorithmic performance and the reason why the algorithms function are vital. In reality, such model interpretability (i.e., providing which phenotypes are driving the predictions) is vital for persuading medical practitioners to take the steps suggested by the predictive system (e.g., prescription of a specific medication, potentially high risk of developing a certain disease). Many of these obstacles present numerous opportunities and future research options for increasing the pitch.

5. Conclusions

Deep learning is an area of computer science that is very busy and growing quickly. Building a good deep learning model for an application is getting harder and harder because there are so many problems caused by the fact that there are so much complicated data.

Although there are still challenges to be tackled and deep learning is still in its infancy, it has shown amazing learning potential. It remains an active research area in the field of future AI. The most notable developments in deep learning and their applications to many fields have been discussed in this paper.

This article gives a clear overview of deep learning technology, which is important for AI and data science. It starts with an overview of how ANNs have changed over time and then moves on to more recent deep learning methods and breakthroughs in a range of fields. Then, we dive into deep neural network modelling and the most important methods in this space. In addition to this, we have provided a classification that takes into account the wide range of deep learning tasks and their varied applications.

Deep learning is different from traditional machine learning and data mining in that it can take very detailed representations of data from very large datasets. This has resulted in many fantastic answers to pressing practical issues. Data-driven modelling that is appropriate for the features of the raw data is essential for any deep learning technique to achieve success. Before a system can aid in intelligent decision-making, it must be educated using data and knowledge specific to the intended application through the use of complex learning algorithms. Applications and research fields that have found success with deep learning are outlined in the paper. These include picture object recognition, biometrics, natural language processing, and clinical imaging.

Finally, the problems that need to be solved, possible research directions, and outlook for the field have been highlighted. Even though deep learning is seen as a black-box solution for many applications because of its poor reasoning and interpretability, fixing the issues or future elements that are mentioned could lead to new generations of deep learning models and smarter systems. The researchers may do more thorough analyses, yielding more credible and plausible results as a result. The overall direction of this study's advanced analytics is encouraging, and it can be used as a resource for further study and practical application in related fields.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arel, I.; Rose, D.C.; Karnowski, T.P. Deep machine learning—A new frontier in artificial intelligence research [research frontier]. *IEEE Comput. Intell. Mag.* **2010**, *5*, 13–18. [[CrossRef](#)]
2. Benos, L.; Tagarakis, A.C.; Dolias, G.; Berruto, R.; Kateris, D.; Bochtis, D. Machine Learning in Agriculture: A Comprehensive Updated Review. *Sensors* **2021**, *21*, 3758. [[CrossRef](#)]
3. Huang, J.; Chai, J.; Cho, S. Deep learning in finance and banking: A literature review and classification. *Front. Bus. Res. China* **2020**, *14*, 13. [[CrossRef](#)]
4. Gambella, C.; Ghaddar, B.; Naoum-Sawaya, J. Optimization problems for machine learning: A survey. *Eur. J. Oper. Res.* **2021**, *290*, 807–828. [[CrossRef](#)]
5. Vuong, Q. Machine Learning for Robotic Manipulation. 2021. Available online: <https://arxiv.org/abs/2101.00755v1> (accessed on 11 April 2023).
6. Yuan, F.-G.; Zargar, S.A.; Chen, Q.; Wang, S. Machine learning for structural health monitoring: Challenges and opportunities. *Sens. Smart Struct. Technol. Civ. Mech. Aersp. Syst.* **2020**, *11379*, 1137903. [[CrossRef](#)]
7. Kubat, M. An Introduction to Machine Learning. In *An Introduction to Machine Learning*; Springer International Publishing: Cham, Switzerland, 2017; pp. 321–329.
8. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)]

9. Deng, L. Deep Learning: Methods and Applications. *Found. Trends Signal Process.* **2013**, *7*, 197–387. [CrossRef]
10. Karhunen, J.; Raiko, T.; Cho, K. Unsupervised deep learning: A short review. In *Advances in Independent Component Analysis and Learning Machines*; Academic Press: Cambridge, MA, USA, 2015; pp. 125–142. [CrossRef]
11. Du, K.L.; Swamy, M.N. *Neural Networks and Statistical Learning*, 2nd ed.; Springer Science & Business Media: London, UK, 2019; pp. 1–988.
12. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*; Morgan Kaufmann: Waltham, MA, USA, 2012. [CrossRef]
13. Haykin, S. *Neural Networks and Learning Machines*; Pearson Education USA: Upper Saddle River, NJ, USA, 2008.
14. Ahmad, J.; Farman, H.; Jan, Z. Deep learning methods and applications. In *Deep Learning: Convergence to Big Data Analytics*; SpringerBriefs in Computer Science; Springer: Singapore, 2019; pp. 31–42. [CrossRef]
15. Deep Learning Techniques: An Overview | SpringerLink. Available online: https://link.springer.com/chapter/10.1007/978-981-15-3383-9_54 (accessed on 11 March 2023).
16. Srinivas, M.; Sucharitha, G.; Matta, A. *Machine Learning Algorithms and Applications*; Wiley: Hoboken, NJ, USA, 2021. [CrossRef]
17. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [CrossRef]
18. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
19. Hassanien, A.E.; Chang, K.C.; Mincong, T. (Eds.) *Advanced Machine Learning Technologies and Applications*; Springer Nature: Singapore, 2021; Volume 1141. [CrossRef]
20. Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
21. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* **2014**, *4*, 3320–3328. [CrossRef]
22. Cireşan, D.C.; Meier, U.; Masci, J.; Gambardella, L.M.; Schmidhuber, J. High-Performance Neural Networks for Visual Object Classification. *arXiv* **2011**, arXiv:1102.0183. [CrossRef]
23. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst.* **2019**.
24. Zhang, Z.; Cui, P.; Zhu, W. Deep Learning on Graphs: A Survey. *IEEE Trans. Knowl. Data Eng.* **2020**, *34*, 249–270. [CrossRef]
25. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
26. Shrestha, A.; Mahmood, A. Review of Deep Learning Algorithms and Architectures. *IEEE Access* **2019**, *7*, 53040–53065. [CrossRef]
27. Bengio, Y. Learning Deep Architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–127. [CrossRef]
28. Mathew, A.; Amudha, P.; Sivakumari, S. Deep learning techniques: An overview. *Adv. Intell. Syst. Comput.* **2021**, *1141*, 599–608.
29. Deng, L. A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Trans. Signal Inf. Process.* **2014**, *3*, e2. [CrossRef]
30. Osisanwo, F.Y.; Akinsola, J.E.T.; Awodele, O.; Hinmikaiye, J.O.; Olakanmi, O.; Akinjobi, J. Supervised Machine Learning Algorithms: Classification and Comparison. *Int. J. Comput. Trends Technol.* **2017**, *48*, 128–138. [CrossRef]
31. Nasteski, V. An overview of the supervised machine learning methods. *Horizons. B* **2017**, *4*, 51–62. [CrossRef]
32. Panigrahi, A.; Chen, Y.; Kuo, C.C.J. Analysis on Gradient Propagation in Batch Normalized Residual Networks. *arXiv* **2018**, arXiv:1812.00342. [CrossRef]
33. Kumari, K.; Yadav, S. Linear regression analysis study. *J. Pr. Cardiovasc. Sci.* **2018**, *4*, 33. [CrossRef]
34. Du, K.-L.; Swamy, M.N.S. Support Vector Machines. In *Neural Networks and Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 593–644. [CrossRef]
35. Swapna, M.; Sharma, Y.K.; Prasad, B. CNN Architectures: Alex Net, Le Net, VGG, Google Net, Res Net. *Int. J. Recent Technol. Eng.* **2020**, *8*, 953–959. [CrossRef]
36. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* **2014**, *63*, 139–144. [CrossRef]
37. Xu, W.; Sun, H.; Deng, C.; Tan, Y. Variational Autoencoders for Semi-supervised Text Classification. In Proceedings of the 31st AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; pp. 3358–3364. [CrossRef]
38. Kameoka, H.; Li, L.; Inoue, S.; Makino, S. Supervised determined source separation with multichannel variational autoencoder. *Neural Comput.* **2019**, *31*, 1891–1914. [CrossRef] [PubMed]
39. Li, Y. Deep Reinforcement Learning: An Overview. *arXiv* **2017**, arXiv:1701.07274. [CrossRef]
40. Paliwal, M. Deep Reinforcement Learning. *Smart Innov. Syst. Technol.* **2022**, *273*, 136–142. [CrossRef]
41. Arulkumaran, K.; Deisenroth, M.P.; Brundage, M.; Bharath, A.A. A Brief Survey of Deep Reinforcement Learning. *IEEE Signal Process. Mag.* **2017**, *34*, 26–38. [CrossRef]
42. Mnih, V.; Badia, A.P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Silver, D.; Kavukcuoglu, K. Asynchronous Methods for Deep Reinforcement Learning. In Proceedings of the 33rd International Conference on Machine Learning, ICML 2016, New York, NY, USA, 19–24 June 2016; Volume 4, pp. 2850–2869. [CrossRef]
43. Van Hasselt, H.; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-learning. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 2094–2100. [CrossRef]
44. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, *529*, 484–489. [CrossRef]

45. Naeem, M.; Paragliola, G.; Coronato, A. A reinforcement learning and deep learning based intelligent system for the support of impaired patients in home treatment. *Expert Syst. Appl.* **2021**, *168*, 114285. [CrossRef]
46. Reynolds, D. Gaussian Mixture Models. *Encycl. Biom.* **2009**, *741*, 659–663. [CrossRef]
47. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
48. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef]
49. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
50. Hinton, G.E. Deep belief networks. *Scholarpedia* **2009**, *4*, 5947. [CrossRef]
51. Goyal, P.; Pandey, S.; Jain, K. Introduction to Natural Language Processing and Deep Learning. In *Deep Learning for Natural Language Processing*; Springer: New York, NY, USA, 2018; pp. 1–74. [CrossRef]
52. Taye, M.M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation* **2023**, *11*, 52. [CrossRef]
53. 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**, *33*, 6999–7019. [CrossRef]
54. Tomè, D.; Monti, F.; Baroffio, L.; Bondi, L.; Tagliasacchi, M.; Tubaro, S. Deep Convolutional Neural Networks for pedestrian detection. *Signal Process. Image Commun.* **2016**, *47*, 482–489. [CrossRef]
55. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
56. Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 420. [CrossRef] [PubMed]
57. Najafabadi, M.M.; Villanustre, F.; Khoshgoftaar, T.M.; Seliya, N.; Wald, R.; Muharemagic, E. Deep learning applications and challenges in big data analytics. *J. Big Data* **2015**, *2*, 1. [CrossRef]
58. Coelho, I.M.; Coelho, V.N.; Luz, E.J.D.S.; Ochi, L.S.; Guimarães, F.G.; Rios, E. A GPU deep learning metaheuristic based model for time series forecasting. *Appl. Energy* **2017**, *201*, 412–418. [CrossRef]
59. Serin, G.; Sener, B.; Ozbayoglu, A.M.; Unver, H.O. Review of tool condition monitoring in machining and opportunities for deep learning. *Int. J. Adv. Manuf. Technol.* **2020**, *109*, 953–974. [CrossRef]
60. Singh, P.; Manure, A. *Learn TensorFlow 2.0*; APress: Berkeley, CA, USA, 2020. [CrossRef]
61. Gad, A.F. TensorFlow: A Guide to Build Artificial Neural Networks Using Python Build Artificial Neural Networks Using TensorFlow Library with Detailed Explanation of Each Step and Line of Code. Available online: https://www.researchgate.net/publication/321826020_TensorFlow_A_Guide_To_Build_Artificial_Neural_Networks_Using_Python (accessed on 12 March 2023).
62. Shafiq, M.; Gu, Z. Deep Residual Learning for Image Recognition: A Survey. *Appl. Sci.* **2022**, *12*, 8972. [CrossRef]
63. Wasnik, P.; Raja, K.B.; Ramachandra, R.; Busch, C. Assessing face image quality for smartphone based face recognition system. In Proceedings of the 5th International Workshop on Biometrics and Forensics, IWBF 2017, Coventry, UK, 4–5 April 2017. [CrossRef]
64. Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. Applications of Deep Learning in Biomedicine. *Mol. Pharm.* **2016**, *13*, 1445–1454. [CrossRef]
65. Li, L.; Zhao, Y.; Jiang, D.; Zhang, Y.; Wang, F.; Gonzalez, I.; Valentin, E.; Sahli, H. Hybrid Deep Neural Network—Hidden Markov Model (DNN-HMM) based speech emotion recognition. In Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013, Geneva, Switzerland, 2–5 September 2013; pp. 312–317. [CrossRef]
66. Vetrekar, N.; Raja, K.B.; Ramachandra, R.; Gad, R.; Busch, C.; Ramachandra, R.; Gad, R.; Busch, C. Multi-spectral imaging for robust ocular biometrics. In Proceedings of the 2018 International Conference on Biometrics, ICB 2018, Gold Coast, QLD, Australia, 20–23 February 2018; pp. 195–201. [CrossRef]
67. Er Schwenk, H.O.G. Continuous Space Translation Models for Phrase-Based Statistical Machine Translation. 2012, pp. 1071–1080. Available online: <http://www.lium.univ-lemans.fr/~cslm> (accessed on 12 March 2023).
68. Tang, D.; Wei, F.; Qin, B.; Liu, T.; Zhou, M. Coooolll: A Deep Learning System for Twitter Sentiment Classification. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; Association for Computational Linguistics: Beijing, China, 2014; pp. 208–212.
69. Lin, M.S.; Tang, C.G.Y.; Kom, X.J.; Eyu, J.Y.; Xu, C. Building a Natural Language Processing Model to Extract Order Information from Customer Orders for Interpretative Order Management. In Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, Kuala Lumpur, Malaysia, 7–10 December 2022; pp. 81–86. [CrossRef]
70. Wang, F.-Y.; Miao, Q.; Li, X.; Wang, X.; Lin, Y. What Does ChatGPT Say: The DAO from Algorithmic Intelligence to Linguistic Intelligence. *IEEE/CAA J. Autom. Sin.* **2023**, *10*, 575–579. [CrossRef]
71. Du, H.; Teng, S.; Chen, H.; Ma, J.; Wang, X.; Gou, C.; Li, B.; Ma, S.; Miao, Q.; Na, X.; et al. Chat with ChatGPT on Intelligent Vehicles: An IEEE TIV Perspective. In *IEEE Transactions on Intelligent Vehicles*; IEEE: New York, NY, USA, 2023; pp. 1–7. [CrossRef]

72. Sedhain, S.; Menon, A.K.; Sanner, S.; Xie, L. Autorec: Autoencoders meet collaborative filtering. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; pp. 111–112. [[CrossRef](#)]
73. Salakhutdinov, R.; Mnih, A.; Hinton, G. Restricted Boltzmann machines for collaborative filtering. *ACM Int. Conf. Proc. Ser.* **2007**, *227*, 791–798. [[CrossRef](#)]
74. He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; Chua, T.S. Neural Collaborative Filtering. In Proceedings of the 26th International World Wide Web Conference, WWW 2017, Perth, Australia, 3–7 April 2017; pp. 173–182. [[CrossRef](#)]
75. Han, J.; Zhang, Z.; Mascolo, C.; Andre, E.; Tao, J.; Zhao, Z.; Schuller, B.W. Deep Learning for Mobile Mental Health: Challenges and recent advances. *IEEE Signal Process. Mag.* **2021**, *38*, 96–105. [[CrossRef](#)]
76. Zou, J.; Zhang, Q. UbiEi-Edge: Human Big Data Decoding Using Deep Learning on Mobile Edge. In Proceedings of the Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Guadalajara, Mexico, 1–5 November 2021; pp. 1861–1864. [[CrossRef](#)]
77. Lane, N.D.; Georgiev, P. Can Deep Learning Revolutionize Mobile Sensing? In Proceedings of the HotMobile 2015—16th International Workshop on Mobile Computing Systems and Applications, Santa Fe, NM, USA, 12–13 February 2015; pp. 117–122. [[CrossRef](#)]
78. Estonilo, C.G.; Festijo, E.D. Evaluation of the Deep Learning-Based m-Health Application Using Mobile App Development and Assessment Guide. In Proceedings of the 2022 IEEE 12th Annual Computing and Communication Workshop and Conference, CCWC 2022, Las Vegas, NV, USA, 26–29 January 2022; pp. 435–440. [[CrossRef](#)]
79. Liu, L.; Xu, J.; Huan, Y.; Zou, Z.; Yeh, S.-C.; Zheng, L.-R. A Smart Dental Health-IoT Platform Based on Intelligent Hardware, Deep Learning, and Mobile Terminal. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 898–906. [[CrossRef](#)]
80. Liu, S.; Liu, S.; Cai, W.; Pujol, S.; Kikinis, R.; Feng, D. Early diagnosis of Alzheimer’s disease with deep learning. In Proceedings of the IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014, Beijing, China, 29 April–2 May 2014; pp. 1015–1018. [[CrossRef](#)]
81. Brosch, T.; Tam, R. Manifold Learning of Brain MRIs by Deep Learning. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8150, pp. 633–640. [[CrossRef](#)]
82. Prasoon, A.; Petersen, K.; Igel, C.; Lauze, F.; Dam, E.; Nielsen, M. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8150, pp. 246–253. [[CrossRef](#)]
83. Yoo, Y.; Brosch, T.; Traboulsee, A.; Li, D.K.; Tam, R. Deep Learning of Image Features from Unlabeled Data for Multiple Sclerosis Lesion Segmentation. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2014; Volume 8679, pp. 117–124. [[CrossRef](#)]
84. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **2016**, *316*, 2402–2410. [[CrossRef](#)] [[PubMed](#)]
85. Alipanahi, B.; Delong, A.; Weirauch, M.T.; Frey, B.J. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* **2015**, *33*, 831–838. [[CrossRef](#)] [[PubMed](#)]
86. Malhotra, C.; Kotwal, V.; Dalal, S. Ethical framework for machine learning. In Proceedings of the 10th ITU Academic Conference Kaleidoscope: Machine Learning for a 5G Future, ITU K, Santa Fe, Argentina, 26–28 November 2018. [[CrossRef](#)]
87. Mageswaran, G.; Nagappan, S.D.; Hamzah, N.; Brohi, S.N. Machine Learning: An Ethical, Social & Political Perspective. In Proceedings of the 4th International Conference on Advances in Computing, Communication and Automation, ICACCA, Subang Jaya, Malaysia, 26–28 October 2018. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.